AARHUS UNIVERSITY

Coversheet

This is the accepted manuscript (post-print version) of the article.

Contentwise, the post-print version is identical to the final published version, but there may be differences in typography and layout.

How to cite this publication

Please cite the final published version:

Brandt, U. S., & Svendsen, G. T. (2019). How robust is the welfare state when facing open borders? An evolutionary game-theoretic model. *Public Choice*, 178(1-2), 179–195.

<https://doi.org/10.1007/s11127-018-0622-5>

Publication metadata

Title:	How robust is the welfare state when facing open borders? An evolutionary game-theoretic model
Author(s):	Urs Steiner Brandt, Gert Tinggaard Svendsen
Journal:	<i>Public Choice</i> , 178(1-2), 179–195
DOI/Link:	https://doi.org/10.1007/s11127-018-0622-5
Document version:	Accepted manuscript (post-print)

General Rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

**How robust is the welfare state when facing open borders?
An evolutionary game-theoretic model**

Urs Steiner Brandt

Department of Environmental and Business Economics, University of Southern
Denmark, Niels Bohrs Vej 9, 6700 Esbjerg. Denmark,
usb@sam.sdu.dk

&

Gert Tinggaard Svendsen

Department of Political Science, Aarhus University,
Bartholins Allé 7, DK, 8000 Aarhus C, Denmark
gts@ps.au.dk

Abstract

How robust is the welfare state when confronting open borders? To answer that question, we develop an evolutionary game-theoretic model combined with an ingroup-outgroup model. The simulations reveal that welfare states in general will transform into low-welfare societies unless the societies in time generate a sufficiently large amount of social recognition of the reciprocators in such a crisis. The recognition implies that the “always cooperators” in favour of supportive policies towards free riders need to step down and hand over privileges to those willing to reciprocate, namely the “willing punishers”. The open-border society is modelled by letting a small amount of random types enter the society each year. Interestingly, it is not the defectors who compromise high-welfare societies. Instead, it is the excessive presence of cooperators who crowd out the reciprocators, thus making society increasingly vulnerable to free riding. This accentuates the need for timely recognition and actions against the risk of moving towards a low-welfare society.

Keywords: Evolutionary game theory; welfare state; free riding; cooperation; migration.

JEL codes: D60, H11, P52, Z13

Acknowledgements: An earlier version of this paper was presented at the European Public Choice Meeting in Budapest, April 2017, and we deeply thank the other participants. Furthermore, a special note of thanks goes to three anonymous reviewers and the editor for helpful comments.

How robust is the welfare state when facing open borders? An evolutionary game-theoretic model

Abstract

How robust is the welfare state when confronting open borders? To answer that question, we develop an evolutionary game-theoretic model combined with an ingroup-outgroup model. The simulations reveal that welfare states in general will transform into low-welfare societies unless the societies in time generate a sufficiently large amount of social recognition of the reciprocators in such a crisis. The recognition implies that the “always cooperators” in favour of supportive policies towards free riders need to step down and hand over privileges to those willing to reciprocate, namely the “willing punishers”. The open-border society is modelled by letting a small amount of random types enter the society each year. Interestingly, it is not the defectors who compromise high-welfare societies. Instead, it is the excessive presence of cooperators who crowd out the reciprocators, thus making society increasingly vulnerable to free riding. This accentuates the need for timely recognition and actions against the risk of moving towards a low-welfare society.

Keywords: Evolutionary game theory; welfare state; free riding; cooperation; migration.

JEL codes: D60, H11, P52, Z13

1. Introduction

The welfare state is well-described in the literature (see, e.g., Esping-Andersen 1990; Bertilsson and Hjorth-Andersen 2009; van Kersbergen and Mannow 2014; Bergh and Bjørnskov 2016). Such a state is unique, as a substantial fraction of income is redistributed from the ‘lucky’ to the ‘unlucky’ (Fong et al. 2006; Lindbeck 2008; Jensen and Svendsen 2011). Thus, the welfare state is perceived by many as a pleasant society with free access to healthcare and education and a relatively generous basic income, even in the presence of unemployment.

However, the more redistributive is the welfare state, the stronger is the incentive to free-ride, as one can let others work and pay taxes, while the free-rider obtains ‘money for nothing’. Without proper economic incentives, citizens can freely choose to free ride on social benefits and enjoy without contributing (Lindbeck 1995; Paldam 2004; Andersen 2004; Nannestad 2004; Nannestad and Green-Pedersen 2008; Svendsen and Svendsen 2016). One can receive social benefits without paying for them, and taxpayers will then subsidize the free riders.

Lindbeck et al. (1999, 2003) identify that the inherent moral hazard problem of social insurance schemes may be mitigated by strong social norms against living off benefits financed by others, in particular when eligibility for the welfare state’s is not caused by bad luck, but by free riding. The current analysis takes that view as its starting point by considering that the welfare state will be vulnerable if free riders are not punished. We conclude that a stable society is possible only if a sufficiently large proportion of the population is willing to punish free riders, even when it is privately costly to do so.

Since the welfare state basically facilitates ‘giving money to strangers’, it requires social control in the sense that people in general are expected to work and pay taxes when capable of doing so (Jensen and Svendsen 2011). Consequently, social punishment and the maintenance of a common norm for cooperation by those willing to punish may reduce the temptation to cheat, defect and free ride. However, what is missing in the literature is what happens to social control when the welfare state encounters open borders and a potential inflow of people to the society. Can it survive? Of course, an economic incentive exists to migrate from poor to rich welfare countries, i.e., ‘welfare tourism’ or ‘welfare magnets’ (Bergh and Bjørnskov 2016; Paldam 2004; Nannestad 2004; Borjas 1999). A number of empirical papers have addressed migrants’ use of welfare programs, e.g., Schultz-Nielsen (2017), Bratsberg et al. (2014), Sarvimäki (2011), Barrett and McCarthy (2008), Rowthorn (2008) and Nannestad (2007).

To our knowledge, however, the problem has not yet been addressed theoretically. Thus, our research question is ‘How robust is the welfare state when borders are open?’ To answer that question, we develop a model adding socioeconomic dimensions to an evolutionary game-theoretic model. Here, the political arena is described by an ingroup-outgroup model, such that society’s ingroup can earn additional payoffs through ingroup favouritism. Furthermore, the inflow of types into the society enables us to derive conditions under which a high-welfare society can be sustained with a continuous inflow of a small number of new players. Of specific interest is to derive conditions whereby a society consisting of only ‘good players’ can be stable against the potential growth of ‘free riders’. In particular, if the society takes

excessively long to identify the pending threat, it may not be able to prevent the inevitable transition from a high-welfare to a low-welfare regime.

