Czech Economic Review

Acta Universitatis Carolinae Oeconomica

Volume IX

Numbers 1–3

Charles University

Faculty of Social Sciences

Institute of Economic Studies

2015

Czech Economic Review

ACTA UNIVERSITATIS CAROLINAE OECONOMICA

Published by the Institute of Economic Studies, Faculty of Social Sciences, Charles University, http://ies.fsv.cuni.cz. All papers available at http://cer.fsv.cuni.cz.

Printed and distributed by Karolinum Press, http://www.karolinum.cz.

Contact address: Institute of Economic Studies, Faculty of Social Sciences, Charles University, Opletalova 26, 110 00 Prague 1, Czech Republic.

MK ČR E 18587
ISSN 1802-4696 (Print), ISSN 1805-9406 (Online)

Czech Economic Review

ACTA UNIVERSITATIS CAROLINAE OECONOMICA

Volume IX Numbers 1–3

Contents

Editorial

Over the last two years, our journal has witnessed many substantial changes. First, effective from January 2015, the journal has merged with *Bulletin of the Czech Econometric Society*. The two journals co-existed in the past with a very similar focus, and the merger served to concentrate the efforts and focus of academics in closely related fields of economic theory and econometrics. The aim of the merged journal was to promote the advancement of economic theory in its relation to statistics and mathematics and to serve as a focal point for the Central European community of scholars.

The merged journal has continued with the title and ISSN of the Czech Economic Review (ISSN 1802-4696 Print, ISSN 1805-9406 Online), and became a joint venture of Charles University and the Czech Econometric Society. As in the past, the journal explicitly attracted relevant articles in both theoretical and quantitative economics, with the main interests in microeconomic theory, pure and applied game theory, macroeconomic theory, mathematical finance, and also in research on the boundaries between economics, statistics and applied mathematics.

In addition, to facilitate communication between authors, reviewers and editors, we had launched a manuscript submission site through which we handled all new submissions. The site worked on an Open Journal Systems platform.

Yet, late in the year of 2015 and also in early 2016, we have observed that a growing specialization in the academia has motivated the community of prospective authors from Central and Eastern Europe to submit their work more into specialized journals than into a general interest journal of our kind. We consider this development a positive trend. The shift of interests reflects growing aspirations to contribute to truly universal academic knowledge (even if in more narrow journals), which is entirely consistent with our perspective of the field.

Facing a low number of high-quality submissions received over the last two years and deciding to not attract articles on the basis of regional focus, we have decided to complete Volume 9 and then give the journal a break.

Therefore, we here kindly ask all prospective authors to submit their work to other respectable field or general interest journals. Our journal website will be maintained in the future and all published articles as well as other lists will continue to be available in open access to all scholars and interested readers.

We would like to take this opportunity to thank all scholars who have contributed to the journal for their help, care, and genuinely professional attitude. Our main thanks belong to the previous Editors, previous and current Editorial Board members, and especially to dozens of reviewers from various universities across the world, coming mainly from departments of economics, applied mathematics, operations research, and political science.

Jozef Baruník, Martin Gregor, Milan Vlach

Czech Economic Review 9 (2015) 7–35
Acta Universitatis Carolinae Oeconomica

Climate Change, Collapse and Social Choice Theory

Norman Schofield*

Received 1 April 2015; Accepted 1 June 2015

Abstract The enlightenment was a philosophical project to construct a rational society without the need for a supreme being. It opened the way for the creation of market democracy and rapid economic growth. At the same time economic growth is the underlying cause of climate change, and we have become aware that this may destroy our civilization. The principal underpinning of the enlightenment project is the *general equilibrium theorem* (GET) of Arrow and Debreu (1954), asserting the existence of a Pareto optimal price equilibrium. Arrow's work in social choice can be interpreted as an attempt to construct a more general social equilibrium theorem. The current paper surveys recent results in social choice which suggests that chaos rather than equilibrium is generic. We also consider models of belief aggregation similar to Condorcet's Jury Theorem and mention Penn's theorem on existence of a belief equilibrium. However, it is suggested that a belief equilibrium with regard to the appropriate response to climate change depends on the creation of a fundamental social principle of "guardianship of our planetary home." It is suggested that this will involve conflict between entrenched economic interests and ordinary people, as the effects of climate change make themselves felt in many countries.

Keywords The enlightenment, climate change, black swan events, dynamical models
JEL classification H10

1. Introduction

In this essay I shall develop my earlier argument, in Schofield (2011), that human decision making is essentially *chaotic*. What I mean by this term is that contrary to general equilibrium theory (Arrow and Debreu 1954) there is no reason, in general, to assume that we can make wise choices. What Israel (2002) calls the *Radical Enlightenment* of the Eighteenth Century was based on the assumption that it was possible to establish a rational society, opposed to monarchy, religion and the church (see the other works by Israel 2006, 2010, 2014). Radical enlighteners included Thomas Jefferson, Thomas Paine and James Madison. They believed that society could be based on constitutional principles, leading to the "probability of a fit choice." Implicit in the Radical Enlightenment was the belief, originally postulated by Spinoza, that individuals could find moral bases for their choices without a need for a divine creator. An ancillary belief was that the economy would also be rational and that the principles of the Radical Enlightenment would lead to material growth and the eradication of poverty and misery. I shall argue here that this enlightenment project has recently had to face two profoundly troubling propositions. First are the results of Arrovian social choice theory. These very

* Center in Political Economy, Washington University in Saint Louis, 1 Brookings Drive, Saint Louis MO 63130, USA. Phone: 3149355630, E-mail: schofield.norman@gmail.com.

abstract results suggest that no process of social choice can be "rational," in the sense of giving rise to a state of affairs that is acceptable to all. Second, recent events suggest that the market models that we have used to guide our economic actions are deeply flawed. In contrast to the Radical enlighteners, both Hume and Burke believed that people would need religion and nationalism to provide a moral compass to their lives. As Putnam and Campbell (2010) have noted, religion today is as important as it has ever been in the U.S. Recent models of U.S. Elections (Schofield and Gallego 2011) show that religion is a key dimension of politics that divides voters one from another.

A consequence of the Industrial Revolution, that followed on from the Radical Enlightenment, has been the unintended consequence of *climate change*. Since this is the most important policy dimension that the world economy currently faces, this paper will address the question whether we are likely to be able to make wise social choices to avoid future catastrophe. To guide us in this, I believe we need a moral compass founded on religious principles.

1.1 The Radical Enlightenment

A fundamental principle of the enlightenment was the utilization of mathematics to uncover nature. Hawking and Mlodinow (2010) argue for a strong version of this universal mathmatical principle, called *model-dependent realism*, citing the recent developments in mathematical physics and cosmology.

They argue that it is only through a mathematical model that we can properly perceive reality. However, this mathematical principle faces two philosophical difficulties. One stems from the Gödel (1931)-Turing (1937) undecidability theorems. The first theorem asserts that mathematics cannot be both complete and consistent, so there are mathematical theories that in principle cannot be verified. Turing's work, though it provides the basis for our computer technology also suggests that not all programs are computable. These two results suggest that our mathematical models of climate and the economy will be fundamentally uncertain. The second problem is associated with the notion of *chaos* or *catastrophe*.

Since the early work of Garrett Hardin (1968) the "tragedy of the commons" has been recognised as a global prisoners' dilemma. In such a dilemma no agent has a motivation to provide for the collective good. In the context of the possibility of climate change, the outcome is the continued emission of greenhouses gases like carbon dioxide into the atmosphere and the acidification of the oceans. There has developed an extensive literature on the n-person prisoners' dilemma in an attempt to solve the dilemma by considering mechanisms that would induce cooperation (see for example Hardin 1971, 1982; Taylor 1976, 1982; Axelrod and Hamilton 1981; Axelrod 1981, 1984; Kreps et al. 1982; Margolis 1982).

The problem of cooperation has also provided a rich source of models of evolution, building on the early work by Trivers (1971) and Hamilton (1964, 1970). Nowak (2011) provides an overview of the recent developments. Indeed, the last twenty years has seen a growing literature on a game theoretic, or mathematical, analysis of the evolution of social norms to maintain cooperation in prisoners' dilemma like situations. Gintis (2000), for example, provides evolutionary models of the cooperation through strong reciprocity and internalization of social norms.[1] The anthropological literature provides much evidence that, from about 500,000 years ago, the ancestors of *homo sapiens* engaged in cooperative behavior, particularly in hunting and caring for offspring and the elderly.[2] On this basis we can infer that we probably do have very deeply ingrained normative mechanisms that were crucial, far back in time, for the maintenance of cooperation, and the fitness and thus survival of early hominids.[3] These normative systems will surely have been modified over the long span of our evolution.

Current work on climate change has focussed on how we should treat the future. For example Stern (2007, 2009), Collier (2010) and Chichilnisky (2009a,b) argue essentially for equal treatment of the present and the future. Dasgupta (2005) points out that how we treat the future depends on our current estimates of economic growth in the near future.

The fundamental problem of climate change is that the underlying dynamic system is extremely complex, and displays many positive feedback mechanisms (see the discussion in Schofield 2011). The difficulty can perhaps be illustrated by Figure 1. It is usual in economic analysis to focus on Pareto optimality. Typically in economic theory, it is assumed that preferences and production possibilities are generated by convex sets. However, climate change could create non-convexities. In such a case the Pareto set will exhibit stable and unstable components. Figure 1 distinguishes between a domain A, bounded by stable and unstable components P_1^s and P^u, and a second stable component P_2^s. If our actions lead us to an outcome within A, whether or not it is Paretian, then it is possible that the dynamic system generated by climate could lead to a catastrophic destruction of A itself. More to the point, our society would be trapped inside A as the stable and unstable components merged together.

Our society has recently passed through a period of economic disorder, where "black swan" events, low probability occurrences with high costs, have occurred with some regularity. Recent discussion of climate change has also emphasized so called "fat-tailed climate events" again defined by high uncertainty and cost (Weitzman 2009; Chichilnisky 2010).[4] The catastrophic change implied by Figure 1 is just such a black swan event. The point to note about Figure 1 is everything would appear normal until the evaporation of A.

Cooperation could in principle be attained by the action of a hegemonic leader such as the United States as suggested by Kindleberger (1973) and Keohane and Nye (1977). In Section 2 we give a brief exposition of the prisoners' dilemma and illustrate how hegemonic behavior could facilitate international cooperation. However, the analysis suggests that in the present economic climate, such hegemonic leadership is unlikely.

---

[1] Strong reciprocity means the punishment of those who do not cooperate.

[2] Indeed, White et al. (2009) present evidence of a high degree of cooperation among very early hominids dating back about 4MYBP (million years before the present). The evidence includes anatomical data which allows for inferences about the behavioral characteristics of these early hominids.

[3] Gintis cites the work of Robson and Kaplan (2003) who use an economic model to estimate the correlation between brain size and life expectancy (a measure of efficiency). In this context, the increase in brain size is driven by the requirement to solve complex cooperative games against nature.

[4] See also Chichilnisky and Eisenberger (2010) on other catastophic events such as collision with an asteroid.

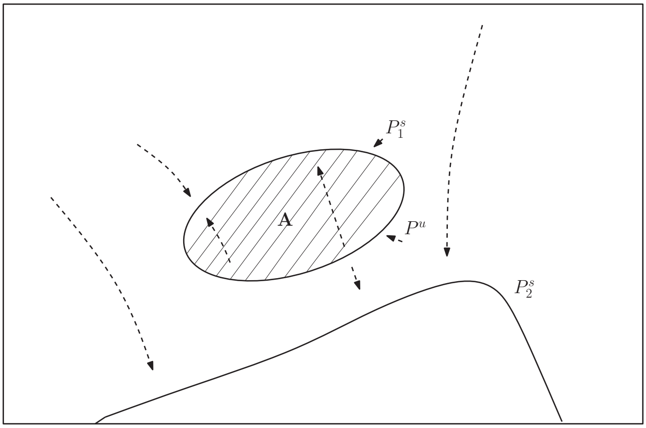


Figure 1. Stable and unstable components of the global Pareto set

Analysis of games such as the prisoner's dilemma usually focus on the existence of a Nash equilibrium, a vector of strategies with the property that no agent has an incentive to change strategy. Section 3 considers the family of equilibrium models based on the Brouwer (1912) fixed point theorem, or the more general result known as the Ky Fan theorem (Fan 1961) as well as the application by Bergstrom (1975, 1992) to prove existence of a Nash equilibrium and market equilibrium.

Section 4 considers a generalization of the Ky Fan Theorem, and argues that the general equilibrium argument can be interpreted in terms of particular properties of a preference field, H, defined on the tangent space of the joint strategy space. If this field is continuous, in a certain well-defined sense, and "half-open" then it will exhibit a equilibrium. This half-open property is the same as the non-empty intersection of a family of dual cones. We mention a theorem by Chichilnisky (1995) that a necessary and sufficient condition for market equilibrium is that a family of dual cones also has non-empty intersection.

However, preference fields that are defined in terms of coalitions need not satisfy the half-open property and thus need not exhibit equilibrium. For coalition systems, it can be shown that unless there is a collegium or oligarchy, or the dimension of the space is restricted in a particular fashion, then there need be no equilibrium. Earlier results by McKelvey (1976), Schofield (1978), McKelvey and Schofield (1987) and Saari (1997) suggested that voting can be "non-equilibriating" and indeed "chaotic" (see Schofield 1977, 1980a,b).[5]

Kauffman (1993) commented on "chaos" or the failure of "structural stability" in the following way.

> "One implication of the occurrence or non-occurrence of structural stability is that, in structurally stable systems, smooth walks in parameter

[5] In a sense these voting theorems can be regarded as derivative of Arrow's Impossiblity Theorem (Arrow 1951). See also Arrow (1986).

> space must [result in] smooth changes in dynamical behavior. By contrast, chaotic systems, which are not structurally stable, adapt on uncorrelated landscapes. Very small changes in the parameters pass through many interlaced bifurcation surfaces and so change the behavior of the system dramatically."

Chaos is generally understood as sensitive dependence on initial conditions whereas *structural stability* means that the qualitative nature of the dynamical system does not change as a result of a small perturbation.[6] I shall use the term *chaos* to mean that the trajectory taken by the dynamical process can wander anywhere.[7]

An earlier prophet of uncertainty was, of course, Keynes (1936) whose ideas on "speculative euphoria and crashes" would seem to be based on understanding the economy in terms of the qualitative aspects of its coalition dynamics (see Minsky 1975, 1986 and Keynes's earlier work in 1921). An extensive literature has tried to draw inferences from the nature of the recent economic events. A plausible account of market disequilibrium is given by Akerlof and Shiller (2009) who argue that

> "...the business cycle is tied to feedback loops involving speculative price movements and other economic activity—and to the talk that these movements incite. A downward movement in stock prices, for example, generates chatter and media response, and reminds people of longstanding pessimistic stories and theories. These stories, newly prominent in their minds, incline them toward gloomy intuitive assessments. As a result, the downward spiral can continue: declining prices cause the stories to spread, causing still more price declines and further reinforcement of the stories."

It would seem reasonable that the rise and fall of the market is due precisely to the coalitional nature of decision-making, as large sets of agents follow each other in expecting first good things and then bad. A recent example can be seen in the fall in the market after the earthquake in Japan, and then recovery as an increasing set of investors gradually came to believe that the disaster was not quite as bad as initially feared.

Since investment decisions are based on these uncertain evaluations, and these are the driving force of an advanced economy, the flow of the market can exhibit singularities, of the kind that recently nearly brought on a great depression. These singularities associated with the bursting of market bubbles are time-dependent, and can be induced by endogenous belief-cascades, rather than by any change in economic or political fundamentals (Corcos et al. 2002).

Similar uncertainty holds over political events. The fall of the Berlin Wall in 1989 was not at all foreseen. Political scientists wrote about it in terms of "belief cascades" (Karklins and Petersen 1993; Lohmann 1994; see also Bikhchandani et al. 1992) as

[6] The theory of chaos or complexity is rooted in Smale's fundamental theorem (Smale 1966) that structural stability of dynamical systems is not "generic" or typical whenever the state space has more than two dimensions.

[7] In their early analysis of chaos, Li and Yorke (1975) showed that in the domain of a chaotic transformation f it was possible for almost any pair of positions (x,y) to transition from x to $y = f^r(x)$, where f^r means the r times reiteration of f.

the coalition of protesting citizens grew apace. As the very recent democratic revolutions in the Middle East and North Africa suggest, these coalitional movements are extremely uncertain.[8] In particular, whether the autocrat remains in power or is forced into exile is as uncertain as anything Keynes discussed. Even when democracy is brought about, it is still uncertain whether it will persist (see for example Carothers 2002 and Collier 2009).

Section 5 introduces the Condorcet (1994 [1795]) Jury Theorem. This theorem suggests that majority rule can provide a way for a society to attain the truth when the individuals have common goals. Schofield (2002, 2006) has argued that Madison was aware of this theorem while writing Federalist X (Madison 1999 [1787]) so it can be taken as perhaps the ultimate justification for democracy. However, models of belief aggregation that are derived from the Jury Theorem can lead to belief cascades that bifurcate the population. In addition, if the aggregation process takes place on a network, then centrally located agents, who have false beliefs, can dominate the process (Golub and Jackson 2010).

In Section 6 we introduce the idea of a belief equilibrium, and then go on to consider the notion of "punctuated equilibrium" in general evolutionary models. Again however, the existence of an equilibrium depends on a fixed point argument, and thus on a half-open property of the "cones" by which the developmental path is modeled. This half-open property is equivalent to the existence of a social direction gradient defined everywhere. In Section 7 we introduce the notion of a "moral compass" that may provide a teleology to guide us in making wise choices for the future, by providing us with a social direction gradient. Section 8 concludes.

2. The Prisoners' Dilemma, Cooperation and Morality

> "For before constitution of Sovereign Power ... all men had right to all things; which necessarily causeth Warre." (Hobbes 2009 [1651])

Kindleberger (1973) gave the first interpretation of the international economic system of states as a "Hobbesian" prisoners' dilemma, which could be solved by a leader, or "hegemon."

> "A symmetric system with rules for counterbalancing, such as the gold standard is supposed to provide, may give way to a system with each participant seeking to maximize its short-term gain ... But a world of a few actors (countries) is not like [the competitive system envisaged by Adam Smith] ... In advancing its own economic good by a tariff, currency depreciation, or foreign exchange control, a country may worsen the welfare of its partners by more than its gain. Beggar-thy-neighbor tactics may lead to retaliation so that each country ends up in a worse position from having pursued its own gain ... This is a typical non-zero sum game, in which any

[8] The response by the citizens of these countries to the demise of Osama bin Laden on May 2, 2011, is in large degree also unpredictable.

> player undertaking to adopt a long range solution by itself will find other countries taking advantage of it . . . ”

In the 1970s, Robert Keohane and Joseph Nye (1977) rejected “realist” theory in international politics, and made use of the idea of a hegemonic power in a context of “complex interdependence” of the kind envisaged by Kindleberger. Although they did not refer to the formalism of the prisoners’ dilemma, it would appear that this notion does capture elements of complex interdependence. To some extent, their concept of a hegemon is taken from realist theory rather than deriving from the game-theoretic formalism.

The essence of the theory of hegemony in international relations is that if there is a degree of inequality in the strengths of nation states then a hegemonic power may maintain cooperation in the context of an *n*-country prisoners’ dilemma. Clearly, the British Empire in the 1800s is the role model for such a hegemon (Ferguson 2002).

Hegemon theory suggests that international cooperation was maintained after World War II because of a dominant cooperative coalition. At the core of this cooperative coalition was the United States; through its size it was able to generate collective goods for this community, first of all through the Marshall Plan and then in the context first of the post-world war II system of trade and economic cooperation, based on the Bretton Woods agreement and the Atlantic Alliance, or NATO. Over time, the United States has found it costly to be the dominant core of the coalition, in particular, as the relative size of the U.S. economy has declined. Indeed, the global recession of 2008–2010 suggests that problems of debt could induce “begger thy neighbor strategies,” just like the 1930s.

The future utility benefits of adopting policies to ameliorate these possible changes depend on the discount rates that we assign to the future. Dasgupta (2005) gives a clear exposition of how we might assign these discount rates. Obviously enough, different countries will in all likelihood adopt very different evaluations of the future. Developing countries like the BRICs (Brazil, Russia, India and China) will choose growth and development now rather than choosing consumption in the future.

An extensive literature over the last few years has developed Adam Smith’s ideas as expressed in the *Theory of Moral Sentiments* (1984 [1759]) to argue that human beings have an innate propensity to cooperate. This propensity may well have been the result of co-evolution of language and culture (Boyd and Richerson 2005; Gintis 2000).

Since language evolves very quickly (McWhorter 2001; Deutcher 2006), we might also expect moral values to change fairly rapidly, at least in the period during which language itself was evolving. In fact there is empirical evidence that cooperative behavior as well as notions of fairness vary significantly across different societies.[9] While there may be fundamental aspects of morality and “altruism,” in particular, held in common across many societies, there is variation in how these are articulated. Gazzaniga (2008) suggests that moral values can be described in terms of various *modules*: reciprocity, suffering (or empathy), hierarchy, in-group and outgroup coalition, and

[9] See Henrich et al. (2004, 2005), which reports on experiments in fifteen “small-scale societies,” using the game theoretic tools of the “prisoners’ dilemma,” the “ultimatum game,” etc.

purity/disgust. These modules can be combined in different ways with different emphases. An important aspect of cooperation is emphasized by Burkart et al. (2009) and Hrdy (2011), namely cooperation between man and woman to share the burden of child rearing.

It is generally considered that hunter-gatherer societies adopted egalitarian or "fair share" norms. The development of agriculture and then cities led to new norms of hierarchy and obedience, coupled with the predominance of military and religious elites (Schofield 2010).

North (1990), North et al. (2009) and Acemoglu and Robinson (2006) focus on the transition from such oligarchic societies to open access societies whose institutions or "rules of the game," protect private property, and maintain the rule of law and political accountability, thus facilitating both cooperation and economic development. Acemoglu et al. (2009) argue, in their historical analyses about why "good" institutions form, that the evidence is in favor of "critical junctures" (see also Acemoglu and Robinson 2008). For example, the "Glorious Revolution" in Britain in 1688 (North and Weingast 1989), which prepared the way in a sense for the agricultural and industrial revolutions to follow (Mokyr 2005, 2010; Mokyr and Nye 2007) was the result of a sequence of historical contingencies that reduced the power of the elite to resist change. Recent work by Morris (2010), Fukuyama (2011), Ferguson (2011) and Acemoglu and Robinson (2011) has suggested that these fortuitous circumstances never occurred in China and the Middle East, and as a result these domains fell behind the West. Although many states have become democratic in the last few decades, oligarchic power is still entrenched in many parts of the world.[10]

At the international level, the institutions that do exist and that are designed to maintain cooperation, are relatively young. Whether they succeed in facilitating cooperation in such a difficult area as climate change is a matter of speculation. As we have suggested, international cooperation after World War II was only possible because of the overwhelming power of the United States. In a world with oligarchies in power in Russia, China, and in many countries in Africa, together with political disorder in almost all the oil producing counties in the Middle East, cooperation would appear unlikely.

To extend the discussion, we now consider more general theories of social choice.

3. Existence of a Choice

The above discussion has considered a very simple version of the prisoner's dilemma. The more general models of cooperation typically use variants of evolutionary game theory, and in essence depend on proof of existence of Nash equilibrium, using some version of the Brouwer's fixed point theorem (Brouwer 1912).

Brouwer's theorem asserts that any continuous function $f : B \to B$ from the finite dimensional ball, B (or indeed any compact convex set in $\mathbb{R}^w$) into itself, has the *fixed point property*. That is, there exists some $x \in B$ such that $f(x) = x$.

[10] The popular protests in North Africa and the Middle East in 2011 were in opposition to oligarchic and autocratic power.

We will now consider the use of variants of the theorem, to prove existence of an equilibrium of a general choice mechanism. We shall argue that the condition for existence of an equilibrium will be violated if there are cycles in the underlying mechanism.

Let W be the set of alternatives and let X be the set of all subsets of W. A *preference correspondence*, P, on W assigns to each point $x \in W$, its *preferred set* $P(x)$. Write $P\colon W \to X$ or $P\colon W \twoheadrightarrow W$ to denote that the image of x under P is a set (possibly empty) in W. For any subset V of W, the restriction of P to V gives a correspondence $P_V : V \twoheadrightarrow V$. Define $P_V^{-1} : V \twoheadrightarrow V$ such that for each $x \in V$,

$$P_V^{-1}(x) = \{y : x \in P(y)\} \cap V.$$

The sets $P_V(x), P_V^{-1}(x)$ are sometimes called the *upper* and *lower* preference sets of P on V. When there is no ambiguity we delete the suffix V. The *choice* of P from W is the set

$$C(W,P) = \{x \in W\colon P(x) = \varnothing\}.$$

Here $\varnothing$ is the empty set. The choice of P from a subset, V, of W is the set

$$C(V,P) = \{x \in V\colon P_V(x) = \varnothing\}.$$

Call C_P a *choice function* on W if $C_P(V) = C(V,P) \neq \varnothing$ for every subset V of W. We now seek general conditions on W and P which are sufficient for C_P to be a choice function on W. Continuity properties of the preference correspondence are important and so we require the set of alternatives to be a topological space.

Definition 1. Let W, Y be two topological spaces. A correspondence $P\colon W \twoheadrightarrow Y$ is

(i) *Lower demi-continuous* (*ldc*) iff, for all $x \in Y$, the set

$$P^{-1}(x) = \{y \in W\colon x \in P(y)\}$$

is open (or empty) in W.

(ii) *Acyclic* if it is impossible to find a cycle $x_t \in P(x_{t-1})$, $x_{t-1} \in P(x_{t-2}), \ldots, x_1 \in P(x_t)$.

(iii) *Lower hemi-continuous* (*lhc*) iff, for all $x \in W$, and any open set $U \subset Y$ such that $P(x) \cap U \neq \varnothing$ there exists an open neighborhood V of x in W, such that $P(x') \cap U \neq \varnothing$ for all $x' \in V$.

Note that if P is *ldc* then it is *lhc*.

We shall use lower demi-continuity of a preference correspondence to prove existence of a choice.

We shall now show that if W is compact, and P is an acyclic and *ldc* preference correspondence $P\colon W \twoheadrightarrow W$, then $C(W,P) \neq \varnothing$. First of all, say a preference correspondence $P\colon W \twoheadrightarrow W$ satisfies the *finite maximality property* (FMP) on W iff for every finite set V in W, there exists $x \in V$ such that $P(x) \cap V = \varnothing$.

Lemma 1. (Walker 1977) *If W is a compact, topological space and P is an ldc preference correspondence that satisfies FMP on W, then $C(W,P) \neq \varnothing$.*

This follows readily, using compactness to find a finite subcover, and then using FMP.

Corollary 1. *If W is a compact topological space and P is an acyclic, ldc preference correspondence on W, then $C(W,P) \neq \varnothing$.*

As Walker (1977) noted, when W is compact and P is *ldc*, then P is acyclic iff P satisfies FMP on W, and so either property can be used to show existence of a choice. A second method of proof is to show that C_P is a choice function to substitute a convexity property for P rather than acyclicity.

Definition 2.

(i) If W is a subset of a vector space, then the *convex hull* of W is the set, $Con[W]$, defined by taking all convex combinations of points in W.

(ii) W is *convex* iff $W = Con[W]$. (The empty set is also convex.)

(iii) W is *admissible* iff W is a compact, convex subset of a topological vector space.

(iv) A preference correspondence $P\colon W \twoheadrightarrow W$ is *semi-convex* iff, for all $x \in W$, it is the case that $x \notin Con(P(x))$.

Fan (1961) has shown that if W is admissible and P is *ldc* and semi-convex, then $C(W,P)$ is non-empty.

Theorem 1 (Choice Theorem). (Fan 1961, Bergstrom 1975) *If W is an admissible subset of a Hausdorff topological vector space, and $P\colon W \twoheadrightarrow W$ a preference correspondence on W which is ldc and semi-convex then $C(W,P) \neq \varnothing$.*

The proof uses the KKM lemma due to Knaster, Kuratowski and Mazurkiewicz (1929).

The original form of the Fan Theorem made the assumption that $P\colon W \twoheadrightarrow W$ was *irreflexive* (in the sense that $x \notin P(x)$ for all $x \in W$) and *convex*. Together these two assumptions imply that P is semi-convex. Bergstrom (1975) extended Fan's original result to give the version presented above (see also Shafer and Sonnenschein 1975).

Note that the Fan Theorem is valid without restriction on the dimension of W. Indeed, Aliprantis and Brown (1983) have used this theorem in an economic context with an infinite number of commodities to show existence of a price equilibrium. Bergstrom (1992) also showed that when W is finite dimensional then the Fan Theorem is valid when the continuity property on P is weakened to *lhc* and used this theorem to show existence of a Nash equilibrium of a game $G = \{(P_1,W_1),\ldots,P_i,W_i),\ldots,(P_n,W_n) : i \in N\}$. Here the i^{th} stategy space is finite dimensional W_i and each individual has a preference P_i on the joint strategy space $P_i : W^N = W_1 \times W_2 \ldots \times W_n \twoheadrightarrow W_i$. The Fan Theorem can be used, in principle to show existence of an equilibrium in complex economies with externalities. Define the Nash improvement correspondence by $P_i^* : W^N \twoheadrightarrow W^N$ by $y \in P_i^*(x)$ whenever $y = (x_1,\ldots,x_{i-1},x_i^*,\ldots,x_n)$, $x = (x_1,\ldots,x_{i-1},x_i,\ldots,x_n)$, and $x_i^* \in P_i(x)$. The joint Nash improvement correspondence is $P_N^* = \cup P_i^* : W^N \twoheadrightarrow W^N$. The Nash equilibrium of a game G is a vector $\mathbf{z} \in W^N$ such that $P_N^*(\mathbf{z}) = \varnothing$. Then the Nash equilibrium will exist when P_N^* is *ldc* and semi-convex and W^N is admissible.

4. Dynamical Choice Functions

We now consider a *generalized preference field* $H : W \twoheadrightarrow TW$, on a manifold W. TW is the tangent bundle above W, given by $TW = \cup\{T_xW : x \in W\}$, where T_xW is the tangent space above x. If V is a neighborhood of x, then $T_VW = \cup\{T_xW : x \in V\}$ which is locally like the product space $\mathbb{R}^w \times V$. Here W is locally like $\mathbb{R}^w$.

At any $x \in W$, $H(x)$ is a *cone* in the tangent space T_xW above x. That is, if a vector $v \in H(x)$, then $\lambda v \in H(x)$ for any $\lambda > 0$. If there is a smooth curve, $c : [-1,1] \to W$, such that the differential $\frac{dc(t)}{dt} \in H(x)$, whenever $c(t) = x$, then c is called an *integral curve* of H. An integral curve of H from $x = c(o)$ to $y = \lim_{t\to 1} c(t)$ is called an *H-preference curve* from x to y. In this case we write $y \in \mathbb{H}(x)$. We say y is reachable from x if there is a piecewise differentiable $H-$preference curve from x to y, so $y \in \mathbb{H}^r(x)$ for some reiteration r. The preference field H is called *S-continuous* iff the inverse relation $\mathbb{H}^{-1}$ is *ldc*. That is, if x is reachable from y, then there is a neighborhood V of y such that x is reachable from all of V. The *choice* $C(W,H)$ of H on W is defined by

$$C(W,H) = \{x \in W : H(x) = \varnothing\}.$$

Say $H(x)$ is semi-convex at $x \in W$, if either $H(x) = \varnothing$ or $0 \notin Con[H(x)]$ in the tangent space T_xW. In the later case, there will exist a vector $v' \in T_xW$ such that $(v' \cdot v) > 0$ for all $v \in H(x)$. We can say in this case that there is, at x, a *direction gradient d* in the cotangent space T_x^*W of linear maps from T_xW to $\mathbb{R}$ such that $d(v) > 0$ for all $v \in H(x)$. If H is S-continuous and half-open in a neighborhood, V, then there will exist such a continuous direction gradient $d : V \to T^*V$ on the neighborhood V.[11]

We define

$$Cycle(W,H) = \{x \in W : H(x) \neq \varnothing, 0 \in ConH(x)\}.$$

Definition 3. The *dual* of a preference field $H : W \twoheadrightarrow TW$ is defined by $H^* : W \twoheadrightarrow T^*W : x \twoheadrightarrow \{d \in T_x^*W : d(v) > 0 \text{ for all } v \in H(x) \subset T_xW\}$. For convenience if $H(x) = \varnothing$ we let $H^*(x) = T_xW$. Note that if $0 \notin ConH(x)$ iff $H^*(x) \neq \varnothing$. We can say in this case that the field is *half-open* at x.

In applications, the field $H(x)$ at x will often consist of some family $\{H_j(x)\}$. As an example, let $u : W \twoheadrightarrow \mathbb{R}^n$ be a smooth utility profile and for any coalition $M \subset N$ let

$$H_M(u)(x) = \{v \in T_xW : du_i(x)(v) > 0,\ \forall i \in M\}.$$

If $\mathbb{D}$ is a family of *decisive* coalitions, $\mathbb{D} = \{M \subset N\}$, then we define

$$H_{\mathbb{D}}(u) = \cup H_M(u) : W \twoheadrightarrow TW.$$

Then the field $H_{\mathbb{D}}(u) : W \twoheadrightarrow TW$ has a dual $[H_{\mathbb{D}}(u)]^* : W \twoheadrightarrow T^*W$ given by $[H_{\mathbb{D}}(u)]^*(x) = \cap[H_M(u)(x)]^*$ where the intersection at x is taken over all $M \in \mathbb{D}$ such that $H_M(u)(x) \neq \varnothing$. We call $[H_M(u)(x)]^*$ the *co-cone of* $[H_M(u)(x)]^*$. It then follows that at $x \in$

[11] i.e. $d(x)(v) > 0$ for all $x \in V$, for all $v \in H(x)$, whenever $H(x) \neq \varnothing$.

$Cycle(W, H_{\mathbb{D}}(u))$ then $0 \in Con[H_{\mathbb{D}}(u)(x)]$ and so $[H_{\mathbb{D}}(u)(x)]^* = \varnothing$. Thus

$$Cycle(W, H_{\mathbb{D}}(u)) = \{x \in W : [H_{\mathbb{D}}(u)]^*(x) = \varnothing\}.$$

The condition that $[H_{\mathbb{D}}(u)]^*(x) = \varnothing$ is equivalent to the condition $\cap[H_M(u)(x)]^* = \varnothing$ and was called the *null dual condition* (at x). Schofield (1978) has shown that $Cycle(W, H_{\mathbb{D}}(u))$ will be an open set and contains cycles so that a point x is reachable from itself through a sequence of preference curves associated with different coalitions. This result was an application of a more general result.

Theorem 2 (Dynamical Choice Theorem). (Schofield 1978) *For any S-continuous field H on compact, convex W, then*

$$Cycle(W, H) \cup C(W, H) \neq \varnothing.$$

Proof. If $x \in Cycle(W, H) \neq \varnothing$ then there is a *piecewise differentiable H-preference cycle* from x to itself. If there is an open path connected neighborhood $V \subset Cycle(W, H)$ such that $H(x')$ is open for all $x' \in V$ then there is a *piecewise differentiable H-preference curve* from x to x'. □

(Here piecewise differentiable means the curve is continuous, and also differentiable except at a finite number of points.) The proof follows from the previous choice theorem. The trajectory is built up from a set of vectors $\{v_1, \ldots, v_t\}$ each belonging to $H(x)$ with $0 \in Con[\{v_1, \ldots, v_t\}]$. If $H(x)$ is of full dimension, as in the case of a voting rule, then just as in the model of chaos by Li and York(1975), trajectories defined in terms of H can wander anywhere within any open path connected component of $Cycle(W, H)$. Chichilnisky (1997a) has obtained an analogous result.

Theorem 3 (Chichilnisky Theorem). (Chichilnisky 1997a) *The limited arbitrage condition $\cap[H_i(u)]^* \neq \varnothing$ is necessary and sufficient for existence of a competitive equilibrium.*

Chichilnisky (1993, 1997b) also defined a topological obstruction to the non-emptiness of this intersection and showed the connection with the existence of a social choice equilibrium.

For a voting rule, $\mathbb{D}$ it is possible to guarantee that $Cycle(W, H_{\mathbb{D}}) = \varnothing$ and thus that $C(W, H_{\mathbb{D}}) \neq \varnothing$. We can do this by restricting the dimension of W.

Extending the equilibrium result of the Nakamura Theorem (Nakamura 1979) to higher dimension for a voting rule faces a difficulty caused by Bank's Theorem. We first define a *fine* topology on smooth utility functions (Hirsch 1976; Schofield 1999, 2003).

Definition 4. Let $(U(W)^N, T_1)$ be the topological space of smooth utility profiles endowed with the the C^1-topology.

In economic theory, the existence of isolated price equilibria can be shown to be "generic" in this topological space (Debreu 1970, 1976; Smale 1974a,b). In social choice no such equilibrium theorem holds. The difference is essentially because of the coalitional nature of social choice.

Theorem 4 (Banks Theorem). *For any non-collegial $\mathbb{D}$, there exists an integer $w(\mathbb{D})$ such that $dim(W) > w(\mathbb{D})$ implies that $C(W, H_{\mathbb{D}}(u)) = \varnothing$ for all u in a dense subspace of $(U(W)^N, T_1)$ so $Cycle(W, H_{\mathbb{D}}(u)) \neq \varnothing$ generically.*

Proof. This result was essentially proved by Banks (1995), building on earlier results by Plott (1967), Kramer (1973), McKelvey (1979), Schofield (1983a,b), McKelvey and Schofield (1987). See for related analyses. Indeed, it can be shown that if $dim(W) > w(\mathbb{D})+1$ then $Cycle(W, H_{\mathbb{D}}(u))$ is generically dense (Schofield 1984). The integer $w(\mathbb{D})$ can usually be computed explicitly from $\mathbb{D}$. For majority rule with n odd it is known that $w(\mathbb{D}) = 2$ while for n even, $w(\mathbb{D}) = 3$. □

Although the Banks Theorem formally applies only to voting rules, Schofield (2010) argues that it is applicable to any non-collegial social mechanism, say $H(u)$ and can be interpreted to imply that

$$Cycle(W, H(u)) \neq \varnothing \text{ and } C(W, H(u)) = \varnothing$$

is a generic phenomenon in coalitional systems. Because preference curves can wander anywhere in any open component of $Cycle(W, H(u))$, Schofield (1979) called this *chaos*. It is not so much the sensitive dependence on initial conditions, but the aspect of indeterminacy that is emphasized. On the other hand, existence of a hegemon, as discussed in Section 2, suggests that $Cycle(W, H(u))$ would be constrained in this case.

Richards (1990) has examined data on the distribution of power in the international system over the long run and and presents evidence that it can be interpreted in terms of a chaotic trajectory. This suggests that the metaphor of the nPD in international affairs does characterise the ebb and flow of the system and the rise and decline of hegemony.

It is worth noting that the early versions of the Banks Theorem were obtained in the decade of the 1970s, a decade that saw the first oil crisis, the collapse of the Bretton Woods system of international political economy, the apparent collapse of the British economy, the beginning of social unrest in Eastern Europe, the revolution in Iran, and and the second oil-crisis (Caryl 2011). Many of the transformations that have occurred since then can be seen as changes in beliefs, rather than preferences. Models of belief aggregation are less well developed than those dealing with preferences.[12] In general models of belief aggregation are related to what is now termed Condorcet's Jury Theorem, which we now introduce.

5. Beliefs and Condorcet's Jury Theorem

The Jury Theorem formally only refers to a situation where there are just two alternatives $\{1, 0\}$, and alternative 1 is the "true" option. Further, for every individual, i, it is the case that the probability that i picks the truth is ρ_{i1}, which exceeds the probability, ρ_{i0}, that i does not pick the truth. We can assume that $\rho_{i1} + \rho_{i0} = 1$, so obviously $\rho_{i1} > \frac{1}{2}$. To simplify the proof, we can assume that ρ_{i1} is the same for every individual, thus $\rho_{i1} = \alpha > \frac{1}{2}$ for all i. We use $\chi_i (= 0 \text{ or } 1)$ to refer to the choice of individual i,

[12] Results on belief aggregation include Penn (2009) and McKelvey and Page (1986).

and let $\chi = \Sigma_{i=1}^{n} \chi_i$ be the number of individuals who select the true option 1. We use Pr for the probability operator, and E for the expectation operator. In the case that the electoral size, n, is odd, then a majority, m, is defined to be $m = \frac{n+1}{2}$. In the case n is even, the majority is $m = \frac{n}{2} + 1$. The probability that a majority chooses the true option is then

$$\alpha_{maj}^{n} = \Pr[\chi \geq m].$$

The theorem assumes that voter choice is *pairwise independent*, so that $\Pr(\chi = j)$ is simply given by the binomial expression $\binom{n}{j} \alpha^j (1-\alpha)^{n-j}$.

A version of the theorem can be proved in the case that the probabilities $\{\rho_{i1} = \alpha_i\}$ differ but satisfy the requirement that $\frac{1}{n}\Sigma_{i=1}^{n} \alpha_i > \frac{1}{2}$. Versions of the theorem are valid when voter choices are not pairwise independent (Ladha and Miller 1996).

Theorem 5 (The Jury Theorem). *If* $1 > \alpha > \frac{1}{2}$, *then* $\alpha_{maj}^{n} \geq \alpha$, *and* $\alpha_{maj}^{n} \to 1$ *as* $n \to \infty$.

Proof. For both n being even or odd, as $n \to \infty$, the fraction of voters choosing option 1 approaches $\frac{1}{n}E(\chi) = \alpha > \frac{1}{2}$. Thus, in the limit, more than half the voters choose the true option. Hence the probability $\alpha_{maj}^{n} \to 1$ as $n \to \infty$. □

Laplace also wrote on the topic of the probability of an error in the judgement of a tribunal. He was concerned with the degree to which jurors would make just decisions in a situation of asymmetric costs, where finding an innocent party guilty was to be more feared than letting the guilty party go free. As he commented on the appropriate rule for a jury of twelve, “I think that in order to give a sufficient guarantee to innocence, one ought to demand at least a plurality of nine votes in twelve” (Laplace 1951[1814], p. 139). Schofield (1972a,b) considered a model derived from the Jury Theorem where uncertain citizens were concerned to choose an ethical rule which would minimize their disappointment over the the likely outcomes. He showed that majority rule was indeed optimal in this sense.

Models of belief aggregation extend the Jury Theorem by considering a situation where individuals receive signals, update their beliefs and make an aggregate choice on the basis of their posterior beliefs (Austen-Smith and Banks 1996). Models of this kind can be used as the basis for analysing correlated beliefs.[13] and the creation of belief cascades (Easley and Kleinberg 2010).

Schofield (2002, 2006) has argued that Condorcet’s Jury Theorem provided the basis for Madison’s argument in Federalist X (Madison 1999 [1787]) that the judgments of citizens in the extended Republic would enhance the “probability of a fit choice.” However, Schofield’s discussion suggests that belief cascades can also fracture the society in two opposed factions, as in the lead up to the Civil War in 1860.[14]

There has been a very extensive literature recently on cascades (Gleick 1987; Buchanan 2001, 2003; Gladwell 2002; Johnson 2002; Barabasi 2003, 2010; Strogatz 2004; Watts 2002, 2003; Surowiecki 2005; Ball 2004; Christakis and Fowler 2011),

[13] Schofield 1972 a,b; Ladha 1992, 1993; Ladha and Miller 1996.

[14] Sunstein (2006, 2011) also notes that belief aggregation can lead to a situation where subgroups in the society come to hold very disparate opinions.

but it is unclear from this literature whether cascades will be equilibriating or very volatile. In their formal analysis of cascades on a network of social connections, Golub and Jackson (2010) use the term *wise* if the process can attain the truth. In particular they note that if one agent in the network is highly connected, then untrue beliefs of this agent can steer the crowd away from the truth. The recent economic disaster has led to research on market behavior to see if the notion of cascades can be used to explain why markets can become volatile or even irrational in some sense (Acemoglu et al. 2010; Schweitzer et al. 2009). Indeed the literature that has developed in the last few years has dealt with the nature of herd instinct, the way markets respond to speculative behavior and the power law that characterizes market price movements (see, for example, Mandelbrot and Hudson 2004; Shiller 2003, 2005; Taleb 2007; Barbera 2009; Cassidy 2009; Fox 2009). The general idea is that the market can no longer be regarded as efficient. Indeed, as suggested by Ormerod (2001) the market may be fundamentally chaotic.

"Empirical" chaos was probably first discovered by Lorenz (1962, 1963) in his efforts to numerically solve a system of equations representative of the behavior of weather. A very simple version is the non-linear vector equation

$$\frac{dx}{dt} = \begin{bmatrix} dx_1 \\ dx_2 \\ dx_3 \end{bmatrix} = \begin{bmatrix} -a_1(x_1 - x_2) \\ -x_1x_3 + a_2x_1 - x_2 \\ x_1x_2 - a_3x_3 \end{bmatrix},$$

which is chaotic for certain ranges of the three constants, a_1, a_2, a_3.

The resulting "butterfly" portrait winds a number of times about the left hole (as in Figure 2), then about the right hole, then the left, etc. Thus the "phase portrait" of this dynamical system can be described by a sequence of winding numbers $(w_l^1, w_k^1, w_l^2, w_k^2,$ etc.). Changing the constants a_1, a_2, a_3 slightly changes the winding numbers. Note that Figure 2 is in three dimensions. The butterfly wings on left and right consist of infinitely many closed loops. An illustration of the butterfly is the chaotic trajectory of the Artemis Earth Moon orbiter (which can be found at nasa.gov-artemis orbiter). The butterfly is also called the Lorentz "strange attractor."A slight perturbation of this dynamic system changes the winding numbers and thus the qualitative nature of the process. Clearly this dynamic system is not structurally stable, in the sense used by Kauffmann (1993). The metaphor of the butterfly gives us pause, since all dynamic systems whether models of climate, markets, voting processes or cascades may be indeterminate or chaotic.

6. The Edge of Chaos

A *dynamic belief equilibrium* at τ for a society N_τ is a fixed point of a transformation in the beliefs of the society. Although the space will be infinite dimensional, if the domain and range of this transformation are restricted to *equicontinous* functions (Pugh 2002), then the domain and range will be compact. Penn (2009) shows that if the domain and range are convex then a generalized version of Brouwer's fixed point theorem can be applied to show existence of such a dynamic belief equilibrium. This notion of

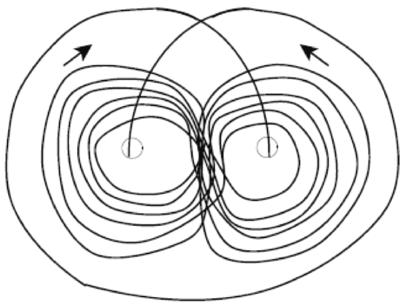


Figure 2. The butterfly

equilibrium was first suggested by Hahn (1973) who argued that equilibrium is located in the mind, not in behavior.

However, the choice theorem suggests that the validity of Penn's result will depend on how the model of social choice is constructed. For example Corcos et al. (2002) consider a formal model of the market, based on the reasoning behind Keynes's "beauty contest" (Keynes 1936). There are two coalitions of "bulls" and "bears." Individuals randomly sample opinion from the coalitions and use a *critical* cutoff-rule. For example if the individual is bullish and the sampled ratio of bears exceeds some proportion then the individual flips to bearish. The model is very like that of the Jury Theorem but instead of guaranteeing a good choice the model can generate chaotic flips between bullish and bearish markets, as well as fixed points or cyclic behavior, depending on the cut-off parameters. Taleb's argument (Taleb 2007) about black swan events can be applied to the recent transformation in societies in the Middle East and North Africa that resemble such a cascade (Taleb and Blyth 2011). As in the earlier episodes in Eastern Europe, it would seem plausible that the sudden onset of a cascade is due to a switch in a critical coalition.

The notion of "criticality" has spawned in enormous literature particularly in fields involving evolution, in biology, language and culture (see for example Cavallli-Sforza and Feldman 1981; Bowles et al. 2003). Bak and Sneppen (1993) refer to the self organized critical state as the

> "… 'edge of chaos' since it separates a frozen inactive state from a 'hot' disordered state … The mechanism of evolution in the critical state can be thought of as an exploratory search for better local fitness, which is rarely successful, but sometimes has enormous effect on the ecosystem."

Flyvbjerg et al. (1993) go on to say

> "… species sit at local fitness maxima … and occasionally a species jumps to another maximum [in doing so it] may change the fitness landscapes of other species which depend on it … Consequently they immediately jump

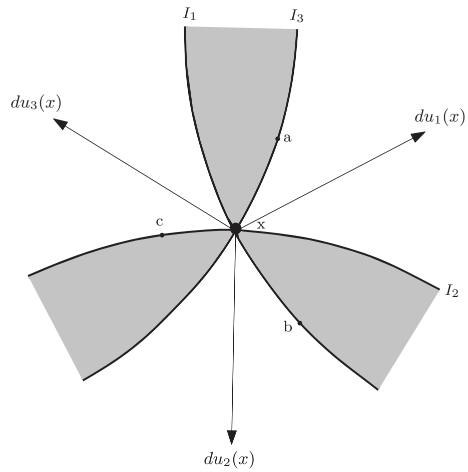


Figure 3. Cycles in a neighborhood of *x*

> to new maxima. This may affect yet another species in a chain reaction, a *burst* of evolutionary activity."

This work was triggered by the earlier ideas on "punctuated equilibrium" by Eldredge and Gould (1972) (see also Eldredge 1976; Gould 1976).

The point to be emphasized is that the evolution of a species involves bifurcation, a splitting of the pathway. We can refer to the bifurcation as a *catastrophe* or a *singularity*. The portal or door to the singularity may well be characterized by chaos or uncertainty, since the path can veer off in many possible directions, as suggested by the bifurcating cones in Figures 3 and 4. At every level that we consider, the bifurcations of the evolutionary trajectory seem to be locally characterized by chaotic domains. I suggest that these domains are the result of different coalitional possibilities. The fact that the trajectories can become indeterminate suggests that this may enhance the exploration of the fitness landscape.

A more general remark concerns the role of climate change. Climate has exhibited chaotic or catastrophic behavior in the past.[15] There is good reason to believe that human evolution over the last million years can only be understood in terms of "bursts" of sudden transformations (Nowak 2011) and that language and culture co-evolve through group or coalition selection (Cavallli-Sforza and Feldman 1981). Calvin (2003) suggests that our braininess was cause and effect of the rapid exploration of the fitness landscape in response to climatic forcing. For example Figure 2 in the earlier paper (Schofield 2011) showed the rapid changes in temperature over the last 100,000 years. It was only in the last period of stable temperature, the "holocene," during the last 10,000 years that agriculture was possible. One danger of the current climate change

[15] Indeed as I understand the dynamical models, the chaotic episodes are due to the complex interactions of dynamical processes in the oceans, on the land, in weather, and in the heavens. These are very like interlinked *coalitions* of non-gradient vector fields.

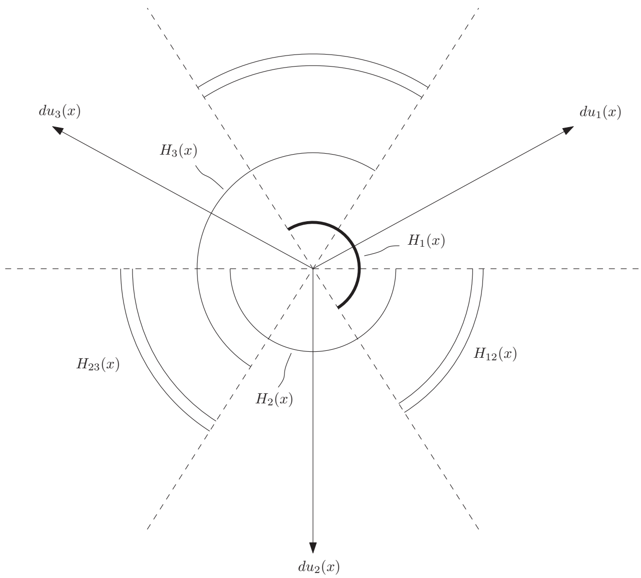


Figure 4. The failure of half-openness of a preference field

is not just the possibility of a rise in temperature but that climate itself could become chaotic, destroying the possibility of agriculture and causing the collapse of our civilization.

Stringer (2012) calls the theory of rapid evolution during a period of chaotic climate change "the Social Brain hypothesis." The cave art of Chauvet, in France dating back about 36,000 years suggests that belief in the supernatural played an important part in human evolution. Indeed, we might speculate that the part of our mind that enhances technological/mathematical development and that part that facilitates social/religious belief are in conflict with each other (this is suggested by Kahneman 2011). We might also speculate that market behavior is largely driven by what Keynes termed *speculation*, namely the largely irrational changes of *mood* (Casti 2010). Figure 1 in the earlier paper (Schofield 2011) gave an illustration of the swings in the U.S. stock market over the last 80 years. This figure certainly suggested that the stock market does not tend to equilibriate.

7. A Moral Compass

If we accept that moral and religious beliefs are as important as rational calculations in determining the choices of society, then depending on models of preference aggregation will not suffice in helping us to make decisions over how to deal with climate change. Instead, I suggest a moral compass, derived from current inferences made about the nature of the evolution of intelligence on our planetary home. The anthropic principle reasons that the fundamental constants of nature are very precisely tuned so that the universe contains matter and that galaxies and stars live long enough to allow for the creation of carbon, oxygen etc, all necessary for the evolution of life itself.[16] Gribbin (2011) goes further and points out that not only is the sun unusual in having the characteristics of a structurally stable system of planets, but the earth is fortunate in being protected by Jupiter from chaotic bombardment but the Moon also stabilizes our planet's orbit.[17] In essence Gribbin gives good reasons to believe that our planet may well be the only planet in the galaxy that sustains intelligent life. If this is true then we have a moral obligation to act as guardians of our planetary home. Parfit (2011) argues

> "What matters most is that we rich people give up some of our luxuries, ceasing to overheat the Earth's atmosphere, and taking care of this planet in other ways, so that it continues to support intelligent life. If we are the only rational animals in the Universe, it matters even more whether we shall have descendants during the billions of years in which that would be possible. Some of our descendants might live lives and create worlds that, though failing to justify past suffering, would give us all, including those who suffered, reason to be glad that the Universe exists." (Parfit 2011, p. 419)

8. Collapse

In the aftermath of the Great Recession, many authors have argued that the institutions that served the west as it industrialized are no longer effective (Ferguson 2012; Oreskes and Conway 2014). An earlier argument by Tainter and Renfrew (1988) on the basis of a review of the anthropological and archeological literature on the collapse of complex societies is that all such societies develop increasing complex institutions, and that complexity itself induces increasing marginal cost. Without any doubt the institutions of capitalism have become more complex over time. Such complexity can be seen in the Limits to Growth models of Meadows et al. (1972, 2012). If we regard "complexity" as the "rules of the game" then it is certainly plausible that the behavior of such a game will be located at the edge of chaos, as suggested above and thus subject to catastrophe. The logic of this theory is that we face the collapse of the American hegemony, with the end of the period of cheap energy and resources. It is also possible that China will become the new hegemon. China has been able to grow rapidly, benefitting from the positive marginal advantages of western economic institutions. Although China faces many problems associated with population, urbanization and environmental degradation, there are also indications that it is concerned to devise entirely new institutions that may help it to continue to prosper. Indeed it is possible that the continued development of China will usher in a completely new world order

[16] As Smolin (2007) points out, the anthropic principle has been adopted because of the experimental evidence that the expansion of the universe is accelerating. Indeed it has led to the hypothesis that there is an infinity of universes all with different laws. An alternative inference is the the principle of intelligent design. My own inference is that we require a teleology as proposed in the conclusion.

[17] The work by Poincare in the late 19th century focussed on the structural stability of the solar system and was the first to conceive of the notion of chaos.

that will be entirely different from the system of nation-states that developed from the post-enlightenment dominance of the West. It has also been suggested that the agricultural revolution that occurred at the beginning of the Holocene was accompanied by an ideological revolution associated with a belief in our ability to manipulate Nature for our own ends (Seddon 2014). This ideology can be seen as a precursor to enlightenment beliefs. Perhaps we need a new system of morality based on post-Holocene virtues appropriate to the age we live in rather than to enlightenment "rationality."

9. Conclusion

Even if we believe that markets are well behaved, there is no reason to infer that markets are able to reflect the social costs of the externalities associated with production and consumption. Indeed Gore (2006) argues that the globalized market place, what he calls *Earth Inc* has the power and inclination to maintain business as usual. If this is so, then climate change will undoubtedly have dramatic adverse effects, not least on the less developed countries of the world.[18]

In principle we may be able to rely on a version of the Jury Theorem (Rae 1960; Schofield 1972a,b; Sunstein 2009), which asserts that majority rule provides an optimal procedure for making collective choices under uncertainty. However, for the operation of what Madison called a "fit choice" it will be necessary to overcome the entrenched power of capital. Although we now disregard Marx's attempt at constructing a teleology of economic and political development,[19] we are in need of a more complex over-arching and evolutionary theory of political economy, embodying a system of morality that will go beyond the notion of equilibrium and might help us deal with the future.[20]

Acknowledgment This paper is based on research supported by NSF grant 0715929.

References

Acemoglu, D. and Robinson, J. (2006). *Economic origins of dictatorship and democracy.* Cambridge, Cambridge University Press.

Acemoglu, D. and Robinson, J. (2008). Persistence of power, elites, and institutions. *American Economic Review*, 98, 267–293.

Acemoglu, D. and Robinson, J. (2011). *Why Nations Fail.* London, Profile Books.

[18] Zhang et al. (2007) and Hsiang et al. (2013) have provided quantitative analyses of such adverse effects in the past. See also Parker (2013) for an historical account of the effect of climate change in early modern Europe.

[19] See Sperber (2014) for a discussion of the development of Marx's ideas, in the context of 19th century belief in the teleology of "progress" or the advance of civilization. The last hundred years has, however, made it difficult to hold such beliefs.

[20] The philosopher Nagel (2012) argues that without a teleology of some kind, we are left with Darwinian evolutionary theory, which by itself cannot provide a full explanation of what we are and where we are going. See also Taylor (2007) and Bellah (2011).

Acemoglu, D., Johnson, S., Robinson, J. and Yared, P. (2009). Reevaluating the modernization hypothesis. *Journal of Monetary Economics*, 56, 1043–1058.

Acemoglu, D., Ozdaglar, A. and Tahbaz-Salehi, A. (2010). Cascades in networks and aggregate volatility. Cambridge MA, The National Bureau of Economic Research, NBER Working Paper No. 16516.

Akerlof, G. A. and Shiller, R. J. (2009). *Animal Spirits*. Princeton, Princeton University Press.

Aliprantis, C. D. and Brown, D. J. (1983). Equilibria in markets with a Riesz space of commodities. *Journal of Mathematical Economics*, 11, 189–207.

Arrow, K. J. (1951). *Social Choice and Individual Values*. New Haven, Yale University Press.

Arrow, K. (1986). Rationality of self and others in an economic system. *Journal of Business*, 59, S385–S399.

Arrow, K. and Debreu, G. (1954). Existence of an equilibrium for a competitive economy. *Econometrica*, 22, 265–90.

Austen-Smith, D. and Banks, J. (1996). Information Aggregation, Rationality, and the Condorcet Jury Theorem. *American Political Science Review*, 90, 34–45.

Axelrod, R. (1981). The emergence of cooperation among egoists. *American Political Science Review,* 75, 306–318.

Axelrod, R. (1984). *The evolution of cooperation*. New York, Basic.

Axelrod, R. and Hamilton, W. D. (1981). The evolution of cooperation. *Science*, 211, 1390–1396.

Bak, P. and Sneppen, K. (1993). Punctuated equilibrium and criticality in a simple model of evolution. *Physics Review Letters*, 71(24), 4083–4086.

Ball, P. (2004). *Critical Mass*. New York, Ferrar, Strauss and Giroux.

Banks, J. S. (1995). Singularity theory and core existence in the spatial model. *Journal of Mathematical Economics*, 24, 523–536.

Barabasi, A.-L. (2003). *Linked*. New York, Plume.

Barabasi, A.-L. (2010). *Bursts*. New York, Dutton.

Barbera, R. (2009). *The Cost of Capitalism: Understanding Market Mayhem*. New York, McGraw Hill.

Bellah, R. N. (2011) *Religion in Human Evolution*. Cambridge MA, Belknap Press.

Bergstrom, T. (1975). The existence of maximal elements and equilibria in the absence of transitivity. Typescript, University of Michigan.

Bergstrom, T. (1992). When non-transitive relations take maxima and competitive equilibrium can't be beat. In Neuefeind, W. and Riezman, R. (eds.), *Economic Theory and International Trade.* Berlin, Springer.

Bikhchandani, S., Hirschleifer, D. and Welsh, I. (1992). A theory of fads, fashion, custom, and cultural change as information cascades. *Journal of Political Economy*, 100, 992–1026.

Bowles, S., Choi, J.-K. and Hopfensitz, A. (2003). The co-evolution of individual behaviors and social institutions. *Journal of Theoretical Biology*, 223, 135–147.

Boyd, J. and Richerson, P. J. (2005). *The Origin and Evolution of Culture*. Oxford, Oxford University Press.

Brouwer, L. E. J. (1912). Uber abbildung von mannigfaltikeiten. *Math Analen*, 71, 97–115.

Buchanan, M. (2001). *Ubiquity*. New York, Crown.

Buchanan, M. (2003). *Nexus*. New York, Norton.

Burkhart, J. M., Hrdy, S. B. and van Schaik, C. P. (2009). Cooperative breeding and human cognitive evolution. *Evolutionary Anthropology*, 18, 175–186.

Calvin, W. H. (2003). *The Ascent of Mind*. New York, Bantam.

Carothers, T. (2002). The end of the transition paradigm. *Journal of Democracy*, 13, 5–21.

Caryl, C. (2011). *Strange Rebels*. New York, Basic.

Cassidy, J. (2009). *How Markets Fail: The Logic of Economic Calamities*. New York, Farrar, Strauss and Giroux.

Casti, J. L. (2010). *Mood Matters*. New York, Copernicus.

Cavallli-Sforza, L. and Feldman, M. (1981). *Cultural Transmission and Evolution*. Princeton, NJ, Princeton University Press.

Chichilnisky, G. (1993). Intersecting families of sets and the topology of cones in economics. *Bulletin of the American Mathematical Society,* 29, 189–207.

Chichilnisky, G. (1995). Limited arbitrage is necessary and sufficient for the existence of a competitive equilibrium with or without short sales. *Economic Theory,* 5, 79–107.

Chichilnisky, G. (1997a). Limited arbitrage is necessary and sufficient for the existence of a equilibrium. *Journal of Mathematical Economics*, 28, 470–479.

Chichilnisky, G. (1997b). Market arbitrage, social choice and the core. *Social Choice and Welfare*, 14, 161–198.

Chichilnisky, G. (2009a). The topology of fear. *Journal of Mathematical Economics*, 45, 807–816.

Chichilnisky, G. (2009b). Avoiding extinction: Equal treatment of the present and the future. LAMETA, University of Montpellier, Working Paper No. 09-07.

Chichilnisky, G. (2010). The foundations of statistics with black swans. *Mathematical Social Science*, 59, 184–192.

Chichilnisky, G. and Eisenberger, P. (2010). Asteroids: Assessing catastrophic risks. *Journal of Probability and Statistics*, Article ID 954750.

Christakis, N. and Fowler, J. H. (2011). *Connected*. New York, Back Bay.

Collier, P. (2009). *Wars, Guns and Votes*. New York, Harper.

Collier, P. (2010). *The Plundered Planet.* Oxford, Oxford University Press.

Condorcet, N. (1994 [1785]). Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix. Paris, Imprimerie Royale. Translated in part in McLean, I. and Hewitt, F. (eds.), *Condorcet: Foundations of Social Choice and Political Theory.* Aldershot UK, Edward Elgar Publishing.

Corcos, A., Eckmann, J.-P., Malaspinas, A., Malevergne, Y. and Sornette, D. (2002). Imitation and contrarian behavior: Hyperbolic bubbles, crashes and chaos. *Quantitative Finance*, 2, 264–281.

Dasgupta, P. (2005). Three conceptions of intergenerational Justice. In Lillehammer, H. and Mellor, D. H. (eds.), *Ramsey's Legacy*. Oxford, Clarendon Press.

Debreu, G. (1970). Economies with a finite number of equilibria. *Econometrica*, 38, 387–392.

Debreu, G. (1976). The application to economics of differential topology and global analysis: regular differentiable economies. *American Economic Review*, 66, 280–287.

Deutscher, G. (2006). *The Unfolding of Language*. New York, Holt.

Easley, D. and Kleinberg, J. (2010). *Networks, crowds and markets*. Cambridge, Cambridge University Press.

Eldredge, N. (1976). Differential evolutionary rates, *Paleobiology*, 2, 174–177.

Eldredge, N. and Gould, S. J. (1972). Punctuated Equilibrium. In Schopf, T. (ed.), *Models of Paleobiology*. New York, Norton.

Fan, K. (1961). A generalization of Tychonoff's fixed point theorem. *Math Annalen*, 42, 305–310.

Ferguson, N. (2002). *Empire: The Rise and Demise of the British World Order*. London, Penguin Books.

Ferguson, N. (2012). *Civilization*. London, Penguin.

Flyvbjerg, H., Sneppen, K. and, Bak, P. (1993). A mean field theory for a simple model of evolution. *Physics Review Letters*, 71, 4087–4090.

Fox, J. (2009). *The Myth of the Rational Market.* New York, Harper.

Fukuyama, F. (2011). *The Origins of Political Order*. New York, Ferrar, Strauss and Giroux.

Gazzaniga, M. (2008). *Human*. New York, Harper.

Gintis, H. (2000). Strong reciprocity and human sociality. *Journal of Theoretical Biology*, 206, 169–179.

Gladwell, M. (2002). *The Tipping Point.* New York, Back Bay.

Gleick, J. (1987). *Chaos: Making a new science*. New York, Viking.

Gödel, K. (1931). Uber formal unentscheidbare Satze der Principia Mathematica und verwandter Systeme. *Monatschefte fur Mathematik und Physik*, 38, 173–98. Translated as "On Formally Undecidable Propositions of *Principia Mathematica* and Related Systems," in van Heijenoort, J. (ed.), *Frege and Gödel: Two Fundamental Texts in Mathematical Logic.* Cambridge, MA, Harvard University Press.

Golub, B. and Jackson, M. (2010). Naive learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics*, 2, 112–149.

Gore, A. (2006). *The Future.* New York, Random House.

Goud (1976)

Gribbin, J. (2011). *Alone in the Universe*. New York, Wiley.

Hahn, F. (1973). *On the Notion of Equilibrium in Economics.* Cambridge, Cambridge University Press.

Hamilton, W. (1964). The genetical evolution of social behavior I and II. *Journal of Theoretical Biology*, 7, 1–52.

Hamilton, W. (1970). Selfish and spiteful behavior in an evolutionary model. *Nature*, 228, 1218–1220.

Hardin, G. (1968 [1973]). The tragedy of the commons. In Daly, H. E. (ed.), *Towards a Steady State Economy*. San Francisco, Freeman.

