Institutions, Behavior and Evolution.

Dissertation zur Erlangung des akademischen Grades

Doktor der Wirtschafts- und Sozialwissenschaften (Dr. rer. pol.)

Vorgelegt am Fachbereich Wirtschaftswissenschaften

der Universität Kassel

von

Stephan Müller

aus

Meißen

Zierenberg

2014

Erstgutachter

Professor Dr. Georg von Wangenheim

Zweitgutachter

Professor Dr. Björn Frank

Drittgutachter

Professor Dr. Guido Bünstorf

Tag der Disputation: 11.06.2015

Erklärung

Hiermit versichere ich, dass ich die vorliegende Dissertation selbstständig, ohne unerlaubte Hilfe Dritter angefertigt und andere als die in der Dissertation angegebenen Hilfsmittel nicht benutzt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten oder unveröffentlichten Schriften entnommen sind, habe ich als solche kenntlich gemacht. Dritte waren an der inhaltlich-materiellen Erstellung der Dissertation nicht beteiligt; insbesondere habe ich hierfür nicht die Hilfe eines Promotionsberaters in Anspruch genommen. Kein Teil dieser Arbeit ist in einem anderen Promotions- oder Habilitationsverfahren verwendet worden.

Zierenberg, 2. September 2014 ____

1. Einleitung

1.1. Motivation

Präferenzen stellen ein zentrales Element der Ökonomik im konzeptionellen Rahmen der Analyse menschlichen Verhaltens dar. Präferenzen beziehen sich auf eine Menge von Annahmen die Rangfolge von Wahlmöglichkeiten betreffend. Formal mathematisch wurde das Konzept erstmalig von Frisch (1926) formuliert und später von Arrow (1951) perfektioniert. Präferenzrelationen mit denen Präferenzen modelliert werden, bilden die Grundlage der Rational-Choice Theorie. Rational Choice ist definiert durch die Bestimmung der Wahlmöglichkeiten und die anschließende Wahl der besten Alternative anhand bestimmter Konsistenzbedingungen. Rational-Choice Theorie basiert auf einer stark vereinfachten Beschreibung des Wahlproblems (Ziele und Nebenbedingungen). Das Wahlentscheidungen von vielen weiteren Faktoren abhängen wird beispielsweise von psychologischen Theorien propagiert und von zahlreichen Laborexperimenten unterstützt (Hogarth und Einhorn 1992; Hoffman et al. 1994; Kahneman und Frederick 2002, um nur einige zu nennen). Derartige Faktoren beinhalten zum Beispiel die Art und die Reihenfolge wie Informationen zur Verfügung gestellt werden. Die empirischen Befunde in der Ökonomik und in psychologischen Experimenten welche den Vorhersagen der Rational-Choice Theorie widersprechen, haben intensive Forschungen auf dem Gebiet der Entscheidungstheorie nach sich gezogen und eine Vielzahl alternativer Verhaltensmodelle wurde entwickelt. Das Verstehen individuellen Verhaltens ist eine der grundlegenden Aufgaben ökonomischer Forschung. Meine Dissertation trägt zur Bewältigung dieser Aufgabe insofern bei, als das die Entstehung bestimmter charakterisierender Merkmale von Präferenzen und deren Auswirkung auf das Verhalten ein übergeordnetes Thema der drei konstituierenden Artikel dieser Arbeit bildet. Neben diesem verbindenden Themenkomplex, der für mich von zentralem Interesse in meiner Forschung ist, ist jeder der Artikel durch eine konkrete Forschungsfrage motiviert. Dies wird in den nächsten Abschnitten genauer herausgestellt.

Der methodologische Individualismus und das Konzept des Homo Oeconomicus bilden die Basis der traditionellen ökonomischen Theorie. Präferenzen und der zulässige Handlungsraum bestimmen in diesem Rahmen das individuelle Verhalten. Zur Operationalisierung der grundlegenden Annahmen der Rationalität und der Eigennutzen-Orientierung wird angenommen, dass Präferenzen im Zeitverlauf stabil sind. Zeitinvariante Präferenzen sollen dabei allerdings kein deskriptives Modell für real existierende Individuen darstellen. Selbst wenn Präferenzen kurzfristig fix wären, stellt sich die Frage wie und warum bestimmte Charakteristika wie Altruismus, Risiko- oder Verlustaversion, die sich wiederholt in Experimenten zeigten, in den Menschen „eingepflanzt“ wurden. Die evolutionäre Perspektive bietet einen möglichen Analyserahmen für die Beantwortung dieser Fragen. Innerhalb dieses Rahmens versuchen Ökonomen besondere Aspekte menschlicher Präferenzen zu rationalisieren. Spätestens mit den bedeutenden Arbeiten von Fehr und Schmidt (1999) und Bolton und Ockenfels (2000) stellt das Konzept der Ungleichheitsaversion als eine Form von Präferenzen, welche die Situation andere Individuen in die Bewertung der eigenen mit einbezieht, eine prominente Erklärung für zahlreiche empirische und experimentelle Ergebnisse dar, welche von der Vorhersage der traditionellen ökonomischen Theorie abweichen. Aufgrund der zunehmenden Bedeutung bedarf es einer Rationalisierung derartiger Präferenzen, da sie sonst lediglich eine ad-hoc Anpassung herkömmlicher Präferenzen darstellen, um empirische Befunde besser erklären zu können. Güth und Napel (2006) weisen darauf hin, dass derartige Präferenzen insbesondere mit der physischen

Notwendigkeit nach materielle Bedürfnisbefriedigung in einer Welt knapper Ressourcen vereinbar sein sollten. Mit anderen Worten, derartige Präferenzen sollten aus evolutionärer Sicht rationalisierbar sein. Die Autoren argumentieren überzeugend, dass jede Art von Untersuchung, die sich mit der Evolution von Präferenzen beschäftigt in einem Rahmen ausgeführt werden sollte, der alle Klassen menschlicher Interaktion enthält, da sonst spiel-spezifische Ergebnisse erzielt werden, deren Verallgemeinerbarkeit zumindest fragwürdig ist. Sie bezeichnen eine solche Umwelt als das „game of life". In dem Artikel 'The evolution of inequality aversion in a simplified game of life' unternehme ich einen ersten Schritt diese Notwendigkeit zu erfüllen.

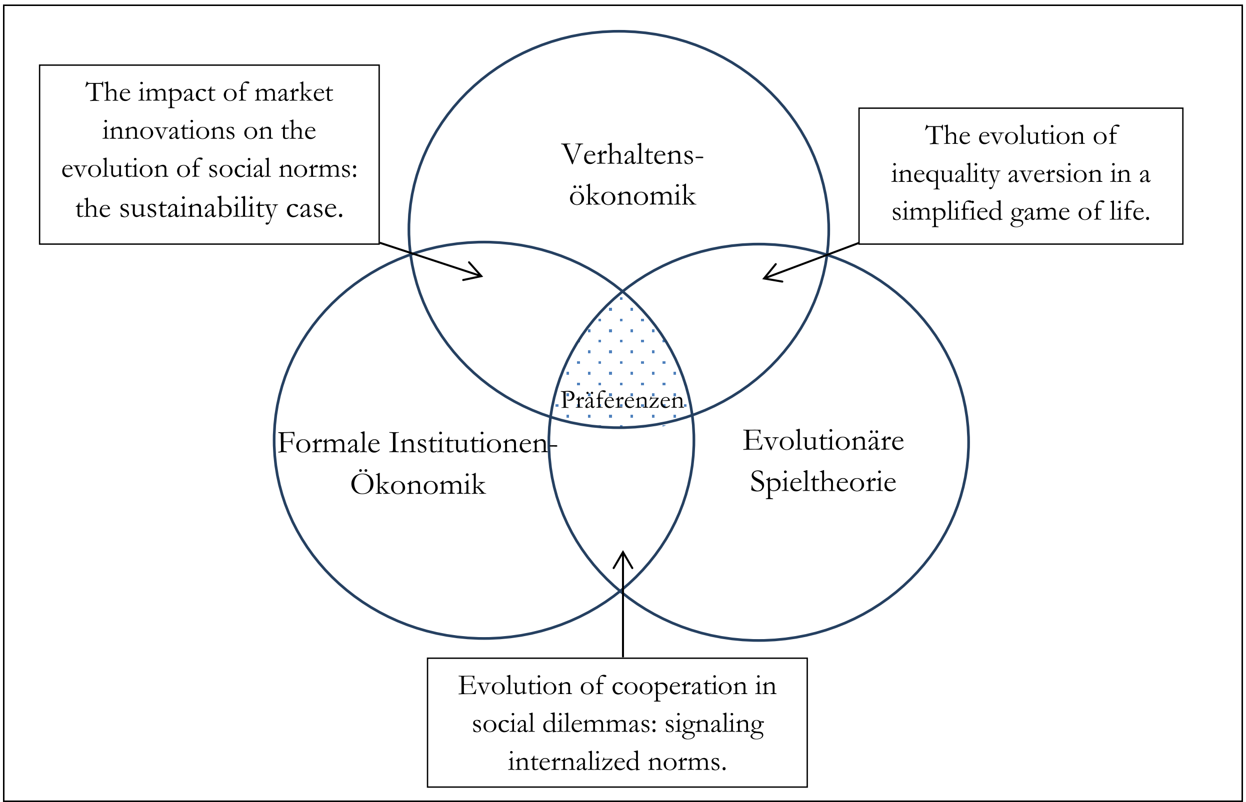
Präferenzen werden jedoch nicht nur durch evolutionäre Kräfte geformt, die langfristig wirken und den genetisch fixierten Teil der Präferenzen verändern. Formelle und informelle Institutionen ermöglichen und beschränken individuelles Handeln. Werden informelle Institutionen internalisiert, so werden sie Teil der Präferenzen eines Individuums. Die Internationalisierung sozialer Normen findet in viel kürzerer Frist statt und hängt von zahlreichen Faktoren ab. Ich verstehe unter sozialen Normen innere Handlungs"empfehlungen" sich in einer bestimmten Art und Weise zu verhalten. Externe Anreize wie die mit einer Handlung verbundenen Kosten und Erträge können sozialen Normen komplementär oder substitutiv gegenüberstehen. Diese Kosten und Erträge sind Teil der zuvor erwähnten externen Beschränkungen des Handelns. In modernen Volkswirtschaften findet der Großteil menschlicher Interaktionen auf Märkten statt. Somit materialisiert sich die Mehrzahl der Verhaltensrestriktionen auf Märkten. Durch die Aggregation individuellen Verhaltens, aggregieren Märkte auch individuelle externe Effekte. Viele der heutigen Umweltprobleme werden durch das Konsumverhalten privater Haushalte verursacht. Individualverkehr, Nutzwärme und Nahrungsmittelproduktion tragen substantiell zur Emission von CO_2 und andere Umweltschadstoffe bei. Die Wahl zwischen mehr oder minder zur Umweltverschmutzung beitragender Produkte hängt von der Verfügbarkeit derartiger Produkte und sozialen Normen und anderer Institutionen ab. Lösungen von Umweltproblemen hängen somit nicht nur von Produktinnovationen sondern auch von den bestehenden sozialen Normen ab, wobei erstere die Märkte um nachhaltige Produkte erweitern und letztere nachhaltigen Konsum fördern. In Anbetracht des Einflusses sozialer Normen auf individuelle Präferenzen ist es augenscheinlich, dass Märkte und soziale Normen nicht getrennt voneinander untersucht werden können. Die existierende Literatur enthält zahlreiche Studien über die Wechselwirkungen zwischen Märkten und soziale Normen in beide Richtungen – wie soziale Normen Märkte beeinflussen und wie Märkte soziale Normen (z.B. Hong und Kacperczyk 2009; Johnson 2004; Ek und Soderholm 2008; Fehr und Gächter 2001 und Gneezy und Rustichini 2000). Sämtliche dieser Ansätze beschränken sich auf die monetären Anreize, die von Märkten gegeben werden und deren Regulierung. Dies allerdings reduziert Märkte auf ihre Preis-Mengen-Dimension und vernachlässigt völlig deren Innovationskapazität. Die Produktvariationen, die aus nachhaltigkeitsrelevanten Innovationen hervorgehen sind ein wichtiges Element in der Wechselwirkung zwischen Märkten und Normen. Diese Wechselwirkung steht im Zentrum des Artikels "The impact of market innovations on the evolution of social norms: the sustainability case".

Der Aspekt der Nachhaltigkeit weist auf die allgemeine Klasse sozialer Dilemmata hin. Soziale Dilemmata zeichnen sich dadurch aus, dass individuell rationales Verhalten zu kollektiv irrationalen Ergebnissen führt. Ökonomen haben in Experimenten (für einen Überblick der

Ergebnisse siehe Roth 1995) und Feldbeobachtungen (Fey und Meier 2004; Cunha und Augenblick 2014) Kooperation in dem Sinne beobachtet, dass die kollektive Irrationalität zumindest teilweise überwunden wird. Aus Sicht der klassischen ökonomischen Perspektive ist dies überraschend insbesondere da kooperatives Verhalten auch dann zu beobachten ist, wenn jetziges Verhalten keinerlei Auswirkungen auf zukünftige Interaktionen hat (Cooper et al. 1996). Die zahlreichen Erklärungsansätze basieren gewöhnlich auf mindestens einer von zwei Einschränkungen. Die erste Einschränkung besteht darin, dass Erklärungsversuche strukturierte Populationen untersuchen, in denen Interaktionen nicht vollständig anonym sind sondern Individuen die Möglichkeit haben Informationen über das Verhalten andere oder deren Identität zu sammeln und zu verarbeiten. Die zweite Einschränkung zeigt sich darin, dass Erklärungsansätze von der nicht motivierten Fähigkeit sozialer Normen ausgehen, individuelle Handlungsräume zu beschränken, insbesondere hinsichtlich des Missbrauchs von Bestrafungsmechanismen. Der Artikel "Evolution of cooperation in social dilemmas: signaling internalized norms." präsentiert eine neue Erklärung kooperativen Verhaltens welche ohne beide Einschränkungen auskommt.

1.2. Einordnung der Dissertation: der weite Blickwinkel

Die drei Artikel, welche meine Dissertation bilden, setzen drei Gebiete der Wirtschaftswissenschaften in Beziehung: Verhaltensökonomik, formale Institutionenökonomik und evolutionäre Spieltheorie. Die folgende Abbildung veranschaulicht dies grafisch.



Verhaltensökonomik: Das Einbeziehen psychologischer, kognitiver und emotionaler Faktoren ist das charakterisierende Merkmal der Verhaltensökonomik. Die Verhaltensökonomik versucht dadurch den Erklärungsgehalt der Wirtschaftswissenschaften zu erhöhen. Der Versuch die Wirtschaftswissenschaften auf eine realistischere psychologische Basis zu stellen impliziert jedoch nicht die Ablehnung neoklassischer Ansätze. Tatsächlich werden in den meisten Artikeln auf diesem Gebiet nur ein oder zwei Annahmen der Standardtheorie angepasst, um einen größeren Grad psychologischen Realismus zu erreichen. Camerer und Loewenstein (2004) weisen darauf hin, dass es nichts im Kern neoklassischer Theorie gibt, was besagt, dass Individuen keine Rücksicht auf Fairness nehmen oder das riskante Ereignisse linear gewichtet werden sollten. Einige der zuvor genannten Modifikationen schwächen diese vereinfachenden Annahmen ab. Andere Abwandlungen des neoklassischen Rahmens berücksichtigen kognitive Schranken des Menschen. Derartige Annahmen beziehen sich auf das was Herbert Simon „prozedurale Rationalität" nennt (Simon 1976). Unter methodologischen Gesichtspunkt griff die Verhaltensökonomik anfänglich vor allem auf experimentelle Ergebnisse zurück. In der jüngeren Forschung finden auch Feldexperimente (Gneezy und Rustichini 2004) und Computersimulationen (Angeletos et al. 2001) Anwendung.

Laut Camerer und Loewenstein (2004) kann die behavioristische Forschung zu menschlichen Entscheidungen, welche die Hauptquelle der Verhaltensökonomik hinsichtlich relevanter psychologischer Aspekte darstellt, in zwei Kategorien klassifiziert werden: Einschätzung und Wahl. In der ersten Kategorie geht es im Kern darum wie Menschen Wahrscheinlichkeitseinschätzungen treffen. Wie Menschen zwischen verschiedenen Alternativen wählen ist Gegenstand der zweiten.

Der Artikel zum Einfluss von Marktinnovationen lässt psychologische Aspekte in den Verbreitungsprozess sozialer Normen einfließen. Nachdem ein Produkt, welches sich durch ein relativ hohes Maß an Normkompatibilität auszeichnet, am Markt angeboten wird, verändert sich der Verbreitungsprozess in zweierlei Hinsicht. Erstens sind die Marktteilnehmer nun in der Lage im Einklang mit der sozialen Norm zu konsumieren, was vor der Innovation nicht möglich war. Zweitens ermöglicht die neue Produktvariation soziale Einflüsse wie den Druck zur Verhaltenskonformität (Boyd und Richerson 1985) ihre Wirkung zu entfalten.

Der Artikel zur Evolution von Ungleichheitsaversion beinhaltet verhaltensökonomische Elemente da diese Präferenzeigenschaft ein realistischeres Bild menschlichen Verhaltens zeichnet. Das Papier liefert eine evolutionäre Grundlage für eine Erklärungsvariable in der behavioristischen Forschung zu menschlichen Entscheidungen.

Evolutionäre Spieltheorie: Bis heute gibt es keine Übereinkunft darüber was genau mit der evolutionären Perspektive auf dem Gebiet der Wirtschaftswissenschaften gemeint ist. Witt (2008) reflektiert für die Verhaltensökonomik über die drei Ebenen wissenschaftlichen Arbeitens: die ontologische, die heuristische und die methodologische Ebene. Im Folgenden werde ich kurz seine Erkenntnisse wiedergeben, da dies hilfreich sein wird, die beiden Artikel mit hohem evolutionsökonomischen Gehalt einzuordnen. Auf der ontologischen Ebene identifiziert Witt (2008) einerseits den monistischen, anderer den dualistische Standpunkt. Im ersten wird unterstellt, dass die ökonomische Sphäre und die Nature in Wechselwirkung stehen. Die dualistische Betrachtungsweise verneint diese Sichtweise und behandelt die ökonomische und die biologische Evolutionsdynamiken als Teil von einander getrennter Sphären der Realität. Auf heuristischer Ebene unterscheidet Witt (2008) die verallgemeinerte darwinistische Heuristik und allgemeine evolutionäre Heuristik. Im ersten Fall werden die drei Prinzipien der Evolution, welche durch abstrakte Reduktion aus Darwins' Theorie zur natürlichen Auslese folgen (Variation, Vererbung und Selektion), zur Konzeptualisierung der Evolution von Technologien (Ziman 2000), der Wissenschaft (Hull 2001), der Wirtschaft (Nelson 1995) und anderen herangezogen. Die allgemeine evolutionäre Heuristik basiert nicht auf einer Analogie zwischen ökonomischen und biologischen Evolutionsdynamiken, sondern auf einem allgemeinen Evolutionskonzept. Dieses Konzept beschreibt Evolution als einen Prozess der Eigentransformation mit der endogenen Erzeugung von Neuheiten und der bedingten Verbreitung als konstituierende Merkmale (Witt 2003, Kap. 1). Die zwei ontologischen und die zwei heuristischen Standpunkte ermöglichen die Einordnung evolutionsökonomischer Ansätze in eine 2x2 Matrix.

Laut Witt (2008) lassen sich die Anwendungen evolutionärer Spieltheorie auf dem Gebiet der Wirtschaftswissenschaften entlang dieser Dimensionen unterscheiden und haben im Wesentlichen zwei Interpretationen. In der ersten Interpretation werden Selektionsprozesse beschreibende Modelle der Evolutionsbiologie in ökonomischen Kontexten angewandt um Lernprozesse abzubilden (siehe Brenner 1999, Kap. 6). Diese Interpretation bedient sich der heuristischen Strategie eine Analogie zwischen biologischer Adaption und ökonomischer Adaption durch nicht-kognitive Lernprozesse. Aus ontologischer Sicht beantwortet die Analogiekonstruktion regelmäßig die Fragen danach ob und wie ökonomische Prozess in Verbindung mit der naturalistischen Fundierung menschlichen Verhaltens stehen (Witt 2008). Die erste Interpretation entspricht somit der Position in der zuvor erwähnten 2x2 Matrix zur Strukturierung der Evolutionsökonomik hinsichtlich des ontologischen Standpunkts und der angewendeten heuristischen Strategie, welche der Neo-Schumpeter'schen Synthese von Nelson und Winter (1982) entspricht. Die zweite Interpretation basiert nicht auf einer Analogie sondern der biologische Kontext ist dabei von direkter Bedeutung für Anwendungen in den Wirtschaftswissenschaften. Autoren die diese Position vertreten gehen davon aus, dass die grundlegenden Charakteristika menschlichen Verhaltens genetische verankert sind und somit am besten vom Standpunkt der natürlichen Auslese verstanden werden können. Zu diesen grundlegenden Eigenschaften menschlichen Verhaltens zählen Altruismus, Fairness und Moral (siehe bspw. Güth und Yaari 1992a; Binmore 1998; Gintis 2007). Diese direkte Übertragung von der Biologie auf die Wirtschaftswissenschaften setzt offenbar eine monistische Ontologie voraus. Witt (2008) argumentiert, dass die heuristische Strategie, welche in der Forschung, die der zweiten Interpretation folgt, Anwendung findet, einige Gemeinsamkeiten mit Hayek's Theorie zur gesellschaftlichen Evolution teilt.

Unter dieser Betrachtung, folgt der Artikel zur Evolution von Ungleichheitsaversion, welcher in Kapitel 4 vorgestellt wird, der zweiten Interpretation, wohingegen die Arbeit, die sich mit der Signalisierung internalisierter Normen beschäftigt, der ersten folgt. Ich untersuche die Evolution von Ungleichheitsaversion in einer Umwelt, die ich als vereinfachtes „game of life" (Güth und Napel 2006) bezeichne. Diese Umwelt vereint drei Klassen von Spielen, die repräsentativ für die Mehrheit menschlicher Interaktionen sind. Der Artikel zur Evolution von Kooperation betrachtet ein konkretes Spiel, das Gefangenendilemma, und untersucht die Signalisierung einer Kooperationsnorm als einen Mechanismus zur Förderung von Kooperation.

Formelle Institutionenökonomik: Der Begriff der „Institutionsökonomik" findet für eine Vielzahl ökonomischer Ansätze und Schulen Anwendung. In der Regel bezieht er sich auf den Bereich der Ökonomik, welcher in der Tradition von Thorstein Veblen, John R. Commons, und Wesley Mitchell steht. In den letzten Jahren hat sich der Begriff der „Neuen Institutionenökonomik" etabliert. Dieser Begriff bezieht sich auf ökonomische Forschung in der Tradition des Transaktionskostenansatzes von Ronald Coase, Oliver Williamson, und Douglas North. In jüngster Vergangenheit wird dieser Begriff oft um spieltheoretische Ansätze zur Evolution gesellschaftlicher Konventionen und manchmal auch um Institutionen im Verständnis der Österreichischen Schule erweitert (siehe Rutherford 2001). In dieser Dissertation folge ich dem spieltheoretischen Ansatz. Um diesen Ansatz von der traditionellen Institutionenökonomik und den meist nicht-formalen Ansätzen der Neuen Institutionenökonomik zu unterscheiden, bezeichne ich ihn als formelle Institutionenökonomik. In der den spieltheoretischen Rahmen anwendenden Literatur können zwei Ansätze zur Definition des Begriffs „Institution" unterschieden werden: der Gleichgewichtsansatz und der Spielregelansatz. Im Gleichgewichtsansatz stellt der Gleichgewichtscharakter individuellen Verhaltens das zentrale definierende Element von Institutionen dar. Meist übersetzt sich der Gleichgewichtscharakter in ein Stabilitätskonzept. Einige Autoren ziehen Konzepte der evolutorischen Spieltheorie (Sugden 1986; Sugden 1989; Young 1998; Aoki 2000; Bowles 2000), andere die Theorie wiederholter Spiele der klassischen Spieltheorie heran (Greif 1989; Greif 1997; Greif 1998, Milgrom et al. 1990; Calvert 1995). Der Spielregelansatz versteht Institutionen als externe Faktoren, welche den Strategienraum und die Auszahlungen des Spiels formen (North 1990; Hurwicz 1993; Hurwicz 1996). Es besteht eine große Lücke zwischen der weiten Definition in der Neuen Institutionenökonomik und der engen Definition innerhalb des Gleichgewichts- und des Spielregelansatzes.

Diese Lücke wird durch den indirekten evolutorischen Ansatz von Güth und Yaari (1992) verkleinert. Innerhalb dieses Ansatzes wird zwischen Verhaltenspayoffs und Fitness- oder materiellen Payoffs unterschieden. Die Fitnesspayoffs einer bestimmten Verhaltensweise sind entscheidend für die Verbreitung dieser innerhalb der Population. Verhaltenspayoffs spiegeln die innere Bewertung dieser Fitnesspayoffs wider. Die Verhaltenspayoffs sind relevant für die Entscheidungsfindung, haben aber keinen Einfluss auf die Adoptionsrate für diese Strategie durch andere Individuen.

Der indirekte evolutorische Ansatz ermöglicht damit eine komplexere Modellierung von Institutionen, die über eine einfache Verhaltensregelmäßigkeit oder die Regeln eines Spiels hinausgeht. Insbesondere die informelle Institution einer sozialen Norm kann durch diesen Ansatz abgebildet werden. Somit liefert dieser Ansatz einen ersten Schritt um die zuvor erwähnte Lücke hinsichtlich der Komplexität unterschiedlicher Definitionen von „Institutionen" zu verringern. Dies ist der Grund dafür, warum dieser Ansatz in den beiden Artikeln, die eine evolutorische Perspektive einnehmen, Anwendung findet (Kap. 3 und 4).

1.3. Einordnung der Dissertation: der enge Blickwinkel

In diesem Abschnitt werde ich jeden der Artikel in die bestehende Literatur einordnen und die Forschungsfragen herausarbeiten. Wie bereits erwähnt bilden „Präferenzen" das übergreifende Thema meiner Dissertation.