The fundamental mechanism that drives our model is based on evolutionary game-theoretic reasoning. According to that reasoning, society consists of three main types of players: the Always Cooperator, the Always Defector and the Willing Punisher. Players engage in a prisoner's dilemma game, and those gaining above-average payoffs increase their proportion in society. Ostrom (2000) writes that experimental economics overall supports that three such types exist in social dilemma games. Other analyses also adopt this three-player setup, e.g., Imhof et al. (2005). While evolutionary game theory originally relates to biological processes, its relevance for analysing and understanding social science questions has long been realized (see, e.g., Morgen et al. 2012).

An important issue in the relevant literature is how cooperation can be sustained in repeated games. Many contributions address the issue of why cooperation emerges between self-interested agents, e.g., in culturally heterogeneous polities (Gächter et al. 2010). An extensive overview of evolutionary explanations of cooperation is provided by Nowak (2012). The main conclusion there is that, with no additional mechanisms in place, defectors will outcompete the other player types. One such class of mechanism consists of various types of punishments for norm-violators (Fehr and Gächter 2000; Fehr and Fischbacher 2004; Oliver 1980). A general distinction exists between direct and indirect punishment. The first such mechanism involves the direct punishment of those behaving antisocially. Conversely, indirect punishment implies excluding players known to behave “antisocially” from future beneficial engagements (Balafoutas et al. 2014).

Direct punishment is the costlier for the punisher for several reasons. It may take time to identify the norm-violators and actually to implement the punishment; the punisher also encounters the risk of retaliation. Indirect punishment also implies bearing an extra cost, as mutually beneficial trade arrangements with the excluded player no longer are possible. In our evolutionary game-theoretic setting, the punishment mechanism is embedded in the behaviour that defines ‘willing punishers’ and is classified as a direct punishment. The crucial issue is whether the presence of willing punishers will be crowded out by cooperators, defectors, or both (Brandt and Svendsen 2010; Poulsen and Svendsen 2005). If so, the welfare state cannot survive in the long run.

Compared to the traditional evolutionary game theory models, and as an important extension of Brandt and Svendsen (2010), we add a stylized social dimension described as ingroup-outgroup dynamics (see, e.g., Chen and Li 2009), which fits our three-type society nicely and introduces ingroup favouritism (Scheepers et al. 2006; Chen and Chen 2011; Fu et al. 2012). Here, the ingroup members have access to additional resources, which increases their relative fitness. The ingroup is determined by the number of defectors in the society; as long as sufficiently few of them exist, the cooperators constitute the ingroup. The larger is the population proportion of defectors, however, the more difficult it is for the cooperators to retain their privileged status, as ‘political justification costs’ will increase (Facchini 2016). Eventually, the ingroup’s privileges shift to the willing punishers.

Our model is inspired by the seminal work on the limits to growth of Meadows et al. (1972). Herein, the authors specify three conditions for a system to overshoot and collapse: exponential growth, stably operating boundaries within the system, and

delayed responses. Exponential growth will materialize in the population proportion of ‘always defect’ types; the boundaries determine the largest proportion of defector types that the society can support (without shifting to a low-payoff society). The response delay is represented by the population proportion of Always Defectors necessary to recognize the pending welfare-state sustainability problem and the willingness to respond to it by supporting the Willing Punishers. In our model, strong incentives are created and simulated by the presence of the Willing Punishers: they guarantee sufficient social pressure or policy outcome influence such that defectors do not multiply. Such behaviour is modelled by the reduction in defectors’ payoffs associated with punishment.

The analysis also provides new insights into another strand of literature. Mutations represent one factor that can block increases in the population proportion of defectors. Traulsen et al. (2009) state that if mutation will sustain all possible strategies at a high frequency in a population regardless of fitness, it is sufficient to disfavour defectors if punishment is targeted only at them. That result is challenged by Hauser et al. (2014), who model a situation wherein punishment targets all groups and is not tailored narrowly toward any one of them. Thus, since both Always Cooperators and Always Defectors likewise are subject to punishment, cooperation is not promoted by such a broad-based approach. We interpret mutations as exogenous inflows of the three player types, thereby proposing another channel through which the effect of mutations can be undermined by introducing ingroup bias and favouritism. The theoretical model addresses all types of welfare states (Esping-Andersen 1990). However, defection is more likely the more redistribution taking place in society, i.e., the Scandinavian welfare state would be the most vulnerable type.

In the following, Section 2 first introduces a socioeconomic evolutionary model and then explores how social control can change the model's predictions. Simulations follow in Section 3. Finally, Section 4 gives the conclusion.

2. The socio-economic evolutionary model

The building blocks of the model are as follows: Overall, society consists of a population that can be divided into three groups according to player types, which differ according to norms. The relative growth of the three groups depends on an internal mechanism (depending on the relative payoff of that group) and on an external mechanism (e.g., migration to that society). Income originates from two types of economic activity: engaging in private free-market exchange or receiving social benefits financed by other. Income is redistributed by a lump sum tax imposed by the group that holds the prevailing political view or controls the prevailing narrative; i.e., a political system exists wherein two main political narratives compete for the right to control the tax system.

2.1 Player types and payoffs

As a point of departure, the society consists of three types of individuals, which are defined by their attitudes towards cooperation and responses to non-cooperators. The structure follows Ostrom (2000), wherein two types of norm-using stereotypes exist together with a third group called rational egoists. Norm-using players always choose a pre-defined behavioural rule. This rule of behaviour or norm is also known as 'describing' how a specific group demonstrates conforming behaviour (Dong et al.

2015). The rational egoists, however, behave as described in game theory. This implies in our setting that they always free-ride and never cooperate.

Before defining the three main player types in detail, it is useful to describe how income is generated in the model. We assume that individuals can engage in two activities to earn income, either in free-market exchange or by not participating in markets whereby the only source of income is social welfare.¹ Free market exchange is undertaken in a prisoner's dilemma situation (see Table 1). We make two fundamental assumptions at this point: The players do not know which type they encounter in advance. Moreover, all player types will be revealed to the encounters once the interaction is completed.

[Table 1 about here]

The three types are denoted as 'Always Cooperators', 'Willing Punishers' and 'Always Defectors'. Both Always Cooperators and Willing Punishers enter and engage in person-to-person relationships with good intentions by choosing cooperation. The resulting payoffs are either c (if meeting an Always Cooperator or a Willing Punisher) or a lower one of a (from meeting an Always Defector).