Hardin, R. (1971). Collective action as an agreeable *n*-prisoners' dilemma. *Behavioral Science*, 16, 472–481.

Hardin, R. (1982). *Collective Action*. Baltimore, MD, Johns Hopkins University Press.

Hawking, S. and Mlodinow, L. (2010). *The Grand Design*. New York, Random.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E. and Gintis, H. (2004). *Foundations of human sociality*. Oxford, Oxford University Press.

Henrich, J. et al. (2005). "Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral Brain Science*, 28, 795–855.

Hirsch, M. (1976). *Differential Topology.* Berlin, Springer.

Hobbes, T. (2009 [1651]). *Leviathan; or the matter, forme, and power of a commonwealth, ecclesiastical and civil* (ed.: Gaskin, J. C. A.). Oxford, Oxford University Press.

Hrdy, S. B. (2011). *Mothers and Others: The Evolutionary Origins of Mutual Understanding.* Cambridge MA, Harvard University Press.

Hsiang, S., Burke, M. and Miguel, E. (2013). Quantifying the influence of climate on human conflict. *Science express*, DOI: 10.1126/science.1235367.

Israel, J. (2002). *Radical Enlightenment.* Oxford, Oxford University Press.

Israel, J. (2006). *Enlightenment Contested.* Oxford, Oxford University Press.

Israel, J. (2010). *Revolution of the Mind.* Princeton NJ, Princeton University Press.

Israel, J. (2014). *Revolutionary Ideas.* Princeton NJ, Princeton University Press.

Johnson, S. (2002). *Emergence*. New York, Scribner.

Kahneman, D. (2011). *Thinking Fast and Slow*. New York, Ferrar, Strauss and Giroux.

Karklins, R. and Petersen, R. (1993). Decision calculus of protestors and regime change: Eastern Europe 1989. *Journal of Politics*, 55, 588–614.

Kauffman, S. (1993). *The Origins of Order.* Oxford, Oxford University Press.

Keohane, R. and Nye, R. (1977). *Power and Interdependence*. New York, Little Brown.

Keynes, J. M. (1921). *Treatise on Probability*. London, Macmillan.

Keynes, J. M. (1936). *The general theory of employment, interest and money*. London, Macmillan.

Kindleberger, C. (1973). *The World in Depression 1929–1939*. Berkeley, CA, University of California Press.

Knaster, B., Kuratowski, K. and Mazurkiewicz, S. (1929). Ein beweis des fixpunktsatzes fur n-dimensionale simplexe. *Fund Math*, 14, 132–137.

Kramer, G. H. (1973). On a class of equilibrium conditions for majority rule. *Econometrica*, 41, 285–297.

Kreps, D. M., Milgrom, P., Roberts, J. and, Wilson, R. (1982). Rational cooperation in the finitely repeated prisoners' dilemma. *Journal of Economic Theory*, 27, 245–252.

Ladha, K. (1992). Condorcet's jury theorem, free speech and correlated votes. *American Journal of Political Science*, 36, 617–74.

Ladha, K. (1993). Condorcet's jury theorem in the light of de Finetti's theorem: Majority rule with correlated votes. *Social Choice and Welfare*, 10, 69–86.

Ladha, K. and Miller, G. (1996). Political discourse, factions and the general will: correlated voting and Condorcet's jury theorem. In Schofield, N. (ed.), *Collective decision making*. Boston, Kluwer.

Laplace, P. S. (1951 [1814]). *Essai Philosophique sur les Probabilités.* Paris, Gauthiers-Villars. (*A philosophical essay on probabilities*, translated by F. Truscott and F. Emory, New York, Dover.)

Li, T. Y. and Yorke, J. A. (1975). Period three implies chaos. *Mathematics Monthly*, 82, 985–992.

Lohmann, S. (1994). The dynamics of Information cascades. *World Politics*, 47, 42–101.

Lorenz, E. N. (1962). The statistical prediction of solutions of dynamical equations. *Proceedings of the International Symposium on Numerical Weather Prediction*, Tokyo.

Lorenz, E. N. (1963). Deterministic non-periodic flow. *Journal of Atmosphere Science*, 130–141.

Madison, J. (1999 [1787]). Federalist X. In Rakove, J. (ed.), *Madison: Writings*. New York, Library Classics.

Mandelbrot, B. and Hudson, R. (2004). *The (Mis)behavior of Markets*. New York, Perseus.

Margolis, H. (1982). *Selfishness, altruism and rationality*. Cambridge, Cambridge University Press.

McKelvey, R. D. (1976). Intransitivities in multidimensional voting models and some implications for agenda control. *Journal of Economic Theory*, 12, 472–482.

McKelvey, R. D. (1979). General conditions for global intransitivities in formal voting models. *Econometrica*, 47, 1085–1112.

McKelvey, R. D. and Page, T. (1986). Common Knowledge, Consensus and Affregrate Information. *Econometrica*, 54, 109–127.

McKelvey, R. D. and Schofield, N. (1987). Generalized symmetry conditions at a core point. *Econometrica*, 55, 923–933.

McWhorter, J. (2001). *The Power of Babel*. New York, Holt.

Meadows, D. H., Meadows, D. L., Randers, J. and Behrens, W. W. (1972). *Limits to growth*. New York, Signet.

Meadows, D. H., Meadows, D. L. and Randers, J. (2012). *Limits to Growth: The Thirty year Update*. New York, Chelsea Green.

Minsky, H. (1975). *John Maynard Keynes.* New York, Columbia University Press.

Minsky, H. (1986). *Stabilizing an Unstable Economy.* Yale University Press, New Haven.

Mokyr, J. (2005). The intellectual origins of modern economic growth. *Journal of Economic History*, 65, 285–351.

Mokyr, J. (2010). *The Enlightened Economy: An Economic History of Britain 1700–1850.* New Haven, CT, Yale University Press.

Mokyr, J. and Nye, V. C. (2007). Distributional coalitions, the Industrial Revolution, and the origins of economic growth in Britain. *Southern Economic Journal*, 74, 50–70.

Morris, I. (2010). *Why the West Rules.* New York, Ferrar, Strauss and Giroux.

Nagel, T. (2012). *Mind and Cosmos*. Oxford, Oxford University Press.

Nakamura, K. (1979). The vetoers in a simple game with ordinal preference. *International Journal of Game Theory*, 8, 55–61.

North, D. C. (1990). *Institutions, Institutional Change and Economic Performance.* Cambridge, Cambridge University Press.

North, D. C. and Weingast, B. R. (1989). Constitutions and commitment: The evolution of institutions governing public choice in seventeenth century England. *Journal of Economic History*, 49, 803–832.

North, D. C., Wallis, B. and Weingast, B. R. (2009). *Violence and Social Orders: A Conceptual Framework for Interpreting Recorded Human History.* Cambridge, Cambridge University Press.

Nowak, M. (2011). *Supercooperators*. New York, Free Press.

Oreskes, N. and Conway, E. M. (2014). *The Collapse of Western Civilization*. New York, Columbia University Press.

Ormerod, P. (2001). *Butterfly Economics*. New York, Basic.

Parfit, D. (2011). *On What Matters.* Oxford, Oxford University Press.

Parker, G. (2013). *Global Crisis.* New Haven, CT, Yale University Press.

Penn, E. (2009). A model of far-sighted voting. *American Journal of Political Science*, 53, 36–54.

Plott, C. R. (1967). A notion of equilibrium and its possibility under majority rule. *American Economic Review*, 57, 787–806.

Pugh, C. C. (2002). *Real Mathematical Analysis.* Berlin, Springer.

Putnam, R. D. and Campbell, D. E. (2010). *American Grace: How Religion Divides and Unites US.* New York, Simon and Schuster.

Rae, D. W. (1960). Decision Rules and Individual Values in Constitutional Choice. *American Political Science Review*, 63, 40–56.

Richards, D. (1990). Is strategic decision making chaotic? *Behavioral Science*, 35, 219–232.

Robson, A. J. and Kaplan, H. S. (2003). The evolution of human life expectancy and intelligence in hunter-gatherer economies. *American Economic Review*, 93, 150–169.

Saari, D. (1997). The Generic existence of a core for q-Rules. *Economic Theory,* 9, 219–260.

Schofield, N. (1972a). Is majority rule special? In Niemi, R. G. and Weisberg, H. F. (eds.), *Probability Models of Collective Decision-Making.* Columbus, OH, Charles E. Merrill Publishing.

Schofield, N. (1972b). Ethical decision rules for uncertain voters. *British Journal of Political Science*, 2, 193–207.

Schofield, N. (1977). The Logic of catastrophe. *Human Ecology*, 5, 261–271.

Schofield, N. (1978). Instability of simple dynamic games. *Review of Economic Studies*, 45, 575–594.

Schofield, N. (1979). Rationality or chaos in social choice. *Greek Economic Review*, 1, 61–76.

Schofield, N. (1980a). Generic properties of simple Bergson-Samuelson welfare functions. *Journal of Mathematical Economics*, 7, 175–192.

Schofield, N. (1980b). Catastrophe theory and dynamic games. *Quality Quantity*, 14, 529–545.

Schofield, N. (1983a). Equilibria in simple dynamic games. In Pattanaik, P. and Salles, M. (eds.), *Social Choice and Welfare*. Amsterdam, North Holland, 269–284.

Schofield, N. (1983b). Generic instability of majority rule. *Review of Economic Studies*, 50, 695–705.

Schofield, N. (1984c). Classification theorem for smooth social choice on a manifold. *Social Choice and Welfare*, 1, 187–210.

Schofield, N. (1999). The C^1-topology on the space of smooth preferences. *Social Choice and Welfare*, 16, 445–470.

Schofield, N. (2002). Evolution of the constitution. *British Journal of Political Science,* 32, 1–20.

Schofield, N. (2003). *Mathematical methods in economics and social choice*. Berlin, Springer.

Schofield, N. (2006). *Architects of Political Change*. Cambridge, Cambridge University Press.

Schofield, N. (2010). Social orders. *Social Choice and Welfare,* 34, 503–536.

Schofield, N. (2011). Is the political economy stable or chaotic? *Czech Economic Review*, 5, 76–93.

Schofield, N. and Gallego, M. (2011). *Leadership or chaos*. Berlin, Springer.

Schweitzer, F. et al. (2009). Economic networks: The new challenges. *Science*, 325, 422–425.

Shafer, W. and Sonnenschein, H. (1975). Equilibrium in abstract economies without ordered preferences. *Journal of Mathematical Economics*, 2, 245–248.

Seddon, C. (2014). *Humans*. New York, Granville.

Shiller, R. (2003). *The New Financial Order*. Princeton, NJ, Princeton University Press.

Shiller, R. (2005). *Irrational Exuberance*. Princeton, NJ, Princeton University Press.

Smale, S. (1966). Structurally stable systems are not dense. *American Journal of Mathematics*, 88, 491–496.

Smale, S. (1974a). Global analysis and economics IIA: Extension of a theorem of Debreu. *Journal of Mathematical Economics*, 1, 1–14.

Smale, S. (1974b). Global analysis and economics IV: Finiteness and stability of equilibria with general consumption sets and production. *Journal of Mathematical Economics*, 1, 119–127.

Smith, A. (1984 [1759]). *The Theory of Moral Sentiments*. Indianapolis, IN, Liberty Fund.

Sperber, J. (2014). *Karl Marx: A nineteenth century life*. New York, Liveright.

Smolin, L. (2007). *The Trouble with Physics*. New York, Houghton Mifflin.

Strogatz, S. (2004). *Sync*. New York, Hyperion.

Stern, N. (2007). *The Economics of Climate Change*. Cambridge, Cambridge University Press.

Stern, N. (2009). *The Global Deal*. New York, Public Affairs.

Stringer, C. (2012). *Lone Survivors*. London, Macmillan.

Sunstein, C. R. (2006). *Infotopia*. Oxford, Oxford University Press.

Sunstein, C. R. (2009). *A Constitution of Many Minds*. Princeton, NJ, Princeton University Press.

Sunstein, C. R. (2011). *Going to extremes*. Oxford, Oxford University Press.

Surowiecki, J. (2005). *The Wisdom of Crowds*. New York, Anchor.

Tainter, J. A. and Renfrew, C. (1988). *The Collapse of Complex Societies*. Cambridge, Cambridge University Press.

Taleb, N. N. (2007). *The Black Swan*. New York, Random.

Taleb N. N. and Blyth, M. (2011). The black swan of Cairo. *Foreign Affairs*, 90(3), 33–39.

Taylor, C. (2007). *A Secular Age*. Cambridge MA, Belknap Press.

Taylor, M. (1976). *Anarchy and Cooperation*. London, Wiley.

Taylor, M. (1982). *Community, Anarchy and Liberty.* Cambridge, Cambridge University Press.

Trivers, R. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, 46, 35–56.

Turing, A. (1937). On Computable Numbers with an Application to the Entscheidungs Problem. *Proceedings of the London Mathematical Society*, 42, 230–265. (Reprinted in Jack Copeland, *The Essential Turing,* 1st edition. Oxford, The Clarendon Press.)

Walker, M. (1977). On the existence of maximal elements. *Journal of Economic Theory*, 16, 470–474.

Watts, D. (2002). A simple model of global cascades on random networks. *Proceedings of the National Academy of Science*, 99, 5766–5771.

Watts, D. (2003). *Six Degrees*. New York, Norton.

Weitzman, M. (2009). Additive damages, fat-tailed climate dynamics, and uncertain discounting. *Economics*, 3, 1–22.

White, T. D. et al. (2009). Ardipithicus ramidus and the paleobiology of early hominids, *Science*, 326, 64–86.

Zhang, D. D., Brecke, P., Lee, H. F., He, Y.-Q. and Zhang, J. (2007). Global climate change, war, and population decline in recent human history. *Proceedings of the National Academy of Science*, 104(49), 19214–19219.

Czech Economic Review 9 (2015) 36–48
Acta Universitatis Carolinae Oeconomica

Participation and Solidarity in Redistribution Mechanisms

José-Manuel Giménez-Gómez* **and Josep E. Peris****

Received 24 February 2015; Accepted 10 June 2015

Abstract Following Bossert (1995), we consider a model where personal income depends on two different characteristics: skills and effort. Luttens (2010) introduces claims that individuals have over aggregate income and that only depend on the effort they exert. Moreover, he proposes redistribution mechanisms in which *solidarity* is based on changes in a lower bound on what every individual deserves according to these claims: the so-called *minimal rights* (O'Neill 1982). A debatable consequence in one of Luttens' mechanisms is that "the poorest individuals might up with a negative income" (Luttens 2010); that is, this mechanism does not satisfy *participation*, which turns out to be incompatible with *claims feasibility*, under Luttens' assumptions. We present a new solidarity axiom that is compatible both with *participation* and *claims feasibility*, and we provide a mechanism satisfying these properties and our *new additive solidarity axiom*. Moreover, our mechanism satisfies additional properties, as *priority*, or *respect of minimal rights*.

Keywords Redistribution mechanism, minimal rights, solidarity, participation, claims feasibility
JEL classification C71, D63, D71

1. Introduction

We suppose that inequalities in welfare among individuals in a society are determined by two different factors or characteristics: *skills* and *effort*. The difference between these characteristics is that individuals are completely responsible for the inequalities due to differences in effort, that do not deserve compensation (effort reflects, for instance, the number of hours that a person decides to work). Nevertheless, there are other circumstances, which are beyond the control of the individuals, that deserve compensation (different innate skills or talents, economic status, historical inequality due to race, gender etc.).

The aim of fair income redistribution is to guarantee an equal income for individuals exerting the same effort (*the principle of compensation*) and to perform equal income transfers to individuals with equal skills (*the principle of natural reward*). It is well known that, in many contexts, a redistribution mechanism satisfying both the principle of compensation and the principle of natural reward simultaneously does not exist. As a result, the literature has concentrated on dealing with such trade-off between both principles. Notably, most contributions have opted for a weakening of the principle of natural reward (see, for instance, Fleurbaey 1994; Bossert 1995; Bossert and Fleurbaey

* Universitat Rovira i Virgili, Dep. d'Economia and CREIP, Av. Universitat 1, 43204 Reus, Spain. Tel.: (+34) 977 759 850, E-mail: josemanuel.gimenez@urv.cat.
** Corresponding author. Universitat d'Alacant, Dep. de Mètodes Quantitatius i Teoria Econòmica, 03080 Alacant, Spain. E-mail: peris@ua.es.

1996; Iturbe-Ormaetxe 1997; Tungodden 2005). Other approaches have opted for strengthening compensation with respect to the principle of solidarity, a principle with a long tradition in the theory of justice.[1] In accordance, Bossert and Fleurbaey (1996) and Iturbe-Ormaetxe (1997) present two modified versions of the solidarity axiom (*additive solidarity* and *multiplicative solidarity*) to characterize the egalitarian equivalent mechanisms (Pazner and Schmeidler 1978) and the family of proportionally adjusted equivalent mechanisms (Iturbe-Ormaetxe 1997), respectively. A central notion on fairness is the *No-Envy* property, suggested by Foley (1967) and analyzed by Kolm (1972), Panzer and Schmeidler (1974) and Varian (1974).

We use the quasi-linear model developed in Bossert (1995), with utility functions taking the form

$$u_i = x_i + v(y_i, z_i),$$

where $v(y_i, z_i)$ is described as agent pre-tax income, x_i is an income transfer, and u_i is the final income after redistribution. As usual in this model, we consider that the total amount to be distributed is $\Omega = 0$; that is, we consider a *redistribution problem*.

An important feature of this model is that the redistribution of resources is based only on the set of characteristics that deserve compensation. Furthermore, such compensation is assigned according to a *solidarity* basis (Rawls 1971): *changes in these characteristics should affect each individual's final utility in the same direction*. Accordingly, Bossert (1995) proposes the property of *additive solidarity*, which is based on the idea that individuals should benefit equally from variations in the skills profile.

In a recent paper, Luttens (2010) includes a new element into the model, by defining a *claims* function that depends on the individual's effort z but does not depend on the individual's skill y, "hence, two individuals with identical effort, but different skills (different pre-tax incomes) have identical claims in the redistribution problem" (Luttens 2010). Within this approach, Luttens makes a bridge between the conflicting claims literature and the Bossert's taxation model (Bossert 1995) with quasi-linear preferences previously mentioned. Moreover, he proposes a lower bound in the individuals' welfare based on their claims function, namely the *minimal rights lower bound*,[2] and defines a strengthening of the additive solidarity principle: *an income gain (loss), generated by a change in the skills profile, is shared on the basis of the information contained in changes of this lower bound*.

The redistribution mechanism proposed by Luttens (2010) fails to respect the minimal rights lower bound on which it is based; that is, for some individuals the income after redistribution might be lower than her *minimal right*. Moreover, as the author suggests, "a debatable property is that the poorest individuals might end up with a negative income after redistribution when R (the aggregate income) is sufficiently low." This property corresponding to the notion of *participation* (Maniquet 1998) "captures the idea of protecting high-skilled agents in the sense that the low-skilled agent compensation should not be carried out by imposing on the former agents so long labor time that they end up worse off than if they withdrew from the economy" (see also

[1] See Fleurbaey and Maniquet (2011) for a comprehensive summary of this literature. We follow this paper for notation and definitions.

[2] This bound was introduced by O'Neill (1982) in the context of claims problems.

Fleurbaey and Maniquet 2011). In other contexts, this non-negativity condition has also been adopted. For instance, "in production models, non-negativity (*participation*) takes care of the "slavery" problem, by avoiding situations in which an agent would rather opt out of the economy than participate in the production process" (Fleurbaey 2008). Given a change in the skill profile, a planner can perform a redistribution by following a *solidarity* property, but this should not come at the cost of an agent ending up with a negative income after redistribution. Once an agent ends up with a zero income after redistribution, she no longer needs to take part in performing solidarity. Luttens (2010), in order to solve this problem, defines an alternative mechanism satisfying *participation*, at the expense of losing *claims feasibility*.

We are interested in keeping both properties along with the axiom of *respect of the minimal rights*. Then, we propose a refinement of Luttens' mechanisms which makes compatible *participation* and *claims feasibility* by weakening the *solidarity* condition. Note that *minimal rights* suppose a very weak notion of guarantee: it requires that each individual receives at least what is left of the resources after the other claims have been fully compensated, or zero if this amount is negative. So, if the claims are high enough with respect to the aggregate income, no individual guarantee exists at all.

The paper is organized as follows. In Section 2, we present the model and introduce the basic definitions. Section 3 proposes and characterizes our respect of minimal rights-based egalitarian mechanism. Some final remarks close the paper in Section 4. An appendix gathers the proof of our characterization result and the independence of the axioms used in this characterization.

2. The model

2.1 Fair monetary compensation model

Let us denote by $N = \{1, \ldots, n\}$ the finite population of size $n \geq 2$. Individuals are distinguished by two characteristics: *skill* and *effort*. Differences in skill elicit compensation. The individual's skill is represented by a real non-negative number $y \in \mathbb{Y}$, where $\mathbb{Y}$ is an interval of $\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$. The skills profile is the vector $y_N = (y_1, \ldots, y_n) \in \mathbb{Y}^n$. The effort is also denoted by a non-negative real number $z \in \mathbb{Z}$, where $\mathbb{Z}$ is an interval of $\mathbb{R}_+$. The effort profile is $z_N = (z_1, \ldots, z_n) \in \mathbb{Z}^n$. So, each individual is identified by the pair of non-negative real numbers $(x_i, y_i) \in \mathbb{Y} \times \mathbb{Z}$, specifying her skill and effort, respectively. Without loss of generality, throughout the work we assume that individuals are ranked with respect to the effort they exert: $z_1 \geq z_2 \geq \ldots \geq z_n$. An economy consists of the pair of skill and effort profiles, $e = (y_N, z_N) \in \mathbb{Y}^n \times \mathbb{Z}^n$. Let $\mathcal{E}$ denote the set of economies, $\mathcal{E} \subseteq \mathbb{Y}^n \times \mathbb{Z}^n$.

Given an economy $e = (y_N, z_N) \in \mathcal{E}$, a *pre-tax income* function (identical for all individuals), $v : \mathbb{Y} \times \mathbb{Z} \to \mathbb{R}_+$, associates to each individual (y_i, z_i) a monetary income $v(y_i, z_i)$ that depends on her skill and effort. It is supposed that function v is strictly increasing in y. The total sum of pre-tax incomes is denoted by $R = \sum_{i \in N} v(y_i, z_i)$.

Differences in individuals' skills are compensated by an amount x_i of a transferable resource (money). Differences in effort do not elicit compensation. An allocation $x_N = (x_1, \ldots, x_n) \in \mathbb{R}^n$ is the vector defined by transferable resources x_i. We assume that the total amount to be distributed is $\Omega = 0$, so that we are looking at a redistribution problem (total subsidies coincide with total taxes). Then, an allocation is *feasible* whenever $\sum_{i\in N} x_i = 0$.

We assume, as in Luttens (2010), that individuals, because of the effort they exert, have some *claim* on the total pre-tax income R. Let $g : \mathbb{Z} \to \mathbb{R}_{++}$ be the *claims function* that assigns to each individual (y_i, z_i) a claim $g(z_i)$ that depends on the individual's effort only. We assume that function $g(z)$ is continuous and strictly increasing in z. We denote the total sum of claims by $C = \sum_{i\in N} g(z_i)$, and $C_{-i} = \sum_{j\neq i\in N} g(z_j)$. The redistribution problem will be a conflicting claims problem whenever $C > R$. Let us denote the claims vector of an economy $e = (y_N, z_N)$ by $\hat{\mathbf{g}} = (g(z_1), g(z_2), \ldots, g(z_n))$.

A (*redistribution*) mechanism is a function $S : \mathcal{E} \times \mathbb{R}^n \to \mathbb{R}^n$ such that for all $e \in \mathcal{E}$, and any claims vector $\hat{\mathbf{g}}$, $S(e, \hat{\mathbf{g}})$ is a feasible allocation, that is

$$\sum_{i\in N} S_i(e, \hat{\mathbf{g}}) = 0.$$

It is assumed, as in Bossert (1995), that individuals preferences are characterized by (quasi-linear) utility functions, $u : \mathbb{R} \times \mathbb{Y} \times \mathbb{Z} \to \mathbb{R}$, which are defined as follows:

$$u(x_i, y_i, z_i) = x_i + v(y_i, z_i).$$

Utility represents the final income after redistribution. It is clear that, as $\sum_{i\in N} x_i = 0$,

$$\sum_{i\in N} u(x_i, y_i, x_i) = \sum_{i\in N} v(y_i, z_i) = R.$$

As the compensation x_i each agent receives will depend on the claims vector, we shall denote the utility function of individual (y_i, z_i) by $u_i(e, \hat{\mathbf{g}})$,

$$u_i(e, \hat{\mathbf{g}}) = x_i(\hat{\mathbf{g}}) + v(y_i, z_i).$$

We now introduce some notation that will be helpful. Given two economies $e = (y_N, z_N)$ and $e' = (y'_N, z_N)$, which only differ in skills profiles, and any claims vector $\hat{\mathbf{g}}$, changes in any function or variable are denoted by the difference operator Δ. Then, $\Delta h = h(e', \hat{\mathbf{g}}) - h(e, \hat{\mathbf{g}})$. Note that, since the effort is the same in both economies, the claims vector also coincides. This notation will be used to represent changes in the utility function, u, the minimal rights vector, m, as well as changes in the total pre-tax income, R (although R does not depends on the claims vector $\hat{\mathbf{g}}$).

2.2 Axioms

Before introducing the axioms, we provide the definition of the minimal rights lower bound (O'Neill 1982). This bound guarantees to each individual the amount that is left when the rest of the individuals have received their claim, or zero if this amount is negative. Associated with this bound, *respect of minimal rights* states that each individual should receive at least her minimal right (this axiom is a consequence of efficiency, non-negativity and claims boundedness combined (Thomson 2003)).

Definition 1. *Minimal rights* (O'Neill 1982). For each economy $e = (y_N, z_N) \in \mathcal{E}$, and each claims vector $\hat{\mathbf{g}}$, the minimal rights vector, $m \in \mathbb{R}^n_+$, is defined by $m_i = m_i(e, \hat{\mathbf{g}}) = \min\left\{g(z_i), [R - C_{-i}]_+\right\}$, $i \in N$, where $[a]_+ = \max\{0, a\}$.

Axiom 1. *Respect of minimal rights* (*RMR*). For each economy $e = (y_N, z_N) \in \mathcal{E}$, each claims vector $\hat{\mathbf{g}}$, and each $i \in N$, $u_i(e, \hat{\mathbf{g}}) \geq m_i(e, \hat{\mathbf{g}})$.

The following axiom, *participation*, states that no individual can incur losses, i.e. when R converges to zero, all incomes should also converge to zero. Moreover, note that before redistribution the pre-tax income of each individual $v_i = v(y_i, z_i)$ is non-negative, so after redistribution this condition should be maintained. Note that this property is implied by the axiom *respect of minimal rights*.

Axiom 2. *Participation* (*P*, Maniquet 1998). For each economy $e = (y_N, z_N) \in \mathcal{E}$, each claims vector $\hat{\mathbf{g}}$, and each $i \in N$, $u_i(e, \hat{\mathbf{g}}) \geq 0$.

Next, we present the axioms characterizing our mechanism. The first one, *claims feasibility* is a standard assumption which requires that when the aggregate pre-tax income (resources) equals the aggregate claim, then each individual's utility equals her claim.

Axiom 3. *Claims Feasibility* (*CF*). For each economy $e = (y_N, z_N) \in \mathcal{E}$, and each claims vector $\hat{\mathbf{g}}$, if $R = C$, then $u_i(e, \hat{\mathbf{g}}) = g(z_i)$ for all $i \in N$.

The following property (Luttens 2010) requires an equal treatment of two individuals in the allocation of the extra resources when their minimal rights change equally.

Definition 2. *Additive Solidarity for equal changes in minimal rights* (AS^*, Luttens 2010). Given two economies $e = (y_N, z_N), e' = (y'_N, z_N) \in \mathcal{E}$, that only differ in skills profiles, and a claims vector $\hat{\mathbf{g}}$, if $\Delta m_i = \Delta m_j$, then $\Delta u_i = \Delta u_j$.

Finally, the following axiom establishes that the changes in the resources should be shared among those individuals with changes in their minimal rights.

Axiom 4. *Priority* (*PRI*, Luttens 2010). Given two economies $e = (y_N, z_N), e' = (y'_N, z_N) \in \mathcal{E}$, that only differ in skills profiles, and a claims vector $\hat{\mathbf{g}}$, if $N_1 = \{i \in N : \Delta m_i \neq 0\} \neq \varnothing$, and $\Delta m_i = \Delta R$, for all $i \in N_1$, then $\sum_{i \in N_1} \Delta u_i = \Delta R$; or, equivalently, $\Delta u_j = 0$, for all $j \notin N_1$.

Luttens (2010) proposes and characterizes two different mechanisms. One of them with the axioms AS^*, *PRI* (he names both axioms together as *minimal rights-based solidarity*) and *CF*; the other mechanism with AS^*, *PRI* and *P*. As a consequence, *minimal rights-based solidarity*, *claims feasibility* and *participation* are incompatible axioms. So, when imposing Luttens' *minimal rights-based solidarity* together with *claims feasibility*, a redistribution mechanism also fails the *respect minimal rights* axiom. Luttens argues that "this incompatibility is due to AS^* rather than *priority*." We agree on this argument and in order to obtain compatibility we just modify the AS^* axiom.

The following example will be useful to better observe what happens with the *minimal rights-based solidarity* axioms when *participation* is required.

Example 1. Let us consider an economy e, with $n = 4$ individuals such that $\hat{\mathbf{g}} = (80, 70, 50, 30)$. As previously mentioned, *participation* implies $u_i = 0$ for all i, when $R' = 0$. In this case, the *minimal rights* vector is $m = (0, 0, 0, 0)$. If, due to a change in the skills profile, we now have $R = 150$, then all *minimal rights* are still null, so axiom AS^* implies an equal sharing of the extra resources, that is $u_i = 37.5$ for all i. Then, individual 4 ends up at a welfare level that is above what she deserves (her *claim*) and the other individuals are below their claims. Thus, *claims feasibility* is not met.

What is the problem in the above example? In our opinion, the "problem" occurs when the *minimal right* equals to zero, and it is originated because of the $[\,]_+$ operator that appears in the definition of *minimal rights*. To see this, observe that whenever $R \leq 150$ all minimal rights are null, but there is a significant difference before applying the $[\,]_+$ operator:

$$m_1 = \min\left\{80, [0]_+\right\} = 0, \qquad m_2 = \min\left\{70, [-10]_+\right\} = 0$$
$$m_3 = \min\left\{50, [-30]_+\right\} = 0, \qquad m_4 = \min\left\{40, [-50]_+\right\} = 0.$$

In order to prevent this situation (individuals with a negative income after redistribution, or individuals above their claims, when it is not possible to satisfy all claims), we present the following modification of the *additive solidarity* axiom, which introduces an equal treatment of two individuals when their *minimal rights* change equally and both individuals have a positive income after redistribution.

Axiom 5. *Additive Solidarity for equal significant changes in minimal rights* (AS^{**}). Given two economies $e = (y_N, z_N), e' = (y'_N, z_N) \in \mathcal{E}$, that only differ in skills profiles, and a claims vector $\hat{\mathbf{g}}$, such that $R' < R$, if $\Delta m_i = \Delta m_j$, then $u_j(e', \hat{\mathbf{g}}) = \left[u_j(e, \hat{\mathbf{g}}) + \Delta u_i\right]_+$.

It is important to note that whenever $u_j(e', \hat{\mathbf{g}}) > 0$, then AS^{**} coincides with AS^*. So, the modification of this axiom only has relevance for values of R sufficiently low. Specifically, when an individual already receives a zero income after redistribution, that individual should no longer be affected by a further deterioration of the skills profile in society. As mentioned in Luttens (2010) "our ethical intuition may lead us to consider a minimal amount of redistribution, that we at least want to perform. Suppose that the poorest in society could not satisfy their basic needs when they receive a negative income after redistribution. Society wants to exclude this possibility in every situation by incorporating the requirement of a non-negative income after redistribution for all individuals in the construction of the redistribution mechanism."

The following example shows the different result we obtain, by applying AS^{**} instead of AS^*, in the situation of Example 1.

Example 2. (continues from Example 1) Note that now, when applying AS^{**} ($R' = 0$, $R = 150$), an egalitarian distribution is not necessarily obtained for R, since from $\mathbf{u} = (0, 0, 0, 0)$ not necessarily all increments must be equal in order to fulfill AS^{**}. In fact, to obtain the utility vector in $R = 150$, we can begin with $R_0 = 230$, and *claims feasibility* implies $\mathbf{u} = (80, 70, 50, 30)$. In this case, the *minimal rights* vector is $m = (80, 70, 50, 30)$. If, due to a change in the skills profile, we now have $R_1 = 200$, then $m = (50, 40, 20, 0)$, so axiom AS^{**} implies an equal sharing of the lost resources, that is $\mathbf{u} = (72.5, 62.5, 42.5, 22.5)$. Now, consider that the resources are $R_2 = 180$, then $m = (30, 20, 0, 0)$, so by AS^{**} and PRI, we obtain $\mathbf{u} = (65.83, 55.83, 35.83, 22.5)$. When $R_3 = 160$, then $m = (10, 0, 0, 0)$, so AS^{**} and PRI imply $\mathbf{u} = (55.83, 45.83, 35.83, 22.5)$. Finally, for $R_4 = 150$, we know $m = (0, 0, 0, 0)$, and AS^{**} and PRI imply that $\mathbf{u} = (45.83, 45.83, 35.83, 22.5)$ which differs from the egalitarian proposal obtained under AS^*.

3. A respect-of-minimal-rights egalitarian mechanism

This section provides an alternative to Luttens' mechanisms, which is based on the fulfillment of the respect of minimal rights axiom. Our main result characterizes this mechanism in terms of *claims feasibility*, *priority* and our new axiom of *additive solidarity for equal significant changes in minimal rights*. It must be noticed that given an economy e and a claims vector $\hat{\mathbf{g}}$, associated to any redistribution mechanism S, the utility each individual obtains is

$$u_i(e, \hat{\mathbf{g}}) = S(e, \hat{\mathbf{g}}) + v(y_i, z_i).$$

Observe that the axioms are formulated in terms of utilities, instead of the redistribution mechanism.

Our mechanism has an egalitarian objective, which is constrained in terms of the minimal right of each individual: that is, individuals with identical minimal rights increase their utility in the same level. Figure 1 shows how slopes of such individuals coincide. Observe that when the aggregate pre-tax income R is large enough $(R \geq C_{-4})$, then all individuals increase their utility at the same rate.

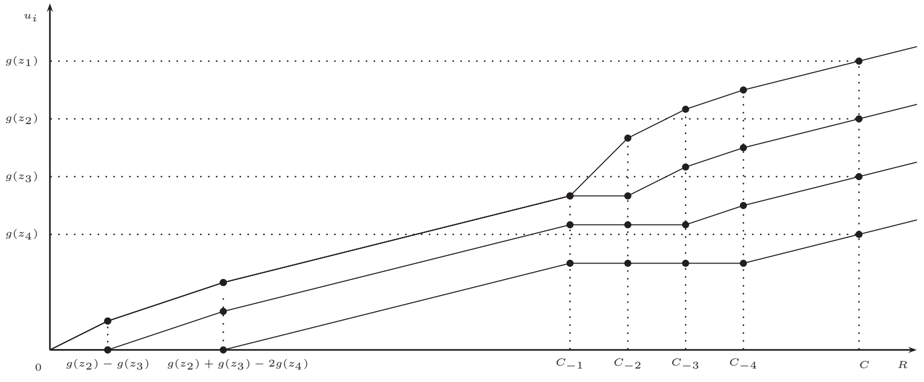


Figure 1. Utilities provided by S_{RMRE} (respect-of-minimal-rights egalitarian mechanism). The top utility level corresponds to the individual with greatest effort, u_1, then u_2, u_3 and u_4.

Definition 3. The *respect-of-minimal-rights* egalitarian mechanism S_{RMRE} allocates resources for each $e \in \mathcal{E}$, each claims vector $\hat{\mathbf{g}}$ and each $i \in N$, as follows:

$$(x_i)_{S_{RMRE}} = -v(y_i, z_i) + d_i(\hat{\mathbf{g}}, R),$$

where $d_i(\hat{\mathbf{g}}, R)$ is defined by:

(i) $R \geq C_{-n}$:

$$d_i(\hat{\mathbf{g}}, R) = g(z_i) + \frac{R - C}{n} \qquad \forall i \in N$$

(ii) $C_{-k} \leq R \leq C_{-(k+1)}$:

$$d_i(\hat{\mathbf{g}}, R) = \begin{cases} d_i(\hat{\mathbf{g}}, C_{-(k+1)}) & \forall i \geq k+1 \\ d_i(\hat{\mathbf{g}}, C_{-(k+1)}) + \dfrac{R - C_{-(k+1)}}{k} & \forall i < k+1 \end{cases}$$

(iii) $G_{n-1} \leq R \leq C_{-1}$:

$$d_i(\hat{\mathbf{g}}, R) = d_i(\hat{\mathbf{g}}, C_{-1}) + \frac{R - C_{-1}}{n} \qquad \forall i \in N$$

(iv) $G_{n+1-k} \leq R \leq G_{n+2-k}$, $\quad k = 2, 3, \ldots, n-2$,

$$d_i(\hat{\mathbf{g}}, R) = \begin{cases} 0 & \forall i \geq n+2-k \\ d_i(\hat{\mathbf{g}}, G_{n+2-k}) + \dfrac{R - G_{n+2-k}}{k} & \forall i < n+2-k \end{cases}$$

where $G_s = \sum_{i=2}^{s-1} g(z_i) - (s-2)g(z_{s-1})$.

Theorem 1. *A redistribution mechanism S coincides with S_{RMRE} if and only if S satisfies CF, AS** and PRI.*

Proof. See Appendix A1.

The independence of the axioms that appear in Theorem 1 is shown in Appendix A2. The following proposition, which can be straightforwardly obtained from the proof of Theorem 1, highlights the fact that our solution meets the boundedness on which it is based. On the other hand, as we have mentioned, *respect of minimal rights* implies the *participation* property.

Proposition 1. *S_{RMRE} satisfies RMR and P.*

4. Conclusions

In this paper we have analyzed redistribution problems by means of a lower bound on what individuals deserve. We have modified the mechanism proposed by Luttens (2010) so that our proposal not only makes *claims feasibility* and *participation* compatible, but also it fulfills the bound on which is based: *respect of minimal rights* (*RMR*). Our proposal behaves as the *CF*-mechanism of Luttens for a large level of resources. But we obtain that for a small level of resources, (i) no-one can incur a negative income, and (ii) no-one can receive more than their claim when the resources are not enough to satisfy the aggregate claim (*claim-boundedness*), two usual requirements in conflicting claims problems. Figure 1 shows how our mechanism works, where the horizontal and vertical axes represent different levels of the resources and the total income received by each individual, in a four-individual problem, respectively.

Finally an interesting ongoing issue is to analyze the behavior of this kind of egalitarian mechanism whenever other lower bounds considered in the literature are used.

Acknowledgement We acknowledge the comments and valuable suggestions of two anonymous referees and the Editor that have substantially improved the paper. This has been partially supported by the Spanish Ministry of Economy and Competitiveness funds under Project ECO2013-43119 and by Universitat Rovira i Virgili, Banco Santander and Generalitat de Catalunya under the project 2011LINE-06.

References

Bossert, W. (1995). Redistribution Mechanisms Based on Individual Factors. *Mathematical Social Sciences*, 29, 1–17.

Bossert, W. and Fleurbaey, M. (1996). Redistribution and Compensation. *Social Choice and Welfare*, 13, 343–355.

Fleurbaey, M. (1994). On Fair Compensation. *Theory and Decision*, 36, 277–307.

Fleurbaey, M. (2008). *Fairness, Responsibility, and Welfare*. Oxford, Oxford University Press.

Fleurbaey, M. and Maniquet, F. (2011). Compensation and Responsibility. In Arrow, A. K., and Sen, K. (eds.), *Handbook of Social Choice and Welfare*, vol. 2. Amsterdam, Elsevier/New Holland, p. 507–604.

Foley, D. K. (1967). Resource Allocation and the Public Sector. *Yale Economics Essays*, 7(1), 45–98.

Iturbe-Ormaetxe, I. (1997). Redistribution and Individual Characteristics. *Review of Economic Design*, (3), 45–55.

Kolm, S. C. (1972). *Justice et équité*. Paris, Editions du CNRS.

Luttens, R. I. (2010). Minimal Rights Based Solidarity. *Social Choice and Welfare*, 34(1), 47–64.

Maniquet, F. (1998). An Equal Right Solution to the Compensation-Responsibility Dilemma. *Mathematical Social Sciences*, 35(2), 671–687.

O'Neill, B. (1982). A Problem of Rights Arbitration from the Talmud. *Mathematical Social Sciences*, 2(4), 345–371.

Panzer, E. and Schmeidler, D. (1974). A Difficulty in the Concept of Fairness. *Review of Economic Studies*, 41(3), 441–443.

Pazner, E. and Schmeidler, D. (1978). Egalitarian Equivalent Allocations: A New Concept of Economic Equity. *Quaterly Journal of Economics*, 92, 671–687.

Rawls, J. (1971). *A Theory of Justice*. Cambridge, MA, Harvard University Press.

Thomson, W. (2003). Axiomatic and Game-Theoretic Analysis of Bankruptcy and Taxation Problems: A Survey. *Mathematical Social Sciences*, 45(3), 249–297.

Tungodden, B. (2005). Responsibility and Redistribution: The Case of First Best Taxation. *Social Choice and Welfare*, 24, 33–44.

Varian, H. (1974). Equity, Envy and Efficiency. *Journal of Economic Theory*, (9), 63–91.

Appendix

A1. Proof of Theorem 1

Given an economy $e = (y_N, z_N)$, and a claims vector $\hat{\mathbf{g}}$, we define the economy $e' = (y'_N, z_N)$ where y'_N is chosen such that $R' = C$. Since the efforts does not change, the claims vector of this new economy is $\hat{\mathbf{g}}$. Note that, for each $i \in N$, $m_i(e', \hat{\mathbf{g}}) = g(z_i)$. Hence the CF axiom implies that the "initial income" is $u_i(e', \hat{\mathbf{g}}) = g(z_i)$, for each $i \in N$. One of the following situations occurs:

(i) $R \geq C$

For each $i \in N$, $R - C_{-i} \geq C - C_{-i} = g(z_i)$. Thus, $m_i(e, \hat{\mathbf{g}}) = g(z_i)$, and $m_i(e, \hat{\mathbf{g}}) - m_i(e', \hat{\mathbf{g}}) = 0$. By AS^{**}, $u_i(e, \hat{\mathbf{g}}) = g(z_i) + \frac{R-C}{n}$, which coincides with (i) of Definition 3.

(ii) $C_{-n} \leq R < C$

For each $i \in N$, $R - C_{-i} = (R - C_{-n}) + (C_{-n} - C_{-i}) = (R - C_{-n}) + (g(z_i) - g(z_n)) = (R - C) + g(z_i) < g(z_i)$, and $R - C_{-i} \geq C_{-n} - C_{-i} = g(z_i) - g(z_n) \geq 0$. Thus, $m_i(e, \hat{\mathbf{g}}) = (R - C) + g(z_i)$. Hence, $m_i(e, \hat{\mathbf{g}}) - m_i(e', \hat{\mathbf{g}}) = R - R' = R - C$. By AS^{**}, $u_i(e, \hat{\mathbf{g}}) = g(z_i) + \frac{R-C}{n}$, which coincides with (i) of Definition 3.

(iii) $C_{-(n-1)} \leq R < C_{-n}$

Consider now the economy $e' = (y'_N, z_N)$ where y'_N is chosen such that $R' = C_{-n}$. Since the efforts does not change, the claims vector of this new economy is $\hat{\mathbf{g}}$. From (ii) we know that $u_i(e', \hat{\mathbf{g}}) = g(z_i) + \frac{C_{-n}-C}{n}$. For each $i \in N$, $R - C_{-i} = (R - C_{-n}) + (C_{-n} - C_{-i}) = (R - C_{-n}) + (g(z_i) - g(z_n)) = (R - C) + g(z_i) < g(z_i)$. For each $i \leq n-1$, $R - C_{-i} \geq 0$, and $R - C_{-n} < 0$. Thus, $m_i(e, \hat{\mathbf{g}}) = (R - C) + g(z_i)$ for each $i \leq n-1$ and $m_n(e, \hat{\mathbf{g}}) = 0$. Hence, $m_i(e, \hat{\mathbf{g}}) - m_i(e', \hat{\mathbf{g}}) = (R - C) + g(z_i) - (C_{-n} - C + g(z_i)) = R - C_{-n}$, for each $i \leq n-1$ and $m_n(e, \hat{\mathbf{g}}) - m_n(e', \hat{\mathbf{g}}) = 0$. By PRI and AS^{**}, $u_i(e, \hat{\mathbf{g}}) = u_i(e', \hat{\mathbf{g}}) + \frac{R-C_{-n}}{n-1}$, for each $i \leq n-1$ and $u_n(e, \hat{\mathbf{g}}) = u_n(e', \hat{\mathbf{g}})$ which coincides with (ii) of Definition 3.

(iv) $C_{-(k-1)} \leq R < C_{-k}$, $k = 2, 3, \ldots, n-1$

The proof of this case is completely analogous to that in (iii), just by considering the economy $e' = (y'_N, z_N)$ where y'_N is chosen such that $R' = C_{-k}$.

(v) $G_{n-1} \leq R < C_{-1}$

We consider the economy $e' = (y'_N, z_N)$ where y'_N is chosen such that $R' = C_{-1}$. Since the efforts does not change, the claims vector of this new economy is $\hat{\mathbf{g}}$. For each $i \in N$, $R - C_{-i} = (R - C_{-1}) + (C_{-1} - C_{-i}) = (R - C_{-1}) + (g(z_i) - g(z_1)) < 0$. Thus, $m_i(e, \hat{\mathbf{g}}) = 0$. Hence, $m_i(e, \hat{\mathbf{g}}) - m_i(e', \hat{\mathbf{g}}) = 0$. By AS^{**}, $u_i(e, \hat{\mathbf{g}}) = u_i(e', \hat{\mathbf{g}}) + \frac{R-C_{-1}}{n}$, which coincides with (iii) of Definition 3.

(vi) $G_{n+1-k} \leq R < G_{n+2-k}$, for $k = 2, 3, \ldots, n-2$

Finally, consider the economy $e' = (y'_N, z_N)$ where y'_N is chosen such that $R' = G_{n+2-k}$. Since the efforts does not change, the claims vector of this new economy is $\hat{\mathbf{g}}$. For each $i \in N$, $R - C_{-i} < 0$. Thus, $m_i(e, \hat{\mathbf{g}}) = 0$. Hence, $m_i(e, \hat{\mathbf{g}}) - m_i(e', \hat{\mathbf{g}}) = 0$. By AS^{**}, $u_i(e, \hat{\mathbf{g}}) = u_i(e', \hat{\mathbf{g}}) + \frac{R - G_{n+2-k}}{k}$, for each $i \leq n+2-(k+1)$ and $u_i(e, \hat{\mathbf{g}}) = 0$, otherwise, which coincides with (iv) of Definition 3.

Finally, it is immediate to observe that the utility functions obtained from the redistribution mechanism S_{RMRE} satisfy the required axioms. □

A2. Axioms Independence

In the next examples we show that the axioms in Theorem 1 are independent. In all of them, there are three individuals which are decreasingly ordered, as usually, that is, $g_1 = g(z_1) \geq g_2 = g(z_2) \geq g_3 = g(z_3)$.

(i) **A mechanism fulfilling** *CF*, AS^{**} **and not** *PRI*.

Let S_a the mechanism defined by:

(a) If $R < C$ and $m_3 < m_2 = m_1 < g_2$, then $u_i = \frac{R}{3}$.

(b) $S_a = S_{RMRE}$, otherwise.

It is clear that *CF* and AS^{**} are fulfilled.

Nonetheless, S_a does no satisfy *PRI*, as the next numerical example shows. Consider two economies $e = (y_N, z_N)$, $e' = (y'_N, z_N) \in \mathcal{E}$, such that $R = 3$, $R' = 12$, and $\hat{\mathbf{g}} = (9, 9, 1)$. Then, $u(e, \hat{\mathbf{g}}) = (1, 1, 1)$, $m(e, \hat{\mathbf{g}}) = (0, 0, 0)$, and $u(e', \hat{\mathbf{g}}) = (4, 4, 4)$, $m(e', \hat{\mathbf{g}}) = (1, 1, 0)$. So S_a does not satisfy *PRI*.

(ii) **A mechanism fulfilling** *CF*, *PRI* **and not** AS^{**}.

Let S_b the mechanism defined by:

$$u_1 = R - u_2 - u_3, \quad u_2 = m_2, \quad u_3 = m_3.$$

It is clear that S_b satisfies *CF* and *PRI*.

Nonetheless, S_b does no satisfy AS^{**} as the next numerical example shows. Consider two economies $e = (y_N, z_N)$, $e' = (y'_N, z_N) \in \mathcal{E}$, such that $R = 12$, $R' = 15$, and $\hat{\mathbf{g}} = (10, 9, 1)$. Then, $u(e, \hat{\mathbf{g}}) = (11, 1, 0)$, $m(e, \hat{\mathbf{g}}) = (2, 1, 0)$, and $u(e', \hat{\mathbf{g}}) = (11, 4, 0)$, $m(e', \hat{\mathbf{g}}) = (5, 4, 0)$. So S_b does not satisfy AS^{**}.

(iii) **A mechanism fulfilling** *PRI*, AS^{**} **and not** *CF*.

Let S_c the mechanism defined by:

(a) If $m_3 < m_2 = m_1 \leq g_2$, then $u_3 = g_3$, and $u_2 = u_1 = \frac{R - g_3}{2}$.

(b) If $m_3 < m_2 = g_2 < m_1$, then $u_3 = g_3$, $u_2 = \frac{3g_2 - g_3}{2}$, and $u_1 = R - u_2 - u_3$.

(c) $S_c = S_{RMRE}$, otherwise.

It is clear that *PRI* and AS^{**} are fulfilled.

Nonetheless, S_c does no satisfy *CF* as the next numerical example shows. Consider an economy $e = (y_N, z_N) \in \mathcal{E}$, such that $R = 21$, and $\hat{\mathbf{g}} = (14, 6, 1)$. Then, $m(e, \hat{\mathbf{g}}) = (14, 6, 1)$ and $u(e, \hat{\mathbf{g}}) = (11.5, 8.5, 1)$. So S_c does not satisfy CF.

Czech Economic Review 9 (2015) 49–73
Acta Universitatis Carolinae Oeconomica

Robust Turnpikes Deduced by the Minimum-Time Needed toward Economic Maturity

Darong Dai*

Received 13 October 2014; Accepted 21 September 2015

Abstract In the paper, a one-sector neoclassical model with stochastic growth has been constructed. The major goal of the study is to characterize relevant mathematical properties of efficient development paths for underdeveloped economies. Since economic maturity is a reasonable objective, we mainly focus on the long-run features of economic development. Indeed, the notion of economic maturity is well-defined in the model, and also a thorough characterization of the minimum time needed toward economic maturity is offered with intuitive interpretations discussed. Moreover, it is confirmed that the capital-labor ratio corresponding to the state of economic maturity provides us with a robust turnpike of the optimal path of capital accumulation.

Keywords Stochastic growth, economic maturity, minimum-time objective, asymptotic turnpike theorem, neighborhood turnpike theorem, robustness
JEL classification C60, E13, E22

1. Introduction

When concerning the issue of economic development for underdeveloped economies, the principle of maximum speed is widely employed. In reality, the Germany and Japan after World War II and China after 1978s (see Song et al. 2011) are typical examples.

Alternatively, provided the existence of maximum sustainable terminal path consumption per capita (or von Neumann path consumption per capita), which would be regarded as the state of economic maturity in a certain sense, the major goal of people and government is to choose appropriate or optimal savings strategy and fiscal policies, respectively, such that the state of economic maturity can be reached as soon as possible. Indeed, the underlying motivation of the present exploration, which is in line with Kurz (1965), is to derive conditions under which the specified economy can reach the maximum terminal path in a minimum time. In particular, we analyze the economy before reaching economic maturity, and hence we focus on underdeveloped economies and leave those economies having reached economic maturity to future research.

Although we focus on a one-sector neoclassical aggregate growth model (see Solow 1956; Cass 1965; Dai 2013, 2014a, 2014c), the present study extends Kurz's analyses in the following ways. First, we consider an economy lying in a persistently non-stationary environment. Second, nature (or social planner) is incorporated into the

* Texas A&M University, College Station, Department of Economics, TX, 77843, United States. E-mail: dai496@tamu.edu.

macroeconomic model, and the endogenous savings rate and the minimum time form a sub-game perfect Nash equilibrium (SPNE) of the stochastic dynamic game between the nature and the representative agent. Third, the minimum time needed to reach economic maturity is completely characterized by the maximum sustainable level of terminal path capital-labor ratio (i.e., the state corresponding to economic maturity), and also the terminal path of capital-labor ratio provides us with a robust turnpike (i.e., the equilibrium path of capital accumulation will robustly converge to this terminal path in an asymptotic sense or will spend almost all time staying in a neighborhood of the terminal path). In addition, rather than letting the terminal capital-labor ratio be exogenously given or prescribed as in Kurz (1965), Samuelson (1965) and Cass (1966), the maximum sustainable level of terminal path consumption per capita (or capital-labor ratio) is endogenously determined in the present model.

The rest of the paper is organized as follows. In Section 2, the basic model is constructed, and some necessary assumptions and definitions, especially the definitions of economic maturity and the minimum time needed to economic maturity, are introduced. Section 3 is the major part of the paper, where both Asymptotic Turnpike Theorem and Neighborhood Turnpike Theorem are established. Section 4 proves robustness of the turnpike theorems established in Section 3, i.e., we assert the existence of a robust turnpike deduced by the minimum-time needed to economic maturity. There is a brief concluding section, where we have discussed possible extensions of the basic framework. All proofs, unless otherwise noted in the text, appear in the Appendix.

2. The environment

Here, and throughout the paper, we consider a one-sector neoclassical model with stochastic growth. As usual, we employ the following neoclassical production function

$$Y(t) = F\left(K(t), L(t)\right), \tag{1}$$

which is a strictly concave function and exhibits constant returns to scale (CRS) with $K(t)$ denoting the aggregate capital stock and $L(t)$ representing the labor force (or population size in some cases). Thus, we have the following law of motion of capital accumulation

$$\frac{dK(t)}{dt} = F\left(K(t), L(t)\right) - \delta K(t) - C(t), \tag{2}$$

where δ, an exogenously given constant, denotes the depreciation rate and $C(t)$ stands for aggregate consumption in period t.

Suppose that $(B(t),\ 0 \leq t \leq T)$ stands for a standard Brownian motion defined on the following filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\leq t\leq T}, \mathbb{P})$ with $\mathbb{F} \equiv \{\mathcal{F}_t\}_{0\leq t\leq T}$ the $\mathbb{P}$-augmented filtration generated by $(B(t),\ 0 \leq t \leq T)$ with $\mathcal{F} \equiv \mathcal{F}_T$ for $\forall T > 0$, i.e., the underlying stochastic basis satisfies the well-known usual conditions. Then, based upon the given probability space and following Merton (1975) and Dai (2014a), we define the following law of motion for labor force

$$dL(t) = nL(t)dt + \sigma L(t)dB(t) \tag{3}$$

subject to $B(0) = 0$ almost surely (hereafter a.s.)-$\mathbb{P}$ and $\sigma \in \Re_0 \equiv \Re - \{0\}$, a constant. Thus, combining (2) with (3) and applying Itô's rule lead us to

$$dk(t) = \left[s(k(t)) f(k(t)) - \left(\delta + n - \sigma^2\right) k(t)\right] dt - \sigma k(t) dB(t) \tag{4}$$

with $k(0) \equiv k_0 > 0,\ k(t) \equiv K(t)/L(t),\ f(k(t)) \equiv F(K(t), L(t))/L(t) = F\left(\frac{K(t)}{L(t)}, 1\right)$, $s(k(t)) \equiv 1 - \frac{c(t)}{f(k(t))}$ and $c(t) \equiv C(t)/L(t)$ denoting the initial capital-labor ratio, capital-labor ratio, per capita output, savings per unit output and per capita consumption, respectively, at time t. Specifically, for the SDE of capital-labor ratio given by (4), Chang and Malliaris (1987) proved the following theorem.

Theorem 1. *If the production function f is strictly concave, continuously differentiable on $[0, \infty)$, $f(0) = 0$, and $\lim_{k(t)\to\infty} f'(k(t)) \equiv \lim_{k(t)\to\infty} \frac{df(k(t))}{dk(t)} = 0$, then there exists a unique solution to* (4).

Thus, we directly give the following assumption for simplicity.

Assumption 1. *The assumptions or conditions given by Theorem 1 are fulfilled throughout the current paper.*

2.1 Economic Maturity

It is assumed that the economy consists of $L(t)$ identical individuals in period t, each of whom possesses perfect foresight. We thus suppose that there is a representative agent with the following objective function:

$$\mathbb{E}_{t_0}\left[\int_{t_0}^{\tau} e^{-\rho(t-t_0)} U_1\left((1 - s(k(t))) f(k(t))\right) dt + e^{-\rho(\tau - t_0)} U_2\left(f(k(\tau))\right)\right], \tag{5}$$

where $\mathbb{E}_{t_0}$ denotes the expectation operator depending on $\mathcal{F}_{t_0}$ with $t_0 \geq 0$, $0 < \rho < 1$ represents the discount factor, $\tau \equiv \tau(\omega) \in \mathcal{T} \equiv \{\mathcal{F}\text{-stopping times}\}$ for $\omega \in \Omega$, and $U_1(\cdot)$, $U_2(\cdot)$ are strictly increasing, strictly concave instantaneous utility functions of per capita consumption and per capita output, respectively, with the well-known Inada conditions satisfied.

It is easily seen that the criterion defined by (5) is widely used in existing literature, including the macroeconomic studies. Nevertheless, $\tau \equiv \tau(\omega)$ is usually pre-specified and deterministic, e.g., $\tau(\omega) \equiv T > 0$ for all $\omega \in \Omega$ and any exogenously given constant $0 < T \leq \infty$. Noting that τ truly implies interesting and important economic implications in accordance to Kurz (1965) and Dai (2012), we will extend Kurz's work by introducing nature or social planner into the present macroeconomic model. The nature will actually choose an admissible value $\tau^* \equiv \tau^*(\omega)$ so that (5) is maximized. Formally, we give the following definition.

Definition 1. The *stochastic dynamic game* Γ between the nature and the representative agent proceeds according to the following timing structure:

Stage 1: Taking the remaining parameters as given, the nature will determine an optimal stopping time $\tau^*(\omega) \in \mathcal{T}$ such that the criterion in (5) is maximized subject to constraint (4) (i.e., this is essentially an optimal stopping problem).

Stage 2: Given the knowledge of the game structure as well as $\tau = \tau^*(\omega) \in \mathcal{T}$, the representative agent chooses an optimal savings strategy $s^*(k(t), \tau^* - t_0)$ such that the criterion defined in (5) is maximized subject to constraint (4).

Then, following the classical Backward Induction Principle, we formulate:

Problem 1. The representative agent will find a savings policy $s^*(k(t), \tau - t_0)$ so as to

$$\max \mathbb{E}_{t_0}\left[\int_{t_0}^{\tau} e^{-\rho(t-t_0)} U_1\left(\left(1 - s\left(k(t)\right)\right) f\left(k(t)\right)\right) dt + e^{-\rho(\tau-t_0)} U_2\left(f\left(k(\tau)\right)\right)\right]$$

subject to the SDE of capital-labor ratio in (4), for $\forall \tau \in \mathcal{T}$.

If Problem 1, the modified Ramsey (1928) problem, has a solution, we obtain the optimal path of capital-labor ratio as follows:

$$dk(t) = \left[s^*\left(k(t), \tau - t_0\right) f\left(k(t)\right) - \left(\delta + n - \sigma^2\right) k(t)\right] dt - \sigma k(t) dB(t). \tag{6}$$

And we put:

Problem 2. The optimization problem facing the nature is to find a stopping rule $\tau^*(\omega) \in \mathcal{T}$ so as to

$$\sup \mathbb{E}_{t_0}\left[\int_{t_0}^{\tau} e^{-\rho(t-t_0)} U_1\left(\left(1 - s^*\left(k(t), \tau - t_0\right)\right) f\left(k(t)\right)\right) dt + e^{-\rho(\tau-t_0)} U_2\left(f\left(k(\tau)\right)\right)\right]$$

subject to the SDE of capital-labor ratio given by (6).

Remark 1.

(i) It is especially worth emphasizing that Problem 2 can also be modified by focusing entirely upon the final state as that of Radner (1961) and Dai (2012). That is, the criterion of preference facing the nature is given by

$$\mathbb{E}_{t_0}\left[e^{-\rho(\tau-t_0)} U_2\left(f\left(k(\tau)\right)\right)\right],$$

which, in general, will result in a new turnpike. Nevertheless, we argue that similar turnpike theorems can be established for the new turnpike.

(ii) In particular, one may notice certain similarity of the present approach to the literature studying endogenous lifetime or endogenous longevity in growth models (see Chakraborty 2004; de la Croix and Ponthiere 2010, and among others), obvious differences, nevertheless, exist between the both, especially when referring to economic intuition and implications behind formal models. Existing studies focus on OLG models and health-investment behaviors while the current exploration emphasizes issues of macroeconomic development, namely, the characterization of economic maturity for underdeveloped economies and the corresponding characteristics of their optimal capital-accumulation paths.

(iii) The *maximum sustainable capital-labor ratio* corresponding to the state of economic maturity as well as the *minimum-time needed to economic maturity* is endogenously determined by using stochastic optimal stopping theory that is widely applied in mathematical finance. However, in Kurz's (1965) study, the target or the maximum sustainable level of terminal path capital-labor ratio is exogenously specified, and the corresponding minimum time problem is expressed as: For any given initial capital-labor ratio, to chose strategies so that the prescribed target can be reached as soon as possible. As a result, the major contribution of the present approach can be expressed as follows: first, we endogenously determine the terminal path of capital-labor ratio as well as the minimum time needed to reach economic maturity; second, we maximize the welfare of the representative agent in solving the minimum-time objective problem.

(iv) It follows from the specification of Problem 2 that we focus on the episode before reaching economic maturity as concentrated in Kurz (1965), Samuelson (1965) and Cass (1966). Put it differently, the present framework is suitable for the studies concerning underdeveloped economies.

Thus, if Problem 2 has a solution, we get the optimal stopping time $\tau^*(\omega) \in \mathcal{T}$, which actually defines the *minimum time needed toward economic maturity*. Also, $(\tau^*(\omega), s^*(k(t), \tau^*(\omega) - t_0))$ forms the sub-game perfect Nash equilibrium (SPNE) of the stochastic dynamic game Γ given by Definition 1.

Remark 2. It is especially worth mentioning that we define the standard of economic maturity from the perspective of economic welfare, which is of course reasonable in the current model economy. Notice that the state of economic maturity for any given economy should imply a peak state that yields the highest level of economic welfare,[1] we argue that the minimum time needed to economic maturity is well-defined.

Finally, noting that we do not focus on the endogenous savings behavior of the representative agent and also the explicit formula of the minimum time needed to economic maturity in the current paper, we can directly put:

Assumption 2. *Let Problem 1 and Problem 2 be solvable, i.e., we can find at least one optimal savings policy $s^*(k(t), \tau^*(\omega) - t_0)$ and at least one minimum time $\tau^*(\omega) \in \mathcal{T}$ needed toward economic maturity. Moreover, let there exist a constant capital-labor ratio $0 < k^* < \infty$ such that the optimal stopping rule is characterized by $\tau^*(\omega) \equiv \inf\{t \geq 0; k(t) = k^*\} < \infty$ a.s.-$\mathbb{P}$.*

[1] We, of course, admit that there are many other standards that can characterize the state of economic maturity. Nevertheless, we argue that *economic welfare* will always be the appropriate choice when noting that the major goal of economic growth and economic development is to improve the economic welfare of the people for any modern economies. And in order to make things easier and tractable, we focus on the highest level of economic welfare, and this assumption is, however, without any loss of generality in the underlying economy.

Remark 3.

(i) In fact, Problem 1 can be solved by employing stochastic dynamic programming, and Merton (1975) proved the existence of optimal savings policy in a quite similar case. On the other hand, Problem 2 can also be solved under certain conditions, and one can refer to Karatzas and Wang (2001), Jeanblanc et al. (2004), and Øksendal and Sulem (2005) for more details. The major goal of the present exploration is to confirm that k^* defines a robust turnpike, which is certainly deduced by economic maturity based on the above constructions.

(ii) Assumption 2 ensures the existence of turnpikes from the viewpoint of mathematical techniques. We, however, emphasize that the existence can be taken for granted in reality. In other words, for any developed economy, it experienced the state of economic maturity in history. Thus, the existence of the state of economic maturity is relatively easily ensured in reality.

3. Turnpike theorems

Now, based on Assumption 2, we get

$$\begin{aligned} dk(t) &= \left[s^*\left(k(t), \tau^* - t_0\right) f\left(k(t)\right) - \left(\delta + n - \sigma^2\right) k(t)\right] dt - \sigma k(t) dB(t) \\ &\equiv \varphi\left(k(t)\right) dt + \psi\left(k(t)\right) dB(t) \end{aligned} \tag{7}$$

subject to $k(0) \equiv k_0 > 0$, a deterministic constant. And also,

$$\tau^*(\omega) \equiv \inf\{t \geq 0; k(t) = k^*\} < \infty \text{ a.s.-}\mathbb{P} \tag{8}$$

for some endogenously given constant $0 < k^* < \infty$. We are to show that k^* exhibits turnpike property providing the above assumptions.

Theorem 2 (**Asymptotic Turnpike Theorem**).[2] *Provided the SDE of capital-labor ratio defined in (7) and the minimum time needed to economic maturity given by (8), then we always get that $k(t)$ converges in $L^1(\mathbb{P})$ and the corresponding limit belongs to $L^1(\mathbb{P})$. In particular, if we have $\varphi\left(k(t)\right) = 0$ a.s.-$\mathbb{P}$, i.e., $s^*\left(k(t), \tau^* - t_0\right) f\left(k(t)\right) = (\delta + n - \sigma^2)k(t)$ a.s.-$\mathbb{P}$, it uniformly converges to k^* a.s.-$\mathbb{P}$, or equivalently,*

$$\lim_{t' \to \infty} \mathbb{P}\left(\bigcup_{t=t'}^{\infty} [|k(t) - k^*| \geq \varepsilon]\right) = 0, \quad \forall \varepsilon > 0.$$

Proof. See Appendix A.

Remark 4. By applying supermartingale property to confirm the corresponding convergence, Joshi (1997) studies the turnpike theory in a stochastic aggregate growth

[2] This proof brings the idea from Dai (2012). Our turnpike theorems satisfy the classical characteristics, i.e., any optimal paths stay within a small neighborhood of the turnpike almost all the time and the turnpike is independent of initial conditions (see McKenzie 1976; Yano 1984; Dai 2014c).

model in which stochastic environments as independent variables are directly and exogenously incorporated into the production function. However, one may easily tell the difference between Joshi's method and our proof. Moreover, it is argued that the essential requirement in Theorem 2 can be easily met thanks to the volatility term σ.