Ich betrachte soziale Normen als wichtigen individuelle Präferenzen formenden Faktor. Der Artikel, der in Kap. 2 vorgestellt wird, untersucht die Wechselwirkung zwischen Märkten und dem Verbreitungsprozess von sozialen Normen. Dabei fließt die größere psychologische Realitätsnähe der Verhaltensökonomik in den Analyserahmen der formalen Institutionenökonomik ein. Darüber hinaus liefert der Artikel Erkenntnisse für die industrieökonomische Forschung hinsichtlich der Wirkung sozialer Normen auf Produktmärkte. Der Einfluss von sozialen Normen auf Märkte wurde aus theoretischer, empirischer und experimenteller Sicht untersucht. Mehrere Versuche wurden unternommen, um norm-motiviertes Verhalten in die neoklassische Konsumtheorie einzubauen (siehe z.B. Nyborgs et al. 2006; Brekke et al. 2003). Trotz dieser Bemühungen gibt es keine auf norm-motivierten Verhalten basierende allgemeine oder partielle Gleichgewichtstheorie, was ursächlich dafür sein könnte, dass die Mehrheit der Forschung auf diesem Gebiet empirischer Natur ist. Wie soziale Normen einen bestimmten Typ von Märkten, die Finanzmärkte, beeinflussen, wird durch Hong und Kacperczyk (2009) und Johnson (2004) untersucht. Kim (2007) zeigt, dass Normen auch für die Märkte privater Eigentumsrechte relevant sind. Eine Reihe von Fehr et al. (1998) durchgeführten Experimenten zu Wettbewerbsmärkten und bilateralen Verhandlungen weisen darauf hin, dass Wettbewerb nur einen begrenzten Einfluss auf Marktergebnisse hat, wenn die Norm der Reziprozität wirksam ist. Die Rolle für die Nachfrage nach „grüner“ Elektrizität des psychologischen Bedürfnisses eine positive Selbstwahrnehmung als sozial verantwortungsbewusste Person zu erhalten, wurde von Ek und Soderholm (2008) untersucht.

Die Forschung zum Einfluss von Märkten auf die Evolution von Normen beschäftigt sich hauptsächlich mit der Analyse der Beziehungen zwischen normgeleiteter intrinsischer Motivation und markt- oder preisgeleiteter extrinsischer Motivation. Es gibt empirische Befunde (Fehr und Gächter 2001), welche belegen, dass Anreizverträge reziprozitätsgeleitete freiwillige Kooperation verdrängen. Einen Überblick über diesen Teil der Literatur, der sich mit Verdrängungseffekten beschäftigt, geben Frey und Jegen (2001). Darüber hinaus gibt es auch theoretische Forschung. Benabou und Tirole (2006) entwickeln eine Theorie prosozialen Verhaltens in der Belohnungen und Strafen Zweifel über die wahren Motive guter Taten wecken. Dies kann zur teilweisen oder völligen Verdrängung prosozialen Verhaltens führen. Die Wechselwirkung zwischen sozialen Normen und ökonomischen Anreizen in Unternehmen wird durch Huck et al. (2012) modelliert. Die Arbeit von Bohnet et al. (2001) enthält sowohl ein theoretisches Modell als auch Beobachtungen aus Laborexperimenten. Sie untersuchen die Verbindung zwischen der Durchsetzbarkeit von Verträgen und individueller Leistungserbringung. Ihre Ergebnisse zeigen, dass Vertrauenswürdigkeit durch schwache Durchsetzung gefördert und bei mittlerer Durchsetzung verdrängt wird. Diesen Ansätzen ist gemein, dass sie sich auf marktinduzierte monetäre Anreize beschränken. Dies jedoch reduziert Märkte auf ihren Preis-Mengen Aspekt und lässt deren Innovationsvermögen völlig außer Acht. Produktvariationen die durch derartige Innovation hervorgebracht werden sind ein wichtiger Baustein der Markt-Norm-Beziehung. Diese Forschungslücke hinsichtlich der Wechselwirkung zwischen Produktinnovationen und der Verbreitung sozialer Normen wird in Kapitel 2 behandelt.

Das dritte Kapitel stellt ein Papier vor, welches die Sichtweise der formellen Institutionenökonomik auf den Analyserahmen der evolutorischen Spieltheorie überträgt. Es hat eine konkrete soziale Norm und deren Potential ein soziales Dilemma zu lösen zum Inhalt, die Norm sich kooperativ zu verhalten. Genauer gesagt, beschäftigt sich die Arbeit mit dem Rätsel über die Entstehung von Kooperation in großen, unstrukturierten Populationen in einem Umfeld in dem nicht-kooperatives Verhalten individuell rational ist. Die meisten Erklärungsansätze weisen eine oder beide von zwei Einschränkungen auf. Entweder werden strukturierte Populationen untersucht oder soziale Normen haben die nicht begründete Fähigkeit den individuellen Handlungs- oder Strategienraum zu beschränken.

Bezüglich der ersten Art der Beschränkung verdienen einige Zweige der Literatur eine besondere Aufmerksamkeit. Die Theorie der Verwandtenselektion stellt die Kooperation unter Individuen, die genetisch in enger Beziehung stehen, ins Zentrum der Betrachtung (Hamilton 1964a, 1964b), wohingegen Theorien zur direkten Reziprozität auf Kooperationsanreize egoistischer Individuen in wiederholten Interaktionen fokussieren (Trivers 1971; Axelrod 1984). Im Falle unendlicher Wiederholung innerhalb einer Gruppe sei auf Taylor (1976) oder Mordecaï (1977) und die Folk-Theoreme von Rubinstein (1979) oder Fudenberg und Maskin (1986) verwiesen. Im Falle unbestimmter Wiederholung, siehe Kreps et al. (1982). Theorien indirekter Reziprozität und kostenverursachende Signalisierung zeigen Kooperation in größeren Gruppen entstehen kann, falls es den kooperierenden Gruppenmitgliedern gelingt eine Reputation aufzubauen (Nowak and Sigmund 1998; Wedekind and Milinski 2000; Gintis et al. 2001).

Bezüglich der zweiten Einschränkung sei auf die frühen Arbeiten von Hirshleifer und Rasmusen (1989) und Witt (1986) verwiesen, welche Bestrafungen nur dann erlauben, wenn eine Norm verletzt wurde. Sethi (1996) erlaubt alle möglichen Strategien, die Bestrafung entweder von der Missachtung oder der Befolgung einer Norm abhängig machen. Allerdings erfolgt darüber hinaus eine exogene Spaltung und damit Strukturierung der Population in Individuen deren Verhalten im klassischen Sinne rational ist und in solche, deren Verhalten durch Routinen bestimmt ist, die sich langsam an die Umweltzustände anpassen.

Das Papier zur Evolution von Kooperation in sozialen Dilemmas eröffnet eine alternative Erklärung für die Entstehung von Kooperation, die nicht von diesen beiden Restriktionen abhängt.

Während sich Kapitel 3 mit sozialen Dilemmas als wichtige Klasse menschlicher Interaktion beschäftigt, untersucht der in Kapitel 4 vorgestellte Artikel die Evolution von Ungleichheitsaversion in einer vereinfachten gemischten Umwelt, die drei Klassen menschlicher Interaktion vereint: ein soziales Dilemma, ein Koordinationsproblem und ein Verteilungsproblem. Das Konzept der Ungleichheitsaversion spielt eine wichtige Rolle in der behavioristischen Forschung zu menschlichen Entscheidungen. In Kapitel 4 wird der Analyserahmen der evolutorischen Spieltheorie auf diese spezifische Verhaltensdeterminante angewandt. In der Vergangenheit wurde die Evolution von Präferenzen in stark vereinfachten Umwelten, die durch ein konkretes Spiel beschrieben werden, untersucht (z.B. Huck und Oechssler 1999; Koçkesen et al. 2000a, 2000b und Sethi und Somanathan 2001). In jüngerer Vergangenheit wurden Versuche unternommen die Evolution von Präferenzen in komplexeren Umwelten zu untersuchen. Güth und Napel (2006) analysieren wie das Persönlichkeitsmerkmal der Ungleichheitsaversion in einer Umgebung evolviert, welche zwei oft untersuchte Spiele vereint: das Ultimatum-Spiel und das Diktator-Spiel. Poulson und Poulson (2006) untersuchen die Evolution sozialer Präferenzen in einer Umwelt, die sich aus einem simultanen und einem sequentiellen Gefangenendilemma zusammensetzt. Die Ergebnisse dieser Arbeiten wiesen darauf hin, dass die Erkenntnisse die auf Untersuchungen von Umwelten basieren, die durch ein einziges Spiel repräsentiert werden, mit Vorbehalt zu bewerten sind, da die Ergebnisse eine signifikante Veränderung erfahren können, sobald komplexere Umwelten betrachten werden. Dieser Sachverhalt zeigt eine Forschungslücke auf, zu deren Schließung Kapitel 3 beiträgt, indem es einen allgemeinen Rahmen zur Analyse der Evolution von Präferenzen vorschlägt und diesen auf die soziale Präferenz der Ungleichheitsaversion anwendet.

Eine Voraussetzung für die Untersuchung der Evolution von Präferenzen im Rahmen des „game of life" ist die Strukturierung der unendlichen Menge an potentiellen Spielen, worin das zweite Ziel der Arbeit besteht. Es gibt Grund zur Annahme, dass menschliches Verhalten nicht spielspezifisch ist, sondern Gemeinsamkeiten für ganze, sehr allgemeine Klassen von Spielen zeigt (siehe Yamagishi et al. 2013; Ashraf et al. 2006; Blanco et al. 2011; Chaudhuri und Gangadharan 2007 und Slonim und Garbarino 2008). Dies macht Hoffnung, dass sich für evolutionäre Studien zu Präferenzen die überwältigende Komplexität der realen Welt auf diese Klassen reduzieren lässt. Zahlreiche Autoren teilen die Ansicht, dass es zwei fundamental verschiedene Arten gesellschaftlicher Probleme gibt (siehe bspw. Sugden 1986; Milgrom et al. 1990), Koordinationsprobleme und soziale Dilemmas. Schotter (1981), Ullmann-Margalit (1977) und andere sind der Auffassung, dass es neben diesen beiden Klassen (mindestens) eine dritte Art von sozialen Problemen gibt, Verteilungsprobleme. Ein Verteilungsproblem zeichnet sich durch eine asymmetrische Verteilung der gleichgewichtigen Auszahlungen aus. Das „game of life" wie ich es vorschlagen werde, umfasst diese drei Klassen von Interaktionen.

1.4. Wissenschaftlicher Beitrag und zentrale Ergebnisse

In diesem Abschnitt werden der wissenschaftliche Beitrag und die zentralen Erkenntnisse jedes einzelnen Artikels hervorgehoben.

Kapitel 2 "The impact of market innovations on the evolution of norms: the sustainability case." beschäftigt sich mit der in 1.3 identifizierten Forschungslücke: die Wechselwirkung zwischen innovativen Produktinnovationen und der Evolution sozialer Normen. Um diese Wechselwirkung zu untersuchen wird in dem Artikel eine neue Dimension der Interaktion von Märkten und Normen entwickelt, die über das Wechselspiel monetärer und nicht-monetärer Anreize einer bestimmten Handlungsweise zu folgen, hinausgeht: die Innovation materieller Güter als Katalysator der Normevolution. Die Katalysatorfunktion der Innovation basiert auf zwei psychologische Faktoren, die Teil des Modells zur Normadoption sein werden. Der Produktmarkt wird als Cournot-Oligopol modelliert mit einer exogenen Anzahl an Firmen, deren Entscheidung darüber die Produktinnovation in ihr Produktionsportfolio mit aufzunehmen endogenisiert wird.

Das Modell zur Normadaption erweitert die bestehende Literatur zur Evolution sozialer Normen auf dreierlei Weise. Erstens, das Modell bezieht die Wirkung von Produktinnovationen auf den Prozess der Normadoption mit ein. Zweitens, der Artikel wird untersuchen wie ein Konformitätsbias im Konsum materieller Güter die Adoption idealistischer Normen beeinflusst. Drittens, das Papier wird veranschaulichen wie die Marktstruktur durch ihre Auswirkung auf die Marktergebnisse die Normdynamik beeinflusst. Dadurch trägt die Arbeit zum Verständnis darüber bei wie die Evolution sozialer Normen von Marktaktivitäten abhängt.

Zwei Fragen wird nachgegangen: Erstens wird untersucht werden wie eine Innovation, die sich durch ihren relativ höheren Grad an Normeinhaltung auszeichnet, die Verbreitung der Norm verändert. Zweitens wird die Wirkung der Markdynamik auf die Evolution der sozialen Norm hinsichtlich der Existenz und Multiplizität der Gleichgewichte analysiert. Hinsichtlich der ersten Frage wird sich zeigen, dass die Innnovation die Normverbreitung erhöht, wenn (1) der Konformitätsbias schwach ist oder genug Individuen bereits vor der Innovation die Norm internalisiert hatten und (2) der Anstieg der individuellen Nachfrage nach dem der Norm entsprechenden Produkt der aus der Normadoption resultiert den korrespondierenden Effekt auf die Nachfrage nach dem die Norm verletzenden Produkt hinreichend stark übersteigt. Diese Bedingungen werden restriktiver je weniger Firmen im Markt aktiv sind, da der notwendige Gewinnanstieg um ein zusätzliches Unternehmen zum Eintritt in den Markt des innovativen Produkts zu bewegen, steigt.

Bezüglich der zweiten Frage wird die Untersuchung zeigen, dass multiple Gleichgewichte nicht nur dann resultieren können, wenn es sich bei der Normadoption um einen frequenzabhängigen Meinungsbildungsprozess mit positiver Rückkopplung handelt, sondern das multiple Gleichgewichte auch dann in Erscheinung treten können, wenn die Normadoption vom beobachteten Marktverhalten, insbesondere vom Anteil des normkompatiblen Konsums, abhängt. Die direkte positive Rückkopplung kann schwächer ausfallen, wenn multiple Gleichgewichte gleichzeitig durch einen Konformitätsbias im Konsum materieller Güter unterstützt werden. Es wird sich zeigen, dass der Effekt der Norm auf die Nachfrage nach dem der Norm entsprechenden Produkt im Vergleich zum Effekt auf die Nachfrage nach dem die Norm verletzenden Produkt weder zu hoch noch zu niedrig sein darf, damit multiple Gleichgewichte entstehen. Im Artikel wird die Marktstruktur als eine zweite Quelle für die Multiplizität von Gleichgewichten diskutiert. Die Endlichkeit der Anzahl der im Markt aktiven Unternehmen bedingt Unstetigkeiten in der Anzahl der Unternehmen. Dies hat Unstetigkeiten in der Markt-Norm-Dynamik zur Folge (siehe Abb.Figure 2-4)). Es wird sich jedoch herausstellen, dass diese Rückkopplung bereits bestehende positive Frequenzabhängigkeiten als Quelle für die Multiplizität der Gleichgewichte zwar verstärken kann, aber kaum allein multiple Gleichgewichte verursachen kann.

Diese Ergebnisse haben Konsequenzen für Politiker, die als mittelfristiges Ziel auf dem Weg zur langfristig angestrebten Reduktion der Umweltverschmutzung auf eine größere Verbreitung sozialer Normen abzielen. Unter anderem wird diskutiert werden, dass der Konformitätsbias so groß sein kann, dass er die Verbreitung der Norm verhindert. Vorwiegend in diesem Fall erscheint die politische Interferenz mit Marktprozessen (und Normbildung) angemessen. Wird von politischer Seite eine Multiplizität von Gleichgewichten aufgrund positiver Rückkopplungen im Prozess der Normverbreitung vermutet und zeichnet die Struktur des neuen Marktes die Züge eines kleinen Oligopol oder gar eines Monopols, dann sollten politische Maßnahmen, die darauf abzielen ein Gleichgewicht mit geringer Normverbreitung zu überwinden, umfangreich und längerfristig wirksam sein. Politische Maßnahmen, welche den Effekt der Norm auf die Nachfrage ändern, sollten nur dann Anwendung finden, wenn die Norm bereits weit verbreitet ist. Sollte dies nicht der Fall sein, so wird die Wirkung nicht nur durch die geringe Anzahl an Individuen reduziert, die möglichweise auf die politischen Maßnahmen reagieren, sondern auch durch die potentielle Wiederbelebung von zumindest einem gewissen Grad an kognitiver Dissonanz, die dann entsteht wenn man eine Norm internalisiert hat, aber nicht dieser entsprechend konsumiert.

Kapitel 3 "Evolution of cooperation in social dilemmas: signaling internalized norms." leistet einen Beitrag zur Erklärung des Phänomens der Kooperation in großen, unstrukturierten Gesellschaften (z.B. Axelrod und Hamilton 1981; Fudenberg et al. 2012). Der Beitrag liegt in der Entwicklung eines alternativen Mechanismus zur Unterstützung von Kooperation in einer Umwelt (Gefangenendilemma), in der nicht-kooperatives Verhalten aus materieller Sicht individuell rational ist. In einer solchen Umwelt kann Kooperation weder durch irgendeine Form wiederholter Interaktion noch durch soziale Normen herbeigefügt werden, die auf Sanktionen basieren, die in zukünftigen Interaktionen auferlegt werden. Selbst internalisierte Normen, d.h. Normen, welche den wahrgenommenen Nutzen aus kooperativen oder nicht-kooperativen Verhalten beeinflussen, können das Dilemma in unstrukturierten Populationen nicht überwinden, es sei denn –und dies ist der alternative Mechanismus– die Individuen sind in der Lage die Eigenschaft ein Normträger zu sein, zu signalisieren. Wenn internalisierte Normen einfach existieren ohne die Möglichkeit diese zu signalisieren oder bei anderen zu erkennen, dann würden diese die Normträger veranlassen zu kooperieren und von anderen ausgenutzt zu werden. Somit hätten Normträger einen klaren evolutorischen Nachteil, der zum Verschwinden der Norm führen würde. Nur wenn Internalisierung der Norm glaubhaft kommuniziert werden kann, mag sich das Bild ändern, da unter diesen Umständen Verhalten auf das erwartete Verhalten anderer konditioniert werden kann.

Ist die Signalisierung kostenlos, so reduziert sie sich zu cheap talk und wird keinen Einfluss auf den evolutorischen Nachteil der Normträger haben. Signalisierung wird demnach mit Kosten verbunden sein und Individuen, welche die Norm angenommen und solche die nicht, mögen sich in den Signalisierungskosten unterscheiden. Im Artikel wird ein Theorem präsentiert, welches notwendige und hinreichende Bedingungen für vollständige oder teilweise Kooperation in einem stabilen Gleichgewicht angibt. Diese Bedingungen nehmen Bezug auf den Unterschied in den Signalisierungskosten zwischen kooperativen und opportunistischen Individuen, auf die Stärke der Kooperationsnorm und auf die Modelparameter des Gefangenendilemmas, d.h. der Anreiz zu Defektieren und der „sucker's payoff". Es ergeben sich mehrere interessante Ergebnisse. Erstens, obschon der exakte Wert des Verhaltensparameters, der den internen Bias zugunsten gegenseitiger Kooperation misst, nicht relevant hinsichtlich der Konsequenz für das Verhalten jedes einzelnen ist, spielt der Wert und dessen Relation zum Defektionsanreiz eine Rolle hinsichtlich der Existenz von Gleichgewichten mit teilweiser Kooperation. Genauer gesagt, je stärker die innere Motivation für kooperatives Verhalten, desto weniger restriktiv sind die Bedingungen für den Unterschied in den Signalisierungskosten. Zweitens, für die Koexistenz von kooperativen und defektierenden Individuen in einem stabilen Gleichgewicht ist es nicht notwendig, dass die Signalisierungstechnologie vollständig den Defektionsanreiz aufhebt. Da dies für viele Ansätze notwendig ist, die auf einer Form unfreiwilliger Umverteilung (z.B. Bestrafung) basieren, kann der Anwendung findende Ansatz Kooperation in mehr Fällen motivieren als die umverteilungsbasierten. Es wird sich des Weiteren zeigen, dass sich die Spanne an Signalisierungskosten der nicht-kooperativen Individuen, welche teilweise oder vollständige Kooperation erlaubt, schwach in der Stärke der sozialen Norm gegenseitiger Kooperation vergrößert. Schließlich wird sich herausstellen, dass die Menge an Paaren aus Signalisierungskosten des nicht-kooperativen Typs und der Stärke der Kooperationsnorm, welche teilweise oder vollständige Kooperation ermöglichen, strikt mit den Signalisierungskosten des kooperativen Typs wächst und sich strikt mit dem „sucker's payoff" und dem Anreiz defektiv auf kooperatives Verhalten zu reagieren, verkleinert.

Kapitel 4 stellt den Artikel "The evolution of inequality aversion in a simplified game of life." vor. Spätestens mit den grundlegenden Arbeiten von Fehr und Schmidt (1999) und Bolton und Ockenfels (2000) avancierte die soziale Präferenz in Form einer Ungleichheitsaversion zu einer bedeutende Erklärung zahlreicher empirischer und experimenteller Befunde, die von den Vorhersagen der klassischen ökonomischen Theorie abweichen. Die steigende Relevanz als Erklärungskonzept verlangt nach einer Rationalisierung derartiger Präferenzen, da sie ansonsten als reine ad hoc Anpassung der Präferenzen gewertet werden könnten, um empirische Beobachtungen zu erklären. Güth und Napel (2006) weisen darauf hin, dass derartige Präferenzen insbesondere mit der physischen Notwendigkeit vereinbar sein sollten, in einer Welt die durch Ressourcenknappheit charakterisiert ist, nach materieller Entlohnung zu streben und um diese zu kämpfen. Mit anderen Worten, derartige Präferenzen sollten sich aus evolutorischer Sicht erklären lassen. Der Argumentation folgend, dass Untersuchungen zur Evolution von Präferenzen in einer Umwelt erfolgen sollten, die im besten Fall alle relevanten Klassen von Spielen umfasst, werde ich eine bestimmte Struktur eines vereinfachten „game of life" vorschlagen. Wie bereits erwähnt beinhaltet das vereinfachte „game of life" wie es definieren werde, ein symmetrisches Dilemma, ein symmetrisches und striktes Koordinationsproblem und ein striktes Verteilungsproblem. Hernach werde ich die Evolution einer konkreten Ausprägung sozialer Präferenzen, die der Ungleichheitsaversion, in eine 2x2 vereinfachten „game of life" untersuchen.

Das vereinfachte „game of life", welches drei besonders wichtige Klassen menschlicher Interaktion beinhaltet, zeigt einerseits wie erwartet eine größere Variation möglicher Gleichgewichtsverteilungen im Vergleich zu den ein einzige Spielklasse umfassende Umwelten. Insbesondere erfahren die überraschend starken Vorhersagen der Einzelspielbetrachtung eine Relativierung. Der globale evolutorische Vorteile ungleichheitsaverser Spieler im Rahmen eines Dilemmas und der globale evolutorische Nachteil in fast allen Fällen für ungleichheitsaverse und im Verteilungsproblem begünstigte Individuen werden relativiert. Dies gilt insbesondere für den Fall, in dem die Wechselwirkung eines Dilemmas und eines Verteilungsproblems ein lokal stabiles Gleichgewicht unterstützen, in dem nur ungleichheitsaverse Individuen existieren. Dann überträgt sich dies auf das vereinfachte „game of life", d.h. Ungleichheitsaversion kann sich auch unter den im Verteilungsproblem begünstigten Individuen etablieren. Andererseits wird sich zeigen, dass die erwartete Stabilisierung innerer Gleichgewichte in denen relative ungleichheitsaverse und relativ opportunistische Individuen koexistieren nur dann auftritt, wenn diese Stabilisierung bereits im Koordinationsproblem für sich genommen erfolgt.

Zusammenfassend, meine Dissertation wird sich mit individuellen Präferenzen beschäftigen, dem zentralen Konzept in der Ökonomik zur Modellierung menschlichen Verhaltens. Genauer gesagt, werden die drei konstituierenden Artikel bestimmte Charakteristika von Präferenzen betrachten, die Auswirkungen auf das Verhalten haben können. Zwei Arten der Strukturierung von Präferenzen werden untersucht, die soziale Präferenz der Ungleichheitsaversion und Strukturen, die aus der Internalisierung sozialer Normen hervorgehen. Meine Dissertation wird die Verbreitung und die Auswirkungen dieser Charakteristika untersuchen. Zwei Kräfte, welche für die Verbreitung dieser Präferenzmerkmale von Bedeutung sind, werden dabei berücksichtigt: evolutionäre und psychologische Kräfte. Meine Arbeit leistet einen Beitrag zum besseren Verständnis der Entstehungsbedingungen und der Konsequenzen möglicher Erklärungsansätze für beobachtetes menschliches Verhalten, welche auf bestimmte Strukturen individueller Präferenzen zurückgreifen.

Acknowledgment

I want to thank my supervisor Prof. Dr. Georg von Wangenheim. My thesis is based on many hours of thinking, calculating and writing. Quite a substantial fraction of that time I directly worked together with Georg, and I really enjoyed it. I cannot remember one working session which was not encouraging and fruitful. I learned a great deal of economics and academics while working with him. Apart from this intellectual aspect I really appreciated the social and cooperative atmosphere we created. He was and is an important mentor and colleague of mine.

I am grateful to my parents who always encouraged me in my thirst for knowledge and were very patient regarding my disputability from the early childhood on. I am very thankful for their emotional and material support in particular during my studies, first in economics and in mathematics thereafter.

I am also thankful to the members of the economic community for valuable and inspiring discussions on topics of my thesis. In particular I want to thank Ivo Bischoff, Guido Bünstorf, Christian Cordes, Dirk Fornahl, Björn Frank, Werner Güth, Claudia Keser, Steffen Kolb, Bertrand Munier, Lennard Munzer, Stefan Napel, and Georg Schwesinger. I am also grateful to Claudia Keser for granting me a huge amount of academic freedom and for her patience regarding the completion of my thesis.