However, in addition to the initial encounter, we introduce a post-encounter phase wherein players can punish other players perceived as norm-violators. Whether that option is exercised depends on the player's type. While the Always Cooperators accepts the initial outcome (payoff a in Table 1), the Willing Punishers does not and punishes the Always Defectors by adopting a costly penalty.² One rationale for Always

¹ We use the two-person prisoner's dilemma game (Table 1) as a reference for how individuals interact. Although highly stylized, the game has been applied widely in modelling actual human interactions.

² The experiment of Quervain et al. report a positive correlation between activity in certain parts of the brain and the willingness to incur costs to punish. The fact that people differ in those brain activities may explain why only a fraction of subjects are willing to punish in experiments.

Cooperators to accept defection is that those types perceive defectors as people who cannot contribute economically because of social- or health-related issues and therefore need to be helped. Such acceptance can be consistent with expressive behaviour, i.e., the self-interested quest for utility through acts and declarations that confirm a person's identity (Hillman 2010).

However, the Willing Punishers are not willing to accept that defectors receive benefits, so they will sacrifice own payoffs by relinquishing some of their income to punish the Always Defectors. The punishment is so that the resulting payoff for both the Willing Punisher and the Always Defector is $p < a$.³

The last and third type of player is the Always Defector, which in Ostrom (2000) are called rational egoists. Such types do not participate in income-generating activities and will receive an implicit welfare transfer of $d > c$ from the Always Cooperator; however, they are punished by the Willing Punisher and receive $p < 0$. Therefore, we can expand the payoff matrix to handle with the description above as being the relevant round-specific payoff realizations for the three types. The payoffs from the encounters are described in Table 2. The table reads as follows; each cell containing a payoff represents a row type meeting a column type and shows the payoff for the row type. For example, when an Always Cooperator meets an Always Defector, the payoff to the Always Cooperator is a .

[Table 2 about here]

This set-up allows us to derive the expected payoffs for each member of the three groups in a given period. In any period, t , society consists of P_t^{AC} persons of the

³ Let $\pi^i(i, j)$ be payoff of type i from meeting type j , while let $\pi^j(i, j)$ be type j 's payoff from meeting type i . The direct payoffs are $\pi^{WP}(WP, AD) = a$ and $\pi^{AD}(WP, AD) = d$. Punishment costs for the Willing Punisher type are $p^{WP} > 0$ and the effects of punishment costs on Always Defectors are $p^{AD} > p^{WP}$. For simplicity, we assume that $a - p^{AD} = d - p^{WP} = p$.

Always Cooperators type, P_t^{WP} persons of the Willing Punishers type and P_t^{AD} persons of Always Defector type. The total number of individuals in the society at time t is $P_t = P_t^{AC} + P_t^{WP} + P_t^{AD}$. We define x_t^i as the population proportion of type i , where $x_t^i = \frac{P_t^i}{P_t}$. We further define b^{ij} as the payoff for type i when encountering type j , where $i, j = \{\text{Always Cooperator, Always Defector, Willing Punisher}\}$. The expected payoff for a type i individual is given by:

$$\pi_t^i = b^{ii} \cdot x_t^i + b^{ij} \cdot x_t^j + b^{ik} \cdot x_t^k$$

$i, j, k = \{\text{Always Cooperator, Always Defector, Willing Punisher}\}$ and $i \neq j \neq k$. The average payoff in society in any period is:

$$\pi_t^A = \pi_t^{AC} \cdot x_t^{AC} + \pi_t^{WP} \cdot x_t^{WP} + \pi_t^{AD} \cdot x_t^{AD}$$

2.2 The composition of the society over time

The internal growth mechanism in our model is based on evolutionary game theory such that the three types compete for resources, and those types obtaining the largest proportion of the resources will have the most offspring.⁴

In its most general form, the population dynamics we use can be written as follows: the number of group i members in period t is determined by the number of individuals of the same type in the previous period plus changes depending on internal

⁴ Definition of an evolutionary stable strategy. The two strategies i, j , can be depicted as follows. Assume that strategy j has a very low frequency when it begins attempting to invade the population. Furthermore, assume that strategy i is stable if $E(i, i) > E(j, i)$ or $E(i, i) = E(j, i)$ and $E(i, j) > E(j, j)$.

(related to the configuration of payoffs in the proceeding period, i.e., π_{t-1}) and two external factors ($z = z_1, z_2$):

$$P_t^i = P_{t-1}^i + g(\pi_{t-1}, z)$$

The population dynamics we use is described by the following equation:

$$P_t^i = P_{t-1}^i + \beta \cdot (\pi_{t-1}^i - \pi_{t-1}^A) + g(z), \quad \beta > 0$$

The number of individuals belonging to a specific population type at period t depends to a larger extent on their number in the preceding period (P_{t-1}^i). Owing to replicator dynamics, the number is adjusted by the relative success the given type had in the previous period measured by $\beta \cdot (\pi_{t-1}^i - \pi_{t-1}^A)$. Finally, an exogenous factor increases the number of any group in every period (independent of the size of the group in the proceeding period).

The *internal* growth mechanism of the population is modelled by using a replicator function wherein types with above-average payoffs increase proportionately, while those receiving below-average payoffs decline proportionately. Therefore, the growth rate of a given type depends on its relative fitness compared to the average fitness in the population at a given time.

Several papers apply the logic of evolutionary dynamics to social phenomena (see, e.g., Morgan et al. 2012). Here, the replicator function is interpreted as above-average payoffs gaining popularity and, hence, are copied in a social learning process. The same methodology will also be the most reasonable means of interpreting the replicator dynamics in our paper. Therefore, because being a specific type no longer is

only a biological trait, but socially learned behaviour, we assume that the players need to regain their reputation in each round, i.e., their group membership. The only way this assumption is possible in our setting is by engaging the prisoner's dilemma described in Table 1. Moreover, we assume that playing the prisoner's dilemma game perfectly reveals all players' types in the given round. The second part, $g(z)$, contains two additional and separate mechanisms by which the groups can increase their relative sizes over time.

One of the main drivers of the results in this paper is that in each period, an equal number of individuals belonging to each group is added to the society in addition to the internal growth of the population. The first part, $g(z_1)$, summarizes the idea of an open society where in each period an equal number ($I_t > 0$) of the three types enter the population, independently of the payoffs in the previous period. Hence, in addition to the change in the numbers and proportions of types caused by the replicator process, an equal number of Always Cooperators, Willing Punishers and Always Defectors exogenously emerge in the society every year. This is interpreted as migration into the country.