However, if $\varphi(k(t)) \neq 0$, we can define a new process $\theta(t)$ by

$$\varphi(k(t)) = \theta(t)\psi(k(t))$$

for almost all (hereafter a.a.) $(t,\omega) \in [0,T] \times \Omega$. Then we can put

$$Z(t) \equiv \exp\left\{ -\int_0^t \theta(s)dB(s) - \frac{1}{2}\int_0^t \theta^2(s)ds \right\}.$$

Define a new measure $\mathbb{Q}$ on $\mathcal{F}_T$ by

$$d\mathbb{Q}(\omega) = Z(T)d\mathbb{P}(\omega),$$

i.e., $Z(T)$ is the Radon-Nikodym derivative. In what follows, we first introduce the following assumption:

Assumption 3. *At least one of the following two conditions holds:*

(i) $\mathbb{E}[Z(T)] = 1$.

(ii) *The following Novikov Condition holds, i.e.,*

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T \theta^2(t)dt\right)\right] < \infty \ \textit{for} \ \ 0 \leq T < \infty.$$

Thus, based upon Assumption 3 and according to the Girsanov Theorem, we get that $\mathbb{Q}$ is a probability measure on $\mathcal{F}_T$, $\mathbb{Q}$ is equivalent to $\mathbb{P}$ and $k(t)$ is a martingale w. r. t. $\mathbb{Q}$ on the stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q})$. Using Girsanov Theorem again, we claim that the process

$$\hat{B}(t) \equiv \int_0^t \theta(s)ds + B(t), \quad \forall t \in [0,T]$$

is a Brownian motion w. r. t. $\mathbb{Q}$ with $\hat{B}(0) = B(0) = 0$ a.s., and expressed in terms of $\hat{B}(t)$ we can get

$$dk(t) = \psi(k(t))\, d\hat{B}(t) \tag{9}$$

subject to $k(0) \equiv k_0 > 0$, a deterministic constant.

Next, based on (9) and similar to (8), we, by slightly modifying Assumption 2, have

$$\hat{\tau}^*(\omega) \equiv \inf\left\{t \geq 0; k(t) = \hat{k}^*\right\} < \infty \ \text{ a.s.-}\mathbb{Q} \tag{10}$$

for some endogenously determined $0 < \hat{k}^* < \infty$.

Here, the operation of changing probability measure is mainly technically motivated. These turnpike properties rely on a martingale feature of the equilibrium path of capital accumulation. Before changing the probability measure, we just focus attention on a special case where the equilibrium path of capital accumulation is already a martingale process. If we relax the assumption used in such a special case, we then need to change the probability measure to obtain a martingale process by making use of the well-known Girsanov Theorem. Therefore, employing similar proof as that of Theorem 2, we can establish:

Theorem 3 (**Asymptotic Turnpike Theorem**). *Provided the SDE of capital-labor ratio defined in (9) and the minimum time needed to economic maturity given by (10), then we always get that $k(t)$ converges in $L^1(\mathbb{Q})$ and the corresponding limit belongs to $L^1(\mathbb{Q})$. Specifically, it uniformly converges to $\hat{k}^*$ a.s.-$\mathbb{Q}$, or equivalently,*

$$\lim_{t'\to\infty} \mathbb{Q}\left(\bigcup_{t=t'}^{\infty}\left[\left|k(t)-\hat{k}^*\right| \geq \varepsilon\right]\right) = 0, \quad \forall \varepsilon > 0.$$

Now, we proceed to prove the neighborhood turnpike theorem. We do this by first giving the following assumption.

Assumption 4. *Let $k(t) \in \Re_{++} \equiv [0,\infty]$, which is the one point compactification of $\Re$ at infinity with the induced topology, $\forall t \geq 0$. Also, there exists a unique invariant Borel probability measure π defined on $\Re_{++}$ such that $\pi\left[bd\left(\Re_{++}\right)\right] \equiv \pi\left[\{0\}\bigcup\{+\infty\}\right] = 0$, where $bd\left(\Re_{++}\right)$ denotes the boundary of $\Re_{++}$. We particularly denote by $\hat{\pi}$ the Borel probability measure corresponding to the SDE defined in (9).*

Remark 5. Mirman (1972) and Dai (2014c) construct a one-sector growth model with uncertain technology, i.e., random variables, which are assumed to be independent and identically distributed, are directly introduced into the neoclassical production function, thereby resulting in a discrete-time Markov process of capital stock. Specifically, Mirman defines Borel probability measure on the Borel sets of non-negative real line by using the corresponding probability transition function of the Markov process. Moreover, Theorem 2.1 of Mirman confirms that there exists a stationary probability measure that has no mass at either zero or infinity. In contrast, the present paper constructs continuous time Markov process of capital-labor ratio. Nonetheless, one can still prove that there exists a unique invariant Borel probability measure satisfying the requirements of Assumption 4 under relatively weak conditions. For more details, one may refer to Theorem 2.1 of Imhof (2005), Theorem 3.1 of Benaïm et al. (2008) and Theorem 5 of Schreiber et al (2011). The present paper omits the corresponding proof to economize on the space.

As a consequence, the following theorem is derived.

Theorem 4 (**Neighborhood Turnpike Theorem**).[3] *Provided assumptions of Theorem 2 are fulfilled and also Assumption 4 holds, we can get that there exists a constant*

[3] This proof brings the method employed by Imhof (2005) and Dai (2012).

$\Sigma > 0$ *such that for* $\forall \alpha > 0$ *with* $\alpha > \Sigma$,

(i) $\mathbb{E}\left[\tau_{\overline{B}_\alpha(k^*)}(\omega)\right] \leq \dfrac{dist\,(k_0,k^*)}{\alpha-\Sigma}$,

(ii) $\pi\left[\overline{B}_\alpha(k^*)\right] \geq 1-\frac{\Sigma}{\alpha} \equiv 1-\varepsilon$,

where

$$\begin{aligned} B_\alpha(k^*) &\equiv \{k(t)\in\Re_{++}; |k(t)-k^*| < \alpha, t\geq 0\}, \\ \tau_{\overline{B}_\alpha(k^*)}(\omega) &\equiv \inf\left\{t\geq 0; k(t)\in \overline{B}_\alpha(k^*) \equiv clB_\alpha(k^*)\right\}, \text{ and} \\ dist\,(k_0,k^*) &\equiv k^*\log(k^*/k_0) \end{aligned}$$

for $(k^* >)k_0 \equiv k(0) > 0$.

Proof. See Appendix B.

In particular, this result just considers the case with $k^* > k_0$. Definitely, we can obtain similar result for the case with $k^* < k_0$ through redefining the distance function as $dist\,(k_0,k^*) \equiv k_0\log(k_0/k^*)$.

Similarly, we derive the following theorem.

Theorem 5 (**Neighborhood Turnpike Theorem**). *Provided assumptions of Theorem 3 are fulfilled and also Assumption 4 holds, we can get that there exists a constant* $\hat{\Sigma} > 0$ *such that for* $\forall\hat{\alpha} > 0$ *with* $\hat{\alpha} > \hat{\Sigma}$,

(i) $\mathbb{E}^{\mathbb{Q}}\left[\hat{\tau}_{\overline{B}_{\hat{\alpha}}(\hat{k}^*)}(\omega)\right] \leq \dfrac{dist\left(k_0,\hat{k}^*\right)}{\hat{\alpha}-\hat{\Sigma}}$,

(ii) $\hat{\pi}\left[\overline{B}_{\hat{\alpha}}(\hat{k}^*)\right] \geq 1-\frac{\hat{\Sigma}}{\hat{\alpha}} \equiv 1-\hat{\varepsilon}$,

where

$$\begin{aligned} B_{\hat{\alpha}}\left(\hat{k}^*\right) &\equiv \left\{k(t)\in\Re_{++}; \left|k(t)-\hat{k}^*\right| < \hat{\alpha}, t\geq 0\right\}, \\ \hat{\tau}_{\overline{B}_{\hat{\alpha}}(\hat{k}^*)}(\omega) &\equiv \inf\left\{t\geq 0; k(t)\in \overline{B}_{\hat{\alpha}}(\hat{k}^*) \equiv clB_{\hat{\alpha}}\left(\hat{k}^*\right)\right\}, \text{ and} \\ dist\left(k_0,\hat{k}^*\right) &\equiv \hat{k}^*\log\left(\hat{k}^*/k_0\right) \end{aligned}$$

for $(\hat{k}^* >)k_0 \equiv k(0) > 0$.

Remark 6. Theorem 4 shows that the Borel probability measure π will place nearly all mass close to the turnpike k^*. Similarly, Theorem 5 reveals that the corresponding probability distribution $\hat{\pi}$ will place almost all mass close to the new turnpike $\hat{k}^*$. Indeed, Theorems 4 and 5 demonstrate the turnpike property from both time dimension and space dimension, i.e., in the sense of Markov time as well as invariant probability distribution, which of course will provide us with a much more complete characterization of the neighborhood turnpike property when compared with existing studies (see McKenzie 1976; Bewley 1982; Yano 1984; Dai 2014c).

What's the potential application of our theoretical result? It seems hard to see any direct application of our abstract assertion, we, however, will offer the following implication to reveal the potential practical-value of our theoretical argument. The finding in Theorem 5 leads us to a much more comprehensive philosophy when we are motivated to comparatively analyze capital accumulation within different economic systems. For example, for two economies with different levels of economic maturity, e.g., the first one is relatively higher than the second one. Hence, we usually claim that the first one will economically dominate the second one. Nonetheless, our result argues that this conclusion is really hasty and hence may not be comprehensive, and it even does not make any sense. Why? When we attempt to evaluate the potential of capital accumulation for different economies, we should simultaneously consider efficiency from the time aspect, e.g., the first economy may take 15 years to reach its neighborhood efficiency, whereas the second one only takes 5 years. In other words, Theorem 5 confirms that both the height of our goal and the speed leading toward our goal are equivalently crucial from the perspective of evaluating economic efficiency.

Not only that, we are encouraged to add the following comment for Theorem 5. It is worthwhile indicating that there exists an intriguing relation between our major result and the concept of flexibility. In fact, we understand the concept of flexibility under the current background like this: It captures the dynamic trade-off between evaluation accuracy and sustainable economic incentive. To be exact, the selected scope or radius of the given neighborhood completely determined by the exogenous parameter $\hat{\alpha}$ reflects the underlying flexibility of the evaluation mechanism proposed by Theorem 5. In particular, if we are to pursue a relatively high goal of economic maturity, then we can properly extend the given neighborhood; symmetrically, if the goal is relatively low, then we can proportionally narrow the neighborhood. Therefore, we are kept in a subtle balance between the *evaluation accuracy* and the *sustainable economic incentive*. As is broadly recognized, accuracy is important because it reveals useful information of the real macroeconomic process and meanwhile avoids any unnecessary overconfidence, while economic incentive is sustainable only when there are external encouragements from real accomplishments. In sum, policy makers should carefully sustain such a balance. It, therefore, can be regarded as an insightful lesson policy makers might have learned from our theoretical model.

4. Robustness

Before establishing the formal assertion, we first give the following definition.

Definition 2 (**Robust Turnpike**). For a turnpike of any given equilibrium path of capital accumulation, we call it a *robust turnpike* if any perturbed equilibrium path of capital accumulation asymptotically converges to it as the perturbation term (or vector) approaches zero (or a zero vector).

It follows from (7) that

$$dk(t) = \varphi(k(t))\,dt + \psi(k(t))\,dB(t) \equiv k(t)\varphi_0(k(t))\,dt + k(t)\psi_0(k(t))\,dB(t). \tag{11}$$

Now, we introduce the following SDE:

$$d\tilde{k}(t) = \tilde{\varphi}\left(\tilde{k}(t)\right)dt + \tilde{\psi}\left(\tilde{k}(t)\right)dB(t) \equiv \tilde{k}(t)\tilde{\varphi}_0\left(\tilde{k}(t)\right)dt + \tilde{k}(t)\tilde{\psi}_0\left(\tilde{k}(t)\right)dB(t), \quad (12)$$

where we have provided the following assumption.

Assumption 5. *For any $\xi > 0$,*

$$\sup_{k,\tilde{k}>0}\left|\varphi_0(k) - \tilde{\varphi}_0(\tilde{k})\right| \vee \sup_{k,\tilde{k}>0}\left|\psi_0(k) - \tilde{\psi}_0(\tilde{k})\right| \leq \xi.$$

That is to say, (12) defines the ξ- perturbation of (11).

Moreover, we need the following assumption.

Assumption 6. *There exist constants $\phi, \tilde{\phi}$ and $\phi_0 < \infty$ such that*

$$|\varphi(k)k| \vee |\psi(k)|^2 \leq \phi\,|k|^2, \quad \left|\tilde{\varphi}(\tilde{k})\tilde{k}\right| \vee \left|\tilde{\psi}(\tilde{k})\right|^2 \leq \tilde{\phi}\left|\tilde{k}\right|^2,$$

$$\textit{and} \quad \sup_{k>0}|\varphi_0(k))|^2 \vee \sup_{k>0}|\psi_0(k)|^2 \leq \phi_0 \quad \textit{for } \forall k > 0 \textit{ and } \forall \tilde{k} > 0.$$

Remark 7. One can easily find that Assumption 6 is truly reasonable thanks to Assumption 1. Assumption 6 is indeed without any loss of generality and is just for the sake of convenience in the following proofs.

Lemma 1. *Provided the above assumptions hold, there exist constants $e(k_0, p, T) < \infty$ and $\tilde{e}(k_0, p, T) < \infty$ such that*

(i) $\mathbb{E}\left[\sup_{0\leq t\leq T}|k(t)|^p\right] \leq e(k_0, p, T)$*, and*

(ii) $\mathbb{E}\left[\sup_{0\leq t\leq T}\left|\tilde{k}(t)\right|^p\right] \leq \tilde{e}(k_0, p, T)$

for $k(0) = \tilde{k}(0) = k_0 > 0, \forall T > 0, \forall p \in \mathbb{N}$ (the set of natural numbers), and $p \geq 2$.

Proof. See Appendix C.

Specifically, even if we do not rely on the above assumptions, we can still get the following result:

Lemma 2. *If both $k(t)$ and $\tilde{k}(t)$ are martingales w. r. t. $\mathbb{P}$, then there exist constants $\eta < \infty$ and $\tilde{\eta} < \infty$ such that*

(i) $\mathbb{E}\left[\lim_{T\to\infty}\sup_{0\leq t\leq T}|k(t)|^2\right] < \eta$*, and*

(ii) $\mathbb{E}\left[\lim_{T\to\infty}\sup_{0\leq t\leq T}\left|\tilde{k}(t)\right|^2\right] < \tilde{\eta}$

for $k(0) = \tilde{k}(0) = k_0 > 0$.

Proof. See Appendix D.

Now, we can derive the following proposition.

Proposition 1. *Provided the above assumptions hold and suppose that* $k(0) = \tilde{k}(0) = k_0 > 0$*, then we get*

$$\mathbb{E}\left[\lim_{T\to\infty} \sup_{0\le t\le T} \left|k(t) - \tilde{k}(t)\right|^2\right] \to 0$$

as $\xi \to 0$.

Proof. See Appendix E.

Remark 8. It should be pointed out that in the proof of Proposition 1 we have implicitly used the following assumptions, namely, the speed for ξ to approach zero is much faster than the speed for time T to approach infinity and also $0 \times \infty \equiv 0$. Moreover, we can get the same conclusion by taking the limit as $\xi \to 0$ first and then taking the limit along $T \to \infty$.

Accordingly, the following theorem is established.

Theorem 6 (Robust Turnpike). *Provided assumptions of Theorems 2 and 4 are fulfilled,* k^* *is a robust turnpike.*

Proof. To prove the robustness, one just needs to combine Theorem 2 with Proposition 1 (or combine Theorem 4 with Proposition 1) and also use the following fact:

$$\left|\tilde{k}(t) - k^*\right|^2 = \left|\tilde{k}(t) - k(t) + k(t) - k^*\right|^2 \le 2\left[\left|\tilde{k}(t) - k(t)\right|^2 + |k(t) - k^*|^2\right].$$

Thus, we leave the details to the interested reader.

Similarly, one can also arrive at the following result.

Theorem 7 (Robust Turnpike). *Provided assumptions of Theorems 3 and 5 are fulfilled,* $\hat{k}^*$ *is a robust turnpike.*

Remark 9. Theorems 6 and 7 have confirmed the asymptotic stability of turnpikes k^* and $\hat{k}^*$, respectively. To summarize, by noticing that our theorems show that the optimal path of capital accumulation will *robustly* converge to the corresponding turnpike *in the sense of uniform topology*, we argue that the current study indeed extends existing turnpike theorems (see Scheinkman 1976; McKenzie 1983; Yano 1998; Dai 2014c) to much stronger cases. This would be regarded as one contribution of the present paper.

5. Concluding remarks

In the current exploration, we are encouraged to study the economic maturity of a one-sector neoclassical model with stochastic growth. To the best of our knowledge, we, for the first time, provide a relatively complete characterization of the minimum time needed to economic maturity for any underdeveloped economy and further show that the corresponding capital-labor ratio exhibits both asymptotic turnpike property and neighborhood turnpike property under reasonable conditions. In other words, the optimal path of capital accumulation (or the equilibrium path of capital accumulation) will uniformly and robustly converge to the turnpike capital-labor ratio or will spend almost all the time staying in any given neighborhood of the turnpike capital-labor ratio under weak conditions and in a persistently non-stationary environment.

Noting that we assume very general forms of preference for the representative agent and production technology for the firm, one can apply the present framework to study different macroeconomic models with stochastic economic growth. Indeed, the present basic model can be naturally extended to other cases, including multi-sector models, heterogeneous-agent models or dynamic general equilibrium models (e.g., Bewley 1982; Yano 1984, and among others). Finally, as an interesting conjecture, the present framework can be extended to include multiple priors via applying the theory developed by Riedel (2009).

Acknowledgement Helpful comments from two anonymous referees are gratefully acknowledged.

References

Benaïm, M., Hofbauer, J. and Sandholm, W. H. (2008). Robust Permanence and Impermanence for the Stochastic Replicator Dynamic. *Journal of Biological Dynamics*, 2, 180–195.

Bewley, T. (1982). An Integration of Equilibrium Theory and Turnpike Theory. *Journal of Mathematical Economics*, 10, 233–267.

Bomze, I. M. (1991). Cross Entropy Minimization in Uninvadable States of Complex Populations. *Journal of Mathematical Biology*, 30, 73–87.

Cass, D. (1965). Optimum Growth in an Aggregative Model of Capital Accumulation. *Review of Economic Studies*, 32, 233–240.

Cass, D. (1966). Optimum Growth in an Aggregative Model of Capital Accumulation: A Turnpike Theorem. *Econometrica*, 34, 833–850.

Chakraborty, S. (2004). Endogenous Lifetime and Economic Growth. *Journal of Economic Theory*, 116, 119–137.

Chang, F.-R. and Malliaris, A. G. (1987). Asymptotic Growth under Uncertainty: Existence and Uniqueness. *Review of Economic Studies*, 54, 169–174.

Dai, D. (2012). Stochastic Versions of Turnpike Theorems in the Sense of Uniform Topology. *Annals of Economics and Finance*, 13, 389–431.

Dai, D. (2013). Cooperative Economic Growth. *Economic Modelling*, 33, 407–415.

Dai, D. (2014a). A Golden Formula in Neoclassical-Growth Models with Brownian-Motion Shocks. *Scottish Journal of Political Economy*, 61, 211–228.

Dai, D. (2014b). The Long-Run Behavior of Consumption and Wealth Dynamics in Complete Financial Market with Heterogeneous Investors. *Journal of Applied Mathematics*, 482314, 1–16.

Dai, D. (2014c). On the Turnpike Property of the Modified Golden Rule. *Czech Economic Review*, 8, 26–32.

de la Croix, D. and Ponthiere, G. (2010). On the Golden Rule of Capital Accumulation under Endogenous Longevity. *Mathematical Social Sciences*, 59, 227–238.

Higham, D. J., Mao, X. R. and Stuart, A. M. (2003). Strong Convergence of Euler-Type Methods for Nonlinear Stochastic Differential Equations. *SIAM Journal on Numerical Analysis*, 40, 1041–1063.

Imhof, L. A. (2005). The Long-Run Behavior of the Stochastic Replicator Dynamics. *Annals of Applied Probability*, 15, 1019–1045.

Jeanblanc, M., Lakner, P. and Kadam, A. (2004). Optimal Bankruptcy Time and Consumption/Investment Policies on an Infinite Horizon with a Continuous Debt Repayment until Bankruptcy. *Mathematics of Operations Research*, 29, 649–671.

Joshi, S. (1997). Turnpike Theorem in Nonconvex and Nonstationary Environment. *International Economic Review*, 38, 225–248.

Karatzas, I. and Shreve, S. E. (1991). *Brownian Motion and Stochastic Calculus*. Volume 113 of Graduate Texts in Mathematics, 2nd edition. New York, Springer-Verlag.

Karatzas, I. and Wang, H. (2001). Utility Maximization with Discretionary Stopping. *SIAM Journal on Control and Optimization*, 39, 306–329.

Kurz, M. (1965). Optimal Paths of Capital Accumulation under the Minimum Time Objective. *Econometrica*, 33, 42–66.

McKenzie, L. (1976). Turnpike Theory. *Econometrica*, 44, 841–865.

McKenzie, L. (1983). Turnpike Theory, Discounted Utility, and the von Neumann Facet. *Journal of Economic Theory*, 30, 30–52.

Merton, R. C. (1975). An Asymptotic Theory of Growth under Uncertainty. *Review of Economic Studies*, 42, 375–393.

Mirman, L. J. (1972). On the Existence of Steady State Measures for One Sector Growth Models with Uncertain Technology. *International Economic Review*, 13, 271–286.

Øksendal, B. and Sulem, A. (2005). *Applied Stochastic Control of Jump Diffusions*. Berlin, Springer-Verlag.

Radner, R. (1961). Paths of Economic Growth that are Optimal with Regard only to Final States: A Turnpike Theorem. *Review of Economic Studies*, 28, 98–104.

Ramsey, F. P. (1928). A Mathematical Theory of Saving. *Economic Journal*, 38, 543–559.

Riedel, F. (2009). Optimal Stopping with Multiple Priors. *Econometrica*, 77, 857–908.

Samuelson, P. A. (1965). A Catenary Turnpike Theorem Involving Consumption and the Golden Rule. *American Economic Review*, 55, 486–496.

Scheinkman, J. (1976). On Optimal Steady States of n-Sector Growth Models when Utility is Discounted. *Journal of Economic Theory*, 12, 11–20.

Schreiber, S. J., Benaïm, M. and Atchadé, K. A. S. (2011). Persistence in Fluctuating Environments. *Journal of Mathematical Biology*, 62, 655–683.

Solow, R. M. (1956). A Contribution to the Theory of Economic Growth. *Quarterly Journal of Economics*, 70, 65–94.

Song, Z., Storesletten, K. and Zilibotti, F. (2011). Growing Like China. *American Economic Review*, 101, 202–241.

Yano, M. (1984). The Turnpike of Dynamic General Equilibrium Paths and Its Insensitivity to Initial Conditions. *Journal of Mathematical Economics*, 13, 235–254.

Yano, M. (1998). On the Dual Stability of a von Neumann Facet and the Inefficacy of Temporary Fiscal Policy. *Econometrica*, 66, 427–451.

Appendix

A. Proof of Theorem 2

Put $\varphi(k(t)) = 0$ in (7), then we find that $k(t)$ will be a martingale w. r. t. $\mathbb{P}$. Thus, by using Doob's Martingale Inequality, we obtain

$$\mathbb{P}\left(\sup_{0\le t\le T}|k(t)| \ge \lambda\right) \le \frac{1}{\lambda}\mathbb{E}\left[|k(T)|\right] = \frac{k_0}{\lambda} \text{ for } \forall\lambda > 0,\ \forall T > 0. \tag{A1}$$

Without loss of generality, we put $\lambda = 2^m$ for $m \in \mathbb{N}$, then

$$\mathbb{P}\left(\sup_{0\le t\le T}|k(t)| \ge 2^m\right) \le \frac{1}{2^m}k_0 \text{ for } \forall m \in \mathbb{N} \text{ and } \forall T > 0.$$

Using the well-known Borel-Cantelli Lemma, we arrive at

$$\mathbb{P}\left(\sup_{0\le t\le T}|k(t)| \ge 2^m i.m.m\right) = 0 \text{ for } \forall T > 0,$$

in which $i.m.m$ represents "infinitely many m." So, for a.a. (almost all) $\omega \in \Omega$, there exists $\overline{m}(\omega) \in \mathbb{N}$ such that

$$\sup_{0\le t\le T}|k(t)| < 2^m \text{ a.s. (almost surely) for } m \ge \overline{m}(\omega) \text{ and } \forall T > 0,$$

hence,

$$\lim_{T\to\infty}\sup_{0\le t\le T}|k(t)| \le 2^m \text{ a.s. for } m \ge \overline{m}(\omega).$$

Consequently, $k(t) = k(t,\omega)$ is uniformly bounded for $t \in [0,T]$, $\forall T > 0$ and a.a. $\omega \in \Omega$. Thus, it is ensured that $k(t) = k(t,\omega)$ converges a.s.-$\mathbb{P}$ and the limit belongs to space $L^1(\mathbb{P})$ thanks to Doob's Martingale Convergence Theorem. Moreover, by applying Kolmogorov's (or Chebyshev's) Inequality, we get

$$\mathbb{P}\left(\sup_{0\le t\le T}|k(t)|\ge\lambda\right)\le\frac{1}{\lambda^2}var\left[|k(T)|\right] \text{ for } \forall 0<\lambda<\infty \text{ and } \forall T>0.$$

It follows from (A1) that

$$\frac{1}{\lambda^2}var\left[|k(T)|\right]\le\frac{k_0}{\lambda}\ \Leftrightarrow\ var\left[|k(T)|\right]\le\lambda k_0 \text{ for } \forall T>0. \tag{A2}$$

Noting that $var\left[|k(T)|\right]=\mathbb{E}\left[|k(T)|^2\right]-(k_0)^2$ for $\forall T>0$, we get by (A2)

$$\mathbb{E}\left[|k(T)|^2\right]\le(\lambda+k_0)k_0<\infty \text{ for } \forall 0<\lambda<\infty \text{ and } \forall T>0,$$

which yields

$$\sup_{T\ge 0}\mathbb{E}\left[|k(T)|^2\right]\le(\lambda+k_0)k_0<\infty.$$

Hence, by applying Doob's Martingale Convergence Theorem again, $k(t)=k(t,\omega)$ converges in $L^1(\mathbb{P})$.

Furthermore, it is easily seen that $k(t)-k^*$ is also a martingale w.r.t. $\mathbb{P}$. Thus, applying the Doob's Martingale Inequality again implies that

$$\mathbb{P}\left(\sup_{0\le t\le T}|k(t)-k^*|\ge\varepsilon\right)\le\frac{1}{\varepsilon}\mathbb{E}\left[|k(T)-k^*|\right] \text{ for } \forall\varepsilon>0 \text{ and } \forall T>0. \tag{A3}$$

Provided that $\tau^*(\omega)\equiv\inf\{t\ge 0;k(t)=k^*\}<\infty$ a.s.-$\mathbb{P}$ given by (8), we see that there exists $\beta>0$ such that the martingale inequality in (A3) still holds for $\forall\tau(\omega)\in B_\beta(\tau^*(\omega))\equiv\{\tau(\omega)\in\mathcal{T};|\tau(\omega)-\tau^*(\omega)|\le\beta\}$ by using Doob's Optional Sampling Theorem. Then, we get that $k(\tau)-k^*$ is uniformly bounded on the compact set $B_\beta(\tau^*(\omega))$ by applying Heine-Borel Theorem and Weierstrass Theorem. Therefore, we, without any loss of generality, set up $\beta=2^{-m}$ for $\forall m\in\mathbb{N}$. Employing the continuity of martingale w.r.t. time t for any given $\omega\in\Omega$ and $\forall\tau_m\in B_\beta(\tau^*(\omega))\equiv B_{2^{-m}}(\tau^*(\omega))$ and using the Lebesgue Dominated Convergence Theorem, we get

$$\limsup_{m\to\infty}\mathbb{P}\left(\sup_{0\le t\le\tau_m}|k(t)-k^*|\ge\varepsilon\right)\le\frac{1}{\varepsilon}\limsup_{m\to\infty}\mathbb{E}\left[|k(\tau_m)-k^*|\right]=0$$

almost surely. And this implies that

$$\limsup_{m\to\infty}\mathbb{P}\left(\sup_{0\le t\le\tau_m}|k(t)-k^*|<\varepsilon\right)\ge 1 \text{ a.s.-}\mathbb{P}.$$

Letting $\varepsilon=2^{-m_0}$ for $\forall m_0\in\mathbb{N}$, we get

$$\limsup_{m\to\infty}\mathbb{P}\left(\sup_{0\le t\le\tau_m}|k(t)-k^*|<2^{-m_0}\right)=1 \text{ a.s.-}\mathbb{P} \text{ for } \forall m_0\in\mathbb{N}.$$

It follows from Fatou's Lemma that

$$\mathbb{P}\left(\sup_{0\leq t\leq\tau^*(\omega)}|k(t)-k^*|<2^{-m_0}\right)=1 \text{ a.s.-}\mathbb{P} \text{ for } \forall m_0\in\mathbb{N}.$$

Then, applying Borel-Cantelli Lemma again implies that

$$\mathbb{P}\left(\sup_{0\leq t\leq\tau^*(\omega)}|k(t)-k^*|<2^{-m_0} i.m.m_0\right)=1,$$

where $i.m.m_0$ stands for "infinitely many m_0." So for a.a. $\omega\in\Omega$, there exists $\overline{m}_0(\omega)\in\mathbb{N}$ such that

$$\sup_{0\leq t\leq\tau^*(\omega)}|k(t)-k^*|<2^{-m_0} \text{ a.s. for } \forall m_0\geq\overline{m}_0(\omega).$$

That is,

$$\limsup_{m_0\to\infty}\sup_{0\leq t\leq\tau^*(\omega)}|k(t)-k^*|\leq 0 \text{ a.s.-}\mathbb{P},$$

which yields

$$\limsup_{\tau^*(\omega)\to\infty}\sup_{0\leq t\leq\tau^*(\omega)}|k(t)-k^*|\leq 0 \text{ a.s.-}\mathbb{P}.$$

That is to say,

$$\mathbb{P}\left(\bigcup_{m=1}^{\infty}\bigcap_{t'=0}^{\infty}\bigcup_{t=t'}^{\infty}\left[|k(t)-k^*|\geq\frac{1}{m}\right]\right)=0.$$

Equivalently, for $\forall m\in\mathbb{N}$, we arrive at

$$\mathbb{P}\left(\bigcap_{t'=0}^{\infty}\bigcup_{t=t'}^{\infty}\left[|k(t)-k^*|\geq\frac{1}{m}\right]\right)=0,$$

i.e., for $\forall\varepsilon>0$, we have

$$\lim_{t'\to\infty}\mathbb{P}\left(\bigcup_{t=t'}^{\infty}\left[|k(t)-k^*|\geq\varepsilon\right]\right)=0,$$

which gives the desired assertion. □

B. Proof of Theorem 4

Given the SDE defined by (7), we can define the following characteristic operator of $k(t)$:

$$\mathcal{A}g(k_0)=\varphi(k_0)\frac{\partial g}{\partial k_0}(k_0)+\frac{1}{2}\psi^2(k_0)\frac{\partial^2 g}{\partial(k_0)^2}(k_0)$$

for any $k_0 \equiv k(0) > 0$. We now define Kullback-Leibler type distance (see Bomze 1991; Imhof 2005) between k_0 and k^* as follows:

$$g(k_0) \equiv dist(k_0, k^*) \equiv k^* \log\left(\frac{k^*}{k_0}\right) \geq 0.$$

Then we get

$$\mathcal{A}g(k_0) = \left[-\varphi(k_0) + \frac{1}{2k_0}\psi^2(k_0)\right]\frac{k^*}{k_0}.$$

By Theorem 2, we find that there exists $T_0 < \infty$ such that

$$\sup_{0\leq t\leq T}|k(t)-k^*| < \mu \quad \text{for } \forall\mu > 0 \text{ and } \forall T \geq T_0.$$

Thus, we have

$$\mathcal{A}g(k_0) \leq \left[-\varphi(k_0) + \frac{1}{2k_0}\psi^2(k_0)\right]\frac{k^*}{k_0} + \mu - |k(t)-k^*| \equiv \Sigma - |k(t)-k^*|. \tag{B1}$$

Define some new notations:

$$B_\alpha(k^*) \equiv \{k(t) \in \Re_{++}; |k(t)-k^*| < \alpha, t \geq 0\},$$

$$\tilde{\tau}(\omega) \equiv \tau_{\overline{B}_\alpha(k^*)}(\omega) \equiv \inf\left\{t \geq 0; k(t) \in \overline{B}_\alpha(k^*) \equiv clB_\alpha(k^*)\right\},$$

where $\overline{B}_\alpha(k^*)$ denotes the closure of $B_\alpha(k^*)$. Suppose $\alpha > \Sigma$ for every $k(t) \notin \overline{B}_\alpha(k^*)$, i.e., $k(t) \in \overline{B}_\alpha^C(k^*)$, we then get

$$\mathcal{A}g(k_0) \leq -\alpha + \Sigma$$

by using (B1). Thus, by making use of Dynkin's formula,

$$0 \leq \mathbb{E}[g(k(t\wedge\tilde{\tau}))] = g(k_0) + \mathbb{E}\left[\int_0^{t\wedge\tilde{\tau}} \mathcal{A}g(k(s))\,ds\right] \leq g(k_0) + (\Sigma-\alpha)\mathbb{E}[t\wedge\tilde{\tau}(\omega)].$$

Since $t\wedge\tilde{\tau} \uparrow \tilde{\tau}$ as $t \to \infty$, applying Lebesgue Monotone Convergence Theorem results in

$$0 \leq g(k_0) + (\Sigma-\alpha)\mathbb{E}[\tilde{\tau}(\omega)],$$

which produces

$$\mathbb{E}\left[\tau_{\overline{B}_\alpha(k^*)}(\omega)\right] = \mathbb{E}[\tilde{\tau}(\omega)] \leq \frac{g(k_0)}{\alpha-\Sigma} = \frac{dist(k_0,k^*)}{\alpha-\Sigma},$$

as required in (i). Moreover, for some constant $W > g(k_0)$, set up

$$\tau_W = \tau_W(\omega) \equiv \inf\{t \geq 0; g(k(t)) = W\}.$$

Thus, by making use of Dynkin's formula and inequality (B1),

$$\begin{aligned} 0 \leq \mathbb{E}\left[g\left(k\left(t \wedge \tau_W\right)\right)\right] &= g\left(k_0\right)+\mathbb{E}\left[\int_0^{t \wedge \tau_W} \mathcal{A} g(k(s))\, d s\right] \\ &\leq g\left(k_0\right)-\mathbb{E}\left[\int_0^{t \wedge \tau_W}\left|k(s)-k^*\right| d s\right]+\Sigma \mathbb{E}\left[t \wedge \tau_W(\omega)\right] . \end{aligned}$$

If $W \rightarrow \infty$, we get $t \wedge \tau_W(\omega) \rightarrow t$. By applying Lebesgue Bounded Convergence Theorem and Levi Lemma, we are led to

$$0 \leq g\left(k_0\right)-\mathbb{E}\left[\int_0^t\left|k(s)-k^*\right| d s\right]+\Sigma t,$$

which yields

$$\mathbb{E}\left[\frac{1}{t} \int_0^t\left|k(s)-k^*\right| d s\right] \leq \frac{g\left(k_0\right)}{t}+\Sigma.$$

Thus, we have

$$\limsup_{t \rightarrow \infty} \mathbb{E}\left[\frac{1}{t} \int_0^t\left|k(s)-k^*\right| d s\right] \leq \Sigma. \tag{B2}$$

If we let $\chi_{\overline{B}_\alpha^C(k^*)}(k(t))$ denote the indicator function of set $\overline{B}_\alpha^C(k^*)$, then by (B2) and Assumption 4, we arrive at

$$\begin{aligned} \pi\left[\overline{B}_\alpha^C(k^*)\right] &= \limsup_{t \rightarrow \infty} \mathbb{E}\left[\frac{1}{t} \int_0^t \chi_{\overline{B}_\alpha^C(k^*)}(k(s))\, d s\right] \\ &\leq \limsup_{t \rightarrow \infty} \mathbb{E}\left[\frac{1}{t} \int_0^t \frac{\left|k(s)-k^*\right|}{\alpha} d s\right] \leq \frac{\Sigma}{\alpha}, \end{aligned}$$

which implies that

$$\pi\left[\overline{B}_\alpha(k^*)\right] \geq 1-\frac{\Sigma}{\alpha} \equiv 1-\varepsilon,$$

which gives the desired assertion in (ii). □

C. Proof of Lemma 1

Applying Itô's rule to (11) produces

$$|k(t)|^2 = |k_0|^2 + 2\int_0^t \varphi(k(s))\,k(s)ds + \int_0^t |\psi(k(s))|^2 ds + 2\int_0^t \psi(k(s))\,k(s)dB(s).$$

By using Assumption 6 we get that for $t_1 \in [0,T]$ and some constant $e \equiv e(p,T) < \infty$ (which may be different from line to line throughout this proof),

$$\sup_{0\le t\le t_1} |k(t)|^p \le e\left\{ |k_0|^p + \left[\int_0^{t_1} \phi\,|k(s)|^2 ds\right]^{\frac{p}{2}} + \sup_{0\le t\le t_1}\left|\int_0^t k(s)\psi[k(s)]\,dB(s)\right|^{\frac{p}{2}}\right\}.$$

It follows from Cauchy-Schwarz Inequality (Dai 2014b) that

$$\sup_{0\le t\le t_1} |k(t)|^p \le e\left\{ |k_0|^p + \int_0^{t_1} |k(s)|^p ds + \sup_{0\le t\le t_1}\left|\int_0^t k(s)\psi[k(s)]\,dB(s)\right|^{\frac{p}{2}}\right\}.$$

Taking expectations on both sides and applying the Burkholder-Davis-Gundy Inequality (see Karatzas and Shreve 1991, p. 166) show that

$$\mathbb{E}\left[\sup_{0\le t\le t_1} |k(t)|^p\right] \le e\left\{ |k_0|^p + \int_0^{t_1} \mathbb{E}\left[|k(s)|^p\right] ds + \mathbb{E}\left[\int_0^{t_1} |k(s)|^2\,|\psi(k(s))|^2 ds\right]^{\frac{p}{4}}\right\}. \tag{C1}$$

Now, using the Young Inequality (see Higham et al. 2003), Assumption 6, and Rogers-Hölder Inequality (Dai 2014b) reveals that

$$\begin{aligned}
\mathbb{E}\left[\int_0^{t_1} |k(s)|^2\,|\psi(k(s))|^2 ds\right]^{\frac{p}{4}} &\le \mathbb{E}\left[\sup_{0\le t\le t_1} |k(t)|^{\frac{p}{2}} \left(\int_0^{t_1} |\psi(k(s))|^2 ds\right)^{\frac{p}{4}}\right] \\
&\le \frac{1}{2e}\mathbb{E}\left[\sup_{0\le t\le t_1} |k(t)|^p\right] + \frac{e}{2}\mathbb{E}\left[\int_0^{t_1} |\psi(k(s))|^2 ds\right]^{\frac{p}{2}} \\
&\le \frac{1}{2e}\mathbb{E}\left[\sup_{0\le t\le t_1} |k(t)|^p\right] + \frac{e}{2}\phi^{\frac{p}{2}}\mathbb{E}\left[\int_0^{t_1} |k(s)|^2 ds\right]^{\frac{p}{2}}
\end{aligned}$$

$$\leq \frac{1}{2e}\mathbb{E}\left[\sup_{0\leq t\leq t_1}|k(t)|^p\right] + \frac{e}{2}\phi^{\frac{p}{2}}T^{\frac{p-2}{2}}\mathbb{E}\left[\int_0^{t_1}|k(s)|^p\,ds\right].$$

Substituting this into (C1) yields

$$\mathbb{E}\left[\sup_{0\leq t\leq T}|k(t)|^p\right] \leq e\left\{|k_0|^p + \int_0^T \mathbb{E}\left[|k(t)|^p\right]dt\right\}.$$

Thus, by applying the following fact (see Higham et al. 2003):

$$\mathbb{E}\left[|k(t)|^p\right] \leq e\left(1+|k_0|^p\right),$$

we arrive at

$$\mathbb{E}\left[\sup_{0\leq t\leq T}|k(t)|^p\right] \leq e(k_0,p,T) < \infty,$$

which gives the desired result in (i). Noting that the proof of (ii) is quite similar to that of (i), we omit it. And this completes the whole proof. □

D. Proof of Lemma 2

By using Doob's Martingale Inequality, we obtain

$$\mathbb{P}\left(\sup_{0\leq t\leq T}|k(t)|\geq\lambda\right) \leq \frac{1}{\lambda}\mathbb{E}\left[|k(T)|\right] = \frac{k_0}{\lambda} \text{ for } \forall 0<\lambda<\infty \text{ and } \forall T>0. \tag{D1}$$

Similarly, by applying Kolmogorov's (or Chebyshev's) Inequality, we get

$$\mathbb{P}\left(\sup_{0\leq t\leq T}|k(t)|\geq\lambda\right) \leq \frac{1}{\lambda^2}var\left[|k(T)|\right] \text{ for } \forall 0<\lambda<\infty,\ \forall T>0.$$

It follows from (D1) that

$$\frac{1}{\lambda^2}var\left[|k(T)|\right] \leq \frac{k_0}{\lambda} \;\Leftrightarrow\; var\left[|k(T)|\right] \leq \lambda k_0 \text{ for } \forall T>0. \tag{D2}$$

Noting that $var\left[|k(T)|\right] = \mathbb{E}\left[|k(T)|^2\right] - (k_0)^2$ for $\forall T>0$, we get by (D2)

$$\mathbb{E}\left[|k(T)|^2\right] \leq (\lambda+k_0)k_0 < \infty \text{ for } \forall 0<\lambda<\infty \text{ and } \forall T>0, \tag{D3}$$

which implies that $k(t)$ is a square-integrable martingale. We define:

$$\zeta \equiv |k(t)|, \quad \zeta^* \equiv \|k(t)\|_\infty \equiv \sup_{0\leq s\leq t}|k(s)| \text{ and } \|k(t)\|_2 \equiv \left\{\mathbb{E}\left[|k(t)|^2\right]\right\}^{\frac{1}{2}}.$$

Thus, by applying Doob's Martingale Inequality and Fubini Theorem, we arrive at the following result for some constant $N < \infty$:

$$\begin{aligned}
\mathbb{E}\left[|\zeta^* \wedge N|^2\right] &= 2\int_0^\infty x\mathbb{P}\left(\zeta^*(\omega)\wedge N \geq x\right)dx \\
&\leq 2\int_0^\infty \int_{\{\zeta^*(\omega)\wedge N\geq x\}} \zeta(\omega)d\mathbb{P}(\omega)dx \\
&= 2\int_0^\infty \int_\Omega \zeta(\omega)\chi_{\{\zeta^*(\omega)\wedge N\geq x\}}d\mathbb{P}(\omega)dx \\
&= 2\int_\Omega \zeta(\omega)\int_0^{\zeta^*(\omega)\wedge N} dxd\mathbb{P}(\omega) \\
&= 2\int_\Omega \zeta(\omega)\left(\zeta^*(\omega)\wedge N\right)d\mathbb{P}(\omega) \\
&= 2\mathbb{E}\left[\zeta\left(\zeta^*\wedge N\right)\right].
\end{aligned}$$

It follows from Rogers-Hölder Inequality that

$$\|\zeta^* \wedge N\|_2^2 = \mathbb{E}\left[|\zeta^* \wedge N|^2\right] \leq 2\|\zeta\|_2 \|\zeta^* \wedge N\|_2,$$

which produces

$$\|\zeta^* \wedge N\|_2 \leq 2\|\zeta\|_2.$$

Noting that $\mathbb{E}\left[|\zeta^* \wedge N|^2\right] \leq N^2 < \infty$, and hence applying Lebesgue Dominated Convergence Theorem leads us to

$$\|\zeta^*\|_2 \leq 2\|\zeta\|_2 \quad \Leftrightarrow \quad \|\zeta^*\|_2^2 \leq 4\|\zeta\|_2^2,$$

namely,

$$\mathbb{E}\left[\sup_{0\leq s\leq t}|k(s)|^2\right] \leq 4\mathbb{E}\left[|k(t)|^2\right] \leq 4(\lambda + k_0)k_0 < \infty \;\text{ for }\; \forall t \geq 0$$

by using the inequality given by (D3). Accordingly, a canonical application of Lebesgue Monotone Convergence Theorem (or Levi Lemma) gives the required assertion in (i). The proof of (ii) is similar to that of (i), we hence omit it. Therefore, the whole proof is complete. □

E. Proof of Proposition 1

Provided the SDEs defined in (11) and (12), it follows from Lemma 1 that for $\forall 2 \leq p < \infty$ and $\forall T > 0$ there exists some constant $W < \infty$ such that

$$\mathbb{E}\left[\sup_{0\leq t\leq T}|k(t)|^p\right] \vee \mathbb{E}\left[\sup_{0\leq t\leq T}\left|\tilde{k}(t)\right|^p\right] \leq W, \tag{E1}$$

where, by using Assumption 1, we can have

$$k(t) = k_0 + \int_0^t k(s)\varphi_0(k(s))\,ds + \int_0^t k(s)\psi_0(k(s))\,dB(s),$$

$$\tilde{k}(t) = k_0 + \int_0^t \tilde{k}(s)\tilde{\varphi}_0\left(\tilde{k}(s)\right)ds + \int_0^t \tilde{k}(s)\tilde{\psi}_0\left(\tilde{k}(s)\right)dB(s).$$

Moreover, we put $|k(t)| \vee \left|\tilde{k}(t)\right| \le \overline{W} < \infty$ for $\forall t \ge 0$; otherwise, we just consider $k(t) \wedge \overline{W}$ and $\tilde{k}(t) \wedge \overline{W}$ instead of $k(t)$ and $\tilde{k}(t)$, respectively, to get the desired result by sending $\overline{W}$ to infinity and using Lebesgue Dominated Convergence Theorem. In what follows, we first define the following stopping times:

$$\tau_{\overline{W}} \equiv \inf\left\{t \ge 0; |k(t)| \ge \overline{W}\right\}, \quad \tilde{\tau}_{\overline{W}} \equiv \inf\left\{t \ge 0; \left|\tilde{k}(t)\right| \ge \overline{W}\right\}, \quad \tau^*_{\overline{W}} \equiv \tau_{\overline{W}} \wedge \tilde{\tau}_{\overline{W}}.$$

By using the Young Inequality (see Higham et al. 2003) and for any $R > 0$,

$$\begin{aligned}
&\mathbb{E}\left[\sup_{0\le t\le T}\left|k(t)-\tilde{k}(t)\right|^2\right] \\
&= \mathbb{E}\left[\sup_{0\le t\le T}\left|k(t)-\tilde{k}(t)\right|^2 \chi_{\left\{\tau_{\overline{W}}>T,\tilde{\tau}_{\overline{W}}>T\right\}}\right] \\
&\quad +\mathbb{E}\left[\sup_{0\le t\le T}\left|k(t)-\tilde{k}(t)\right|^2 \chi_{\left\{\tau_{\overline{W}}\le T, or \tilde{\tau}_{\overline{W}}\le T\right\}}\right] \\
&\le \mathbb{E}\left[\sup_{0\le t\le T}\left|k\left(t\wedge\tau^*_{\overline{W}}\right)-\tilde{k}\left(t\wedge\tau^*_{\overline{W}}\right)\right|^2 \chi_{\left\{\tau^*_{\overline{W}}>T\right\}}\right] \\
&\quad +\frac{2R}{p}\mathbb{E}\left[\sup_{0\le t\le T}\left|k(t)-\tilde{k}(t)\right|^p\right] + \frac{1-\frac{2}{p}}{R^{\frac{2}{p-2}}}\mathbb{P}\left(\tau_{\overline{W}}\le T, or \tilde{\tau}_{\overline{W}}\le T\right).
\end{aligned} \tag{E2}$$

It follows from (E1) that

$$\mathbb{P}\left(\tau_{\overline{W}}\le T\right) = \mathbb{E}\left[\chi_{\left\{\tau_{\overline{W}}\le T\right\}}\frac{\left|k\left(\tau_{\overline{W}}\right)\right|^p}{\overline{W}^p}\right] \le \frac{1}{\overline{W}^p}\mathbb{E}\left[\sup_{0\le t\le T}|k(t)|^p\right] \le \frac{W}{\overline{W}^p}.$$

Similarly, one can get $\mathbb{P}\left(\tilde{\tau}_{\overline{W}}\le T\right) \le W/\overline{W}^p$. So,

$$\mathbb{P}\left(\tau_{\overline{W}}\le T, or \tilde{\tau}_{\overline{W}}\le T\right) \le \mathbb{P}\left(\tau_{\overline{W}}\le T\right) + \mathbb{P}\left(\tilde{\tau}_{\overline{W}}\le T\right) \le \frac{2W}{\overline{W}^p}.$$

Moreover, we obtain by (E1)

$$\mathbb{E}\left[\sup_{0\le t\le T}\left|k(t)-\tilde{k}(t)\right|^p\right] \le 2^{p-1}\mathbb{E}\left[\sup_{0\le t\le T}\left(|k(t)|^p + \left|\tilde{k}(t)\right|^p\right)\right] \le 2^p W.$$

Hence, (E2) becomes

$$\begin{aligned}
&\mathbb{E}\left[\sup_{0\le t\le T}\left|k(t)-\tilde{k}(t)\right|^2\right] \\
&\quad\le \mathbb{E}\left[\sup_{0\le t\le T}\left|k\left(t\wedge\tau^*_{\overline{W}}\right)-\tilde{k}\left(t\wedge\tau^*_{\overline{W}}\right)\right|^2\right] + \frac{2^{p+1}RW}{p} + \frac{2(p-2)W}{pR^{\frac{2}{p-2}}\overline{W}^p}.
\end{aligned} \tag{E3}$$

By making use of Cauchy-Bunyakovsky-Schwarz Inequality, we get

$$\begin{aligned}
\left|k\left(t\wedge\tau_W^*\right)-\tilde{k}\left(t\wedge\tau_W^*\right)\right|^2 &= \left|\int_0^{t\wedge\tau_W^*}\left[k(s)\varphi_0\left(k(s)\right)-\tilde{k}(s)\tilde{\varphi}_0\left(\tilde{k}(s)\right)\right]ds\right.\\
&\quad\left.+\int_0^{t\wedge\tau_W^*}\left[k(s)\psi_0\left(k(s)\right)-\tilde{k}(s)\tilde{\psi}_0\left(\tilde{k}(s)\right)\right]dB(s)\right|^2\\
&\le 2\left\{T\int_0^{t\wedge\tau_W^*}\left[\left|k(s)\varphi_0\left(k(s)\right)-\tilde{k}(s)\tilde{\varphi}_0\left(\tilde{k}(s)\right)\right|^2\right]ds\right.\\
&\quad\left.+\left|\int_0^{t\wedge\tau_W^*}\left[k(s)\psi_0\left(k(s)\right)-\tilde{k}(s)\tilde{\psi}_0\left(\tilde{k}(s)\right)\right]dB(s)\right|^2\right\}\\
&\le 4\left\{T\int_0^{t\wedge\tau_W^*}\left[\left|k(s)\varphi_0\left(k(s)\right)-\tilde{k}(s)\varphi_0\left(k(s)\right)\right|^2\right]ds\right.\\
&\quad+T\int_0^{t\wedge\tau_W^*}\left[\left|\tilde{k}(s)\varphi_0\left(k(s)\right)-\tilde{k}(s)\tilde{\varphi}_0\left(\tilde{k}(s)\right)\right|^2\right]ds\\
&\quad\left.+\left|\int_0^{t\wedge\tau_W^*}\left[k(s)\psi_0\left(k(s)\right)-\tilde{k}(s)\tilde{\psi}_0\left(\tilde{k}(s)\right)\right]dB(s)\right|^2\right\}.
\end{aligned}$$

Taking expectations on both sides and using Itô's Isometry (Dai, 2014b), we have for $\forall\tau\le T$:

$$\begin{aligned}
&\mathbb{E}\left[\sup_{0\le t\le\tau}\left|k\left(t\wedge\tau_W^*\right)-\tilde{k}\left(t\wedge\tau_W^*\right)\right|^2\right]\\
&\le 4\left\{T\mathbb{E}\left[\int_0^{t\wedge\tau_W^*}\left|k(s)-\tilde{k}(s)\right|^2\left|\varphi_0\left(k(s)\right)\right|^2ds\right]\right.\\
&\quad+T\mathbb{E}\left[\int_0^{t\wedge\tau_W^*}\left|\tilde{k}(s)\right|^2\left|\varphi_0\left(k(s)\right)-\tilde{\varphi}_0\left(\tilde{k}(s)\right)\right|^2ds\right]\\
&\quad\left.+\mathbb{E}\left[\int_0^{t\wedge\tau_W^*}\left|k(s)\psi_0\left(k(s)\right)-\tilde{k}(s)\tilde{\psi}_0\left(\tilde{k}(s)\right)\right|^2ds\right]\right\}\\
&\le 8\left\{T\phi_0\mathbb{E}\left[\int_0^{t\wedge\tau_W^*}\left|k(s)-\tilde{k}(s)\right|^2ds\right]+T\xi^2\mathbb{E}\left[\int_0^{t\wedge\tau_W^*}\left|\tilde{k}(s)\right|^2ds\right]\right.\\
&\quad+\mathbb{E}\left[\int_0^{t\wedge\tau_W^*}\left|k(s)\psi_0\left(k(s)\right)-\tilde{k}(s)\psi_0\left(k(s)\right)\right|^2ds\right]\\
&\quad\left.+\mathbb{E}\left[\int_0^{t\wedge\tau_W^*}\left|\tilde{k}(s)\psi_0\left(k(s)\right)-\tilde{k}(s)\tilde{\psi}_0\left(\tilde{k}(s)\right)\right|^2ds\right]\right\}\\
&\le 8\left\{T\phi_0\mathbb{E}\left[\int_0^{t\wedge\tau_W^*}\left|k(s)-\tilde{k}(s)\right|^2ds\right]+T\xi^2\mathbb{E}\left[\int_0^{t\wedge\tau_W^*}\left|\tilde{k}(s)\right|^2ds\right]\right.\\
&\quad\left.+\phi_0\mathbb{E}\left[\int_0^{t\wedge\tau_W^*}\left|k(s)-\tilde{k}(s)\right|^2ds\right]+\xi^2\mathbb{E}\left[\int_0^{t\wedge\tau_W^*}\left|\tilde{k}(s)\right|^2ds\right]\right\}
\end{aligned}$$

$$
\begin{aligned}
&= 8\left\{(T+1)\phi_0\mathbb{E}\left[\int_0^{t\wedge\tau^*_{\overline{W}}}\left|k(s)-\tilde{k}(s)\right|^2 ds\right]+(T+1)\xi^2\mathbb{E}\left[\int_0^{t\wedge\tau^*_{\overline{W}}}\left|\tilde{k}(s)\right|^2 ds\right]\right\} \\
&\leq 8\left\{(T+1)\phi_0\int_0^T \mathbb{E}\left[\sup_{0\leq t_0\leq s}\left|k\left(t_0\wedge\tau^*_{\overline{W}}\right)-\tilde{k}\left(t_0\wedge\tau^*_{\overline{W}}\right)\right|^2\right]ds+T(T+1)\overline{W}^2\xi^2\right\},
\end{aligned}
$$

where we have used Assumptions 5 and 6. Hence, applying Gronwall's Inequality (see Higham et al. 2003; Dai 2014b) gives rise to

$$
\mathbb{E}\left[\sup_{0\leq t\leq\tau}\left|k\left(t\wedge\tau^*_{\overline{W}}\right)-\tilde{k}\left(t\wedge\tau^*_{\overline{W}}\right)\right|^2\right]\leq 8T(T+1)\overline{W}^2\exp\left[8(T+1)\phi_0\right]\xi^2.
$$

Inserting this into (E3) leads us to

$$
\mathbb{E}\left[\sup_{0\leq t\leq T}\left|k(t)-\tilde{k}(t)\right|^2\right]\leq 8T(T+1)\overline{W}^2\exp\left[8(T+1)\phi_0\right]\xi^2+\frac{2^{p+1}RW}{p}+\frac{2(p-2)W}{pR^{\frac{2}{p-2}}\overline{W}^p}.
$$

Hence, for $\forall\varepsilon>0$, we can choose R and $\overline{W}$ such that

$$
\frac{2^{p+1}RW}{p}\leq\frac{\varepsilon}{3}\quad\text{and}\quad\frac{2(p-2)W}{pR^{\frac{2}{p-2}}\overline{W}^p}\leq\frac{\varepsilon}{3}.
$$

And for any given $T>0$, we put ξ such that

$$
8T(T+1)\overline{W}^2\exp\left[8(T+1)\phi_0\right]\xi^2\leq\frac{\varepsilon}{3}.
$$

In consequence, for $\forall\varepsilon>0$, we obtain

$$
\mathbb{E}\left[\sup_{0\leq t\leq T}\left|k(t)-\tilde{k}(t)\right|^2\right]\leq\frac{\varepsilon}{3}+\frac{\varepsilon}{3}+\frac{\varepsilon}{3}=\varepsilon.
$$

Notice the arbitrariness of ε, we can employ Levi Lemma to produce the desired result. This proof is accordingly complete. □

Czech Economic Review 9 (2015) 74–90
Acta Universitatis Carolinae Oeconomica

Asymmetric Information, Bank Lending and Implicit Contracts: Differences between Banks

Juha-Pekka Niinimäki*

Received 5 August 2015; Accepted 2 October 2015

Abstract This paper studies asymmetric information on banks, relationship lending and switching costs. According to the classic theory of relationship banking asymmetric information on borrower types causes an informational lock-in by borrowers: good borrowers are tied to their banks. This paper shows that an informational lock-in effect occurs even if borrowers are identical. Asymmetric information on banks generates an informational lock-in for borrowers. A borrower is tied to the initial bank even if it charges higher loan interest. The borrower is not ready to leave the bank and take a risk that the new bank proves to be even worse.

Keywords Asymmetric information, banking, relationship lending, bank competition, switching costs
JEL classification G21

1. Introduction

Innovative articles by Sharpe (1990), Rajan (1992) and von Thadden (2004) design the theory of relationship lending that provides a theoretical explanation for actual long-term bank-firm relationships. The relationship lending theory has had a groundbreaking impact on the banking literature. Customer relationships arise between banks and firms (i.e. borrowers) because, in the process of lending, the bank that does the actual lending to a firm learns more about that borrower's characteristics than other banks. An important consequence of this asymmetric evolution of information is the potential creation of ex post, or temporary, monopoly power: the existing bank has an information advantage over potential competitors at the refinancing stage. The monopoly power allows the bank to capture some of the rents generated by its old borrowers. Due to competition, however, the rents are eroded through low introductory loan interest offers to all firms in their initial period, precisely when banks know the least about the firms. Banks lend to new borrowers at interest rates which initially generate expected losses. The relationship lending theory suggests that firms stay with the same bank because high quality firms are, in the sense, informationally captured. This is due to the difficulties firms face in conveying information about their superior performance to other banks. Adverse selection makes it difficult for one bank to attract another bank's good borrowers without also drawing the less desirable ones as well.

* University of Turku, Department of Economics, 20014 Turku, Finland. Phone: +358023335404, E-mail: juhnii@utu.fi.

This paper suggests an alternative explanation for long-term lending relationships and lock-in effects. To highlight the deviation from the classic theory of relationship lending, we assume that borrowers are identical. On the contrary, banks are different. Some banks are good at helping firms boost their returns as in Boot and Thakor (2000) and Song and Thakor (2007), but bad banks cannot offer this kind of support. The bank type is private information. In the process of lending, the firm that actually borrows from a bank learns more about that bank's characteristics than other firms. The borrower learns whether the bank is good or bad.[1] A consequence of this asymmetric evolution of information is the creation of monopoly power. Even if the firm learns to know the bank type in the process of borrowing, the monopoly power is channeled to the bank. The monopoly power allows a good bank to capture some of the rents generated by an old borrower. Because of competition, the rents are lost through low introductory interest rates offered to all firms in the initial period. The model suggests that firms stay with the same good bank because they are informationally captured. This is due to the difficulties firms face in distinguishing between good and bad banks. Adverse selection makes it difficult for a good bank's borrower to seek loan offers from other good banks without attracting offers from the less desirable banks as well. Therefore, the borrower stays with the same familiar bank even if it charges higher loan interest. The borrower will not risk taking a lower interest offer from another bank which may later prove to be a bad bank.

Consequently, both in this model and in the original relationship lending theory (e.g. Sharpe 1990, Rajan 1992, von Thadden 2004) an old borrower is informationally captured in a bank and yields profit for it, whereas a new borrower is unprofitable. In the original theory the lock-in effect is based on asymmetric information on borrower types when banks are identical. Now the lock-in effect is based on asymmetric information on bank types when borrowers are identical. The motivation of our paper is to extend the relationship lending theory by showing a new type of lock-in effect in banking, which generates the same type of interest structure as the original theory.

Empirical research supports (i) the existence of hold-up costs in banking and (ii) the interest structure of relationship lending (i.e. new borrowers pay less loan interest than old ones). Barone et al. (2011) document evidence on hold-up problem in Italy. Banks discriminate between new and old borrowers by charging lower interest rates on the former. The discount amounts to about 44 basis points and is equal to 7% of the average interest rate. Switching costs are higher for single bank firms. On average being a primary bank in the previous period increases the probability of being the main lender by about 70%. The estimated effect is larger for single-bank firms (about 80%) but is also significantly sizeable for multiple-lender enterprises (40%). These findings are supported by Ioannidou and Ongena (2010). They find that a loan granted by a new outside bank carries a loan rate that is significantly (89 basis points) lower than the rates on comparable new loans from the firm's current inside banks, and 87 basis points lower than the rates on comparable new loans that the outside bank currently

[1] Manove et al. (2001) introduced into the literature on asymmetric information credit contracts the idea that researchers could consider different banks instead of different borrowers. We will thank a referee who informed us on this.

extends to its existing customers. Ioannidou and Ongena (2010, p. 1848) go on: "We also find that when the firm switches, the outside bank is willing to decrease loan rates by another 36 basis points … The combined reduction of 122 basis points comprises almost one-tenth of the average observed loan cost of 13.4%. However, a year and a half after a switch, the new bank starts hiking up the loan rates—even if the firm's condition has not deteriorated. Rates increase slowly at first but eventually at a clip of more than 30 basis points per year." In sum, an outside bank initially decreases the loan rate but eventually raises it. These findings are consistent with the original theory of relationship lending and our alternative theory. For evidence on lock-in effects in banking, see also Schenone (2010) and Kano et al. (2011).

We study relationship lending in which a bank can improve the expected output of the borrower's project. For example, Boot and Thakor (2000) and Song and Thakor (2007) examine this type of relationship lending. Is this type of assumption realistic? Boot and Thakor (2000) give three examples on this type of help. First, a bank can provide additional financing to a liquidity-constraint firm after receiving inside information about the firm. Second, a bank can restructure the debt of a financially distressed firm by reducing its near-term repayment obligation in exchange for a higher repayment later. Third, a bank may finance many firms in the same industry. This creates specific information for the bank which can provide this information to borrowers. Scott (1986) surveys the evidence on this type of help and draws the following conclusions:

> "Many commercial banks, for example, routinely provide both financial and managerial advice to business firms … banks indicated that they made special efforts to accommodate small business borrowers by providing financial counseling, and referrals to technical and management assistance as non-fee services. As part of their cash management services, most commercial banks now offer comprehensive analysis of customer receipts and disbursements, as well as credit information, market analysis, financial management assistance and production advice." Scott (1986, p. 948–949)

Obviously, the level of this kind of help may vary between banks. Alternatively, it is possible to interpret the bank's service in a different way. In a loan contract a borrower commits to numerous covenants. The firm, for example, cannot sell its property during the loan period or replace existing key persons. The firm must meet several accounting-based ratios. The covenants may limit the growth of the firm and prevent investments in new industries. The covenants reduce the lender's risk but rigid enforcement may cause severe problems to the borrower. In reality, loan covenants are usually flexible and the borrower may receive a waiver from the lender. The lender evaluates the case and waives the right to enforce the contract if the borrower benefits from the waiver and if the decision does not add to the lender's risk. Obviously, a rational borrower expects the covenants to be flexible. Consider now two banks: Bank G and Bank B. Bank B minimizes credit losses though strict covenants and later grants no waivers even if this causes severe problems to the borrowers. Bank G also favors tight loan covenants but grants several waivers at a later date after a careful analysis. Under asymmetric information the bank type is unobservable. Hence, a borrower faces the risk that its bank proves to represent type B. Finally, the effects of unexpected external shocks vary between banks. Lo (2014) provides examples in which exposed banks reduced lending volume and increased loan spreads substantially more than other banks after a crisis. Thus, a borrower faces a risk that its bank represents a type, which has a high probability to reduce lending in the future. Gopalan et al. (2011, p. 1335) examine reasons for new bank relationships and report: "Our findings suggest that firms form new banking relationships to expand their access to credit ..." On average, firms obtain higher loan amounts when they form new banking relationships, while small firms also experience and increase in sales growth, capital expenditures and leverage. Hence, differences between lending policies drive borrowers to switch banks. This result supports our theory.

Finally, we review novel research on relationship lending. This literature is so numerous that it is possible to review only a small part of it. Since Boot (2000) and Freixas and Rochet (2008) survey prior research extensively, we focus on new research. To begin, there exist few new theoretical articles which apply relationship banking. Repullo and Suarez (2013) examine bank capital regulation using a model of relationship lending. Relationship lending and the transmission of monetary policy are investigated by Hachem (2011). Niinimäki (2014) extends the relationship lending theory to bank regulation and Niinimäki (2015) extends this theory to loan collateral. The magnitude of interesting empirical research is large. Chang et al. (2014) present evidence from China. The bank's relationship information (soft information) is statistically and economically significant in forecasting loan defaults. This information contributes the most significant improvement in default prediction, more than four times larger than the improvement arising from the firm-specific hard information. Agarwal et al. (2011) investigate home equity loans and lines-of-credit applications in U.S.A. The analysis confirms the importance of soft information and suggests that its use can be effective in reducing overall portfolio credit losses. Uchida et al. (2012) discover that more soft information tends to be accumulated when loan officer turnover is less, and when loan officer contact is frequent. These findings from Japan support the vision that the "relationship" in relationship lending is the loan officer–entrepreneur relationship, not the bank-entrepreneur relationship. Uchida (2011) explores lending decisions in Japan. Banks stress three factors: the relationship factor, the financial statement factor, and the collateral/guarantee factor. The relationship factor is crucial for small banks and under intensive competition. Cotugno et al. (2013) find that in Italy a strong bank-borrower relationship mitigates credit rationing. Distance has negative impact on credit availability. Cenni et al. (2015) report that in Italy the probability of credit rationing increases with the number of lenders and decreases with the length of the relationship with the main bank. Debt concentration with the main bank affects positively small firms. Fiordelisi et al. (2014) also utilize Italian data and discover that a closer and longer bank-borrower relationship decreases the probability of the borrower's financial distress. Geršl and Jakubík (2011) find that in Czech Republic the level of a bank's credit risk decreases with the share of relationship loans in the bank's portfolio. The research problems and results of these articles differ from our paper.