Finally, I want to thank my beloved wife and my daughter. We as a family are the foundation of our daily wellbeing and the realization of long run aspirations. I am very thankful that my wife and my daughter enabled me to pursue my goal to write a doctoral thesis. I dedicate my thesis to myself in humility and in absolute awareness of the great luck of having such an opportunity.

Contents

List of Figures

List of Tables

1. Introduction

1.1. Motivation

In Economics preferences are the central concept in the framework to analyze human behavior. They refer to the set of assumptions concerned with the ordering of choice-alternatives. These alternatives can incorporate aspects of uncertainty or intertemporal issues. A mathematical model of preference relations was first written down by Frisch (1926) and brought to perfection by Arrow (1951). A preference relation which models preferences is the foundation of rational choice theory in economics. Rational choice is defined to mean the process of determining the set of options to choose from and then selecting the most preferred one according to some consistent criterion. Rational choice theory is based on a rather sparse description of the choice problem (objectives and constraints). That choice depends on many more factors is for instance proposed by psychological theories and supported by many laboratory experiments (Hogarth and Einhorn 1992; Hoffman et al. 1994; Kahneman and Frederick 2002, to name a few). Such factors include for instance the way or the order information is presented. The empirical failings of rational choice theory in economic and psychological experiments have triggered intense research in that field and many alternative models have been proposed. To understand individual behavior is one of the fundamental tasks for economic research. My thesis contributes to this research agenda. The emergence of certain particularities of preferences and their impact on behavior are the overarching theme of the constituting three articles. Beyond this unifying theme, which is of central interest for me each of the articles is motivated by a more narrow research question. This will be exemplified in the next paragraphs.

Methodological individualism and the concept of homo oeconomicus form the basis for standard economic theory. Preferences and restrictions of the action space determine individual behavior. To operationalize the basic assumptions of rationality and narrow self-interest of homo oeconomicus preferences are assumed to be stable over time. Stable preferences are not meant to be a descriptive model for real individuals though. Even if preferences are stable in the short run, the questions arise why and how certain features like altruism, risk aversion, loss aversion, repeatedly confirmed in experiments were, "implemented" into humans. The evolutionary perspective offers one framework to answer these questions. Within this framework economists try to rationalize specific aspects of human preferences. At the latest with the seminal work of Fehr and Schmidt (1999) and Bolton and Ockenfels (2000) other-regarding preferences in the form of inequality aversion have become a prominent explanation for many empirical and experimental findings which deviate from the prediction of standard economic theory. Increasing importance of other-regarding preferences in behavioral economics and other fields' calls for a rationalization for such preferences, otherwise it may be regarded as an ad-hoc adjustment of preferences to explain empirical results. As Güth and Napel (2006) point out such preferences should in particular be compatible with the physical necessity to strive and compete for material rewards in an environment characterized by the scarcity in resources. In other words such preferences ought to be rationalizable from an evolutionary point of view. The authors convincingly argue that any type of study concerned with the evolution of preferences need to be carried out in an environment that comprises all of the classes of human interaction since otherwise game-specific results are obtained whose generalizability is at least questionable. They refer to such an environment as the 'game of life'. In the article 'The evolution of inequality aversion in a simplified game of life' I make a first step towards fulfilling this requirement.

However, preferences are not only shaped by evolutionary forces, which work on a long time scale and transform the genetically encoded part of preferences. Formal and informal institutions enable and constrain actions of individuals. If informal institutions like social norms become internalized they become part of the preferences of an individual. The adoption of a social norm works on a much shorter time scale and depends on multiple factors. In my understanding social norms are inner "recommendation" to act in a certain way. They may be complemented or substituted by external incentives like cost and benefits associated with a chosen action. These cost and benefits are thus part of the aforementioned external restrictions to behavior. In modern economies a large part of human interaction takes place in market environments. Hence most of the behavioral constraints materialize in markets. By aggregating individual behavior markets also aggregate individual external effects. Many of today's environmental problems stem from private consumption patterns. Individuals consume transportation, heating and food, all of which cause substantial emissions of CO_2 and other pollutants. Preferences for choosing more or less polluting products and services are shaped by their availability as well as by social norms and other institutions. Thus, solutions to mitigate environmental problems depend not only on product innovation, but also on the presence of social norms, with the former enriching markets with sustainable products, and the latter supporting sustainable consumption. When recognizing that social norms influence preferences, it becomes apparent that markets and social norms cannot be treated separately. The existing literature has widely studied the interrelation between markets and social norms in both directions – how social norms affect markets and how markets affect social norms (e.g. Hong and Kacperczyk 2009; Johnson 2004; Ek and Soderholm 2008; Fehr and Gächter 2001 and Gneezy and Rustichini 2000). All of these approaches are limited to monetary incentives provided by markets and their regulation. This, however, reduces markets to their price-quantity aspect and completely neglects their innovation capacity. The product variation due to such innovations is an important element of the market-norm interaction. It is this interrelation that is focused on in the article "The impact of market innovations on the evolution of social norms: the sustainability case".

The aspect of sustainability points to the more general class of problems of social dilemmas. Social dilemmas are characterized by the property that individually rational behavior leads to collectively irrational outcomes. Economists however observe in experiments (for a survey see Roth 1995) and in the field (Fey and Meier 2004; Cunha and Augenblick 2014) cooperation in the sense that this collective irrationality is at least partially circumvent. This is puzzling from the traditional economic perspective in particular as cooperative behavior emerges even in the absence of any shadow of future interaction (Cooper et al. 1996). Attempts to solve the puzzle are abundant but have thus far commonly relied on one or both of two restrictions. The first restriction is that explanations have focused on structured populations, in which interactions are not completely anonymous but allows individuals to collect and process information about past behavior of others and about their identity. The second restriction is that explanations have depended on an unexplained ability of social norms to restrict the individuals' action or strategy spaces, in particular, with respect to the abuse of punishment. The article "Evolution of cooperation in social dilemmas: signaling internalized norms." presents a new explanation for cooperation that avoids both restrictions.

1.2. Positioning of the thesis: the broad perspective

The three articles constituting my thesis are related to three fields of economics: Behavioral Economics, Formal Institutional Economics and Evolutionary Game Theory. Figure 1-1 illustrates the relation graphically.

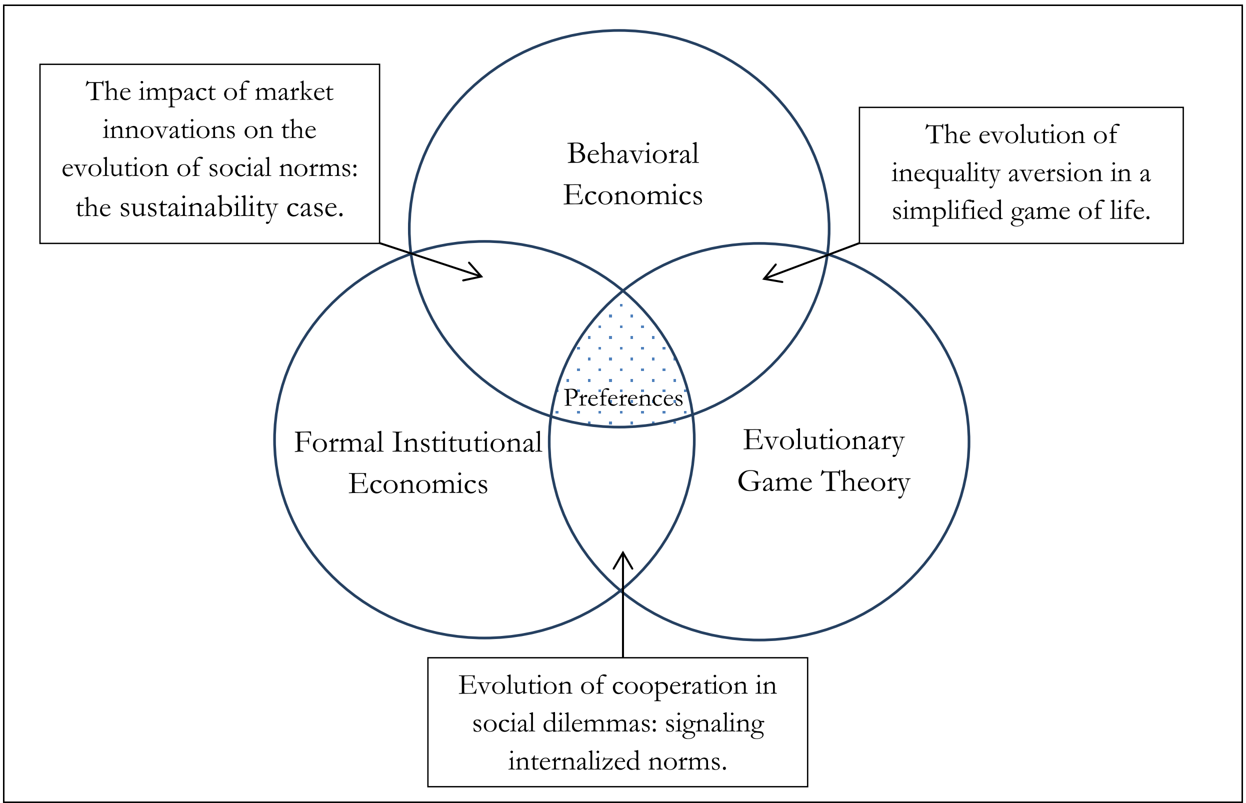


Figure 1-1: The fields of economics the articles relate to.

Behavioral Economics: Incorporating psychological, cognitive and emotional factors is the distinctive feature of behavioral economics. It thereby tries to increase the explanatory power of economics. The attempt to set the discipline of economics on more realistic psychological grounds does not imply a rejection of the neoclassical approach. Indeed most of the papers in this field modify only one or two assumptions in standard theory to achieve a higher degree of psychological realism. As Camerer and Loewenstein (2004) point out there is nothing in core neoclassical theory that specifies that people should not take fairness into consideration or that they should weight risky outcomes in a linear fashion. Some of the aforementioned modifications relax these simplifying assumptions. Other modifications take cognitive limitations of humans into account. These assumptions refer to what Herbert Simon calls “procedural rationality” (Simon 1976). From a methodological point of view behavioral economics initially made primarily use of experimental data. More recently field experiments (Gneezy and Rustichini 2004) and computer simulations (Angeletos et al. 2001) have been utilized.

According to Camerer and Loewenstein (2004) behavioral decision research as the primary source for behavioral economics with respect to psychological aspects to be incorporated can be classified into two categories: judgment and choice. How people estimate probabilities is at the core of the former. The latter is concerned with how people select among actions.

The paper on the impact of market innovation incorporates psychological aspects of the adoption process of a social norm. After a new product or service which is characterized by a relatively high degree of norm compliance has entered the market, the process of norm adoption changes in two ways. First, an individual is now able to consume in accordance with his or her norm, which could not have happened before the innovation. Second, the new variety allows the social influences like the pressure to behave in conformity to others (Boyd and Richerson 1985) to enter the scene.

The paper on the evolution of inequality aversion is related to the field of behavioral economics as it is concerned with inequality aversion which is considered to give rise to a more realistic view of human behavior. It provides an evolutionary foundation of an explanatory variable in behavioral decision research.

Evolutionary Game Theory: Until today there is no common agreement about what is meant by the "evolutionary" viewpoint in economics. Witt (2008) reflects for the field of evolutionary economics on the three levels of scientific reasoning, the ontological level, the heuristic level and the methodological level and classifies the existing differing views. In what follows I will briefly restate his insights which will helpful to position the two articles which carry a strong evolutionary notion. With respect to the ontological stance Witt (2008) identifies on the one hand the monistic view and on the other the dualistic view. The former assumes that the economic sphere and nature are connected spheres with potentially interdependent processes. The latter explicitly rejects this view and treats economic and biological evolutionary processes as being part of disconnected spheres of reality. On the heuristic level Witt (2008) distinguishes the generalized Darwinian heuristic and the generic evolutionary heuristic. The former applies the three principles of evolution that follow by abstract reduction from the Darwinian theory of natural selection (blind variation, selection, and retention) to conceptualizing the evolution of technology (Ziman 2000), science (Hull 2001), the economy (Nelson 1995) and others. The latter is not based on an analogy between economic and biological evolutionary processes, but by a generic concept of evolution. This concept characterizes evolution as a process of self-transformation with the endogenous generation of novelty and its contingent dissemination as its constituting elements (Witt 2003, Chap.1). The two ontological stances and the two heuristic strategies allow the different approaches in evolutionary economics to be represented in a 2x2 matrix.

According to Witt (2008) applications of evolutionary game theory to the discipline of economics can be distinguished along these dimensions and have essentially two interpretations. In the first interpretation models from evolutionary biology which describe the selection process are applied to an economic context to model learning processes (see Brenner 1999, Chap.6). This interpretation makes use of the heuristic strategy of assuming an analogy between adaption in biology and adaption in economics through non-cognitive learning. From an ontological perspective the analogy construction typically does explicitly answer the question of whether, and how, the economic processes connect to the naturalistic foundation of human behavior (Witt 2008). Hence it parallels the entry that is beset with the Neo-Schumpeterian synthesis of Nelson and Winter (1982) in the aforementioned 2x2 matrix structuring the field of evolutionary economics with respect to the ontological stance and the applied heuristic strategy. The second interpretation is not based on an analogy but the biological context is directly relevant for the application in economics. Authors taking up this position claim that very basic features of human behavior are genetically encoded and can therefore be best understood from the viewpoint of natural selection. Among such basic features of human behavior are altruism, fairness or morality (see. e.g. Güth and Yaari 1992a; Binmore 1998; Gintis 2007). This direct transfer from biology to economics obviously presupposes a monistic ontology. Witt (2008) argues the heuristic strategy applied in the research following the second interpretation has some similarities with Hayek's theory of societal evolution.

In this view the article on the evolution of inequality aversion which is presented in Chapter 4 follows the second interpretation, whereas the article that deals with the signaling of internalized norms follows the first interpretation. I study the evolution of inequality aversion in an environment to what I refer as a simplified "game of life" (Güth and Napel 2006). This environment comprises three classes of games which are representative for most of human interactions. The article on the evolution of cooperation focuses on one particular game, the Prisoners' Dilemma, and studies the signaling of a cooperative norm as a mechanism to foster cooperation.

Formal Institutional Economics: The term "institutional economics" has been applied to a variety of economic approaches and schools of thought. Most of the time it refers to economics in the tradition of Thorstein Veblen, John R. Commons, and Wesley Mitchell. In recent years the term "new institutional economics" has become well-established. This term refers to economics in the tradition of the transactions cost approach of Ronald Coase, Oliver Williamson, and Douglas North. Nowadays the term is often extended to incorporate game theoretic approaches to the evolution of social convention, and sometimes to institutions in the tradition of the Austrian School (see Rutherford 2001). In this thesis I follow the game theoretic approach. To distinguish it from the traditional institutional economics and the primarily non-formal approaches of new institutional economics I refer to it as formal institutional economics. In the literature applying a game theoretic framework to institutions two approaches defining the term can be distinguished: the equilibrium approach and the rules-of-the-game approach. In the equilibrium approach the core defining element of institutions is the equilibrium character of individual actions. Most of the time the equilibrium character translates into a notion of stability. Some authors rely on concepts of evolutionary game theory (Sugden 1986, 1989; Young 1998; Aoki 2000; Bowles 2000), others use the theory of repeated games of standard game theory (Greif 1989, 1997, 1998; Milgrom et al. 1990; Calvert 1995). The rules-of-the-game approach treats institutions as an external factor shaping the actions spaces and payoffs of the game (North 1990; Hurwicz 1993, 1996). There is a huge gap between the wide definition in the field of new institutional economics and the narrow definition within the equilibrium approach and the rules-of-the-game approach.

This gap is narrowed down by the indirect evolutionary approach pioneered by Güth and Yaari (1992b). Within the indirect evolutionary approach behavioral and fitness or material payoffs are distinguished. Fitness-payoffs of a certain behavior are relevant for the diffusion of that particular strategy among individuals. Behavioral payoffs reflect the inner evaluation of those fitness payoffs. These payoffs are relevant for decision making but have no influence on the adoption rate of that strategy by other agents.

The indirect evolutionary approach allows to model institutions in a more complex way beyond a simple regularity in behavior or the rules of a game. In particular the informal institution of a social norm may be modeled in this way. It thus provides a first step towards reducing the aforementioned gap in the complexity of the different definitions. For this reason I will apply the indirect evolutionary approach in the two articles that take an evolutionary perspective (Chapter 3 and 4).

1.3. Positioning of the thesis: the narrow perspective

In this section I will briefly relate each of the articles to the existing literature and identify the research questions being addressed. As mentioned earlier "preferences" constitute the overarching theme of my thesis.

I consider social norms as important entities shaping individual preferences. The article presented in Chapter 2 studies the interdependency of markets and the adoption process of a social norm. It applies the psychological realism of behavioral economics to the analytical framework of formal institutional economics. Furthermore it provides some insight for the industrial organization literature with respect to the impact of social norms on product markets. The influence of social norms on markets has been studied from theoretical, empirical and experimental perspectives. Several attempts have been made to incorporate norm-motivated behavior into neoclassical consumer theory (see e.g. Nyborgs et al. 2006; Brekke et al. 2003). Despite these attempts there is no general or partial equilibrium theory based on norm-motivated behavior, what may explain why most research in the field is empirical. How social norms influence a particular type of market, the financial market is studied by Hong and Kacperczyk (2009) and Johnson (2004). Kim (2007) shows that norms are also relevant for markets of private property rights. A series of competitive-market and bilateral-bargaining experiments carried out by Fehr et al. (1998) indicate that competition has a rather limited effect on market outcomes if the norm of reciprocity is operative. The role of the psychological need to maintain a positive self-image as a socially responsible person on the demand for "green" electricity is studied by Ek and Soderholm (2008).

The research on the impact of markets on the evolution of norms primarily deals with the analysis of the relationship of norm-driven intrinsic motives and market- or price-driven extrinsic motives. There is empirical support (Fehr and Gächter 2001) that incentive contracts crowd out reciprocity-driven voluntary cooperation. A first survey of this stream of empirical literature on crowding-out effects was carried out by Frey and Jegen (2001). There is also theoretical research. Benabou and Tirole (2006) provide a theory of pro-social behavior where rewards or punishments create doubt about the true motives for which good deeds are performed. Consequently this may lead to partial or even total crowding-out of pro-social behavior. The interplay of social norms and economic incentives in firms is modeled by Huck et al. (2012). A study that provides both a theoretical model and evidence form the laboratory is performed by Bohnet et al. (2001). They study the connection between contract enforceability and individual performance. The results show that trustworthiness is "crowded in" with weak enforcement and "crowded out" with medium enforcement. All of these approaches are limited to monetary incentives provided by markets and their regulation. This, however, reduces markets to their price-quantity aspect and completely neglects their innovation capacity. The variation due to such innovations is an important element of the market-norm interaction. This gap in research on the interdependency between innovative variation of products and the evolution of social norms is addressed in Chapter 2.

Chapter 3 presents a paper that applies the formal institutional economics' perspective on institutions to the framework of evolutionary game theory. It analysis a particular social norm, a norm to behave cooperatively, and its potential to resolve social dilemmas. More precisely it deals with the puzzle of the emergence of cooperation in large, unstructured societies in an environment where non-cooperative behavior is individual rational. As mentioned before most attempts rely on one or both of two restrictions. Either it is structured population which is analyzed or social norms are given the unmotivated ability to restrict the individuals' action or strategy spaces.

With respect to the first group of restrictions, some strands of the literature deserve special mention. The theory of kin selection focuses on cooperation among individuals who are genetically closely related (Hamilton 1964a, 1964b), whereas theories of direct reciprocity focus on incentives to cooperate in repeated interactions of self-interested individuals (Trivers 1971; Axelrod 1984). For infinite repetition within one group, see Taylor (1976) or Mordecaï (1977) and for Folk-Theorem-type of results Rubinstein (1979) or Fudenberg and Maskin (1986). For indefinite repetition, see Kreps et al. (1982). The theories of indirect reciprocity and costly signaling show how cooperation in larger groups can emerge when those cooperating can build a reputation (Nowak and Sigmund 1998; Wedekind and Milinski 2000; Gintis et al. 2001).

In terms of the second set of exclusions, reference is made to early papers of Hirshleifer and Rasmusen (1989) and Witt (1986) that allow for punishment only after a norm has been violated. Sethi (1996) allows for all possible strategies which condition punishment on either the violation of or compliance with a norm. However, he then adds structure to the society by introducing some exogenous division of the population – the behavior of some individuals is rational, and for the rest it is determined by routines that are slowly adapted to their environment.

The article on the evolution of cooperation in social dilemmas offers an alternative explanation for the emergence of cooperation that does not depend on these two restrictions.

Whereas Chapter 3 deals with the important class of human interactions of social dilemmas, the article presented in Chapter 4 studies the evolution of inequality aversion in a simplified compound environment which comprises three classes of human interaction: a social dilemma, a problem of coordination and a problem of distribution. The concept of inequality aversion plays an important role in behavioral decision research. In Chapter 4 the framework of evolutionary game theory is applied to this particular behavioral determinant. In the past the evolution of preferences has been studied in highly artificial single-game environments (e.g. Huck and Oechssler 1999; Koçkesen et al. 2000a, 2000b and Sethi and Somanathan 2001). More recently, some attempts were made to analyze the evolution of preferences in more complex environments. Güth and Napel (2006) analyze how the personal characteristic of inequality aversion evolves in a setting containing two well-studied and characteristic games: the Ultimatum game and the Dictator game. Poulsen and Poulsen (2006) study the evolution of other-regarding preferences in an environment that comprises a simultaneous and a sequential Prisoners' Dilemma. The results of the latter suggest that the results of single-game environments should be treated with caution because they demonstrate a significant change in results once more complex environments are analyzed. The gap in research being identified, Chapter 3 provides a step towards conceptualizing a framework for evolutionary studies of preferences and applies this framework to the other-regarding preference of inequality aversion.

A prerequisite for the analysis of the evolution of preferences in the game of life is the structuring of the infinite set of potential games, which is the second aim of the paper. There is evidence that human behavior is not game-specific, but acts of men are similar in entire, quite general classes of games (see Yamagishi et al. 2013; Ashraf et al. 2006; Blanco et al. 2011; Chaudhuri and Gangadharan 2007 and Slonim and Garbarino 2008). This raises hope that the overwhelming complexity of the real world might be reducible to these classes when the evolution of preferences is considered. Many authors implicitly or explicitly share and express the view that there are two fundamentally different societal problems (see e.g. Sugden 1986; Milgrom et al. 1990), problems of coordination and social dilemmas. Apart from these two classes, Schotter (1981), Ullmann-Margalit (1977) and others share the view that there is (at least) a third type of social problems, one of distributive nature. A problem of distribution is characterized by unequal payoffs in equilibrium. The notion of a game of life which I will suggest comprises these three classes of games.

1.4. Contributions and main conclusions

This section highlights the contribution and presents the main conclusions of each of the articles.

Chapter 2 "The impact of market innovations on the evolution of norms: the sustainability case." is concerned with the gap in research identified in 1.3: the interdependency between innovative variation of products and the evolution of social norms. To analyze this interdependency the paper will introduce a new dimension to the interaction between markets and norms beyond the interplay of monetary and non-monetary incentives to act in a certain way: the innovation of material goods as a catalyst of norm evolution. The catalytic function of the innovation is based on two psychological forces being incorporated in a model of norm adoption. The product market is modeled by a Cournot-oligolpoly with a fixed number of firms which decision whether or not to add the innovative product to their production portfolio is endogenized.

The model extends the existing literature on the evolution of social norms in three ways. First, the model incorporates the influence of a product innovation on the process of norm adoption. Second, the paper will analyze how a conformity bias in the consumption of material goods affects the adoption of idealistic norms. Third, the paper will demonstrate how market structure, through its impact on market outcomes, may influence norm dynamics. The paper will thereby add to the understanding of how the evolution of norms depends on market activities.

Two questions will be pursued. First, it will be studied how an innovation that differs with respect to the level of norm compliance modifies the dissemination of a norm. Second, it will be investigated the effect of market dynamics on the evolution of the norm with respect to the existence and stability of the equilibria. Concerning the first question, the innovation increases the norm diffusion if (1) the conformity bias is weak or enough individuals already bear the norm prior to the innovation and (2) the increase of individual demand for the norm-compliant product variant resulting from norm adoption exceeds the corresponding effect on the demand for the norm-violating variant by a sufficient degree. These conditions become more restrictive when fewer firms are in the market, since then the required increase in profits to induce an additional incumbent to produce the innovative product increases.

With respect to the second question, the analysis will reveal that multiple norm equilibria may not only result if norm adoption is a frequency-dependent opinion formation process with direct positive feedback loops. But multiple norm equilibria may also emerge if norm adoption depends on observed market behavior, in particular, on the proportion of norm compliant consumption. The direct positive feedback loop may be weaker when multiple equilibria are also supported by a conformity bias in consumption of material goods. It will turn out that the effect of the norm on the demand for the norm-compliant variant may be neither too high nor too low as compared to the effect on demand for the norm-violating product for multiplicity to arise. In the paper a second possible source of multiplicity of norm equilibria will be discussed, the market structure. The number of operating firms in the market being finite introduces discontinuities in the number of firms also operating in the market for the innovative product. Consequently the market-norm dynamics shows discontinuities what may generate multiple equilibria (see Figure 2-4). It will turn out, though, that this feedback loop may reinforce already existing positive frequency dependency as source of multiplicity of equilibria, and will rarely induce multiple equilibria on its own.

The results have consequences for policy makers aiming at a higher dissemination of the social norm as an intermediate goal to ultimately achieving the greater goal of reducing environmental pollution. Among others it will be discussed that the conformity bias may be so strong that it hinders the dissemination of the innovation. It is mainly in these cases where political interference with market forces (and norm formation) is appropriate. If policy suspects the existence of multiple equilibria due to positive feedback loops in the norm formation process and the market structure on the new market is a small oligopoly or even a monopoly, then policies aiming at overcoming equilibria of little norm adoption have to be strong and patient. Political measures which alter the effect that the norm imposes on demand should only be implemented when norm adoption is already wide spread. If it is not, the effect is not only diminished by the small number of individuals who may react to the policy measure, but also by a possible reintroduction of at least some cognitive dissonance from having the norm but not complying with it.