We can now prove that the exogenous emergence of a small number of Always Defectors is sufficient to guarantee that, over time, the proportion of Willing Punishers falls below a threshold value $\bar{X}^{WP}$. Below that threshold, expansion of the Always Defectors in the society cannot be contained and their population proportion will move to 1.⁵

⁵ Proof that a proportion of Willing Punishers exist in society, denoted by $\bar{X}^{WP}$, below which $X_t^{AD} \rightarrow 1$. Let X_t^{AD} be small, denoted by ε_t . We have that $\Pi_t^{AC} - \Pi_t^{WP} = \varepsilon_t(d - p) > 0$. Therefore, over time, $\frac{X_t^{AC}}{X_t^{WP}}$ increases. If we assume that ε_t is extremely small such that the average payoff in society is approximately c , then we have that $\Pi_t^{AD} = dX_t^{AC} + pX_t^{WP}$, while $\Pi_t^A = c$. $\Pi_t^{AD} = \Pi_t^A$ if $d(1 - X_t^{WP}) + pX_t^{WP} = c \Rightarrow c \Rightarrow \bar{X}_t^{WP} = \frac{c-d}{p-d}$. If $d = 4, d = 6, p = -1 \Rightarrow \bar{X}^{WP} \cong 0.28$.

[Figure 1 about here]

An example of such a scenario is depicted in Figure 1. As shown, once $\bar{X}^{WP} \cong 0.28$, the proportion of the Always Defectors rapidly increases and outcompetes the two other types. Note that the initial population consists only of Always Cooperators and Willing Punishers where we typically begin with equal proportions of each in the simulations to come.

2.3 The political context and tax redistribution

Next, the *exogenous* population growth mechanism, referred to as (z_2), relates to the political system and the three types' ability to control a redistributive tax system. Our model is inspired by the ideas of ingroup-outgroup divisions and recurring ingroup-outgroup biases and favouritism dynamics, which we expand on to encompass the political-economic system in our model.

Scheepers et al. (2006, p. 359) defines ingroup bias and favouritism as 'The tendency to over-evaluate or favour one's own group (the ingroup) and/or to under-evaluate or derogate to which one does not belong (the outgroup)'. In the literature, ingroup favouritism has been identified as significant (Chen and Chen 2011) and as operating through several channels (see, e.g., Fu et al. 2012 for an overview of types and evolution of ingroup favouritism). Specific types of ingroup favouritism are labour market ingroup favouritism, e.g., in the form of preferential hiring of ingroup members (Dickinson et al. 2017) and ingroup members' access to or outgroup members' exclusion from powerful formal and informal networks (Bayer et al. 2008).

As we have defined three specific types of individuals in our society, it is natural to assume that three groups exist in society, each representing one of the three types.

The assumption now is that only one group can be an ingroup at a time and therefore be eligible for favouritism. That assumption is operationalized in the model by assuming that the prevailing ingroup has the right to collect taxes and redistribute the resulting revenue to its own members such that each ingroup member receives an additional payoff not available to outgroup members. The additional payoff translates into better fitness and relatively more offspring. Hence, the interpretation is that the ingroup has access to additional resources.

The next question that naturally arises is when does a group qualify to be an ingroup? In our model, the only threat to society is that the Always Defectors obtain a foothold in society, thus depressing the average payoff. However, recall that two distinct reactions to the Always Defectors are available to the other groups. We call them reaction narratives.

The Always Cooperators narrative is founded on generosity and meeting everyone with inclusive and supportive behaviour. Those supporting the Always Cooperators narrative believe that supportiveness and inclusion are keys to providing and sustaining a high average payoff. The Willing Punishers narrative is based on the ideas of reciprocity and being responsible. While cooperation is met with cooperation, defection is met with punishment. It follows that, according to the Willing Punishers narrative, the threat of retaliation and punishment is the main mechanism for eradicating free riding behaviour, thus securing high wealth to society.

Combining the above, two competing narratives in our model establish the best response to defection. The winning narrative is the one that dominates public understanding and therefore rationalizes the existing social order.⁶ It follows that the

⁶ For more on governance narratives, see, e.g., Bevir and Rhodes (2010) and Hajer and Wagenaar (2003).

group adopting the winning narrative also is dominant and, hence, can control the tax system.

The most important assumption is that the ingroup is the group whose narrative is closest to the observed scenario explained below. As long as the actual proportion of Always Defectors in society is below an exogenously defined threshold level (which we denote as $\bar{X}_t^{AD}$, or $\bar{X}^{AD}$), the Always Cooperators are the ingroup. One could state that that group claims the moral high ground because of its inclusive and supportive behaviour, which fits the observed structure of society best (= almost no defection). We call this the prevailing narrative.

If the proportion of Always Defectors rises, the narrative of inclusion and supportiveness becomes harder to justify (Facchini 2016).⁷ As soon as the observed proportion of Always Defectors is above $\bar{X}^{AD}$, the Willing Punishers' narrative becomes acceptable. That narrative is based on punishing defectors and therefore is radically different from the hitherto predominant narrative. The ingroup shifts from the Always Cooperators group to the Willing Punishers group. If the observed proportion of Always Defectors again declines, the punishment narrative becomes costlier to justify, and once X_t^{AD} falls below $\bar{X}^{AD}$, the ingroup switches again.

The size of $\bar{X}^{AD}$ is determined by how rapidly society identifies the expansion of Always Defectors and becomes willing to change the policy (narrative). The slower is

⁷ Facchini (2016) describes a feedback model wherein an individual believes that action A will result in outcome B. However, if information invalidates that belief, a state of cognitive dissonance occurs, which will open a 'window' of opportunity for reform. Denial or withholding of information can postpone that opportunity. If, over time, mounting evidence contradicts the initial belief, the justification cost increases. Again, according to Facchini (2016), individuals are inclined to change their ideologies when justification costs become excessively high. Such costs will rise if a growing number of individuals oppose the initial beliefs, if scientific authorities condemn those beliefs or simply when the cost of withholding evidence is excessively high. In our setting, changing ideology simply means supporting another narrative. See also Kaplan et al. (2016) for a neurological explanation of such political confirmation bias.

that process, the larger is $\bar{X}^{AD}$. As addressed previously in the introduction, the consequence can be one of two: either the process can no longer be halted, or the costs of doing so increasingly will rise. Note that policies (policy changes mediated by shifting narratives) are triggered only by observed events and not by rational (forward-looking) analysis and prognosis.⁸

The reward from belonging to the privileged ingroup in a society is the capability to redistribute wealth to its own group. To keep the analysis as general as possible, we assume that the ingroup's members have access to a redistributive tax system of the following type: each member of society is subject to a uniform lump-sum tax.