The paper is organized as follows. Section 2 introduces the economy. Section 3 investigates the operations of good and bad banks under symmetric and asymmetric information. Section 4 draws conclusions.

2. Economy

Consider a risk-neutral economy with N banks and N borrowers (=firms), where N approaches infinity. Banks and firms maximize their expected returns. The economy has two periods: period 1 and period 2. Banks can raise unlimited quantity of deposits by paying (gross) interest r on them. Here r is the risk-free interest rate of the economy.

Firms and projects A firm can undertake an investment project in each period. A project lasts for a period and requires a unit of investment input. If a project is successful, it produces Y units. If unsuccessful, the output is zero. With bank counseling a project succeeds with certainty. Without counseling it succeeds with probability p. A firm invests effort e in a project at the start of a period with certainty. The firm has no wealth of its own and it borrows a unit of capital from one bank for a period.

Banks Banks raise deposits, grant loans and may counsel borrowers. Banks have other returns which make them risk free. To shorten the study, we do not model these returns. We focus entirely on loans. Two bank types exist: good and bad. A good bank has the capacity to counsel borrowers and thus raise the probability of project success from p to 1. This type of relationship lending in which banks counsel borrowers and thereby boost the probability of project success is similar to Boot and Thakor (2000). As in Boot and Thakor (2000) we assume that the counseling capacity entails cost C to a bank in each period. More precisely, good banks may, for example, hire managers for two periods to counsel borrowers. Since good banks have already purchased the counseling capacity, they bear these sunk costs with certainty. Since the counseling process raises the probability of project success from p to 1, a good bank can increase loan repayments through counseling. Hence, good banks are always motivated to counsel borrowers. A bad bank cannot counsel borrowers and their projects succeed with probability p. The bank type is fixed: a bank is either good in each period or bad in each period. The bank type is private information and unobservable to outsiders. However, during the loan period the borrower learns the type of the bank but is unable to communicate this fact to other borrowers (this information becomes private to the two parties in the relationship—the bank and the borrower). The bank type causes a risk to a borrower. If the bank proves to be good, the borrower's project succeeds with certainty. If bad, the project succeeds with probability $p < 1$. Even if the bank type is unobservable to outsiders the shares of good and bad banks are commonly known. The share of good banks in the economy is λ and the share of bad banks is $1-\lambda$. Banks compete for borrowers and are able to charge different loan interest rates from (i) new borrowers, (ii) old borrowers of other banks and (iii) their own old borrowers (which they already know from period 1). A borrower is free to switch banks after period 1. We make the following assumptions.

Assumption 1. *With counseling, a project is profitable to a firm:*

$$\pi_G(r+C) = Y - r - C - e > 0$$

Assumption 2. *Without counseling, a project is unprofitable to a firm even if the loan interest is at the minimum level:*

$$\pi_B(r) = p(Y - r) - e < 0$$

It is easy to see that bad banks have a negative contribution to an overall wealth of the economy. A bad bank can attract borrowers only because they do not observe its type. The share of bad banks is assumed to be sufficiently low in the economy and thus firms optimally seek for loans in period 1: $\lambda\pi_G(R_1) + (1-\lambda)\pi_B(R_1) > 0$ or $[\lambda + (1-\lambda)\ p](Y - R_1) - e > 0$. Here R_1 denotes the loan interest rate of period 1. We find out its exact value in Section 3. The following assumption clarifies the model.

Assumption 3. *Borrowers, who seek for a new bank, take up positions evenly in active banks.*

If each bank is active in period 1 the number of borrowers in each bank is one.

Assumption 4. *A borrower knows in period 2 the banks that were active in period 1.*

Assumption 4 is necessary for the following reason. Bad banks are more profitable in period 2 than in period 1. Their returns may be negative in period 1. In this case bad banks maximize their life-time returns by operating only in period 2. This is impossible in the model, because the share of bad banks must be $1-\lambda$ in each period. Assumption 4 forces bad banks to operate in each period. If a bank operates only in period 2, this reveals it bad type to borrowers and it cannot attract borrowers. Under assumption 4, bad banks are ready to earn negative returns in period 1 if their returns are positive in period 2 and the life-time returns are non-negative. Then each bank is active in period 1.

Assumption 5. *The cost of the counseling capacity, C, meets $C \geq \underline{C} > 0$.*

Here $\underline{C}$ is the minimum cost so that bad banks can operate in each period. That is, their life-time returns are non-negative when $C \geq \underline{C} > 0$. We detail $\underline{C}$ later.

The assumptions of Sharpe (1990): Our paper applies most of the standard assumptions of the classic relationship lending theory, e.g. Sharpe (1990). First, only short-term loan contracts for a period are possible. Banks cannot make any commitments in period 1 regarding the loan contracts of period 2. Second, a firm consumes the profit of period 1 after the period. The profit cannot be pledged as loan collateral in period 2. Third, a firm cannot borrow from many banks in a period. Fourth, the realized project output is unobservable to outsiders.[2] This assumption ensures in the classic relationship lending theory that outsiders cannot utilize the borrower's materialized project output in period 1 to update information on him. As a result, the

[2] The fact that the project output is totally unobservable to outsiders is one important case in Sharpe (1990).

inside bank of period 1 has more information on him than outsiders. The inside bank is also motivated to hide information on loan repayments so that competing banks cannot bid away its best borrowers. This assumption—the project output is unobservable to outsiders—is needed in our model, because the borrowers of good banks always succeed in their projects. A loan loss reveals that the bank represents the bad type. Hence, bad banks are motivated to hide realized project outputs and loan losses so they that can conceal their true type and attract borrowers.[3] By adopting these assumptions our paper follows the tradition of the relationship lending theory. Yet, we do not need the Sharpe's assumption that a borrower does not know his type in period 1. Now each borrower and each bank knows its type but bank type is unobservable to borrowers. The timing of the model is as follows.

1. At the start of period 1 each bank offers a short-term loan interest for period 1. In this context banks rationally anticipate the expected return from these borrowers in period 2.
2. During period 1 each borrower learns the type of its bank (be it good or bad).
3. Project outputs materialize and firms repay the loans to the banks which pay back deposits.
4. The start of period 2. Banks announce loan rate offers to the borrowers of other banks.
5. Each bank makes loan offers to its old borrowers who borrowed from the bank in period 1.
6. The firms choose their banks, borrow capital and invest the capital in the projects.
7. Projects mature at the end of period 2. Firms repay loans to banks which pay back deposits.

3. Bank operations in period 1 and period 2

Section 3 consists of the following parts. Subsection 3.1 examines a benchmark case under perfect information.[4] The rest of the section analyzes banking under asymmetric information. Subsection 3.2 outlines the operations of a good bank whereas subsection 3.3 focuses on a bad bank.

3.1 Benchmark: perfect information

Bank type is observable in each period. In period 2 bad banks cannot attract borrowers, because a loan from a good bank provides more output to a borrower. Good banks compete for borrowers. Competition pushes the loan rate down to the banks' zero profit level, $r+C$. The profit of a good bank is zero in period 2 and has no effect on the

[3] For alternative ways to hide loan losses, see Niinimaki (2012).
[4] Assumption 3 is not applicable in this subsection.

lending decision in period 1. In period 1 the scenario is the same as in period 2. Bank types are observable and bad banks cannot attract borrowers. Good banks compete for borrowers and the loan interest rate is $r+C$. The return of a good bank is again zero. Consequently, under perfect information the scenario is simple. Bad banks cannot attract borrowers. Good banks charge loan interest $r+C$ in each period and earn zero profit from every loan in each period. Now we turn to banking under asymmetric information.

3.2 Asymmetric information: a good bank

At the start of period 2 each borrower knows the type of its initial bank during period 1. First we explore the optimal strategy of a good bank in period 2. Thereafter we study it in period 1.

Period 2. A borrower knows that if he continues the same lending relationship in period 2, a project succeeds with certainty. Let R_g denote the loan interest offer of the initial good bank for period 2. Alternatively, the borrower can find a new bank for period 2. Unfortunately, the borrower cannot observe the type of the new bank. With probability λ the type is good and with probability $1-\lambda$ it is bad. Let R_o denote the loan interest offer of outside banks (good and bad) for period 2. The borrower continues the initial loan relationship also in period 2 if $Y-R_g \geq \lambda(Y-R_o)+(1-\lambda)p(Y-R_o)$. The L.H.S. displays the firm's return under the initial bank relationship. The R.H.S. indicates the firm's expected return if it switches its bank. The first (second) term on the R.H.S. denotes the case in which the new bank is good (bad). Obviously, λ and $1-\lambda$ define the prior probabilities of good and bad banks in the economy. The borrower learns the type of one bank in period 1. This implies that the borrower which learned the information is able to update its probability assessment of the chances of each of remaining $N-1$ banks being of good or bad type. When the type of the first bank is good, the updated probabilities for other banks are following. A new bank is good with probability $(\lambda N-1)/(N-1)$ and bad with probability $(1-\lambda)N/(N-1)$. Yet, since N is assumed to approach infinity in the economy, these posterior probabilities approach λ and $1-\lambda$.[5] Hence, the updated probabilities on the R.H.S. are identical to the prior probabilities. The breakeven interest offer of outside good banks meets $R_o = r+C$. Outside bad banks make the same offer, because they must mimic good outside banks to be able to attract borrowers for period 2. An outside bad bank cannot attract a borrower if its true type surfaces. From $Y-R_g \geq \lambda(Y-R_o)+(1-\lambda)p(Y-R_o)$ it is easy to solve the optimal (= maximal) interest offer of a initial good bank to its old borrower in period 2:

$$R_g = R_o + (1-\lambda)(1-p)(Y-R_o). \tag{1}$$

Here R_g exceeds the offer of outside banks, R_o. The second term on the R.H.S. shows the interest premium, which increases with the share of bad banks in the economy,

[5] We are grateful to an anonymous referee, who showed us the need to update probabilities in the second period.

$1-\lambda$, and with the probability of a project failure, $1-p$. The premium also increases with the profitability of a successful project, $Y-R_o$. Intuitively, the firm can change its bank after period 1. The change is risky, because the new bank may prove to be bad, the project may fail and the firm may lose the return from the project. Therefore, the borrower avoids outside banks and is motivated to continue the initial bank relationship. The initial bank is rational and recognizes the motivation. It can charge high loan interest in period 2 and still retain the lending relationship. Now (1) reveals the maximal loan interest so that the borrower does not switch its bank. In addition, the interest rate must be so low that the firm is ready to start the project. The interest rate has such an upper limit, $\overline{R}_g$, that $Y-\overline{R}_g-e=0$. To see this, assume the following scenario. When outside banks offer interest R_o the borrower's expected return is negative in period 2 if he switches a bank after period 1 $E(\pi_2(R_0))=\lambda(Y-R_o-e)+(1-\lambda)[p(Y-R_o)-e]<0$. This can be rewritten (given $R_o=r+C$) as follows

$$E(\pi_2(R_0))=\lambda(Y-r-C)+(1-\lambda)p(Y-r-C)-e<0. \tag{2}$$

If (2) is true a borrower is not ready to switch banks after period 1 because a new bank would be bad with a high probability. The switch is unprofitable even if outside banks charge minimum interest on loans, $R_o=r+C$. If (2) is true and if the borrower's initial bank is bad, the borrower leaves the loan market after period 1. If (2) is true and the initial bank is good, it can raise the interest rate of period 2 to the upper limit, $\overline{R}_g$, and the borrower continues the initial bank relationship. If (2) is not true, it is possible to switch banks after period 1 and the initial good bank charges interest R_g, which is sufficiently low to prevent the switch. We can express the good bank's optimal loan interest in period 2 as follows $R_g^*=\min(\overline{R}_g,R_g)$. An old borrower yields a positive return to a bank in period 2

$$\Pi_{2g}=R_g^*-r-C>0. \tag{3}$$

Period 1. Good banks compete for new borrowers and anticipate correctly the rent from lending relationships in period 2. Competition pushes the loan interest of period 1, R_1, down to such a level that the bank return is zero during the whole lending relationship

$$R_1-r-C+\delta\Pi_{2g}=0. \tag{4}$$

Here $\delta=1/r$ is a discount factor and R_1 is so low that the bank return from a new borrower is negative in period 1, $R_1<r+C$.[6] Good banks make low introductory loan offers to establish valuable lending relationships. A conclusion follows.

Proposition 1. *The hold-up problem is present. Old borrowers are tied to good banks in period 2 and these banks can charge high interest from old borrowers. Old borrowers are profitable for good banks. Competition for new borrowers makes the bank return from new entrants negative. The expected bank return during the whole lending relationship is zero.*

[6] The discount factor is not necessary in the model. It is possible that $\delta=1$.

3.3 Asymmetric information: a bad bank

A firm is unwilling to borrow from bad banks, because a loan from a good bank is more profitable. To attract a borrower, a bad bank mimics good banks and hides its true type. We study first period 2 and then period 1.

Period 2. The borrower of period 1 learns the bank type during the period and is not ready to continue the lending relationship in period 2, because the expected return from a project with a bad bank is negative for the borrower. Two alternatives result.

(i) If $E(\pi_2(R_0)) < 0$ in (2) the borrower exits from the loan market after period 1. He abandons the initial bank, which is bad. He will not switch a bank, because a new bank is bad with a high probability. Since all borrowers of bad banks exit from the loan market, no borrower searches for a new bank in period 2. A bad bank cannot attract a new borrower in period 2. Hence, bad banks have no borrowers in period 2 although they offer loans.

(ii) If $E(\pi_2(R_0)) > 0$ in (2) each firm whose initial bank was bad finds a new bank for period 2. The new bank may prove to be good or bad. Hence, a bad bank, which losses its initial borrower after period 1, can attract a new borrower in period 2. The bad bank's expected return from a loan unit to a new borrower in period 2 is $\Pi_{2b} = pR_0 - r$ units. Here $R_0 = r + C$ is the same loan interest as above when outside good banks aim to attract initial borrowers from other banks in period 2. Bad banks must offer the same loan interest.

Period 1. A bad bank offers the same loan interest, $R_1 = r + C - \delta\Pi_{2g}$, as good banks. The expected return of the bad bank from a loan unit, $\Pi_{1b} = pR_1 - r$, can be restated as

$$\Pi_{1b} = pC - (1-p)r - \delta p\Pi_{2g}. \tag{5}$$

The expected life-time return of a bad bank adds up to

$$\Pi_{12} = \begin{cases} \Pi_{1b} + \delta(1-\lambda)\Pi_{2b} & \text{if} \quad E(\pi_2(R_o)) \geq 0 \\ \Pi_{1b} & \text{if} \quad E(\pi_2(R_o)) < 0. \end{cases} \tag{6}$$

Consider scenario $\Pi_{12} = \Pi_{1b} + \delta(1-\lambda)\Pi_{2b}$. The borrowers of period 1 leave bad banks after the period. Banks aim to attract new borrowers for period 2. The total number of switching borrowers who leave their initial bad banks after period 1 and search for a new bank is $(1-\lambda)N$. These switching borrowers take their positions evenly in N banks (Assumption 3). As a result, few good banks obtain a new borrower in period 2 and have two borrowers in period 2. The rest of the good banks do not obtain a new borrower in period 2 and have then only one borrower. Recall that a good bank always retains its initial lending relationship during the second period. Few bad banks obtain a new borrower in period 2 and have one borrower in this period. The rest of the bad banks have no borrowers in period 2. Recall that a bad bank always loses its initial lending relationship after period 1. As a result, the expected number of borrowers in a bad bank is $1-\lambda$ during period 2. Consider now scenario $\Pi_{12}=\Pi_{1b}$. Since all borrowers of bad banks exit from the loan market after period 1, no borrower searches for a new bank in period 2. A bad bank cannot attract new borrowers in period 2 although it offers loans. Hence, bad banks have borrowers only in period 1.

It is easy to observe that the return in period 2 exceeds the return of period 1, $\Pi_{2b}>\Pi_{1b}$. A bad bank is more willing to operate in period 2. Whether $\Pi_{12}=\Pi_{1b}$ or $\Pi_{12}=\Pi_{1b}+\delta(1-\lambda)\Pi_{2b}$ in (6), Π_{12} increases with C and Π_{12} is negative when $C=0$ and positive when C is sufficient. There exists $\underline{C}$ so that Π_{12} is zero. Therefore, when $C\geq\underline{C}$, bad banks can operate in each period and earn non-negative life-time returns. It is possible that a bad bank is willing to operate only in period 2 if $\Pi_{2b}>0>\Pi_{1b}$. This is impossible, because borrowers recognize the banks that operate only in period 2 (Assumption 4). The choice to operate only in period 2 reveals the bad type. Assumption 5 ensures that the characteristics of the economy are such that everyone acts "correctly" in the model. That is, bad banks are willing to participate in loan markets in each period, because their life-time returns are non-negative.

It is now possible to sum the findings in an environment that meets the assumptions. Bad banks offer the same loan interest rates as good banks. The lending relationships of bad banks are short-term (one period). When a borrower learns that his bank is bad, he abandons it. A short-term lending relationship may be profitable to a bad bank even if it is unprofitable to a good bank, because the operation costs are lower for bad banks. They avoid cost C. In period 2, for example, outside banks attract borrowers from the initial banks by offering loan interest $R_0=r+C$. This loan interest generates zero return to outside good banks but positive expected return, pR_0-r, to outside bad banks if C is sufficient. In period 1, outside banks offer $R_1=r+C-\delta\Pi_{2g}$ to new borrowers. The interest rate is so low that good banks bear losses in period 1, but the expected return of a bad bank, pR_1-r, is positive if C is sufficient. Given the assumptions of the model, $C\geq\underline{C}$, bad banks can operate and earn non-negative life-time returns. The non-negative return of a bad bank is based on asymmetric information, which makes it possible for them to attract borrowers. Each lending relationship with a bad bank is a mistake from the borrower's point of view. A conclusion follows.

Proposition 2. *A bad bank operates in each period and it has short-term lending relationships. Three alternative scenarios are possible. Firstly, a bad bank may earn positive return in each period, because it has a lighter cost structure than good banks. Secondly, a bad bank makes a negative return in the first period but the return is positive in the later period. The life-time return is non-negative. Thirdly, a bad bank may earn positive (or zero) return in the first period and it has no borrowers in the later period.*

The first two scenarios are possible when $E(\pi_2(R_0))>0$. The last scenario occurs when $E(\pi_2(R_0))<0$. Bad banks lose initial borrowers after period 1 and these borrowers do not search for a new bank in period 2. The borrowers of good banks continue their initial bank relationships with good banks even if these charge high loan interest $\overline{R}_g$ in period 2. Outside banks cannot attract borrowers from the initial lenders. Hence, bad banks have no borrowers in period 2 although they offer loans. Appendix gives a numerical example, which clarifies Propositions 1 and 2.

4. Conclusions

The question of whether relationship lending provides the lender an information monopoly, which the lender exploits to extract rents from its lock-in borrowers, has captured the interest of many academics and practitioners. In this paper we take a novel approach on this question. The paper presents a model in which asymmetric information on bank types generates a lock-up effect even when the borrowers are identical. A borrower is tied to a good bank, because he does not want to risk accepting a lower loan interest offer from another bank which may represent the bad bank type. Hence, old borrowers yield high returns for good banks. The banks compete fiercely for new borrowers in order to establish valuable lending relationships. As a result, new borrowers are unprofitable for the good banks whereas old borrowers yield profit. The expected life-time return from a lending relationship is zero to a good bank. A bad bank attempts to hide its true type in order to be able to attract borrowers. These banks have only short-term lending relationships. When a borrower learns that his bank is bad, he switches banks. Yet, a bad bank may be more profitable than a good bank owing to its light cost structure.

The model is based on incomplete contracts between banks and borrowers. Banks cannot commit to detailed policies during the loan period. As a result, the bank's actual policy during the loan period may generate losses to the borrower. The contribution of a bad bank to the overall wealth of the economy is negative. When the share of bad banks in the economy is sufficient, the expected project output is negative. If bank regulators can create a method to acquire information on true lending strategies of banks and make this information public, asymmetric information on lending strategies mitigates and borrowers learn to avoid bad banks. That is, regulators ought to improve the transparency of banks.

The classic theory of relationship lending (e.g. Sharpe 1990, Rajan 1992, von Thadden 2004) examines banks which have asymmetric information on borrowers. Our study is based on assumption that identical borrowers have asymmetric information on banks' type. Is our assumption realistic? Banks are usually more transparent than small firms. Yet, it is natural to assume that the bank's ability to boost borrower's output is unobservable to outsiders. It is also natural to assume that outsiders cannot observe the bank's policy regarding loan covenants. Thus, there is asymmetric information on some aspects of banks. In the future it would be interesting to study a case in which there is asymmetric information on both banks' type and borrowers' type.

We simplify the model by assuming that borrowers cannot contact each other and communicate on bank types. In reality communication may be possible but it is not credible. It is impossible to prove the bank's ability to boost the project output later to other borrowers. It is also difficult to prove the bank's policy regarding loan covenants to outsiders. An unsuccessful borrower may blame his bank afterwards even when the quality of the bank service has been good.

In banking sector switching costs are also important from a macroeconomic point of view. They may decrease price elasticity in credit markets so that the transmission of policy rate changes to retail interest rate dynamics may exhibit some form of sluggishness because banks may not find it profitable to adjust their loan offers frequently.

In addition, strong lending relationships can reduce the negative effects of a crisis on the availability of firms to access credit.

As to the empirical implications of the model, empirical evidence on switching costs, lock-in effects and loan interest rates supports our findings. Yet, the origin of the switching costs differs in our model from the classic theory of relationship banking (e.g. Sharpe 1990, Rajan 1992, von Thadden 2004). In our model asymmetric information on banks' type creates switching costs whereas in the classic theory of relationship lending asymmetric information on borrowers' type generates switching costs. Therefore, in the empirical research it is necessary to find out the true origin for switching costs.

It is possible to reinterpret the model. Initial and later periods may represent "good times" and "bad times" within an economic cycle.[7] Santos and Winton (2008, p. 316) find evidence that "... during recessions banks raise their rates more for bank-dependent borrowers than for those with access to public bond markets. Further analysis suggests that much of this is due to informational hold-up effects rather than to greater risk of bank-dependent borrowers versus those with bond market access." The results are supported by the findings of Mattes et al. (2013, p. 177): "We find that information monopolies exist in periods of economic contraction: only weak banks raise their spreads above the level that is justified by the credit risk for borrowers with a high cost of switching lenders." Furthermore, Asea and Blomberg (1998) report evidence that banks change their lending standards (e.g. loan spread)—from tightness to laxity—systematically over the business cycle. Bernanke et al. (1996) survey abundant empirical evidence on banks' tight lending policy during recessions. This evidence provides some support for our model. It is difficult for a bank-dependent borrower to switch banks during recessions. A rational bank knows this and charges high interest on these borrowers during recessions. A long lending relationship with the initial bank ensures that a borrower receives a loan (e.g. Cotugno et al. 2013, Cenni et al. 2015).

The paper offers a strongly simplified model which presents the basic concept. The model includes a few assumptions. Although most of the assumptions are the same as in the classic theory of relationship banking (e.g. Sharpe 1990, Boot and Thakor 2000), it would be worthwhile to design a more sophisticated model in the future and drop the assumptions step by step.

Acknowledgement The author is indebted to editor Martin Gregor and two anonymous referees for numerous useful comments and suggestions.

References

Agarwal, S., Ambrose, B., Chomsisengphet, S. and Liu, C. (2011). The role of soft information in a dynamic contract setting: evidence from the home equity credit market. *Journal of Money, Credit and Banking,* 43(4), 633–655.

[7] We are grateful to an anonymous referee who suggested this option to reinterpret the model to us.

Asea, P. and Blomberg, B. (1998). Lending cycles. *Journal of Econometrics*, 83, 89–128.

Barone, G., Felici, R. and Pagnini, M. (2011). Switching costs in the local credit markets. *International Journal of Industrial Organization*, 29, 694–704.

Bernanke, B., Gertler, M. and Gilchrist, S. (1995). The financial accelerator and the flight to quality. *Review of Economics and Statistics*, 78(1), 1–15.

Boot, A. (2000). Relationship banking: what do we know? *Journal of Financial Intermediation*, 9(1), 7–25.

Boot, A. and Thakor, A. (2000). Can relationship banking survive competition? *Journal of Finance*, 55(2), 679–713.

Cenni, S., Monferra, S., Salotti, V., Sangiorgi, M. and Torluccio, G. (2015). Credit rationing and relationship lending. Does firm size matter? *Journal of Banking and Finance,* 53, 249–265.

Chang, C., Liao, G., Yu, X. and Ni, Z. (2014). Information from relationship lending: Evidence from loan defaults in China. *Journal of Money, Credit and Banking*, 46(6), 1225–1257.

Cotugno, M., Monferra, S. and Sampagnaro, G. (2013). Relationship lending, hierarchical distance and credit tightening: Evidence from the financial crisis. *Journal of Banking and Finance,* 37, 1372–1385.

Fiordelisi, F., Monferra, S. and Sampagnaro, G. (2014). Relationship lending and credit quality. *Journal of Financial Services Research*, 46, 295–315.

Freixas, X. and Rochet, J.-C. (2008). *Microeconomics of Banking*. Cambridge, MIT Press.

Geršl, A. and Jakubík, P. (2011). Relationship lending in emerging markets: Evidence from the Czech Republic. *Comparative Economic Studies*, 53, 575–596.

Gopalan, R., Udell, G. and Yerramilli, V. (2011). Why do firms form new banking relationships? *Journal of Financial and Quantitative Analysis*, 46(5), 1335–1365.

Hachem, K. (2011). Relationship lending and the transmission of monetary policy. *Journal of Monetary Economics*, 58, 590–600.

Ioannidou, V. and Ongena, S. (2010). Time for a change: Loan conditions and bank behavior when firms switch banks. *Journal of Finance*, 65(5), 1847–1877.

Kano, M., Uchida, H., Udell, G. and Watanabe, W. (2011). Information verifiability, bank organization, bank competition and bank-borrower relationships. *Journal of Banking and Finance* 35, 934–954.

Lo, A. (2014). Do declines in bank health affect borrowers' voluntary disclosures? Evidence from international propagation of banking shocks. *Journal of Accounting Research*, 52(2), 541–581.

Manove, M., Padilla, J. and Pagano, M. (2001). Collateral versus project screening: A model of lazy banks. *RAND Journal of Economics*, 32(4), 726–744.

Mattes, J., Steffen, S. and Wahrenburg, M. (2013). Do information rents in loan spreads persist over business cycles? *Journal of Financial Services Research,* 143, 175–195.

Niinimäki, J.-P. (2012). Hidden loan losses, moral hazard and financial crises. *Journal of Financial Stability,* 8(1), 1–14.

Niinimäki, J.-P. (2014). Relationship lending, bank competition and financial stability. *Czech Economic Review*, 8(3), 102–124.

Niinimäki, J.-P. (2015). The optimal allocation of alternative collateral assets between different loans. *North American Journal of Economics and Finance*, 34, 22–41.

Rajan, R. (1992). Insiders and outsiders: the choice between informed and arm's-length debt. *Journal of Finance*, 47(4), 1367–1400.

Repullo, R. and Suarez, J. (2013). The procyclical effects of bank capital regulation. *Review of Financial Studies*, 26(2), 452–490.

Santos, J. and Winton, A. (2008). Bank loans, bonds, and information monopolies across the business cycle. *Journal of Finance*, 48(3), 1315–1359.

Schenone, C. (2010). Lending relationships and information rents: Do bank exploit their information advantages? *Review of Financial Studies*, 23(3), 1149–1199.

Scott, R. (1986). A relational theory of secured financing. *Columbia Law Review*, 86(5), 901–977.

Sharpe, S. A. (1990). Asymmetric information, bank lending and implicit contracts: A stylized model of customer relationships. *Journal of Finance,* 55, 1069–1087.

Song, F. and Thakor, A. (2007). Relationship banking, fragility, and the asset-liability matching problem. *Review of Financial Studies*, 20(5), 2129–2177.

Uchida, H., Udell, G. and Yamori, N. (2012). Loan officers and relationship lending to SMEs. *Journal of Financial Intermediation*, 21, 97–122.

von Thadden, E.-L. (2004). Asymmetric information, bank lending and implicit contracts: The winner's curse. *Finance Research Letters*, 1, 11–23.

Appendix

The Appendix provides a numerical example, which shows that the considered restrictions are not mutually inconsistent (i.e. it shows that we are not considering empty set of solutions). We have three cases, which are based on the cases of Proposition 2. In each case a good bank earns negative returns in period 1, positive returns in period 2 and the life-time returns are zero.

Case I. This case represents the first case of Proposition 2, in which a bad bank earns positive returns in each period. Consider the following economy: $Y = 3$, $r = 1$, $p = 0.97$, $\lambda = 0.9$, $C = 0.04$, $e = 1.95$. Assumption 1 is met because $\pi_G(r+C) = 0.01 > 0$. Assumption 2 is also satisfied because $\pi_B(r) = -0.01$. In period 2 outside banks offer interest $R_0 = r + C = 1.04$ to the old borrowers of other banks. From (2) we observe that the borrowers of bad banks can switch banks after period 1 because $E(\pi_2(R_0)) = 0.00412 > 0$. Now (1) implies that the initial good bank of period 1 can retain its old borrowers by charging interest $R_g = 1.04588$ that exceeds the costs of the loan, 1.04. From (3) in which $R_g^* = R_g$ we observe that an old borrower yields a positive profit to its initial bank in period 2, $\Pi_{2g} = R_g - r - C = 0.00588$. The loan interest of period 1 can be solved from (4): it is $R_1 = r + C - \Pi_{2g} = 1 + 0.04 - 0.00588 = 1.03412$ and does not cover the costs of the loan, 1.04. Hence, a lending relationship is unprofitable to a good bank in period 1 and profitable in period 2. Now we focus on bad banks. In period 2 a loan provides expected return $\Pi_{2b} = pR_0 - r = 0.97 \times 1.04 - 1 = 0.0088$ to a bad bank. From (5) we obtain the expected return from a loan in period 1, $\Pi_{1b} = 0.0030964$. Thus, loans are profitable to bad banks in each period.

Let us turn to borrowers. In period 2 a borrower, who retains his original lending relationship with a good bank earns $Y - R_g - e = 3 - 1.04588 - 1.95 = 0.00412$, and is ready to participate in the loan markets. The L.H.S. of (2) is positive, 0.00412. Hence, the borrowers, who had a bad bank in period 1, switch banks after period 1 and have a positive expected return in period 2. Now we know that each borrower participates in the loan markets in period 2. In period 1 a borrower is ready to seek for a loan if $[\lambda + p(1-\lambda)](Y - R_1) - e > 0$. In this economy we have $[0.9 + 0.97 \times 0.1](3 - 1.03412) - 1.95 \approx 0.01 > 0$. Borrowers seek for loans in period 1. Hence, borrowers and banks participate in the loan markets in each period.

Case II. This case represents the second case of Proposition 2. A bad bank earns negative expected return in period 1 and positive expected return in period 2. The expected life-time return is positive. Consider a change in the economy of Case I. Now we have $C = 0.0365$. Assumption 1 is satisfied because $\pi_G(r+C) = 0.0135 > 0$. Assumption 2 is also met since $\pi_B(r) = -0.01$. In period 2 banks offer interest $R_0 = r + C = 1.0365$ to the old borrowers of other banks. From (2) we observe that the borrowers of bad banks can switch banks after period 1 because $E(\pi_2(R_0)) = 0.0076095 > 0$. We learn from (1) that the initial good bank of period 1 can retain its old borrowers in period 2 by charging interest $R_g = 1.0423905$ that exceeds the costs of the loan, 1.0365. From (3) in which $R_g^* = R_g$ we observe that an old borrower yields a positive profit to its initial good bank in period 2, $\Pi_{2g} = R_g - r - C = 0.0058905$. The interest of period 1 can be solved from (4): it is $R_1 = r + C - \Pi_{2g} = 1 + 0.0365 - 0.0058905 = 1.0306095$ and does not cover the costs of the loan, 1.0365. A lending relationship is unprofitable to a good bank in period 1 and profitable in period 2. Next we turn to bad banks. In period 2 a loan provides expected profit $\Pi_{2b} = pR_0 - r = 0.97 \times 1.0365 - 1 = 0.005405$. From (5) we obtain the expected return from a loan in period 1, $\Pi_{1b} = -0.000308785$. The expected life-time profit of the bad bank can be seen from (6), $\Pi_{12} = -0.000308785 + (1-0.9) \times 0.005405 \approx 0.00023$. Here $1-0.9$ reveals that a bad bank receives a new borrower in period 2 with probability 10 percent. Bad banks earn negative expected return in period 1 and positive expected return in period 2.

Let us turn to borrowers. In period 2 a borrower, who retains his original lending relationship with a good bank, earns $Y - R_g - e = 3 - 1.0423905 - 1.95 = 0.0076095$, and is ready to participate in the loan markets. The R.H.S. of (2) is positive, 0.0076095. Hence, the borrowers, who had a bad bank in period 1, switch banks after period 1 and earn a positive expected return in period 2. We have shown that each borrower participates in the loan markets in period 2. In period 1 a borrower is ready to seek for a loan if $[\lambda + p(1-\lambda)](Y - R_1) - e > 0$. In the current numeric example this implies $[0.9 + 0.97 \times 0.1](3 - 1.0306095) - 1.95 \approx 0.013 > 0$. Borrowers seek for loans in period 1. Hence, borrowers and banks participate in the loan markets in each period.

Case III. We present the last case of Proposition 2. The borrowers of bad banks leave loan markets after period 1. Hence, bad banks have no borrowers in period 2 but these banks are profitable in period 1. To show this we change the economy of Case I a bit. Now we have $e = 1.955$. Assumption 1 is met because $\pi_G(r + C) = 0.005 > 0$. Assumption 2 is also satisfied because $\pi_B(r) = -0.015$. In period 2 banks offer interest $R_0 = r + C = 1.04$ to the old borrowers of other banks. From (2) we observe that the borrowers of bad banks leave loan markets after period 1 because $E(\pi_2(R_0)) = -0.00088 < 0$. From (1) we see that the initial good bank of period 1 can retain its old lending relationships by charging interest $R_g = 1.04588$. Yet, this interest rate exceeds the maximal interest $\overline{R}_g = Y - e = 1.045$. Given $R_g^* = \min(\overline{R}_g, R_g)$, a good bank offers interest 1.045 to its old borrowers in period 2. This interest rate exceeds the costs of the loan, 1.04. From (3) (in which $R_g^* = 1.045$) we observe that an old borrower yields a positive return to its initial bank in period 2, $\Pi_{2g} = 1.045 - 1 - 0.04 = 0.005$. The loan interest of period 1 can be solved from (4): it is $R_1 = r + C - \Pi_{2g} = 1.035$ and it does not cover the costs of the loan, 1.04. A lending relationship is unprofitable to a good bank in period 1 and profitable in period 2. Consider now bad banks. They do not have borrowers in period 2. Now (5) shows the bad bank's expected return from a loan in period 1, $\Pi_{1b} = 0.00395$. Hence, bad banks make profits in period 1 and have no borrowers in period 2. The borrowers of good banks undertake projects in period 2 and earn zero profit. In period 1 a borrower seeks for a loan if $[\lambda + p(1-\lambda)](Y - R_1) - e > 0$. In this economy we have $[0.9 + 0.97 \times 0.1](3 - 1.035) - 1.955 \approx 0.004 > 0$. Borrowers seek for loans in period 1.

Czech Economic Review 9 (2015) 91–103
Acta Universitatis Carolinae Oeconomica

Insurance-Markets Equilibrium with Double Indivisible Labor Supply

Aleksandar Vasilev*

Received 4 December 2015; Accepted 15 December 2015

Abstract This note describes the lottery- and insurance-market equilibrium in an economy with both private and public sector employment and non-convex labor supply. In addition, when households are constrained to search for jobs only in a certain sector, the framework requires that there should be separate insurance markets: a public and a private sector one, which would pool the unemployment risk of the corresponding group of households. The unemployment insurance market segmentation is a new result in the literature and a direct consequence of the non-convexity of the labor supply in each sector and the sorting effect of the sector-type shock introduced in the model setup.

Keywords indivisible labor, public employment, insurance
JEL classification H31, J21

1. Introduction

Changes in hours account for approximately two-thirds of the cyclical output volatility in the standard real business cycle model (Cooley and Prescott 1995, Kydland 1995). Those hours, however, are assumed to be supplied in the private sector only, and thus the private-public sector labor choice is ignored. While this might be a reasonable assumption for the US economy, it comes in a stark contrast with the European Union (EU) evidence—after all, central governments in EU countries are the biggest employers at a national level, and public employment is a significant share of total employment.

This note adds to the literature by distinguishing between the two types of labor supply decisions by focusing on the fact that most of the volatility in hours is driven by volatility in employment. That is, most workers in Europe are employed full-time and in addition, only very rarely move between public and/or private sector, as documented in Gomes (2012).[1] Thus, the non-convex labor supply decisions (either work a full week on a job, or not work at all) in both sectors are taken under scrutiny, and the note will try to uncover whether this double binary labor supply decision could provide new implications for business cycle fluctuations in different EU member states.

* CERGE-EI Affiliate Fellow, Prague; Department of Economics, American University in Bulgaria, 1 Georgi Izmirliev Sq., Blagoevgrad 2700, Bulgaria. Tel.: 00 359 73 888 482, E-mail: avasilev@aubg.edu.

[1] In the setup, we model this lack of mobility between sectors via a shock process that sorts workers into a private-sector or a public-sector type.

In an earlier paper, Vasilev (2015a) extends Rogerson's (1988) and Hansen's (1985) static setup by augmenting it with a public sector, and introducing a shock that determines each household's type to be either private-sector or public-sector. The households then search for work in the sector corresponding to their type. Vasilev (2015a) then aggregates over individual households' utility functions, and finds that the resulting utility representation features constant, but different disutility of labor in the two sectors. The aggregate utility function then can not only accommodates the fact that average public sector wages feature a significant mark-up over private sector ones, as documented in Vasilev (2015b), but also allows for an additional transmission and propagation mechanism of shocks through the endogenous public sector labor choice.

In contrast to this earlier study, the focus of the present note falls on the lottery- and insurance-market equilibrium for the setup in Vasilev (2015a). When households in the setup are constrained to search for jobs only in one of the two sectors, in equilibrium there should be separate insurance firms: one for the public sector and another for the private sector, where each insurance company would pool the unemployment risk of the corresponding group of workers. This insurance market segmentation is an important new result in this literature, and is due to the presence of the double non-convexity, as well as the sorting effect of the sector-type shock in the model setup.

2. Model setup

The model follows Vasilev (2015a). The theoretical setup is a static economy, where agents face a non-convex decision in a two-sector economy. There is a large number of identical one-member households, indexed by i and distributed uniformly on the $[0,1]$ interval. The households will be assigned a sector "type," and after the type is revealed, each one decides whether to work in that sector or not. In the exposition below, we will suppress the index i to save on notation.

2.1 Households

Each household maximizes the following utility function

$$\max_{\{c,h^p,h^g\}} \left\{ \ln(c^\eta + S^\eta)^{\frac{1}{\eta}} + \alpha \ln(1 - h^p - h^g) \right\}, \tag{2.1}$$

where c, S, h^p, h^g denote household's private consumption, consumption of the public good, hours worked in the private sector, and hours worked in the government sector. The parameter $\alpha > 1$ measures the relative weight of leisure in the utility function. Total consumption is a Constant Elasticity of Substitution (CES) aggregation of private consumption and consumption of government services, where $\eta > 0$ measures the degree of substitutability between the two types of consumption.[2]

[2] The separability of consumption and leisure is not a crucial assumption for the results that follow. A more general, non-separable, utility representation, does not generate new results, while significantly complicates the algebraic derivations, and thus interferes with model tractability.

Each household is endowed with 1 unit of time that can be allocated to work in the private sector, work in the government sector, or leisure, so $h^p + h^g \leq 1$. Labor supply in each sector is discrete $h^p \in \{0, \overline{h}^p\}$, $h^g \in \{0, \overline{h}^g\}$, where $\overline{h}^p \leq 1$, $\overline{h}^g \leq 1$, and $\overline{h}^p + \overline{h}^g > 1$. In other words, working full-time in both sectors is infeasible, as it takes more than the total time available. Thus, the paper is consistent with Gomes (2014), who assumes that looking for a job will follow a "directed search" process: Each household decides in each period whether to go to the public or to the private sector (or, alternatively, is assigned a "sector type"). This process is stochastic and has two realizations. The probability of going to the private sector (or being a "private-sector type") is

$$q = \frac{H^p}{H^p + H^g}, \tag{2.2}$$

where uppercase letters denote aggregate quantities, i.e. H^p denotes aggregate hours in the private sector, and H^g are the aggregate hours worked in the public sector. Then the probability of being a public sector type is

$$1 - q = \frac{H^g}{H^p + H^g}. \tag{2.3}$$

This process is *i.i.d.* across individuals, so the Law of Large Number holds: At the aggregate level, q share of the households will be private sector type (and thus each household of this type would thus choose $h^g = 0$, as it searches for work only in the private sector), and $1-q$ share will be public sector type (and thus each household of this type would thus choose $h^p = 0$, as it searches for work only in the public sector).[3] Once the particular sector-type is determined, each household decides on its labor supply accordingly. Note that the setup is quite general and allows for different wage rates per hour worked in the two sectors.

In addition to the work income, households hold shares in the private firm and receive profit share π, with $\int_0^1 \pi di = \Pi$.[4] Income is subject to a lump-sum tax t, where $\int_0^1 t di = T$. Therefore, each household's budget constraint is

$$c^j \leq w^j h^j + \pi - t, \quad j = p, g. \tag{2.4}$$

Households act competitively by taking the wage rates $\{w^p, w^g\}$, aggregate outcomes $\{S, H^p, H^g\}$ and lump-sum taxes $\{T\}$ as given. Each household chooses $\{c^j, h^p, h^g\}$ to maximize (2.1) s.t. (2.2)–(2.4).

[3] So the labor supply choice in a sector different from the type of the respective household is degenerate, as it will never be positive.

[4] This is a technical assumption which would guarantee a positive consumption to either of the two types, even if they choose not to work in their sector.

2.2 Firms

Next, there is a single firm producing a homogeneous final consumption good, which uses labor H^p as the only input. The production function is given by

$$Y = F(H^p),\ F' > 0,\ F'' < 0,\ F'(\overline{H}^p) = 0,$$

where the last assumption is imposed to proxy a capacity constraint. The firm takes $\{w^p\}$, aggregate outcomes $\{S, H^g\}$ and policy variable $\{T\}$ as given, and chooses $\{H^p\}$ to

$$\max_{H^p} F(H^p) - w^p H^p \quad \text{s.t } H^p \geq 0.$$

2.3 Government

The public authority hires H^g employees to provide public services, which are paid $w^g = \gamma w^p$, with $\gamma \geq 1$, as in the EU, average public sector wages feature a mark-up over private sector ones (Vasilev 2015b). The production function of non-market public services is as follows:

$$S = S(H^g),\ S' > 0,\ S'' < 0,\ S'(\overline{H}^g) = 0,$$

where the last assumption guarantees that not all "public-sector types" will work in the production of the public good. The public sector wage bill is financed by levying a lump-sum tax T on all households, or $w^g H^g = T$. The government takes H^g as given, and sets w^g, as a fixed gross mark-up above w^p, while T is residually chosen to ensure budget balance.

Vasilev (2015a) establishes that in equilibrium, given an initial realization of a type-specific shock, a fraction λ^p of the private-sector-type households would be working in the private sector, where c_w^p denotes consumption of those working, and c_n^p denotes consumption of those not working. Similarly, a fraction λ^g of the public-sector-type households would be working in the public sector and consuming c_w^g, while the public-sector types will be enjoying c_n^g. Alternatively, the workers would be participating in a sector-specific lottery with the proportions representing the probability of being selected for work. Conditional on the sector type, a household would receive the same income in expected terms.

Alternatively, we can introduce insurance markets, and allow households to buy insurance, which would allow them to equalize the actual income received, conditional on the sector-type. Given the difference in the wages and hours worked across sectors, segmented insurance markets are needed in order to provide actuarially fair insurance.

2.4 Insurance markets

Insurance markets is segmented, with one representative company per sector.[5] Insurance costs q^j per unit, $j = p, g$, and provides one unit of income if the household is not

[5] The insurance market segmentation is a direct effect of the discreteness of the labor supply in each sector and the sorting done by the sector-type shock.

working. We can think of insurance as bonds that pay out only in case the household is not chosen for work. Thus, household will also choose the quantity of insurance to purchase $b^j,\ j = p,g$. With sector types, the setup requires that the insurance market is segmented, with public sector insurance market insuring only public-sector-type households, and the private sector insurance market insuring only private-sector-type households.

Without segmentation, insurance will not be actuarially fair, one of the groups will face better odds versus price, the company will not be able to break even, and/or at least one type of households will not be able to buy full insurance, which would completely smooth consumption across employment states, given the non-convexity constraint of labor supply.

As pointed out in Hansen (1985), the plausibility of this insurance market segmentation result depends crucially on the fact that probabilities λ^p and λ^g are perfectly observable to everyone, and that the contracts written are perfectly enforceable. Also, who has won and who has lost the lottery is assumed to be perfect knowledge. Lastly, everyone will always announce truthfully the same λ^p (λ^g) to the private (public) firm and the private-sector (public-sector) insurance company.

2.4.1 Private-sector insurance company

The private-sector insurance company maximizes profit. The company only services private-sector types. It receives revenue if a private-sector-type household is working and makes payment if it is not. More specifically, the proportion of people working in the private sector contribute towards the unemployment benefits pool, which are then distributed of benefits to the unemployed in that sector. The amount of insurance sold in the private sector is a solution to the following problem: Taking $q^{p*}(i)$ as given, $b^{p*}(i)$ solves

$$\max_{b^p} \lambda^{p*}(i) q^{p*}(i) b^p - [1-\lambda^{p*}(i)] b^p.$$

With free entry profits are zero, hence

$$\lambda^{p*}(i) q^{p*}(i) b^p - [1-\lambda^{p*}(i)] b^p = 0.$$

This condition implicitly clears the insurance market for each individual in the private sector.

2.4.2 Public-sector insurance company

The public-sector insurance company also maximizes profit. The company only services public-sector types. It receives revenue if a public-sector-type household is working and makes payment if it is not. More specifically, the proportion of people working in the public sector contribute towards the unemployment benefits pool, which are then distributed of benefits to the unemployed in that sector. The amount of insurance sold in the public sector is a solution to the following problem: Taking $q^{g*}(j)$ as given,

$b^{g*}(j)$ solves

$$\max_{b^g} \lambda^{g*}(j)q^{g*}(j)b^g - [1-\lambda^{g*}(j)]b^g$$

With free entry profits of the insurance company operating in the public sector are also zero since

$$\lambda^{g*}(j)q^{g*}(j)b^g - [1-\lambda^{g*}(j)]b^g = 0.$$

This implicitly clears the insurance market for each individual of a public sector type.

In the next section, the equilibrium with lotteries and no insurance markets is presented and discussed first, and then the setup is extended to incorporate a regime with insurance.

3. Decentralized Competitive Equilibrium (DCE) with lotteries

3.1 Definition of the DCE with lotteries

A competitive equilibrium with lotteries in the private and public sector for this economy is a list:

$$\left(c_w^{p*}(i), c_n^{p*}(i), \lambda^{p*}(i)\right), \left(c_w^{g*}(j), c_n^{g*}(g), \lambda^{g*}(j)\right), h^{f*}, w^{p*}, w^{g*}, p^*, \pi^*$$

s.t.

(i) Private-sector consumer maximization – taking w^{p*}, p^{p*}, π^* as given, for each private-sector type household i, $c_w^{p*}(i)$, $c_n^{p*}(i)$, $\lambda^{p*}(i)$ solve:[6]

$$\max_{\lambda^p, c_w^p, c_n^p} \lambda^p(i)\left\{\ln[(c_w^p)^\eta + S^\eta]^{1/\eta} + \alpha\ln(1-\overline{h}^p)\right\} + \\ +(1-\lambda^p(i))\left\{\ln[(c_n^p)^\eta + S^\eta]^{1/\eta} + \alpha\ln(1)\right\}$$

$$\text{s.t.} \quad p^*[\lambda^p(i)c_w^p + (1-\lambda^p(i))c_n^p] = w^{p*}\overline{h}^p\lambda^p(i) + \pi^* - t, \\ c_w^p \geq 0,\ c_n^p \geq 0,\ 0 < \lambda^p(i) < 1.$$

(ii) Public-sector consumer maximization – taking w^{g*}, p^{p*}, π^* as given, for each public-sector type household j, $c_w^{g*}(j)$, $c_n^{g*}(j)$, $\lambda^{g*}(i)$ solve

$$\max_{\lambda^g, c_w^g, c_n^g} \lambda^g(j)\left\{\ln[(c_w^g)^\eta + S^\eta]^{1/\eta} + \alpha\ln(1-\overline{h}^g)\right\} + \\ +(1-\lambda^g(j))\left\{\ln[(c_n^g)^\eta + S^\eta]^{1/\eta} + \alpha\ln(1)\right\}$$

[6] Note that now when they trade lotteries the outcome is no longer deterministic. Now consumers maximize expected utility, i.e. if a private sector type is chosen to work with probability λ^p, that individual will get expected income $\lambda^p w^p \overline{h}^p$.

$$\begin{aligned} \text{s.t.} \quad & p^*[\lambda^g(j)c_w^g + (1-\lambda^g(j))c_n^g] = w^{g*}\overline{h}^g\lambda^g(j) + \pi^* - t, \\ & c_w^g \geq 0,\ c_n^g \geq 0,\ 0 < \lambda^g(j) < 1. \end{aligned}$$

(iii) Firm maximization – taking p^*, w^{p*} as given, h^{f*} solves

$$\begin{aligned} \max\nolimits_h \quad & p^* f(h) - w^{p*}h \\ \text{s.t.} \quad & h \geq 0, \end{aligned}$$

and

$$\pi^* = p^* f(h^{f*}) - w^{p*}h^{f*}.$$

(iv) Government – taking p^*, w^{p*}, and π^* as given, government provides public services according to the following production function

$$S = S(\lambda^{g*}\overline{h}^{g*})$$

The government sets $w^{g*} = \gamma w^{p*}$. Finally, T is residually set to ensure

$$w^{g*}\lambda^{g*}\overline{h}^g = T.$$

(v) Market clearing:

$$\begin{aligned} \int_i \lambda^{p*}(i)\overline{h}^p di &= h^{f*}, \\ \int_i \left[\lambda^{p*}(i)c_w^{p*}(i) + (1-\lambda^{p*}(i))c_n^{p*}(i)\right] di + & \\ + \int_j \left[\lambda^{g*}(j)c_w^{g*}(j) + (1-\lambda^{g*}(j))c_n^{g*}(j)\right] dj &= f(h^{f*}). \end{aligned}$$

3.2 Characterization of the DCE with lotteries

Private-sector types problem is:

$$\begin{aligned} L = \lambda^p(i)\Big\{\ln[(c_w^p)^\eta + S^\eta]^{1/\eta} + \alpha\ln(1-\overline{h}^p)\Big\} + (1-\lambda^p(i))\Big\{\ln[(c_n^p)^\eta + S^\eta]^{1/\eta}\Big\} \\ -\mu\Big\{p^*\lambda^p(i)c_w^p + p^*(1-\lambda^p(i))c_n^p - w^{p*}\overline{h}^p\lambda^p(i) - \pi^* + t\Big\} \end{aligned}$$

FOCs:

$$c_w^p: \qquad \lambda^p(i)\frac{1}{[(c_w^p)^\eta + S^\eta]}(c_w^p)^{\eta-1} = \mu p^* \lambda^p(i) \tag{3.1}$$

$$c_n^p: \qquad (1-\lambda^p(i))\frac{1}{[(c_n^p)^\eta + S^\eta]}(c_n^p)^{\eta-1} = \mu p^*(1-\lambda^p(i)) \tag{3.2}$$

$$\begin{aligned}\lambda^p(i): \qquad & \left\{\ln[(c_w^p)^\eta + S^\eta]^{1/\eta} + \alpha\ln(1-\overline{h}^p)\right\} - \left\{\ln[(c_n^p)^\eta + S^\eta]^{1/\eta}\right\} \\ & -\mu\left\{p^* c_w^p - p^* c_n^p - w^{p*}\overline{h}^p\right\} = 0 \end{aligned} \tag{3.3}$$

(3.1) and (3.2) show that $c_w^p = c_n^p$, $\forall i$. Also, $\lambda^p(i) = \lambda^p$, $\forall i$. Then (3.3) simplifies to

$$\alpha\ln(1-\overline{h}^p) = -\mu w^{p*}\overline{h}^p.$$

Hence,

$$w^{p*} = f'(\lambda^{p*}\overline{h}^{p*}) = \frac{\alpha\ln(1-\overline{h}^p)[(c_w^p)^\eta + S^\eta]}{(c_w^p)^{\eta-1}\overline{h}^p}.$$

This equation is a discrete version of the marginal product of labor equals the marginal rate of substitution. It implicitly characterizes the optimal λ^p.

Public-sector types problem:

$$\begin{aligned} L = \lambda^g(j)&\left\{\ln[(c_w^g)^\eta + S^\eta]^{1/\eta} + \alpha\ln(1-\overline{h}^g)\right\} + (1-\lambda^g(j))\left\{\ln[(c_n^g)^\eta + S^\eta]^{1/\eta}\right\} \\ & -\nu\left\{p^*\lambda^g(j)c_w^g + p^*(1-\lambda^g(j))c_n^g - w^{g*}\overline{h}^g\lambda^g(j) - \pi^* + t\right\} \end{aligned}$$

FOCs:

$$c_w^g: \qquad \lambda^g(j)\frac{1}{[(c_w^g)^\eta + S^\eta]}(c_w^g)^{\eta-1} = \nu p^* \lambda^g(j) \tag{3.4}$$

$$c_n^g: \qquad (1-\lambda^g(j))\frac{1}{[(c_n^g)^\eta + S^\eta]}(c_n^g)^{\eta-1} = \nu p^*(1-\lambda^g(j)) \tag{3.5}$$

$$\begin{aligned}\lambda^g(j): \qquad & \left\{\ln[(c_w^g)^\eta + S^\eta]^{1/\eta} + \alpha\ln(1-\overline{h}^g)\right\} - \left\{\ln[(c_n^g)^\eta + S^\eta]^{1/\eta}\right\} \\ & -\nu\left\{p^* c_w^g - p^* c_n^g - w^{g*}\overline{h}^g\right\} = 0 \end{aligned} \tag{3.6}$$

(3.4) and (3.5) show that $c_w^g = c_n^g$, $\forall j$. Also, $\lambda^g(j) = \lambda^g$, $\forall j$. Then (3.6) simplifies to

$$\alpha\ln(1-\overline{h}^g) = -\nu w^{p*}\overline{h}^p.$$

Hence,

$$w^{g*} = \gamma f'(\lambda^{p*}\overline{h}^{p*}) = \frac{\alpha \ln(1-\overline{h}^{g})[(c_w^g)^\eta + S^\eta]}{(c_w^g)^{\eta-1}\overline{h}^{g}}.$$

This equation is a also the discrete version of the marginal product of labor equals the marginal rate of substitution. In this case it implicitly characterizes the optimal λ^g. Note that it is optimal for the benevolent government point of view to choose randomly λ^p, λ^g and to introduce uncertainty. With randomization, choice sets are convexified, and thus market completeness is achieved.

Since a household of either type can be chosen to work with some probability, the households are exposed to risk. Hence it would be optimal to have insurance. The government can then sell employment lotteries, and individuals will buy insurance to cover the risk of being unemployed (not being chosen for work). With insurance, however, the employer pays wage to individuals only if they work. That is, instead of working with expected income, we will work with actual income. This allows to extend the commodity space in the model framework and include insurance markets.

4. Decentralized Competitive Equilibrium with lotteries and insurance markets

4.1 Definition of the DCE with insurance markets

A competitive equilibrium with lotteries and unemployment insurance is a list

$$\big(c_w^{p*}(i), c_n^{p*}(i), \lambda^{p*}(i), b^{p*}(i)\big), \big(c_w^{g*}(j), c_n^{g*}(g), \lambda^{g*}(j), b^{p*}(j)\big), h^{f*}, w^{p*}, w^{g*}, p^*,$$
$$q^{p*}, q^{g*}, \pi^*$$

s.t.

(i) Private-sector-type household maximization – taking w^{p*}, p^{p*}, π^* as given, for each private-sector type household i, $c_w^{p*}(i), c_n^{p*}(i), \lambda^{p*}(i), b^*(i)$ solve

$$\max_{\lambda^p, c_w^p, c_n^p} \lambda^p(i)\Big\{\ln[(c_w^p)^\eta + S^\eta]^{1/\eta} + \alpha \ln(1-\overline{h}^p)\Big\} +$$
$$+(1-\lambda^p(i))\Big\{\ln[(c_n^p)^\eta + S^\eta]^{1/\eta} + \alpha\ln(1)\Big\}$$

$$\text{s.t.} \qquad \begin{aligned} & p^* c_w^p + b^p q^{p*}(i) = w^{p*}\overline{h}^p + \pi^*, \\ & p^* c_n^p = b^p + \pi^*, \\ & c_w^p \geq 0,\ c_n^p \geq 0,\ 0 < \lambda < 1, \end{aligned}$$

or

$$p^* c_w^p + p^* q^{p*} c_n^p = w^{p*}\overline{h}^p + (1+\pi^*) q^{p*}.$$

Foe each household in the private sector, there are two states: a household is buying unemployment insurance when working, receiving a payout when not working, hence in equilibrium $b^{p*} = \lambda^{p*} w^{p*} \overline{h}^{p*}$.

(ii) Public-sector-type household maximization – taking w^{g*}, p^{p*}, π^* as given, for each public-sector-type household j, $c_w^{g*}(j), c_n^{g*}(j), \lambda^{g*}(j), b^{g*}(j)$ solve

$$\max_{\lambda^g, c_w^g, c_n^g} \lambda^g(j)\Big\{\ln[(c_w^g)^\eta + S^\eta]^{1/\eta} + \alpha \ln(1-\overline{h}^g)\Big\} + \\ +(1-\lambda^g(j))\Big\{\ln[(c_n^g)^\eta + S^\eta]^{1/\eta} + \alpha \ln(1)\Big\}$$

$$\begin{aligned} \text{s.t.} \qquad & p^* c_w^g + b^g q^{g*}(i) = w^{g*}\overline{h}^g + \pi^*, \\ & p^* c_n^g = b^g + \pi^*, \\ & c_w^g \geq 0,\ c_n^g \geq 0,\ 0 < \lambda^g < 1, \end{aligned}$$

or

$$p^* c_w^p + p^* q^{p*} c_n^p = w^{p*}\overline{h}^p + (1+\pi^*) q^{p*}.$$

Foe each household in the public sector, there are two states: a household is buying unemployment insurance when working, receiving a payout when not working, hence in equilibrium $b^{g*} = \lambda^{g*} w^{g*} \overline{h}^{g*}$.

(iii) Firm maximization – taking p^*, w^* as given, h^{f*} solves

$$\begin{aligned} \max_h \qquad & p^* f(h) - w^{p*} h, \\ \text{s.t.} \qquad & h \geq 0, \end{aligned}$$

and

$$\pi^* = p^* f(h^{f*}) - w^{p*} h^{f*}.$$

(iv) Insurance companies. Insurance markets is segmented, with one company per sector.

(a) Private sector. Taking $q^{p*}(i)$ as given, $b^{p*}(i)$ solves

$$\max_{b^p} \lambda^{p*}(i) q^{p*}(i) b^p - (1-\lambda^{p*}) b^p$$

With free entry profits are zero, hence

$$\lambda^{p*}(i) q^{p*}(i) b^p - (1-\lambda^{p*}(i)) b^p = 0.$$

This implicitly clears the insurance market for each individual in the private sector.

(b) Public sector. Taking $q^{g*}(j)$ as given, $b^{g*}(j)$ solves

$$\max_{b^g} \lambda^{g*}(j)q^{g*}(j)b^g - (1-\lambda^{g*}(j))b^g$$

With free entry profits of the insurance company operating in the public sector are also zero since

$$\lambda^{g*}(j)q^{g*}(j)b^g - (1-\lambda^{g*}(j))b^g = 0.$$

This implicitly clears the insurance market for each individual of a public sector type.

(v) Government – taking p^*, w^{p*}, and π^* as given, government provides public services according to $S = S(\lambda^{g*}\overline{h}^{g*})$. The government sets $w^{g*} = \gamma w^{p*}$, and taxes T are residually set to ensure

$$w^{g*}\lambda^{g*}\overline{h}^{g} = T.$$

(vi) Market clearing.

$$\int_i \lambda^{p*}(i)\overline{h}^p di = h^{f*},$$

$$\int_i [\lambda^{p*}(i)c_w^{p*}(i) + (1-\lambda^{p*}(i))c_n^{p*}(i)]di + \int_j [\lambda^{g*}(j)c_w^{g*}(j) + (1-\lambda^{g*}(j))c_n^{g*}(j)]dj = f(h^{f*}).$$

4.2 Characterization of the DCE with insurance markets

Private sector consumer problem:

$$\begin{aligned}\max_{\lambda^p, c_w^p, c_n^p, b^p(i)} \quad & \lambda^p(i)\Big\{\ln[(c_w^p)^\eta + S^\eta]^{1/\eta} + \alpha\ln(1-\overline{h}^p)\Big\} + \\ & +(1-\lambda^p(i))\Big\{\ln[(c_n^p)^\eta + S^\eta]^{1/\eta} + \alpha\ln(1)\Big\} \\ \text{s.t.} \quad & p^*c_w^p + p^*q^{p*}c_n^p = w^{p*}\overline{h}^p + \pi^* + q^{p*}\pi^*.\end{aligned}$$

Normalize $p^* = 1$.

$$\begin{aligned} c_w^p: \quad & \lambda^p \frac{(c_w^p)^{\eta-1}}{[(c_w^p)^\eta + S^\eta]} = p\mu \\ c_n^p: \quad & (1-\lambda^p)\frac{(c_n^p)^{\eta-1}}{[(c_n^p)^\eta + S^\eta]} = pq^p\mu \end{aligned}$$

Optimal λ^p ($\lambda^p(i) = \lambda^p$, $\forall j$) is implicitly characterized by the zero-profit condition from the private sector insurance company:

$$\frac{\lambda^p}{1-\lambda^p} = \frac{1}{q^p} \tag{4.1}$$

The price of insurance depends on probability of the event you are insuring against. We cannot force $q^{p*}(i) = q^{p*}$ although ex post that would indeed be the case. For the insurance firms, profits are linear in q^p. This implies that profits cannot be positive or negative in equilibrium. Zero profits in the private sector insurance market then mean $q^p = \frac{1-\lambda^p}{\lambda^p}$. A common interpretation for both insurance companies is that this price of the insurance is the odds ratio, or the ratio of probabilities of the two events.

Combining then with the FOCs for state-contingent consumption, we obtain that $c_w^p = c_n^p$, $\forall i$. That is, private-sector-type households buy full insurance to smooth consumption perfectly.

Similarly, for the public sector consumers:

$$
\begin{aligned}
\max_{\lambda^g, c_w^g, c_n^g}, b^g(j)\lambda^g(j)\Big\{ \ln[(c_w^g)^\eta + S^\eta]^{1/\eta} + \alpha \ln(1-\overline{h}^g) \Big\} + \\
+(1-\lambda^g(j))\Big\{ \ln[(c_n^g)^\eta + S^\eta]^{1/\eta} + \alpha \ln(1) \Big\} \\
\text{s.t.} \qquad c_w^g + q^{g*} c_n^g = w^{g*}\overline{h}^g + \pi^* + q^{g*}\pi^*.
\end{aligned}
$$

$$
\begin{aligned}
c_w^g : \qquad & \lambda^g \frac{(c_w^g)^{\eta-1}}{[(c_w^g)^\eta + S^\eta]} = p\nu \\
c_n^g : \qquad & (1-\lambda^g) \frac{(c_n^g)^{\eta-1}}{[(c_n^g)^\eta + S^\eta]} = pq^g\nu
\end{aligned}
$$

Optimal λ^g ($\lambda^g(j) = \lambda^g$, $\forall j$) is implicitly characterized by the zero-profit condition from the public sector insurance company:

$$
\frac{\lambda^g}{1-\lambda^g} = \frac{1}{q^g}
$$

Combining then with the FOCs for state-contingent consumption, we obtain that $c_w^g = c_n^g$, $\forall j$. Also, $\lambda^g(j) = \lambda^g$, $\forall j$. That is, public-sector-type households buy also full insurance to equalize consumption in the two states (employed *vs.* unemployed). In particular, when income is stochastic, i.e., it is uncertain whether the individual will be employed, we need insurance markets for each sector type. In this economy there is no uncertainty (after the types are revealed) but it is optimal to introduce insurance markets. This is because of the non-convexity of the choice set, which is similar to having incomplete markets. Lotteries can then be introduced to achieve market completeness. Therefore, randomization may be optimal in a non-convex environment even though there is no aggregate uncertainty.

5. Conclusions

This note describes the lottery and insurance market equilibrium in an economy with both private and public sector jobs and non-convex labor supply decision faced by the workers. In addition, when households are constrained to search for jobs only in a certain sector, the framework requires that there should be separate insurance markets: public and private sector one, which would pool the risk of the corresponding group of workers. In equilibrium, conditional on the sector-type, each household would fully insure against the uncertainty in terms of the employment status (but cannot insure against the “type” shock). The unemployment insurance market segmentation is a direct effect of the discreteness of the labor supply in each sector and the sorting done by the sector-type shock.

The plausibility of the result derived in the paper depends crucially on the fact that probabilities λ^p and λ^g are perfectly observable to everyone, and that the contracts written are perfectly enforceable. Also, who has won and who has lost the lottery is assumed to be perfect knowledge. Lastly, everyone will always announce truthfully the same λ^p (λ^g) to the private (public) firm and the private-sector (public-sector) insurance company. Therefore, whether and how this insurance-market segmentation can be implemented in reality is not entirely clear at this point.

Acknowledgement The author would like to thank the Editor and an anonymous referee for the valuable comments.

References

Cooley, T. and Prescott, E. (1995). Economic Growth and Business Cycles. In Cooley, T. (ed.), *Frontiers of Business Cycle Research*. Princeton NJ, Princeton University Press.

Gomes, P. (2014). Optimal Public Sector Wages. *The Economic Journal*, 125, 1425–1451

Hansen, G. (1985). Indivisible Labor and the Business Cycle. *Journal of Monetary Economics*, 16, 309–328.

Kydland, F. (1995). Business Cycles and Aggregate Labor Market Fluctuations. In Cooley, T. (ed.), *Frontiers of Business Cycle Research*. Princeton NJ, Princeton University Press.