Chapter 3 "Evolution of cooperation in social dilemmas: signaling internalized norms." contributes to solving the puzzle of the emergence of cooperation in large, unstructured societies (e.g. Axelrod and Hamilton 1981; Fudenberg et al., 2012). It contributes to the literature on the emergence of cooperation by offering an alternative mechanism to foster cooperation in an environment (Prisoners' Dilemma) where non-cooperative behavior is materially individual rational. In such an environment, cooperation cannot be induced by any form of repeated interaction nor by social norms based on sanctions to be inflicted in later interactions. Even internalized norms, i.e. norms that alter the perceived utility from acting in a cooperative or uncooperative way, will not help to overcome a dilemma in an unstructured society, unless – and this is the alternative mechanism – individuals are able to signal their property of being a norm bearer. If internalized norms simply exist while lacking the possibility of being signaled or screened for, they would induce norm bearers to cooperate and be exploited by others. Hence, norm bearers would have a clear evolutionary disadvantage so that norm adoption would vanish. Only when internalization of the norm can be communicated in a reliable way, may the scenario change, because behavior may then be conditioned on the expected behavior of others.

If signaling is costless then signaling is reduced to cheap talk and will not alter the evolutionary disadvantage of norm-bearers. Thus signaling will be costly and individuals bearing the norm and those who don't may have different signaling cost. The paper will present a theorem that states necessary and sufficient conditions for full or partial cooperation to be prevalent in a stable equilibrium. These conditions will refer to the difference in signaling cost between the cooperative and the opportunistic type, the extent of the cooperative norm and the model parameters of the Prisoners' Dilemma, i.e. the temptation to defect and the sucker's payoff. Several interesting results will be obtained. First, it is true that the exact size of the behavioral parameter measuring the internal bias in favor of mutual cooperation is not important for the behavioral consequence for each individual. However, when it comes to the presence of stable equilibria characterized by partial cooperation its size and its relation to the incentive to defect do become relevant. More precisely, the stronger the inner motive to cooperate is, the less restrictive are the conditions on the spread in signaling cost. Second, for cooperative agents to coexist with defecting agents in a stable equilibrium, it is not necessary that the signaling technology fully cancels the incentive to defect. Since this would be necessary for many corresponding results that are based on some sort of involuntary redistribution (e.g. punishment), the applied approach may explain cooperation in more cases than the latter approaches. Furthermore, the range of signalling cost for the defective individuals allowing for partial or full cooperation is weakly increasing in the strength of the social norm for mutual cooperation. Finally, the set of pairs of signalling cost for the defective type and level of cooperative norm allowing for partial or full cooperation is strictly increasing in signalling cost for the cooperative type and strictly decreasing in the sucker's payoff and the incentive to defect on cooperation.

Chapter 4 presents the article "The evolution of inequality aversion in a simplified game of life." At the latest since the seminal work of Fehr and Schmidt (1999) and Bolton and Ockenfels (2000) an other-regarding preference in the form of inequality aversion has become a prominent explanation for many empirical and experimental findings which departure from the prediction of standard economic theory. The increasing importance calls for a rationalization for such preferences, otherwise it may be regarded as a rather ad-hoc adjustment of preferences to explain empirical results. As Güth and Napel (2006) point out such preferences should in particular be compatible with the physical necessity to strive and compete for material rewards in an environment characterized by the scarcity in resources. In other words such preferences ought to be rationalizable from an evolutionary point of view. Following the argument of the necessity to analyze the evolution of preference in an environment that comprises at best all relevant classes of games individuals engage in, I will suggest a particular notion of a simplified game of life. The simplified game of life as I will define it comprises three classes of games: a symmetric dilemma, a symmetric and strict problem of coordination and a strict problem of distribution. Then I will analyze the evolution of a particular type of other-regarding preference that of inequality aversion in the 2x2 simplified game of life.

The simplified game of life that comprises three major important types of human interaction, on the one hand as expected gives rise to a greater variety in potential equilibrium distributions of preferences than the single environments. In particular the surprisingly strong predictions for the single environments are put into perspective. The global evolutionary advantage of inequality-averse players in the dilemma and the global disadvantage in almost all cases for inequality-averse individuals who are favored in the problem of distribution experience significant qualification. In particular whenever the interplay of the dilemma and the problem of distribution allows for a locally stable equilibrium with only inequality-averse players then this transfers to the simplified game of life, i.e. inequality aversion may also be present among individuals who are favored in the problem of distribution. On the other hand the expected stabilization of inner equilibria in which relatively inequality-averse individuals and relatively selfish individuals coexist occurs only if the problem of coordination shows the same feature.

In summary my thesis will deal with individual preferences, the central concept in economics to model behavior. More precisely the constituting articles will be concerned with certain characteristics of preferences which may have behavioral consequences. Two types of structuring of preferences will be incorporated, the other-regarding preference of inequality aversion and the structures that emerge from the internalization of social norms. My thesis will analyze the dissemination and the impact of these characteristics. Two forces relevant for the dissemination of these particularities of preferences in a population will be considered: evolutionary and psychological forces. The thesis contributes to a better understanding of the conditions of emergence and the consequences of potential explanations for observed human behavior which rely on certain structures of individual preferences.

2. The impact of market innovations on the evolution of norms: the sustainability case.

2.1. Abstract

It is widely accepted among economists that institutions and in particular social norms as an important category of informal institutions do matter. Social norms matter in many economic situations, particularly for markets. The economic literature has studied the interrelation between markets and social norms in both directions – how social norms affect markets and how markets affect social norms. In the latter, markets are reduced to their price-quantity aspect and innovation is completely neglected. Our paper introduces a new dimension to the interaction between markets and norms beyond the interplay of monetary and non-monetary incentives to act in a certain way: the innovation of material goods as a catalyst of norm evolution. We analyze how the evolution of a social norm may be affected by product innovation, which adds to the variation of products with respect to their level of norm compliance. We derive necessary and sufficient conditions for a) a positive impact of the innovation on the level of norm adoption and b) for multiplicity of norm equilibria. In concluding, we discuss several policy implications.

Keywords: Consumer Behavior – Social Norms – Evolutionary Economics – Sustainability – Innovation

JEL Classifications: A13, D02, D11, Q01, Q55

2.2. Introduction

Many of today's environmental problems stem from private consumption patterns. Individuals consume transportation, heating and food, leaving a significant carbon footprint. Preferences for choosing more or less polluting variants of these products and services are shaped by their availability as well as social norms and other institutions. Thus, solutions to mitigate environmental problems depend not only on product innovation, but also on the presence of social norms, with the former enriching markets with sustainable products, and the latter supporting sustainable consumption. When recognizing that social norms influence preferences, it becomes apparent that markets and social norms cannot be treated separately.

The existing literature has widely studied the interrelation between markets and social norms in both directions – how social norms affect markets and how markets affect social norms. The influence of social norms on markets has been studied from theoretical, empirical and experimental perspectives. With respect to theory, there have been various attempts to incorporate norm-motivated behavior into neoclassical consumer theory (see e.g. Nyborgs et al. 2006; Brekke et al. 2003). Or social norms are treated as a prerequisite for working market systems (e.g. Platteau 1994)[1]. However, there is no general or partial equilibrium theory based on

[1] For a normative theory of social norms in market economies, see Bergsten (1985).

norm-motivated behavior[2]. This may explain why most research in the field is empirical. Hong and Kacperczyk (2009) and Johnson (2004) for instance study the impact of norms on financial markets. Kim (2007) finds support for the relevance of norms for the market pricing of private property rights. A series of competitive-market and bilateral-bargaining experiments carried out by Fehr et al. (1998) indicate that competition has a rather limited effect on market outcomes if the norm of reciprocity is operative. The impact of wanting to maintain a positive self-image as a socially responsible person on the demand for "green" electricity is studied by Ek and Soderholm (2008). Johnson (2004) develops a framework using evidence from central Kenya for the relationship between gender norms and financial markets, i.e. the demand and access to financial services.

The research on the impact of markets on the evolution of norms primarily deals with the analysis of the relationship of norm-driven intrinsic motives and market- or price-driven extrinsic motives. Fehr and Gächter (2001) provide empirical support for incentive contracts crowding out reciprocity-driven voluntary cooperation. In a similar vein, Gneezy and Rustichini (2000) present results from a field study that contradict any deterrence hypothesis. A first survey of this stream of empirical literature on crowding-out effects was carried out by Frey and Jegen (2001). With respect to theory, Benabou and Tirole (2006) provide a theory of pro-social behavior where rewards or punishments create doubt about the true motives for which good deeds are performed, and hence, may lead to partial or even total crowding-out of pro-social behavior. Huck et al. (2012) provide a model of the interplay of social norms and economic incentives in a firm in which crowding-out of social incentives may occur. Bohnet et al. (2001) study the connection between contract enforceability and individual performance, both theoretically and in the laboratory. They find that trustworthiness is "crowded in" with weak enforcement and "crowded out" with medium enforcement. All of these approaches are limited to monetary incentives provided by markets and their regulation. This, however, reduces markets to their price-quantity aspect and completely neglects their innovation capacity. The variation due to such innovations is an important missing element of the market-norm interaction.

In this paper we try to close that gap by focusing on the interdependence between innovative variation of products and the process of norm-adoption. To understand the explanatory potential of the interdependence, consider a market where at the pre-innovation stage, the individual characteristic of having adopted a specific norm is not observable, neither by observation of the individual itself or its general behavior, nor by observation of its consumption behavior. Obviously, the latter presupposes that products or services fail to differ with respect to their norm compliance. After a new product or service which is characterized by a relatively high degree of norm compliance has entered the market, the process of norm adoption changes in two ways. First, an individual is now able to consume in accordance with his or her norm, which could not have happened before the innovation. The innovation thereby directly facilitates the adoption of the norm by reducing potential cognitive dissonances that would occur if a norm adopter consumes in contradiction to his or her norm. We call this event *cognitive bias*. Second, the

[2] For a discussion of an extension of Walrasian economics by social norms and psychological dispositions see Bowles and Gintis (2000). For a multi-agent simulation model on the psychological factors like need for identity on market dynamics, see Janssen and Jager (2001).

new variety allows the *conformity bias* (Boyd and Richerson 1985) and other social influences (Cialdini and Goldstein 2004) to enter the scene. The consumption of the (old) norm-violating product and of the (new) norm-complying product will hence become more attractive, the more other individuals still, or already consume the respective product.

In our model, we address both of these two elements of the link between product innovation and the evolution of social norms. To achieve this goal, we consider a market in which consumers are heterogeneous with respect to their norm-dependent and product-specific demand and the producers' product-portfolios heterogeneity evolves endogenously. The equilibrium of this market depends on the share of consumers who have adopted the norm to which the new product complies. Conversely, the innovation and the equilibrium ratio of norm-complaint to norm-violating consumption affects the norm-adoption process via the two biases we introduced in the preceding paragraph. Since the equilibria of markets strongly depend on market structure, and markets for innovative products are highly susceptible to monopoly or oligopoly power, we control for market structure. We do so by opposing the two cases of a discrete number of firms and a continuum of producers of the innovative product.

The link between the process of norm adoption and the market may only be relevant if the product or service is sufficiently important for individuals in terms of the time spent with it, money spent on it, utility drawn from it, social status connected to it etc. since otherwise, cognitive dissonances would be too weak to have a major impact. For our analysis, we therefore employ *e-mobility* as the innovation and *sustainable transportation* as the norm. In 2010, German households spent around two-thirds of their income on the following four categories: housing, water, electricity, gas and other fuels (30.8%); transportation (13.2%[3]); leisure, entertainment and culture (11.6%) and food including non-alcoholic beverages (10.4%). Of these four categories, only the expenditure for transportation and food reflect the attitude towards sustainable consumption in an observable way.[4] According to an extensive study on mobility in Germany conducted by the *infas* Institute for Applied Social Sciences and the *DLR* German Aerospace Centre in 2008 (MiD 2008, p.21), a mobile person spent on average 1.5h a day on traveling excluding regular travel time associated with a job, e.g. as a bus driver. Almost 60 % of that time (about 54 minutes) is assigned to private transportation. In summary, the car is expensive, important, omnipresent and relevant for sustainable consumption and therefore a product with a high potential for a conformity bias and cognitive dissonances for norm-adopters.

Our analysis, however, is not limited to this case. We include two other examples that illustrate the wider relevance of our approach. Consider first the technological innovation of *social networks* such as Facebook or Twitter and the norm *share yourself* (opinions, activities, etc.) in opposition to the norm *protect your privacy*. Prior to social networking, individuals willing to share their lives with a wider public audience could not live in accordance to their norm. In contrast, privacy-loving individuals were able to conceal most of their information. *Protect your privacy* was the prevalent norm in many countries. When internet services such as Facebook or Twitter entered the market, some individuals could start living according to their norm, *share yourself*. The innovation has

[3] More than 85% of these expenditures are spent on private transportation.
[4] Exceptions are things like solar panels for the accommodation category or the attendance of a pro-environment concert.

caused a complete reversal of the social norm. The second example is the innovation of *ecological food* and the norm of *sustainable and healthy consumption*. Today, almost all large supermarket chains include ecological food on their shelves, many being branded directly by the supermarkets themselves. With this innovation, people concerned with sustainability, health and also with the conditions of livestock breeding can live in accordance to this norm, and have become a large minority.

To make our argument precise, in the remainder of the paper, we proceed as follows. In Section 2, we introduce the model. Assumptions and notation are presented in 2.1. In 2.2, we derive the market equilibrium for a given share of norm-adopters and a given number of firms operating on innovative and traditional markets and then deduce the equilibrium number of firms supplying the innovative market. We then turn to studying the dynamics of norm adoption in 2.3. Results are summarized in Section 3. Policy implications are discussed in Section 4 and Section 5 concludes.

2.3. The Model

We consider a market where demand is characterized by a large number of consumers, who differ only with respect to their having adopted a particular consumption-related norm. The commodity traded on the market may occur in two specifications, one in compliance with the norm and one in violation thereof. We base our argument on a specific example, the market for automobiles and the norm of sustainable transportation, with electric cars as the norm-compliant variant and gasoline cars as the norm-violating variant. However, as we have already argued in the introduction, the argument extends to other examples as well.

To make identification of the two consumer groups easy, we call those consumers who have adopted the norm-*adopters* and those who did not, *hedonists*. $t \in \{a, h\}$ identifies the type of consumers in the natural way, while $v \in \{e, g\}$ identifies the variant of the norm-compliant (electric-powered) and, respectively, the norm-violating (gasoline-powered) variant of the commodity automobiles. For simplicity, both variants of the commodity are imperfect substitutes for each other and the slopes of demand curves as well as substitutability are assumed to be independent of the type of the consumer. With the simplification of linearity, and p^e and p^g denoting the prices of electric and gasoline cars, respectively, demand per consumer can be written as

$$x_t^v\left(p^e, p^g\right) = \chi_t^v - \kappa p^v + \lambda p^{\neg v} \text{ with } v \neq \neg v \in \{e, g\},\ \chi_t^v > 0 \text{ and } \kappa > \lambda > 0, \qquad (2.1)$$

for those price combinations which induce strictly positive quantities. For simplicity, we concentrate on these combinations and leave other cases to further research:

Assumption 1 $\min\left(x_a^e\left(p^e, p^g\right), x_a^g\left(p^e, p^g\right), x_h^e\left(p^e, p^g\right), x_h^g\left(p^e, p^g\right)\right) > 0$.

We refer to χ_t^v as the zero-price consumption of variant v by type t. To reflect that electric cars comply with the norm of sustainable transportation to a larger degree than gasoline cars, we state the following,

Assumption 2 If prices of the two variants of the commodity are identical ($p^e = p^g$), then the difference between consumption of the norm-compliant variant and of the norm-violating variant will be larger for the norm-adopters than for the hedonists: $x_a^e(\tilde{p},\tilde{p}) - x_a^g(\tilde{p},\tilde{p}) > x_h^e(\tilde{p},\tilde{p}) - x_h^g(\tilde{p},\tilde{p})$.

Corollary $\chi_a^e - \chi_a^g > \chi_h^e - \chi_h^g$.

We will later make use of the *effect of norm adoption on individual demand* for electric cars and for gasoline cars, $\Delta^e \equiv \chi_a^e - \chi_h^e$ and $\Delta^g \equiv \chi_a^g - \chi_h^g$, respectively, where the former is obviously larger than the latter due to the Corollary.

If we normalize the number of consumers to unity and write q as the proportion of consumers who have adopted the norm, market demands for the two product variants is:

$$\begin{aligned} X^e &= qx_a^e + (1-q)x_h^e = q\chi_a^e + (1-q)\chi_h^e - \kappa p^e + \lambda p^g \\ X^g &= qx_a^g + (1-q)x_h^g = q\chi_a^g + (1-q)\chi_h^g - \kappa p^g + \lambda p^e \end{aligned} \tag{2.2}$$

or equivalently, the system of inverse demand functions:

$$\begin{aligned} p^e &= \frac{1}{\kappa^2 - \lambda^2}\left(\left(q\chi_a^e + (1-q)\chi_h^e\right)\kappa + \left(q\chi_a^g + (1-q)\chi_h^g\right)\lambda - \kappa X^e - \lambda X^g\right) \\ p^g &= \frac{1}{\kappa^2 - \lambda^2}\left(\left(q\chi_a^g + (1-q)\chi_h^g\right)\kappa + \left(q\chi_a^e + (1-q)\chi_h^e\right)\lambda - \kappa X^g - \lambda X^e\right) \end{aligned} \tag{2.3}$$

On the supply side, we assume myopic profit maximization[5] on a simple Cournot oligopoly market for both variants of the commodity with constant marginal production costs of c^g and c^e for gasoline-powered and electric cars, respectively. We assume that the number of suppliers on the market for gasoline cars is given exogenously by n. The number m of suppliers on the market for electric cars is given by the maximum number of producers who can profitably produce for both markets when adding the second production line, entailing a fixed cost of k. Note that the oligopoly market may turn into a monopoly market. For consistency with the simplifications on the demand side, we exclude by assumption the absence of electric car producers.

We assume that markets find their equilibrium fast enough to neglect the specific dynamics when investigating the norm dynamics. In other words, we make use of the method of adiabatic

[5] We believe that profits, especially in large incorporations, are the main concerns of decision makers.

elimination[6] which allows us to include markets into the norm dynamics only by their equilibria, which may, of course, depend on the current level of norm adoption.

Finally, we assume that the dynamics of norm adoption and norm abandonment is a Markov process driven by randomly assigned moments in which each individual may adopt or abandon the norm. Whether it does may depend on the current state of the society with respect to norm adoption and norm-related market behavior. The dynamics of the proportion of individuals having adopted the norm, q, is thus given by

$$\dot{q} = (1-q)\pi_{h\to a} - q\pi_{a\to h} \tag{2.4}$$

where the transition rates $\pi_{h\to a}$ and $\pi_{a\to h}$ are the expected number of adoptions and abandonments of the norm per individual and per time unit.[7] This approximate equation of motion is standard in population dynamics[8] and is highly intuitive. The change in the share is simply the difference in the inflow and outflow. The inflow (outflow) is the product of the share of hedonists (norm-adopters) and the rate of transition from hedonists to adopters (adopters to hedonists).

In order to clearly identify the effect of the market innovation on the norm dynamics, we assume that norms may not be inferred from consumption behaviour and is not observable when no product variant compliant with the norm exists. The transition rates are then independent of the current proportion of norm adoption in society and any parameters relating to the (non-existent) market for the norm compliant variant of the commodity:

$$\pi^o_{a\to h} = \sigma_h \text{ and } \pi^o_{h\to a} = \sigma_a \text{, where } \sigma_h > 0 \text{ and } \sigma_a > 0 \text{ are constants.} \tag{2.5}$$

If the norm-compliant variant of the product enters the market, it will have two effects on the transition rates, a cognitive dissonance effect and a conformity bias effect. The former is caused by the possibility to behave according to the norm. It makes adopting the norm easier and being a norm adopter less repelling. We capture this idea in the formal presentation of the dynamics by increasing the norm adoption rate by a factor $(1+CB)$ and lowering the rate by which norm holders abandon it by a factor $(1-CB)$, where CB is the *reduction* in cognitive dissonances from having the norm but not complying with it. We assume $CB<1$ to ensure that the transition rates remain positive.

The conformity bias has a similar effect on norm adoption and norm abandonment. Once the norm-compliant variant of the product enters the market, individual consumers may observe whether their consumption conforms to the majority of consumers. Acting against the majority

---

[6] The method was introduced under this label by Haken (1977) for the synergetic approach of aggregation of dynamics of micro-data to the dynamics of macro-data. It has been introduced to economics e.g. by Weidlich and Haag (1983). The basic idea of the method may, however, already be found in Samuelson's "*Foundations*" (1947).

[7] Strictly speaking, the transition rates are the limits of the expected number of transitions per second, when we consider ever shorter time intervals (similar to the speed of a car being measured in miles per hour, but measured for a specific point in time, not for an entire hour).

[8] For example, see Weidlich and Haag (1983).

implies dissonances, which will be larger when the majority is larger. An individual is more likely to adopt the norm if norm-compliant behavior reflects the consumption pattern of the majority, i.e. if the ratio of electric cars to gasoline cars exceeds unity, then the transition rate towards norm adoption should increase relative to the pre-innovation level. If the opposite is true with respect to $\frac{X^e}{X^g}$, then the abandonment should be facilitated.[9] If $\alpha \in (0,1)$ measures the relative weight on the conformity bias, the post-innovation rates of transition can be written as follows:

$$\pi_{h\to a} = \sigma_a \left\lfloor \alpha(1+CB)+(1-\alpha)\frac{X^e}{X^g} \right\rfloor \text{ and } \pi_{a\to h} = \sigma_h \left\lfloor \alpha(1-CB)+(1-\alpha)\frac{X^g}{X^e} \right\rfloor. \tag{2.6}$$

Thus, the dynamics of the proportion of norm-adopters becomes:

$$\dot{q} = \underbrace{\alpha\big((1-q)\sigma_a - q\sigma_h\big)}_{\text{pre-innovation dynamics (linear)}} + \underbrace{\alpha CB\big((1-q)\sigma_a + q\sigma_h\big)}_{\text{cognitive bias (linear)}} + \underbrace{(1-\alpha)\left((1-q)\sigma_a\frac{X^e}{X^g} - q\sigma_h\frac{X^g}{X^e}\right)}_{\text{conformity bias (non-linear)}}. \tag{2.7}$$

The market-norm dynamics described in equation (2.7) completes the model. The equilibria for the model will be discussed in the following sections.

2.4. Equilibria

2.4.1. Market equilibrium

To find the equilibria of the norm-cum-market system described in the previous section, we first determine the market equilibrium and then turn to the dynamic part (Section 2.4.2).

As oligopolists, each producer $i \in \{1,2,\ldots,n\}$ maximizes $\max\{\hat{\Pi}_i, \tilde{\Pi}_i\}$, with $\hat{\Pi}_i = p^g \hat{x}_i^g - c^g \hat{x}_i^g$ and $\tilde{\Pi}_i = p^g \tilde{x}_i^g + p^e \tilde{x}_i^e - c^g \tilde{x}_i^g - c^e \tilde{x}_i^e - k$ over his production quantities $\hat{x}_i^g$, $\tilde{x}_i^g$ and $\tilde{x}_i^e$.

Proposition 2-1 For each share of norm-adopters $q \in [0,1]$ and each number $m \in \{0,\ldots,n\}$ of firms producing the innovative product, there is a unique equilibrium in the Cournot oligopoly game.

The proof follows Okuguchi and Szidarovszky (1990) and is given in Appendix A, as are all other proofs for this paper.

Taking the derivatives of $\tilde{\Pi}_i$ for *m* producers of both variants with respect to $\tilde{x}_i^g$ and $\tilde{x}_i^e$ yields two first order conditions which entail

$$\tilde{x}_i^g = (p^g - c^g)\kappa - (p^e - c^e)\lambda \text{ and } \tilde{x}_i^e = (p^e - c^e)\kappa - (p^g - c^g)\lambda. \tag{2.8}$$

[9] We neglect the possibility of having a conformity bias that affects consumption directly. This allows us to concentrate on the effects of the conformity bias on norm adoption and abandonment. We conjecture that this has no qualitative effects because the conformity bias affecting consumption directly should only reinforce the effects of the norm-related conformity bias.

Similarly, the derivative of $\hat{\Pi}_i$ for the $n-m$ producers of gasoline cars only with respect to $\hat{x}_i^g$ yields a first order condition which simplifies to

$$\hat{x}_i^g = \left(p^g - c^g\right)\left(\kappa^2 - \lambda^2\right)/\kappa . \tag{2.9}$$

Summing up all x_i^g and all x_i^e yields

$$\begin{aligned} X^e &= \sum_{i=1}^{m} \tilde{x}_i^e = m\left(\left(p^e - c^e\right)\kappa - \left(p^g - c^g\right)\lambda\right) \\ X^g &= \sum_{i=1}^{m} \tilde{x}_i^g + \sum_{i=1}^{n-m} \hat{x}_i^g = \frac{n\kappa^2 - (n-m)\lambda^2}{\kappa}\left(p^g - c^g\right) - m\left(p^e - c^e\right)\lambda \end{aligned} \tag{2.10}$$

Inserting p^e and p^g from equation (2.3) and solving for X^e and X^g gives the market equilibrium quantities

$$\begin{aligned} X^{e*} &= \frac{m}{m+1}\left(q\chi_a^e + (1-q)\chi_h^e - \kappa c^e + \lambda c^g\right) \\ X^{g*} &= \frac{n}{n+1}\left(q\chi_a^g + (1-q)\chi_h^g - \kappa c^g + \lambda c^e\right) + \frac{n-m}{(m+1)(n+1)}\left(q\chi_a^e + (1-q)\chi_h^e - \kappa c^e + \lambda c^g\right)\frac{\lambda}{\kappa} \end{aligned} \tag{2.11}$$

As it is obvious from equations (2.8) and (2.9), the equilibrium is symmetric in the sense that each firm of the same type (only conventional cars or both variants of cars) produces the same quantities. Indeed from Proposition 1 we know that this equilibrium is unique.