When everyone in the society is subject to this lump-sum tax, τ , the ingroup, i , is entitled to the whole tax revenue, which reads $R_t^i = \tau \cdot P_t^{TOT}$. The ingroup redistributes the tax revenue by an equal-sharing rule, such that each ingroup member receives $\frac{\tau}{x_t^i} > \tau$. We assume budget neutrality such that all tax-revenue is spent in the same period. It follows that the tax system does not affect average payoff in a period. The tax system will on the other hand generate additional payoff for the ingroup members and positively affect this group's relative growth in the next period. The tax system (at the group-level) has the following redistribution:

$$R_t^{AC} = \begin{cases} R^{AC} > 0 & \text{if } X_t^{AD} \leq \bar{X}^{AD} \\ 0 & \text{if } X_t^{AD} > \bar{X}^{AD} \end{cases}$$

$$R_t^{WP} = \begin{cases} R^{WP} > 0 & \text{if } X_t^{AD} > \bar{X}^{AD} \\ 0 & \text{if } X_t^{AD} \leq \bar{X}^{AD} \end{cases}$$

⁸ See Brandt (2014) for an explanation and consequences of such a policy process.

$$R_t^{AD} = 0$$

We do not explicitly model how the size of τ is determined. Instead we focus on the size of R_t^i , the total amount of revenue that the ingroup is able redistribute to itself. The size of R_t^{AC} is ultimately determined by the strength of the narrative, i.e. how pervasive the narrative materializes itself in the society and its institutions. R_t^{AC} appears as the reference value by which we compare R_t^{WP} , thus making it possible to identify how much transformation of a narrative is needed to stop the welfare state from collapsing.

As long as the actual proportion of defectors is below $\bar{X}^{AD}$, the Always Cooperators form the ingroup and hold the privilege of obtaining tax redistribution. Conversely, if $X_t^{AD} > \bar{X}^{AD}$, the narrative shifts, and the Willing Punishers control the redistributive tax system. For simplicity, we assume that the Always Defectors will not have access to ingroup favouritism. Hence, the general expected payoff is

$$E\Pi_t^i = X_t^{AC}\pi(i; AC) + X_t^{WP}\pi(i; WP) + X_t^{AD}\pi(i; AD) + R_t^i/P_t^i - \tau$$

3. Simulations

We are now in position to describe the population dynamics. Of special interest are questions related to the effect of the tax redistribution system, the delay factor, the composition of types over time, and the evolution of social welfare, i.e., the society-average payoff.

We begin with a reference scenario that will serve as a benchmark for evaluating how changes in the model affect the abovementioned factors. The reference scenario

storyline is as follows: we have a society that consists of an equal proportion of Always Cooperators and Willing Punishers. Since all are cooperating, welfare is high and society-average welfare is at c . A small inflow of individuals, including Always Defectors; however, the presence of sufficiently many Willing Punishers implies that those types cannot multiply. Consequently, the prevailing narrative supports the Always Cooperators, and this generates additional payoffs for that group. We arbitrarily set $\bar{X}^{AD} = 0.1$, while $R^{AC} = 0.5$ and $R^{WP} = 0$. The game begins with 100 Always Cooperators and 100 Always Defectors. The inflow of individuals in each round is five of each type. The beta parameter determines the speed of change as it relates the average fitness of any player type to changes in the numbers of offspring in the subsequent period. Thus, an equal proportion of each player type, based on identities and not fitness, enters society each year.

The result is presented on the left of Figure 2. Average society welfare is calculated on the basis of 1000 periods. As shown, average society welfare initially is high; however, once the proportion of Willing Punishers falls below the $\bar{X}^{WP}$ threshold, the Always Defectors begin to increase their proportions. When the proportion of the Always Defectors exceeds the $\bar{X}^{AD}$ threshold, where the Always Cooperators no longer own the right to receive additional benefits from the tax system, no brake prevents a society populated exclusively by Always Defectors, except for the small inflow of the other types in each period. Welfare falls to its lowest possible level in this society, and no cooperation occurs. Note that an increase in β would hasten the transition towards a population consisting solely of Always Defectors.

The first adjustment we propose is to incorporate a positive R^{WP} . The question is whether a size of R^{WP} exists that can counteract an increase of the Always Defectors.

The simulation implies that, as soon as the proportion of Always Defectors is above $\bar{X}^{AD} = 0.1$, the right to own the distributional tax system shifts from the Always Cooperators to the Willing Punishers. If R^{WP} is sufficiently high relative to R^{AC} , a second possible development in society is feasible. At a value of $R^{WP} = 1.0$, the reward for the Willing Punishers is sufficiently large to sustain those types in the population to an extent that reduces the size of Always Defectors back below $\bar{X}^{AD} = 0.1$; thereafter, a new cycle emerges. A version of such a society is presented on the right side of Figure 2. That society exhibits a cyclical pattern where the Always Cooperator group and the Willing Punisher group alternate between being the ingroup or the outgroup, such that a relatively high-welfare society is sustained.

[Figure 2 about here]

The simulations generally can result in two states: a high-average welfare state where the Always Defector growth rate will be reversed and one wherein the two other types are eliminated once the Always Defectors increase in proportion. The parameter values determining which state will occur are determined by the relative sizes of R^{WP} , R^{AC} and $\bar{X}^{AD}$. Note, however, that the effects on welfare and population proportions will vary in the same state for various parameter values. From the simulations, it can be derived that a threshold value of R^{WP} exists above which the increase in Always Defectors can be reversed.

In Tables 3 and 4, we further examine the relationship between the sizes of R^{WP} and R^{WP} necessary to counter the increase in Always Defectors. In Table 3, we adjust R^{WP} such that mode 1 remains in place, while in Table 4, we set $R^{WP} = 1.5$.

[Table 3 about here]

In Figure 3, we depict two of the scenarios from Table 3 to illustrate how the dynamics evolve.