Rogerson, R. (1988). Indivisible Lotteries, Lotteries and Equilibrium. *Journal of Monetary Economics*, 21, 3–16.

Vasilev, A. (2015a). Aggregation with a Double Non-Convex Labor Supply Decision: Indivisible Private- and Public-Sector Hours. Submitted to the *Journal of Economic Development*, under review.

Vasilev, A. (2015b). Macroeconomic Effects of Public-Sector Unions. *LABOUR*, 29(2), 101–126.

Czech Economic Review 9 (2015) 104–134
Acta Universitatis Carolinae Oeconomica

Sentiment Cyclicality

Orlando Gomes*

Received 2 April 2015; Accepted 22 February 2016

Abstract The paper investigates the dynamics of a model of sentiment switching. The model is built upon rumor propagation theory and it is designed to uncover, for a given population, the social process through which optimistic individuals might become pessimistic or the other way around. The outcome is a scenario of perpetual motion with the shares of optimistic and pessimistic agents varying persistently over time. On a second stage, the cyclical sentiments setup is attached to a mechanism of formation of expectations based on the notion of optimized rationality, leading to a description of the macro economy in which aggregate output and inflation exhibit sentiment driven fluctuations. The proposed model contributes to a recent strand of macroeconomic literature that recovers the Keynesian notions of animal spirits, market sentiments and waves of optimism and pessimism.

Keywords sentiments, animal spirits, business cycles, rumor propagation, New-Keynesian macroeconomics, optimized rationality
JEL classification E32, E37, D83, D84

1. Introduction

This paper merges two strands of scientific literature with the objective of offering a behavioral interpretation about observed aggregate business fluctuations. The here relevant lines of thought include, on one hand, the recent contributions on the macroeconomic role of animal spirits and, on the other hand, rumor spreading theory, a widely debated theme in various disciplinary fields and a theme that can be easily adapted to a setting of sentiment propagation.

The first part of the paper describes the dynamics of sentiment switching. The sources of sentiment changes reside exclusively on social interaction and, therefore, this is a model of pure animal spirits where features outside the scope of the economy (as confidence, fairness, antisocial behavior, and other behavioral elements mentioned in detail in Akerlof and Shiller 2009) determine the mood with which agents face economic decisions. In a second stage, sentiment dynamics are integrated into a baseline macro model. Particularly, the New-Keynesian macroeconomic framework, that involves a dynamic IS equation, the New-Keynesian Phillips curve and a monetary policy Taylor rule, is used to this end. In the proposed setting, waves of optimism and pessimism will determine economic outcomes by exacerbating the fluctuations caused by random shocks and by contributing to generate well defined periods of expansion

* Lisbon Accounting and Business School (ISCAL/IPL) & Business Research Unit (UNIDE/ISCTE-IUL), Av. Miguel Bombarda 20, 1069-035 Lisbon, Portugal. E-mail: omgomes@iscal.ipl.pt.

and recession in the economy.

A relevant feature of the setup is that adding sentiment changes to the model of the aggregate economy will require a departure from pure rational expectations and the adoption of an optimized rationality procedure, according to which agents have to weigh whether the effort and the cost of collecting information to generate an accurate forecast is compensated by the benefits of producing such correct forecast. In the presence of information acquisition costs, rational expectations are replaced by a heuristic rule under which agents believe macroeconomic variables will approach the defined policy target (if they are optimistic) or depart from such target (if they are pessimistic).

The results and the discussion in this paper are aligned with what the latest and most influential research in macroeconomics suggests. The reader will be able to identify a strong coincidence between the main ideas that will be put forward and the arguments advanced, for instance, in the conclusion of the paper by Angeletos et al. (2015, p. 25):

> "By relying on a particular solution concept together with complete information, standard macroeconomic models impose a rigid structure on how agents form beliefs about endogenous economic outcomes and how they coordinate their actions. In this paper, by contrast, (...) we augmented DSGE models with a tractable form of higher-order belief dynamics that (...) captures a certain kind of waves of optimism and pessimism about the short-term outlook of the economy. We believe that this adds to our understanding of business-cycle phenomena (...)."

And they continue (Angeletos et al. 2015, p. 26):

> "These findings naturally raise the question of where the drop in confidence during a recession, or more generally the waves of optimism and pessimism in the agents' beliefs about one another's actions come from. Having treated the 'confidence shock' as exogenous, we can not offer a meaningful answer to this question. This limitation, however, is not specific to what we do in this paper: any formal model must ultimately attribute the business cycle to some exogenous trigger, whether this is a technology shock, a discount-rate shock, a financial shock, or even a sunspot."

While agreeing with the first statement, that systematic shocks on sentiments help in understanding business cycles, the analysis in this paper goes deeper in the sense that instead of treating sentiment fluctuations as completely exogenous and inexpugnable, it offers an explanation for waves of animal spirits that is supported on social interaction across agents holding different 'views of the world'.

The remainder of the paper is organized as follows. Section 2 undertakes a brief tour across the relevant literature. Section 3 adapts the rumor spreading framework to allow for the possibility of changing sentiments. In Section 4, it is shown how a slight and reasonable change in the proposed setup can result into everlasting oscillations in the shares of optimists and pessimists. Section 5 introduces the optimized rationality concept and describes how optimistic and pessimistic agents form expectations. Section 6 applies the previously presented expectation rules to a New-Keynesian benchmark model, revealing that the setup is adequate to explain business-cycle persistence. Finally, Section 7 concludes.

2. The literature: a brief tour

2.1 Animal spirits and contemporaneous macroeconomic thought

Mainstream economic theory bases most of its analysis on a strict notion of rationality. Because agents are, allegedly, capable of optimally processing available information, aggregate fluctuations are interpreted as the mere outcome of the response of utility maximizing agents to supply side shocks. This interpretation on the sources of business cycles is not unanimous; in fact, in the last few years, an increasing number of macroeconomists began exploring different routes. The turning point can be traced back to Kocherlakota (2010) who, with an insightful reflection about the state of macroeconomics, was able to convince the scientific community that the frequency and depth of observed business fluctuations cannot be explained solely on the basis of exogenous shocks on technology, preferences or policy. Surely, other drivers of aggregate cyclical motion exist.

The quest for such drivers has led macroeconomists to recover and focus attention on some Keynesian notions and ideas, namely those attached to market sentiments, animal spirits and other psychological factors that shape the decision-making process of economic agents. The issue is not whether these notions are relevant to characterize human behavior (they certainly are!), but how one can integrate them in the benchmark macro models without losing the relevant contribution that the dynamic stochastic general equilibrium framework currently gives for the understanding of the functioning of the aggregate economy.

Meaningful studies going on the direction mentioned in the above paragraphs include De Grauwe (2011, 2012), Milani (2011), Bidder and Smith (2012), Franke (2012), Angeletos and La'O (2013), Bofinger et al. (2013) and Lengnick and Wohltmann (2013). Although the adopted approaches differ, the cited references all share a desire to incorporate a behavioral component into the macro theory of short-run fluctuations.

In De Grauwe (2011, 2012), Bofinger et al. (2013) and Lengnick and Wohltmann (2013), it is considered that economic agents use simple rules, called heuristics, in order to predict future values of relevant macro variables, in the context of the New-Keynesian macro model. Combining these heuristics with an evolutionary approach that contemplates a discrete choice selection mechanism, this class of models triggers the generation of endogenous waves of optimism and pessimism that allow to replicate with a reasonable degree of precision observable business cycles. Two points about this approach are worth stressing. First, animal spirits are viewed as a way to guarantee the existence of a true decentralization in market decisions; under rational expectations, agents are identical, endowed with unlimited cognitive capabilities and, therefore, there must be a coincidence between decentralized decisions and the choices of a representative agent. Animal spirits open the door to behavior heterogeneity and to a richer set of potential outcomes. Second, as emphasized by Paul DeGrauwe, one must be careful about the way in which departures relatively to full rationality are introduced into macro models; it is necessary to avoid that everything becomes possible, as the result of assumptions that are, eventually, unreasonable and hard to reconcile with a rigorous scientific analysis. In this specific context, it is claimed that the mentioned problem is solved once the evolutionary learning process is attached to the model.

The strategy followed by Milani (2011) is different. The New-Keynesian model is, again, used to discuss departures from full rationality and from the formation of purely rational expectations, however the approach is now based on an explicit learning device. Agents directly exploit historical series with the goal of understanding the true law of motion of the relevant economic indicators. As they collect information they will learn, but this learning process might not be immaculate, in the sense that it might not lead to a convergence to a rational expectations equilibrium. Instead, waves of optimism and pessimism might subsist over time.

In Bidder and Smith (2012), the motivation is the same, i.e., to highlight the importance of animal spirits in the analysis of macroeconomic phenomena, but the psychological driver of the departure from strict rationality differs from the previously mentioned. Specifically, the sentiment waves are the outcome of a peculiar behavioral aspect which is the fear of model misspecification. Agents have knowledge on the true model of the economy but they are concerned with the possibility of such model being distorted in some way, leading to an overly pessimistic interpretation of the reality. In Franke (2012), rational expectations are, once more, set aside, in this case in favor of a forecasting mechanism based on the use of an average opinion index built upon the revealed sentiments of the population. This study furnishes a micro foundation for the presence of animal spirits in the macro economy, which incorporates a herding component, and emphasizes the cyclical nature of the trajectories followed by the macro variables in the specified scenario.

Finally, Angeletos and La'O (2013) also propose a business cycle theory, constructed in turn of the notions of animal spirits and market sentiments. These authors, however, intentionally preserve rational expectations. They introduce a communication constraint by assuming that trade is random and decentralized. This is all that is required to generate waves of optimism and pessimism in a model that, otherwise, is of a neoclassical nature. Trading frictions that limit communication are the key element, in this view, underlying the formation of animal spirits. This study accommodates market sentiments and self-fulfilling beliefs in macro theory without abandoning rational expectations, competitive markets and equilibrium uniqueness.

2.2 From rumors to sentiments

Sentiments of optimism and pessimism arise in the mind of the individuals most of the times as the outcome of a social interaction process, in which positive and negative feelings are shared across a given population. Therefore, it seems reasonable to associate sentiment switching processes to the literature on rumor spreading, namely the part of this literature that characterizes the propagation of rumors in a similar way relatively to the modeling of infectious diseases. Such contributions go back to Daley and Kendall (1964, 1965) and Maki and Thompson (1973), who have made the first relevant characterization of a rumor spreading mechanism. In the last few years a clear resurgence of this theme emerged, with meaningful extensions of the basic model being presented.

The benchmark rumor propagation model classifies individual agents into three categories: susceptible, spreaders and stiflers. Susceptible individuals are those who ignore the rumor but may be 'infected' when entering in contact with someone who knows the rumor. Spreaders are the ones that have acquired knowledge on the rumor and transmit it to others. And stiflers are the individuals who know the rumor, have spread it in the past, but no longer propagate it (see, e.g., Cintron-Arias, 2006, for a basic version of the model).

The rumor setup has evolves essentially in two directions. A first group of authors, including Thompson et al. (2003), Huo et al. (2012), Zhao et al. (2012) and Wang et al. (2013), have introduced changes on the typology of agents participating in the rumor spreading process, namely including passive and active individuals (who differ in their propensity to contact others), and on the nature of the relations, through the consideration of trust mechanisms, forgetting and remembering processes and incubation periods.

A second group of changes over the original model relates to the structure of interaction. In the original framework, the topology of the underlying social interaction network across which the rumor spreads is overlooked, i.e., it is implicitly assumed that we are in the presence of a homogeneously mixing population: anyone can interact with another agent and these meetings occur randomly. Pastor-Santorras and Vespignani (2004) and Nekovee et al. (2007) approach the rumor spreading problem in scenarios of complex social networks; specifically, they analyze rumor propagation in the following types of networks: random graphs, uncorrelated scale-free networks and scale-free networks with assortative degree correlations. Similarly, Zanette (2002) applies rumor propagation to a specific network topology, namely small-world networks, which are social networks that are highly clustered and for which the distance between any two nodes is on average very small as compared to the total number of nodes and links.

First built with the purpose of characterizing a simple process of rumor spreading in a homogeneously mixing population, the rumor propagation model has been, as described, sophisticated in various directions that, basically, have added new types of agents and have alerted to the need of exploring more complex interaction scenarios. The framework is also useful, as we shall see, to approach sentiment switching.

In our specific setting, in which agents are exposed to sentiments of optimism and pessimism, an optimistic agent may be susceptible to turn into a pessimist if she enters in contact with an agent in the other category. In that case, she eventually becomes a spreader of the pessimistic feeling and, after a given period of time, she is likely to turn into a stifler. Pessimistic stiflers will then, eventually, become susceptible of turning optimistic again, and this process will tend to repeat itself endlessly. As a result, the proposed model of social interaction implies a circular flow on the motion of the shares of optimistic and pessimistic individuals, for the assumed population. Under reasonable and logical conditions, this flow of individuals from one group to the other may be such that the number of optimists and pessimists does not remain constant over time. In fact, we will show that it might fluctuate endlessly, following a cyclical movement.

3. Sentiment switching

3.1 The rumor propagation framework

Consider a discrete notion of time, $t = 0, 1, \ldots$, and a population of individuals organized under the form of a homogeneous social interaction network. This network is composed by nodes and by links connecting the nodes. It is assumed that each node j has an identical number of k links to other nodes; for simplicity, we will normalize the value of k to 1. In the context of rumor propagation, each node j in the network corresponds to an individual that may belong, at date t, to one of three categories: susceptible or ignorants, spreaders and stiflers; the respective shares are x_t, y_t and z_t.

In Nekovee et al. (2007), it is demonstrated how interacting Markov chains and the law of mass action can be used to represent the dynamics of the interaction process. In a $k = 1$ homogeneous network, in which a meeting between an ignorant and a spreader triggers a transition of the ignorant to the spreader state with probability $\lambda \in (0, 1]$, and a meeting between a spreader and another spreader or a stifler implies a transition of the first to the stifler state with probability $\sigma \in (0, 1]$, such dynamics are presentable under the form of a system of three difference equations:

$$\begin{cases} x_{t+1} - x_t = -\lambda x_t y_t \\ y_{t+1} - y_t = \lambda x_t y_t - \sigma y_t (y_t + z_t) \\ z_{t+1} - z_t = \sigma y_t (y_t + z_t) \end{cases} \tag{1}$$

The 3-dimensional system (1) can be displayed in a compact 2-D form, given that $x_t + y_t + z_t = 1$. Selecting variables x_t and y_t as the endogenous variables of the new system, it comes:

$$\begin{cases} x_{t+1} - x_t = -\lambda x_t y_t \\ y_{t+1} - y_t = [(\lambda + \sigma) - \sigma x_t] y_t \end{cases} \tag{2}$$

Despite its apparent simplicity, system (2) encloses an intricate dynamic behavior. The single substantive feature one draws from the respective analysis is that the number of spreaders falls to zero as time goes to infinity. The steady-state distribution of individuals across the categories of ignorants and stiflers is not determinable in generic form, because such allocation will be dependent on the initial state (x_0, y_0, z_0). When linearizing system (2) in the vicinity of a hypothetical steady-state point, one observes that one of the eigenvalues of the respective Jacobian matrix is equal to 1, and therefore the system rests over a bifurcation line, delivering an unconventional transitional dynamics outcome.

Rumor spreading studies tend to distinguish between two kinds of equilibria (see,

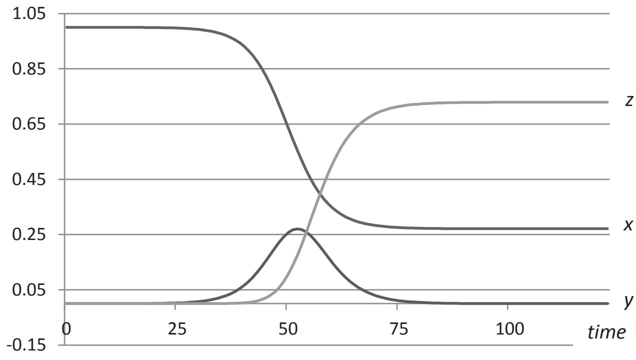


Figure 1. Time trajectories of the shares of susceptible, spreader and stifler individuals ($\lambda = 0.25$, $\sigma = 0.33$)

e.g., Huo et al. 2012). The rumor-free equilibrium corresponds to the case where $(x^*, y^*, z^*) = (1, 0, 0)$. This occurs, under the presented specification, only for $\lambda = 0$, i.e., when the rate of rumor spreading is zero. All other possible steady-state results can be designated rumor-endemic equilibria; these results are such that $(x^*, y^*, z^*) = (x^*, 0, 1 - x^*)$, $\forall \lambda, \sigma \in (0, 1)$.[1]

Figure 1 displays the typical trajectories of x_t, y_t and z_t. The figure is drawn for $\lambda = 0.25$ and $\sigma = 1/3$. At the starting date, almost all individuals are ignorant about the rumor; a single spreader is necessary to begin the rumor's dissemination. As time unfolds, the share of susceptible ignorants falls, the number of spreaders increases and some spreaders start changing to the stifler position. After a given threshold, the share of spreaders begins to fall, as the passage of spreaders to stiflers turns stronger than the transition from the susceptible state to the spreader state. In the long-term, all spreaders switch to stiflers and the population will be grouped into two classes: those who never heard the rumor, x^*, and those that know the rumor, have spread it in the past but no longer disseminate it, z^*.

The rumor propagation model, as described above, is sufficiently flexible to be adapted in a multiplicity of directions. In what follows, the model is modified and transformed in a framework where waves of optimism and pessimism may be subject to discussion.

3.2 The sentiment propagation framework

In this subsection, the above rumor propagation apparatus is adapted to deal with sentiments that might influence the aggregate outcome of economic relations. Only two types of sentiments are allowed for, namely optimism and pessimism. The share of optimistic agents will be denoted by ω_t; the share of pessimists is $1 - \omega_t$. Now, six

[1] See Piqueira (2010) for further insights on the study of the transitional dynamics of the benchmark rumor propagation model.

categories of agents will populate the economy:

(i) Optimists susceptible of being 'infected' with a negative feeling: x_t^{ω}.

(ii) Spreaders of negative feelings: $y_t^{1-\omega}$. These are previous optimists, who were 'infected' with a negative sentiment and start spreading it.

(iii) Stiflers 'infected' with a negative sentiment: $z_t^{1-\omega}$. This part of the population is composed by previous optimists, that became spreaders of a negative feeling and that continue to be pessimists after they stop spreading the rumor underlying such sentiment.

(iv) Pessimists susceptible of being 'infected' with a positive feeling: $x_t^{1-\omega}$.

(v) Spreaders of positive sentiments: y_t^{ω}.

(vi) Stiflers who are optimists: z_t^{ω}.

All the shares presented in the previous list respect to percentages of the whole population and, therefore, $x_t^{\omega}+y_t^{\omega}+z_t^{\omega}+x_t^{1-\omega}+y_t^{1-\omega}+z_t^{1-\omega}=1$. From the stated arguments, it also follows that $\omega_t \equiv x_t^{\omega}+y_t^{\omega}+z_t^{\omega}$ and $1-\omega_t \equiv x_t^{1-\omega}+y_t^{1-\omega}+z_t^{1-\omega}$.

A model similar to the plain ignorant-spreader-stifler paradigm of the last subsection can be adapted to this new setting. The main difference is that now we have a closed circuit, where two types of states are achievable: at each time moment, agents can only be one of two things: optimists or pessimists. The implementation of the idea of a closed circuit requires one further assumption: stiflers (both optimists and pessimists) become susceptible of being infected with the opposite feeling, with a probability $\theta \in (0,1)$. The relevant system of difference equations is now a 6-dimensional system, although one of the dimensions can be suppressed because the sum of the endogenous variables is equal to 1. The list of equations is:

$$\left\{\begin{array}{l}
x_{t+1}^{\omega}-x_t^{\omega}=-\lambda_{\omega}x_t^{\omega}y_t^{1-\omega}+\theta_{\omega}z_t^{\omega}\\
y_{t+1}^{1-\omega}-y_t^{1-\omega}=\lambda_{\omega}x_t^{\omega}y_t^{1-\omega}-\sigma_{\omega}y_t^{1-\omega}\left(y_t^{1-\omega}+z_t^{1-\omega}\right)\\
z_{t+1}^{1-\omega}-z_t^{1-\omega}=\sigma_{\omega}y_t^{1-\omega}\left(y_t^{1-\omega}+z_t^{1-\omega}\right)-\theta_{1-\omega}z_t^{1-\omega}\\
x_{t+1}^{1-\omega}-x_t^{1-\omega}=-\lambda_{1-\omega}x_t^{1-\omega}y_t^{\omega}+\theta_{1-\omega}z_t^{1-\omega}\\
y_{t+1}^{\omega}-y_t^{\omega}=\lambda_{1-\omega}x_t^{1-\omega}y_t^{\omega}-\sigma_{1-\omega}y_t^{\omega}\left(y_t^{\omega}+z_t^{\omega}\right)\\
z_{t+1}^{\omega}-z_t^{\omega}=\sigma_{1-\omega}y_t^{\omega}\left(y_t^{\omega}+z_t^{\omega}\right)-\theta_{\omega}z_t^{\omega}
\end{array}\right. \tag{3}$$

In system (3), we have allowed for the possibility of different rates λ, σ and θ for the spreading of each of the two types of sentiments. As we will see below, considering that they are identical simplifies the analysis of the steady-state results. In order to maintain the analysis at a general level, for now we assume that they might differ.

Steady-state properties of (3) significantly diverge from what one has characterized concerning (1). In the current case, steady-state results are independent of the initial state and might, under particular conditions, be explicitly presented. Furthermore, the steady-state values, including the shares of spreaders, are all non-zero values, meaning that we have a dynamic steady-state: there will always be, at each period t, a portion of agents who spread negative sentiments and a portion of agents who spread positive sentiments. This is the direct outcome of our closed circuit assumption, that makes optimists converted to pessimists to become susceptible of being again 'infected' with an optimistic sentiment.

Let $\mathbf{v}_{t+1} = V(\mathbf{v}_t)$, with $\mathbf{v} = (x^{\omega}, y^{\omega}, z^{\omega}, x^{1-\omega}, y^{1-\omega} z^{1-\omega})$, be a compact representation of the system of difference equations (3) and define $E = \{\mathbf{v}^* : \mathbf{v}^* - V(\mathbf{v}^*) = 0\}$ as the set of steady-state values attached to this group of equations.

Proposition 1. *The steady-state equilibrium point $\mathbf{v}^* \in E$ exists and it is unique.*

Proof. See Appendix.

One should remark that the steady-state point is unique under the assumption that rumor-free equilibria are excluded from the analysis, i.e., that at least one of the following conditions holds: $y_0^{\omega} \neq 0$ or $y_0^{1-\omega} \neq 0$.

Although one cannot determine $\mathbf{v}^*$ explicitly for generic values of the various rates involved in the analysis, this becomes possible under constraint $\lambda \equiv \lambda_{\omega} = \lambda_{1-\omega}$, $\sigma \equiv \sigma_{\omega} = \sigma_{1-\omega}$, $\theta \equiv \theta_{\omega} = \theta_{1-\omega}$. In this case, the following result is derived.

Proposition 2. *For common parameter values θ, λ, σ, the steady-state point $\mathbf{v}^*$ corresponds to vector*

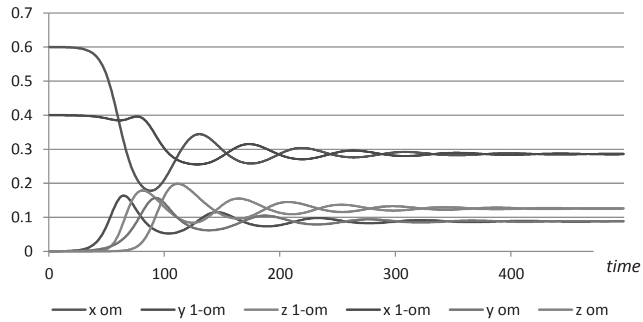
$$\begin{bmatrix} x^* \\ y^* \\ z^* \end{bmatrix} = \begin{bmatrix} \frac{\sigma}{2(\sigma+\lambda)} \\ \frac{\theta\lambda}{2\theta(\sigma+\lambda)+\lambda\sigma} \\ \frac{\lambda^2\sigma}{2(\sigma+\lambda)[2\theta(\sigma+\lambda)+\lambda\sigma]} \end{bmatrix},$$

where $x^* \equiv (x^{\omega})^* = \left(x^{1-\omega}\right)^*$; $y^* \equiv (y^{\omega})^* = \left(y^{1-\omega}\right)^*$; $z^* \equiv (z^{\omega})^* = \left(z^{1-\omega}\right)^*$.

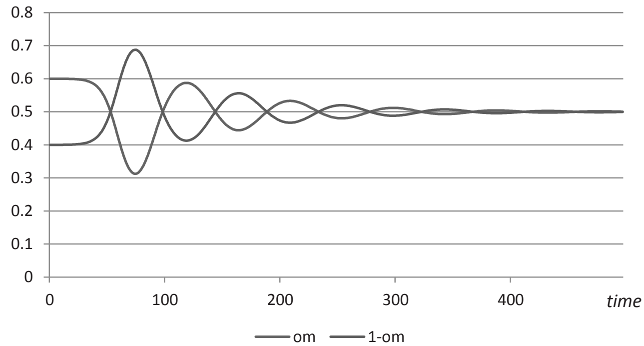
Proof. See Appendix.

If variables in vector $\mathbf{v}^*$ converge to the steady-state, then the shares of optimistic and pessimistic agents will remain constant after the transient phase is completed. In the long-term there will exist six classes of individuals: those who are optimists (pessimists) and ignore any rumor that can change their sentiments, those who are pessimists (optimists) and spread this sentiment, and those who are pessimists (optimists), do not spread the sentiment and are not susceptible of being 'infected' by the other sentiment.

Figures 2 and 3 display illustrative time trajectories for the dynamics of the sentiment-switching model for specific values of parameters. For the construction of Figure 2, it is assumed $\lambda_{\omega} = \lambda_{1-\omega} = 0.25$; $\sigma_{\omega} = \sigma_{1-\omega} = 1/3$; $\theta_{\omega} = \theta_{1-\omega} = 0.05$. The upper panel presents the time trajectories of the six categories of agents; as one observes, the values of variables oscillate around the steady-state as they approach it. Furthermore, compared with Figure 1, it is evident that the number of spreaders, for each class of sentiment, never falls to zero; as spreaders become stiflers, some previously susceptible individuals become spreaders. This dynamic process is possible because the susceptible category continuously receives individuals that no longer spread the respective sentiment. Note, as well, that, according to the result in Proposition 2, values of x^*, y^* and z^* are identical for both sentiments. The lower panel represents



Susceptible-spreaders-stiflers time trajectories



Optimists-pessimists time trajectories

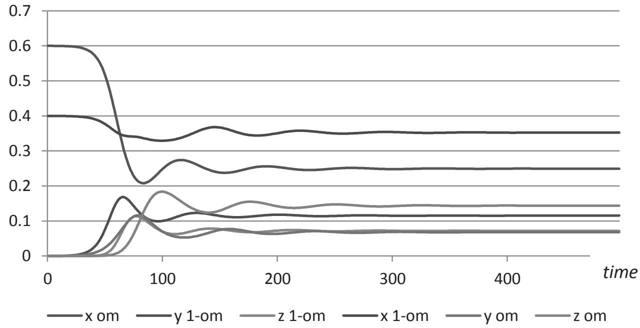
Figure 2. Sentiment-switching dynamics ($\lambda_\omega = \lambda_{1-\omega} = 0.25$; $\sigma_\omega = \sigma_{1-\omega} = 0.33$; $\theta_\omega = \theta_{1-\omega} = 0.05$)

the shares of optimists and pessimists; the symmetry triggered by the coincidence in parameter values implies that $\omega^* = 1 - \omega^* = 0.5$.

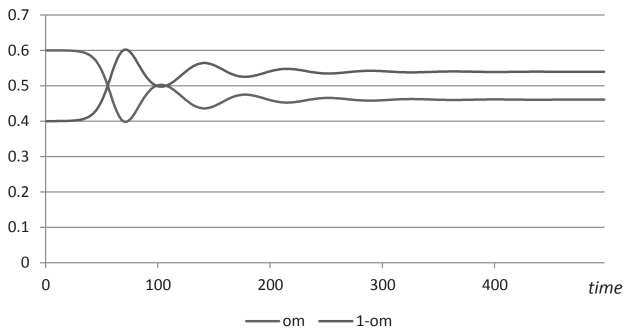
Figure 3 is generated for different parameter values of the various rates under each sentiment. In particular, the example takes $\lambda_\omega = 0.25$; $\lambda_{1-\omega} = 0.3$, $\sigma_\omega = 1/3$; $\sigma_{1-\omega} = 0.5$; $\theta_\omega = 0.05$; $\theta_{1-\omega} = 0.1$. Differences in parameters annulate the steady-state symmetry and make the number of optimists differ, in the long-run, relatively to the number of pessimists. In this particular case, $\omega^* = 0.46068$, $1 - \omega^* = 0.53932$.

4. The cyclicality mechanism

Sentiment switching, as characterized in the previous section, generates time series for the shares of optimists and pessimists that exhibit an oscillatory movement. As time unfolds, however, such cycles tend to diminish their intensity and fade away as the values of variables converge to their steady-state positions. In this section, we introduce an additional assumption, which allows for the cyclical motion of the sentiment shares to be perpetuated in time.



Susceptible-spreaders-stiflers time trajectories



Optimists-pessimists time trajectories

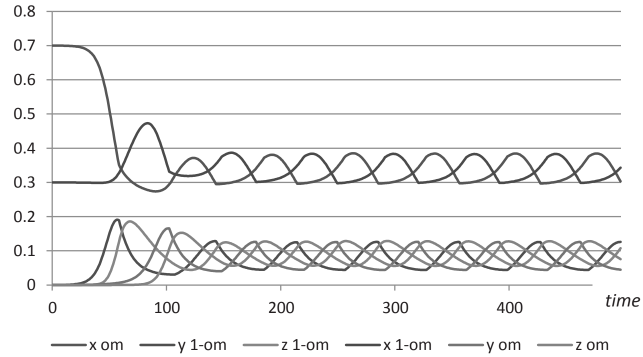
Figure 3. Sentiment-switching dynamics ($\lambda_\omega = 0.25$; $\lambda_{1-\omega} = 0.3$, $\sigma_\omega = 0.33$; $\sigma_{1-\omega} = 0.5$; $\theta_\omega = 0.05$; $\theta_{1-\omega} = 0.1$)

The new assumption requires maintaining the values of parameters σ_ω, $\sigma_{1-\omega}$, θ_ω and $\theta_{1-\omega}$ constant, but to allow λ_ω and $\lambda_{1-\omega}$ to take different values in two different circumstances. Specifically, we consider that the groups of susceptible agents are able to observe the rates of infection and to separate two cases, the one in which the growth rate of sentiment spreading is non negative and the opposite case. Susceptible agents will react as follows:

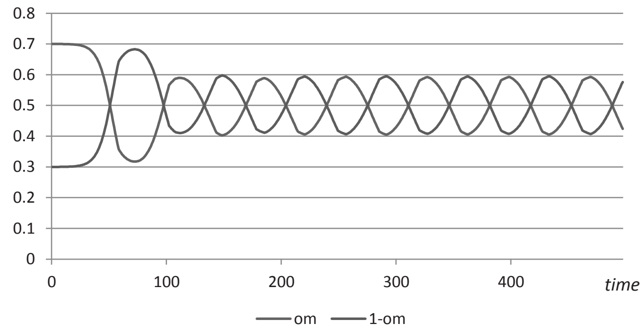
(i) If the growth rate at which optimistic/pessimistic sentiments are spread is positive or zero, then the rate at which pessimists/optimists are infected with a positive/negative sentiment is high;

(ii) If the growth rate at which optimistic/pessimistic sentiments are spread is negative, then the rate at which pessimists/optimists are infected with a positive/negative sentiment is low.

This mechanism translates the idea that the strength of sentiment spreading influences how susceptible the susceptible individuals are. They are more susceptible if



Susceptible-spreaders-stiflers time trajectories



Optimists-pessimists time trajectories

Figure 4. Sentiment cycles ($\lambda_0^{\omega} = \lambda_0^{1-\omega} = 0.1$; $\lambda_1^{\omega} = \lambda_1^{1-\omega} = 0.25$; $\sigma_{\omega} = \sigma_{1-\omega} = 0.33$; $\theta_{\omega} = \theta_{1-\omega} = 0.05$)

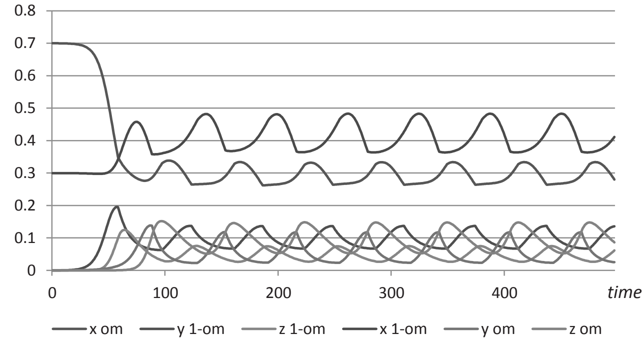
the sentiment is propagating at an increasing rate. Analytically, the described process might be formulated in the following form:

$$\lambda_{1-\omega,t} = \begin{cases} \lambda_0^{1-\omega} \text{ if } \gamma_{t-1}^{\omega} < 0 \\ \lambda_1^{1-\omega} \text{ if } \gamma_{t-1}^{\omega} \geq 0 \end{cases}, \ \gamma_t^{\omega} = \left(y_t^{\omega} - y_{t-1}^{\omega}\right)/y_{t-1}^{\omega}, \ \lambda_0^{1-\omega} < \lambda_1^{1-\omega}$$

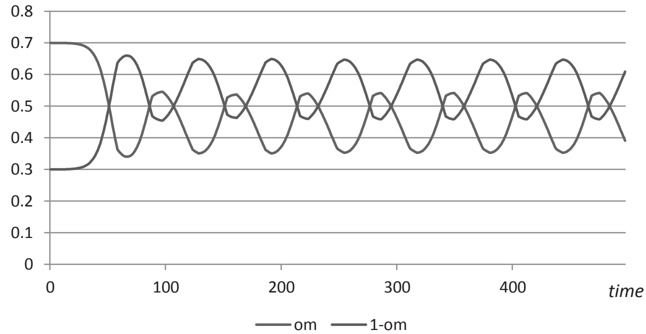
and

$$\lambda_{\omega,t} = \begin{cases} \lambda_0^{\omega} \text{ if } \gamma_{t-1}^{1-\omega} < 0 \\ \lambda_1^{\omega} \text{ if } \gamma_{t-1}^{1-\omega} \geq 0 \end{cases}, \ \gamma_t^{1-\omega} = \left(y_t^{1-\omega} - y_{t-1}^{1-\omega}\right)/y_{t-1}^{1-\omega}, \ \lambda_0^{\omega} < \lambda_1^{\omega}$$

The reaction of susceptible individuals to the observed spreading rate triggers a perpetual cyclical movement on the shares of ignorants, spreaders and stiflers for both pessimists and optimists and, as a result, the number of optimists and pessimists will



Susceptible-spreaders-stiflers time trajectories



Optimists-pessimists time trajectories

Figure 5. Sentiment cycles ($\lambda_0^{\omega} = \lambda_0^{1-\omega} = 0.1$; $\lambda_1^{\omega} = 0.25$; $\lambda_1^{1-\omega} = 0.3$; $\sigma_{\omega} = 0.33$; $\sigma_{1-\omega} = 0.5$; $\theta_{\omega} = 0.05$; $\theta_{1-\omega} = 0.1$)

be continuously changing. Figures 4 and 5 illustrate this process for the same array of parameter values σ_{ω}, $\sigma_{1-\omega}$, θ_{ω} and $\theta_{1-\omega}$ as the one used to draw Figures 2 and 3. The only change is in the values of λ; now, we take $\lambda_0^{\omega} = \lambda_0^{1-\omega} = 0.1$, $\lambda_1^{\omega} = \lambda_1^{1-\omega} = 0.25$, in the first case, and $\lambda_0^{\omega} = \lambda_0^{1-\omega} = 0.1$, $\lambda_1^{\omega} = 0.25$, $\lambda_1^{1-\omega} = 0.3$ in the second case.

The adaptation of the rumor propagation model to the sentiment-switching process with a cyclicality mechanism exemplifies how sentiments of optimism and pessimism might spread regardless from economic conditions. There are periods in which the majority of the agents adopts an optimistic view of the world just because this sentiment is being propagated faster than the opposite sentiment. Under the proposed process, this situation tends to be reversed after some time periods, making pessimistic feelings to dominate in a given time interval; then, optimistic feelings take over again as dominant, and this process continues indefinitely.

Now that we have described the mechanism of aggregate mood swings that occur in a context of social interaction, next sections will integrate this behavioral process into a simple macro model.

5. Optimistic and pessimistic expectations

Sentiments might play a fundamental role on the process of formation of expectations, namely when some kind of departure relatively to the benchmark of rational expectations is considered, i.e., when some sort of bounded rationality is taken into account. Effectively, in order to proceed with the analysis it is now introduced a less than perfect forecasting rule. At this respect, we follow Brock et al. (2006), Dudek (2010) and Gomes (2012), who consider a device of 'optimized rationality', according to which the information required to form educated expectations is costly and agents have to weigh the benefits of generating accurate expectations against the cost associated to the acquisition and to the treatment of relevant information.

Agents will be interested in forming expectations about two variables: the inflation rate, π_t, and the output gap, g_t. Agents ignore, at period t, the values these variables will take in the subsequent period, $t+1$, but they can collect information in order to improve the reliability of the expectations. Information acquisition is costly. Each individual may acquire a predictor of a given quality; the better the quality, the more it will cost. When purchasing a predictor of quality $q_t \in (0,1)$, the individual will be acquiring a signal v_t. The exact shape of the signal depends on the type of agent, optimistic or pessimistic, one is considering. Specifically, the following signals are available to be acquired:

(i) Signal on future inflation, acquired by an optimistic agent:

$$v_t^{\omega,\pi} = \begin{cases} \pi_{t+1}, \text{ with probability } q_t^{\omega,\pi} \\ \pi_t - \varepsilon(\pi_t - \overline{\pi}), \text{ with probability } 1 - q_t^{\omega,\pi}, \ \ \varepsilon > 0 \end{cases} \tag{4}$$

When acquiring, at period t, a signal $v_t^{\omega,\pi}$, through the purchase of a predictor of quality $q_t^{\omega,\pi}$, one of two outcomes is possible: the signal will reveal the true value of the inflation rate with a probability $q_t^{\omega,\pi}$; the same signal will be totally uninformative with a probability $1-q_t^{\omega,\pi}$. An uninformed agent will make the following forecast for the inflation rate at period $t+1$: because the agent is optimistic, she will believe that the inflation rate will converge towards a socially known and accepted target value $\overline{\pi}$. This target might be, for instance, the objective set by the central bank to guarantee price stability. Hence, the expectation formed by the optimistic agent regarding future inflation is

$$E_t^{\omega}\left(\pi_{t+1}|v_t^{\omega,\pi}\right) = q_t^{\omega,\pi}\pi_{t+1} + \left(1-q_t^{\omega,\pi}\right)\left[\pi_t - \varepsilon(\pi_t - \overline{\pi})\right]. \tag{5}$$

(ii) Signal on future inflation, acquired by a pessimistic agent:

$$v_t^{1-\omega,\pi} = \begin{cases} \pi_{t+1}, \text{ with probability } q_t^{1-\omega,\pi} \\ \pi_t + \varepsilon(\pi_t - \overline{\pi}), \text{ with probability } 1 - q_t^{1-\omega,\pi} \end{cases} \tag{6}$$

A pessimistic agent, as an optimistic one, will be capable of predicting the true value of the inflation rate with a probability that corresponds directly to the quality of the predictor. However, if the agent is unable to produce the accurate forecast, what occurs with a probability $1-q_t^{1-\omega,\pi}$, then she will take the pessimistic attitude, which is, in this case, to believe that the inflation rate will diverge from the target value. The same parameter ε is considered in (4) and (6) in order to maintain a symmetry between the behavior of optimists and pessimists. In this case, the individual expectation is

$$E_t^{1-\omega}\left(\pi_{t+1}|v_t^{1-\omega,\pi}\right)=q_t^{1-\omega,\pi}\pi_{t+1}+\left(1-q_t^{1-\omega,\pi}\right)\left[\pi_t+\varepsilon(\pi_t-\overline{\pi})\right]. \tag{7}$$

Signals with a similar structure can be built for the output gap. Let $\overline{g}$ be the target defined by public authorities for this aggregate and recognized by the population as such; denote by η the rate at which uninformed agents expect a convergence (if they are optimists) or a divergence (if they are pessimists) relatively to the respective target value.

(iii) Signal on future output gap, acquired by an optimistic agent:

$$v_t^{\omega,g}=\begin{cases} g_{t+1}, \text{ with probability } q_t^{\omega,g} \\ g_t-\eta(g_t-\overline{g}), \text{ with probability } 1-q_t^{\omega,g},\ \ \eta>0 \end{cases} \tag{8}$$

(iv) Signal on future output gap, acquired by a pessimistic agent:

$$v_t^{1-\omega,g}=\begin{cases} g_{t+1}, \text{ with probability } q_t^{1-\omega,g} \\ g_t+\eta(g_t-\overline{g}), \text{ with probability } 1-q_t^{1-\omega,g} \end{cases} \tag{9}$$

The respective expectations are:

$$E_t^{\omega}\left(g_{t+1}|v_t^{\omega,g}\right)=q_t^{\omega,g}g_{t+1}+\left(1-q_t^{\omega,g}\right)\left[g_t-\eta(g_t-\overline{g})\right] \tag{10}$$

$$E_t^{1-\omega}\left(g_{t+1}|v_t^{1-\omega,g}\right)=q_t^{1-\omega,g}g_{t+1}+\left(1-q_t^{1-\omega,g}\right)\left[g_t+\eta(g_t-\overline{g})\right] \tag{11}$$

Next, we must approach how probabilities reflecting the quality of the signal are determined. At each date t, agents intend to purchase an optimal predictor, i.e., a predictor that delivers the best possible balance between the accuracy of the forecast and the minimization of information acquisition and processing costs. In this case, optimists and pessimists will, respectively, solve the following optimality problems:

$$\begin{aligned}\min_{q_t^{\omega,\pi},q_t^{\omega,g}} U_t^{\omega} &= \frac{1}{2}\left[E_t^{\omega}\left(\pi_{t+1}|v_t^{\omega,\pi}\right)-\pi_{t+1}\right]^2+ \\ &+\frac{1}{2}a\left[E_t^{\omega}\left(g_{t+1}|v_t^{\omega,g}\right)-g_{t+1}\right]^2+C\left(q_t^{\omega,\pi},q_t^{\omega,g}\right),\end{aligned} \tag{12}$$

and

$$\begin{aligned}\min_{q_t^{1-\omega,\pi},q_t^{1-\omega,g}} U_t^{1-\omega} &= \frac{1}{2}\left[E_t^{1-\omega}\left(\pi_{t+1}|v_t^{1-\omega,\pi}\right)-\pi_{t+1}\right]^2+ \\ &+\frac{1}{2}a\left[E_t^{1-\omega}\left(g_{t+1}|v_t^{1-\omega,g}\right)-g_{t+1}\right]^2+C\left(q_t^{1-\omega,\pi},q_t^{1-\omega,g}\right)\end{aligned} \tag{13}$$

In (12) and (13), parameter $a > 0$ represents the weight given to output stabilization relatively to price stability in the agents' objective functions, and functions $C(\cdot)$ translate the costs of acquisition of each one of the predictors. Convex cost functions are taken:

$$C(q_t^{\omega,\pi}, q_t^{\omega,g}) = \frac{1}{2}\psi\left[\left(q_t^{\omega,\pi}\right)^2 + \left(q_t^{\omega,g}\right)^2\right],\ \psi \geq 0 \tag{14}$$

$$C(q_t^{1-\omega,\pi}, q_t^{1-\omega,g}) = \frac{1}{2}\psi\left[\left(q_t^{1-\omega,\pi}\right)^2 + \left(q_t^{1-\omega,g}\right)^2\right] \tag{15}$$

The solutions of problems (12) and (13) are:

$$\frac{\partial U_t^{\omega}}{\partial q_t^{\omega,\pi}} = 0 \quad \Leftrightarrow \quad q_t^{\omega,\pi} = \frac{\left[\pi_{t+1} - \pi_t + \varepsilon(\pi_t - \overline{\pi})\right]^2}{\psi + \left[\pi_{t+1} - \pi_t + \varepsilon(\pi_t - \overline{\pi})\right]^2} \tag{16}$$

$$\frac{\partial U_t^{1-\omega}}{\partial q_t^{1-\omega,\pi}} = 0 \quad \Leftrightarrow \quad q_t^{1-\omega,\pi} = \frac{\left[\pi_{t+1} - \pi_t - \varepsilon(\pi_t - \overline{\pi})\right]^2}{\psi + \left[\pi_{t+1} - \pi_t - \varepsilon(\pi_t - \overline{\pi})\right]^2} \tag{17}$$

$$\frac{\partial U_t^{\omega}}{\partial q_t^{\omega,g}} = 0 \quad \Leftrightarrow \quad q_t^{\omega,g} = \frac{a\left[g_{t+1} - g_t + \eta(g_t - \overline{g})\right]^2}{\psi + a\left[g_{t+1} - g_t + \eta(g_t - \overline{g})\right]^2} \tag{18}$$

$$\frac{\partial U_t^{1-\omega}}{\partial q_t^{1-\omega,g}} = 0 \quad \Leftrightarrow \quad q_t^{1-\omega,g} = \frac{a\left[g_{t+1} - g_t - \eta(g_t - \overline{g})\right]^2}{\psi + a\left[g_{t+1} - g_t - \eta(g_t - \overline{g})\right]^2} \tag{19}$$

Optimal predictors (16) to (19) reflect the importance of information acquisition costs in forming expectations. Costless information ($\psi = 0$) implies $q = 1$ for every predictor, meaning that perfect foresight prevails. As the value of the cost parameter increases, the quality of the signal will fall and the perfect foresight outcome becomes progressively less probable. Although it is possible to compute optimal predictors, as presented above, these depend on future values of the inflation rate and of the output gap that are not known at date t (the predictors are used precisely because such values are not known with anticipation!). To circumvent this obstacle, various approaches are possible; Brock et al. (2006), for instance, resort to the concept of managerial perfect foresight equilibrium, while Dudek (2010) considers the possibility of computing an average of all the available signals. The approach we follow is simpler; it is considered that agents know the perfect foresight steady-state $(\widehat{\pi}, \widehat{g})$ and, in order to save effort and cognitive resources, they adopt a constant in time predictor where observable values of variables give place to the perfect foresight steady-state values. The inflation rate and the output gap $(\widehat{\pi}, \widehat{g})$ depend on the specific macro structure of the economy.[2] For an economy that is hypothetically resting in the defined steady-state, predictors (16) and (17) are identical,

$$\widehat{q}^{\omega,\pi} = \widehat{q}^{1-\omega,\pi} = \frac{\left[\varepsilon(\widehat{\pi} - \overline{\pi})\right]^2}{\psi + \left[\varepsilon(\widehat{\pi} - \overline{\pi})\right]^2}, \tag{20}$$

[2] These values are presented, in explicit form, in the next section, for the New-Keynesian model.

$$\widehat{q}^{\omega,g}=\widehat{q}^{1-\omega,g}=\frac{a\left[\eta\left(\widehat{g}-\overline{g}\right)\right]^{2}}{\psi+a\left[\eta\left(\widehat{g}-\overline{g}\right)\right]^{2}}. \tag{21}$$

Reconsider now expectations (5), (7), (10) and (11). By replacing the predictor values (16) and (17) in them, one obtains explicit expressions for each of the relevant expectations,

$$E_t^{\omega}\left(\pi_{t+1}|v_t^{\omega,\pi}\right)=\frac{\left[\varepsilon(\widehat{\pi}-\overline{\pi})\right]^{2}\pi_{t+1}+\psi\left[\pi_t-\varepsilon\left(\pi_t-\overline{\pi}\right)\right]}{\psi+\left[\varepsilon(\widehat{\pi}-\overline{\pi})\right]^{2}}, \tag{22}$$

$$E_t^{1-\omega}\left(\pi_{t+1}|v_t^{1-\omega,\pi}\right)=\frac{\left[\varepsilon(\widehat{\pi}-\overline{\pi})\right]^{2}\pi_{t+1}+\psi\left[\pi_t+\varepsilon\left(\pi_t-\overline{\pi}\right)\right]}{\psi+\left[\varepsilon(\widehat{\pi}-\overline{\pi})\right]^{2}}, \tag{23}$$

$$E_t^{\omega}\left(g_{t+1}|v_t^{\omega,g}\right)=\frac{a\left[\eta\left(\widehat{g}-\overline{g}\right)\right]^{2}g_{t+1}+\psi\left[g_t-\eta\left(g_t-\overline{g}\right)\right]}{\psi+a\left[\eta\left(\widehat{g}-\overline{g}\right)\right]^{2}}, \tag{24}$$

$$E_t^{1-\omega}\left(g_{t+1}|v_t^{1-\omega,g}\right)=\frac{a\left[\eta\left(\widehat{g}-\overline{g}\right)\right]^{2}g_{t+1}+\psi\left[g_t+\eta\left(g_t-\overline{g}\right)\right]}{\psi+a\left[\eta\left(\widehat{g}-\overline{g}\right)\right]^{2}}. \tag{25}$$

Observe, for expectations (22) to (25), that the absence of information costs implies a return to perfect foresight, i.e., $\psi=0\Rightarrow E_t^{\omega}\left(\pi_{t+1}|v_t^{\omega,\pi}\right)=E_t^{1-\omega}\left(\pi_{t+1}|v_t^{1-\omega,\pi}\right)=\pi_{t+1}$ and $E_t^{\omega}\left(g_{t+1}|v_t^{\omega,g}\right)=E_t^{1-\omega}\left(g_{t+1}|v_t^{1-\omega,g}\right)=g_{t+1}$.

Since we are interested in dealing with aggregate expectations, we have to compute the weighted average expectations in the economy. Given the shares of optimists and pessimists that populate the economy at date t, computed according to what was established in Section 4, such expectations are

$$E_t(\pi_{t+1})=\omega_t E_t^{\omega}\left(\pi_{t+1}|v_t^{\omega,\pi}\right)+(1-\omega_t)E_t^{1-\omega}\left(\pi_{t+1}|v_t^{1-\omega,\pi}\right), \tag{26}$$

$$E_t(g_{t+1})=\omega_t E_t^{\omega}\left(g_{t+1}|v_t^{\omega,g}\right)+(1-\omega_t)E_t^{1-\omega}\left(g_{t+1}|v_t^{1-\omega,g}\right). \tag{27}$$

The final expressions of the inflation rate and of the output gap expectations are obtained by replacing (22) and (23) into (26), and (24) and (25) into (27). They are,

$$E_t(\pi_{t+1})=\frac{\left[\varepsilon(\widehat{\pi}-\overline{\pi})\right]^{2}\pi_{t+1}+\psi\left[\pi_t+\varepsilon(1-2\omega_t)(\pi_t-\overline{\pi})\right]}{\psi+\left[\varepsilon(\widehat{\pi}-\overline{\pi})\right]^{2}}, \tag{28}$$

$$E_t(g_{t+1})=\frac{a\left[\eta\left(\widehat{g}-\overline{g}\right)\right]^{2}g_{t+1}+\psi\left[g_t+\eta(1-2\omega_t)(g_t-\overline{g})\right]}{\psi+a\left[\eta\left(\widehat{g}-\overline{g}\right)\right]^{2}}. \tag{29}$$

Note, also on the aggregate level, that if $\psi=0$, then $E_t(\pi_{t+1})=\pi_{t+1}$ and $E_t(g_{t+1})=g_{t+1}$.

6. Application to the New-Keynesian macro model

In this section, a characterization of the long-term dynamics of the New-Keynesian model is undertaken, taking into account expectation formation rules (28) and (29). We consider a reduced form of the model, which contemplates two difference equations, describing the demand-side and the supply-side of the economy.[3] The two equations are a dynamic IS curve that establishes the common opposite sign relation between the real interest rate, r_t, and the output gap,

$$g_t = -\varphi r_t + E_t(g_{t+1}) + \mu_t, \ \ \varphi > 0, \tag{30}$$

and a New-Keynesian Phillips curve,

$$\pi_t = \kappa g_t + \beta E_t(\pi_{t+1}) + \upsilon_t, \ \ \kappa, \beta \in (0,1). \tag{31}$$

Parameter β is the discount factor and κ measures the degree of price stickiness; the lower the value of κ, the stickier prices are. Variables μ_t and υ_t correspond to white noise disturbances that influence, respectively, demand and supply. The real interest rate is given by the Fisher equation, $r_t = i_t - E_t(\pi_{t+1})$, with i_t the nominal interest rate; and monetary policy is implemented through a standard Taylor rule,

$$i_t = \rho i_{t-1} + (1-\rho)\left\{\phi_\pi\left[E_t(\pi_{t+1}) - \overline{\pi}\right] + \phi_g g_t\right\}, \ \rho \in (0,1), \phi_\pi > 1, \phi_g \geq 0. \tag{32}$$

In equation (32), parameter ρ translates policy inertia. Values ϕ_π and ϕ_g are policy parameters. Condition $\phi_\pi > 1$ guarantees, under this model's specification, the determinacy of the model, $\forall \phi_g \geq 0$.

Our goal is not to pursue a thorough investigation of the model's dynamics; instead, we will concentrate the analysis in the steady-state. First, we derive the perfect foresight steady-state equilibrium.

Proposition 3. *A perfect foresight steady-state equilibrium for the New-Keynesian macro model composed by equations (30), (31) and (32) exists, it is unique and it is given by the pair of values*

$$(\widehat{\pi}, \widehat{g}) = \left(\frac{\phi_\pi}{\phi_\pi - 1 + \frac{1-\beta}{\kappa}\phi_g}\overline{\pi}; \frac{\phi_\pi}{\frac{\kappa}{1-\beta}(\phi_\pi - 1) + \phi_g}\overline{\pi} \right).$$

Proof. Solve the system (30)–(32) under conditions $\widehat{\pi} \equiv \pi_t = E_t(\pi_{t+1})$, $\widehat{g} \equiv g_t = E_t(g_{t+1})$, $\mu_t = \upsilon_t = 0$. □

Observe that, as long as the target inflation rate is positive, the values of $\widehat{\pi}$ and $\widehat{g}$ will also be positive, given the condition $\phi_\pi > 1$. If the central bank aims at a zero inflation rate, the perfect foresight equilibrium implies that not only the inflation rate but also the output gap are equal to zero. The system of equations allows, as well, to determine the steady-state value of the nominal interest rate, under conditions of

[3] See Clarida et al. (1999) and Woodford (2003), for details on the New-Keynesian model.

perfect foresight, which is $\widehat{i} = \widehat{\pi}$; i.e., in the perfect foresight equilibrium, the real interest rate is equal to zero.

Next, we need to compute the steady-state not under perfect foresight but under the sentiment expectations derived in the previous section. As in De Grauwe (2011) we remark that the microfoundations of this model were built under the implicit assumption that the expectations are rational and that one should be careful when extrapolating the analysis of the reduced form of the model to a scenario of bounded rationality; as in the mentioned paper, we follow the arguments in Evans and Honkapohja (2001), in order to consider it an admissible assumption. We define (π^*, g^*) as the steady-state that will hold under the following long-term expectations,

$$E_t(\pi^*) = \pi^* + (1 - 2\omega_t)\frac{\psi\varepsilon(\pi^* - \overline{\pi})}{\psi + [\varepsilon(\widehat{\pi} - \overline{\pi})]^2}, \tag{33}$$

$$E_t(g^*) = g^* + (1 - 2\omega_t)\frac{\psi\eta(g^* - \overline{g})}{\psi + a[\eta(\widehat{g} - \overline{g})]^2}. \tag{34}$$

Expectations (33) and (34) are steady-state expectation values withdrawn from (28) and (29) under conditions $\pi^* \equiv \pi_{t+1} = \pi_t$ and $g^* \equiv g_{t+1} = g_t$. These long-run expectations have interesting features. Expectations will coincide with observed steady-state values (what implies long-term perfect foresight) in four possible scenarios: (i) absence of information costs ($\psi = 0$); (ii) neutral sentiments ($\varepsilon = 0; \eta = 0$); (iii) coincidence between target values and steady-state levels ($\pi^* = \overline{\pi}; g^* = \overline{g}$); (iv) identical number of pessimists and optimists ($\omega_t = 1/2$). In the above expectations, we maintain the time subscript because share ω_t is subject, under the assumption introduced in Section 4, to perpetual motion.

Proposition 4. *The steady-state equilibrium under sentiment cyclicality exists, it is unique and it is the pair of values*

$$\begin{bmatrix} \pi^* \\ g^* \end{bmatrix} = \begin{bmatrix} \dfrac{\left\{\varphi\left[(\phi_\pi - 1)\Theta_t + \phi_\pi\right] + (\Lambda_t - \varphi\phi_g)\frac{\beta}{\kappa}\Theta_t\right\}\overline{\pi} - \Lambda_t\overline{g}}{\varphi(\phi_\pi - 1)(1 + \Theta_t) + (\varphi\phi_g - \Lambda_t)\frac{1}{\kappa}[1 - \beta(1 + \Theta_t)]} \\ \frac{1}{\kappa}[1 - \beta(1 + \Theta_t)]\,\pi^* + \frac{\beta}{\kappa}\Theta_t\overline{\pi} \end{bmatrix}$$

with $\Theta_t \equiv (1 - 2\omega_t)\frac{\psi\varepsilon}{\psi + [\varepsilon(\widehat{\pi} - \overline{\pi})]^2}$ *and* $\Lambda_t \equiv (1 - 2\omega_t)\frac{\psi\eta}{\psi + a[\eta(\widehat{g} - \overline{g})]^2}$.

Proof. Solve the system (30)–(32) under conditions $\pi^* \equiv \pi_t = \pi_{t+1}$, $g^* \equiv g_t = g_{t+1}$, $\mu_t = \upsilon_t = 0$, and with expectations given by (33) and (34). $\square$

Under $\Theta_t = \Lambda_t = 0$, we confirm that $(\pi^*, g^*) = (\widehat{\pi}, \widehat{g})$.

The comparison between the two steady-state results highlights essentially that cyclical sentiments can transform an otherwise fixed-point steady-state into a regular fluctuations long-term scenario. However, our setup is not fully deterministic and we might consider that demand and supply shocks continue to hit the economy in the long-run. In what follows, we numerically simulate the long-term outcome, comparing the rational expectations setup with the one that assumes sentiments. For such, we rewrite the steady-state results without the removal of the exogenous disturbances. We have:

(i) Rational expectations long-run outcome,

$$\begin{bmatrix} \widehat{\pi} \\ \widehat{g} \end{bmatrix}_{(\mu,\upsilon)} = \begin{bmatrix} \dfrac{\phi_\pi \overline{\pi} + \frac{\phi_g}{\kappa}\upsilon_t + \frac{1}{\varphi}\mu_t}{\phi_\pi - 1 + \frac{1-\beta}{\kappa}\phi_g} \\ \dfrac{\phi_\pi \overline{\pi} - \frac{\phi_\pi - 1}{1-\beta}\upsilon_t + \frac{1}{\varphi}\mu_t}{\frac{\kappa}{1-\beta}(\phi_\pi - 1) + \phi_g} \end{bmatrix}$$

(ii) Sentiment expectations long-run outcome,

$$\begin{bmatrix} \widehat{\pi} \\ \widehat{g} \end{bmatrix}_{(\mu,\upsilon)} = \begin{bmatrix} \dfrac{\left\{\varphi[(\phi_\pi - 1)\Theta_t + \phi_\pi] + \left(\Lambda_t - \varphi\phi_g\right)\frac{\beta}{\kappa}\Theta_t\right\}\overline{\pi} - \Lambda_t \overline{g} + \left(\varphi\phi_g - \Lambda_t\right)\frac{1}{\kappa}\upsilon_t + \mu_t}{\varphi(\phi_\pi - 1)(1+\Theta_t) + \left(\varphi\phi_g - \Lambda_t\right)\frac{1}{\kappa}[1-\beta(1+\Theta_t)]} \\ \frac{1}{\kappa}[1-\beta(1+\Theta_t)]\pi^* + \frac{\beta}{\kappa}\Theta_t\overline{\pi} - \frac{1}{\kappa}\upsilon_t \end{bmatrix}$$

The nature of the shocks is straightforward to understand from the rational expectations case: positive cost-push shocks rise inflation and lower output; positive demand shocks rise inflation and make effective output to increase as well, relatively to the potential level. In order to address business cycles dynamics, we will concentrate the analysis on the output gap series. Under rational expectations, the only source of fluctuations is the random realizations of the disturbance variables; in the sentiment scenario, an additional source emerges: sentiment cyclicality. The example that follows allows to illustrate how waves of optimism and pessimism imply a change on the interpretation one can make about long-term fluctuations.

The parameter values selected for the analysis are displayed in Table 1. Those which have to do directly with the macro model specification (the first row of values) are withdrawn from Woodford (2003, p. 341, 285); the others are reasonable and plausible values, that do not interfere significantly with the qualitative nature of the results.

Table 1. Parameter values

$\beta = 0.99; \varphi = 6.25; \kappa = 0.024; \phi_\pi = 2;$
$\overline{\pi} = 0.02; \overline{g} = 0.01; \psi = 1; \varepsilon = 0.175; \eta = 0.2; a = 0.25$
$\mu_t \sim N(0; 2.5 \times 10^{-7}); \upsilon_t \sim N(0; 2.5 \times 10^{-7})$

There is a parameter missing in Table 1. It is the monetary policy parameter associated with real stabilization. This is because the parameter has an important role in determining the results to be obtained and, therefore, to illustrate its relevance we will work with three different values: $\phi_g = 0$, $\phi_g = 0.25$ and $\phi_g = 0.5$.

Figures 6 and 7 display the long-term trajectories of the output gap, comparing the rational expectations and the sentiment cycles outcomes. Each figure corresponds to each one of the cases depicted in Figures 4 and 5 (recall that the difference between the

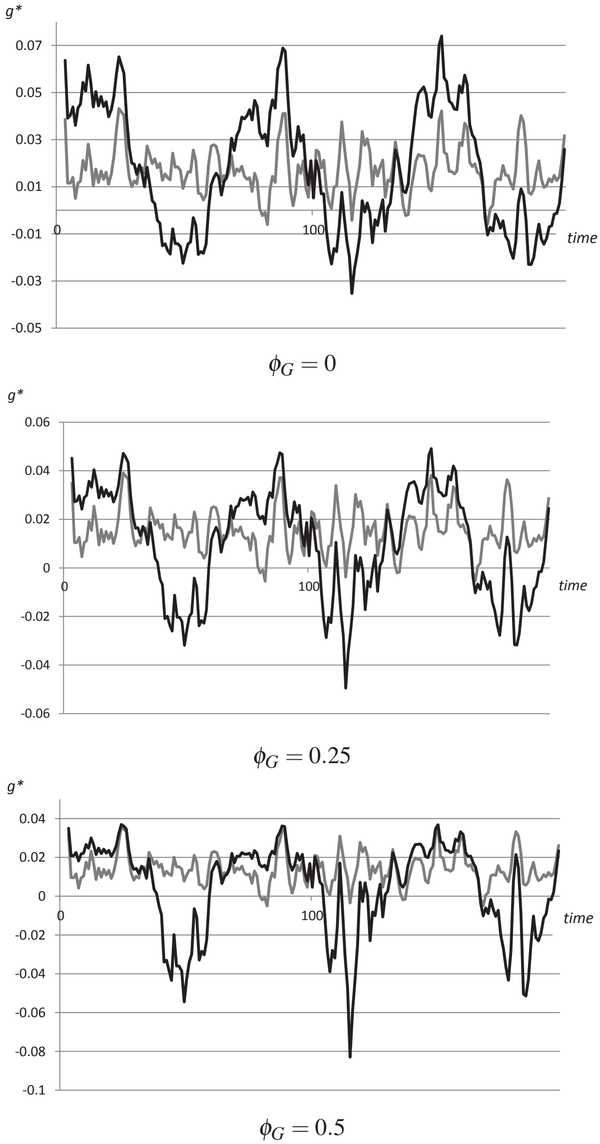


Figure 6. Output gap time trajectory: sentiment propagation example 1

two has to do with the values of parameters in the susceptible-spreader-stifler framework). Each figure has three panels that represent, from up to bottom, the cases $\phi_g = 0$, $\phi_g = 0.25$ and $\phi_g = 0.5$. In each figure, 200 time periods are assumed. In order to smooth the fluctuations, the presented trajectories are displayed as trend lines over the

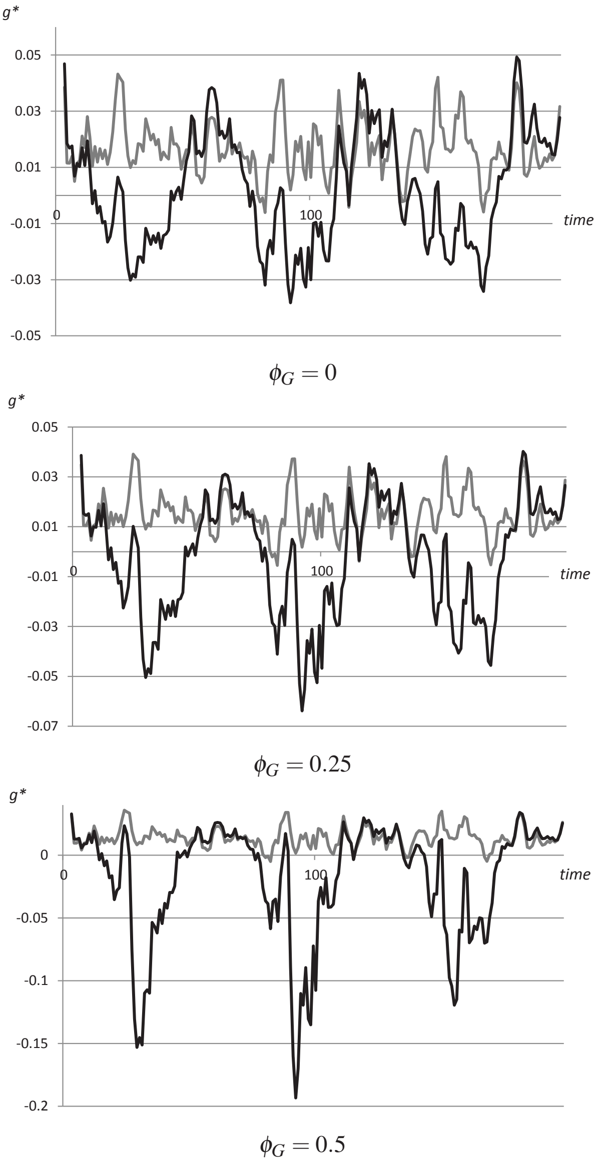


Figure 7. Output gap time trajectory: sentiment propagation example 2

original time-series taking a 4-period moving average. The darker lines correspond to the trajectories of the output gap under sentiment cyclicality; the brighter ones correspond to the rational expectations outcome.