The market entry equilibrium in terms of the equilibrium number of firms operating in both markets is given by the condition of equal payoffs. Due to indivisibility, the equilibrium number of firms active also on the market for e-mobility, m^{eq}, corresponds to the integer part of m^* solving $\tilde{\Pi}_i = \Pi_i$ with $\tilde{x}_i^g, \tilde{x}_i^e, \hat{x}_i^g$ given by (2.8) and (2.9) and p^e and p^g by inserting X^{e*}, X^{g*} from (2.11) into (2.3). m^{eq} is thus given by:

$$m^{eq} = \min\left\{n, \max\left\{0, \text{integerpart}\left(m^*\right)\right\}\right\} \text{ where } m^* = \frac{q\chi_a^e + (1-q)\chi_h^e - \kappa c^e + \lambda c^g}{\sqrt{k\kappa}} - 1 \tag{2.12}$$

Note that the condition on m^{eq} to be of integer value will cause discontinuity in equilibrium prices and quantities at levels of q that induce a change in the value of m^{eq}. The number of firms serving both markets in equilibrium is increasing in the weighted willingness to pay for e-mobility and in the weighted cost differential between conventional cars and electric cars. The number of firms is decreasing in fixed costs k. Notably, the equilibrium number of firms producing both products is independent of the total number of firms n. We further note the following:

Lemma 2-1 The number of firms m^* is monotonically increasing in the share of norm-adopters if and only if $\chi_a^e > \chi_h^e$, i.e. if and only if the effect of the norm adoption on individual demand for electric cars is positive $\left(\Delta^e > 0\right)$.

Hence, if the sustainable-transportation norm is accompanied with a reduced overall demand for individual mobility, then an increasing share of norm-adopters may induce a larger number of producers of electric cars. This is true only if the reduction in the demand for transportation exclusively affects the demand for gasoline cars, which has to be partially substituted by an increased demand for electric cars. Lemma 2 will be helpful in Section 2.4.2.2 when we study the impact of the discontinuity of m^{eq} on the number of stable equilibria.

Having derived the number of firms serving both markets, we can now determine the quantities emerging if the expansion of firms on the e-mobility market is endogenous as $\hat{X}^e = X^{e*}\big|_{m=m^{eq}}$ and $\hat{X}^g = X^{g*}\big|_{m=m^{eq}}$. For expositional simplicity, we will heavily make use of the continuous version of m for the moment:

$$\begin{aligned}
\tilde{X}^e &= X^{e*}\big|_{m=m^*} = \left((1-q)\chi_h^e + q\chi_a^e - \kappa c^e + \lambda c^g\right) - \sqrt{k}\sqrt{\kappa} = \theta^e + \Delta^e q \\
\tilde{X}^g &= X^{g*}\big|_{m=m^*} \\
&= \frac{n}{n+1}\left((1-q)\chi_h^g + q\chi_a^g - \kappa c^g + \lambda c^e\right) - \frac{1}{n+1}\left((1-q)\chi_h^e + q\chi_a^e - \kappa c^e + \lambda c^g\right)\frac{\lambda}{\kappa} + \frac{\lambda}{\kappa}\sqrt{k}\sqrt{\kappa} \\
&= \frac{n}{n+1}\left(\theta^g + \Delta^g q\right) - \frac{1}{n+1}\left(\theta^e + \Delta^e q\right)\frac{\lambda}{\kappa}
\end{aligned} \tag{2.13}$$

where the tilde denotes the simplification based on the continuous version of m and the two terms

$$\theta^e = \chi_h^e - \kappa c^e + \lambda c^g - \sqrt{k}\sqrt{\kappa} > 0 \text{ and } \theta^g = \chi_h^g - \kappa c^g + \lambda c^e + \sqrt{k}\sqrt{\kappa}\frac{\lambda}{\kappa} > 0 \tag{2.14}$$

facilitate notation in the remainder of the paper. Before we turn to the analysis of the norm dynamics, we briefly study the total demand for private transportation:

$$\begin{aligned}
\tilde{X}^g + \tilde{X}^e &= \frac{n}{n+1}\left((1-q)\chi_h^g + q\chi_a^g + (1-q)\chi_h^e + q\chi_a^e - (\kappa-\lambda)\left(c^e + c^g\right)\right) \\
&\quad + \frac{1}{n+1}\left((1-q)\chi_h^e + q\chi_a^e - \kappa c^e + \lambda c^g\right)\left(1 - \frac{\lambda}{\kappa}\right) - \sqrt{k}\sqrt{\kappa}\left(1 - \frac{\lambda}{\kappa}\right) \\
&= \frac{n}{n+1}\left(\theta^e + \theta^g + \left(\Delta^e + \Delta^g\right)q\right) + \frac{1}{n+1}\left(\theta^e + \Delta^e q\right)\left(1 - \frac{\lambda}{\kappa}\right)
\end{aligned} \tag{2.15}$$

Total demand for individual transportation is a linear function in the share of norm-adopters. Neglecting a factor of proportionality close to 1, it increases (decreases) if the effect of norm adoption on the individual demand for electric cars (Δ^e) is larger (smaller) than the opposite effect on the individual demand for conventional cars ($-\Delta^g$).[10] The precise condition is:

$$\frac{\partial\left(\tilde{X}^e+\tilde{X}^g\right)}{\partial q}\gtreqless 0 \Leftrightarrow \Delta^e \gtreqless \frac{n}{n+1-\frac{\lambda}{\kappa}}\left(-\Delta^g\right). \tag{2.16}$$

2.4.2. Norm equilibrium

We now turn to the evolution of the share q in the population carrying a norm to consume in a sustainable way. As our model is fully specified, we can refine our research question concerning the impact of an innovation of a relative norm-compliant product variant on the evolution of a social norm shaping the preference for the good considered. We will address two questions in detail. First, what is the impact of an innovation that differs with respect to the level of norm compliance on the dissemination of a norm. Second, what is the effect of the market dynamics on the evolution of the norm with respect to the existence and stability of equilibria.

In the pre-innovation stage where transition rates are given by the constants defined in equation (2.5), the dynamics of equation (2.4) has an easy-to-calculate stable and unique equilibrium at

$$q^o=\sigma_a/\left(\sigma_h+\sigma_a\right). \tag{2.17}$$

When the innovation enters the market, transition rates now change depending on the equilibrium quantities of the different product variants and as given in equation (2.6). In the following paragraphs, we analyse the effects of three phenomena with respect to the two aforementioned questions. We first study the interplay of the cognitive dissonance bias and the conformity bias, and then turn to the discontinuity resulting from the fact that firms interact in an oligopoly.

2.4.2.1. Cognitive Bias and Conformity Bias

In order to understand the interplay of cognitive dissonance bias and conformity bias, we neglect the requirement that the number of firms supplying the norm-compliant variant of the product is an integer, and base our argument on the continuous version of the equilibrium number of such firms as defined by m^* in equation (2.12). Obviously, this requires assuming (for the moment) that the demand for electric vehicles by hedonists is large enough to keep X^e as defined by equation (2.10) strictly positive. In order to clearly differentiate between the continuous-m^* version of the model from the version with the discrete m^{eq}, we write $\dot{\mathring{q}}$ instead of $\dot{q}$ whenever

---

[10] Note that $\frac{n}{n+1-\frac{\lambda}{\kappa}}\approx 1$ for sufficiently large *n* and if the cross price "elasticity" is sufficiently close to the direct price "elasticity", i.e. if the two types of goods are very close substitutes.

we use $\tilde{X}^e$ and $\tilde{X}^g$ instead of $\hat{X}^e$ and $\hat{X}^g$ in equation (2.7). To guarantee differentiability of $\dot{\tilde{q}}$ we will further assume that $m^* \in [1,n]$.

This translates into a pair of inequalities:

$m^* \in [1,n] \Leftrightarrow \frac{1}{\sqrt{k\kappa}}\left(\theta^e + q\Delta^e\right) \in [1,n] \Leftrightarrow \frac{\theta^e}{\sqrt{k\kappa}} \in [1,n] \wedge \frac{1}{\sqrt{k\kappa}}\left(\theta^e + \Delta^e\right) \in [1,n]$, or equivalently,

$$\sqrt{k\kappa} \le \theta^e \le n\sqrt{k\kappa} \quad \wedge \quad \sqrt{k\kappa} - \theta^e \le \Delta^e \le n\sqrt{k\kappa} - \theta^e . \tag{2.18}$$

We will neglect this condition in the following paragraphs since its inclusion, while straightforward, would unnecessarily complicate the notation. So far, the reader should keep in mind that the number of firms n should be sufficiently high and fixed set up cost k should be sufficiently small. We will return to this issue in Section 2.6.

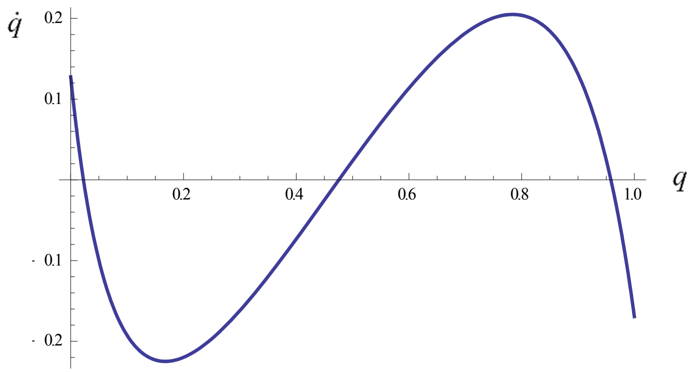


Figure 2-1: market-norm dynamics.

$\theta^e = 0.1,\ \theta^g = 1,\ \Delta^e = 0.75,\ \Delta^g = -0.65,\ \text{n} = 4,\ \lambda/\kappa = 0.8,\ \sigma_a = \sigma_h = 1,\ \alpha = 0$

Neglecting the conformity bias ($\alpha = 1$), inspection of equation (2.7) shows that cognitive bias shifts the norm dynamics upwards and turns it counterclockwise, increasing the equilibrium level of norm adoption. The conformity bias changes the motion of the norm adoption proportion described in equation (2.7) from a linear function to an s-shaped function with at most one increasing branch in the middle (see Figure 2-1):

Lemma 2-2 Assume that χ_h^e and χ_a^g are large enough to guarantee that $\tilde{X}^e$ and $\tilde{X}^g$ as defined by equation (13) are strictly positive for all $q \in [0,1]$. Then:

1. $\dot{\tilde{q}}\big|_{q=0} > \alpha\left(1 + CB\right)\sigma_a > 0$ and $\dot{\tilde{q}}\big|_{q=1} < \alpha\left(CB - 1\right)\sigma_h < 0$;
2. Any value of $\dot{\tilde{q}}$ is reached for at most three different $q \in [0,1]$; and
3. $\Delta^e \le \frac{\theta^e}{\theta^g}\Delta^g$ implies $\frac{d\left(\tilde{X}^e/\tilde{X}^g\right)}{dq} \le 0$, which in turn implies $\frac{d\dot{\tilde{q}}}{dq} < 0$.

The intuition behind claims 1 and 2 is simple. Claim 1 is obvious when X^e and X^g are strictly positive. Claim 2 follows from the fact that X^e and X^g are linear in q and thus solving equation (2.7) for q for any given value of $\dot{\tilde{q}}$ is tantamount to solving a polynomial of degree three. The first implication of Claim 3 follows from the fact that the denominator of the derivative $\frac{d\left(\tilde{X}^e/\tilde{X}^g\right)}{dq}$ is strictly positive and the numerator is given by:

$$\begin{aligned}
&\frac{d\tilde{X}^e}{dq}\tilde{X}^g-\frac{d\tilde{X}^g}{dq}\tilde{X}^e\\
&=\left(\Delta^e\left(\frac{n}{n+1}\left(\theta^g+\Delta^g q\right)-\frac{1}{n+1}\left(\theta^e+\Delta^e q\right)\right)\frac{\lambda}{\kappa}\right)-\left(\frac{n}{n+1}\Delta^g-\frac{1}{n+1}\Delta^e\frac{\lambda}{\kappa}\right)\left(\theta^e+\Delta^e q\right)\\
&=\frac{n}{n+1}\left[\theta^g\Delta^e-\theta^e\Delta^g\right]
\end{aligned}\tag{2.19}$$

The second implication of Claim 3 follows from the observation that all three terms summed up in

$$\begin{aligned}
\frac{d\dot{\tilde{q}}}{dq}&=-\alpha\left(\sigma_a\left(1+CB\right)+\sigma_h\left(1-CB\right)\right)-\left(1-\alpha\right)\left(\sigma_a\frac{\tilde{X}^e}{\tilde{X}^g}+\sigma_h\frac{\tilde{X}^g}{\tilde{X}^e}\right)\\
&+\left(1-\alpha\right)\left(\frac{\left(1-q\right)\sigma_a}{\left(\tilde{X}^g\right)^2}+\frac{q\sigma_h}{\left(\tilde{X}^e\right)^2}\right)\left(\frac{d\tilde{X}^e}{dq}\tilde{X}^g-\frac{d\tilde{X}^g}{dq}\tilde{X}^e\right)
\end{aligned}\tag{2.20}$$

are negative if $\frac{d\left(\tilde{X}^e/\tilde{X}^g\right)}{dq}\le 0$.

As a consequence of Claim 1 of Lemma 3, $\dot{\tilde{q}}$ must have at least one branch declining in q. Claim 2 of the lemma then implies that there is at most one increasing branch. Such an increasing branch is a necessary condition for multiple inner equilibria of the market-norm dynamics. Hence, a direct consequence of Claim 3 is the following:

Corollary If the market-norm dynamics has multiple (two) stable inner equilibria then $\tilde{X}^e/\tilde{X}^g$ increases strongly in q for all $q\in\left[0,1\right]$, i.e. $\Delta^e>\Delta^g\,\theta^e/\theta^g$.

Figure 2-1 illustrates the possibility of multiple equilibria. In the following section, we look at the conditions and thereby at the parameter set that gives rise to this phenomenon. With the assumption of strictly positive demand, the roots of (2.7) are equivalent to the roots of (2.21).

$$\begin{aligned}
&\dot{q}\equiv X^eX^g\dot{\tilde{q}}=\alpha\tilde{X}^e\tilde{X}^g\cdot\\
&\left(\left(1+CB\right)\sigma_a-q\left(\left(1+CB\right)\sigma_a+\left(1-CB\right)\sigma_h\right)\right)+\left(1-\alpha\right)\left(\left(1-q\right)\sigma_a\left(\tilde{X}^e\right)^2-q\sigma_h\left(\tilde{X}^g\right)^2\right)
\end{aligned}\tag{2.21}$$

The dynamics given by (2.21) is a polynomial of degree 3 and has two stable inner equilibria in the unit interval if and only if it has two extreme points with a negative functional value at the minimum and a positive functional value at the maximum. Note that if there are two extreme points $q^{Low}<q^{High}$, then $\dot{q}\left(q^{High}\right)>0$ implies $q^{High}<1$ and $\dot{q}\left(q^{Low}\right)<0$ implies $q^{Low}>0$ by inspection of (2.7), given strictly positive demand. Given $\dot{q}\left(q^{High}\right)>0$ and $\dot{q}\left(q^{Low}\right)<0$, the fact that $\dot{q}(0)>0, \dot{q}(1)<0$ implies that q^{Low} is the minimum and q^{High} is the maximum.

Hence, only the two conditions with respect to the existence of two extrema and the sign condition at the extrema points remain. Since demand is linear in the share of norm-adopters, the conditions of positive demand amount to: $0<\theta^e<\frac{n}{\lambda}\theta^g$ and $-\theta^e<\Delta^e<-\theta^e+\frac{n}{\lambda}\left(\Delta^g+\theta^g\right)$.

The binding constraints are therefore given by: $\dot{q}\left(q^{Low}\right)<0$, $\dot{q}\left(q^{High}\right)>0$,$0<\theta^e<\frac{n}{\lambda}\theta^g$ and $-\theta^e<\Delta^e<-\theta^e+\frac{n}{\lambda}\left(\Delta^g+\theta^g\right)$. It turns out that only $\dot{q}\left(q^{Low}\right)<0$, $\dot{q}\left(q^{High}\right)>0$ and $\Delta^e<-\theta^e+\frac{n}{\lambda}\left(\Delta^g+\theta^g\right)$ depend on Δ^e and Δ^g.

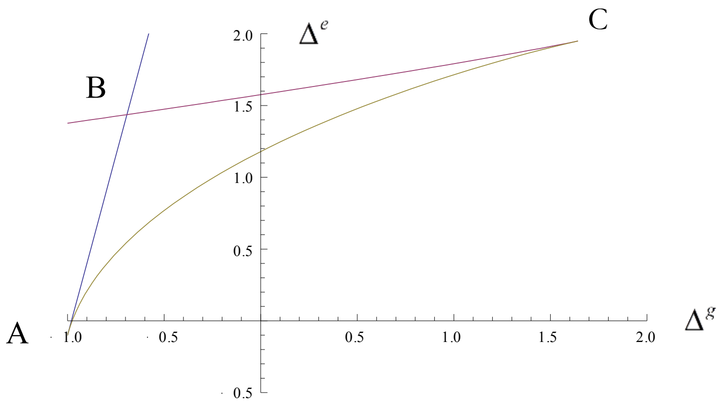


Figure 2-2: Range of multiple equilibria: blue line: $\tilde{X}^g(1)=0$, $\tilde{X}^g(1)>0$ to the right of the blue line; red line: $\Delta^{e,Min}(\Delta g)$ upper bound of Δ^e allowing for multiple equilibria; yellow line: $\Delta^{e,Max}(\Delta g)$ lower bound of Δ^e allowing for multiple equilibria. $\theta^e=0.1$, $\theta^g=1$, $\sigma_h/\sigma_a=1$, $n=4$, $\alpha=0$, $\lambda/\kappa=4/5$.

Therefore, if we study the parameter region of Δ^e and Δ^g such that multiple equilibria exist, only these three conditions are relevant, given that the values for the other parameters satisfy the remaining inequalities ($0<\theta^e<\frac{n}{\lambda}\theta^g$). Figure 2-2 gives an illustrative example.

The intuition behind having an upper and lower limit for Δ^e is simple. If Δ^e were too large, $X^e(q)$ increases too quickly relative to $X^g(q)$ that $\dot{\tilde{q}}(q)$ increases at $q=0$ or the minimum of $\dot{\tilde{q}}(q)$ is above the $\dot{\tilde{q}}=0$-axis. If Δ^e were too small, $X^g(q)$ declines rapidly relative to $X^e(q)$ that $\dot{\tilde{q}}(q)$ never increases or only has a minimum but no maximum, or has a maximum which remains below the $\dot{\tilde{q}}=0$-axis. In our application, a relatively large Δ^e implies that norm adoption has such a strong effect on the market equilibrium amount of norm compliant consumption that the growth in this consumption (possibly at the cost of norm violating consumption) reinforces the norm very quickly. This happens at such a pace that norm adoption is always self-reinforcing until the number of individuals not having adopted the norm becomes very small. If, on the other hand, Δ^e is very small, then norm adoption has too little of an effect on norm compliant consumption to become self-reinforcing.

In the next section, we derive sufficient conditions for multiple equilibria to exist. If we look at Figure 2-2, it appears that these three conditions define a triangular region. In what follows, we will derive the vertices of that region and reformulate the two differential equations $\dot{q}\left(q^{Low}\right)<0$, $\dot{q}\left(q^{High}\right)>0$ as differential equation for $\Delta^e\left(\Delta^g\right)$.

Given strictly positive demand (2.7) gives rise to a fixed point equation:

$$\dot{q}=\alpha\left((1-q)\sigma_a-q\sigma_h\right)+\alpha CB\left((1-q)\sigma_a+q\sigma_h\right)+(1-\alpha)\left((1-q)\sigma_a\frac{X^e}{X^g}-q\sigma_h\frac{X^g}{X^e}\right)=0$$

$$\overset{\alpha\neq 1}{\Leftrightarrow}$$

$$(1-q)\sigma_a\frac{X^e}{X^g}-q\sigma_h\frac{X^g}{X^e}=\frac{\alpha}{1-\alpha}\sigma_a(1+CB)-q\frac{\alpha}{1-\alpha}\left(\sigma_a(1+CB)+\sigma_h(1-CB)\right)\equiv\gamma+\beta q$$

$$\Leftrightarrow \tag{2.22}$$

$$\left(\frac{X^e}{X^g}\right)^2=\frac{q}{(1-q)}\underbrace{\sigma}_{\equiv\sigma_h/\sigma_a}+\frac{(\gamma+\beta q)}{(1-q)}\frac{X^e}{X^g}$$

$$\underset{z(q)\equiv\frac{X^e}{X^g}}{\Leftrightarrow}$$

$$\left(z(q)\right)^2=\frac{q}{(1-q)}\sigma+\frac{(\gamma+\beta q)}{(1-q)}z(q)\Rightarrow$$

At $q=q^{extr.}$ such that $\dot{q}'\left(q^{extr.}\right)=0$ this gives a fixed point equation in Δ^e,Δ^g :

$$\dot{q}\left(q^{extr.}\left(\Delta^e,\Delta^g\right)\right)=0$$

We take the total derivative with respect to Δ^e,Δ^g and apply the envelope theorem[11].

$$\left[2z\left(q^{extr.}\right)\left.\frac{\partial z}{\partial\Delta^e}\right|_{q=q^{extr.}}-\frac{(\gamma+\beta q)}{(1-q)}\left.\frac{\partial z}{\partial\Delta^e}\right|_{q=q^{extr.}}\right]d\Delta^e+\left[2z\left(q^{extr.}\right)\left.\frac{\partial z}{\partial\Delta^g}\right|_{q=q^{extr.}}-\frac{(\gamma+\beta q)}{(1-q)}\left.\frac{\partial z}{\partial\Delta^g}\right|_{q=q^{extr.}}\right]d\Delta^g=0$$

$$\Leftrightarrow$$

$$\frac{n+1}{n}\frac{X^g}{X^e}+\frac{\lambda}{n}=\frac{d\Delta^g}{d\Delta^e}\Leftrightarrow\frac{d\Delta^g}{d\Delta^e}=\frac{\Delta^g q^{extr.}+\theta^g}{\Delta^e q^{extr.}+\theta^e}>0$$

11 $\frac{\partial\dot{q}\left(q^{extr.}\left(\Delta^e,\Delta^g\right)\right)}{\partial\text{parameter}}=\left.\frac{\partial\dot{q}(q)}{\partial\text{parameter}}\right|_{q=q^{extr.}}$

Together with initial conditions: $\left(\Delta e,\Delta g\right)\left|\dot{q}\left(q^{Max}\left(\Delta e,\Delta g\right)\right)=0\right.$ and $\left(\Delta^e,\Delta^g\right)\left|\dot{q}\left(q^{Min}\left(\Delta^e,\Delta^g\right)\right)=0\right.$ the differential equation $\frac{d\Delta^g}{d\Delta^e}=\frac{\Delta^g q^{extr.}+\theta^g}{\Delta^e q^{extr.}+\theta^e}$ gives rise to two boundary functions: $\Delta^{e,Min}\left(\Delta^g\right),\Delta^{e,Max}\left(\Delta^g\right)$.

Definition All $\left(\Delta^e,\Delta^g\right)$ pairs that satisfy the following three conditions define the parameter region such that multiple equilibria exist: (1) $\Delta^e<-\theta^e+\frac{n}{\lambda}\left(\Delta^g+\theta^g\right)$, (2) $\Delta^e>\Delta^{e,Min}\left(\Delta^g\right)$, (3) $\Delta^e<\Delta^{e,Max}\left(\Delta^g\right)$. We will refer to this set as the *multiple equilibria set* (MES).

Before we continue, we will state some observations based on $\frac{d\Delta^e}{d\Delta^g}=\frac{1}{\frac{n+1}{n}\frac{X^g}{X^e}+\frac{\lambda}{n}}$ that will be helpful in the course of our argument:

(1) The slopes of $\Delta^{e,Min}\left(\Delta^g\right),\Delta^{e,Max}\left(\Delta^g\right)$ are positive and smaller than the slope of the third constraint $\Delta^e<-\theta^e+\frac{n}{\lambda}\left(\Delta^g+\theta^g\right)$.

(2) By corollary 4 $\frac{d\Delta^g}{d\Delta^e}=\frac{\Delta^g q+\theta^g}{\Delta^e q+\theta^e}$ is ceteris paribus decreasing in q

(3) At point A, the relevant constraints have the same slope.

We are able to determine the coordinates for points A and B (Figure 2-2) analytically. For better readability, Table 2-1 below presents the results for $\alpha=0$. Note that there exist multiple equilibria if and only if $\left(\Delta^e\right)^B>\left(\Delta^e\right)^A$. As mentioned before, the dynamics given by (2.7) consist of a linear and nonlinear term, the latter is weighted with $1-\alpha$. Intuitively, one would expect that α, the weight of the linear term, must be sufficiently small so that the nonlinear term dominates the dynamics and for some parameter constellations multiple equilibria might arise. It indeed turns out that there exists a unique threshold value for α, such that multiple equilibria are possible. Its derivation is deferred to Appendix A. The value and its properties are summarized in the next lemma.

Lemma 2-3 For $0<\theta^e<\frac{n}{\lambda}\theta^g$ there exists an unique

$$\alpha^{crit.}=\frac{(1-CB)(n+1)\theta^e+2\tau-\sqrt{\left((1-CB)(n+1)\theta^e\right)^2+4\tau^2\left(1-CB^2\right)}}{2\left(CB^2\tau+(n+1)(1-CB)\theta^e\right)}\text{, with } \tau\equiv n\theta^g-\lambda\theta^e$$

such that MES is non-empty if and only if $\alpha<\alpha^{crit.}$.