[Figure 3 about here]

[Table 4 about here]

Note first that, for any R^{AC} , if the value of R^{WP} is below the values shown in the brackets in the first column of Table 1, society inevitably ends up in mode 2. Table 4 highlights that the larger is R^{AC} (interpreted as stronger ingroup favouritism), the smaller is the proportion of Always Cooperators on average in the society, the lower their average payoff and the lower the average welfare. Finally, to reverse the growth of the Always Defectors, the larger the R^{WP} needs to be. If the latter is not possible, the society will end in a low-welfare mode. The same conclusion emerges from Table 4.

The explanation for the foregoing results is that a larger R^{WP} results in a faster reduction in Willing Punishers and, correspondingly, a larger increase in the Always Cooperators, as can be observed in Figure 3. Consequently, given that it takes longer before the Willing Punishers can recover, a longer phase exists wherein the Always Defectors are present in larger numbers.

We summarize these findings in the following results:

RESULT 1: For any R^{AC} and a fixed $\bar{X}^{AD}$, there exists a threshold value of R^{WP} below which the society moves towards a low-welfare state and above which the high-welfare society is sustained.

RESULT 2: The larger R^{AC} , the larger needs R^{WP} to be to secure the high-welfare society.

RESULT 3: When R^{WP} is set exactly to reverse the increase in the proportion of Always Defectors, then the larger will be ingroup favouritism (measured by R^{AC}):

- The lower the total welfare in society
- The lower the average welfare for the Always Cooperators
- The smaller, on average, the proportion of Always Cooperators in society

The response time to identify and avoid the emerging threat to the high-welfare society is critical. In our model, that time is captured by the size of $\bar{X}^{AD}$, i.e., the level of Always Defectors in the society that shifts tax-reward privileges from the Always Cooperators to the Willing Punishers. Calculations of the effect are presented in Tables 5 and 6.

[Table 5 about here]

The simulations reported in Table 5 tell the story of how important the delay factor is for sustaining a high-welfare society. When the amounts of redistributions are non-variable, an overly slow response inevitably will result in a case wherein the increase in Always Defectors cannot be reversed once it sets off. In the example presented in Table 5, the parameter values are based on reference scenarios. Therefore, if the imminent crisis is not responded to by changing the prevailing narrative in society before more than 12.5% of the population consists of Always Defectors, the high-welfare society collapses. However, a non-trivial relationship exists between R^{AC} and R^{WP} and how large $\bar{X}^{AD}$ needs to be before the high-welfare state breaks down.

The implication from the results above is presented in Table 6. The table shows that the greater is the failure to recognize the imminent crisis as measured by $\bar{X}^{AD}$, the larger the response needs to be to avoid catastrophe. If that scenario is not possible, the collapse of the high-welfare society cannot be stopped.

[Table 6 about here]

Compared to Table 5, Table 7 shows the effect of high levels of $\bar{X}^{AD}$ counteracted by sufficiently high R^{WP} to avert the potential collapse. Several conclusions can be drawn from Table 7. First, as $\bar{X}^{AD}$ increases above 0.1, the necessary R^{WP} rises increasingly. Second, although the collapse is avoided, a few negative consequences remain for the Always Cooperators. The Always Cooperators generally will receive a smaller average payoff, and their proportion in society falls the larger is $\bar{X}^{AD}$. Finally, society inevitably experiences lower average welfare.

[Table 7 about here]

Explaining why a welfare loss occurs when $\bar{X}^{AD}$ is increased relates to the point when the growth of the Always Defectors is reversed. The larger is $\bar{X}^{AD}$, the larger will be the population proportion of Always Defectors (at least temporarily). That result has a detrimental effect on the average payoffs of the Always Cooperators and society as a whole. This finding leads to the final result:

RESULT 4: The size of the delay factor is very important for the sustainability of a high-welfare society. The larger the delay factor, $\bar{X}^{AD}$, the less likely the transition to a low-welfare society can be averted. Although it is possible to reverse the dynamic, a welfare loss will still be sustained.

4. Conclusion

The main research question herein was to ask how robust the welfare state is in the presence of open borders. The main answer given by our model is that welfare states in general will transform into low-welfare societies if incentives are not met by proper policies. The transition to a low-welfare society can be reversed only if three main conditions are met.

First, there must be sufficiently fast recognition and acceptance of the problem once the proportion of Always Defectors increases. Second, the Always Cooperators must be willing to abandon their privileges. Third, society must be willing to generate sufficiently large income to support the Willing Punishers in a crisis. If any one of these three conditions is not met, society will inevitably be transformed into a low-welfare state. Therefore, our research question concerning the robustness of a generous welfare state and whether it can survive in a globalized open-border society also is explained by the same three conditions. A welfare state can exist in such an environment only if it has the capacity and willingness to accept strong and timely measures against defectors.

That conclusion hinges on several assumptions, the most important of which is that the driving force for describing norm formation in a society can be captured by the evolutionary process and the three norm types identified by Ostrom (2000), namely Always Cooperators, Always Defectors and Willing Punishers. The second important assumption is related to the manner in which political power is distributed in society. We assume that changes in prevailing norms affect policy outcomes, either through changes in policies of ruling governments, wherein the justification costs of not changing policies increase sharply.

Our model generates the well-known phenomenon of political cycles in the case where the high-welfare situation is sustained and tells the following story: in good times with no or minimal defection, the narrative of the Always Cooperators dominates, and average social welfare is at its highest, while the proportion of Willing Punishers is reduced over time by ingroup discrimination against outgroups. Such a state cannot persist, as it is threatened eventually by a rising proportion of Always Defectors, and welfare will decline. The political power then shifts to the Willing Punishers who

increase in numbers and implement tough policies against defectors. Owing to more widespread and costly punishment and the relatively large population proportion of defectors, average society welfare falls. However, over time, as the proportion of defectors is reduced, welfare increases again. Once the immediate threat to the high-welfare society is over, political power again shifts to the supporting and inclusive cooperative narrative, and another cycle begins.

In our model, the only mechanism for controlling defection is the willing punishers' costly retribution. Other mechanisms to limit the number of defectors do exist; such mechanisms can be employed by the always cooperators, e.g., in the form of social engineering and social education that, it can be hoped, transform defectors into non-defectors. If such mechanisms are brought into use, the robustness of the high-welfare society may be increased in the future.