Both figures show an evident result: the way sentiment cyclicality impacts on aggregate fluctuations is strongly influenced by the value of parameter ϕ_g. Time trajectories in Figure 6 differ from the ones in Figure 7 for just one fundamental reason: the number of pessimists is, on average, larger than the number of optimists in the case of Figure 7 and, thus, the output gap is, on average, a lower value on each of the three displayed examples. Concentrating the attention on the trajectories provided by Figure 6, note the following; when monetary authorities show no concern with real stabilization, sentiment cycles exacerbate both periods of expansion and periods of contraction of the economy, relatively to the benchmark of rational expectations. This introduces a more pronounced cyclical movement on a time series that otherwise follows a relatively erratic behavior. As we increase the value of ϕ_g, a relevant phenomenon occurs: the introduction of waves of optimism and pessimism do not generate periods of remarkable expansions relatively to the case of rational expectations; however, it allows for the occurrence of strong recessions, in which the trajectory of the output gap departs significantly from what the rational expectations analysis would predict.

Therefore, through the inspection of the trajectories, we find both a source of strong recessions and a policy recommendation to avoid them: strong recessions are the result of a an output stabilization effort on the part of the central bank; in order to avoid them, monetary authorities should concentrate on the price stability goal.[4]

Synthesizing, cycles of large amplitude are the result of a series of events that, once combined, can lead to strong recessions; they are:

(i) The social interaction process that transforms optimists into pessimists and the opposite, in a recurrent way over time;

(ii) Information costs, that prevent individuals from gaining access, under optimal conditions, to the knowledge required to formulate rational expectations;

(iii) Price stickiness, which is the main foundation on which the New-Keynesian model and, in particular, the New-Keynesian Phillips curve is built upon;

(iv) A misdirected monetary policy effort, that puts too much weight on output stabilization.

To gain further insights on the role of waves of optimism and pessimism over the benchmark New-Keynesian macro model, let us now simultaneously consider both policy goals: price stability and real stabilization. The following monetary policy objective function is adapted from Geraats (1999),

$$L_t = -\frac{1}{2}\left[(\pi_t^* - \overline{\pi}) \times 100\right]^2 + bf\left((g_t^* - \overline{g}) \times 100\right). \tag{35}$$

[4] In de Grauwe (2011), it is suggested that the presence of animal spirits will imply that inflation targeting monetary policy may no longer be optimal and that output stabilization could in fact improve welfare. Angeletos and La'O (2013) present arguments in the opposite direction: strategic uncertainty coming from animal spirits will contribute to an ineffective policy and, thus, the monetary authority should refrain pursuing measures that go beyond its main assignment, which is to guarantee price stability. The second interpretation is closer to the line of reasoning and to the results in this paper.

In expression (35), π_t^* and g_t^* furnish long-term values for inflation and output gap in the sentiment case; their time series are the ones displayed in Figure 6.[5] Parameter $b \geq 0$ reflects the weight of output stabilization as a policy goal, relatively to the price stability objective. Function f is such that $f'' < 0$ and $f''' > 0$, what signifies that negative deviations from the output gap target are more penalized than positive deviations, from the point of view of the central bank's objective. This is the same as saying that the central bank has a strong dislike for recessions. An admissible functional form, which obeys to the specified conditions, is

$$f\left((g_t^* - \overline{g}) \times 100\right) = 1 - \exp\left[-\frac{1}{2}(g_t^* - \overline{g}) \times 100\right] - \frac{1}{2}(g_t^* - \overline{g}) \times 100. \tag{36}$$

Note that $f(0) = 0$, i.e., when the value of the output gap coincides with the target, then the contribution of the output gap to the objective value L_t is zero. Observe, as well, that $f < 0$ for $g_t^* \neq \overline{g}$, i.e., $f(0)$ is the maximum value of f.

Given objective function (35) and the previously assumed parameter values, to which we add $b = 0.048$ (Woodford 2003, p. 431), one can make an inspection about the role that both policy parameters, ϕ_π and ϕ_g, have in allowing for a desirable policy result. Figure 8 draws the relation between the value of parameter ϕ_g and an average of the value of L over 200 long-term periods, $\langle L_{200} \rangle$. Three lines are displayed, for different values of the other parameter, $\phi_\pi = 1.95$, $\phi_\pi = 2$, $\phi_\pi = 2.05$. The results are evident: in order to maximize its utility, the central bank will have to choose between policies according to the following requisites,

(i) The larger the value of ϕ_π, i.e., the more aggressive monetary policy is in terms of promoting price stability, the higher is the obtained utility;

(ii) Real stabilization policy measures may enhance the utility outcome if it is applied with moderation. The figure indicates that, for each value of ϕ_π, the value of ϕ_g that maximizes the average value of L is located around $\phi_g = 0.2$.

Therefore, for a central bank that has, as policy goals, price stability and the avoidance of strong recessions, the effectiveness of its policy is best achieved by adopting an aggressive attitude relatively to price stability and by addressing, as well, output gap stabilization concerns, although changes in the interest rate to respond to output gap fluctuations should be relatively moderate.

Finally, we compare, for a specific policy value $\phi_\pi = 2$, the relation between ϕ_g and $\langle L \rangle$ under sentiment cycles and rational expectations. It is evident that the less intense fluctuations of the rational expectations case generate a better fit relatively to the designed policy goals, as shown in Figure 9. Thus, prior to specific policy actions, authorities should address another challenge: how can animal spirits be attenuated. If animal spirits refer, as pointed out in the introduction, to confidence, fairness and social attitudes, society should direct its efforts to promote ethical principles and education in

[5] The analysis is now restricted to a single case, namely the one in which identical parameter values for the sentiment switching setup are assumed.

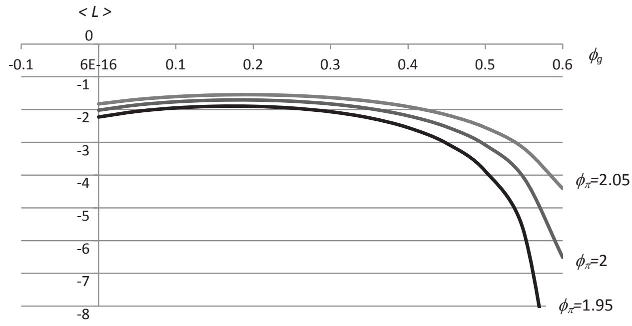


Figure 8. Utility of the central bank for different policy parameter values

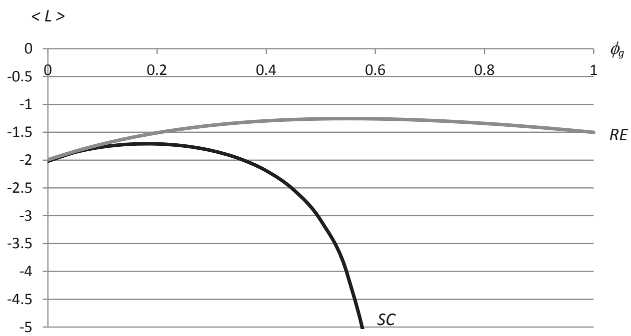


Figure 9. Utility of the central bank – comparison between rational expectations and sentiment cycles

order to attenuate the intensity of the sentiment switching that underlies the observed fluctuations on economic aggregates.

7. Conclusion

This paper proposed a foundation for the persistence of fluctuations in the aggregate sentiment level. Waves of optimism and pessimism alternate as the result of a fully deterministic dynamic process in which pessimists become optimists and optimists become pessimists under a susceptible-spreader-stifler sequence.

The cyclical nature of animal spirits, as discussed, can be introduced into a typical macroeconomic model in order to justify, at least partially, observed business cycles. The compatibility between the sentiment framework and a description of the macro environment requires some sort of departure relatively to the rational expectations paradigm. In this specific case, we consider a setting where the information required to form accurate predictions about future events is costly and, thus, agents' expectations may deviate from perfect foresight; when this occurs, agents will be optimistic or pessimistic about the future performance of the economy, with the shares of optimists and pessimists determined by the characterized rumor propagation framework.

The setup suggests that, in the long-term, observed fluctuations are strongly determined by sentiment switching with origins in social interaction. In this sense, the study supports the Keynesian view on animal spirits, that interprets business cycles as the outcome of forces that have to do with mass psychology much more than with concrete economic phenomena. Business cycles are the result of uncontrollable behavioral factors, and there is not much public authorities can do to avoid cyclical movements in sentiments, except contributing to a society based on fairness, confidence and social collaboration and cohesion.

However, the same is not true in what concerns the way sentiments shape expectations and impact on macro variables. Adequate policies to reduce the effect of systematic sentiment changes over the performance of the economy are essentially those that (i) reduce the cost of information acquisition; (ii) establish reasonable and realistic policy targets; (iii) develop monetary policy measures, by manipulating policy parameters, that might fight the undesirable consequences of natural sentiment fluctuations.

The analysis also suggested that, in the context of the New-Keynesian model, a strong effort to stabilize output may be counterproductive and may generate or perpetuate strong recessions. This conclusion is in syntony with a neoclassical interpretation of monetary policy intervention (i.e., the central bank should concentrate exclusively on its price stability mandate and avoid real stabilization measures that are often ineffective), what places the analysis in this paper in a same class as Angeletos and La'O (2013): although a Keynesian cornerstone is added to the discussion, the implications of the analysis are typically neoclassical, with observed cycles being largely determined by uncontrollable sentiment fluctuations that are due to interaction and communication frictions that cannot be successfully mitigated through direct stabilization policy intervention.

Acknowledgement I would like to acknowledge the helpful comments of two anonymous referees. The usual disclaimer applies.

References

Akerlof, G. A. and Shiller, R. J. (2009). *Animal Spirits: How Human Psychology Drives the Economy, and Why It Matters for Global Capitalism*. Princeton, NJ: Princeton University Press.

Angeletos, G. M., Collard, F. and Dellas, H. (2015). Quantifying Confidence. London, Center for Economic Policy Research, *CEPR discussion paper* No. 10463.

Angeletos, G. M. and La'O, J. (2013). Sentiments. *Econometrica*, 81, 739–779.

Bidder, R. M. and Smith, M. E. (2012). Robust Animal Spirits. *Journal of Monetary Economics*, 59, 738–750.

Bofinger, P., Debes, S.,Gareis, J. and Mayer, E. (2013). Monetary Policy Transmission in a Model with Animal Spirits and House Price Booms and Busts. *Journal of Economic Dynamics and Control,* 37, 2862–2881.

Brock, W. A., Dindo, P. and Hommes, C. H. (2006). Adaptive Rational Equilibrium with Forward Looking Agents. *International Journal of Economic Theory*, 2, 241–278.

Cintron-Arias, A. (2006). *Modeling and Parameter Estimation of Contact Processes*. Cornell University, Dissertation Thesis.

Clarida, R., Gali, J. and Gertler, M. (1999). The Science of Monetary Policy: A New Keynesian Perspective. *Journal of Economic Literature*, 37, 1661–1707.

Daley, D. J. and Kendall, D. G. (1964). Epidemics and Rumors. *Nature*, 204, 1118.

Daley, D. J. and Kendall, D. G. (1965). Stochastic Rumours. *Journal of the Institute of Mathematics and Its Applications*, 1, 42–55.

De Grauwe, P. (2011). Animal Spirits and Monetary Policy. *Economic Theory*, 47, 423–457.

De Grauwe, P. (2012). Booms and Busts in Economic Activity: A Behavioral Explanation. *Journal of Economic Behavior and Organization*, 83, 484–501.

Dudek, M. K. (2010). A Consistent Route to Randomness. *Journal of Economic Theory*, 145, 354–381.

Evans, G. and Honkapohja, S. (2001). *Learning and Expectations in Macroeconomics*. Princeton, NJ: Princeton University Press.

Franke, R. (2012). Microfounded Animal Spirits in the New Macroeconomic Consensus. *Studies in Nonlinear Dynamics and Econometrics*, 16, 1–41.

Geraats, P. M. (1999). Inflation and Its Variation: An Alternative Explanation. UC Berkeley, Center for International and Development Economics Research, Working Paper No. qt56b2g3vn.

Gomes, O. (2012). Rational Thinking under Costly Information – Macroeconomic Implications. *Economics Letters*, 115, 427–430.

Huo, L., Huang, P. and Guo, C. X. (2012). Analyzing the Dynamics of a Rumor Transmission Model with Incubation. *Discrete Dynamics in Nature and Society*, article ID 328151.

Kocherlakota, N. R. (2010). Modern Macroeconomic Models as Tools for Economic Policy. Federal Reserve Bank of Minneapolis, *the Region*, May, 5–21.

Lengnick, M. and Wohltmann, H.-W. (2013). Agent-Based Financial Markets and New Keynesian Macroeconomics: A Synthesis. *Journal of Economic Interaction and Coordination*, 8, 1–32.

Maki, D. P. and Thompson, M. (1973). *Mathematical Models and Applications, with Emphasis on Social, Life, and Management Sciences*. Englewood Cliffs, NJ: Prentice-Hall.

Milani, F. (2011). Expectation Shocks and Learning as Drivers of the Business Cycle. *Economic Journal*, 121, 379–401.

Nekovee, M., Moreno, Y., Bianconi, G. and Marsili, M. (2007). Theory of Rumor Spreading in Complex Social Networks. *Physica A*, 374, 457–470.

Pastor-Santorras, R. and Vespignani, A. (2004). *Evolution and Structure of the Internet: a Statistical Physics Approach*. Cambridge, UK: Cambridge University Press.

Piqueira, J. R. C. (2010). Rumor Propagation Model: An Equilibrium Study. *Mathematical Problems in Engineering,* article ID 631357.

Thompson, K., Estrada, R. C., Daugherty, D. and Cintron-Arias, A. (2003). A Deterministic Approach to the Spread of Rumors. Cornell University, Department of Biological Statistics & Computational Biology, Technical Report BU-1642-M.

Wang, Y. Q., Yang, X. Y., Han, Y. L. and Wang, X. A. (2013). Rumor Spreading Model with Trust Mechanism in Complex Social Networks. *Communications in Theoretical Physics*, 59, 510–516.

Woodford, M. (2003). *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton, NJ: Princeton University Press.

Zanette, D. H. (2002). Dynamics of Rumor Propagation on Small-World Networks. *Physical Review E*, 65, 041908.

Zhao, L. J., Wang, J. J., Chen, Y. C., Wang, Q., Cheng, J. J. and Cui, H. X. (2012). SIHR Rumor Spreading Model in Social Networks. *Physica A*, 391, 2444–2453.

Appendix

Proof of Proposition 1

Applying equilibrium condition $\mathbf{v}^* - V(\mathbf{v}^*) = 0$ to system (3), the following chain of equalities will hold in the steady-state,[6]

$$\begin{aligned}
\theta_\omega (z^\omega)^* &= \lambda_\omega (x^\omega)^* \left(y^{1-\omega}\right)^* = \sigma_\omega \left(y^{1-\omega}\right)^* \left[\left(y^{1-\omega}\right)^* + \left(z^{1-\omega}\right)^*\right] \\
&= \theta_{1-\omega} \left(z^{1-\omega}\right)^* = \lambda_{1-\omega} \left(x^{1-\omega}\right)^* (y^\omega)^* = \sigma_{1-\omega} (y^\omega)^* \left[(y^\omega)^* + (z^\omega)^*\right]
\end{aligned} \tag{37}$$

From (37), it is straightforward the computation of the following equilibrium relations,

$$\frac{(z^\omega)^*}{\left(z^{1-\omega}\right)^*} = \frac{\theta_{1-\omega}}{\theta_\omega} \tag{38}$$

$$(z^\omega)^* = \frac{\sigma_{1-\omega} \left[(y^\omega)^*\right]^2}{\theta_\omega - \sigma_{1-\omega} (y^\omega)^*} \tag{39}$$

$$\left(z^{1-\omega}\right)^* = \frac{\sigma_\omega \left[\left(y^{1-\omega}\right)^*\right]^2}{\theta_{1-\omega} - \sigma_\omega \left(y^{1-\omega}\right)^*} \tag{40}$$

$$(x^\omega)^* = \frac{\sigma_\omega}{\lambda_\omega} \left[\left(y^{1-\omega}\right)^* + \left(z^{1-\omega}\right)^*\right] \tag{41}$$

$$\left(x^{1-\omega}\right)^* = \frac{\sigma_{1-\omega}}{\lambda_{1-\omega}} \left[(y^\omega)^* + (z^\omega)^*\right] \tag{42}$$

Solving (39) and (40) with respect to $(y^\omega)^*$ and $\left(y^{1-\omega}\right)^*$, respectively, replacing the results into (41) and (42), and making use of relation (38), one can display steady-state values $(y^\omega)^*$, $\left(y^{1-\omega}\right)^*$, $(x^\omega)^*$ and $\left(x^{1-\omega}\right)^*$ as depending solely on $(z^\omega)^*$. The expressions are:

$$(y^\omega)^* = \left(\frac{\sqrt{1 + \frac{4\theta_\omega}{\sigma_{1-\omega}(z^\omega)^*}} - 1}{2}\right)(z^\omega)^* \tag{43}$$

$$\left(y^{1-\omega}\right)^* = \frac{\theta_\omega}{\theta_{1-\omega}} \left(\frac{\sqrt{1 + \frac{4(\theta_{1-\omega})^2}{\sigma_\omega \theta_\omega (z^\omega)^*}} - 1}{2}\right)(z^\omega)^* \tag{44}$$

[6] Condition $y_0^\omega \neq 0 \vee y_0^{1-\omega} \neq 0$ is implicitly assumed.

$$(x^{\omega})^* = \frac{\sigma_\omega}{\lambda_\omega}\frac{\theta_\omega}{\theta_{1-\omega}}\left(\frac{\sqrt{1+\frac{4(\theta_{1-\omega})^2}{\sigma_\omega\theta_\omega(z^{\omega})^*}}+1}{2}\right)(z^{\omega})^* \tag{45}$$

$$\left(x^{1-\omega}\right)^* = \frac{\sigma_{1-\omega}}{\lambda_{1-\omega}}\left(\frac{\sqrt{1+\frac{4\theta_\omega}{\sigma_{1-\omega}(z^{\omega})^*}}+1}{2}\right)(z^{\omega})^* \tag{46}$$

Next, we apply condition $(x^{\omega})^* + \left(x^{1-\omega}\right)^* + (y^{\omega})^* + \left(y^{1-\omega}\right)^* + (z^{\omega})^* + \left(z^{1-\omega}\right)^* = 1$. This is equivalent to

$$\begin{aligned}
&\frac{\sigma_\omega}{\lambda_\omega}\frac{\theta_\omega}{\theta_{1-\omega}}\left(\frac{\sqrt{1+\frac{4(\theta_{1-\omega})^2}{\sigma_\omega\theta_\omega(z^{\omega})^*}}+1}{2}\right)(z^{\omega})^* + \frac{\sigma_{1-\omega}}{\lambda_{1-\omega}}\left(\frac{\sqrt{1+\frac{4\theta_\omega}{\sigma_{1-\omega}(z^{\omega})^*}}+1}{2}\right)(z^{\omega})^* \\
&+\left(\frac{\sqrt{1+\frac{4\theta_\omega}{\sigma_{1-\omega}(z^{\omega})^*}}-1}{2}\right)(z^{\omega})^* + \frac{\theta_\omega}{\theta_{1-\omega}}\left(\frac{\sqrt{1+\frac{4(\theta_{1-\omega})^2}{\sigma_\omega\theta_\omega(z^{\omega})^*}}-1}{2}\right)(z^{\omega})^* \\
&+(z^{\omega})^* + \frac{\theta_\omega}{\theta_{1-\omega}}(z^{\omega})^* = 1
\end{aligned} \tag{47}$$

Equation (47) can be rearranged and presented under the form,

$$\begin{aligned}
(z^{\omega})^* \quad = \quad & 2\left[\frac{\theta_\omega}{\theta_{1-\omega}}\left(1+\frac{\sigma_\omega}{\lambda_\omega}\right)\left(\sqrt{1+\frac{4\left(\theta_{1-\omega}\right)^2}{\sigma_\omega\theta_\omega\left(z^{\omega}\right)^*}}+1\right)+\right. \\
&\left.\left(1+\frac{\sigma_{1-\omega}}{\lambda_{1-\omega}}\right)\left(\sqrt{1+\frac{4\theta_\omega}{\sigma_{1-\omega}\left(z^{\omega}\right)^*}}+1\right)\right]^{-1}.
\end{aligned} \tag{48}$$

Although one cannot solve, in its generic form, equation (48) in order to encounter an explicit expression for the steady-state value of z_t^{ω}, it is possible to confirm that a unique $(z^{\omega})^*$ is the solution of (48). Let $F[(z^{\omega})^*]$ represent the r.h.s. of the equation. The evaluation of the properties of F reveal the following: $\forall\theta_\omega$, $\theta_{1-\omega}$, λ_ω, $\lambda_{1-\omega}$, σ_ω, $\sigma_{1-\omega} \in (0,1)$: $\lim_{(z^{\omega})^*\to 0}F[(z^{\omega})^*] = 0$; $\lim_{(z^{\omega})^*\to 1}F[(z^{\omega})^*] \in (0,1)$; $F'[(z^{\omega})^*] > 0$ and $F''[(z^{\omega})^*] < 0$. This set of properties directly implies that $F[(z^{\omega})^*]$ intersects line $(z^{\omega})^*$ in the $\left((z^{\omega})^*; F[(z^{\omega})^*]\right)$ locus once and only once for $(z^{\omega})^* \in (0,1)$, what allows to unequivocally state that a unique solution for $(z^{\omega})^*$ exits.

Given the unique solution for $(z^{\omega})^*$, equations (38) and (43)–(46) guarantee that a single 6-dimensional array of steady-state values $\mathbf{v}^*$ exists. □

Proof of Proposition 2

Because parameter values are identical in the cases of pessimistic and optimistic contagion, the model becomes completely symmetric in terms of the associated dynamics between the two rumor spreading processes. In this case, the steady-state values $(x^{\omega})^*$ and $\left(x^{1-\omega}\right)^*$, $(y^{\omega})^*$ and $\left(y^{1-\omega}\right)^*$, $(z^{\omega})^*$ and $\left(z^{1-\omega}\right)^*$ must be identical. Therefore, the following equilibrium relation holds,

$$\theta z^* = \sigma y^*(y^* + z^*) = \lambda x^* y^* \tag{49}$$

From (49), simple algebra conducts to the results in the proposition, after noticing that $2x^* + 2y^* + 2z^* = 1$. □

Czech Economic Review 9 (2015) 135–154
Acta Universitatis Carolinae Oeconomica

Differential Game Approach for International Environmental Agreements with Social Externalities

Lina Mallozzi*, Stefano Patrì, Armando Sacco†**

Received 12 November 2015; Accepted 29 March 2016

Abstract In this work we study an N-player differential game, in which positive social externalities affect the payoffs of the players when they make an agreement. We divide the N players in two homogeneous groups, N_1 developed countries and N_2 developing countries. For the latter, we consider a damage-cost function that evolves in time. We imagine the externalities as the possibility that bilateral or multilateral agreements of various nature are by-products of an International Environmental Agreement (IEA). After the determination of emissions solutions, we use the externalities to investigate whether it is possible to have a self-enforcing agreement on pollution emissions in the short run.

Keywords Differential game, self-enforcing agreement, social externality, asymmetric players
JEL classification C72, C73

1. Introduction

A great part of environmental problems, like global warming, depletion of ozone layer or loss of biological diversity, is related to global commons and, for that, requires global policies. During the last three decades, many times countries organized meetings to find an agreement on pollution control. In 1987 in the Canadian city of Montreal, was ratified the Protocol on Substances that Deplete the Ozone Layer. Since 1995 the United Nations organize yearly conferences, Conferences of Parties, within which in 1997 was signed the Kyoto Protocol for the reduction of Green House Gases (GHG), with the objective to contain the global warming.

From an economic point of view, International Environmental Agreements (IEA) lie within the coordination problem class. A natural approach to this kind of problem is game theory. So, there is an extensive literature on this argument, that approaches the problem both as cooperative and non-cooperative games, both as static and dynamic games. The first attempts set the issue in a static context (see e.g, Carraro and Siniscalco 1993; Barrett 1994), and this way is still followed and gives some important contributions. But pollution is an evolving phenomenon, so a dynamic game approach can lead to more explicative results (see e.g., Casino and Rubio 2005; Rubio and Ulph

* University of Naples Federico II, Faculty of Economics, Via Cintia, Monte S. Angelo, Naples 80126, Italy. E-mail: mallozzi@unina.it.
** University of Rome I, Piazzale Aldo Moro 5, Rome 00185, Italy. E-mail: stefano.patri@uniroma1.it.
† Corresponding author. University of Rome I, Piazzale Aldo Moro 5, Rome 00185, Italy. E-mail: armando.sacco@uniroma1.it.

2007; Breton et al. 2010). Both in static and dynamic context, the literature is divided into two streams: cooperative and non-cooperative games. The main focus of the first stream is to contrast the cooperative and non-cooperative solutions and to show the benefits of cooperation. The real question in these games is how to allocate the payoff among players.

The non-cooperative stream starts from the consideration that there is no supranational authority that can force countries to cooperate, thus players choose non-cooperatively whether join or not in a coalition. In this game it is necessary to specify the concept of stability of the coalition. In line with the stable set definition (see von Neumann and Morgenstern 1944), the terms most commonly used are those of internal and external stability, introduced in d'Aspremont et al. (1983). In few words, these two conditions say that a coalition is stable if none of the members has an incentive to defect from cooperation and none of non-members has an incentive to join. So, the two focus points are research of the solutions (emissions or abatement level) and research of the coalition's dimension.

We consider myopic players, that is to say that economic interests are still too strong than environmental concern. It could be a limited point of view. There are relevant examples, like EU, that put into the foreground the control of emissions. But, with the arrival of the new millennium, the economic center of the world has changed his coordinates, and with it the center of the environmental problems. The great challenge now is to include in emissions reduction process those countries that are not considered developed countries, but that give significant contribution to pollution (e.g, the countries called BRICS, Brazil, Russia, India, China, South Africa). We think that it is not realistic to ask those countries to take care of environment for some kind of farsightedness or consider some kind of punishment for those who do not cooperate. So, we want to try to design an IEA that is profitable.

This paper belongs to the non-cooperative games stream (for a recent surveys see Jørgensen et al. 2010; Long 2010). We have a group of developed countries and a group of developing countries. We suppose homogeneity within each group. We also assume that the payoff of cooperators is affected by social externalities that is a function of the coalition's dimension. We recall this idea from Cabon-Dhersin and Ramani (2006). They start from the evidence that, despite theoretic literature on non-cooperative games supports only small coalition, principal agreements are signed by many countries. Thus, their hypothesis is that when countries have to make the decision to join or not an agreement, they consider all possible earnings due to relations with other countries. In this way, they show that in a static game on abatement level, with N symmetric countries, it is possible to have a grand coalition. Our purpose is to verify whether their results are true also in a dynamic contest with asymmetric players. We work on emissions instead on abatement levels (for a discussion on the duality between emission and abatement level see Diamantoudi and Sartzetakis 2006; Finus 2001) and include the social externalities in an N-player differential game in the framework of Masoudi and Zaccour (2013). Thus, we have two different treatments for environmental concern of the two kinds of players. The idea is that developing countries have an environmental damage-cost that is not full from the outset, but increases in time with the increasing of their cumulative revenue.

Our main focus is on two points:

1. To find the feedback-Nash equilibrium emissions both for cooperators and defectors, and
2. to evaluate the size and the composition of the coalition by using the concept of internal and external stability.

The rest of the paper is organized as follows: In Section 2 we introduce the model. In Section 3 we characterize the cooperative and non-cooperative emission solutions. In Section 4 we discuss the stability of the coalition and show some numerical results and Section 5 concludes.

2. The Model

2.1 Social externalities

Why a country should join an agreement for pollution control? This is the main issue at which environmental economics literature tried to respond in the last decades. Actually, two real issues should be considered. The first is the need to involve developing countries in emissions' reduction process. The second is to consider the IEAs within the complexity of a large number of relationships between countries. The mechanisms most used in economic literature to reach a large coalition are *transfer scheme strategies* (see e.g., Fuentes-Albero and Rubio 2010; Pavlova and de Zeeuw 2013) and *trigger strategies* (see e.g., Hoel and Schneider 1997). The idea of a transfer scheme is that signatories use the gain from cooperation to convince defectors to join in the coalition. On the other side, in the case of trigger strategies, the assumption is that signatories have the power to punish the defectors.

A further way is the *issue linkage* literature (see e.g., Botteon and Carraro 1994; Le Breton and Souberyan 1997; Hübler and Finus 2013), in which the IEA is linked with another agreements, that could be a R&D, or trade or another economic issues. Considering the globalization of relations between countries, we think that transfer scheme and trigger strategies are not credible mechanisms to enlarge an environmental coalition. Issue linkage could be a suitable way to model an IEA, but we think that specifying a kind of side agreement does not allow to take into account several other connection between countries.

Our approach to solve the problem, is to assume that social externalities affect the pay-off of players that decide to join in a coalition. From a mathematical point, we assume that a strictly positive function is added to the pay-off of signatories, and this function does not depend on emission, but only on the number of players in coalition. The principal reason of that choice lies in the great flexibility of the possible interpretation of what this externalities represent. We only request that when countries decide whether to join a coalition, they consider the possible benefits deriving from being in relationships with other countries. A classical example is Russia, that ratified the Kyoto protocol with the hope to have more consideration when its entry in World Trade Organization (WTO) would have been voted. We can also imagine that collateral to environmental agreement, other agreements, like trading, R&D ecc., could be signed (like in issue linkage literature). The immateriality of this social externalities brings with it a certain degree of vagueness, because within this concept we include all the possible networks that countries could establish. Nevertheless, we think that the loss of descriptive power is acceptable, compared to the great flexibility that we gain.

2.2 Functional forms

We consider an N-player differential game, assuming the world divided in two types of countries. So, we have a total number of players $N = N_1 + N_2$, in which N_1 are developed countries and N_2 are developing countries (N, N_1 and N_2 are integer numbers). We use 1 to denote developed countries and 2 to denote developing countries. As usual, we proceed backward.

We first assume that a number k of players join the agreement, while the rest stays out. In particular we take $k = n_1 + n_2$, where n_1 are developed countries and n_2 are developing countries, and k, n_1, n_2 are integer numbers. The assumption that some players that act cooperatively, while other players act non-cooperatively, belongs into partial cooperative games (for a discussion see Mallozzi and Tijs 2008, 2009; Chakrabarti et al. 2011). According to this assumption, we first solve the emission game and then we use the optimal emissions to find the numbers n_1 and n_2 that satisfy the stability conditions.

First of all, due to the fact that emissions are by-product of industrial activities, and assuming that the function which relates emissions and production are smooth and invertible, we can express the production for country i as a function of emission levels. Denoting it with $f_i(e_i)$, being e_i the emission of country i, and assuming that it is a concave and increasing function, that is a standard assumption in literature (see, e.g., Finus 2001; Rubio and Casino 2005; Diamantoudi and Sartzetakis 2006), we can define the productions as

$$f_i(e_i) = \alpha_i e_i - \frac{1}{2} e_i^2,$$

where $\alpha_i > 0$, so that $f_i(e_i)$ is positive for all suitable value of emissions.

Moreover, our hypothesis is that developed countries have a higher degree of interest in environmental issues with respect to developing countries, both for economic and historical motivations. In the end, this is the same approach of the Kyoto protocol.

The point is that a developing country needs to improve its infrastructure, increasing per capita wealth, life expectance, instruction level, etc. In this context, environment is a “luxury good”. In addition, the actual level of stock pollutant cannot be attributed to developing countries.

So, we have a different degree of internalization of the environmental damage-cost, that is given by a different definition of the damage-cost functions. Being $S(t)$ the stock of pollutant at time t, we denote by $D_1(S)$ the cost for developed countries and we assume that it is full from the outset. We choose a linear function of stock of pollutant, that is a not uncommon choice (see Hoel and Schneider 1997; Breton et al. 2010) and it is supported by some empirical estimation (see Labriet and Loulou 2003). In the end, the difference between linear and a more realistic quadratic damage-cost function should be only quantitative, but not qualitative. Then, we assume

$$D_1(S) = \beta_1 S,$$

where $\beta_1 > 0$, so that the damage-cost function is an increasing function of stock of pollutant.

On the other side, for developing countries the full damage-cost is related to the achievement of a preset threshold in terms of cumulative discounted revenues, denoted by $Y_2(t)$. The idea of linking income and environmental quality is not new in literature, there is the well-known environmental Kutznets curve (EKC) and some works, like Shafik and Bandyopadhyay (1992), that support the empirical consistence of this hypothesis. Then, denoting by ρ the rate of time preference, that we assume common to all players, we have $Y_2(t) = \int_0^t f_i(e_i(z))e^{-\rho z}dz$. Given that relation, we can define the time T as the instant ($T > 0$) at which a country of type 2 start to fully account the damage-cost, that is the time at which

$$\int_0^T f_i(e_i(t))e^{-\rho t}dt = \overline{Y_2}$$

is verified, where $\overline{Y_2}$ is the threshold chosen.

Then, for players 2 the damage-cost function is described for any t in two intervals:

$$\begin{cases} d_2(S,t) = \frac{t}{T}\gamma\beta_2 S, & \forall t < T, \\ D_2(S) = \beta_2 S, & \forall t \geq T, \end{cases}$$

where $\beta_2 > 0$. Moreover, we suppose $\gamma \in \{0,1\}$. The case $\gamma = 0$ is equivalent to say that players of type 2 do not allow for pollution at all, until they reach the threshold $\overline{Y_2}$ (that's the spirit of Kyoto protocol). If $\gamma = 1$, we are in the case of gradual internalization of damage cost. The stock of pollutant $S(t)$ is solution of the following differential equation:

$$\dot{S}(t) = \mu\left(\sum_{i=1}^{N} e_i(t)\right) - \delta S(t), \;\; S(0) = S_0,$$

where μ is a positive scaling parameter and δ is the natural rate of absorption of pollution. Here S_0 is the initial value of the pollution.

From now on, we will skip the time argument if there is no risk of ambiguity. We can introduce the payoff functions, that we denote by w_i, $i = 1,2$, given by

$$w_1 = \int_0^\infty (f_1(e_1) - D_1(S))e^{-\rho t}dt,$$

$$w_2 = \int_0^T (f_2(e_2) - d_2(S,t))e^{-\rho t}dt + \int_T^\infty (f_2(e_2) - D_2(S))e^{-\rho(t-T)}dt.$$

The last function that we need to characterize is the social externalities, that we denote by $Ext(n_1,n_2)$. As in Cabon-Dhersin and Ramani (2006), we make the assumption that it is a positive linear function of the coalition size, as follows

$$Ext(n_1,n_2) = s_1 n_1 + s_2 n_2,$$

where we assume the marginal externality s_i to be positive for $i = 1,2$. Obviously, with n_i we state the number of players of kind i that join in coalition, precisely $n_i \in \{0,1,...,N_i\}$, $i = 1,2$.

3. Emission Solution

In this section we characterize the emission solutions both for signatories and defectors. We suppose that we have a set C of signatory players, with n_1 developed countries and n_2 developing countries and a set of defectors, that we denote with NC. Then, the cardinality of the two sets is given by $|C| = n_1 + n_2$ and $|NC| = N_1 - n_1 + N_2 - n_2$. As usual, every player $j \in NC$ maximizes his own welfare, while players in C maximize the joint welfare. Due to homogeneity within groups, we just have to find four emission solutions: two for signatories (called e_1^C and e_2^C), and two for the defectors (called e_1^{NC} and e_2^{NC}).

Thus, the problem for defectors is:

$$\max_{e_1} \int_0^{\infty} (f_1(e_1) - D_1(S))e^{-\rho t} dt, \tag{1}$$

and

$$\max_{e_2} \int_0^{T^{NC}} (f_2(e_2) - d_2(S,t))e^{-\rho t} dt + \int_{T^{NC}}^{\infty} (f_2(e_2) - D_2(S))e^{-\rho (t - T^{NC})} dt \tag{2}$$

$$\text{s.t.} \quad \dot{S}(t) = \mu \left(\sum_1^N e_i(t) \right) - \delta S(t), \;\; S(0) = S_0.$$

For signatories, the joint maximization is:

$$\begin{aligned} \max_{e_1,e_2} \int_0^{T^C} & (n_1 f_1(e_1) + n_2 f_2(e_2) - n_1 D_1(S) - n_2 d_2(S,t) + Ext(n_1,n_2))e^{-\rho t} dt + \\ & \int_{T^C}^{\infty} (n_1 f_1(e_1) + n_2 f_2(e_2) - n_1 D_1(S) - n_2 D_2(S) + Ext(n_1,n_2))e^{-\rho(t-T^C)} dt, \end{aligned} \tag{3}$$

$$\text{s.t.} \quad \dot{S}(t) = \mu \left(\sum_1^N e_i(t) \right) - \delta S(t), \;\; S(0) = S_0.$$

In the maximization problems we call T^{NC} and T^C the instants of time at which a developing country achieves the threshold to become developed, respectively in the cases of defector and signatory. It seems clear from the optimization problems that we consider a feedback emission game, but an open-loop membership game.

3.1 Emissions of defectors

To solve the problem, we use the dynamic programming method. We proceed backward, solving first the problem on $[T^{NC},\infty)$. So, we have to solve first:

$$\max_{e_i}\int_{T^{NC}}^{\infty}\left(\alpha_i e_i-\frac{1}{2}e_i^2-\beta_i S\right)e^{-\rho(t-T^{NC})}dt,$$

$$\text{s.t.}\quad \dot S(t)=\mu\left(n_1e_1^C(t)+n_2e_2^C(t)+(N_1-n_1)e_1^{NC}(t)+(N_2-n_2)e_2^{NC}(t)\right)-\delta S(t),$$

where $S(0)=S_0$. Using Hamilton-Jacobi-Bellman (HJB) equations, we find the solution:

$$e_i^{NC}(t)=\alpha_i-\mu\frac{\beta_i}{\rho+\delta},\qquad i=1,2. \tag{4}$$

Since the functional forms for players of type 1 is the same in the entire period, we have that the emissions of developed countries are in $[0,T^N]$:

$$e_1^{NC}(t)=\alpha_1-\mu\frac{\beta_1}{\rho+\delta}.$$

For developing countries we have to consider the different damage-cost function, so the problem is:

$$\max_{e_2}\int_0^{T^{NC}}\left(\alpha_2e_2-\frac{1}{2}e_2^2-\gamma\frac{t}{T^{NC}}\beta_2S\right)e^{-\rho t}dt,$$

$$\text{s.t.}\quad \dot S(t)=\mu\left(n_1e_1^C(t)+n_2e_2^C(t)+(N_1-n_1)e_1^{NC}(t)+(N_2-n_2)e_2^{NC}(t)\right)-\delta S(t),$$

where $S(0)=S_0$. In this case HJB equation leads us to the emissions, for $t\in[0,T^N]$:

$$\begin{aligned}e_2^{NC}(t)=\alpha_2-\mu\frac{\beta_2}{T^{NC}(\rho+\delta)^2}\Big[&\gamma\big(1+t(\rho+\delta)-e^{(\rho+\delta)(t-T^{NC})}\big)+\\&+T^{NC}(1-\gamma)(\rho+\delta)e^{(\rho+\delta)(t-T^{NC})}\Big].\end{aligned}\tag{5}$$

3.2 Emissions of signatories

As the previous case, we proceed backwards to find the optimal emissions of signatories. First of all we solve the problem in $[T^C,\infty)$:

$$\max_{e_1,e_2}\int_{T^C}^{\infty}\left[n_1\left(\alpha_1e_1-\frac{1}{2}e_1^2\right)+n_2\left(\alpha_2e_2-\frac{1}{2}e_2^2\right)-n_1\beta_1S-n_2\beta_2S+s_1n_1+s_2n_2\right]e^{-\rho(t-T^C)}dt,$$

$$\text{s.t.}\quad \dot S(t)=\mu\left(n_1e_1^C(t)+n_2e_2^C(t)+(N_1-n_1)e_1^{NC}(t)+(N_2-n_2)e_2^{NC}(t)\right)-\delta S(t),$$

where $S(0)=S_0$.

We want to highlight that the externalities has no effects on the determination of the emissions. Using HJB, we determine the solution for signatory countries as:

$$e_i^C(t) = \alpha_i - \mu \frac{n_1\beta_1 + n_2\beta_2}{\rho+\delta}, \quad i = 1,2. \tag{6}$$

In the period $[0, T^C]$, the different damage-cost function implies that the optimal solutions solve:

$$\max_{e_1,e_2} \int_0^{T^C} \left[n_1\left(\alpha_1 e_1 - \frac{1}{2}e_1^2\right) + n_2\left(\alpha_2 e_2 - \frac{1}{2}e_2^2\right) - n_1\beta_1 S - n_2\gamma\frac{t}{T^C}\beta_2 S + s_1 n_1 + s_2 n_2 \right] e^{-\rho t} dt,$$

$$\text{s.t.}\quad \dot{S}(t) = \mu\left(n_1 e_1^C(t) + n_2 e_2^C(t) + (N_1 - n_1)e_1^{NC}(t) + (N_2 - n_2)e_2^{NC}(t)\right) - \delta S(t),$$

where $S(0) = S_0$. The optimal emissions are given by:

$$\begin{aligned} e_i^C(t) = \alpha_i - \mu\frac{n_1\beta_1}{\rho+\delta} - \mu\frac{n_2\beta_2}{T^C(\rho+\delta)^2}\Big[&\gamma\big(1 + t(\rho+\delta) - e^{(\rho+\delta)(t-T^C)}\big) + \\ &+ T^C(1-\gamma)(\rho+\delta)e^{(\rho+\delta)(t-T^C)}\Big], \end{aligned} \tag{7}$$

where $i = 1,2$. Nevertheless, we will see that externalities will effect the payoffs of the players.

4. Stability

In order to approach the stability analysis, we use the concept of self-enforcing agreements. The idea is due to d'Aspremont et al. (1983). They use this concept to study the stability of a cartel, but in several places, it is also used to discuss the stability of environmental agreements. We want to highlight that these conditions are more stringent and there are different papers that try to propose different ways to face the problem (see Finus 2003; Eyckmans and Finus 2006). The basic idea is that a coalition is stable if no one inside has an incentive to defect and no one outside has an incentive to join in. So, called w_i the pay-off of a player i, a coalition of k players is stable if

$$w_i^C(k) \geq w_i^{NC}(k-1), \qquad w_j^{NC}(k) \geq w_j^C(k+1)$$

are verified $\forall i \in C$ and $\forall j \in NC$. First condition is called *internal* stability, while the second is called *external* stability.

In our case, having two different types of players, we need to adapt the definition, requiring that internal and external stability are verified both for developed and developing countries. So, we have to find the values n_1 and n_2 that solve this system of four inequalities:

$$\begin{cases} w_1^C(n_1,n_2) - w_1^{NC}(n_1-1,n_2) \geq 0, \\ w_2^C(n_1,n_2) - w_2^{NC}(n_1,n_2-1) \geq 0, \\ w_1^{NC}(n_1,n_2) - w_1^C(n_1+1,n_2) \geq 0, \\ w_2^{NC}(n_1,n_2) - w_2^C(n_1,n_2+1) \geq 0. \end{cases} \tag{8}$$

As the agreements are revised periodically, we focus our analysis only on the short term, that is the period $[0,T]$. Unfortunately, we are not able to solve the system analytically, due to the complexity and to the nonlinearity of the functions. So, we analyze the problem from a numerical point of view. We make some simulations using the software Wolfram Mathematica. The first step is the calibration of parameters in a benchmark model, that we summarize in Table 1.

Table 1. Benchmark parameters for developed (Type 1) and developing (Type 2) countries

	i=1	i= 2	
α_i	3.38	2.32	$\mu = 0.64$
β_i	0.0031	0.0048	$\delta = 0.0083$
s_i	0.08	0.1	$\rho = 0.01$

To calibrate the parameters of production and damage-cost functions we use the World Bank data set of emissions, expressed in kg per purchasing power parity (PPP) of Gross Domestic Product (GDP). We use the aggregate data of the developed countries and the upper middle income countries, within that we have all the developing countries with a significant industrial structure (e.g., BRICS). Based on this division of the World Bank, we set $N = 10$, with $N_1 = 6$ and $N_2 = 4$.

Moreover our choice of δ and μ is based on Nordhaus (1993). The parameters s_1 and s_2 in the function $Ext(n_1,n_2)$ are chosen sufficiently small not to overestimate the effect of the social externalities. Because by our simulation the starting level of pollution seems to be not relevant for the stability, we can assume $S_0 = 0$. We make two different studies, one with $\gamma = 0$ and one with $\gamma = 1$. In the following, we present only the case latter case, because it is more interesting for the sensitivity analysis and the case $\gamma = 0$ does not present substantial differences.

The first values that we have to compute are T^{NC} and T^C, taking every time a suitable value for $\overline{Y}_2$, which depends on the production function. As we expected, we have always $T^{NC} \leq T^C$, so we focus our simulation on the interval $[0,T^{NC}]$. After that, we solve the differential equations for $S(t)$ in the different configurations required by stability conditions. Having all the elements we need, we can proceed with the simulation of stability. We evaluate the 32 possible coalitions, due to the different combinations of variables n_1 and n_2.

The benchmark model gives an univocal result: the only stable coalition is the grand coalition. What we observed is that all possible coalitions are internally stable, but no one is externally stable, that means that all players have an incentive to join

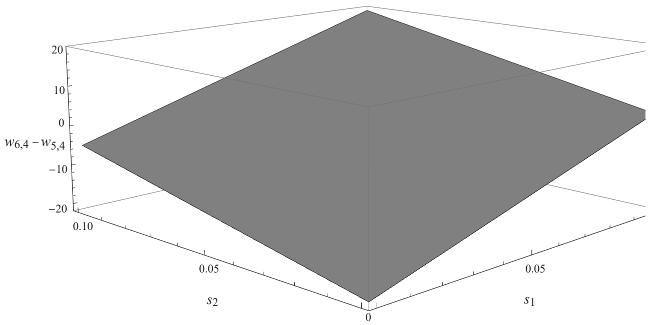


Figure 1. Internal stability for developed countries as function of s_1 and s_2, with $n_1 = 6$ and $n_2 = 4$

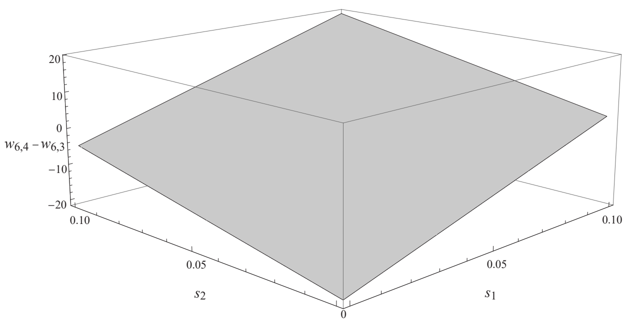


Figure 2. Internal stability for developing countries as function of s_1 and s_2, with $n_1 = 6$ and $n_2 = 4$

the agreement. Another thing we learn from this model is that the grand coalition stops being stable only if the parameters s_1 and s_2 converge to zero. In Figures 1 and 2 we represent the internal stability conditions of the system (8), as function of the two parameters s_1 and s_2.[1] We can see that the coalition is unstable if there are no social externalities or if at least one of the two groups gives little weight to non environmental possibilities given by the agreement. We make the same simulations on smaller coalition and the surfaces that we obtain have a shape very close to that of Figures 1 and 2.

Going forward with our sensitivity analysis, we pass to the parameters of the production function, α_1 and α_2. The results of the simulations say that, while there is no effect due to a change in α_1, the parameter α_2 influences the stability only in the measure of change of T^{NC}, but the grand coalition is still the only stable.

We obtain different results if we test the sensitivity of the model to variations in damage-cost parameters, β_1 and β_2. In this case an increment of the vulnerability to

[1] Clearly, we have not external stability conditions in case of Grand Coalition. We make some simulation on external stability for smaller coalition, and we found that the qualitative results are symmetric to these.

Table 2. Stable coalitions for simultaneous variation of β_1 and β_2

	Marginal costs sensitivity		
	$\beta_1 = 0.0062$ $\beta_2 = 0.0024$	$\beta_1 = 0.0016$ $\beta_2 = 0.0096$	$\beta_1 = 0.0062$ $\beta_2 = 0.0096$
Stable coalitions	$(5,0)$	$(0,4)$	$(0,4)$ $(5,0)$ $(1,3)$ $(2,2)$ $(4,1)$

the damage of pollution can make large coalitions unstable. The first test we make in this sense in taking a double value of β_1, that is $\beta_1 = 0.0062$, given the values of the other parameters. This increment brings to a situation of homogeneity: we still have a unique stable solution, but now the coalition is formed only by developed countries, in fact the solution couple is $(n_1, n_2) = (5,0)$.

The next step is the evaluation of an increment in the value of β_2, so we take $\beta_2 = 0.0096$, given the other parameters as in the benchmark model. The results are symmetric to the previous one: We have only one stable and homogeneous coalition, but in this case it is the one formed only by developing countries, that is $(n_1, n_2) = (0,4)$. The difference is that this coalition consists of all the developing countries, while above we have a coalition formed by all but one developed countries.

We also test the effects of a simultaneous variations of parameters β_1 and β_2. We first try to balance out the growth of one parameters with a reduction of the other. In the first column of the Table 2, we show the results on the stability in the case that β_1 is doubled and β_2 is the half. As we can see, reduction of the latter doesn't bring some compensation in terms of stable coalitions and the only stable coalition is still $(n_1, n_2) = (5,0)$. The same thing happens in the second column of Table 2, in which β_1 decreases and β_2 increases. Also in this case, the only stable coalition is $(n_1, n_2) = (0,4)$. Then, decreasing one of the marginal damage costs can not compensate in any an increase of the other damage cost. More interesting the last column, in which both parameters are doubled. In this case we have an enlargement of the set of stable solutions, which now includes also the mixed coalitions $(1,3)$, $(2,2)$ and $(4,1)$.

5. Conclusions

To develop efficient policies on pollution control, stability of an International Environmental Agreement is the key. In this work we investigated a non-cooperative N-player differential game, in which we divided the world in two types of players, developed and developing countries. For the latter we assumed an evolving damage-cost function, taking into account the particular issues of this countries. Our main contribution is to recall the idea of social externalities, investigated by Cabon-Dhersin and Ramani (2006) in a static context with symmetric players, and to verify its validity in a more general and realistic context.

First of all we characterized the emission solutions both for signatories and defectors, assuming as a date the size of the coalition. After that, we used these solutions to study the stability of the agreement. Due to the non-linearity and the complexity of the model, we analyzed the problem with numerical simulations.

Our benchmark model shows that we can have just one stable coalition, namely a coalition of all players. Moreover, this coalition become unstable only if the parameters s_1 and s_2 converge to zero. Nevertheless, there is a significant sensitivity to the parameters β_1 and β_2, so if they increase, we have smaller stable coalitions. To conclude, we want to highlight two possible ways to extend the model. The first point is to consider a feedback game also for the membership, including the possibility that players join the agreement in different times. The second one is to consider a higher degree of asymmetry, going beyond the assumption of homogeneity within the two groups of countries.

Acknowledgment We would like to thank Fabio Di Dio (SOGEI SPA) for his valuable comments and suggestions.

References

Barrett, S. (1994). Self-Enforcing International Environmental Agreements. *Oxford Economic Papers*, 46, 878–894.

Botteon, M. and Carraro, C. (1994). Strategies for Environmental Negotiations: Issue Linkage with Heterogeneous Countries. In Hanley, N., Folmer, H. (eds.), *Game Theory and Environment*. Cheltenham, Edward Elgar, 181–203.

Breton, M., Sbragia, L. and Zaccour, G. (2010). A Dynamic Model for International Environmental Agreements. *Environmental and Resource Economics*, 45, 25–48.

Cabon-Dhersin, M. L. and Ramani, S. V. (2006). Can Social Externalities Solve the Small Coalition Puzzle in International Environmental Agreements?. *Economics Bulletin*, 17(4), 1–8.

Carraro, C. and Siniscalco, S. (1993). Strategies for International Protection of the Environment. *Journal of Public Economics*, 52, 309–328.

Casino, B. and Rubio, S. J. (2005). Self-Enforcing International Environmental Agreements with a Stock Pollutant. *Spanish Economic Review*, 7, 89–109.

Chakrabarti, S., Gilles, R. P. and Lazarova, E. A. (2011). Strategic Behaviour under Partial Cooperation. *Theory and Decision*, 71(2), 175–193.

d'Aspremont, C., Jacquemin, A., Gabszewicz, J. J. and Weymark, J. A. (1983). On the Stability of Collusive Price Leadership. *The Canadian Journal of Economics*, 16, 17–25.

Diamantoudi, E. and Sartzetakis, E. S. (2006). Stable International Environmental Agreements: An Analytical Approach. *Journal of Public Economic Theory*, 8, 247–263.

Eyckmans, J. and Finus, M. (2006). A Coalition Formation in a Global Warming Game: How the Design of Protocols Affects the Success of Environmental Treaty-Making. *Natural Resource Modeling*, 19(3), 323–358.

Finus, M. (2001). *Game Theory and International Environmental Cooperation*. Cheltenham, Edward Elgar.

Finus, M. (2003). Stability and Design of International Environmental Agreements: The Case of Transboundary Pollution. In Folmer, H. and Tietenberg, T. (eds.), *The International Yearbook of Environmental and Resource Economics 2003/2004: A Survey of Current Issues*. Cheltenham, Edward Elgar, 82–158.

Fuentes-Albero, C. and Rubio, S. (2010). Can International Environmental Cooperation Be Bought? *European Journal of Operational Research*, 202, 255–264.

Hoel, M. and Schneider, K. (1997). Incentives to Participate in an International Environmental Agreement. *Environmental and Resource Economics*, 9, 153–170.

Hübler, M. and Finus, M. (2013). Is the Risk of North-South Technology Transfer Failure an Obstacle to a Cooperative Climate Change Agreement? *International Environmental Agreements: Politics, Law and Economics*, 13(4), 461–479.

Jørgensen, S., Martín-Herrán, G. and Zaccour, G. (2010). Dynamic Games in the Economics and Management of Pollution. *Environmental Modeling and Assessment*, 15(6), 433–467.

Labriet, M. and Loulou, R. (2003). Coupling Climate Changes and GHG Abatement Costs in a Linear Programming Framework. *Environmental Modeling and Assessment*, 8, 261–274.

Le Breton, M. and Souberyan, A. (1997). The Interaction between International Environmental and Trade Policies. In C. Carraro (ed.), *International Environmental Negotiations – Strategic Policy Issues*. Cheltenham, Edward Elgar, 126–149.

Long, N. V. (2010). *Dynamic Games in Economics: A Survey*. Singapore, World Scientific Publishing Company.

Mallozzi, L. and Tijs, S. (2008). Conflict and Cooperation in Symmetric Potential Games. *International Game Theory Review*, 10(3), 1–12.

Mallozzi, L. and Tijs, S. (2009). Coordinating Choice in Partial Cooperative Equilibrium. *Economics Bulletin*, 29(2), 1467–1473.

Masoudi, N. and Zaccour, G. (2013). A Differential Game of International Pollution Control with Evolving Environmental Costs. *Environment and Development Economics*, 18(6), 680–700.

Nordhaus, W. D. (1993). Rolling the Dice: An Optimal Transition Path for Controlling Greenhouse Gases. *Resource and Energy Economics*, 15(1), 27–50.

Pavlova, Y. and de Zeeuw, A. (2013). Asymmetries in International Environmental Agreements. *Environment and Development Economics*, 18, 51–68.

Rubio, S. J. and Ulph, A. (2007). An Infinite-Horizon Model of Dynamic Membership of International Environmental Agreements. *Journal of Environmental Economics and Management*, 54, 296–310.

Shafik, N. and Bandyopadhyay, S. (1992). Economic Growth and Environmental Quality: Time Series and Cross-Country Evidence. Policy, research working papers No. 904. World development report. Washington, DC, World Bank.

von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behaviour*. Princeton, Princeton University Press.

Appendix

A. Emissions of defectors

To solve the problem, we use dynamic programming. For given time T^{NC}, we proceed backward, solving first the problem on $[T^{NC},\infty)$. We have to solve first:

$$\max_{e_i^{NC}} \int_{T^{NC}}^{\infty} \left(\alpha_i e_i^{NC} - \frac{1}{2}(e_i^{NC})^2 - \beta_i S\right) e^{-\rho(t-T^{NC})} dt,$$

$$\text{s.t.} \quad \dot{S}(t) = \mu\left(\sum_1^N e_i(t)\right) - \delta S(t), \;\; S(0) = S_0.$$

If we denote with $v(t,S)$ the value function of the problem, we can write the Hamilton-Jacobi-Bellman (HJB) equation as

$$-v_t = \max_{e_i^{NC}} \left\{ \left(\alpha_i e_i - \frac{1}{2}(e_i^{NC})^2 - \beta_i S\right) e^{-\rho t} + v_S\left(\mu \sum_{j=1}^N e_j - \delta S\right)\right\}. \tag{9}$$

Solving the first order conditions in (9), we obtain an expression for the optimal emissions

$$e_i^{NC}(t) = \alpha_i + \mu v_S e^{\rho t}.$$

Let us assume that value function $v(t,S)$ is linear in S:

$$v(t,S) = (KS+L)e^{-\rho t},$$

with partial derivatives $v_t = -\rho(KS+L)e^{-\rho t}$ and $v_S = Ke^{-\rho t}$. So, emissions for a player i outside coalition are given by

$$e_i^{NC}(t) = \alpha_i + K\mu.$$

In order to find an expression for the parameter K, we substitute v_t, v_S and e_i inside (9):

$$\rho(KS+L)e^{-\rho t} = \left[\alpha_i(\alpha_i + K\mu) - \frac{1}{2}(\alpha_i + K\mu)^2 - \beta_i S\right] e^{-\rho t} + K\left(\mu \sum_{j=1}^N e_j - \delta S\right) e^{-\rho t}$$

$$\rho(KS+L) = \alpha_i^2 + \alpha_i K\mu - \frac{1}{2}\alpha_i^2 - \frac{1}{2}K^2\mu^2 - \alpha_i K\mu - \beta_i S + K\mu \sum_{j=1}^N e_j - K\delta S$$

$$\rho KS + \rho L = -(\beta_i + K\delta)S + \frac{1}{2}\alpha_i^2 - \frac{1}{2}K^2\mu^2 + K\mu \sum_{j=1}^N e_j.$$

By the principle of identity of polynomials, we can write the equation

$$\rho K = -\beta_i - K\delta,$$

from which we have that $K = -\frac{\beta_i}{\rho+\delta}$.

Finally, we have the expression of the emissions of defectors

$$e_i^{NC}(t) = \alpha_i - \mu \frac{\beta_i}{\rho+\delta}, \qquad i = 1,2. \tag{10}$$

We proceed now to find the optimal emissions for $t \in [0, T^{NC}]$. Since the functional forms for players of type 1 is the same in the entire period, we have that the emissions of developed countries are the same in $[0, T^{NC}]$:

$$e_1^{NC}(t) = \alpha_1 - \mu \frac{\beta_1}{\rho+\delta}, \quad \forall t \geq 0.$$

For developing countries we have to consider the different damage-cost function, so the problem is:

$$\max_{e_2^{NC}} \int_0^{T^{NC}} \left(\alpha_2 e_2^{NC} - \frac{1}{2}(e_2^{NC})^2 - \gamma \frac{t}{T^{NC}} \beta_2 S \right) e^{-\rho t} dt,$$

$$\text{s.t.} \quad \dot{S}(t) = \mu \left(\sum_1^N e_i(t) \right) - \delta S(t), \;\; S(0) = S_0.$$

In this case the HJB equation is given by

$$-v_t = \max_{e_2} \left\{ \left(\alpha_2 e_2 - \frac{1}{2} e_2^2 - \gamma \frac{t}{T^{NC}} \beta_2 S \right) e^{-\rho t} + v_S \left(\mu \sum_{j=1}^N e_j - \delta S \right) \right\}. \tag{11}$$

As usual, we derive the first order conditions from maximization in (11)

$$(\alpha_2 - e_2) e^{-\rho t} + v_s \mu = 0 \;\Rightarrow\; e_2^{NC}(t) = \alpha_2 + v_S \mu e^{\rho t}.$$

For the value function $v(t,S)$ we assume

$$v(t,S) = [x(t)S + y(t)] e^{-\rho t},$$

whose partial derivatives are:

$$v_t = [(x'(t) - \rho x(t))S + y'(t) - \rho y(t)] e^{-\rho t} \;\text{ and }\; v_S = x(t) e^{-\rho t}.$$

The optimal emissions are given by

$$e_2^{NC}(t) = \alpha_2 + \mu x(t).$$

Now, we have to substitute v_t, v_S and e_2 inside (11)

$$-[(x'(t)-\rho x(t))S+y'(t)-\rho y(t)]e^{-\rho t}=\left[\alpha_2(\alpha_2+\mu x(t))-\frac{1}{2}(\alpha_2+\mu x(t))^2-\gamma\frac{t}{T^{NC}}\beta_2 S\right]e^{-\rho t}+$$
$$+x(t)\left(\mu\sum_{j=1}^{N}e_j-\delta S\right)e^{-\rho t}.$$

Rearranging with respect to the principle of identity of polynomials, we can write the differential equation

$$\begin{cases}-x'(t)+(\rho+\delta)x(t)=-\gamma\frac{t}{T^{NC}}\beta_2,\\ x(T^{NC})=-\frac{\beta_2}{\rho+\delta}.\end{cases}$$

The solution $x(t)$ is given by

$$x(t)=-\frac{\beta_2}{T^{NC}(\rho+\delta)^2}\left[\gamma\big(1+t(\rho+\delta)-e^{(\rho+\delta)(t-T^{NC})}\big)+\right.$$
$$\left.+e^{(\rho+\delta)(t-T^{NC})}T^{NC}(\rho+\delta)(1-\gamma)\right],$$

and leads us to the developing countries' emissions, for $t\in[0,T^{NC}]$:

$$e_2^{NC}(t)=\alpha_2-\mu\frac{\beta_2}{T^{NC}(\rho+\delta)^2}\left[\gamma\big(1+t(\rho+\delta)-e^{(\rho+\delta)(t-T^{NC})}\big)+\right.$$
$$\left.+T^{NC}(1-\gamma)(\rho+\delta)e^{(\rho+\delta)(t-T^{NC})}\right]. \tag{12}$$

B. Emissions of signatories

As for the defectors' case, we assume as known the time T^C and we proceed backward. First of all solve the problem for $t\in[T^C,\infty)$:

$$\max_{e_1^C,e_2^C}\int_{T^C}^{\infty}\left[n_1\left(\alpha_1 e_1^C-\frac{1}{2}(e_1^C)^2\right)+n_2\left(\alpha_2 e_2^C-\frac{1}{2}(e_2^C)^2\right)-n_1\beta_1 S-n_2\beta_2 S+\right.$$
$$\left.+s_1n_1+s_2n_2\right]e^{-\rho(t-T^C)}dt,$$

$$\text{s.t. } \dot{S}(t)=\mu\left(n_1e_1^C(t)+n_2e_2^C(t)+(N_1-n_1)e_1^{NC}(t)+(N_2-n_2)e_2^{NC}(t)\right)-\delta S(t),$$
$$S(0)=S_0.$$

We want to highlight that the externalities have no effects on the determination of the emissions. Let $v(t,S)$ be the value function, the HJB equation is

$$-v_t = \max_{e_1^C, e_2^C} \Bigg\{ \left[n_1 \left(\alpha_1 e_1^C - \frac{1}{2}(e_1^C)^2 \right) + n_2 \left(\alpha_2 e_2^C - \frac{1}{2}(e_2^C)^2 \right) - n_1\beta_1 S - n_2\beta_2 S \right] e^{-\rho t} +$$
$$+ v_S \left[\mu \left(n_1 e_1^C + n_2 e_2^C + (N_1 - n_1) e_1^{NC} + (N_2 - n_2) e_2^{NC} \right) - \delta S \right] + (s_1 n_1 + s_2 n_2) e^{-\rho t} \Bigg\}. \tag{13}$$

The first order conditions in (13) are given by

$$n_1[(\alpha_1 - e_1^C)e^{-\rho t} + \mu v_S] = 0,$$
$$n_2[(\alpha_1 - e_2^C)e^{-\rho t} + \mu v_S] = 0,$$

from which we can derive the expressions for emissions

$$e_1^C(t) = \alpha_1 + \mu v_S e^{\rho t}, \qquad e_2^C(t) = \alpha_2 + \mu v_S e^{\rho t}.$$

The steps are the same: we choose a guess for value function, then we substitute its partial derivatives and the emissions in equation (13). So, take

$$v(t,S) = (AS + B)e^{-\rho t},$$

whose partial derivatives are: $v_t = -\rho(AS+B)e^{-\rho t}$ and $v_S = Ae^{-\rho t}$. Then optimal emissions are given by

$$e_1^C(t) = \alpha_1 + \mu A, \qquad e_2^C(t) = \alpha_2 + \mu A.$$

Substituting in (13), we can derive an expression for parameter A

$$\rho A = -n_1\beta_1 - n_2\beta_2 - A\delta,$$

so that, $A = -\frac{n_1\beta_1 + n_2\beta_2}{\rho+\delta}$. Then, the emission solutions for cooperative countries, for $t \in [T^C, \infty)$, are given by

$$e_i^C(t) = \alpha_i - \mu \frac{n_1\beta_1 + n_2\beta_2}{\rho + \delta}, \quad i = 1,2. \tag{14}$$

Now, we can proceed to find the feedback Nash equilibrium for signatory players in the interval $[0, T^C]$. The different damage-cost function implies that the cooperative solutions solve:

$$\max_{e_1^C, e_2^C} \int_0^{T^C} \left[n_1 \left(\alpha_1 e_1^C - \frac{1}{2}(e_1^C)^2 \right) + n_2 \left(\alpha_2 e_2^C - \frac{1}{2}(e_2^C)^2 \right) - n_1\beta_1 S - n_2\gamma \frac{t}{T^C} \beta_2 S + \right.$$
$$\left. + s_1 n_1 + s_2 n_2 \right] e^{-\rho t} dt,$$

$$\text{s.t. } \dot{S}(t) = \mu\left(n_1 e_1^C(t) + n_2 e_2^C(t) + (N_1 - n_1)e_1^{NC}(t) + (N_2 - n_2)e_2^{NC}(t)\right) - \delta S(t),$$
$$S(0) = S_0.$$

If we denote with $v(t,S)$ the value function, than the HJB equation is

$$\begin{aligned}-v_t = \max_{e_1^C, e_2^C}\Bigg\{&\left[n_1\left(\alpha_1 e_1^C - \frac{1}{2}(e_1^C)^2\right) + n_2\left(\alpha_2 e_2^C - \frac{1}{2}(e_2^C)^2\right) - n_1\beta_1 S - n_2\gamma\frac{t}{T^C}\beta_2 S\right]e^{-\rho t} + \\ &+ v_S\left[\mu\left(n_1 e_1^C + n_2 e_2^C + (N_1 - n_1)e_1^{NC} + (N_2 - n_2)e_2^{NC}\right) - \delta S\right] + \\ &+ (s_1 n_1 + s_2 n_2)e^{-\rho t}\Bigg\}.\end{aligned} \tag{15}$$

As usual, we compute the first order conditions in (15), to obtain a characterization for signatories' emissions. So

$$n_1[(\alpha_1 - e_1^C)e^{-\rho t} + v_S\mu] = 0,$$
$$n_2[(\alpha_2 - e_2^C)e^{-\rho t} + v_S\mu] = 0,$$

from which $e_i^C(t) = \alpha_i - v_S\mu$, $i = 1,2$.