Furthermore, $\frac{\partial\alpha^{crit.}}{\partial\theta^g} = \frac{\partial\alpha^{crit.}}{\partial n} = -\frac{\theta^e}{\theta^g}\frac{\partial\alpha^{crit.}}{\partial\theta^e} = -\frac{n}{\theta^e}\frac{\partial\alpha^{crit.}}{\partial\lambda}\bigg|_{\substack{\theta^e,\theta^g \\ fixed}} > 0; \frac{\partial\alpha^{crit.}}{\partial\sigma_a} = \frac{\partial\alpha^{crit.}}{\partial\sigma_h} = 0; \frac{\partial\alpha^{crit.}}{\partial CB} > 0$

which implies: $\frac{\partial\alpha^{crit.}}{\partial\chi_h^e} < 0; \frac{\partial\alpha^{crit.}}{\partial\chi_h^g} > 0; \frac{\partial\alpha^{crit.}}{\partial c^e} > 0; \frac{\partial\alpha^{crit.}}{\partial c^g} < 0; \frac{\partial\alpha^{crit.}}{\partial k} > 0$.

In other words, as long as the weight for the non-linear term is sufficiently large, there will always be (Δ^e,Δ^g) pairs such that multiple equilibria exist. With respect to partial effects, Lemma 2-3 states that the required weight for the non-linear term of the dynamics $1-\alpha$ is increasing in maximum willingness to pay for electric cars by hedonists χ_h^e and in the marginal cost for gasoline cars c^g. The required weight decreases in the maximum willingness to pay for gasoline cars χ_h^g, the marginal cost for electric cars c^e and the fixed setup cost *k*. The effects with respect to parameters measuring the price sensitivity are ambiguous. The weight also decreases in the number of firms in the market and in CB measuring the reduction of cognitive dissonances from having adopted the norm but not complying with it.

Point	A	B	C
q	$q^{Max}=1$	$q^{Min}=\frac{4(n+1)^2(\theta^e)^2}{(n\theta^g-\lambda\theta^e)^2+4(n+1)^2(\theta^e)^2}$	$q^{ip}=\frac{(n\theta^g-\lambda\theta^e)}{3(n\Delta^g-\lambda\Delta^e)+4(n\theta^g-\lambda\theta^e)}$
$\begin{pmatrix}\Delta^g\\ \Delta^e\end{pmatrix}$	$\begin{pmatrix}-\theta^g\\ -\theta^e\end{pmatrix}$	$\begin{pmatrix}\frac{\lambda}{n}\left(\frac{\tau}{2(n+1)}\right)^2\frac{1}{\theta^e}-\theta^g\\ \left(\frac{\tau}{2(n+1)}\right)^2\frac{1}{\theta^e}-\theta^e\end{pmatrix}$	/

Table 2-1: Vertices of multiple equilibria set for $\alpha=0$.

The differential equations given by $\frac{d\Delta^g}{d\Delta^e}=\frac{\Delta^g q^{extr.}+\theta^g}{\Delta^e q^{extr.}+\theta^e}$ cannot be solved for analytically. In the following, we present our approximation strategy for $\alpha=0$, such that we can state explicit sufficient conditions for multiple equilibria to exist. Again, the general case can be found in the Appendix A. Note that the values for *q* that corresponds to (Δ^e,Δ^g) pairs that are elements of the graph of $\Delta^{e,Max}(\Delta^g)$ range from $q^C=\frac{(n\theta^g-\lambda\theta^e)}{3(n\Delta^g-\lambda\Delta^e)+4(n\theta^g-\lambda\theta^e)}$ to $q^A=1$. We can use the $(\Delta^g)^B$ as a lower bound for Δ^g and by that, can give a lower bound for q independent of Δ^e and Δ^g, i.e. $\underline{q}=\frac{4(n+1)^2\theta^e}{4(n+1)^2\theta^e+3\lambda(n\theta^g-\lambda\theta^e)}$. The system $\dot{q}(q^C)=0$, $\dot{q}'(q^C)=0$ can be solved for Δ^e and Δ^g as a function of *q*. If we plug in $\underline{q}$, we get as point *D* a (Δ^e,Δ^g) pair on the graph of $\Delta^{e,Max}\left(\Delta^g\right)$ that corresponds to a maximum for the dynamics in (7) that equals $\underline{q}$.

$$\begin{pmatrix}\Delta^e \\ \Delta^g\end{pmatrix}^D = \begin{pmatrix} \dfrac{4\sqrt{3\theta^e\lambda\tau^3} - \theta^e\left(4(n+1)^2\theta^e + 9\lambda\tau\right)}{4(n+1)^2\theta^e} \\ \dfrac{3\underline{q}\left(\Delta^e\right)^D - 2\tau + \sqrt{6\underline{q}\left(\Delta^e\right)^D\left(\left(\Delta^e\right)^D - \theta^e\right) - 9\underline{q}^2 + 6(1-q)\left(\Delta^e\right)^D\theta^e - (n+1)^2\left(\theta^e\right)^2 + \tau^2}}{3n\underline{q}} \end{pmatrix} \quad (2.23)$$

We approximate the upper and lower boundaries by linear functions intersecting point B and D, respectively. Our observation above, that the slope $\Delta^{e,Min}\left(\Delta^g\right)$ is decreasing in q gives us a lower bound for the slope by $\dfrac{\theta^e}{\theta^g}$. Figure 2-3 illustrates our approximation procedure. Note that under our approach, MES is not empty if and only if the area spanned by $X^g(1) > 0$ and the two approximating linear function is non-empty.

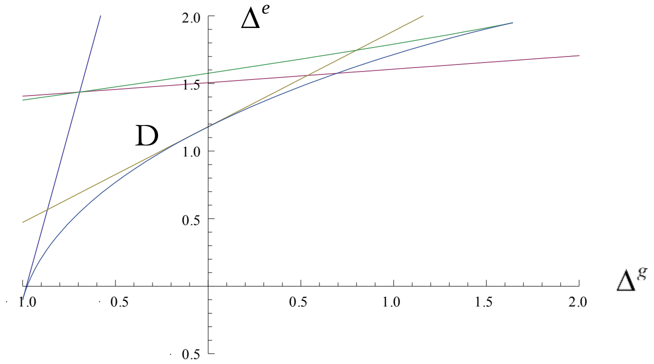


Figure 2-3: Approximation of MES. Red line: approximation of $\Delta^{e,Min}\left(\Delta^g\right)$; yellow line: approximation of $\Delta^{e,Max}\left(\Delta^g\right)$. $\theta^e = 0.1,\ \theta^g = 1,\ \sigma_h/\sigma_a = 1,\ n = 4,\ \alpha = 0,\ \lambda/\kappa = 4/5$.

Lemma 2-4 If $\alpha = 0$ and $0 < \theta^e < \dfrac{n}{\lambda}\theta^g$, $\dot{\tilde{q}} = 0$ has three solutions if (sufficient condition):

$$\left(\Delta^e,\Delta^g\right) \in \left\{ \begin{array}{c} \left(\Delta^e,\Delta^g\right) \middle| \Delta^e < -\theta^e + \dfrac{n}{\lambda}\left(\Delta^g + \theta^g\right), \Delta^e < \dfrac{\theta^e}{\theta^g}\Delta^g + \dfrac{\left(n\theta^g - \lambda\theta^e\right)^3}{4n(n+1)^2\theta^g\theta^e} \\ \Delta^e > \dfrac{\left(\Delta^e\right)^D\underline{q} - \theta^e}{\left(\Delta^g\right)^D\underline{q} - \theta^g}\Delta^g + \left(\Delta^e\right)^D - \dfrac{\left(\Delta^e\right)^D\underline{q} - \theta^e}{\left(\Delta^g\right)^D\underline{q} - \theta^g}\left(\Delta^g\right)^D \end{array} \right\}$$

The effects of market parameter variations on the location of $\dot{\tilde{q}}(q)$ and on the number of equilibria are best understood by observing that they only enter via $\tilde{X}^e$ and $\tilde{X}^g$ into equation (2.7). Since $\dfrac{d\dot{\tilde{q}}(q)}{d\tilde{X}^e} > 0 > \dfrac{d\dot{\tilde{q}}(q)}{d\tilde{X}^g}$, the derivatives are all straight forward. In particular, $\dfrac{d\dot{\tilde{q}}(q)}{d\Delta^e} > 0 > \dfrac{d\dot{\tilde{q}}(q)}{d\Delta^g}$, i.e. the effect of norm adoption on the demand for electric (gasoline) cars has a positive (negative) effect on the growth rate of the share of adopters.

Before studying the effect of the cognitive bias and the transition rates $\sigma_{a,h}$, it is worth mentioning that these parameters are not subject to policy measures. They reflect the dynamics of norm adoption before the innovation takes place. In particular, the cognitive dissonance from having adopted a norm that one cannot comply with is beyond the reach of political measures. Discussing these parameters is thus only relevant for understanding the context in which policy is formulated. Since the transition rates $\sigma_{a,h}$ occur in each and every term of the right-hand side of equation (2.7), it is only their ratio which is relevant. If σ_a/σ_h is small, there will be only a few norm-adopters in the equilibrium before the innovation takes place, in particular, because too much cognitive dissonance is implied by having the norm. After the innovation, small values of σ_a/σ_h imply that the range of Δ^e for which multiple equilibria occur shifts upwards and stretches along the Δ^e-axis.

If the cognitive bias is large, that is, if the innovation removes a lot of cognitive dissonance from norm-adopters, then the innovation tends to have a particularly positive effect on norm adoption. Starting from the pre-innovation equilibrium value of the rate of norm adoption, $q^o = \sigma_a/(\sigma_h + \sigma_a)$ exemplifies the effect of the size of *CB* and its interplay with the conformity bias on which most of our hitherto discussion was concentrated. The following lemma states the necessary and sufficient condition for a positive growth rate in norm adoption at the pre-innovation level.

Lemma 2-5

$$\dot{q}\left(q^o\right) > 0 \Leftrightarrow 2CB\frac{\alpha}{1-\alpha} + \left.\frac{\tilde{X}^e}{\tilde{X}^g}\right|_{q=q^o} - \left.\frac{\tilde{X}^g}{\tilde{X}^e}\right|_{q=q^o} > 0 \qquad (2.24)$$

which may be transformed to $$\Delta^e > \frac{\mu}{\mu+1}\frac{n}{n+1}\left(\frac{\theta^g}{q^o} + \Delta^g\right) + \frac{1}{\mu+1}\left(1-\mu\frac{\lambda/\kappa}{n+1}\right)\frac{\theta^e}{q^o}, \qquad (2.25)$$

where $$\mu = \left(-CB\frac{\alpha}{1-\alpha} + \sqrt{CB^2\frac{\alpha^2}{(1-\alpha)^2}+1}\right)\frac{\lambda/\kappa}{n+1}.$$

Equation (2.25) describes a straight and increasing line, above which $\dot{q}\left(q^o\right)$ is positive so that the innovation induces a growth of norm adoption, while below this line, norm adoption will decline when the innovation occurs. The straight line moves upward if *CB* or α increase.

If $\dot{q}\left(q^o\right) < 0$, then it implies that the positive cognitive bias is offset by a negative conformity bias with a sufficiently large weight α. Obviously, the conformity bias is negative only if at q^o, the market-equilibrium quantity of the norm-compliant variant of the good is less than the corresponding quantity of the norm-violating variant.

If the quantities of the two variants of the good are hardly affected by the number of norm-adopters or the quantity of the norm-compliant variant grows only slightly compared to the quantity of the norm-violating variant, i.e. if the effects of norm adoption on individual demand are small or not too much diverging, then $\dot{q} < 0$ may hold true for all $q \geq q^o$. However, if the effects of norm adoption are strong and induce rapid growth of $\frac{\tilde{X}^e}{\tilde{X}^g} - \frac{\tilde{X}^g}{\tilde{X}^e}$ in *q* (see (2.24)), then

$\dot{\tilde{q}}$ may become positive for some $q \in \left(q^o, 1\right)$ so that a (second) stable equilibrium with a large level of norm adoption is generated by the conformity bias. In the next section, we will enlighten the effects that the discontinuity of the number of firms adds to our discussion of the cognitive and conformity bias.

2.4.2.2. Discontinuity of Firm Number

We now drop the simplifying assumption of continuity of the equilibrium number of firms producing the norm-compliant variant of the product. We first study the effect of the discreteness of this number of firms on the pace at which norm adoption changes and then infer consequences for the number and location of equilibria with reference to the structure of the market of the innovative good.

A helpful first insight is the following:

Lemma 2-6 Except for the discontinuities, where $\dot{q}(q) = \dot{\tilde{q}}(q)$ holds true, we have:

1. $\dot{q}(q) < \dot{\tilde{q}}(q)$ and $\frac{d\dot{q}(q)}{dq} \lessgtr \frac{d\dot{\tilde{q}}(q)}{dq} \Leftrightarrow \Delta^e \gtrless 0$ for all q.

2. Let q_1 and q_2 be two instances of discontinuity of $\dot{q}$ with $q_2 > q_1$. Then:

a. $q_2 - q_1 = \eta \frac{\sqrt{k\kappa}}{\Delta^e}$ where $\eta \in \{1, 2, \ldots\}$

b. $\dot{q}(q_1) - \lim_{q \uparrow q_1}\left(\dot{q}(q)\right) > \dot{q}(q_2) - \lim_{q \uparrow q_2}\left(\dot{q}(q)\right) > 0$ if $\Delta^e > 0$ and

$\dot{q}(q_1) - \lim_{q \downarrow q_1}\left(\dot{q}(q)\right) < \dot{q}(q_2) - \lim_{q \downarrow q_2}\left(\dot{q}(q)\right) < 0$ if $\Delta^e < 0$.

Figure 2-4 visualizes the relationship between $\Delta^e > 0$ and $\dot{\tilde{q}}(q)$ reported in the lemma.

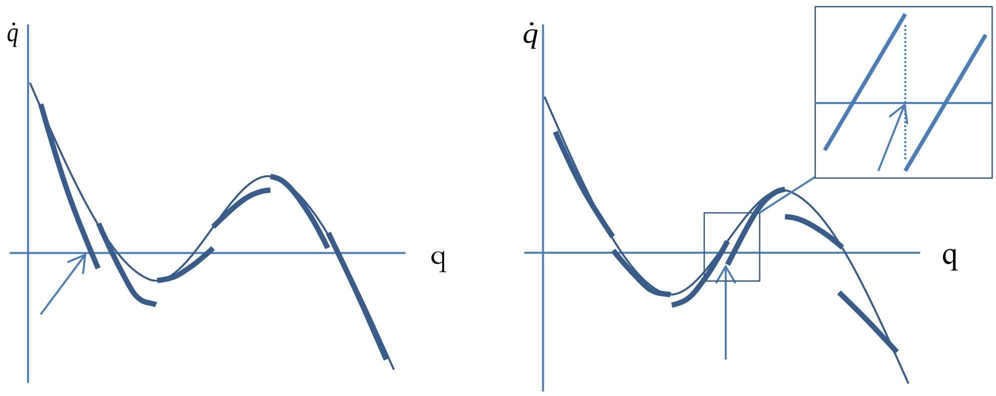


Figure 2-4: Effects of discontinuity on $\dot{q}(q)$. Left: $\Delta^e > 0$, right $\Delta^e < 0$: Additional stable equilibria marked by an arrow.

The discontinuities described in Lemma 2-6 may increase the number of instances at which the sign of $\dot{q}(q)$ changes from positive to negative as q increases, i.e. the number of stable equilibria. It does not reduce this number. The additional stable equilibria may not occur over the entire range of q, but only in those intervals, in which the "jumps" and the slope in the neighborhood of the discontinuities are in opposite directions. Only then may the discontinuities result in additional sign changes. We state the argument more precisely in the following:

Corollary Additional stable equilibria due to the discontinuities of $\dot{q}(q)$ occur if and only if the discontinuities entail additional sign changes of $\dot{q}(q)$. If $\Delta^e > 0$, every additional stable equilibrium is in one of the intervals in which $\dot{q}(q)$ is continuous and which has its lower bound in one of the decreasing branches of $\dot{\tilde{q}}(q)$. If $\Delta^e < 0$, almost all[12] additional stable equilibria occur at discontinuities which form the lower bound of a continuity interval of $\dot{q}(q)$ which is at least partly in the increasing branch of $\dot{\tilde{q}}(q)$.

We note that this corollary implies that with negative Δ^e and a monotonously decreasing function $\dot{\tilde{q}}(q)$, the discontinuity will never induce additional equilibria. The relevance of this insight becomes obvious if one remembers that with negative Δ^e, the existence of an increasing branch of $\dot{\tilde{q}}(q)$ is only possible if Δ^g is sufficiently smaller than Δ^e.

With more stable equilibria, temporary policies are more likely to induce a permanent shift in market structures or market outcomes, but as the larger number of stable equilibria become less distant, such permanent effects of temporary policies tend to be smaller. Much of the discussion in the following section on policy implications is based on this insight.

2.5. Policy implications

The policy implications of our model depend to some degree on the definition of policy goals. Within the realm of environmental policy in general and traffic-emissions policy in particular, policy goals may run the gamut from the dissemination of environment-friendly products over a reduction of particularly polluting products to straight emission reductions. Very often, environmental sustainability and emission reductions may be the final goal, but political activism often involves preliminary targets such as electric cars replacing gasoline cars. General adoption of environmental norms, such as the sustainable-transportation norm we have been using as a running example in our model, may also serve as one of the more immediate goals.

All these goals may be affected by innovation such as electric cars with similar consumption properties as gasoline cars. If the innovation is unrelated to a norm, or if the adoption and abandonment of the norm do not depend on the relative frequency of the consumption of the new, norm-compliant product variant, then there would be few arguments for government support of the new technology, except for the internalization of external effects. However, if the dissemination of the innovation is linked to a norm in the two ways we have described in our model, namely both higher valuation of the new product by norm bearers and the feedback of

[12] The only case in which an additional equilibrium may be in a continuity interval of $\dot{q}(q)$ occurs if $\dot{\tilde{q}}(q)$ has a minimum, this minimum is positive, a continuity interval of $\dot{q}(q)$ embraces this minimum, has an interior minimum which is negative and has positive limits at both bounds.

norm-compliant consumption on the dissemination of the norm, then the introduction of a norm-compliant innovation ceases to have unambiguous effects.

We have discussed the case that the conformity bias may be so strong that it can hinder the dissemination of innovation. In fact, as innovation allows for the observable choice between norm-compliant and norm-violating behavior, the innovation may reduce the number of norm-adopters if it enters the market in small numbers at the beginning, and thereby hinders its own further dissemination into the market. In these cases, it is particularly appropriate for political interference with market forces (and norm formation!). However, policy measures should be carefully chosen. It would be detrimental if policy aimed at (and succeeded in) increasing the influence of the normative sphere on the market by strengthening the conformity bias in society. Such policy measures would only reinforce the innovation-curbing effects of the conformity bias. However, policy should be willing to strongly support innovation in an early stage by improving market parameters in order to shift the market-norm system into the region of attraction of the high level of norm adoption. Only in the long run should such policies be replaced by supporting the conformity bias in order to further shift the "good" equilibrium towards greater norm adoption. The reverse order of these measures may have detrimental effects: the system may be driven to the "bad" equilibrium if it exists, and this may make later successful market interference extremely expensive.

Among the market parameters to be influenced politically, choices should be made according to the dissemination of the norms in the given society. Political measures which alter the effect that the norms impose on demand should only be implemented when norm adoption is wide already. If it is not, the effect is not only diminished by the small number of individuals who may react to the policy measure, but also by a possible reintroduction of at least some cognitive dissonances from having the norm but not complying with it, which in our model would be tantamount to reducing *CB*. The effect would be less norm adoption and thus even less effectiveness of the political instruments. Policies which affect the valuation for the innovative product of both norm-adopters and hedonists in the same way (such as a subsidy for consumption of norm-compliant behavior) or operate on the supply side (such as cost reductions) will of course also have the desired effects, but cannot be tailored to the level of norm adoption.

If the norm compels individuals to use electric mobility rather than to avoid gasoline cars, i.e. if the effect of norm adoption on individual demand for electric cars (Δ^e in our model) is positive, then discontinuity of the number of firms may have to be considered when determining political action to support the innovation of electric cars. In particular, if the number of suppliers is small due to an initially low demand for such cars, discontinuity effects tend to be large. As a consequence, temporary policy measures supporting the innovation are more likely to have permanent effects. In addition, the permanence of the effects is triggered faster than if multiplicity of the equilibria only stems from positive feedback loops in norm formation (in our model, working via the market). However, this permanence cuts both ways. Not only is the return to an initial equilibrium with lower consumption of the innovation avoided, but also further increases in consumption may be hindered. If additional stable equilibria occur on the way from an equilibrium of little consumption to an equilibrium of much consumption, then their regions of attraction may trap the system before it can evolve to the region of attraction of the "best" equilibrium. Hence, if policy suspects the existence of multiple equilibria due to positive feedback loops in the norm formation process and the market structure on the new market is a small oligopoly or even a monopoly, then policies aiming at overcoming equilibria of little norm adoption have to be particularly strong and patient.

2.6. Conclusions

Our paper introduces a new dimension to the interaction between markets and norms beyond the interplay of monetary and non-monetary incentives to act in a certain way: innovation of material goods as a catalyst of norm evolution. This new dimension allows us to incorporate two neglected channels through which markets affect norm evolution. On the one hand, consumption may express the normative attitude of an individual, but only if products vary sufficiently with respect to compliance of the considered norm. On the other hand, observed consumption also exposes an individual to social influence which may reinforce norm adoption or norm abandonment. We have condensed these arguments in a model that extends the existing literature on the evolution of social norms in three ways. First, our model incorporates the influence of a product innovation on the process of norm adoption. Second, we consider how conformity bias in the consumption of material goods affects the adoption of idealistic norms. Third, we demonstrate how market structure, through its impact on market outcomes, may influence norm dynamics. We thereby add to the understanding of how the evolution of norms depends on market activities.

Within our model, we have pursued two questions. First, we studied how an innovation that differs with respect to the level of norm compliance modifies the dissemination of a norm. Second, we investigated the effect of market dynamics on the evolution of the norm with respect to the existence and stability of the equilibria. Concerning the first question, we have derived the necessary and sufficient conditions for an innovation to induce an increasing dissemination of the social norm. The innovation increases the norm diffusion if (1) the conformity bias is weak or enough individuals already bear the norm prior to the innovation and (2) the increase of individual demand for the norm-compliant product variant resulting from norm adoption exceeds the corresponding demand for the norm-violating variant by a sufficient degree. These conditions become more restrictive when fewer firms are in the market, since then the required increase in profits to induce an additional incumbent to produce the innovative product increases.

With respect to the second question, we have shown that multiple norm equilibria may not only result if norm adoption is a frequency-dependent opinion formation process with direct positive feedback loops. But multiplicity may also arise if norm adoption depends on observed market behavior, in particular, on the proportion of norm compliant consumption. The direct positive feedback loop may be weaker when multiple equilibria are also supported by a conformity bias in consumption of material goods. We have further derived sufficient conditions under which the positive effect of norm adoption on individual demand induces multiplicity of equilibria. It turns out that the effect of the norm on the demand for the norm-compliant variant may be neither too high nor too low as compared to the effect on demand for the norm-violating product for multiplicity to arise. We have also discussed a second possible source of multiplicity of norm equilibria, the market structure. In principle, if more suppliers offer a norm-compliant good, they would offer the good at lower prices, thereby facilitating norm compliance and norm adoption.

This could in turn increase the demand for the norm-compliant good and thereby allow more suppliers to enter the market. It turns out, though, that this feedback loop may reinforce already existing positive frequency dependency as source of multiplicity of equilibria, and will rarely induce multiple equilibria on its own.

Based on these results, we have drawn conclusions for policy makers aiming at a higher dissemination of the social norm as an intermediate goal to ultimately achieving the greater goal of reducing environmental pollution. We have discussed the case that the conformity bias may be so strong that it hinders the dissemination of the innovation. It is mainly in these cases where political interference with market forces (and norm formation) is appropriate. If policy suspects the existence of multiple equilibria due to positive feedback loops in the norm formation process and the market structure on the new market is a small oligopoly or even a monopoly, then policies aiming at overcoming equilibria of little norm adoption have to be strong and patient. Political measures which alter the effect that the norm imposes on demand should only be implemented when norm adoption is already wide spread. If it is not, the effect is not only diminished by the small number of individuals who may react to the policy measure, but also by a possible reintroduction of at least some cognitive dissonance from having the norm but not complying with it.

3. Evolution of cooperation in social dilemmas: signaling internalized norms.

3.1. Abstract

Economists have a long tradition of finding that the evolution of cooperation in large, unstructured societies is a puzzle. We suggest a new explanation for cooperation that avoids the restrictions required in most previous attempts. Our explanation deals with the role of internalized norms for cooperation in large unstructured populations. Even internalized norms, i.e. norms that alter the perceived utility from acting in a cooperative or uncooperative way, will not help to overcome a dilemma in an unstructured society, unless individuals are able to signal their property of being a norm bearer. Only when having the norm may be communicated in a reliable way, can the picture change. We derive necessary and sufficient conditions for cooperation to be part of an asymptotically stable equilibrium of an evolutionary dynamics of signaling norm internalization, behavior and norm adoption. These conditions put the signaling costs of norm-adopters and non-adopters, the strength of the social norm and two parameters measuring the cost of cooperation into relation with each other.

Keywords: Evolution - Cooperation – Signaling

JEL Classifications: A13, D02, D21

3.2. Introduction

Despite the obvious advantages of exploiting the good will of others, human beings often cooperate, even in large, unstructured societies. However, cooperation is neither universal nor is it easy to explain. Economists have a long tradition of finding that the evolution of cooperation in large, unstructured societies is a puzzle (e.g. Axelrod and Hamilton 1981; Fudenberg et al. 2012); and in explaining cooperation based on some structure within the population.

Attempts to solve the puzzle are abundant but have thus far commonly relied on one or both of two restrictions. The first restriction is that explanations have focused on structured populations, in which interactions are not completely anonymous but allows individuals to collect and process information about past behavior of others and about their identity. The second restriction is that explanations have depended on an unexplained ability of social norms to restrict the individuals' action or strategy spaces, in particular, with respect to the abuse of punishment.