References

- Andersen, T. M. (2004). Challenges to the Scandinavian welfare model. *European Journal of Political Economy*, 20(3), 743–754.
- Balafoutas, L., Nikiforakis, N., & Rockenbach, B. (2014). Direct and indirect punishment among strangers in the field. *Proceedings of the National Academy of Sciences of the United States of America*, 111(45), 15924–15927. doi: 10.1073/pnas.1413170111. Pubmed:25349390.
- Barrett, A., & McCarthy, Y. (2008). Immigrants and welfare programmes: exploring the interactions between immigrant characteristics, immigrant welfare dependence, and welfare policy. *Oxford Review of Economic Policy*, 24(3), 542-559.
- Bayer, P., Ross, S. L., & Topa, G. (2008). Place of work and place of residence: informal hiring networks and labour market outcomes. *Journal of Political Economy*, 116(6), 1150–1196.
- Bergh, A., & Bjørnskov, C. (2016). Burying the bumblebee once and for all: does big government hurt growth less in high-trust countries? IFN working Paper 1119 (published March 21, 2016).
- Bertilsson, M., & Hjorth-Andersen, C. (2009). The Nordic welfare state. In G. T. Svendsen, G. L. H. Svendsen (Eds.), *Handbook of Social capital: the troika of sociology, political science and economics*. Cheltenham, UK: Edward Elgar Publishing, pp 212–227.
- Bevir, M., & Rhodes, R. A. W. (2010), *The state as cultural practice*. Oxford University Press.
- Borjas, G. J. (1999), *Heaven's door*. Princeton, NJ: Princeton University Press.

- Brandt, U. S. (2014). The implication of extreme events on policy responses. *Journal of Risk Research*, 17(2), 221–240.
- Brandt, U. S., & Svendsen, G. T. (2010). The survival of the Nordic welfare state and social trust. *Nordic Journal of Political Economy*, 36, 1–15.
- Bratsberg, B., Raaum, O., & Røed, K. (2014). Immigrants, labour market performance and social insurance. *Economic Journal*, 124(580), F644–FF683.
- Chen, R., & Chen, Y. (2011). The potential of social identity for equilibrium selection. *American Economic Review*, 101(6), 2562–2589.
- Chen, Y., & Li, S. X. (2009). Group identity and social preferences. *American Economic Review*, 99(1), 431–457.
- de Quervain, D. J.-F., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., & Fehr, E. (2004). The neural basis of altruistic punishment. *Science*, 305(5688), 1254–1258.
- Dickinson, D. L., Masclet, D., & Peterle, E. (2018). Discrimination as favouritism: the private benefits and social costs of in-group favouritism in an experimental labour market. IZA Discussion Papers, No. 10599. *European Economic Review*, 104, 220–236.
- Dong, Y., Li, C., Tao, Y., & Zhang, B. (2015). Evolution of conformity in social dilemmas. *PLOS ONE*, 10(9), e0137435.
- Dreber, A., Rand, D. G., Fudenberg, D., & Nowak, M. A. (2008). Winners don't punish. *Nature*, 452(7185), 348–351.
- Esping-Andersen, G. (1990), *The three worlds of welfare capitalism*. Princeton, NJ: Princeton University Press.

- Facchini, F. (2016). Political ideological shift: A theoretical approach. *Social Science Information*, 55(4), 589–602.
- Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behaviour*, 25(2), 63–87.
- Fehr, E., & Gächter, S. (2000). Cooperation and punishment in Public Goods Experiments. *American Economic Review*, 90(4), 980-994.
- Fong, C. M., Bowles, S., & Gintis, H. (2006). Reciprocity and the welfare state. In J. M. Ythier, S. C. Kolm (Eds.), *Handbook on the economics of giving, reciprocity and altruism*, Vol. 2. Amsterdam: North-Holland/Elsevier, pp 1439–1464.
- Fu, F., Tarnita, C. E., Christakis, N. A., Wang, L., Rand, D. G., & Nowak, M. A. (2012). Evolution of in-group favouritism. *Scientific Reports*, 2, 460.
- Gächter, S., Herrmann, B., & Thöni, C. (2010). Culture and cooperation. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 365(1553), 2651-2661.
- Hajer, M., & Wagenaar, H., editors (2003), *Deliberative policy analysis: understanding governance in the network society*. Cambridge University Press.
- Hillman, A. L. (2010). "Expressive behaviour in economics and politics," *European Journal of Political Economy*. *Elsevier*, 26(4), 403-418.
- Imhof, L. A., Fudenberg, D., & Nowak, M. A. (2005). Evolutionary cycles of cooperation and defection. *Proceedings of the National Academy of Sciences of the United States of America*, 102(31), 10797–10800.
- Jensen, C., & Svendsen, G. T. (2011). Giving money to strangers: European welfare states and social trust. *International Journal of Social Welfare*, 20(1), 3–9.

- Kaplan, J. T., Gimbel, S. I., & Harris, S. (2016). Neural correlates of maintaining one's political beliefs in the face of counterevidence. *Scientific Reports*, 6, 39589.
- Lindbeck, A. (1995). Welfare state disincentives with endogenous habits and norms. *Scandinavian Journal of Economics*, 97(4), 477–494.
- Lindbeck, A. (2008). Prospects for the welfare state. IFN Working Paper, No.731. Research Institute for Industrial Economics. Stockholm.
- Lindbeck, A., Nyberg, S., & Weibull, J. W. (1999). Social norms and economic incentives in the welfare state. *Quarterly Journal of Economics*, 114(1), 1–35.
- Lindbeck, A., Nyberg, S., & Weibull, J. W. (2003). Social norms and welfare state dynamics. *Journal of the European Economic Association*, 1(2–3), 533–542.
- Meadows, D. H., Meadows, D. L., Randers, J., & Behrens, W. W. (1972), *The limits to growth*, Universe Books.
- Morgan, T. J. H., Rendell, L. E., Ehn, M., Hoppitt, W., & Laland, K. N. (2012). The evolutionary basis of human social learning. *Proceedings. Biological Sciences*, Proceedings of the Royal Society of Biology, 279(1729), 653–662.
- Nannestad, P. (2004). Immigration as a challenge to the Danish welfare state? *European Journal of Political Economy*, 20(3), 755–767.
- Nannestad, P. (2007). Immigration and welfare states: A survey of 15 years of research. *European Journal of Political Economy*, 23(2), 512–532.
- Nannestad, P., & Green-Pedersen, C. (2008). Keeping the bumblebee flying: economic policy in the welfare state of Denmark. In E. Albæk, L. C. Eliason, A. S. Nørgaard, H. M. Schwartz (Eds.). *Crisis, miracles, and Beyond: negotiated adaptation of the Danish welfare state*. Aarhus: Aarhus Universitetsforlag, pp 33–74.