We need to give a guess for value function, and we choose, as in the previous cases, a linear function of S

$$v(t,S) = [g(t)S + z(t)]e^{-\rho t},$$

whose partial derivatives with respect to t and S are

$$v_t = [(g'(t) - \rho g(t))S + z'(t) - z(t)]e^{-\rho t}, \qquad v_S = g(t)e^{-\rho t}.$$

The expression of v_S gives us the emission solutions

$$e_i^C = \alpha_i + \mu g(t), \qquad i = 1,2.$$

To conclude the computation of the Nash equilibrium, we need to find an expression for the function $g(t)$. The way is to substitute v_t, v_S and e_i, $i = 1,2$, inside (15). With some algebra, and because of the continuity of value function, we have to solve the dynamical system

$$\begin{cases} -g'(t) + (\rho + \delta)g(t) = n_1\beta_1 + n_2\gamma\frac{t}{T^C}\beta_2, \\ g(T^C) = -\dfrac{n_1\beta_1 + n_2\beta_2}{\rho + \delta}. \end{cases}$$

The system has a unique solution $g(t)$, as follows

$$g(t) = -\frac{n_1\beta_1}{\rho+\delta} - \frac{n_2\beta_2}{T^C(\rho+\delta)^2}\Big[\gamma\big(1+t(\rho+\delta)-e^{(\rho+\delta)(t-T^C)}\big)+ \\ +T^C(1-\gamma)(\rho+\delta)e^{(\rho+\delta)(t-T^C)}\Big].$$

Finally, the emissions for signatory players, when $t \in [0, T^C]$, are given by:

$$e_i^C(t) = \alpha_i - \mu\frac{n_1\beta_1}{\rho+\delta} - \mu\frac{n_2\beta_2}{T^C(\rho+\delta)^2}\Big[\gamma\big(1+t(\rho+\delta)-e^{(\rho+\delta)(t-T^C)}\big)+ \\ +T^C(1-\gamma)(\rho+\delta)e^{(\rho+\delta)(t-T^C)}\Big],$$

where $i = 1, 2$.

Czech Economic Review 9 (2015) 155–168
Acta Universitatis Carolinae Oeconomica

The Iterative Nature of a Class of Economic Dynamics

Shilei Wang*

Received 14 October 2014; Accepted 6 May 2016

Abstract This work aims to demonstrate a rather specific "iterative nature" existing in a class of regular economic dynamics by revisiting two typical economic concepts as informative examples, viz., random utility and stochastic growth. We begin with a formal treatment of discrete dynamical system and its popular derivation, iterated function system, so that a solid foundation could be laid for our analysis of economic dynamics. Two economic systems afterwards are constructed to show how random utility function and stochastic growth in a classical economy could be essentially driven by some iterative elements. Besides, our analyses also implicitly show that a quite complex economic dynamics carrying substantial randomness could basically originate in some fairly simple dynamic principles.

Keywords Dynamical system, iterated function system, random utility function, stochastic growth, chaos
JEL classification C61, D99

1. Introduction

The present paper deals with economic dynamics in a very specific way with a quite general objective yet, that is, characterizing some critical and widely existent nature in a class of economic dynamics. A number of somewhat popular terms are usually adopted to describe economic dynamics, say "complex" and "chaotic" (cf., Goodwin 1990; Lorenz 1993; Tu 1994; Day 1994, 1999), which both convey that the basic mechanism of economic dynamics should be in essence highly hard to capture. In this work, we do not plan to argue this viewpoint, however do plan to see its negation, that some fairly simple economic principles could also generate complex or even chaotic properties.

The economic science used to study static models, and discuss their equilibria and comparative statics thereof. That being said, a great number of dynamic models have been developed, such as bifurcation phenomena in a delayed demand-supply system (cf., Leontief 1934; Kaldor 1934; Ezekiel 1938), chaotic properties in models of optimal economic growth (cf., Day 1983; Benhabib and Nishimura 1985; Boldrin and Montrucchio 1986), and nowadays many investigations on financial market dynamics. Evidently, the literature on economic dynamics, nonlinearity, and complexity is vast and also tends to be diverse, yet there is a lack of closely relevant ones to this article and hence we shall pass such potential references directly to our writing.

* Department of Economics, Università Ca' Foscari Venezia, 30121 Venice, Italy. E-mail: shilei.wang@unive.it.

The technical foundations are written in Section 2 and 3, and they are followed by two economic systems which in some sense are artificial. In Section 4, we study random utility function, and show different approaches of randomness aggregation in time preference. In Section 5, a classical economy driven by consumption and production is reconsidered. We show that a multiplicative shock in that economy could produce a stochastic growth which is determined equivalently by an iterated linear function system on one scaled real interval.

2. Discrete dynamical system

Throughout this article, we will use $\mathbb{R}^+$ to mean the nonnegative real numbers, and use $\mathbb{Z}^+$ and $\mathbb{Z}^-$ to mean the nonnegative and nonpositive integers. For any sets X and Y, $X \times Y$ denotes their Cartesian product. Let the state space and time domain be X and $\mathbb{Z}$, respectively. Suppose the state space X is a metric space with a metric $d : X \times X \to \mathbb{R}^+$.

Definition 1. A *discrete dynamical system* on X is a pair (X, f) with $x_{n+1} = f(x_n)$ for all $x_n, x_{n+1} \in X$ and all $n \in \mathbb{Z}$, where $f : X \to X$ is of class C^0.

The trajectory passing through a state $x \in X$ is

$$\gamma(x) = \{f^n(x) : n \in \mathbb{Z}\}, \tag{1}$$

and its positive and negative semi-trajectories are

$$\gamma^+(x) = \{f^n(x) : n \in \mathbb{Z}^+\}, \quad \gamma^-(x) = \{f^n(x) : n \in \mathbb{Z}^-\}.$$

Evidently, the positive semi-trajectory $\gamma^+(x)$ also represents the motion starting from the state x.

A state x is an *equilibrium state*, if $\gamma(x) = \{x\}$, or $f(x) = x$. A state y is an *ω-limit state* for an initial state x if $\lim_{n\uparrow+\infty} f^n(x) = y$, and the set of all such ω-limit states is called the *ω-limit set* of x, and denoted by $\omega(x)$. A set of states $S \subseteq X$ is *invariant* if $f(S) = S$. Note that any nonempty ω-limit set should be invariant, and thus we have $f\big(\omega(x)\big) = \omega(x)$ for all $\omega(x) \neq \emptyset$.

A set of states $A \subseteq X$ is an *attractor*, if there is a neighborhood $N(A,\varepsilon)$ such that $f\big(N(A,\varepsilon)\big) \subseteq N(A,\varepsilon)$ and

$$\omega\big(N(A,\varepsilon)\big) = \bigcap_{n\in\mathbb{Z}^+} f^n\big(N(A,\varepsilon)\big) = A,$$

but no proper subset of A has such properties (cf., Milnor 1985).

A state x (and also the motion $\gamma^+(x)$) is *periodic*, if there is a $k \in \mathbb{Z}^+$ such that $f^k(x) = x$, and the minimal $k \in \mathbb{Z}^+$ satisfying $f^k(x) = x$ is the *period* of $\gamma^+(x)$. If the period of $\gamma^+(x)$ is 1, then $f(x) = x$, and thus x is actually an equilibrium state. If the period of $\gamma^+(x)$ is $p < +\infty$, then

$$\gamma^+(x) = \big\{x, f(x), \ldots, f^p(x)\big\}.$$

A state x is called *finally periodic*, if there is an $m \in \mathbb{Z}^+$ such that $f^n(x)$ is a periodic state for all $n \geq m$, or equivalently stating, there is some $p \in \mathbb{Z}^+$ such that $f^{n+p}(x) = f^n(x)$ for all $n \geq m$. A state x is called *asymptotically periodic*, if there is a $y \in X$ such that

$$\lim_{n \uparrow +\infty} d\big(f^n(x), f^n(y)\big) = 0. \tag{2}$$

In case the state space $X \subseteq \mathbb{R}$ and it is compact, one would have the following theorem:

Theorem 1. (Li and Yorke 1975) *Suppose X is an interval in $\mathbb{R}$, and $f : X \to X$ is of class C^0. If there exists a motion of period 3 in (X, f), viz., there are three distinct states $x, y, z \in X$ such that $f(x) = y$, $f(y) = z$, and $f(z) = x$, then there is some motion of period n in (X, f) for all $n \in \mathbb{N}$.*

Proof. Let $<_S$ denote Šarkovskii's order on $\mathbb{N}$, then we have

$$3 <_S 5 <_S 7 <_S \cdots <_S 2^n <_S 2^{n-1} <_S \cdots <_S 2^2 <_S 2 <_S 1.$$

By Šarkovskii's (1964) theorem, if (X, f) has a motion of period m, then it must have some motion of period m' with $m <_S m'$. Since $3 <_S n$ for all $n \neq 3$, and there is a motion of period 3 in (X, f), the statement will thus directly follow. □

A generic dynamical system is chaotic if its dynamics sensitively depend on the initial state, and its states are transitive. For the moment, a discrete dynamical system (X, f) is called *chaotic* if it satisfies

(i) for all $x \in X$ and any $\varepsilon > 0$, there is a $\delta > 0$ such that $d\big(f^n(x), f^n(y)\big) > \varepsilon$ for all $y \in N(x, \delta)$ and some $n \in \mathbb{Z}^+$,

(ii) for all $S_1, S_2 \subseteq X$, there is an $n \in \mathbb{Z}^+$ such that $f^n(S_1) \cap S_2 \neq \emptyset$.

In particular, when $X \subseteq \mathbb{R}$ is compact, and f is of class C^0, an alternative definition of chaos can be proposed in the sense of Li and Yorke (1975).

Definition 2. A discrete dynamical system (X, f) is *nonperiodically chaotic*, if there is an uncountable set $S \subseteq X$ such that

(i) $\limsup_{n \uparrow +\infty} d\big(f^n(x), f^n(y)\big) > 0$ for all distinct $x, y \in S$,

(ii) $\liminf_{n \uparrow +\infty} d\big(f^n(x), f^n(y)\big) = 0$ for all distinct $x, y \in S$,

(iii) for all $z \in X$ periodic, $\limsup_{n \uparrow +\infty} d\big(f^n(x), f^n(z)\big) > 0$ for all $x \in S$.

It might be noticed that nonperiodic chaos is a slightly weaker concept than chaos itself. That's to say, if a discrete dynamical system on $X \subseteq \mathbb{R}$ is chaotic, then it must be nonperiodically chaotic as well; but if a discrete dynamical system is nonperiodically chaotic, it may not be chaotic.

The reason behind such an assertion is constructive. It (X, F) is nonperiodically chaotic, then there is at most one asymptotically periodic state in S. Now suppose a state $u \in X$ is not asymptotically periodic, then $\omega(u)$ should have infinitely many states. Let $V \subseteq \omega(u)$ be the (minimally invariant) kernel of $\omega(u)$, and suppose there is some $v \in X$ such that $V = \omega(v)$, which hence again contains infinitely many states. Let $U = X \setminus V$, then $f^n(V) \cap U = \emptyset$ for all $n \in \mathbb{Z}^+$, and therefore V and U are not transitive, which then implies (X, f) is not chaotic.

3. Iterated function system

Let's now consider a collection of contractive functions defined on the state space X with the metric d. Here, a function $f : X \to X$ is called *contractive*, if there is a $\lambda \in (0,1)$ such that $d\big(f(x), f(y)\big) \leq \lambda d(x,y)$ for all $x, y \in X$.

Let I_N denote an index set with N elements for $N \geq 2$ finite. Let

$$F = \{f_i : i \in I_N\},$$

where $f_i : X \to X$ is contractive and of class C^0 for all $i \in I_N$.

Definition 3. The pair (X, F) is called an *iterated function system*, if (X, f_i) is a discrete dynamical system for all $i \in I_N$.

Suppose X is compact, and let $\mathcal{Q}(X)$ denote the collection of all the nonempty compact subsets of X. Then $\mathcal{Q}(X)$ with the Hausdorff metric d_H is a compact metric space, where the Hausdorff metric d_H on $\mathcal{Q}(X)$ can be defined by the metric d on X, i.e., for all $U, V \in \mathcal{Q}(X)$

$$d_H(U,V) = \sup_{u \in U,\, v \in V} \big\{d(u,V), d(v,U)\big\},$$

in which $d(u,V) = \inf_{v \in V} d(u,v)$ and $d(v,U) = \inf_{u \in U} d(v,u)$.

Define a mapping $H : \mathcal{Q}(X) \to \mathcal{Q}(X)$, such that for all $B \in \mathcal{Q}(X)$,

$$H(B) = \bigcup_{i \in I_N} f_i(B). \tag{3}$$

Here, H is called the *Hutchinson operator* (Hutchinson 1981). Moreover, we define H^n by the recursion $H^n = H \circ H^{n-1}$ with $H^0 = \mathrm{id}_{\mathcal{Q}(X)}$, where $n \in \mathbb{Z}$ and $\mathrm{id}_{\mathcal{Q}(X)}$ denotes the identity mapping on $\mathcal{Q}(X)$.

Definition 4. $A \in \mathcal{Q}(X)$ is called an *attractor* of (X, F), if there is a neighborhood $N(A, \varepsilon) \in \mathcal{Q}(X)$ such that

$$H\big(N(A,\varepsilon)\big) \subseteq N(A,\varepsilon), \qquad \bigcap_{n \in \mathbb{Z}^+} H^n\big(N(A,\varepsilon)\big) = A,$$

and no proper subset of A in $\mathcal{Q}(X)$ has such properties.

Theorem 2. *(X, F) has a unique attractor A with $H(A) = A$.*

Proof. For all $f_i \in F$, there is a $\lambda_i \in (0,1)$ such that for all $x, y \in X$,

$$d\big(f_i(x), f_i(y)\big) \leq \lambda_i d(x,y).$$

Let $\lambda = \max_{i \in I_N} \lambda_i$, then $\lambda \in (0,1)$ as well. Note that for all $U, V \in \mathcal{Q}(X)$ we have

$$d_H\big(H(U), H(V)\big) \leq \sup_{i \in I_N} d_H\big(f_i(U), f_i(V)\big) \leq \sup_{i \in I_N} \lambda_i d_H(U,V) \leq \lambda d_H(U,V),$$

thus by the Banach fixed point theorem, there is a unique $A \in \mathcal{Q}(X)$ such that $H(A) = A$, and $\lim_{n\uparrow+\infty} H^n(B) = A$ for all $B \in \mathcal{Q}(X)$. And clearly, there exists a neighborhood $N(A,\varepsilon) \in \mathcal{Q}(X)$ serving as a basin of A.

We then show that any $B \neq A$ in $\mathcal{Q}(X)$ can not be an attractor of (X,F), which would imply A is the unique attractor of (X,F), and thus completes our proof. First of all, any $B \supset A$ can not be an attractor of (X,F), as for all $\varepsilon > 0$

$$\bigcap_{n \in \mathbb{Z}^+} H^n\big(N(B,\varepsilon)\big) \subseteq A \subset B.$$

Next, any $B \subset A$ also can not be an attractor of (X,F), otherwise we would have

$$\lim_{n\uparrow+\infty} H^n\big(N(B,\varepsilon)\big) = B \subset A,$$

a contradiction. □

Now consider the space I_N^ω, and for all $\mu \in I_N^\omega$ we write

$$\mu = (\mu_n,\ n \in \mathbb{N}) = (\mu_1, \mu_2, \ldots, \mu_\omega),$$

where $\mu_n \in I_N$ for all $n \in \mathbb{N}$. The Baire metric between all $\mu, \upsilon \in I_N^\omega$ is

$$d_B(\mu, \upsilon) = 2^{-m},$$

where $m = \min\{n \in \mathbb{N} : \mu_n \neq \upsilon_n\}$. Clearly, (I_N^ω, d_B) is a compact metric space. Let's define a mapping $C : I_N^\omega \times \mathcal{Q}(X) \to \mathcal{Q}(X)$, such that for all $\mu \in I_N^\omega$ and $S \in \mathcal{Q}(X)$,

$$C(\mu, S) = \bigcap_{n \in \mathbb{N}} f_{\mu_\omega} \circ \cdots \circ f_{\mu_{n+1}} \circ f_{\mu_n}(S). \tag{4}$$

Note in addition that the motion of any state $x \in S$ can be expressed as

$$\gamma^+(x) = \{f_{\mu_n} \circ \cdots \circ f_{\mu_2} \circ f_{\mu_1}(x) : n \in \mathbb{N}\}.$$

Suppose $B(A) = N(A,\varepsilon)$ for some $\varepsilon > 0$ is a basin of the attractor A, then for all $S \subseteq B(A)$ and $\mu \in I_N^\omega$, we have $C(\mu, S) \subseteq A$, and hence one can write

$$C\big(I_N^\omega, B(A)\big) = A. \tag{5}$$

It therefore suggests that the attractor A of (X,F) could be practically attained by all the ω-permutations of the transition rules in F.

Suppose there is some probability measure on I_N^{ω}, and in particular, we shall assume it is stationary, so that it can be fully characterized by a discrete probability measure on I_N. Let $\pi : I_N \to [0,1]$ denote such a probability measure, which satisfies $\sum_{i\in I_N} \pi(i) = 1$. As a consequence, at any time a function f_i stands out in F with a probability $\pi(i)$ for all $i \in I_N$.

Definition 5. The triplet (X,F,π) is called an *iterated random function system*.

Let σ_n be a random variable, such that $\mathrm{Prob}(\sigma_n = i) = \pi(i)$ for all $i \in I_N$. The transition function at a time $n \in \mathbb{Z}$ can thus be denoted by a randomly indexed function f_{σ_n}. Let a random variable Z_n denote the stochastic state in the system (X,F,π) at the time $n \in \mathbb{Z}$, then we have

$$Z_{n+1} = f_{\sigma_{n+1}}(Z_n). \tag{6}$$

Suppose the initial time is 0, and the initial state is $x \in X$, then the random motion can be written as

$$\Gamma^{+}(x) = \{Z_n : n \in \mathbb{Z}^{+}\},$$

in which $Z_0 = x$, $Z_1 = f_{\sigma_1}(x)$, and $Z_n = f_{\sigma_n}(Z_{n-1})$ for all $n \geq 2$.

Note that the stochastic process $(Z_n,\ n \in \mathbb{N})$ is in effect a Markov chain, and it is equivalent to the iterated random function system (X,F,π) (cf., Diaconis and Freedman 1999). Suppose $Z_n = z$, then $\mathrm{Prob}(Z_{n+1} \in S)$ for some $S \subseteq X$ takes the following value

$$P(z,S) = \sum_{i\in I_N} \pi(i) \mathbf{1}_S\big(f_i(z)\big), \tag{7}$$

where the characteristic function $\mathbf{1}_S$ is defined as

$$\mathbf{1}_S(x) = \begin{cases} 1 & \text{if } x \in S \\ 0 & \text{if } x \notin S \end{cases}$$

When S is a Borel subset of X, there should be an invariant probability measure ρ, such that

$$\rho(S) = \int_X P(z,S)\, d\rho(z) = \sum_{i\in I_N} \pi(i) \rho\big(f_i^{-1}(S)\big). \tag{8}$$

Here, ρ is called a *π-balanced measure* for (X,F,π), as was proposed by Barnsley and Demko (1985).

Let $R(\rho)$ denote the support of ρ, then $R(\rho) = \{x \in X : \rho(x) \neq 0\}$ and hence

$$R(\rho) = \bigcup_{i\in I_N} f_i\big(R(\rho)\big) = H\big(R(\rho)\big). \tag{9}$$

By Theorem 2, it directly appears that $R(\rho) = A$, and therefore the support of a π-balanced measure for (X,F,π) is exactly the unique attractor A of (X,F) for all π. As a result, we can see that the attractor A of (X,F) can also be attained by a random ω-permutation of transition rules in F.

Example 1. Let's assume $X = [0,1]$, $I_2 = \{a,b\}$, $\pi(a) = \pi(b) = 0.5$, and $F = \{f_a, f_b\}$ with

$$f_a : x \mapsto x/3, \quad f_b : x \mapsto x/3 + 2/3.$$

The iterated random function system (X,F,π) is clearly equivalent to the following autoregressive process,

$$Z_{n+1} = Z_n/3 + \varepsilon_{n+1} \quad (n \in \mathbb{Z}^+), \tag{10}$$

where Z_0 is deterministic, and for all $n \in \mathbb{N}$

$$\mathrm{Prob}(\varepsilon_n = 0) = \mathrm{Prob}(\varepsilon_n = 2/3) = 0.5.$$

The iterated function system (X,F) has a unique attractor A as a Cantor ternary set, that is,

$$A = \Big\{ \sum_{n\in\mathbb{N}} x_n/3^n : (x_n,\ n \in \mathbb{N}) \in \{0,2\}^\omega \Big\}. \tag{11}$$

Let

$$B_k = \Big\{ \sum_{n\geq k} x_n/3^n : (x_n,\ n \geq k) \in \{0,2\}^\omega \Big\} \quad (k \in \mathbb{N}),$$

then $B_1 = A$ and $B_2 = A/3$, and thus

$$f_a(A) = B_2, \quad f_b(A) = B_2 + 2/3,$$

which yield $f_a(A) \cup f_b(A) = B_1$. Recall that $H = f_a \cup f_b$ is the Hutchinson operator for (X,F), one can thus write $H(A) = A$.

In practice, the attractor A can be realized by a random motion with any initial state $x \in [0,1]$ in (X,F,π). There are two cases to consider.

If $x \in A$, then $\Gamma^+(x) = A$ almost surely. If $x \notin A$, then there should be a finite sequence $(x_1, x_2, \ldots, x_k) \in \{0,2\}^k$, such that

$$x = \sum_{n=1}^{k} x_n/3^n + r_k(x),$$

where $r_k(x) \leq 1/3^k$. Clearly, $r_k(x)$ will tend to be 0 when k goes to infinity. Now we have $Z_1 = x/3 + \varepsilon_1 = Y_1 + r_k(x)/3$, where $Y_1 = \sum_{n=1}^k x_n/3^{n+1} + \varepsilon_1 \in A$, and in general, $Z_m = Y_m + r_k(x)/3^m$, where $Y_m \in A$ and $r_k(x)/3^m \leq 1/3^{m+k}$. Evidently, there should be an ℓ such that $Z_\ell \in A$, which suggests $\Gamma^+(Z_\ell) = A$ almost surely.

4. Random utility

Consider a generic agent w in a large group W, and suppose w has a preference relation by nature. Let X denote a decision state space for the group W, and let $\gtrsim$ be a weak order on X such that

(i) either $x \gtrsim y$ or $y \gtrsim x$ for all $x, y \in X$,

(ii) $x \gtrsim y$ and $y \gtrsim z$ implies $x \gtrsim z$ for all $x, y, z \in X$.

So $\gtrsim$ can serve as a *rational* preference relation for w. In particular, we shall assume that there is a utility function $u : X \to \mathbb{R}$ such that for all $x, y \in X$,

$$x \gtrsim y \Longleftrightarrow u(x) \geq u(y).$$

Let $\mathcal{P}(X)$ be the power set of X. A mapping $C : \mathcal{P}(X) \to \mathcal{P}(X)$ is called a *choice function* if $\emptyset \neq C(Y) \subseteq Y$ for all nonempty $Y \in \mathcal{P}(X)$.

If $y \in C(Y)$ for some $y \in Y$, and meanwhile, $u(y) \geq u(x)$ for all $x \in Y$, we say the choice made by w matches to her preference relation. It should be noted that there are two implicit assumptions under this statement, i.e., the choices made by w can be perfectly observed, and w can perfectly identify and also intentionally apply her preference relation. However, it seems that empirical verifications would be unable to simultaneously support these two intertwined assumptions. The reason is that observations about the choices made by w are reasonable only if w does make her choices complying with her preference relation, and on the other hand, the true preference relation of w can be thought of as identifiable only if observations about her choices are perfect.

To overcome such difficulties in empirical verifications of consistency of choice and preference, we have to set one assumption ad hoc true, so that we could verify the other one. To begin with, if the preference relation of w is supposed to be perfectly identifiable and intentionally applied by w herself, then it will become possible to infer it from observations about her choices with some confidence level. This approach appeared in a study on stochastic utility model estimation by Manski (1975).

Let $v : X \to \mathbb{R}$ denote a utility function consistent with observations about the choices made by w. And we shall say $v(x)$ is the observed utility if a choice $x \in X$ has been observed. It thus appears to us that

$$u(x) = v(x) + \varepsilon(x), \tag{12}$$

where $\varepsilon(x)$ denotes a "noise" function that might be independently distributed for all $x \in X$. In particular, the choice x can be assumed to be characterized by n independently observed attributes, $J(x) \in \mathbb{R}^n$, thus $v(x)$ admits a linearly parametric model $v(x) = \beta' J(x)$ for $\beta \in \mathbb{R}^n$. In consequence, we have

$$u(x) = \beta' J(x) + \varepsilon(x), \tag{12$'$}$$

in which the estimation $\hat{\beta}$ is determined by the observed data $J(x)$ for $x \in S$, where $S \subseteq X$ is a certain sample.

On the other hand, if the choices made by w are supposed to be perfectly observed, then we could discuss her identification of the true preference relation. In practice, the true preference relation might be only partially identified by w, but it should not be totally vague to her, even if she had an extremely limited cognitive ability. Suppose w has a collection of observable utility functions, which can represent her identified preference relations in different situations, and all these utility functions have an identical kernel as her invariant knowledge of her true preference relation.

Let I_N be a finite index set with $|I_N| = N \geq 2$, and let $v_i : X \to \mathbb{R}$ be a utility function of w for all $i \in I_N$. Suppose $u : X \to \mathbb{R}$ is the kernel utility function of all v_i for $i \in I_N$. Let $X_u = u(X)$, then $X_u \subseteq \mathbb{R}$. And for all $i \in I_N$, there is a contractive function $f_i : X_u \to X_u$ such that $v_i = fi \circ u$, or $v_i(x) = f_i\big(u(x)\big)$ for all $x \in X$. Clearly, $\{v_i : i \in I_N\}$ on the domain X is now equivalent to $F = \{f_i : i \in I_N\}$ on the domain X_u.

Suppose w makes her choices along the time domain $\mathbb{Z}^+$ in such a way that at each time $t \in \mathbb{N}$, she picks a function $f_i \in F$ to form her utility function

$$u_t = f_i \circ u_{t-1}, \tag{13}$$

where u_{t-1} is her utility function at the time $t-1$. More concretely, at the initial time 0, the utility function of w is set as her kernel utility function, i.e., $u_0(x) = u(x)$, and at the time 1, her utility function is $u_1(x) = f_i\big(u(x)\big)$ for some $i \in I_N$. In general, at any time $t \in \mathbb{N}$, her utility function is $u_t(x) = f_i\big(u_{t-1}(x)\big)$ for some $i \in I_N$. Here, the sequence of utility functions $\big(u_t(x),\, t \in \mathbb{Z}^+\big)$ can be considered as a general extension of a normal discounted utility function series, and in terms of time preference, we actually generalize $(\gtrsim,\, t \in \mathbb{Z}^+)$ to $(\gtrsim_t,\, t \in \mathbb{Z}^+)$, where $\gtrsim_t$ varies across time.

For the moment, we should notice that $(u_t(x),\, t \in \mathbb{Z}^+)$ is completely determined by the iterated function system (X_u, F). By Theorem 2, one can directly see that (X_u, F) has a unique attractor, say $A \subseteq X_u$, such that $A = \bigcup_{i \in I_N} f_i(A)$. It thus suggests that some kernel utilities in A could be reached by w in the long term.

Let π denote a probability measure on I_N, then an iterated random function system (X_u, F, π) will emerge. By (6), we obtain

$$U_{t+1}(x) = f_{\sigma_{t+1}}\big(U_t(x)\big) \quad (t \in \mathbb{Z}^+), \tag{14}$$

where $U_0(x) = u(x)$, and $U_t(x)$ denotes the random utility function at the time t.

If $f_i(x) = \rho_i x$ for all $i \in I_N$, where $\rho_i \in (0,1)$ and $\rho_i \neq \rho_j$ for all distinct $i, j \in I_N$, then (14) will be

$$U_{t+1}(x) = \xi_{t+1} U_t(x) \quad (t \in \mathbb{Z}^+), \tag{14$'$}$$

where $\mathrm{Prob}(\xi_t = \rho_i) = \pi(i)$ for all $i \in I_n$ and all $t \in \mathbb{N}$. Thus at any time $t \in \mathbb{N}$, the random utility function of w is

$$U_t(x) = \Big(\prod_{n=1}^{t} \xi_n\Big) u(x) = \exp\Big(\sum_{n=1}^{t} \log \xi_n\Big) u(x) = \exp\Big(-\sum_{n=1}^{t} \log(1/\xi_n)\Big) u(x).$$

Let $\delta_t = \sum_{n=1}^{t} \log(1/\xi_n)$, then the random utility function of w at $t \in \mathbb{N}$ can be written as

$$U_t(x) = \mathrm{e}^{-\delta_t} u(x). \tag{15}$$

When t goes to infinity, δ_t will approach infinity, and thus $U_t(x)$ will approach zero almost surely for all choice x.

If $f_i(x) = \rho x + r_i$ for all $i \in I_N$, where $\rho \in (0,1)$, $r_i > 0$, and $r_i \neq r_j$ for all distinct $i, j \in I_N$, then (14) will be

$$U_{t+1}(x) = \rho U_t(x) + \theta_{t+1} \quad (t \in \mathbb{Z}^+), \tag{14''}$$

in which once again $\mathrm{Prob}(\theta_t = r_i) = \pi(i)$. At any time $t \in \mathbb{N}$, the random utility function of w then becomes

$$U_t(x) = \rho^t u(x) + \sum_{n=1}^{t} \rho^{t-n} \theta_n. \tag{16}$$

Note that $\rho^t u(x)$ will vanish when t goes to infinity, but the remaining part will not converge almost surely, as a new piece of randomness θ_t will emerge at each time t.

5. Stochastic growth

Consider an economy with a production function $Y = F(K,L)$, where Y, K, L denote the total production, the capital input, and the labor supply in the economy, respectively. Let $y = Y/L$ and $k = K/L$, and suppose $F(K,L)$ is a homogeneous function of degree 1, then $Y/L = F(K/L, 1)$. Define $f(k) = F(K/L, 1)$, thus the production technology of a generic agent w in that economy can be represented by

$$y = f(k) \quad (k \in \mathbb{R}^+). \tag{17}$$

As typically assumed, $f(k)$ should satisfy that for all $k \in \mathbb{R}^+$

$$f'(k) > 0, \quad f''(k) < 0,$$

and the following Inada conditions, which are usually named after K. Inada, but also partly attributed to H. Uzawa (Uzawa 1961),

$$\lim_{k \downarrow 0} f'(k) = +\infty, \quad \lim_{k \uparrow +\infty} f'(k) = 0.$$

Let's now introduce a stochastic factor ξ into the economy, so that the production technology of w can be expressed as

$$y = f(k, \xi) \quad (k \in \mathbb{R}^+). \tag{18}$$

In case k and ξ are separable, we could consider two fundamental cases, i.e., ξ is an additive shock to $f(k)$, or ξ is a multiplicative shock to $f(k)$. Similar to the studies by Mitra et al. (2004), and Mitra and Privileggi (2009), we shall also focus on the latter case, and rewrite the technology (18) as

$$y = \xi f(k) \quad (k \in \mathbb{R}^+), \tag{18'}$$

where $\xi > 0$ is a random variable. In practice, we can assume that the support of ξ is $\{\lambda_i : i \in I_N\}$, where I_N is a finite index set with $|I_N| = N \geq 2$, and there is a probability measure π on I_N, such that $\mathrm{Prob}(\xi = \lambda_i) = \pi(i)$ for all $i \in I_N$.

In addition, the consumption and investment which are both necessary parts of a sustainable economy, are denoted by C and E, thus we should have $Y = C + E$. Let $c = C/L$ and $e = E/L$, then the income identity for w is $y = c + e$. Suppose the economy functions on the time domain $\mathbb{Z}^+$, so that the economic variables all become discretely time-dependent, that is, $y_t, k_t, c_t, e_t, \xi_t$ for $t \in \mathbb{Z}^+$, then the economy can be represented by the following system:

$$\begin{cases} y_t = \xi_t f(k_t) \\ y_t = c_t + e_t \\ k_{t+1} = e_t \end{cases}$$

in which $k_0 \neq 0$ is the initial capital input, and $\xi_t, \xi_{t'}$ are independent for all distinct $t, t' \in \mathbb{Z}^+$.

Suppose w has a stationary utility function in her consumption c which is written as $u(c)$, such that $u'(c) > 0$ and $u''(c) < 0$ for all $c \in \mathbb{R}^+$, and $\lim_{c \downarrow 0} u'(c) = +\infty$, then it clearly appears that $c_t > 0$ at any time $t \in \mathbb{Z}^+$. Assume the time preference of w can be characterized by a regular discounting $\rho \in (0,1)$, then her additive utilities from a deterministic consumption flow $(c_0, c_1, \ldots, c_t)$ for $t \in \mathbb{Z}^+$, can be expressed as $\sum_{n=0}^{t} \rho^n u(c_n)$.

The steady growth path of the economy is thus determined by the equilibrium of the decision-making process for w. In other words, w maximizes

$$\mathrm{E}_0 \sum_{t \in \mathbb{Z}^+} \rho^t u(c_t),$$

subject to $c_t = \xi_t f(k_t) - k_{t+1}$ for all $t \in \mathbb{Z}^+$ with $k_0 > 0$ initially given. Here, we apply E_t to denote the expectation operator at a time $t \in \mathbb{Z}^+$.

Recall that an optimal consumption flow $(c_t,\ t \in \mathbb{Z}^+)$ should satisfy the following Euler equation,

$$u'(c_t) = \rho\, \mathrm{E}_t \left(\xi_{t+1} f'(k_{t+1}) u'(c_{t+1}) \right). \tag{19}$$

Since $k_{t+1} = y_t - c_t$, (19) is equivalent to

$$u'(c_t) = \rho f'(y_t - c_t)\, \mathrm{E}_t \left(\xi_{t+1} u'(c_{t+1}) \right). \tag{19$'$}$$

There should be a real function φ such that $c_t = \varphi(y_t)$ for all c_t in the optimal consumption flow, which yields $k_{t+1} = y_t - \varphi(y_t)$, and thus

$$y_{t+1} = \xi_{t+1} f(k_{t+1}) = \xi_{t+1} f\big(y_t - \varphi(y_t)\big).$$

Let $\psi(y) = f\big(y - \varphi(y)\big)$, then we have the following stochastic growth process:

$$y_{t+1} = \xi_{t+1} \psi(y_t) \quad (t \in \mathbb{Z}^+). \tag{20}$$

Let $X_Y \subseteq \mathbb{R}^+$ be an invariant support set for y_t driven by the above process (20), so that $y_t \in X_Y$ at any $t \in \mathbb{Z}^+$. Define $g_i(y) = \lambda_i \psi(y)$ for all $y \in X_Y$. Let $G = \{g_i : i \in I_N\}$, then the stochastic growth process $(y_t,\ t \in \mathbb{Z}^+)$ as is determined by (20) should be equivalent to the iterated random function system (X_Y, G, π).

Corresponding to the optimal consumption flow, the following optimal capital flow would directly come out,

$$k_{t+1} = y_t - \varphi(y_t) = \xi_t f(k_t) - \varphi\big(\xi_t f(k_t)\big), \tag{21}$$

which can also be supposed to admit an invariant support set $X_K \subseteq \mathbb{R}^+$. Define

$$m_i(k) = \lambda_i f(k) - \varphi\big(\lambda_i f(k)\big),$$

and let $M = \{m_i : i \in I_N\}$, then we have another iterated random function system (X_K, M, π), which in a sense is conjugate to the former (X_Y, G, π).

Example 2. Take $I_N = \{a, b\}$, $f(k) = \sqrt[3]{k}$, and $u(c) = \log c$. Let's suppose $(\xi_t,\ t \in \mathbb{Z}^+)$ is a Bernoulli process with

$$\mathrm{Prob}(\xi_t = \lambda_a) = q, \quad \mathrm{Prob}(\xi_t = \lambda_b) = 1 - q,$$

where $q \in (0, 1)$, and

$$1/\lambda_a^2 < \lambda_b < 1 < \lambda_a < 1/\lambda_b.$$

It thus suggests that the shock is either positive or negative, while the negative shock would not make the economy vanish as $\lambda_b \lambda_a^2 > 1$, and the positive shock would not make it too expansive as $\lambda_a \lambda_b < 1$.

In the optimal consumption flow $(c_t,\ t \in \mathbb{Z}^+)$, we might see that $c_t = (1 - \rho/3) y_t$, which yields $\varphi(y_t) = (1 - \rho/3) y_t$, and thus the optimal capital flow is determined by the formula

$$k_{t+1} = \rho y_t / 3 = \rho \xi_t \sqrt[3]{k_t} / 3.$$

Let $\kappa_t = \log k_t$, then we have

$$\kappa_{t+1} = \kappa_t / 3 + \log \xi_t + \log(\rho/3),$$

which should have an invariant support interval $[\alpha, \beta] \subset \mathbb{R}$.

We now have the following two affine functions:

$$\ell_a(\kappa) = \kappa/3 + \big(\log \lambda_a + \log(\rho/3)\big), \quad \ell_b(\kappa) = \kappa/3 + \big(\log \lambda_b + \log(\rho/3)\big).$$

Let $\Lambda = \{\ell_a, \ell_b\}$, then $([\alpha, \beta], \Lambda)$ is an iterated function system. Notice that

$$\beta/3 + \big(\log \lambda_a + \log(\rho/3)\big) = \beta, \quad \alpha/3 + \big(\log \lambda_b + \log(\rho/3)\big) = \alpha,$$

so $\log \lambda_a + \log(\rho/3) = 2\beta/3$ and $\log \lambda_b + \log(\rho/3) = 2\alpha/3$, and thus $\ell_a(\kappa)$ and $\ell_b(\kappa)$ can be also written as

$$\ell_a(\kappa) = \kappa/3 + 2\beta/3, \quad \ell_b(\kappa) = \kappa/3 + 2\alpha/3,$$

where $\beta > \alpha$ because $\lambda_a > \lambda_b$. Let $z = (\kappa - \alpha)/(\beta - \alpha)$, then Λ on $[\alpha, \beta]$ can be transformed into a pair of functions defined on $[0,1]$, i.e.,

$$Z = \{z/3, z/3 + 2/3\}.$$

It therefore appears that $([\alpha,\beta],\Lambda)$ is equivalent to the iterated function system $([0,1],Z)$. By Example 1, we know that the unique attractor of $([0,1],Z)$ is the Cantor ternary set, and thus the attractor of $([\alpha,\beta],\Lambda)$ should be also a Cantor set, which then conveys that the dynamics of the optimal stochastic growth in the economy should be essentially chaotic.

6. Concluding remarks

In this article, we have demonstrated by example how a plain mechanism could generate a complex economic system with increasing disorder through the iteration process. In reality, economic systems usually themselves show very complicated dynamics which could be observed and recorded. For example, the quote dynamics in a security market are erratic and occasionally trapped in catastrophes. People are inclined to understand such "irregular" phenomena from the statistical viewpoint, that is, the dynamics should be replicated by a certain skeleton with some additional randomness or perturbation. With a flavor of this work, one might perceive a quite different approach to interpret the irregularity, that is, the dynamics might be driven by some deterministic rules, and the irregularity emerges as a form of chaos.

Acknowledgment The author is grateful to two reviewers, whose comments and suggestions should have made the present article more clear and rigorous.

References

Barnsley, M. F. and Demko, S. (1985). Iterated function systems and the global construction of fractals. *Proceedings of the Royal Society of London A*, 399(1817), 243–275.

Benhabib, J. and Nishimura, K. (1985). Competitive equilibrium cycles. *Journal of Economic Theory*, 35(2), 284–306.

Boldrin, M. and Montrucchio, L. (1986). On the indeterminacy of capital accumulation paths. *Journal of Economic Theory*, 40(1), 26–39.

Day, R. H. (1983). The emergence of chaos from classical economic growth. *Quarterly Journal of Economics*, 98(2), 201–213.

Day, R. H. (1994). *Complex Economic Dynamics: Vol. I, An Introduction to Dynamical Systems and Market Mechanisms*. Cambridge, MIT Press.

Day, R. H. (1999). *Complex Economic Dynamics: Vol. II, An Introduction to Macroeconomic Dynamics*. Cambridge, MIT Press.

Diaconis, P. and Freedman, D. (1999). Iterated random functions. *SIAM Review*, 41(1), 45–76.

Ezekiel, M. (1938). The cobweb theorem. *Quarterly Journal of Economics*, 52(2), 255–280.

Goodwin, R. M. (1990). *Chaotic Economic Dynamics*. New York, Oxford University Press.

Hutchinson, J. E. (1981). Fractals and self similarity. *Indiana University Mathematics Journal*, 30(5), 713–747.

Kaldor, N. (1934). A classificatory note on the determinateness of equilibrium. *Review of Economic Studies*, 1(2), 122–136.

Leontief, W. (1934). Delayed adjustment of supply and partial equilibrium. *Journal of Economics*, 5(5), 670–676. (in German)

Li, T.-Y. and Yorke, J. A. (1975). Period three implies chaos. *American Mathematical Monthly*, 82(10), 985–992.

Lorenz, H.-W. (1993). *Nonlinear Dynamical Economics and Chaotic Motion*, 2nd edition. Berlin, Springer-Verlag.

Manski, C. F. (1975). Maximum score estimation of the stochastic utility model of choice. *Journal of Econometrics*, 3(3), 205–228.

Milnor, J. (1985). On the concept of attractor. *Communications in Mathematical Physics*, 99(2), 177–195.

Mitra, T., Montrucchio, L. and Privileggi, F. (2004). The nature of the steady state in models of optimal growth under uncertainty. *Economic Theory*, 23(1), 39–71.

Mitra, T. and Privileggi, F. (2009). On Lipschitz continuity of the iterated function system in a stochastic optimal growth model. *Journal of Mathematical Economics*, 45(1–2), 185–198.

Šarkovskii, A. N. (1964). Coexistence of cycles of a continuous map of the line into itself. *Ukrainian Mathematical Journal*, 16, 61–71. (in Russian)

Tu, P. N. V. (1994). *Dynamical Systems: An Introduction with Applications in Economics and Biology*, 2nd edition. Berlin, Springer-Verlag.

Uzawa, H. (1961). On a two-sector model of economic growth. *Review of Economic Studies*, 29(1), 40–47.

Czech Economic Review 9 (2015) 169–183
Acta Universitatis Carolinae Oeconomica

Reciprocal Equilibria in Link Formation Games

Hannu Salonen*

Received 4 August 2015; Accepted 20 May 2016

Abstract We study non-cooperative link formation games in which players have to decide how much to invest in relationships with other players. A link between two players is formed, if and only if both make a positive investment. The cost of forming a link can be interpreted as the value of privacy. We analyze the existence of pure strategy equilibria and the resulting network structures with tractable specifications of utility functions. Sufficient conditions for the existence of reciprocal equilibria are given and the corresponding network structure is analyzed. Pareto optimal and strongly stable network structures are studied. It turns out that such networks are often complete.

Keywords Link formation games, reciprocal equilibrium, complete network
JEL classification C72, D43

1. Introduction

We study non-cooperative link formation games in which players have to decide how much to invest in relationships with other players. A link between two players is formed, if and only if both make a positive investment. The value of a link depends on the size of investments, and this value can be different for different players. The cost of forming a link can be interpreted as the value of privacy, or the opportunity cost of lost privacy.

Friendships, partnerships, and researchers' collaboration networks are prime examples of situations that could be modeled this way. Friendships could be strong or weak and two people in a relationship could value it differently. Researchers may spend different levels of effort in their joint projects, and they could value their cooperation differently. It is therefore important to understand what kind of factors affect agents' choices in such situations, and how the equilibrium network looks like.

We analyze the existence of pure strategy equilibria and the resulting network structures with tractable specifications of utility functions. Sufficient conditions for the existence of reciprocal equilibria are given and the corresponding network structure is analyzed. Pareto optimal and strongly stable network structures are studied. It turns out that such networks are often complete.

Each player has a fixed amount of a single resource like time or effort that he can invest in relationships with other players and/or use for his own private benefit. The more two players invest in their mutual relationship, the higher is the utility to both

* University of Turku, Department of Economics, 20014 Turku, Finland. Phone: +35823335403, E-mail: hansal@utu.fi.

players from this relationship. Since resources are limited, utility from privacy or from other relationships decreases, and there is a tradeoff between relationships. Decisions are made simultaneously and pure strategy Nash equilibria are searched for.

We show that a reciprocal equilibrium with a complete network (or a network with complete components) exists in many symmetric or anonymous link formation games (Theorems 1 and 2). In such an equilibrium players i and j invest equal amounts in their mutual relationship. Bauman (2015) studies reciprocity of equilibria in symmetric games with strictly concave valuations of privacy and constant returns to scale Cobb-Douglas utilities from relationships.

Network structure in a reciprocal equilibrium depends on players' valuations of privacy. If these valuations are linear functions, then reciprocal equilibria often exhibit *homophily* (Theorem 3): links are more likely to be formed between similar players (Currarini et al. 2009).

Equilibria with a complete network exists under variety of circumstances when reciprocity is not demanded, for example in *semi-symmetric* games with bilateral strategic complements or substitutes (Theorems 4 and 5). In semi-symmetric link formation games players have common preferences over other players as friends. In such cases it is important to understand how the popularity or status of an agent affects his behavior in the network. Salonen (2015) studies the relation between popularity and some well-known network centrality measures in semi-symmetric games.

In the class of models studied in this paper, Pareto optimality of a network structure implies in many cases that network must be complete (Propositon 1). Similarly, strongly stable equilibria (Bloch and Dutta 2009) have often complete networks as well (Proposition 2). It is shown at the end of Section 4.1. that *any* Pareto optimal or strongly pairwise stable equilibrium must have a complete network, when utilities have a strictly concave Cobb-Douglas form.

Completeness of a network sounds rather extreme if the player set is very large. A more moderate interpretation of these results would be that networks consist of completely connected components. Be this as it may, Bloch and Dutta (2009) get results that efficient or strongly stable networks are stars. It is therefore necessary to compare the underlying assumptions of our models.

We assume that players get utility only from private consumption or direct links (relationships) with other players, and that a relationship of two players gives positive utility only if both players have made a positive investment. Bloch and Dutta (2009) assume that players get utility also from indirect connections, i.e. from friends of friends, and that a link between two players is formed even if only one of the players has made a positive investment. In our model two linked players may value the relationship differently, whereas in their model the values are identical.

The model of Bloch and Dutta (2009) may be more natural in situations where links have instrumental value, like communication networks. Since direct links are not absolutely necessary for information transmission, complete networks need not be efficient structures. Our model is perhaps better suited in cases where links have intrinsic value, like friendships. In such cases indirect connections may be very poor substitutes for direct links, and increasing the number of direct links becomes both individually and socially optimal.

There is a large literature of link formation games where the link strength can take only two values: either it is 1 (link is formed) or 0 (link is not formed). Jackson and Wolinsky (1996) is the seminal paper of this strand of literature (see Jackson and Zenou 2015 for a comprehensive review of network games). Cabrales et al. (2011) analyze a linear quadratic game with productive investments and link formation where link strengths can be nonnegative real numbers. Rather than choosing each link intensity separately, a player chooses one real number that describes his socialization effort. Strengths of individual links are then determined jointly, given socialization efforts of all players. The resulting network determines the profitability of productive investments.

In our model players invest in each link separately, and the utility from equal investments in different links may be different. So the links of a player may represent very different relationships with other players, although seemingly a player decides only how to share a homogeneous resource among his friends.

The paper is organized in the following way. The notation is introduced in Section 2. In Section 3 some simple models with Cobb-Douglas functions are analyzed. Main results are stated in Section 4.

2. The Model

A tuple $W = (N, g)$ denotes an unweighted, undirected *network* with a finite node set N and a link set g. The link set g specifies which nodes $i, j \in N$ are directly connected. Such a link may be denoted by $ij \in g$ with the understanding that $ji = ij$. In this paper loops are ignored so $i \neq j$ if i and j are linked. If it is clear what the node set is we may denote a network simply by g.

Given a network $W = (N, g)$ and $i, j \in N$, there exists a *path* between i and j, if there exists nodes $i_0, \ldots, i_K$ such that (i) $i_0 = i, i_K = j$; (ii) $i_k i_{k+1} \in g$ for all $k = 0, \ldots, K-1$; (iii) all nodes are distinct except possibly i_0 and i_K. A network $W = (N, g)$ is *connected* if there exists a path between any two nodes $i, j \in N$.

A subset $A \subset N$ is a *component* of a network $W = (N, g)$, if (i) there exists a path between any two nodes $i, j \in A$; (ii) there are no links between A and $A^c \equiv N \setminus A$. A node set N can always be partitioned into connected components. A network $W = (N, g)$ is connected, if N is a component. A component A is complete, if for every $i, j \in A$ there is a link $ij \in g$. A network $W = (N, g)$ is complete, if N is a complete component.

A *normal form game* $G = \big(N, (S_i)_{i \in N}, (u_i)_{i \in N}\big)$ specifies a player set N, a set of pure strategies S_i and a utility function $u_i : S \longrightarrow \mathbb{R}$ for each player $i \in N$, where $S = \Pi_i S_i$, the product of strategy sets, is the set of *strategy profiles*.

A game G is *symmetric*, if $S_i = S_j$ for all $i, j \in N$, and $u_i(s) = u_j(s')$ for all $i, j \in N$, for all $s, s' \in S$ such that $s_i = s'_j, s_j = s'_i$ and $s_k = s'_k$ for all $k \neq i, j$.

A game G is *anonymous*, $S_i = S_j$ for all $i, j \in N$, and $u_i(s) = u_i(s')$ if the only difference between s and s' is that $s_j = s'_k$ and $s_k = s'_j$ for some $j, k \neq i$.

Given $s \in S$, we may denote $s = (s_i, s_{-i})$ when we want to emphasize that i chooses

s_i. A pure strategy Nash equilibrium is a strategy profile $s \in S$ such that

$$u_i(s_i, s_{-i}) \geq u_i(s'_i, s_{-i}), \forall i \in N, \forall s'_i \in S_i. \tag{1}$$

Given a symmetric game G, a strategy profile s is a *symmetric equilibrium*, if $s_i = s_j$ for all $i, j \in N$.

We study *link formation games* of the following type. The set of pure strategies of player $i \in N$ is

$$S_i = \left\{ s_i \in R^N_+ \mid \sum_j s_{ij} = 1 \right\}.$$

An interpretation is that each player i has one unit of time or effort to be shared with other player j including i himself. The utility function of player i is

$$u_i(s) = \sum_{j \neq i} U_{ij}(s_{ij}, s_{ji}) + V_i(s_{ii}), \tag{2}$$

where $U_{ij} : [0,1]^2 \longrightarrow \mathbb{R}_+$ is a function giving the utility for player i from investments s_{ij}, s_{ji}. The function U_{ij} has the following properties: (i) $U_{ij}(0, s_{ji}) = 0 = U_i(s_{ij}, 0)$ for all s_{ij}, s_{ji}; (ii) U_{ij} is strictly concave and differentiable in s_{ij} for any given $s_{ji} > 0$; (iii) U_{ij} is strictly increasing and continuous on $(0,1] \times (0,1]$.

The function $V_i : [0,1] \longrightarrow \mathbb{R}_+$ tells how much player i values privacy. The function V_i is concave, strictly increasing, differentiable on $(0,1)$, and $V_i(0) = 0$.

In anonymous link formation games $U_{ij} = U_i$ for all $i, j \in N$, and hence $u_i = U_i + V_i$. In symmetric link formation games $u_i = U + V$ for all $i \in N$. [In a link formation game the identity of strategies $s_i = s_j$ is understood so that $s_{ii} = s_{jj}$, $s_{ij} = s_{ji}$, and $s_{ik} = s_{jk}$ for all $k \neq i, j$.]

We say that a link formation game $G = \big(N, (S_i)_{i \in N}, (u_i)_{i \in N}\big)$ is *semi-symmetric*, if there are functions U and V such that

$$u_i(s) = \sum_{j \neq i} p_j U(s_{ij}, s_{ji}) + c_i V(s_{ii}), \forall s \in S, \tag{3}$$

for some parameters $p_j > 0, c_i > 0$, for all $i, j \in N$. So there is a common ordering of players such that player j is considered as a more valuable friend than i, if $p_j > p_i$. The cost parameters c_i reflecting the value of privacy could be player specific.

There are two different interpretations of the network model associated with the game. We may think that the network is undirected and unweighted, and the investments describe the intensity of the relationship, or how much the agents utilize a given link and how much they benefit from it. Alternatively, investment s_{ij} gives the strength of the link from i to j, and players get nonzero utility from a relationship only if both make a positive investment. In this case the network is directed and weighted. To keep notation simple, we gave formal definitions for the undirected unweighted network only.

Next we give some definitions that are needed in the main theorems.

Definition 1 (Bilateral strategic complements). U_{ij} is twice continuously differentiable on $(0,1) \times (0,1)$ with $\partial^2 U_{ij} / \partial s_{ji} \partial s_{ij} > 0$, $i \neq j$.

Bilateral strategic complements imply $\partial^2 u_i/\partial s_{ji}\partial s_{ij} = \partial^2 U_{ij}/\partial s_{ji}\partial s_{ij}$ by (2). Since $s_{ii} = 1 - \sum_{j\neq i} s_{ij}$ the usual strategic complements condition is not satisfied: if s_{ji} increases, then s_{ij} increases but s_{ik} decreases for some $k \neq j$ when i uses a best reply.

Analogously, bilateral strategic substitutes mean $\partial^2 U_{ij}/\partial s_{ji}\partial s_{ij} < 0$ holds on $(0,1)\times(0,1)$, for all players i.

Definition 2 (**Increasing derivative on the diagonal**). A function $U_{ij} : [0,1]^2 \longrightarrow \mathbb{R}_+$ has (strictly) increasing derivative on the diagonal, if

$$\frac{\partial U_{ij}(y,y)}{\partial x_1} (<) \leq \frac{\partial U_{ij}(z,z)}{\partial x_1}, \text{ for all } y < z.$$

If both i and j invest y in their relationship, the marginal utility for i increases in y. If the inequality in Definition 2 is reversed, we say that U_{ij} has (strictly) decreasing derivative on the diagonal. If equality holds for all $y < z$ we say that U_{ij} has constant derivative on the diagonal.

Note that if U_{ij} is (jointly) concave, then it has a decreasing derivative on the diagonal. On the other hand the Cobb-Douglas function $f(x,y) = x^a y^b$ is concave in both arguments separately and has increasing derivative on the diagonal, if $0 < a,b < 1$, and $a+b \geq 1$. If U_{ij} is homogeneous of degree $\alpha \geq 1$ ($0 < \alpha \leq 1$), then U_{ij} has increasing (decreasing) derivative on the diagonal. Homogeneity is clearly a much stronger assumption than increasing and decreasing derivative conditions.

If a game is not symmetric, a symmetric equilibrium need not exist. However, behavior may be nearly symmetric also in non-symmetric games. The following is a pairwise or bilateral symmetry condition that seems natural in the context of friendship networks.

Definition 3 (**Reciprocal equilibrium**). An equilibrium s of a link formation game $G = \big(N, (S_i)_{i\in N}, (u_i)_{i\in N}\big)$ is *reciprocal*, if $s_{ij} = s_{ji}$ for any pair $i,j \in N, i \neq j$ of players.

An *interior equilibrium* s of a link formation game is such that $s_{ii}, s_{ij} > 0$ for all players i,j. The network corresponding to an interior equilibrium is complete. Note that if $s_{ii} = 0$ for some player i, then s cannot be an interior equilibrium. If $s_{ij} > 0$ for all i and for all $j \neq i$, then the network is complete.

3. Examples

Let us first analyze some simple examples based on Cobb-Douglas functions U_{ij}.

Example 1. Let $G = \big(N, (S_i)_{i\in N}, (u_i)_{i\in N}\big)$ be a semi-symmetric game with bilateral strategic complements such that

$$u_i(s) = \sum_{j\neq i} p_j s_{ij}^{\alpha} s_{ji}^{1-\alpha} + c_i\big(1 - \sum_{j\neq i} s_{ij}\big),$$

where $0 < \alpha < 1$, and $c_i, p_j > 0$. For generic values of parameters α, c_i, p_j all equilibria s satisfying $s_{ii} > 0$ for all i are autarkic. That is, $s_{ii} = 1, \forall i$. To see this, suppose that in equilibrium all the values s_{ij}, s_{ji}, s_{ii} and s_{jj} are strictly positive for some players i, j. Then the corresponding first order conditions for players i and j satisfy:

$$\begin{aligned} \alpha p_j s_{ij}^{\alpha-1} s_{ji}^{1-\alpha} &= c_i \\ \alpha p_i s_{ji}^{\alpha-1} s_{ij}^{1-\alpha} &= c_j \end{aligned} \tag{4}$$

These equations imply

$$\alpha^2 p_i p_j = c_i c_j, \tag{5}$$

which does not hold for generic values of α, c_i, p_j. Namely, let $\mathbb{R}^{2n+1}_{++}$ be the set of all strictly positive vectors $x = (\alpha, p_1, c_1, \ldots, p_n, c_n)$. Take any two players i, j. The subset of vectors $x \in \mathbb{R}^{2n+1}_{++}$ such that $\alpha p_i p_j = c_i c_j$ is closed and has an empty interior in $\mathbb{R}^{2n+1}_{++}$. Since there are only finitely many players, the subset B such that equation (5) does *not* hold for any two players i, j is such that the closure of B contains $\mathbb{R}^{2n+1}_{++}$. Hence, generically $\alpha p_i p_j = c_i c_j$ does not hold.

Suppose that for each pair p_t, c_t there is a group N_t of players with these parameters in their utility functions. Then the genericity result above implies that typically links are formed only within each group N_t, if $s_{ii} > 0$ for all players. In this case equilibria exhibit *homophily*: links are formed only between similar players (Currarini et al. 2009). Note however that other kinds of equilibria may exist if $s_{ii} = 0$ for some players.

The game G of Example 1 has constant derivative on the diagonal and a linear V_i function. Theorem 2 below shows that if G is an anonymous game and V_i functions are strictly concave, then an interior *reciprocal* equilibrium often exists.

Let us modify the game G of Example 1 slightly so that interior equilibria exist.

Example 2. Let $G = \big(N, (S_i)_{i \in N}, (u_i)_{i \in N}\big)$ be a game such that

$$u_i(s) = \sum_{j \neq i} p_j s_{ij}^{\alpha} s_{ji}^{\beta} + c_i\big(1 - \sum_{j \neq i} s_{ij}\big),$$

where $0 < \alpha, \beta$, $\alpha + \beta < 1$, and $c_i, p_j > 0$, for all $i, j \in N$. Let $p_{ij} = p_j / c_i$, and the first order conditions for an interior equilibrium for players i, j are:

$$\alpha p_{ij} s_{ij}^{\alpha-1} s_{ji}^{\beta} = 1 \tag{6}$$

$$\alpha p_{ji} s_{ji}^{\alpha-1} s_{ij}^{\beta} = 1 \tag{7}$$

Solving for s_{ij} gives us

$$s_{ij} = \alpha^{1/[1-\alpha-\beta]} \left[p_{ij}^{1-\alpha} p_{ji}^{\beta} \right]^{1/[(1-\alpha)^2 - \beta^2]}, \forall i, j \in N. \tag{8}$$

Note that s_{ij} is an increasing function of both p_{ij} and p_{ji}. If c_j increases, the value of privacy for j increases, and p_{ji} decreases. Then j invests less in his relations with other agents. Consequently, also s_{ij} decreases by bilateral complementarity.

If all players are identical, then $p_{ij} = p/c$ for all i, j, for some p, c. A symmetric interior equilibrium exists if

$$\alpha p < c\left[\frac{1}{n-1}\right]^{1-\alpha-\beta}.$$

As n increases, this inequality holds if p decreases or c increases sufficiently. This holds since in symmetric equilibrium marginal utility from links increases as n increases because $\alpha+\beta<1$. At an interior equilibrium $s_{ii}>0$, and therefore the value of privacy c must increase relative to p.

For a nonsymmetric example, let $n=11, \alpha=1/4, \beta=1/2$, and $p_1=p, p_2=p^2, \ldots, p_n=p^n$ for some $p\in(0,1)$. If $c_i=1$ for all i, then an equilibrium with a complete network is given by

$$s_{ij}=4^{-4}\left[p^{2i+3j}\right]^{4/5}, \tag{9}$$

from which we can compute that

$$s_{ji}=\left[p^{i-j}\right]^{4/5}s_{ij}, \text{ for } j<i.$$

The players who are highly ranked by the society (low i and high p_i) invest less in relationships than lower ranked players. Take $i=4$ and $j=3$. Then $s_{34}=p^{4/5}s_{43}$, and therefore $s_{34}<s_{43}$.

For another numerical example, assume $p_j=1$ for all players j, and $c_1=c, c_2=2c, \ldots, c_n=nc$, for some $c>1/2$, and let the other parameters have the same values as above. Then the following values characterize an equilibrium with a complete network:

$$s_{ij}=4^{-4}\left[i^{-3}j^{-2}c^{-5}\right]^{4/5}, \tag{10}$$

from which we can compute that

$$s_{ji}=\left[\frac{i}{j}\right]^{4/5}s_{ij}.$$

The players with high value of privacy (high c and i) invest less in relationships than players with a low value of privacy. Take $i=4$ and $j=3$. Then $s_{34}=(4/3)^{4/5}s_{43}$, and therefore $s_{34}>s_{43}$.

4. Results

The existence of equilibria is not a problem in our model, since a strategy profile s such that $s_{ii}=1$ and $s_{ik}=0, k\neq i$, for all $i\in N$ is trivially an (autarkic) equilibrium, and also a reciprocal equilibrium. Here is a more interesting existence result for symmetric games. All long proofs are relegated in the Appendix.

Theorem 1. *A symmetric link formation game* $G=\big(N,(S_i)_{i\in N},(u_i)_{i\in N}\big)$ *has a nontrivial reciprocal equilibrium with complete components, if and only if there exists*

$x \in (0,1)$ *such that*

$$\frac{\partial U(x,x)}{\partial x_1} - V'(1-x) \geq 0. \tag{11}$$

Proof. See Appendix.

The condition (11) says it is better to form one reciprocal link than not to form any links with other players.

Reciprocal equilibria may exist also in nonsymmetric games.

Theorem 2. *Let $G = \big(N, (S_i)_{i\in N}, (u_i)_{i\in N}\big)$ be an anonymous link formation game with the following properties: (i) constant derivative on the diagonal, (ii) V_i is strictly concave. Assume also that if all players $i \in N$ have the same utility function u_i, then the corresponding symmetric game would have a symmetric equilibrium s^i such that the resulting network is complete, for any $i \in N$. Then there exists a reciprocal equilibrium such that the resulting network is complete.*

Sketch of a proof. Theorem 1 gives necessary and sufficient conditions for the existence of a reciprocal equilibrium in symmetric games. The idea of the proof is the following. Formulate a symmetric game corresponding to each of the utility functions u_i that players in a (nonsymmetric) game G have. By assumption, each of these games has a symmetric equilibrium with a complete network. A reciprocal equilibrium for G can be recursively constructed from these symmetric equilibria. For details see Appendix. □

Note that if V_i is linear, then Theorem 2 may not hold by Example 1. Theorem fails if V_i is linear even if U_i is assumed to be strictly concave as the following result demonstrates. We say that a node subset C is a *clique*, if there is a link between every two nodes $i, j \in C$.

Theorem 3. *Suppose $G = \big(N, (S_i)_{i\in N}, (u_i)_{i\in N}\big)$ is an anonymous link formation game such that (i) derivative is strictly decreasing on the diagonal, and (ii) V_i is linear and $U_i = U$. If there is a reciprocal interior equilibrium s such that the equilibrium network has a clique C, then players $i \in C$ have the same utility functions $u_i = U + V_i$.*

Proof. By condition (ii), $V_i(s_{ii}) = c_i s_{ii}$ for some constant $c_i > 0$. For each $i \in C$, there is at most one x^i such that $\partial U_i(x^i, x^i)/\partial x_1 = c_i$ by condition (i). For $i \in C$ this equality must hold in the reciprocal equilibrium s since $s_{ii} > 0$. If $c_i \neq c_j$, then $x^i \neq x^j$ because $U_i = U_j$. Therefore if C is a clique in an equilibrium network and $i, j \in C$, then $c_i = c_j$ and hence players in C have the same utility functions. □

Remark 1. Note that Theorem 3 holds also if condition (i) is replaced by the condition that derivative is strictly increasing on the diagonal. Of course, marginal utility from link formation may be so large as compared to the cost parameters c_i, that $s_{ii} = 0$ in equilibrium. Then there could exist reciprocal equilibria with a complete network even if players have different cost parameters c_i.

We show next that if a game has bilateral strategic complements, then with the same or slightly weaker assumptions as in Theorem 3 there exists an equilibrium such that the equilibrium network is complete. By Theorem 3 this equilibrium cannot in general be reciprocal.

Theorem 4. *Suppose $G = \big(N, (S_i)_{i\in N}, (u_i)_{i\in N}\big)$ is a semi-symmetric link formation game with bilateral strategic complements such that: (i) derivative is strictly decreasing on the diagonal; (ii) parameters $p_j > 0$ and $c_i > 0$ of equation* (3) *are taken from compact intervals P and C, respectively; (iii) the function V in* (3) *is linear. Assume also that if all players $i \in N$ would have the same parameters $p \in P, c \in C$, then the corresponding symmetric game would have a symmetric interior equilibrium s. Then there exists an equilibrium with a complete network.*

Sketch of a proof. For each pair $p \in P$ and $c \in C$ of parameters there exists a symmetric interior equilibrium. By assumption (i) these equilibria can be naturally ordered. An equilibrium for G with a complete network can be formed from these symmetric equilibria. For details see Appendix. □

Remark 2. Note that Theorem 4 holds also if condition (i) is replaced by the condition that derivative is strictly increasing on the diagonal. In such a case an interior equilibrium is not stable in the usual best reply dynamics. The assumption of Theorem 4 that derivative is strictly decreasing (or strictly increasing) is critical as demonstrated in Example 1.

For games with bilateral strategic substitutes we have the following.

Theorem 5. *Suppose $G = \big(N, (S_i)_{i\in N}, (u_i)_{i\in N}\big)$ is a semi-symmetric link formation game with bilateral strategic substitutes such that: (i) parameters $p_j > 0$ and $c_i > 0$ of equation* (3) *are taken from compact intervals P and C, respectively; (ii) the function V in* (3) *is linear. If for each p_j and c_i, and for each $z \in (0, 1/(n-1)]$ there exists $x \in (0, 1/(n-1)]$ such that $p_j \partial U(x,z)/\partial x_1 - c_i = 0$, then there exists an equilibrium with a complete network.*

Proof. See Appendix.