With respect to the first group of restrictions, some strands of the literature deserve special mention.[13] The theory of kin selection focuses on cooperation among individuals who are genetically closely related (Hamilton 1964a, 1964b), whereas theories of direct reciprocity focus on incentives to cooperate in repeated interactions of self-interested individuals (Trivers 1971; Axelrod 1984). For infinite repetition within one group, see Taylor (1976) or Mordecaï (1977) and for Folk-Theorem-type of results Rubinstein (1979) or Fudenberg and Maskin (1986). For

[13] A complete literature review lies outside the scope of an introductory section of a journal article, as it would merit a scholarly work in its own right.

indefinite repetition, see Kreps et al. (1982). The theories of indirect reciprocity and costly signaling show how cooperation in larger groups can emerge when those cooperating can build a reputation (Nowak and Sigmund 1998; Wedekind and Milinski 2000; Gintis et al. 2001)[14].

In terms of the second set of exclusions, we point to early papers of Hirshleifer and Rasmusen (1989) and Witt (1986) that allow for punishment only after a norm has been violated. Sethi (1996) allows for all possible strategies which condition punishment on either the violation of or compliance with a norm. However, he then adds structure to the society by introducing some exogenous division of the population – the behavior of some individuals is rational, and for the rest it is determined by routines that are slowly adapted to their environment.

We present a new explanation for cooperation that avoids both restrictions. Our explanation focuses on cooperation in large unstructured populations of individuals whose incentives to use or abuse actions or strategies evolve endogenously from the model. We assume that their behavioral routines adapt to the sum of both objective and subjective payoffs and that their subjective payoffs – which express internalized norms – slowly evolve according to the objective payoffs. This allows us to explain all variation among individuals endogenously and to assume absence of any information on the past behavior of other individuals.

We place our model in an environment that is most unfavorable to cooperation, a completely unstructured society where every interaction occurs among strangers. We do this for two reasons. The first reason is methodological: we want to isolate the impact of internalized norms from other factors that might stabilize cooperation. The other is empirical: we believe that in modern societies a non-negligible part of everyday interactions is characterized by cooperation in dilemma situations although they actually do take place in an unstructured environment (for a survey on experimental evidence see Roth 1995; Cooper et al. 1996).

In such an environment, cooperation cannot be induced by any form of repeated interaction[15] nor by social norms based on sanctions to be inflicted in later interactions. Even internalized norms, i.e. norms that alter the perceived utility from acting in a cooperative or uncooperative way, will not help to overcome a dilemma in an unstructured society, unless – and this is the thrust of the current paper – individuals are able to signal their property of being a norm bearer[16]. If internalized norms simply exist while lacking the possibility of being signaled or screened for, they would induce norm bearers to cooperate and be exploited by others. Hence, norm bearers

---

[14] There are other mechanisms that do not rely on informational aspects. Instead, they are based on restrictions in rationality or on extended strategy spaces. In finitely repeated games, cooperation can, for example, result from bounded complexity of strategies (Neyman 1985), history-dependent payoffs (Janssen et al. 1997) or bounded complexity of beliefs (Harrington, 1987).

[15] Kandori (1992) and Ellison (1994) show that in an environment with similar informational restrictions as in our model, contagious strategies may support cooperation in a social dilemma in an extremely indirect way of repeated interaction. In such strategies, when one player defects in one period, his opponent of that interaction will start to defect from this period onwards, infecting other player who will defect in the future, infecting others and so forth. For any given fixed population size, Kandori (1992) and Ellison (1994) show that cooperation can be sustained in a sequential equilibrium if individuals exhibit enough patience. However, such contagious strategies may only uphold complete cooperation by all individuals in large societies, if patience is nearly infinite. In addition, they are not tolerant with respect to behavioral errors. We therefore do not discuss this approach in detail.

[16] For an empirical paper on the role of costly signaling for the promotion of intragroup cooperation, see Soler (2012).

would have a clear evolutionary disadvantage so that norm adoption would vanish. Only when internalization of the norm can be communicated in a reliable way, may the scenario change, because behavior may then be conditioned on the expected behavior of others.

Within this environment, we borrow two elements from the indirect evolutionary approach (Güth and Yaari 1992 and Güth 1995): first, the idea that internalized norms are nothing more than an internal payoff conditional on the behavior of the individual and its partners, and second, the assumption that the adoption of an internalized norm evolves slowly depending on its effects on material, external payoffs. Our approach is thus closely related to Güth et al. (2000), who analyzes the Game of Trust rather than the Prisoners' Dilemma. The two games are clearly similar since in the Game of Trust, the outcome of the first mover trusting and the second mover reciprocating is Pareto-superior to the unique Nash equilibrium. In Güth's model, evolution allows for heterogeneity with respect to the evaluation of the material outcome such that some agents will reciprocate and some will exploit trust as second movers. By adding the opportunity of partially informative but costly screening of this evaluation to the standard Game of Trust, Güth opens the path to equilibria in which the first mover trusts and the second reciprocates. We carry this approach over to the Prisoners' Dilemma and concentrate on signaling, instead of screening.

In addition to these differences with respect to the interaction environment, we depart from the standard indirect evolutionary approach in a fundamental way concerning the behavioral assumptions. We assume that agents play inherited strategies defining both whether the agents signal their norm internalization and whether they cooperate or not. We thus take the stand of behavioral economics (as it is often reflected in evolutionary game theory) whereas Güth et al. (2000) apply a rational choice approach with agents using Bayesian updating and making rational investment decisions with respect to screening. Our model is thus evolutionary with respect to both norm internalization and behavior, although the speed of the norm internalization dynamics is clearly less than the speed of behavioral adaptation.

In the field of evolutionary biology it has been argued before that signaling may provide way out of social dilemmas where mechanisms such as reputation, reciprocity or assortative matching are absent or fail to work sufficiently (e.g. Wright 1999; Smith and Bliege Bird 2000; Leimar and Hammerstein 2001). Yet only a few of these approaches incorporate a formal model (Gintis et al. 2001). The novelty of our approach is the derivation of the full set of behavioral equilibria, i.e. all separating, pooling and semi-pooling equilibria of the signaling-extended Prisoners' Dilemma. This would be rather a technical note were it not for the implication of a far richer set of rate-of-norm-adoption equilibria that can stabilize cooperation. Notably, the interplay of those multiple behavioral equilibria may stabilize partial cooperation and dissolves the necessity to introduce specific frequency-based evolutionary forces into the dynamics of norm adoption beyond payoff monotonicity (e.g. Gintis et al. 2001 rely on the replicator dynamics).

Sethi (1996) suggests a linkage between his own approach, i.e. mixing optimizing and non-optimizing behavior in an evolutionary game; and the approach taken by Güth and Yaari (1992) and Güth and Kliemt (1994) in which all agents are assumed to optimize given heterogeneous preferences. Both authors establish the existence of games in which preferences for cooperation or fairness are evolutionary stable. Similarity in results despite differences in methodology suggest that the two research approaches are highly complementary Sethi (1996, p. 117). Our results show that the complementarity between these different approaches is limited. We show that there is substantial difference between assuming that norms simply fix a certain behavior, and assuming that norms only create internal incentives to adhere to this behavior. In our case, the parameter measuring the strength of this incentive affects the range of the other parameters for which cooperation may emerge.

The remainder of the paper proceeds as follows. The model is presented in Section 3.3. Since we consider a heterogeneous population composed of norm adopters and non-adopters, we first derive equilibria in each sub-population for which the stable equilibria are presented in Section 3.4. Thereafter, we endogenize heterogeneity and consider equilibria of the two subpopulations in Section 3.5. Section 3.6 collects and presents the requirements for partial or full cooperation being part of a stable evolutionary equilibrium. Section 3.7 concludes.

3.3. The model

The classical Prisoners' Dilemma (PD) is the most prominent and best-studied example of a social dilemma and serves as the basis for our analysis. The PD is played recurrently in an unstructured population. An *unstructured population* is defined by the anonymity of the interaction, i.e. agents process only information on outcomes of their own past interactions. In particular, they process no information on the opponent's identity or on outcomes in games in which they were not involved. To save space, payoff matrices are given from the row player's perspective. The strategy domain is finite, consisting of two strategies, C – "cooperation" and D – "defection". In conformity with the standard evolutionary model, we assume that individuals are randomly matched into pairs with each pair having the same probability in each short time period.[17] Any pair will engage in a one-shot PD game. Table 3-1 below presents the material payoffs of the PD that will be decisive with respect to evolutionary success.

Material payoffs are given by:

	C	D
C	1	$-\beta$
D	$1+\alpha$	0

Table 3-1: Prisoners' Dilemma, where $\alpha > 0, \beta > 0$ and $1+\beta > \alpha$.

A common assumption in evolutionary models that explain the presence of cooperative behavior is that individuals play inherited strategies that may depart from payoff maximizing behavior. Playing non-maximizing strategies in this line of research is then interpreted as norm-guided (e.g. Sethi 1996). This line of argument, however, appears incomplete because while showing that such strategies can be sustained in equilibrium, it lacks motivation behind why an individual would adhere to that particular norm. We believe that individuals will not stick to any behavior that is suboptimal in the current environment. We do not claim that individuals will always do what is best for them from an objective perspective (e.g. maximizes fitness), but we argue that they will

[17] An unstructured population need not necessarily engage in uniform or random matches, but departures from those assumptions significantly complicates analysis without changing the qualitative results since we assume that population is unstructured and remains unstructured. Non-random or non-uniform matching might however increase the chance that structure is introduced into the population.

not commit to suboptimal strategies forever. Hence, in our view, any long-lasting departure from the behavior that maximizes material payoffs needs to be motivated by a valuation of the outcome of behavior that differs from the material payoffs in a substantial way. In other words, norm-guided behavior is not equivalent to an unmotivated commitment to a certain behavior, but it reflects the subjective valuation of the (physical) outcome of the game. Following this reasoning, we rely on (a variant of) the indirect evolutionary approach, pioneered by Güth and Yaari (1992)[18], i.e. we explicitly model cooperative preferences, which determine behavior, and behavior, which in turn determines fitness.

As a particular internalized norm, we focus on the case of a cooperative norm. Players carrying such an internalized preference gain an additional internal payoff if the behavioral outcome of the stage game is mutual cooperation, i.e. (C, C). We assume that there are two types in the population (high and low types). Let λ denote the share of high types in the population and let $m \in \{\underline{m}, \overline{m}\}$ be their preference parameter measuring the attitude towards cooperation, resulting in the internal payoff matrix depicted in Table 3-2 below. As Güth et al. (2000) noted in a different setting, the precise level of *m* is behaviorally irrelevant. All *m*-types for whom the same inequality with respect to α holds, form an equivalence class concerning the implied behavior. We therefore normalize $\underline{m} = 0, \overline{m} > \alpha$.[19] The value of *m* is assumed to be private information of the agent. In the tradition of Harsanyi (1967, 1968a, 1968b), beliefs about the opponent's type are common knowledge. Like Güth and Ockenfels (2005), we adopt the natural assumption that beliefs correspond to actual frequencies of types. Without communication, the impossibility result of Kandori (1992, Proposition 3) applies, which states that the unique equilibrium is characterized by full defection, i.e. everybody always defects.

Communication is modeled as an additional stage prior to the play of the adjusted PD. In that stage, agents can simultaneously send one message concerning their inner motive. Without loss of generality, we assume the message space to be the same as the type space. The message to be a low type corresponds to sending no message and is costless. As in the standard signaling model (Spence 1973) we assume the existence of a social technology which enables individuals to signal their positive attitude towards cooperation by incurring some costs. Furthermore, agents who adopted the norm are supposed to bear lower costs for sending the signal. Let $\overline{k}, \underline{k}$ denote the signaling cost for high types and low types respectively, so that $\overline{k} < \underline{k}$. In the current setup, strategies are given by signal-dependent behavior and the choice of sending the signal or not, e.g. "cooperate if signal is received, deviate if no signal is received and send signal", denoted $CD\overline{m}$. In general, a strategy is denoted by a triple XYm, where the first entry denotes behavior in case of receiving the signal (*C* or *D*), the second denotes behavior in the case of not receiving the signal (*C* or *D*), and the third signifies whether the signal is sent or not ($\overline{m}$ or $\underline{m}$, respectively).

What might such a signal be? To give an illustrative example, consider a situation where individuals elbow their way through a rummage sale. There is a table with one good offered as

[18] The indirect evolutionary approach has also been applied in different strategic settings (ultimatum game, Huck and Oechssler 1999) or to analyze the evolutionary stability of altruistic preferences (Bester and Güth 1998) or of altruistic and spiteful preferences (Possajennikov 2000).

[19] Assuming $\overline{m} > \alpha$ is necessary, since otherwise, defection would still be the dominant strategy for norm-adopters.

two variants, goods A and B. There are also two individuals, one preferring good A, the other preferring good B. However getting both goods is the first best outcome for both individuals. They can behave cooperatively, allowing the other to select their preferred good; or they can try to queue-jump and grab both goods, in which case, the other gets none. If both individuals chose not to cooperate, they will grab one of the goods by chance, leaving them in expectation with a lower utility then in the cooperative state. Hence, this example is structurally equivalent to a PD. In this scenario, the signal often used is to make room for the other person. Such a signal is costly in terms of time, which usually has some monetary equivalent. If this gesture is received by both individuals, this might lead to mutual cooperation. This example is also instructive in demonstrating that signaling in our context is rather part of the behavioral strategy than an act of rational choice. In the light of this example indeed most acts of courtesy may be understood as a signal for a cooperative attitude. The signals are not limited to this aspect though.

Evaluation of material payoffs is given by:

	C	D
C	$1+m$	$-\beta$
D	$1+\alpha$	0

Table 3-2: PD with preference for cooperation.

Based on the basic behavioral actions C and D, for the high types, there are eight signal-dependent strategies $CC\overline{m}, CD\overline{m}, DC\overline{m}, DD\overline{m}$ and $CC\underline{m}, CD\underline{m}, DC\underline{m}, DD\underline{m}$. For the low types, since defection is the dominant behavior, there are only two strategies that reflect their signals, denoted by $D\overline{m}, D\underline{m}$. We will denote the share in the subpopulation of high types playing the strategy $CC\overline{m}$ by $p_{CC\overline{m}}$ and accordingly, for any other strategy. Since low types always defect, we denote their respective shares by $p_{\overline{m}}$ and $p_{\underline{m}}$.

In evolutionary game theory, there are two approaches with respect to capturing the dynamical aspect of evolution. The first one, due to the work of Smith and Price (1973), centers on the concept of an evolutionary stable strategy and is considered as a "static" approach since typically no reference is given to the underlying process by which behavior changes in the population. The second approach does not attempt to define a particular notion of stability. By explicitly modeling the underlying dynamics, all standard stability concepts used in the analysis of dynamical systems can be applied. We will follow the second approach by modeling the dynamics of the according population shares via payoff-monotone dynamics (see e.g. Bendor and Swistak, 1998 for definitions), i.e. if the fitness payoff of a certain strategy is larger than the one of another, the share of the population following the former will increase faster than the share following the latter, or decrease slower. An equilibrium is defined by the dynamics introduced above. An equilibrium is a distribution in the shares of the population playing certain strategies, such that the dynamical process induces no further adjustments, i.e. an equilibrium is a fixed point of the adjustment process. As a stability concept, we will apply the notion of asymptotic stability (see. e.g. Samuelson, 1997 for definitions). An equilibrium of that type must be reconstituted after a small perturbation, which is arbitrary in terms of the composition of mutation-strategies.

As mentioned above, there are eight strategies for high types and two for low types. We assume that the dynamic accommodation of the population shares playing the various strategies is relatively fast compared to the dynamics of the population share of $\bar{m}$-types, i.e. λ.[20] This assumption will simplify analysis of the dynamics and is considered adequate since behavior will adapt faster to differences in payoffs than socially and culturally transmitted norms. We can therefore analyze these processes separately as long as the faster process is stable. More precisely, we apply the mathematical tool of quasi-stationary approximation, or 'adiabatic elimination' (Haken 1977; Weidlich and Haag 1983, used in economics by Samuelson 1947: 320, already) of fast variables to solve the coupled differential equations which describe our system. The system consists, on the one hand, of the differential equations that describe the fast dynamics of various signal-behavior strategies and, on the other hand, of the differential equations that describe the slow dynamics of norm-adoption. The eight strategies for high types and the two for low types amount to ten differential equations, one per share per strategy, yielding nine independent equations since the size of the total population is fixed. Fixing the size of each subpopulation while analyzing the dynamics of behavioral strategies within each subpopulation reduces the number of independent differential equation by one more, seven for the high types and one for low types. We recall that p_{XYm} and p_m denote the shares of strategies *within* the subpopulations so that $\sum_{X,Y,m} p_{XYm} = 1$ with $X,Y \in \{C,D\}$ and $m \in \{\underline{m},\bar{m}\}$ and $p_{\underline{m}} + p_{\bar{m}} = 1$.

Given our assumption on the speed of the dynamic processes, we first derive all the behavioral equilibria for a given proportion λ of individuals with a high internal motivation for (mutual) cooperation, and then analyze whether the implied λ-dynamics can support a fully or partially cooperative state. We call the former equilibria 'p-equilibria' and the latter, 'λ-equilibria'. If they are asymptotically stable with respect to the corresponding p- or λ-dynamics, we say that they are p-stable and λ-stable, respectively. The p-stable equilibria are presented in section 3.4, and λ-stable equilibria are derived in section 3.5.

3.4. Equilibria with Exogenous Proportions of Norm Bearers

For ease of reading, we present only the equilibria and their stability properties and leave the derivation in Appendix B.4 (existence) and B.3 (stability). As in many other cases, we have separating and pooling equilibria, depending on the parameters including λ. There are one p-stable separating and three p-stable pooling equilibria. In the separating equilibrium, the subpopulations of the two types of individuals (high and low internal motivation for cooperation) exhibit homomorphic behavior, whereas behavior of types in the pooling equilibria is heteromorphic. However, there is a third type of equilibria where at least one subpopulation applies both types of signals, so called semi-pooling equilibria. Table 3-3 reports these equilibria.

In the following paragraphs, we will take a closer look at the separating and pooling equilibria. We will refer to the first of these equilibria as the '*cooperative separating equilibrium*', to the second as the '*low pooling cooperative equilibrium*', to the third as the '*low pooling defective equilibrium*' and to the

---

[20] This assumption implies that payoff monotonicity is restricted to the fast and to the slow dynamics, but does not comprise the combination of the two.

fourth as the '*high pooling cooperative equilibrium*'. It turns out that the semi-pooling equilibria with one exception are less important for the implied λ-dynamics and are therefore not further discussed. The exception is the p-stable semi-pooling equilibrium at $\lambda = \frac{\underline{k}}{1+\alpha}$ that will be of relevance for one of the inner λ-stable equilibria. In this semi-pooling equilibrium, high types always play $CD\overline{m}$ and low types are indifferent between sending the signal or not, and therefore $p_{\overline{m}}$ is undefined. The minor importance of all other p-stable semi-pooling equilibria is partly due to their being characterized by strictly negative fitness differentials between high and low types and partly to their limited λ-support (see Figure 3-1 and Figure 3-2).

In the cooperative separating equilibrium, the high types recognize each other and cooperate only among themselves. The intuition behind the fact that the support of this equilibrium has both a lower and an upper is as follows: If there are too few high types, then the cooperative outcome among them cannot compensate for the signaling costs. The higher the signaling costs relative to the (non-material) reward for a cooperative outcome, the higher the required share of high types in the population. If on the other hand, there are too many high types, signaling becomes sufficiently profitable for low types. In other words, if there are enough high types that cooperate when receiving the cooperative signal, it becomes profitable for low types to incur the signaling costs. The higher the signaling cost for low types relative to what can be gained from defection against a cooperative opponent, the higher is the share of high types needed for signaling to become a profitable strategy for low types. The thresholds for the share of high types have a precise economic interpretation. For high types, the cost-benefit ratio from signaling ($\frac{\overline{k}}{1+\overline{m}}$) must be smaller than the probability to gain the benefit (λ). The reverse holds true for low types, i.e. their cost-benefit ratio from signaling must exceed ($\frac{\underline{k}}{1+\alpha}$), the likelihood of gaining the benefit.

In the low pooling cooperative equilibrium, nobody signals and high types cooperate. This equilibrium exists if there are sufficiently many high types. Only then can they compensate for the loss from being cooperative against low types by the cooperative outcome among each other. In other words, if the share of high types falls below a certain threshold, then they will start to prefer defecting when receiving the low signal. Note that this equilibrium is indeed an equilibrium set, since the strategies $CC\underline{m}$ and $DC\underline{m}$ are equivalent in equilibrium. The share of high types required for this to be an equilibrium increases in the sucker's payoff, since cooperative behavior becomes more disadvantageous with increasing (absolute) sucker's payoffs. This threshold, too, has an intuitive meaning. Note that $\overline{m}-\alpha$ (β) measures the incentive to reciprocate cooperative (defective) behavior. In essence, the condition $\frac{\beta}{\beta+\overline{m}-\alpha} < \lambda$, which can be rewritten as $\lambda(\overline{m}-\alpha) > (1-\lambda)\beta$, states that the expected gain from reciprocating cooperative behavior must exceed the expected gain from reciprocating defective behavior.

Type	Involved strategies	Equilibrium	Support	Conditions for existence	Payoff differentials (superscript "f" indicates the difference in fitness payoffs)
Separating	High types cooperate against signal and defect against no signal. Norm holders signal, while others do not signal.				
	$CD\bar{m}$ $\underline{m}$	$p_{CD\bar{m}}=1,\ p_{\underline{m}}=1$	$\frac{\bar{k}}{1+\bar{m}}<\lambda<\frac{\underline{k}}{1+\alpha}$	$\bar{k}<1+\bar{m}$	$\Pi_{\bar{m}}(CD\bar{m})-\Pi_{\underline{m}}(\underline{m})=\lambda(1+\bar{m})-\bar{k}$ $\left(\Pi_{\bar{m}}(CD\bar{m})-\Pi_{\underline{m}}(\underline{m})\right)^{f}=\lambda-\bar{k}$
Low Pooling	High types cooperate, no signal				
	$CC\underline{m}$ $DC\underline{m}$ $\underline{m}$	$p_{CC\underline{m}}+p_{DC\underline{m}}=1$ $p_{\underline{m}}=1$	$\lambda\geq\frac{\beta}{\bar{m}-\alpha+\beta}$		$\Pi_{\bar{m}}(CC\underline{m})-\Pi_{\underline{m}}(\underline{m})=-\left(\lambda(\alpha-\bar{m})+(1-\lambda)\beta\right)$ $\left(\Pi_{\bar{m}}(CC\underline{m})-\Pi_{\underline{m}}(\underline{m})\right)^{f}=-\left(\lambda\alpha+(1-\lambda)\beta\right)<0$
	Complete defection, no signal				
	$CD\underline{m}$ $DD\underline{m}$ $\underline{m}$	$p_{CD\underline{m}}+p_{DD\underline{m}}=1$ $p_{\underline{m}}=1$	$0<\lambda<1$	$p_{CD\underline{m}}\leq\frac{1}{\lambda}\min\left\{\frac{\bar{k}+\beta}{1+\bar{m}+\beta},\frac{\bar{k}}{1+\alpha}\right\}$	$\Pi_{\bar{m}}(CD\underline{m})-\Pi_{\underline{m}}(\underline{m})=0$ $\left(\Pi_{\bar{m}}(CD\underline{m})-\Pi_{\underline{m}}(\underline{m})\right)^{f}=0$
High Pooling	High types cooperate, all signal				
	$CC\bar{m}$ $CD\bar{m}$ $\bar{m}$	$p_{CC\bar{m}}+p_{CD\bar{m}}=1$ $p_{\bar{m}}=1$	$\lambda\geq\max\left\{\frac{\bar{k}}{1+\alpha},\frac{\beta}{\bar{m}-\alpha+\beta},\right\}$	$\underline{k}<1+\alpha$ $\lambda\geq\frac{\underline{k}}{p_{CD\bar{m}}(1+\alpha)}\Rightarrow p_{CD\bar{m}}\geq\frac{\underline{k}}{(1+\alpha)}$	$\Pi_{\bar{m}}(CC\bar{m})-\Pi_{\underline{m}}(\bar{m})=\underline{k}-\bar{k}-\left(\lambda(\alpha-\bar{m})+(1-\lambda)\beta\right)$ $\left(\Pi_{\bar{m}}(CC\bar{m})-\Pi_{\underline{m}}(\bar{m})\right)^{f}=\underline{k}-\bar{k}-\left(\lambda\alpha+(1-\lambda)\beta\right)$

Table 3-3: p-stable equilibria (p-stable semi-pooling equilibria are referred to Appendix B.2)

In the low pooling defective equilibrium, nobody sends the cooperative signal and everybody defects earning a payoff of zero. Again, due to lack of distinguishability in equilibrium, equilibrium is indeed a set where $CD\underline{m}$ and $DD\underline{m}$ might be played by high types. This set of equilibrium reflects the benchmark solution in the underlying game and exists for all population compositions between high types and low types.

In the high pooling cooperative equilibrium, everybody signals and high types cooperate. This equilibrium exists if there are sufficiently many high types. If the latter's proportion is large enough, they can compensate for the loss from being cooperative against low types by the cooperative outcome among each other. In other words, if the share of high types falls beneath a certain threshold, they will then start to prefer to play defective while receiving the low signal. Contrary to the low pooling equilibrium, an additional restriction with respect to the share of high types will arise, reflecting the incentive compatibility for low types to signal. Note that this equilibrium is again an equilibrium set, since the strategies $CC\overline{m}$ and $CD\overline{m}$ are equivalent in equilibrium. The share of high types required for this to be an equilibrium weakly increases in the sucker's payoff and the signaling cost for low types. Since with increasing (absolute) sucker's payoffs, cooperative behavior and sending the signal for low types respectively become more disadvantageous. Here, for low types, the reverse logic applies in comparison to the separating cooperative equilibrium, i.e. for low types to find it worthwhile to signal, their cost-benefit ratio ($\frac{\underline{k}}{1+\alpha}$) must be smaller than the likelihood to profit from signaling (λ). The lower bound stemming from incentive constraint for high types bears the same logic as in the low pooling cooperative equilibrium.