- Nowak, M. A. (2012). Evolving cooperation. *Journal of Theoretical Biology*, 299, 1–8.
- Oliver, P. (1980). Rewards and punishments as selective incentives for collective action: theoretical investigations. *American Journal of Sociology*, 85(6), 1356–1375.
- Ostrom, E. (2000). Collective action and the evolution of social norms. *Journal of Economic Perspectives*, 14(3), 137–158.
- P Hauser, O., A Nowak, M., & G Rand, D. (2014). Punishment does not promote cooperation under exploration dynamics when anti-social punishment is possible. *Journal of Theoretical Biology*, 360, 163–171.
- Paldam, M. (2004). Introduction to the 2003 European Public Choice Society plenary lectures: The Nordic welfare state—success under stress. *European Journal of Political Economy*, 20(3), 739–742.
- Poulsen, A. U., & Svendsen, G. T. (2005). Social capital and endogenous preferences. *Public Choice*, 123(1–2), 171–196.
- Rowthorn, R. (2008). The fiscal impact of immigration on the advanced economies. *Oxford Review of Economic Policy*, 24(3), 560–580.
- Sarvimäki, M. (2011). Assimilation to a welfare state: labour market performance and use of social benefits by immigrants to Finland. *Scandinavian Journal of Economics*, 113(3), no.
- Scheepers, D., Spears, R., Doosje, B., & Manstead, A. S. R. (2006). The social functions of ingroup bias: creating, confirming, or changing social reality. *European Review of Social Psychology*, 17(1), 359 – 396.
- Schultz-Nielsen, M. L. (2017). "Labour market integration of refugees in Denmark." Nordic. *Economic Policy Review* -, 520, 45–72.

- Svendsen, G. L. H., & Svendsen, G. T. (2016), *Trust, social capital and the Scandinavian welfare state: explaining the flight of the bumblebee*. Cheltenham, UK: Edward Elgar Publishing.
- Traulsen, A., Hauert, C., De Silva, H., Nowak, M. A., & Sigmund, K. (2009). Exploration dynamics in evolutionary games. *Proceedings of the National Academy of Sciences of the United States of America*, 106(3), 709–712.
- Van Kersbergen, K., & Manow, P. (2014). *The welfare state*. In. D. Caramani (Ed.). *Comparative Politics*, 3rd ed., pp 520–545. Oxford: Oxford University Press.

Table 1: *Underlying payoff structure in free market exchange*

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	(c, c)	(a, d)
	Defect	(d, a)	(b, b)

Note: Payoff: $d > c > b > a$

Table 2: *Payoff when Row Type Meets Column Type*

	AC	AD	WP
AC	c	a	c
AD	d	b	p
WP	c	p	c

Note: Payoff: $d > c > b > a > p$

AC: Always Cooperates, AD: Always Defects, WP: Willing Punishers

Table 3: *Consequence of varying R^{AC} when R^{WP} is adjusted such that society remains a high-welfare society*

R^{AC} (R^{WP} adjusted)	Share AC	Average profit AC	Average welfare in society
0.25 (0.75)	0.62	3.58	3.57
0.5 (1.0)	0.62	3.46	3.44
0.75 (1.3)	0.60	3.33	3.31
1.00 (1.5)	0.58	3.22	3.20
1.25 (1.5)	0.58	3.13	3.12
1.5 (1.5)	0.59	3.11	3.09

Note: Number in brackets are values of R^{WP} to exactly invert the potential collapse.

Table 4: *Consequence of varying R^{AC} when R^{WP} is set at a (very) high level*

R^{AC} (R^{WP} fixed at 1.5)	Share AC	Average profit AC	Average welfare in society
0.25	0.64	3.57	3.56
0.50	0.67	3.63	3.62
0.75	0.64	3.47	3.46
1.00	0.58	3.22	3.20
1.25	0.58	3.13	3.12
1.50	0.59	3.11	3.09

Table 5: *Changes in $\bar{X}^{AD}$ with $R^{AC} = 0.5$ and $R^{WP} = 1$ fixed*

$\bar{X}^{AD}$	Share AC	Average profit AC	Average welfare in society
0.05	0.68	3.73	3.71
0.075	0.65	3.61	3.60
0.10	0.62	3.46	3.44
0.125	0.09	0.53	1.29

Table 6: Relationship between R^{WP} and Size of $\bar{X}^{AD}$ before high-welfare state collapse

$R^{WP} / (R^{AC} = 0.5)$	Size of $\bar{X}^{AD}$ resulting in high-welfare state collapse
1.0	0.12
1.1	0.13
1.2	0.14
1.3	0.15
1.4	0.16
1.5	0.18

Table 7: Varying $\bar{X}^{AD}$, but adjusting R^{WP} such that the growth of AD-types is inverted

$\bar{X}^{AD}$	Share AC	Average profit AC	Average welfare in society
0.125 (1.4)	0.59	3.38	3.35
0.15 (1.5)	0.60	3.36	3.35
0.20 (1.8)	0.58	3.27	3.26
0.30 (2.4)	0.56	3.10	3.08

Note: Number in brackets are values of R^{WP} to exactly invert the potential collapse.

Figure 1: *Simulation without any tax redistribution*

Figure 2: *The two possible modes that the society will move to*

Mode 1: Domination of AD-types: ~~Low~~low-welfare society (Reference scenario)

Parameter	Average share	Average payoff
$\beta = 5$	$X^{AC} = 0.09$	$\pi^{AC} = 0.51$
$R^{AC} = 0.5$	$X^{WP} = 0.04$	$\pi^{WP} = -0.36$
$R^{WP} = 0.5$	$X^{AD} = 0.87$	$\pi^{AD} = 1.31$
$I_t = 5$	Average welfare in society	
$\bar{X}^{AD} = 0.1$	$\pi^S = 1.28$	

Mode 2: Domination of non-AD-types: high-welfare society

Parameter	Average share	Average payoff
$\beta = 5$	$X^{AC} = 0.62$	$\pi^{AC} = 3.46$
$R^{AC} = 0.5$	$X^{WP} = 0.28$	$\pi^{WP} = 3.43$
$R^{WP} = 1$	$X^{AD} = 0.10$	$\pi^{AD} = 3.23$
$I_t = 5$	Average welfare in society	
$\bar{X}^{AD} = 0.1$	$\pi^S = 3.44$	

Figure 3: *Comparing situations with low and high R^{AC} (cf. simulations from Table 1)*

Situation where $R^{AC} = 0.5$ and $R^{WP} = 1$

Situation where $R^{AC} = 1$ and $R^{WP} = 1.5$