4.1 Efficiency and Stability of Equilibria

We have focused on equilibria such that the corresponding network is complete, or has complete components. It turns out in our framework completeness of equilibrium networks is in many cases closely related to stability and Pareto optimality of equilibria.

Given a game $G = \big(N, (S_i)_{i\in N}, (u_i)_{i\in N}\big)$, a strategy profile s is Pareto optimal, if there is no other profile s' such that $u_i(s') \geq u_i(s)$ for all $i \in N$ and $u_j(s') > u_j(s)$ for some $j \in N$. A network corresponding to a strategy profile s is Pareto optimal, if s is a Pareto optimal strategy profile.

The following result gives conditions under which a Pareto optimal network must be complete. Intuitively, the condition that guarantees completeness of the equilibrium network is that the marginal utility from privacy is less than the marginal benefit from a sufficiently small reciprocal investment.

Proposition 1. *Suppose a link formation game $G = \big(N, (S_i)_{i\in N}, (u_i)_{i\in N}\big)$ is such that for each $i \in N$ and $z^i \in (0,1]$ there exists $x^i \in (0, z^i)$ such that*

$$\frac{\partial U_{ij}(x^i, x^i)}{\partial x_1} > V_i'(z^i - x^i), \forall i, j \in N, i \neq j,$$

and that each U_{ij} is concave. If $s \in S$ is Pareto optimal and $s_{ii} > 0, \forall i \in N$, then $s_{ij}, s_{ji} > 0$, $\forall i, j \in N$.

Proof. Suppose to the contrary that $s \in S$ is Pareto optimal and $s_{ii} > 0, \forall i \in N$, but $s_{ij} = 0$ for some $i, j \in N$. Since $U_{ij}(0, s_{ji}) = 0$ and $U_{ji}(s_{ji}, 0) = 0$, Pareto optimality of s implies that $s_{ji} = 0$. By assumption, there exists $x^i < s_{ii}$ and $x^j < s_{jj}$ such that

$$\frac{\partial U_{ij}(x^i, x^i)}{\partial x_1} > V_i'(s_{ii} - x^i), \ \frac{\partial U_{ji}(x^j, x^j)}{\partial x_1} > V_j'(s_{jj} - x^j).$$

Since U_{ij} and V_i are concave functions, these inequalities hold for every $x \in (0, \min\{x^i, x^j\})$ as well. Given such an x, consider a strategy profile s' that is otherwise like the profile s, except that $s'_{ij} = s'_{ji} = x$, and $s'_{ii} = s_{ii} - x$, $s'_{jj} = s_{jj} - x$. Then $u_i(s') > u_i(s)$ and $u_j(s') > u_j(s)$ while $u_k(s') = u_k(s)$ for all $k \neq i, j$, and therefore s is not Pareto optimal, a contradiction. □

Remark 3. Proposition 1 holds for example when each U_{ij} is a strictly concave Cobb-Douglas function. The functions V_i can then be any concave, strictly increasing functions. Note that Proposition 1 holds also if functions U_{ij} have decreasing derivative on the diagonal, which is a weaker assumption than concavity.

An equilibrium s and the corresponding network are called s is *strongly pairwise stable*, if there is no strategy profile s' such that $u_i(s') > u_i(s)$ and $u_j(s') > u_j(s)$ for some $i, j \in N$, when $s_k = s'_k$ for all $k \in N \setminus \{i, j\}$ (Bloch and Dutta 2009). The following result states conditions such that the network corresponding to a strongly stable equilibrium must be complete.

Proposition 2. *Suppose a link formation game $G = \big(N, (S_i)_{i\in N}, (u_i)_{i\in N}\big)$ is such that for each $i \in N$ and $z^i \in (0,1]$ there exists $x^i \in (0, z^i)$ such that*

$$\frac{\partial U_{ij}(x^i, x^i)}{\partial x_1} > V_i'(z^i - x^i), \forall i, j \in N, i \neq j,$$

and that each U_{ij} is concave. If $s \in S$ is a strongly pairwise stable equilibrium and $s_{ii} > 0, \forall i \in N$, then $s_{ij}, s_{ji} > 0$, $\forall i, j \in N, i \neq j$.

Proof. The proof of Proposition 1 applies here. □

Remark 4. If the functions V_i satisfy $\lim_{z \to 0^+} V_i'(z) = +\infty$, then $s_{ii} > 0$ must hold at any equilibrium.

Remark 5. The main lesson of Propositions 1 and 2 is *not* that Pareto optimal networks are always complete, or that strongly stable equilibria have complete networks. Utilities from some links may be so low that these links are not formed either for efficiency or for equilibrium reasons. The lesson of these propositions is that network structures that have *complete components* often appear as efficient solutions or as equilibrium networks of a strongly pairwise stable equilibrium.

However, one can easily verify that if utility functions have the following Cobb-Douglas form, then *any* Pareto optimal or strongly pairwise stable equilibrium *must* have a complete network:

$$u_i(s) = \sum_{k \neq i} p_{ik} s_{ik}^{a_i} s_{ki}^{b_i} + c_i s_{ii}^{d_i},$$

where all parameters are strictly positive, $a_i + b_i < 1$, and $d_i < 1$.

In network literature efficiency is usually defined by using the utilitarian welfare function: those strategy profiles that maximize the sum of utilities are efficient. While such strategy profiles are Pareto optimal, not all Pareto optimal profiles satisfy this efficiency criterion.

If the functions U_{ij} are concave, then the utility functions u_i are concave on a simplex. In such a case each Pareto optimal strategy profile maximizes a *weighted* sum of players' utilities. The (positive) weights depend on the profile in question. If also the functions V_i satisfy $\lim_{z \to 0^+} V_i'(z) = +\infty$, then $s_{ii} > 0$ must hold at every Pareto optimal s, for all i.

Acknowledgment I thank an anonymous referee for many valuable comments. I'm grateful to Nizar Allouch, Mitri Kitti, Anne van den Nouweland, Matti Pihlava, Nicole Tabasso, Quan Wen, and seminar participants at the PET 2014 meeting in Seattle for discussions and useful comments. I thank the Academy of Finland, the Yrjö Jahnsson Foundation, and the OP-Pohjola Group Research Foundation for financial support.

References

Bauman, L. (2015). Time allocation in friendship networks. Mimeo, Department of Economics, University of Hamburg.

Bloch, F. and Dutta, B. (2009). Communication networks with endogenous link strength. *Games and Economic Behavior*, 66, 39–56.

Cabrales, A., Calvó-Armengol, A. and Zenou, Y. (2011). Social interactions and spillovers. *Games and Economic Behavior*, 72, 339–360.

Currarini, S., Jackson, M. O. and Pin, P. (2009). An economic model of friendship: homophily, minorities, and segregation. *Econometrica*, 77, 1003–1045.

Jackson, M. O. and Wolinsky, A. (1996). A strategic model of economic and social networks. *Journal of Economic Theory*, 71, 44–74.

Jackson, M. O. and Zenou, Y. (2015). Games on networks. In Young, P. and Zamir, S. (eds.), *Handbook of Game Theory with Economic Applications, Vol. 4*. Amsterdam, Elsevier Science, 95–163.

Salonen, H. (2015). Equilibria and centrality in link formation games. *International Journal of Game Theory*, DOI: 10.1007/s00182-015-0514-6.

Appendix

Proof of Theorem 1. ($\Leftarrow$) A reciprocal equilibrium s is nontrivial if $s_{ij} = s_{ji} > 0$ for at least two players i, j. Let $N_1, \dots, N_k$ be the complete components of the equilibrium network. If N_t has $m \geq 2$ members, then there exists $x = s_{ij}, i, j \in N_t$ such that

$$\frac{\partial U(x,x)}{\partial x_1} - V'(1-(m-1)x) \geq 0.$$

Since V is concave, the inequality (11) holds for this x.

($\Rightarrow$) Suppose that inequality (11) holds. Let m be the largest number, $m \leq n$, such that

$$\frac{\partial U(z,z)}{\partial x_1} - V'(1-(m-1)z) \geq 0$$

holds for some $z \in (0, 1/(m-1)]$. Clearly $m \geq 2$. Either there exists $z < 1/(m-1)$ such that this inequality is actually an equality, or else the inequality is satisfied by $z = 1/(m-1)$.

If $m = n$, then $s_{ij} = z$ for all $i, j \in N,\ j \neq i$, is a reciprocal equilibrium with a complete network. If $m < n$, then let k be the largest integer such that $km \leq n$. Choose k disjoint subsets N_t of N such that $|N_t| = m$ for all $t = 1, \dots, k$. If the union of these subsets does not cover N, then let N_{k+1} be the residual subset.

Define $s \in S$ by setting $s_{ij} = z$ for all $i, j \in N_t,\ j \neq i,\ t = 1, \dots, k$ (and also for $i, j \in N_{k+1}$ if this subset is nonempty) defines a nontrivial reciprocal equilibrium such that subsets N_t are complete components of the equilibrium network. □

Proof of Theorem 2. Since G is anonymous, $u_i = U_i + V_i$. By assumption, if all players have the same utility function u_i, there is a symmetric equilibrium s^i such that the resulting network is complete. Since u_i has constant derivative on the diagonal and V_i is strictly concave and increasing, the symmetric equilibrium s^i is unique. If every s^i is such that $s^i_{jk} = 1/(n-1)$, we are done. So we may assume that s^1 is the equilibrium in which $s^1_{ij} = x^1$ takes the smallest value, $i \neq j$, and $x^1 < 1/(n-1)$. Note that there may be another equilibrium s^k such that $s^k_{ij} = x^1$.

Construct a reciprocal equilibrium recursively as follows.

Step 1. Let N_1 be the subset of players for whom the following first order condition holds:

$$\frac{\partial U_i(x^1,x^1)}{\partial x} = V_i'\big(1-(n-1)x^1\big). \tag{A1}$$

By assumption, $|N_1| \geq 1$. If $N_1 = N$, the recursion ends. If $|N_1| < n$, then there exists at least one player for whom the left hand side of equation (A1) is greater than the right hand side.

Step 2. Let $x^2 \in (0,1)$ be the least number such that $x^1 < x^2$ and the following weak inequality is satisfied for at least one player:

$$\frac{\partial U_i(x^2,x^2)}{\partial x_1} \geq V_i'\big(1-n_1x^1-(n-n_1-1)x^2\big). \tag{A2}$$

Since the derivative of U_i is constant on the diagonal and V_i is strictly concave, such an x^2 exists uniquely. Let N_2 be the set of players for whom equation (A2) holds. If $|N_1|+|N_2|=n$, the recursion ends, because $N_1 \cap N_2 = \emptyset$. If $|N_1|+|N_2|<n$, continue the recursion to Step 3. Since there are n players, there is Step k, $k>2$, as follows.

Step k. Let $x^k \in (0,1)$ be the least number such that $x^{k-1} < x^k$ and the following weak inequality is satisfied for at least one player:

$$\frac{\partial U_i(x^k, x^k)}{\partial x} \geq V_i'\big(1 - \sum_{t<k} n_t x^t - (n_k - 1)x^k\big). \tag{A3}$$

By assumption and the previous Steps, such a number x^k exists uniquely. The subset N_k of players for whom equation (A3) holds, satisfies $|N_1|+\cdots+|N_k|=n$ and $\{N_1,\ldots,N_k\}$ is a partition of N.

Given player $i \in N$, let m be such that $i \in N_m$. Define $s_{ij} = x^t$ for all $j \neq i$ such that $j \in N_t$ and $t < m$. For $j \neq i$ such that $j \in N_t$ and $m \leq t$, let $s_{ij} = x^m$. Let $s_{ii} = 1 - \sum_{t<m} n_t x^t - \big[\big(\sum_{m \leq t} n_t\big) - 1\big]x^m$.

By construction s is a reciprocal equilibrium such that the resulting network is complete. □

Proof of Theorem 4. Denote the set of "types" of players by $T = P \times C$. Given any type $t = (p,c) \in T$, if all players had this type, then by assumption there exists a symmetric interior equilibrium s^t satisfying

$$\frac{p}{c}\frac{\partial U(x^t, x^t)}{\partial x_1} = 1$$

where $s^t_{ij} = x^t$ and $s^t_{ii} = 1-(n-1)x^t$ for all $i,j \in N, j \neq i$. Since derivative is strictly decreasing on the diagonal, these symmetric equilibria can be ordered so that $x^t > x^{t'}$ iff $p/c > p'/c'$, where $t=(p,c)$ and $t'=(p',c')$.

Let $\overline{p}$ and $\underline{p}$ be the greatest and least elements, respectively, of the interval P. Define analogously $\overline{c}$ and $\underline{c}$. So the symmetric equilibrium corresponding to the type $\overline{t} = (\overline{p}, \underline{c})$ has the largest x^t, denoted by $\overline{x}$. The symmetric equilibrium corresponding to the type $\underline{t} = (\underline{p}, \overline{c})$ has the least x^t, denoted by $\underline{x}$.

Suppose that there are k different types $t^1,\ldots,t^k \in T$ present in the player set N. Let N_m consists of all players whose type is $t^m, m=1,\ldots,k$.

Let us construct an equilibrium s with a complete network such that players in the same subset N_m treat each other reciprocally.

Step 1. Set $s_{ii} = y^{t^m}$, and $s_{ij} = x^{t^m}$, for all $i,j \in N_m$, for all $m=1,\ldots,k$. Note that the first order conditions of an interior equilibrium are satisfied by these choices. The values x^{t^m} and y^{t^m} are the same as in the symmetric equilibrium s^{t^m}.

Step 2. Take any players $i \in N_m$ and $j \in N_h$, $m \neq h$. Consider a two-person game with strategic complements between i and j. Let $t^m = (p,c)$ and $t^h = (p'c')$. Let b^t denote the best reply function of type $t = t^m, t^h$ against opponent's choices $x \in [\underline{x}, \overline{x}]$. The best replies for these types (unique by strict concavity of $U(\cdot,x)$) satisfy

$$\frac{p'}{c}\frac{\partial U(b^{t^m}(x),x)}{\partial x_1} = 1 = \frac{p}{c'}\frac{\partial U(b^{t^h}(x),x)}{\partial x_1}.$$

If $p'/c = p/c'$, then best replies are the same. If $p'/c < p/c'$, then $b^{t^m}(x) < b^{t^h}(x)$. Since $\underline{p}/\overline{c} \le p'/c < p/c' \le \overline{p}/\underline{c}$, we have also $\underline{x} \le b^t(\underline{x}) \le b^t(\overline{x}) \le \overline{x}$ for both types $t = t^m, t^h$. This holds since bilateral strategic complements implies $b^t(\underline{x}) \le b^t(\overline{x})$ (increasing best reply function). Strictly decreasing derivative on the diagonal implies $\underline{x} \le b^t(\underline{x})$ and $b^t(\overline{x}) \le \overline{x}$, because $b^t(x^t) = x^t$. But then by Tarski's fixed point theorem the mapping $(x_m, x_h) \longrightarrow (b^{t^m}(x_h), b^{t^h}(x_m))$ on $[\underline{x},\overline{x}] \times [\underline{x},\overline{x}]$ has a fixed point (x^{mh}, x^{hm}).

Consider the game between all players in the set $N_m \cup N_h$. Then note that the choices y^{t^m}, x^{t^m}, x^{mh} for players in N_m and the choices y^{t^h}, x^{t^h}, x^{hm} for players in N_h form an equilibrium, since the resource constraints are satisfies by the definition of the symmetric equilibria $s^{t^1}, \dots, s^{t^k}$, and the payoff of any player i is additively separable *w.r.t.* his opponents.

Since the types t^m and t^h were chosen arbitrarily, we have solved an equilibrium for the whole game. To see this, take any player i, and assume that $i \in N_m$. Then his choices satisfy the resource constraint:

$$y^{t^m} + |N_m - 1| x^{t^m} + \sum_{h \neq m} |N_h| x^{mh} = 1.$$

Since the first order conditions for maximum satisfied, we are done. □

Proof of Theorem 5. Let the "type set" be $T = P \times C$. Suppose that there are k different types $t^1, \dots, t^k \in T$. Let N_m consist of all players whose type is $t^m, m = 1, \dots, k$. We construct an equilibrium s such that players in the same subset N_{i_m} behave reciprocally.

Step 1. Suppose $i, j \in N_m$, so they both have the type $t^m = (p^m, c^m)$. Let $b^{t^m}(z)$ denote the unique best reply of either player to $z \in [0, 1/(n-1)]$. By assumption $b^{t^m}(1/(n-1)) \le 1/(n-1)$. If equality holds, then $x^{t^m} = 1/(n-1)$ is a reciprocal equilibrium in the game with player set N_m.

Suppose $b^{t^m}(1/(n-1)) < 1/(n-1)$. Let $I^* = \{z \mid b^{t^m}(y) < y, \forall y \in [z, 1/(n-1)]\}$ and $x^* = \inf I^*$. Note that x^* exists since $1/(n-1) \in I^*$. We want to show that $b^{t^m}(x*) = x^*$.

By bilateral strategic substitutes, $b^{t^m}(1/(n-1) < b^{t^m}(z)$ for all $z \in I^*, z < 1/(n-1)$ and by assumption $b^{t^m}(1/(n-1)) > 0$. By the Theorem of the maximum, the best reply $b^{t^m}(z)$ is a continuous function on $[x^m - \varepsilon, 1/(n-1)]$, for any $\varepsilon > 0$ such that $x^* - \varepsilon > 0$. By continuity, $b^{t^m}(x^*) \le x^*$. Again by continuity and the definition of I^*, this inequality cannot be strict, so $b^{t^m}(x^*) = x^*$. A reciprocal equilibrium in the game with player set N_m is obtained by setting $s_{ij} = x^* \equiv x^{t^m}$ for all $i, j \in N_m, i \neq j$.

Step 2. Suppose $i \in N_m$ and $j \in N_h$, $m \neq h$. Let $t^m = (p, c)$ and $t^h = (p', c')$. If $p'/c = p/c'$, then the choices given in *Step 1* apply. Given $x \in [0, 1/(n-1)]$, the best replies satisfy

$$\frac{p'}{c}\frac{\partial U(b^{t^m}(x),x)}{\partial x_1} = 1 = \frac{p}{c'}\frac{\partial U(b^{t^h}(x),x)}{\partial x_1}.$$

Now $b^{t^m}(x) < b^{t^h}(x)$ because $U(\cdot,x)$ is strictly concave function, and because $p'/c < p/c'$.

Consider the function $f(x) = b^{t^h}(b^{t^m}(x))$ on $[b^{t^m}(1/(n-1)), 1/(n-1)]$. This function is continuous, and $f(x) \leq 1/(n-1)$ for all x. At $x = b_i(1/(n-1))$, $f(x) \geq x$, since both best replies are decreasing functions. Hence there is a fixed point $x^{hm} = f(x^{hm})$. But then x^{hm} is the best reply of player j against $b^{t^m}(x^{hm}) = x^{mh}$, which in turn is the best reply of player i against x^{hm}.

Therefore $s_{ji} = x^{hm}, s_{ij} = x^{mh}$ forms an equilibrium when the player set is $N_m \cup N_h$.

Since the types t^m and t^h were chosen arbitrarily, we have solved an equilibrium for the whole game. To see this, take any player i, and assume that $i \in N_m$. Then his choices satisfy the resource constraint:

$$|N_m - 1|x^{t^m} + \sum_{h \neq m} |N_h| x^{mh} \leq 1.$$

Define $s_{ii} = y^{t^m}$ so that the resource constraint is satisfied as equality. Since the first order conditions for maximum satisfied, we are done. □

Czech Economic Review 9 (2015) 184–201
Acta Universitatis Carolinae Oeconomica

Experimental Results about Linguistic Voting

Manzoor Ahmad Zahid*, **Harrie de Swart****

Received 29 July 2015; Accepted 29 June 2016

Abstract In this paper, we describe the results of experiments in which about 7000 voters in the Netherlands were asked in three different waves to give their most favored party and to give an evaluation on a scale of 0 till 10 of eleven major Dutch parties. We have applied five different voting rules to determine the number of seats each party would obtain in Parliament. Different from what one might think, in general voters had no problem to give an evaluation of eleven major Dutch parties. Interestingly, many voters gave the same evaluation to two or more parties, something they cannot do if they can only vote for one party. Although Majority Judgment has not been designed for a seat distribution in parliament, we describe two possible ways which enable such a distribution.

Keywords Voting experiments, linguistic voting, plurality rule, range voting, approval voting, majority judgment, Borda majority count
JEL classification D71, D72

1. Introduction

As is well known there are many different election mechanisms and the result of an election may depend strongly on the election mechanism used. In order to get an idea what shifts would be caused in the seat distribution in parliament by applying different election mechanisms, we have applied several election mechanisms to the experimental results of three waves in each of which about 7000 voters were asked to mention their most favored party and to give an evaluation of eleven major Dutch parties on a scale of 0 till 10, where 10 stands for ‘excellent’, 9 for ‘very good’, 8 for ‘good’, 7 for ‘very satisfactory’, 6 for ‘satisfactory’, 5 for ‘almost satisfactory’, 4 for ‘unsatisfactory’, 3 for ‘very unsatisfactory’, 2 for ‘poor’, 1 for ‘very poor’ and 0 for ‘extremely poor’. The resulting seat distributions are summarized in Figure A1 in Appendix.

In Section 2 we will give some background information with respect to the experiments. In Sections 3 and 4 the details of the results obtained in the three waves of the experiment will be given. After a short description of the different election mechanisms, i.e., Plurality Rule, Range Voting, Approval Voting, Majority Judgment and the Borda Majority Count, we present for each of these election mechanisms the resulting seat distribution in Dutch parliament with 150 seats.

* COMSATS Institute of Information Technology, Department of Mathematics, Sahiwal Campus, Pakistan. E-mail: manzoor@ciitsahiwal.edu.pk.
** Corresponding author. Erasmus University Rotterdam, Faculty of Philosophy, Burgemeester Oudlaan 50, 3062 PA, the Netherlands. E-mail: deswart@fwb.eur.nl.

Although Majority Judgment has not been devised for a seat distribution of parties in parliament, we describe two ways to adapt Majority Judgment to enable such a seat distribution. One way is described in Subsection 4.3, the other way is what we call the Borda Majority Count, described in Subsection 4.4, in which the different evaluations *excellent, good, acceptable, poor* and *reject* are identified with the numbers 4, 3, 2, 1, 0 respectively.

Finally, we discuss and compare the outcomes under the different election mechanisms.

2. Background of the experiments

In 2006, CentERdata at Tilburg University received major NWO funding for the project: an advanced multi-disciplinary facility for Measurement and Experimentation in the Social Sciences (MESS). This NWO subsidy was instituted by the Cabinet with a view to boosting the Dutch knowledge economy and the climate for innovation in the Netherlands. These funds have been used to establish a new online panel of 5,000 Dutch households: the LISS panel (Longitudinal Internet Studies for the Social sciences). The panel is the core component of the MESS project and is based on a true probability sample of households. The LISS core study consists of 11 projects. Project Number 8, called Politics and Values, is a longitudinal study delivering a broad range of social core information about the panel members. It focuses on politics, social attitudes and values.

The results in this paper are based on the answers of the members of the LISS panel to the following questions in an online survey conducted three times between 2007 to 2011:

– If parliamentary elections were held today, for which party would you vote?

– How sympathetic do you find the political parties? You can assign each party a score between 0 and 10. 0 means that you find the party very unsympathetic, and 10 means that you find the party very sympathetic. If you are not familiar with a party, you can indicate this using the button ‘I don’t know’.

The voters were unaware of the different election mechanisms; only afterwards their votes have been used to determine a seat distribution in Dutch parliament according to different voting mechanisms. So, the word ‘sympathetic’ does not depend on the electoral rule.

The parties in question are:

CDA (Christelijk Democratisch Appel, Christian Democrat Party)
PvdA (Partij van de Arbeid; Labor Party)
VVD (Volkspartij voor Vrijheid en Democratie; Liberal Party)
SP (Socialist Party); GL (Green Left)
D66 (Democraten 66; Social-Liberal party); CU (Christian Union)
SGP (Staatkundig Gereformeerde Partij; Christian Reformed Party)
PVV (Partij voor de Vrijheid; Party for the Freedom, Groep Wilders)

PvdD (Partij voor de Dieren; Party for the Animals)
TON (Trots op Nederland; Proud of the Netherlands; Rita Verdonk)

In wave I, December 2007, the questionnaire was presented to 8204 panel members, and it was completed by 6811 respondents (83%). In wave II, December 2008, the questionnaire was presented to 8289 panel members, and it was completed by 6037 respondents (response percentage 73%). In wave III, December 2009, the questionnaire was presented to 9398 panel members, and it was filled out by 6386 respondents (response percentage 68%).

It is worth noticing that most of the respondents did give an evaluation of all major political parties in the Netherlands on a scale of 0 till 10. This scale is very familiar to all Dutchmen, because it is used at all education institutions. It is frequently thought that persons are not able to give an evaluation of so many parties, but the responses to the second question show that people are able to do so. This confirms the findings of Michel Balinski (Balinski and Laraki 2007b) in his experiment at the 2007 French presidential elections where about 2000 voters were asked to give an evaluation of the twelve presidential candidates.

The results of the answers to the first question, involving Plurality Rule, will be presented in Section 3. In Section 4 we summarize the results of the answers to the second question in waves I, II and III and apply Range Voting, Approval Voting, Majority Judgment and the Borda Majority Count to the data obtained.

To the best of our knowledge there are only few data available concerning linguistic voting. The reason is that the predominant question asked to voters usually is: how do you rank the different candidates? As argued by Balinski and Laraki (Balinski and Laraki 2011), however, the predominant question should be: how do you evaluate the different candidates? From an evaluation of the candidates one may easily deduce a ranking, but not conversely. Hence, evaluations are much more informative than mere rankings. We were surprised to find that data about evaluations by the voters of the different parties were available at CentER data and we know of no other data of this type other than those collected by Balinski and Laraki in their experiments around French presidential elections in Balinski and Laraki (2007b).

3. Question 1

Many nations around the world use the Plurality voting system to determine the outcome of elections, although it is well known there are many objections against this system. In the Netherlands one uses a list system of proportional representation, where each party has a list containing the names of the candidates for that party. Although it is possible to vote for a particular candidate on that list, most voters will just vote for the first candidate on the list, in other words for the party in question. A particular candidate on the list is only sure of a seat if the number of votes he or she obtains passes a certain threshold. If a party is entitled to, say, n seats, then the first n persons on the list obtain a seat in parliament, unless someone lower on the list has already obtained a seat by his own.

Table 1. Results in wave I, II and III for Question 1

Party	Plurality vote			Party seats		
	I	II	III	I	II	III
CDA	885	727	692	30	30	24
PvdA	609	637	506	21	26	18
VVD	417	427	533	14	17	18
SP	628	454	426	21	19	15
GL	339	251	346	11	10	12
D66	162	413	703	05	17	25
CU	240	150	191	08	06	06
SGP	102	091	073	03	03	02
TON	724	143	059	24	06	02
PVV	269	333	681	09	13	24
PvdD	131	088	128	04	03	04
Total	4506	3714	4338	150	150	150

In Table 1 we list the results in wave I, II and III for Question 1: *If parliamentary elections were held today, for which party would you vote?*

We have computed the number of seats for each party by applying Jefferson's method, also known as d'Hondt's method (see Balinski and Young 1982): find a divisor x such that the whole numbers contained in the quotients of the different parties sum to the required total of 150. Each party is given its whole number of seats. The divisor that does the job is 29 for wave I, 23.8 for wave II and 28.1 for wave III.

4. Question 2

Table 2 shows the responses to Question 2: *How sympathetic do you find the political parties? You can assign each party a score between 0 (very unsympathetic) and 10 (very sympathetic).*

In Table 2, 999 stands for 'I do not know'. In the next Subsections we will apply Range Voting, Approval Voting, Majority Judgment and the Borda Majority Count to these data.

4.1 Range Voting

In Range Voting (RV), due to Smith (2015), voters are asked to evaluate the different alternatives on a scale which, for instance, may range from 0 to 99, but also other ranges may be taken. The scores for a particular candidate may be added up or one may take the average of the scores for the candidate in question. The candidate with the highest score or average wins. The larger the range of values, the smaller the probability that a tie will occur. In such an exceptional case one might simply toss a coin. Range Voting has many nice properties (see Smith 2015), but it is very vulnerable for manipulation:

Table 2. Responses to Question 2

Party	10	9	8	7	6	5	4	3	2	1	0	999
CDA-I	35	93	610	1229	1350	1130	698	487	291	154	226	495
II	28	92	514	1113	1161	1066	621	398	234	118	138	523
III	45	92	550	1055	1122	1037	582	502	319	188	222	637
PvdA-I	33	77	404	1104	1403	1236	778	525	361	193	230	454
II	24	70	491	1197	1375	1019	557	344	215	112	116	486
III	38	82	443	1086	1297	1062	579	420	326	197	226	594
VVD-I	14	57	285	768	1178	1288	946	736	458	255	274	539
II	7	40	303	735	1192	1216	744	598	305	143	148	575
III	17	72	362	843	1211	1168	737	534	335	174	200	696
SP-I	81	124	597	969	1220	1034	721	537	362	223	266	664
II	46	123	481	991	1117	995	595	444	264	135	150	665
III	37	119	419	893	1184	1042	624	467	337	196	214	817
GL-I	51	110	444	931	1151	1100	749	571	405	256	270	760
II	34	107	434	942	1083	1046	639	435	299	171	157	659
III	45	129	490	961	1110	972	656	459	318	189	224	796
D66-I	11	52	195	620	1202	1369	838	607	403	273	234	994
II	16	73	372	888	1190	1158	620	416	226	119	112	816
III	35	148	580	1195	1190	995	525	332	215	124	146	864
CU-I	40	78	299	720	1123	1049	768	645	451	314	379	932
II	23	68	202	501	860	1036	822	621	461	302	332	778
III	34	67	231	552	949	1055	745	608	500	337	376	895
SGP-I	44	40	96	207	482	869	895	790	672	477	727	1499
II	36	45	79	192	442	856	832	758	596	410	538	1222
III	31	34	85	214	537	916	817	718	603	434	653	1307
TON-I	89	111	372	539	563	663	522	537	470	447	1227	1258
II	25	20	119	295	423	736	547	615	618	541	1256	811
III	12	15	78	225	443	663	667	637	667	653	1355	934
PVV-I	66	71	220	369	472	519	521	570	630	551	2049	760
II	58	56	165	308	384	539	473	577	544	551	1702	649
III	81	84	269	384	483	488	424	472	458	528	1989	689
PvdD-I	92	91	224	536	791	863	665	639	627	565	808	897
II	74	50	186	378	619	823	562	631	583	573	790	737
III	92	74	232	502	709	830	606	640	551	610	685	818

voters who have a slight preference for *A* over *B* might strategically give 1 point to *B* and 99 to *A* in order to achieve that their favored candidate wins.

In the survey of the LISS panel the range consists of the numbers from 0 till 10. It is worth noticing that many participants gave the same evaluation to different parties.

For each of the eleven parties we have computed the average score and next we have for each wave applied Jefferson's method as described in Section 3 in order to obtain the number of seats for each party. The divisor for wave I is 0.31, for wave II 0.311 and for wave III 0.324. The resulting seat distributions for the three waves are shown in Table 3.

Table 3. Seat distributions using Range Voting

Party	AVG	Seats	Party	AVG	Seats
CDA-I	5.30	17	CU-I	4.56	14
II	5.39	17	II	4.34	13
III	5.19	16	III	4.51	13
PvdA-I	5.07	16	SGP-I	3.40	10
II	5.49	17	II	3.53	11
III	5.41	16	III	3.65	11
VVD-I	4.63	14	TON-I	3.69	11
II	4.93	15	II	2.98	9
III	5.06	15	III	2.82	8
SP-I	5.12	16	PVV-I	3.51	11
II	5.32	17	II	2.77	8
III	5.14	15	III	2.91	8
GL-I	4.93	15	PvdD-I	3.85	12
II	5.19	16	II	3.61	11
III	5.37	16	III	5.05	15
D66-I	4.62	14			
II	5.25	16			
III	5.78	17			

As one can see in wave III, the Plurality Rule attributes many more seats to CDA, D66 and PVV (24, 25 and 24 respectively) than Range Voting does (CDA 16, D66 17 and PVV 8 seats). This may be explained by the fact that relatively many voters have CDA, D66 or PVV as first choice, while at the same time relatively many voters dislike these parties. On the other hand, Range Voting is beneficial for CU (13 seats in wave III), SGP (11 seats) and TON (8 seats) which under the Plurality Rule only receive 6, 2 and 2 seats, respectively in wave III. This may be explained by the fact that there are relatively few voters who have CU, SGP and TON as their first choice, but relatively many voters who appreciate these parties.

4.2 Approval Voting

Approval voting (AV) (Brams 1976; Brams and Fishburn 1978, 1983) is a voting procedure in which voters can vote for, or approve of, as many candidates as they wish. A voter divides the candidates into two groups: those which he or she approves of and those which he or she does not approve of. Candidates who are approved by a voter receive one point, while candidates who are not approved by a voter receive zero points.

Since in the Dutch education system a mark below 6 is considered as insufficient, it seems reasonable to identify approval with a mark between 6 and 10 and disapproval with a mark between 0 and 5. Doing so, Table 4 above shows the election outcomes for the three different waves in our survey.

Table 4. Seat distributions using Approval Voting

Party	Approved Vote			Party Seats		
	I	II	III	I	II	III
CDA	3317	2908	2864	21	20	19
PvdA	3021	3157	2946	19	22	19
VVD	2302	2277	2505	14	16	16
SP	2991	2758	2652	19	19	17
GL	2687	2600	2735	17	18	18
D66	2080	2539	3148	13	18	20
CU	2260	1654	1833	14	11	12
SGP	869	794	901	5	5	6
TON	1674	882	773	10	6	5
PVV	1198	971	1301	7	6	8
PvdD	1734	1307	1609	11	9	10
Total	24133	21847	23267	150	150	150

The seat allocation of the different parties in Table 4 has again been calculated by using Jefferson's method, described in Section 3. The divisor for wave I is 155, for wave II 139 and for wave III 150.

What strikes us is that the traditionally larger parties like CDA and PvdA get more seats under Approval Voting than under Range Voting; the same holds for the parties SP and GL. However, parties like SGP, TON, PVV and PvdD are clearly worse off under Approval Voting than under Range Voting.

4.3 Majority Judgment

Balinski and Laraki (2007a, 2011) ask the voters to give an evaluation of the candidates, like in Range Voting. While from an evaluation of all alternatives one can construct a (weak) preference ordering of the alternatives, conversely, from a given (weak) preference ordering of the alternatives—as assumed in the original Borda Count—one cannot deduce an evaluation of the alternatives. So, an evaluation of the alternatives by an individual voter gives (much) more information than a preference ordering of the alternatives by the voter in question.

In their experiments Balinski and Laraki (2007b) use the grades in the set $\{$*excellent, very good, good, acceptable, poor, reject*$\}$. But in order to decrease the possibilities for manipulation, they do not take the average or the sum of the evaluations as the final result of a candidate, but the (lower) median value of the evaluations. They call their election mechanism Majority Judgment (MJ), and define the *majority grade* $f^{maj}(A)$ of candidate A as the lower median value of the grades assigned by the voters to A. For instance, if A gets the evaluations 2, 5, 7, 8, 9, its majority grade will be 7, and if A gets the evaluations 2, 5, 7, 9, its majority grade will be 5.

Clearly, when the majority grade of A is greater than the majority grade of B, we declare that $A \succ_{maj} B$, i.e., A is socially preferred to B according to Majority Judgment. In their recent paper Balinski and Laraki (2016) explain how to define the social ranking $\succ_{maj}$ also in the case that A and B have the same majority grade. It goes too far to repeat their definition and motivation at this place. Here we restrict ourselves to an alternative definition, $\succ_{mg}$ which is useful in the case of large electorates and which corresponds with the original definition $\succ_{maj}$ in all cases where it gives a decision. Balinski and Laraki (2016) define the *majority gauge* of a candidate A as a triple (p_A, α_A, q_A), where $\alpha_A = f^{maj}(A)$ is the majority grade of A, p_A is the number of grades given to A strictly above its majority grade, q_A is the number of grades given to A strictly below its majority grade.

Now A is socially preferred to B according to the majority gauge, $A \succ_{mg} B$, or $(p_A, \alpha_A, q_A) \succ_{mg} (p_B, \alpha_B, q_B)$, iff $\alpha_A \succ \alpha_B$ or ($\alpha_A = \alpha_B$ and $p_A > max\{p_B, q_A, q_B\}$) or ($\alpha_A = \alpha_B$ and $q_B > max\{q_A, p_A, p_B\}$). So, e.g., (20, *good*, 30) $\succ_{mg}$ (40, *ac*, 10), (30, *good*, 20) $\succ_{mg}$ (25, *good*, 10), and (20, *good*, 22) $\succ_{mg}$ (20, *good*, 25). Balinski and Laraki also show that if $A \succ_{mg} B$, then $A \succ_{maj} B$.

In Table 5 we have translated the LISS panel data which used the evaluations from 10 till 0 into the grades used by Balinski and Laraki (2007b), by identifying 10 with *ex*(cellent), 9 with *vg* (very good), 8 with *go*(od), 7 and 6 with *ac*(ceptable), 5 and 4 with *po*(or), 3, 2, 1 and 0 with *re*(ject), more or less in accordance with the meaning of the marks 10 till 0 in the Dutch education system. We have computed the majority grade of each party and shown it in Table 5 by using boldface digits. In addition, we have indicated the values p_A and q_A for each party A. We did not take into account the voters who said that they could not give an evaluation of the party in question.

To illustrate, in wave III the majority gauge of D66 is (763, *ac*, 2337) and the one of CDA is (687, *ac*, 2850). Because $q_{CDA} > max\{q_{D66}, p_{D66}, p_{CDA}\}$, by definition D66 is socially preferred to CDA according to the majority gauge, D66 $\succ_{mg}$ CDA and hence also D66 $\succ_{maj}$ CDA in wave III.

It is not self evident how one may allocate seats to parties using Majority Judgment. We see two possibilities: the one that is described in Subsection 4.4, identifying the grades $\{$*ex(cellent), go(od), ac(ceptable), po(or), re(ject)*$\}$ with the numbers 4, 3, 2, 1, 0 respectively, and the procedure described below in this Subsection.

The procedure we apply in this subsection is as follows: given a wave, let γ be the highest majority grade of the different parties. In our example, $\gamma = ac$ for all three waves. For each party A let $\beta(A)$ be the number of voters who gave A an evaluation higher or equal to γ. Next apply Jefferson's method described in Section 3 to determine the number of seats of each party, such that the total number of seats is 150. The divisor for wave I is 157, for wave II 140 and for wave III it is 150.

As one can see in Figure A1 in Appendix, using this procedure there are only minor differences between the seat distributions under Approval Voting and the Majority Judgment. This comes as no surprise, since for the seat allocation we have taken into account the number of voters who gave a grade higher than or equal to $\gamma = ac$ which is more or less the number of voters who approved of the party in question. With this procedure for determining the number of seats, in all three waves SGP, TON, PVV and PvdD receive less seats under Majority Judgment than under Range Voting.

4.4 The Borda Majority Count

Let A be an alternative and $\{g_1, g_2, \ldots, g_k\}$ be the set of grades, with $g_1 > g_2 > \ldots > g_k$. Let p_j be the number of voters who gave grade g_j to A, where $j = 1, 2, \ldots, k$. The *Borda Majority Count* BMC(A) of A is defined by BMC(A) := $p_1 \cdot (k-1) + p_2 \cdot (k-2) + \ldots + p_k \cdot 0$.

$$\mathrm{BMC}(A) = \sum_{j=1}^{k} p_j \cdot (k-j)$$

For instance, suppose we have five grades: *ex*(cellent), *go*(od), *ac*(ceptable), *po*(or) and *re*(ject). Then we assign 4 points to grade *ex*, 3 points to grade *go*, 2 points to grade *ac*, 1 point to grade *po* and 0 points to grade *re*. Now suppose that 10 voters evaluate a party A as follows:

ex	*go*	*ac*	*po*	*re*
1	2	3	3	1

Then BMC(A) = $1 \times 4 + 2 \times 3 + 3 \times 2 + 3 \times 1 + 1 \times 0 = 19$. It is illuminating to realize that BMC(A) equals the sum of the cumulative evaluations (numbers) as shown in the following table:

at least	*ex*	*go*	*ac*	*po*
	1	3	6	9

Notice that 1 + 3 + 6 + 9 = 19 = BMC(A). This is explained by the fact that in the last table of cumulative grades the grade *ex* is taken into account 4 times, the grade *go* is taken into account three times, etc.

In order to transform the data from the LISS panel into evaluations in terms of the language just mentioned, i.e. $\{ex, go, ac, po, re\}$, we have identified *ex* with the grades 10 and 9, *go* with 8 and 7, *ac* with 6 and 5, *po* with 4 and 3, and *re* with 2, 1, 0 and 999. The seat distribution among the different parties has been computed by applying Jefferson's method to the Borda Majority Counts of the different parties. The resulting

Table 5. Seat distributions using Majority Judgment

Party	*p*	*ex*	*vg*	*go*	*ac*	*po*	*re*	*q*	$\beta(A)$	#Seats
CDA-I	0738	35	93	610	**2579**	1828	1158	2986	3317	21
II	0634	28	92	514	**2274**	1687	0888	2575	2908	20
III	0687	45	92	550	**2177**	1619	1231	2850	2864	19
PvdA-I	3021	33	77	404	2507	**2014**	1309	1309	3021	19
II	0585	24	70	491	**2572**	1576	0787	2363	3157	22
III	0563	38	82	443	**2383**	1641	1169	2810	2946	19
VVD-I	2302	14	57	285	1946	**2234**	1723	1723	2302	14
II	2277	07	40	303	1927	**1960**	1194	1194	2277	16
III	2505	17	72	362	2054	**1905**	1243	1243	2505	16
SP-I	2991	81	124	597	2189	**1755**	1388	1388	2991	19
II	0650	46	123	481	**2108**	1590	0993	2583	2758	19
III	2652	37	119	419	2077	**1666**	1214	1214	2652	17
GL-I	2687	51	110	444	2082	**1849**	1502	1502	2687	17
II	2600	34	107	434	2025	**1685**	1062	1062	2600	18
III	2735	45	129	490	2071	**1628**	1190	1190	2735	18
D66-I	2080	11	052	195	1822	**2207**	1517	1517	2080	13
II	2539	16	073	372	2078	**1778**	0873	0873	2539	18
III	0763	35	148	580	**2385**	1520	0817	2337	3148	20
CU-I	2260	40	078	299	1843	**1817**	1789	1789	2260	14
II	1654	23	068	202	1361	**1858**	1716	1716	1654	11
III	1833	34	067	231	1501	**1800**	1821	1821	1833	12
SGP-I	2633	44	040	096	0689	1764	**2666**	0000	0869	05
II	0794	36	045	079	0634	**1688**	2302	2302	0794	05
III	0901	31	034	085	0751	**1733**	2408	2408	0901	06
TON-I	1674	89	111	372	1102	**1185**	2681	2681	1674	10
II	2165	25	020	119	0718	1283	**3030**	0000	0882	06
III	2103	12	015	078	0668	1330	**3312**	0000	0773	05
PVV-I	2238	66	071	220	0841	1040	**2511**	0000	1198	07
II	1983	58	056	165	0692	1012	**3374**	0000	0971	06
III	2213	81	084	269	0867	0912	**3447**	0000	1301	08
PvdD-I	1734	92	091	224	1327	**1528**	2639	2639	1734	11
II	1307	74	050	186	0997	**1385**	2577	2577	1307	09
III	1609	92	074	232	1211	**1436**	2486	2486	1609	10

Table 6. Seat distributions using the Borda Majority Count

Party	*ex*	*go*	*ac*	*po*	*re*	BMC	#Seats
CDA-I	128	1839	2480	1185	1166	12,174	19
II	120	1627	2227	1019	1013	10,834	18
III	137	1605	2159	1084	1366	10,765	17
PvdA-I	110	1508	2639	1303	1238	11,545	18
II	094	1688	2394	0901	0929	11,129	18
III	120	1529	2359	0999	1343	10,784	17
VVD-I	071	1053	2466	1682	1526	10,057	15
II	047	1038	2408	1342	1171	9,460	16
III	089	1205	2379	1271	1405	10,000	16
SP-I	205	1566	2254	1258	1515	11,284	17
II	169	1472	2112	1039	1214	10,355	17
III	156	1312	2226	1091	1564	10,103	16
GL-I	161	1375	2251	1320	1691	10,591	16
II	141	1376	2129	1074	1286	10,024	17
III	174	1451	2082	1115	1527	10,328	17
D66-I	063	0815	2571	1445	1904	9,284	14
II	089	1260	2348	1036	1273	9,868	16
III	183	1775	2185	0857	1349	11,284	18
CU-I	118	1019	2172	1413	2076	9,286	14
II	091	0703	1896	1443	1873	7,708	13
III	101	0783	2004	1353	2108	8,114	13
SGP-I	084	0303	1351	1685	3375	5,632	08
II	081	0271	1298	1590	2766	5,323	09
III	065	0299	1453	1535	2997	5,598	09
TON-I	200	0911	1226	1059	3402	7,044	10
II	045	0414	1159	1162	3226	4,902	08
III	027	0303	1106	1304	3609	4,533	07
PVV-I	137	0589	0991	1091	3990	5,388	08
II	114	0473	0923	1050	3446	4,771	08
III	165	0653	0971	0896	3664	5,457	09
PvdD-I	183	0760	1654	1304	2897	7,624	11
II	124	0564	1442	1193	2683	6,265	10
III	166	0734	1539	1246	2664	7,190	11

seat distributions are shown in Table 6. The divisor for wave I is 640.70, for wave II 589 and for wave III it is 600.

The more voters there are, the smaller is the chance of a tie under the Borda Majority Count. Typically, the differences in the seat distribution under Range Voting (Smith 2015), Approval Voting (Brams 1976; Brams and Fishburn 1978, 1983), Majority Judgment (Balinski and Laraki 2007a,b, 2011) and the Borda Majority Count (Zahid and de Swart 2015) highest BMC than others parties. All other parties are almost consistent in their ranks. The main party PvdA has slightly improved his position over CDA. The BMC ranking position, in all waves are as under:

5. About the number of grades

In the LISS panel the voters could give an evaluation of the different parties on a scale from 10 (excellent) to 0 (reject), in other words, the common language was the set of grades $\{10, 9, 8, \ldots, 2, 1, 0\}$ familiar to every voter from the Dutch education system. One may wonder what language is appropriate and whether the outcome of an election depends on the language used. For that reason we have counted the number of voters who used k different grades, for $k = 1, \ldots, 10$. The results are in Table 7.

Table 7. Number of grades used by voters

Voters (%)	Grades
0.62	1
2.44	2
5.83	3
15.25	4
25.31	5
28.41	6
16.23	7
5.12	8
0.78	9
0.01	10

Only 0.01% of the voters used ten different grades to evaluate the parties and most voters (28.41%) used six different grades to evaluate all parties. As is clear from the table, almost half of the voters used 5 or less grades, 77.86% of the voters used six or less different grades and almost 85.2% of the voters used four to seven different grades. This is in line with the experimental results of Balinski and Laraki (2007b), who observed that the six grades (excellent, very good, good, acceptable, poor, reject) in their experiment were sufficient and no more grades were needed. For reasons of symmetry we slightly prefer the language {excellent, good, acceptable, poor, reject}, leaving out the term 'very good', because the term 'acceptable' is then precisely in the middle. In addition, it reduces the possibilities for manipulation, because one may only

Table 8. Frequency of grades

Grade	Percentage of use
0	17.99
1	5.18
2	7.13
3	9.27
4	11.29
5	15.52
6	15.25
7	11.15
8	5.20
9	1.24
10	0.8

reduce the evaluation of a candidate dishonestly by four points, instead of five when Balinski's language is used.

We have also counted how many times each grade has been used. The results are in Table 8. Notice that grades 5 and 6 were used most frequently.

6. Pairwise comparison

The results of pairwise comparisons of parties in percentages have been calculated from the original data in Table 2 obtained in the LISS panel taking the three waves together, and are shown in Table 9.

So, the first number 52 in the first row indicates that 52% of the voters prefer CDA to PvdA. As one can see in this table, in a pairwise comparison the party CDA defeated every other party except D66 and D66 defeated all other parties. Notice that

Table 9. Pairwise comparisons

	CDA	PvdA	VVD	SP	GL	D66	CU	SGP	TON	PVV	PvdD
CDA		52	63	50	51	49	68	73	73	77	67
PvdA	48		60	51	53	49	62	68	69	74	70
VVD	37	40		43	44	41	55	66	73	78	63
SP	50	49	57		53	49	61	67	70	76	73
GL	49	47	56	47		47	62	68	69	74	74
D66	51	51	59	51	53		67	73	72	77	73
CU	32	38	45	39	38	33		74	67	73	61
SGP	27	32	34	33	32	27	26		63	70	53
TON	27	31	27	30	31	28	33	37		70	44
PVV	23	26	22	24	26	23	27	30	30		34
PvdD	33	30	37	27	26	27	39	47	56	66	

Table 10. Percentage of voters giving the same evaluation

	PvdA	VVD	SP	GL	D66	CU	SGP	TON	PVV	PvdD
CDA	31	31	22	16	27	29	21	18	15	19
PvdA		28	29	31	30	26	20	16	14	20
VVD			25	24	29	27	25	21	17	20
SP				42	31	26	25	20	17	23
GL					40	27	25	20	17	25
D66						30	27	21	17	23
CU							43	23	19	24
SGP								30	24	26
TON									48	27
PVV										28

although D66 is the Condorcet winner, the parties CDA and PvdA get more seats when the Plurality Rule is applied (except in wave III). Van Deemen (1993) calls this the *More-Preferred, Less-Seats paradox.*

For each pair of parties we have also calculated from the original data in Table 2, taking the three waves together, what percentage of voters is indifferent between the two parties in question. The results are shown in Table 10.

Notice that almost half of the voters (48%) is indifferent between TON and PVV, which is not surprising if one knows the political landscape in the Netherlands. A similar remark can be made for CU and SGP, but now with 43%. Among CDA, PvdA, VVD and SP, roughly speaking at most 30% of the voters is indifferent between any pair of them.

7. Summary

Balinski and Laraki's Majority Judgment (Balinski and Laraki 2011) asks the voter to give *evaluations* of the alternatives instead of giving a first preference or a ranking of the candidates. In this way, the voter is able to provide much more information than in the traditional framework of social choice theory, which was inspired by Arrow (1963, 1983): in Balinski and Laraki's framework the voter may give the same evaluation to two or more candidates and also is able to express to which degree he prefers one candidate to another one. From an evaluation of the candidates one may deduce a weak ordering or ranking of them, but conversely, one cannot deduce an evaluation of the candidates from a given ranking. In his Majority Judgment this extra information is also used in the aggregation of the individual evaluations to an evaluation by the society. In order to reduce the possibilities for manipulation, Balinski and Laraki take the median value of the evaluations by the voters as the final social evaluation. In experiments they have shown that, contrary to what is frequently thought, voters are quite able to give evaluations of relatively many (about 10) candidates. Their idea of asking the voters for evaluations instead of rankings is inspired by the practice of many contests, for instance of ice-skating. However, in elections for parliament or for choosing a president, to the best of our knowledge, voters are nowhere asked to give their *evaluations* of the different candidates or parties; instead, in most cases they just have to mention one candidate or, at best, a ranking of the candidates.

By taking the median value of the evaluations by the voters as the social outcome, it frequently is the case that several candidates have the same median value and consequently there usually are many ties. Balinski and Laraki propose two tie breaking rules and show that if a candidate A is socially preferred to candidate B according to the majority-gauge, then A is also socially preferred to B according to the majority ranking.

There is a number of examples where application of Majority Judgment yields controversial results. That is, the social outcomes look at first sight counter-intuitive. However, Balinski and Laraki argue in Chapter 16 of their book (Balinski and Laraki 2011) that these surprising results are very reasonable outcomes and after all are not counter-intuitive at all. They only look counter-intuitive at first sight, because we are used to think in the traditional framework of Arrow.

An item not touched by Balinski and Laraki is how their Majority Judgment may be used to give a seat distribution for parties in parliament and it is not immediately clear how this may be done. We present two ways to do so: the first one is described in Subsection 4.3 and the second way is—once the votes have been casted in linguistic terms—by replacing the linguistic grades by appropriate numbers, resulting in what we have called the Borda Majority Count.

In order to avoid the controversial examples, to make the computations for determining the social outcome more simple and in order to be able to compute a seat distribution for parties in parliament, we have made a number of changes in the procedure of Balinski and Laraki:

(i) We use the same set of grades as they do, say $\{$*ex*(cellent), *go*(od), *ac*(ceptable, *po*(or), *re*(ject)$\}$, for reasons of symmetry leaving out the grade *vg* (very good). Voters are asked to evaluate the candidates using these linguistic grades.

(ii) After the voters have casted their votes, *ex* is identified with the number 4, *go* with 3, *ac* with 2, *po* with 1 and *re* with 0.

(iii) Next for each alternative we simply add up the number grades obtained by that alternative, which we call the Borda Majority Count of that alternative.

In this way one obtains one or more winners and a social ranking of the alternatives. The chance that two candidates have the same Borda Majority Count is relatively low, in particular when there are many voters.

We call this procedure the Borda Majority Count (Zahid and de Swart 2015), because on the one hand it reminds us of the Borda Count (Saari 2001, 2008) and on the other hand it reminds us of Majority Judgment. The controversial examples disappear when applying the Borda Majority Count and it becomes easy to apply the Borda Majority Count if one wants to compute a seat distribution for parliament. Although the Borda Majority Count has a number of nice properties, compared with Majority Judgment we also pay a price: it is easy to manipulate. When I know that two candidates

A and B are close competitors, and A is my favorite one, then I may dishonestly give B a very low evaluation. However, the difference for the Borda Majority Count of B will be at most 4, frequently less than 4. In this respect the Borda Majority Count, although a special case of Range Voting (Smith 2015), is less manipulable than Range Voting, where the range of possible numbers usually is (much) larger.

The Borda Majority Count has with the Borda Count in common that they both compute scores of the alternatives, but it differs from the Borda Count because it uses as input evaluations of the candidates instead of rankings, which are much less informative than evaluations. The Borda Majority Count may be conceived as a special case of Range Voting, but it differs from Range Voting by using evaluations in terms of a small set of linguistic expressions, well understood by everyone involved, instead of evaluations in terms of a fairly large set of natural numbers. The Borda Majority Count is similar to Majority Judgment in that both use a common language consisting of a relatively small set of linguistic grades, but it differs from Majority Judgment by not taking the median value of the evaluations given to a candidate by the voters, but by summing up or averaging the numbers associated with the linguistic grades given to the candidate in question.

Anyway, while it is not clear at all how Majority Judgment may be used to give a seat distribution for parties in parliament, the Borda Majority Count seems an appropriate way to do so.

8. Conclusion

We have applied five different election mechanisms to the data of the LISS panel, showing the evaluations by its members of the most well-known Dutch parties on an eleven point scale, ranging from 0 (reject) till 10 (excellent), as familiar from the Dutch education system. In the case of Approval Voting (AV), Majority Judgment (MJ) and the Borda Majority Count (BMC) we had to transform these data to the language of the election mechanism in question, i.e., $\{0, 1\}$ for Approval Voting, $\{0, 1, 2, 3, 4, 5\}$ for Majority Judgement and $\{0, 1, 2, 3, 4\}$ for the Borda Majority Count. Generally speaking, the seat distributions under Range Voting, Approval Voting, Majority Judgement and the Borda Majority Count are more or less similar, except for SGP and TON, which get clearly less seats under AV and MJ than under RV and BMC. Plurality Rule (PR) is clearly beneficial for some parties, like CDA (in all three waves), and to a lesser degree for PvdA, D66 and PVV, while Range Voting and the Borda Majority Count are beneficial to CU, SGP and TON. The last observation may be explained by the fact that these parties may not be approved of by many of the voters, but still obtain a lot of respect by these voters.

More than 50% of the participants used five or six grades. It is striking that the members of the panel clearly were able to give evaluations of the eleven parties involved and many gave different parties the same evaluation. This shows that one should not ask the voters to give a ranking of the parties and that it is not reasonable to ask the voter to select just one party from the list, as is done under the Plurality Rule.

Acknowledgment The authors thank CentERdata and NWO for making available to them the data of project number 8 in the LISS core study.

References

Arrow, K. (1963). *Social Choice and Individual Values*. New York, Wiley.

Arrow, K. (1983) *Collected Papers, Volume 1: Social Choice and Justice*. Cambridge, Harvard University Press.

Balinski, M. and Young, H. P. (1982). *Fair Representation: Meeting the Ideal of One Man, One Vote*. New Haven, Yale University Press.

Balinski, M. and Laraki, R. (2007a). *A Theory of Measuring, Electing and Ranking*. PNAS, 104(21), 8720-8725.

Balinski, M. and Laraki, R. (2007b). *Election by Majority Judgment: Experimental Evidence*. Paris, Ecole Polytechnique, Centre National De La Recherche Scientifique, Cahier No. 2007-28.

Balinski, M. and Laraki, R. (2011). *Majority Judgment*. Cambridge MA, MIT Press.

Balinski, M. and R. Laraki, *Majority Judgment vs Majority Rule*. Paris, Ecole Polytechnique, Centre National De La Recherche Scientifique, Cahier no 2016-04.

Brams, S. (1976). *Paradoxes in Politics*. New York, Free Press.

Brams, S. and Fishburn, P. (1978). Approval Voting. *American Political Science Review*, 72, 831–47.

Brams, S. and Fishburn, P. (1983). *Approval Voting*. Boston, Birkhauser.

Saari, D. (2001). *Decisions and Elections*. New York, Cambridge University Press.

Saari, D. (2008). *Disposing Dictators; Demystifying Voting Paradoxes. Social Choice Analysis*. New York, Cambridge University Press.

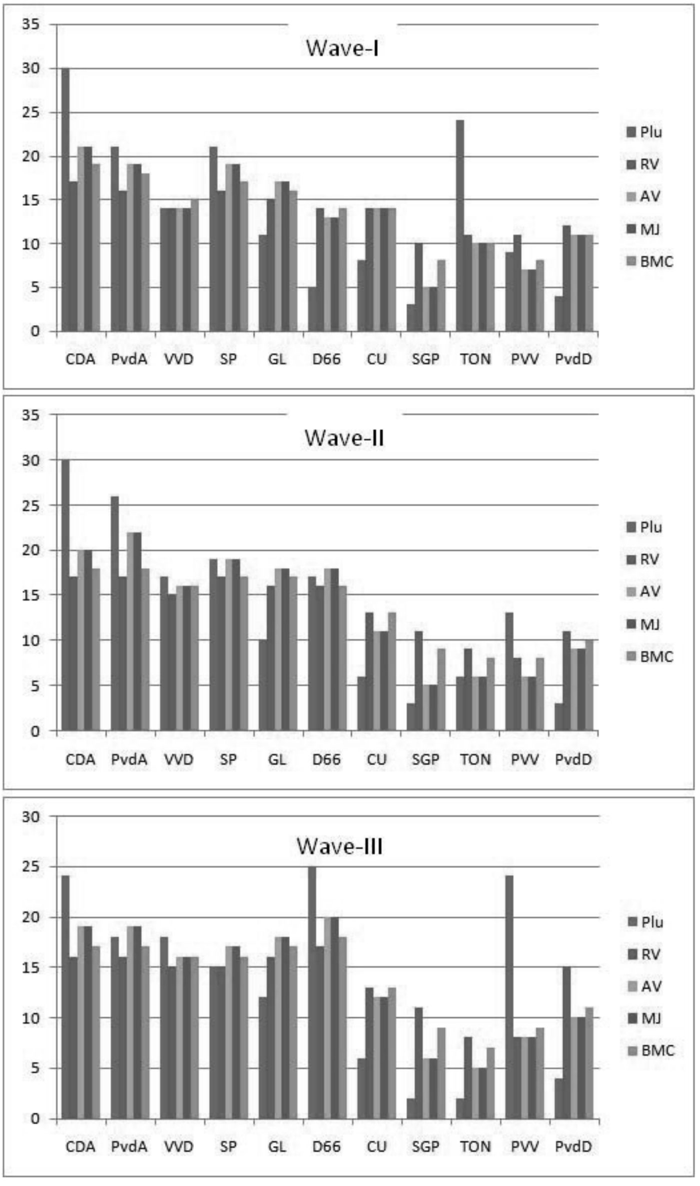
Smith, W. D. (2015) *Range Voting*. rangevoting.org/RangeVoting.html

Van Deemen, A. (1993). Paradoxes of Voting in List Systems of Proportional Representation. *Electoral Studies*, 12, 234–241.

Zahid, M. A. and de Swart, H. (2015). The Borda Majority Count. *Information Sciences*, 295, 429–440.

Appendix

Figure A1. Overview of the resulting seat distributions in wave I, II and III



Czech Journal of Economics and Finance

The journal (ISSN 0015-1920) is an English-language, double-blind refereed academic journal published in Prague by Charles University in association with the Czech National Bank and the Czech Ministry of Finance. It has been published continuously since 1951 (originally as Finance a úvěr), and has recently switched to the open-access format (free online access to all articles). The journal's readers include macroeconomic policymakers, academic researchers, university teachers and students, economists working in the public and private sectors, financial officers, and financial consultants.

Journal highlights

Impact Factor since 1997.
The journal is indexed in the Web of Science (the Social Sciences Citation Index), Current Contents Connect, JEL, ECONLIT, SCOPUS, and ABI Inform. It is also registered in the Ulrich's International Journals Archive. The journal is distributed worldwide by the EBSCOhost Electronic Journals Service and by the ProQuest LLC.
All articles freely accessible online at journal.fsv.cuni.cz.
Hard copies of the journal sold for a fee.
Short time between acceptance for publication and actual posting.
Submission fee: none.

Call for papers

We encourage submission of original unpublished papers written in good English. We focus on monetary economics, public finance, financial economics, and international economics, but are open to high-quality papers from all fields of modern economics.

We prefer empirically oriented papers, but do not exclude review articles or theoretical contributions provided that they are of high quality and relevant to the journal's aims. Among empirical papers, we prefer those relevant to a broad international audience, i.e., covering a range of countries or analyzing topics clearly relevant outside a single country. We also publish more narrowly oriented short articles (4000 words or less).

Contacts

E-mail: redakce@fsv.cuni.cz; Web: journal.fsv.cuni.cz

Kybernetika

Kybernetika, published by Institute of Information Theory and Automation of the Academy of Sciences of the Czech Republic, is the bi-monthly international journal dedicated for rapid publication of high-quality, peer-reviewed research articles in fields covered by its title. Kybernetika traditionally publishes research results in the fields of Control Sciences, Information Sciences, System Sciences, Statistical Decision Making, Applied Probability Theory, Random Processes, Fuzziness and Uncertainty Theories, Operations Research and Theoretical Computer Science, as well as in the topics closely related to the above fields.

The journal was established in 1964. It has been monitored in the Science Citation Index since 1977 and it is indexed in databases of Mathematical Reviews, Current Mathematical Publications, Current Contents ISI Engineering and Computing Technology, SCOPUS, and Zentralblatt MATH. Journal Impact Factor for 2015 was 0.628.

The full texts of all papers published in Kybernetika are available electronically at journal web page **http://www.kybernetika.cz**.

The current issue, **Volume 52, Number 3,** includes the following papers:

Homo Oeconomicus

Homo Oeconomicus (HOEC) is devoted to the study of classical and neoclassical economics, public choice, collective decision-making, law & economics, and philosophy & economics. It focuses on the study of political and social institutions. For further details please see **www.homooeconomicus.org**.

HOEC is a refereed journal. **Submissions** are to be sent electronically (preferably in PDF format) to the Managing Editor:

Manfred J. Holler, e-mail: holler@econ.uni-hamburg.de
University of Hamburg, IAW
Von-Melle-Park 5
D-20146 Hamburg, Germany

HOEC publishes one volume of 500 pages (3–4 issues) a year. The subscription price is EUR 90.00 per volume (including postage and shipment). Please send your orders for back issues and subscriptions to the publisher:

ACCEDO Verlag, e-mail: accedoverlag@web.de
Gnesener Str. 1
D-81929 Munich, Germany

The **electronic archive** of the journal Homo Oeconomicus (full text since volume 17, 2000) is retrievable via www.gbi.de. See also www.accedoverlag.de.

Operations Research and Decisions

The quarterly *Operations Research and Decisions* includes original scientific papers covering the full scope of operations research. Special emphasis is placed on quantitative methods and their application as a tool to solve economic, social and management problems. In particular, the articles published deal with the following subjects:

Mathematical modeling

Forecasting

Econometric and statistic methods

Optimization and simulation

Management information systems and systems supporting decision making

Computer applications in optimization

The papers presented contain substantial theoretical discussions concerning the methodology of operational research and decisions, as well as descriptions of quantitative methods and applications of these techniques to solving practical problems of an economic and social nature.

The quarterly *Operations Research and Decisions* has been published under this title since 1991. It continues the tradition of the journal entitled "Papers on Science and Forecasting" which was edited at Wrocław University of Technology during the eighties. Before 2010, papers were published mainly in Polish and occasionally in English. Since that year the whole journal has been printed in English.

The full texts of articles are freely available from the *Operations Research and Decisions* web site and, since 2005, through the EBSCO database serving the best interests of the scientific community. The journal is abstracted and indexed in: ARIANTA, BazEkon, BazTech, CEJSH, DOAJ, EBSCO, Index Copernicus Journals Master List, Inspec, MathSciNet, RePEc,Thomson Reuters – Emerging Sources Citation Index,Thomson Reuters – Web of Science Journal Citation Report, Ulrichs Periodicals Directory, Zentralblatt MATH.

Journal of Applied Economic Science

Journal of Applied Economic Science (ISSN 1834-6110) is a young economics and interdisciplinary research journal, aimed to publish articles that should contribute to the development of both the theory and practice in the field of Economic Sciences. The journal seeks to promote the best papers and researches in management, finance, accounting, marketing, informatics, decision/making theory, mathematical modelling, expert systems, decision system support, and knowledge representation.

The Journal appeals for experienced and junior researchers, who are interested in one or more of the diverse areas covered by the journal. It is currently published quarterly with three general issues in Winter, Spring, Summer and a special one, in Fall.

The primary aim of the Journal has been and remains the provision of a forum for the dissemination of a variety of international issues, empirical research and other matters of interest to researchers and practitioners in a diversity of subject areas linked to the broad theme of economic, business, management, accounting, finance, information technologies. Subject areas include, but not restricted to:

Quantitative and qualitative research methodology and practice
Strategic, public sector and human resource management
Entrepreneurship and small & medium sized enterprises (SME's)
Marketing and e-business
Economics, finance and financial management
Organizational theory and behavior
Supply chain management
Information systems, technology, innovation and operations management
Educational management and leadership

All papers will first be considered by the Editors for general relevance, originality and significance. If accepted for review, papers will then be subject to double blind peer review. Full author guidelines are available at the website: http://cesmaa.eu/journals/jaes/.

Journal of Applied Economic Science is indexed in CEEOL, EBSCO, RePEc, and SCOPUS databases.