3.5. Endogenous Proportion of Norm Bearers

We now analyze the dynamics of the share of high types in the population for which we assume that the p-dynamic has reached a stable p-equilibrium, as we assumed that inner motives evolve far more slowly than behavioral frequencies. The evolution of the proportion of norm bearers is determined by its relative fitness. Fitness is measured by the material payoffs as presented in Table 3-1. Thus, any preference parameter measuring the evaluation of material payoffs will be neglected when calculating fitness payoffs. Analogous to the derivation of p-equilibria, the differentials in these fitness payoffs among high and low types are the driving force for the evolution of their respective shares. To ease the understanding of the differentials of fitness payoff differentials, we provide some intuition for their size in the relevant p-stable equilibria.

In the cooperative separating equilibrium, both types defect in all interactions, except when two individuals of the high type meet. In this case, they cooperate. The low type will thus always earn a fitness payoff of zero, and the high type will earn a fitness payoff of one with probability λ, i.e. the probability that he interacts with another individual of the high type. Since high types unconditionally bear the signaling cost $\bar{k}$, their expected payoff in the cooperative separating equilibrium is $\lambda-\bar{k}$, which is also the expected difference of fitness payoffs: $\left(\Pi_{\overline{m}}\left(CD\overline{m}\right)-\Pi_{\underline{m}}\left(\underline{m}\right)\right)^{f}=\lambda-\bar{k}$. Obviously, this fitness advantage of the high type grows in the share of high types in the population.

In the two (partially) cooperative pooling equilibria, individuals of the high type cooperate in reaction to the signal they send, and all individuals of the low type copy this signal but still defect.[21] Leaving aside signaling costs for a moment, differences in material payoffs then reflect payoffs of unconditional cooperators and defectors in the underlying PD. More precisely, with probability λ, high types meet their own type and realize the cooperative outcome, earning 1. With the residual probability, they meet a low type and lose β. Low types always defect and only earn positive payoffs when matched with high types, which happens with probability λ and earns them $1+\alpha$. A fitness differential to the advantage of the high types thus cannot result from playing the game itself, but only form sufficiently large differences in signaling cost (see Table 3-3). Obviously, if no signal is sent, as is the case in the low pooling cooperative equilibrium, the fitness payoff of the high type can only be smaller than that of the low type, $\left(\Pi_{\bar{m}}\left(CC\underline{m}\right)-\Pi_{\underline{m}}\left(\underline{m}\right)\right)^{f}=-\left(\lambda\alpha+\left(1-\lambda\right)\beta\right)<0$.

Only in the high pooling cooperative equilibrium, the signaling cost disadvantage of the low type may outweigh the disadvantage of the high type from playing cooperatively in the game, so that the high type earns a higher fitness payoff than the low type, $\left(\Pi_{\bar{m}}\left(CC\bar{m}\right)-\Pi_{\underline{m}}\left(\bar{m}\right)\right)^{f}=\underline{k}-\bar{k}-\left(\lambda\alpha+\left(1-\lambda\right)\beta\right)$.

Obviously, the fitness payoff difference increases (declines) in the share of the high types if defection is more (less) tempting against defection than against cooperation, i.e. if β is larger (smaller) than α. If the proportion of the high type in the population is too small, it is either not worthwhile to mimic the other type, or the chances to meet another high-type individual are so low that cooperation ceases to be the best reaction to the signal sent by all individuals. For these small shares of the high type in the population, the pooling cooperative equilibria break down just like the cooperative separating equilibrium discussed earlier breaks down for shares of the high type that are too large.

In the pooling defective equilibrium, both types always defect without sending signals and thus all earn the same fitness (and behavioral) payoff of zero.

The following two figures depict the differences in material payoffs for the various p-stable equilibria (see Table 3-3).

[21] This implies that the other signal is never sent, which explains why the high type is indifferent between the two behavioural actions C and D to this never-observed signal.

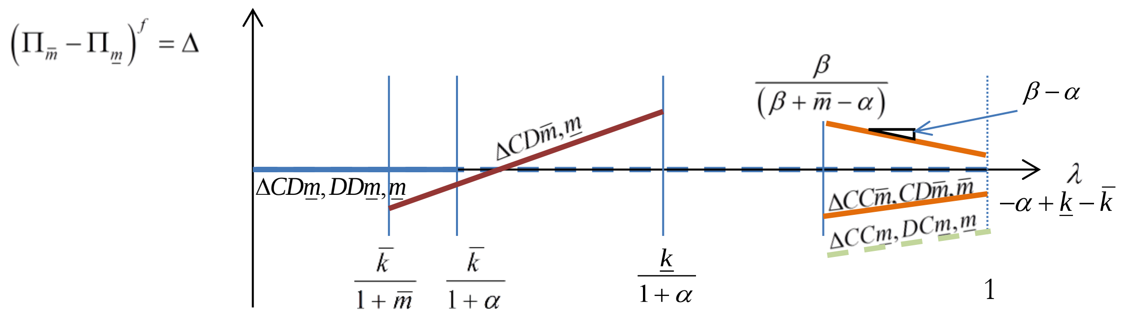


Figure 3-1: Differences in material payoffs for $\frac{\underline{k}}{1+\alpha}<\frac{\beta}{(\beta+\overline{m}-\alpha)}$

Payoff differences for semi-pooling equilibria are neglected since their support lies in the interval $\left(\frac{\beta}{(\beta+\overline{m}-\alpha)},1\right)$ and the difference is strictly negative for all. Hence, their presence will have no important implications for the dynamics of the share of high types.

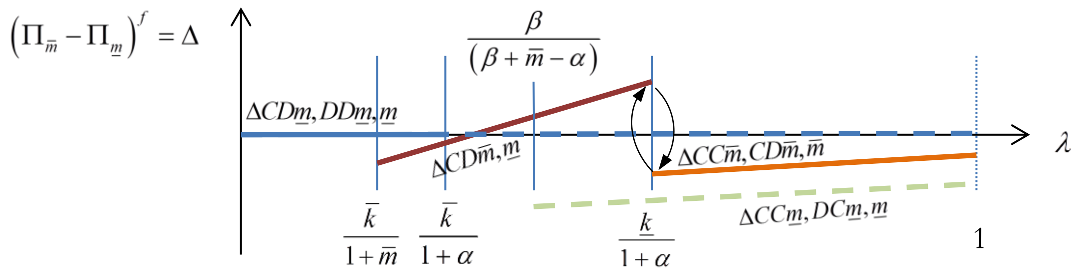


Figure 3-2: Differences in material payoffs for $\frac{\underline{k}}{1+\alpha}\geq\frac{\beta}{(\beta+\overline{m}-\alpha)}$

A stable λ-equilibrium may be realized around one p-stable equilibrium or by the interplay of several such equilibria. We first concentrate on the first case, which we further differentiate into corner equilibria (Lemma 3-1) and inner equilibria (Lemma 3-2) and then turn to the second case (Lemma 3-3).

In the first case, the difference in fitness payoffs between high and low types must vanish to constitute a stationary point at this particular value of the share of high types λ. For stability, in the neighborhood of an equilibrium λ^*, high types must earn strictly more than low types for $\lambda<\lambda^*$ and strictly less for $\lambda>\lambda^*$. In terms of Figure 3-1 and Figure 3-2, the stationary point is a zero of the linear payoff difference for a certain p-stable equilibrium, and stability is equivalent to a negative slope of the payoff difference function. Of course, the requirement with respect to the zero and the slope is only relevant for inner equilibria. At the upper bound of the domain, $\lambda=1$, a strictly positive payoff difference in favor of high types at $\lambda<1$ is necessary and sufficient for having a corner equilibrium. At the lower bound of the domain, $\lambda=0$, a strictly negative payoff difference at $\lambda>0$ is necessary and sufficient for having a corner equilibrium.

We first analyze whether λ-stable equilibria with full cooperation exist. Since only high types may cooperate, this is equivalent to asking whether there is a λ-stable equilibrium at $\lambda=1$ with cooperating high types. Since high types in the low pooling cooperative equilibrium face an evolutionary disadvantage for all population compositions, this p-stable equilibrium cannot induce a stable cooperative λ-equilibrium (partial or full). Hence, there are two potential candidates left, the separating cooperative equilibrium and the high pooling equilibrium. The following lemma states the conditions such that a locally stable equilibrium with only high types present in the population who cooperate with each other exists.

Lemma 3-1 The PD can be fully resolved as a locally λ-stable equilibrium only in two ways:

(1) by the separating cooperative equilibrium if and only if $\underline{k} \geq 1+\alpha$ and $\overline{k} < 1$

(2) by the high pooling cooperative equilibrium if and only if $\underline{k} < 1+\alpha$ and either $\underline{k} - \overline{k} > \alpha$ or $\underline{k} - \overline{k} = \alpha > \beta$.

All proofs are in Appendix B.1.

The existence of fully cooperative equilibria seems surprising at first glance, but a closer look at the stated conditions for their existence reveals how rarely they occur. In the case of the separating cooperative equilibrium, the condition corresponds to a scenario where the signaling cost for low types are so severe that it will never pay for them to signal. More precisely, in a cooperative separating equilibrium with $\lambda = 1$, a single low-type mutant would earn $1+\alpha$ from playing the dominant defective strategy at cost $\underline{k}$. The second qualification $\overline{k} < 1$ stems from the incentive compatibility constraint for high types, since they could always earn zero by not-signaling and exhibiting defective behavior. In the case of the high pooling cooperative equilibrium, the difference in the signaling cost must exceed the material reward of defecting on a cooperative opponent.

The restrictiveness of Lemma 3-1 draws our attention to inner stable equilibria. The only candidate for such a λ-equilibrium supported by only one p-stable equilibrium is one associated with the high pooling cooperative equilibrium at $1 - \frac{\underline{k} - \overline{k} - \alpha}{\beta - \alpha}$. All other equilibria are characterized by either strictly negative or strictly increasing payoff differentials. The high pooling cooperative equilibrium exists and is λ-stable if $1 - \frac{\underline{k} - \overline{k} - \alpha}{\beta - \alpha}$ is inside the λ-support of this equilibrium and the fitness differential decreases in λ, which is the case if $\beta - \alpha < 0$ (see Figure 3-1). Taking these conditions together yields:

Lemma 3-2 The high pooling cooperative equilibrium constitutes an inner λ-stable equilibrium at $1 - \frac{\underline{k} - \overline{k} - \alpha}{\beta - \alpha}$ if and only if: $\frac{\beta}{\overline{m} - \alpha + \beta}\overline{m} < \underline{k} - \overline{k} < \alpha$ and $\beta < \frac{1+\beta}{1+\alpha}\underline{k} - \overline{k}$.

Note that the first condition implies $\beta - \alpha < 0$, which guarantees stability. As expected, the conditions presented in Lemma 3-2 are less restrictive as compared to the requirements for an equilibrium formed only by high types. Looking at the conditions, we observe that the existence of inner stable equilibria requires that the costs of signaling for norm adopters must differ sufficiently from the corresponding costs of non-adopters.

What remains to be studied is whether separating λ-equilibrium constituted by the interplay of several p-equilibria exists. For this to be the case requires: (1) the supports of the p-equilibria need to be adjacent, (2) around the point where the supports are adjacent, the differences of fitness payoffs of the relevant p-equilibria must be positive for less-than-equilibrium shares of high types and negative for more-than-equilibrium shares of high types, and (3) after λ moves from the support of one p-equilibrium to the support of another, the behavioral frequencies have to be within the basin of attraction of the "new" equilibrium if they have been sufficiently close to the "old" equilibrium. In our case, we may have such an equilibrium only at $\lambda = \frac{\underline{k}}{1+\alpha}$ where three equilibria interplay: the separating cooperative equilibrium, a semi-pooling cooperative equilibrium (last row in Appendix B.2), and the high pooling cooperative equilibrium. To facilitate understanding of this argument, we recommend that the reader views Figure 3-2 while reading the following argument.

Condition (1) requires that $\frac{\underline{k}}{1+\alpha} \geq \frac{\beta}{\bar{m}-\alpha+\beta}$ (cf. Table 3-3 and Appendix B). Condition (2) has implications for the fitness differences of the p-stable equilibria. For the cooperative separating p-equilibrium, the fitness difference is given by $\left(\Pi_{\bar{m}}\left(CD\bar{m}\right)-\Pi_{\underline{m}}\left(\underline{m}\right)\right)^{f} = \lambda - \bar{k}$ for $\lambda \leq \frac{\underline{k}}{1+\alpha}$. This difference must be strictly positive at $\lambda = \frac{\underline{k}}{1+\alpha}$, whence $1+\alpha < \frac{\underline{k}}{\bar{k}}$. In other words, the relative disadvantage for low types in terms of signal costs must exceed the relative incentive to defect given the opponent cooperates. Given this inequality and a share of high types sufficiently close to, but lower than $\lambda = \frac{\underline{k}}{1+\alpha}$, the share of the high type increases when the p-dynamics has reached the cooperative separating equilibrium. For the high pooling cooperative equilibrium, the fitness difference is given by $\left(\Pi_{\bar{m}}\left(CC\bar{m}\right)-\Pi_{\underline{m}}\left(\bar{m}\right)\right)^{f} = \underline{k}-\bar{k}-\left(\lambda\alpha+\left(1-\lambda\right)\beta\right)$, which has to be negative. Hence, we get $\frac{\underline{k}}{1+\alpha} < \frac{\bar{k}+\beta}{1+\beta}$.

To see that Condition (3) is satisfied under certain conditions we present our argument in three steps. First, we draw the reader's attention to the fact that for all three of the considered equilibria, we have $p_{CD\bar{m}} + p_{CC\bar{m}} = 1$. This implies that for $\lambda = \frac{\underline{k}}{1+\alpha}$ we have:

$$\begin{aligned} &\Pi_{\bar{m}}\left(CD\bar{m}\right) = \lambda\left(1+\bar{m}\right)-\left(1-\lambda\right)p_{\bar{m}}\beta-\bar{k} \\ &\geq \Pi_{\bar{m}}\left(CC\bar{m}\right) = \lambda\left(1+\bar{m}\right)-\left(1-\lambda\right)\beta-\bar{k} \\ &> \max_{X}\left(\Pi_{\bar{m}}\left(X\right)\right) \quad \text{where} \quad X \in \left\{C,D\right\}^{2}\times\left\{\underline{m},\bar{m}\right\}\setminus\{CD\bar{m},CC\bar{m}\} \end{aligned} \tag{3.1}$$

where the first inequality is strict if $p_{\bar{m}} < 1$ and the second inequality requires $\lambda^{*} \equiv \frac{\underline{k}}{1+\alpha} > \frac{\beta}{\bar{m}-\alpha+\beta} \equiv \tilde{\lambda}$. Hence, continuity of the payoffs and Lipschitz-continuity of the dynamics implies that for all λ sufficiently close to λ^{*} and all sufficiently large $p_{CD\bar{m}} + p_{CC\bar{m}} = 1$ we have $\dot{p}_{CD\bar{m}} + \dot{p}_{CC\bar{m}} = 1$. Hence, once the system is close enough to any of the three relevant p-stable equilibria, and in particular once $p_{CD\bar{m}} + p_{CC\bar{m}}$ has become large enough, $p_{CD\bar{m}} + p_{CC\bar{m}}$ will continue to grow for all $p_{\bar{m}}$. Second, we observe that if $p_{CD\bar{m}}$ is large enough and the p-dynamics is sufficiently fast compared to the λ-dynamics, then λ will always stay close enough to λ^{*} to uphold the validity of the first argument. Third, if $p_{CD\bar{m}} + p_{CC\bar{m}}$ is large enough and thus increases, $\Pi_{\bar{m}}(CD\bar{m}) < \Pi_{\bar{m}}(CC\bar{m})$ only occurs for ever decreasing ranges of large $p_{\bar{m}}$. Hence, for every payoff-monotone dynamic $p_{CC\bar{m}}$ will be smaller after every full cycle and will never again reach its previous maximum level. Hence, $p_{CD\bar{m}}$ will eventually be large enough to ensure the validity of our second argument.

Hence, once our full dynamic system is close enough to λ^* and the λ-dynamic is slow enough, the system will rotate between the separating equilibrium and the high pooling equilibrium in ever smaller cycles (note that this does not necessarily imply that a fixed point is reached because a limit cycle may exist). We summarize all conditions in the following:

Lemma 3-3 If $\frac{\beta}{\bar{m}-\alpha+\beta} < \frac{\underline{k}}{1+\alpha} \underset{\substack{\text{if}=\text{then} \\ \alpha>\beta}}{\leq} \frac{\bar{k}+\beta}{1+\beta}$ and $\bar{k} < \frac{\underline{k}}{1+\alpha}$, an inner λ-stable equilibrium exists at $\lambda = \frac{\underline{k}}{1+\alpha}$, in which (1) high-type individuals cooperate among each other but also with those low-type individuals who signal to be of the high type and (2) the proportion of low-type individuals who signal to be of the high type fluctuates.

Note that the conditions in Lemma 3-2 and Lemma 3-3 are mutually exclusive, i.e. there is at most one stable inner equilibrium.

We have so far not considered the case of $\frac{\underline{k}}{1+\alpha} \leq \frac{\beta}{\bar{m}-\alpha+\beta}$. If there is equality. i.e. $\frac{\underline{k}}{1+\alpha} = \frac{\beta}{\bar{m}-\alpha+\beta}$, an equilibrium of the type discussed in Lemma 3 still exists at $\lambda = \frac{\underline{k}}{1+\alpha}$, but it is unstable (the argument on condition 3 fails). If the inequality is strict ($\frac{\underline{k}}{1+\alpha} < \frac{\beta}{\bar{m}-\alpha+\beta}$), there is a gap between the λ-supports of the separating cooperative equilibrium and the high pooling cooperative equilibrium (see Figure 3-1). In the interval $\left(\frac{\underline{k}}{1+\alpha}, \frac{\beta}{\bar{m}-\alpha+\beta}\right)$, the defective pooling equilibrium is the unique equilibrium. Should the population start at the cooperative separating p-equilibrium with a positive fitness differential, then it will eventually drive the share of high-type individuals beyond the λ-support of this equilibrium so that $p_{\bar{m}}$ starts to grow. Once it grows too much, the strategy $DD\bar{m}$ yields the largest behavioral payoff to high-type individuals while $CD\bar{m}$ yields only the second largest. Hence, the share of always defecting high-type individuals $p_{DD\bar{m}}$ must grow and $p_{CD\bar{m}}$ must decline because the shares of the other strategies (with even lower behavioral payoffs) are already zero. Less cooperation by high-type individuals reduces the advantages that low-type individuals accrue from falsely signaling to be of the high type. Hence, $p_{\bar{m}}$ will eventually decline again. A behavioral equilibrium exists in which only some low-type individuals signal the wrong type and only some high-type individuals cooperate after receiving the high signal while the others always defect, but this equilibrium is not stable (see Appendix B.3). Consequently, $p_{DD\bar{m}}$ will eventually grow large enough to move the population in the attraction region of the defective separating equilibrium, where it will remain. We admit that the evolution may become more complex when $p_{\bar{m}}$ and $p_{DD\bar{m}}$ both become so large that $CD\bar{m}$

becomes less profitable than $DC\bar{m}$. There may then be payoff monotonic dynamics for which $p_{DC\bar{m}}$ starts to grow, although slower than $p_{DD\bar{m}}$. If this happens, false signaling by high types may eventually become reasonable. However, as the low pooling equilibrium with cooperation only of the high types fails to exist in the interval $\left(\frac{k}{1+\alpha}, \frac{\beta}{\bar{m}-\alpha+\beta}\right)$, we conjecture that the population will eventually end up in the defective pooling equilibrium as the unique behavioral equilibrium.

Conjecture If $\frac{k}{1+\alpha} < \frac{\beta}{\bar{m}-\alpha+\beta}$, no λ-stable inner equilibrium exists at $\lambda = \frac{k}{1+\alpha}$.

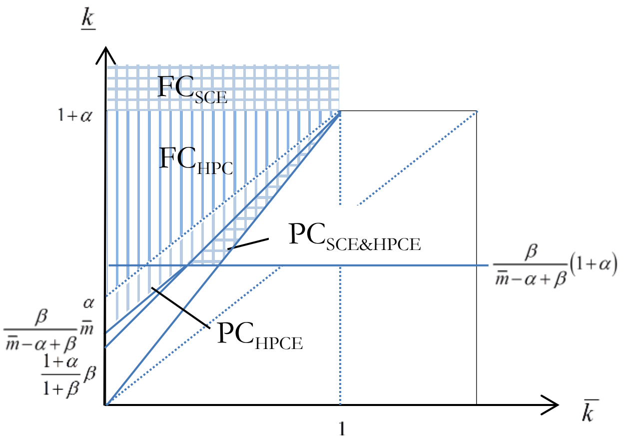


Figure 3-3: Parameter region for partial or full cooperation

Figure 3-3 illustrate the conditions of Lemmas 1 through 3 graphically. For illustrative purposes, we assume $\beta-\alpha<0$ and $\frac{k}{1+\alpha} \geq \frac{\beta}{\bar{m}-\alpha+\beta}$ so that all inner equilibria may exist for some parameter ranges. In Figure 3-3, areas marked by FC and PC represent parameter combinations for which full and partial cooperation occur, respectively. More specifically, the indexes mark parameter ranges for which cooperation is induced by the separating cooperative equilibrium (SCE), the high pooling cooperative equilibrium (HPCE), or the interplay of the two and a semi-pooling equilibrium (SCE&HPCE).

It is worth noting that the strength of the cooperative norm measured by $\bar{m}$ has a direct impact on the parameter set allowing for λ-stable inner equilibria (see Figure 3-3). As $\bar{m}$ gets closer to the incentive to defect α, the parameter region supporting a separating cooperative equilibrium ($PC_{SCE\&HPCE}$) becomes smaller and smaller. Although the exact size of $\bar{m}$ is not important for the behavioral consequence for each individual as long as $\bar{m} > \alpha$ holds true, the exact size of $\bar{m}$ does matter for the size of the parameter range for which evolutionary stable equilibria characterized by partial cooperation exist.

3.6. Collecting requirements for equilibria with cooperation

By combining Lemma 3-1 through Lemma 3-3 from the previous section, we deduce a theorem on cooperation in an unstructured population:

Theorem In an unstructured society, cooperation in a PD may exist and be stable due to the possibility of signaling the existence of inner payoffs for (mutual) cooperation, which do not affect fitness, if the costs of falsely signaling to have such inner payoffs are sufficiently large. These costs must be larger to reach full cooperation than to reach partial cooperation.

In our model, 'sufficiently large' translates to $\underline{k}-\bar{k}>\alpha$ or $\underline{k}-\bar{k}=\alpha>\beta$ for full cooperation (Lemma 3-1). For partial or full cooperation (Lemmas 1 through 3), 'sufficiently large' translates to:

$$\underline{k} \underset{\substack{\geq \\ \text{for last} \\ \text{term}}}{>} (1+\alpha)\max\left\{\min\left\{\frac{\bar{k}+\alpha}{1+\alpha},\frac{\beta}{\bar{m}-\alpha+\beta}\right\},\bar{k}\right\} \text{ if } \alpha\leq\beta \text{ and to} \tag{3.2}$$

$$\underline{k}>(1+\alpha)\max\left\{\min\left\{\frac{\beta\bar{m}}{\bar{m}-\alpha+\beta}+\bar{k},\frac{\beta}{\bar{m}-\alpha+\beta}\right\},\bar{k}\right\} \text{ if } \alpha>\beta. \tag{3.3}$$

Figure 3-4 and Figure 3-5 illustrate the interrelation between the costs for low types to signal falsely and the extent of the inner motive for mutual cooperation. This relation is determined by the various inequality conditions for existence of partial or full cooperation stated in the theorem above. Figure 3-4 and Figure 3-5 reveal the negative relation between these two parameters, i.e. in order to sustain some level of cooperation, lower signalling costs for low-types must be compensated by a higher inner motive for mutual cooperation of the high-types. Here, the aforementioned interdependence of $\bar{m}$ and the presence of cooperative equilibria is directly observable. Although the precise level of $\bar{m}$ is not decisive with respect to its behavioural consequence, its level plays a crucial role with respect to the size of the set of parameters such that partial or full cooperation could be sustained as an equilibrium outcome. Furthermore, we observe that this set of parameters is strictly decreasing in the signalling cost for the high type. Finally, Figure 3-4 and Figure 3-5 show that the chances for cooperation diminish with increasing β. In essence, the riskier or more painful cooperation occurs when matched with defective behaviour, the higher requirements have to be met with respect to signalling costs for low types and the inner motive for mutual cooperation. A mirror argument applies with respect to parameter α, measuring the incentive to defect on cooperation in the underlying game. The following corollary summarizes these insights.

Corollary

(1) The range of signalling cost for the low type allowing for partial or full cooperation is weakly increasing in the social norm for mutual cooperation $\bar{m}$.

(2) The set of $(\underline{k},\bar{m})$-pairs allowing for partial or full cooperation is strictly increasing in signalling cost for the high type $\bar{k}$ and strictly decreasing in the Sucker's payoff β and the incentive to defect on cooperation α.

The theorem reveals that in case of full cooperation, almost always it is only the incentive to defect on a cooperative player α relative to the difference in signalling costs that matters. Whereas for stable partial cooperation, the relation of α and β is relevant. The loss from playing cooperatively on a defective opponent β must be less than what a player could gain from defecting on a cooperative player. Intuitively, this explains the edge of defective players over cooperative players for shares of the latter that exceed the equilibrium level and vice versa. Reflecting on both incentives in case of a partially cooperative equilibrium is also plausible since both behaviors are present in equilibrium, whereas fully cooperative equilibria are characterized by solely cooperative actions. In that case, only the price for cooperation given the monomorphic cooperative behavior α is relevant.

Interdependence between the size of the inner motive and the cost to send a false signal

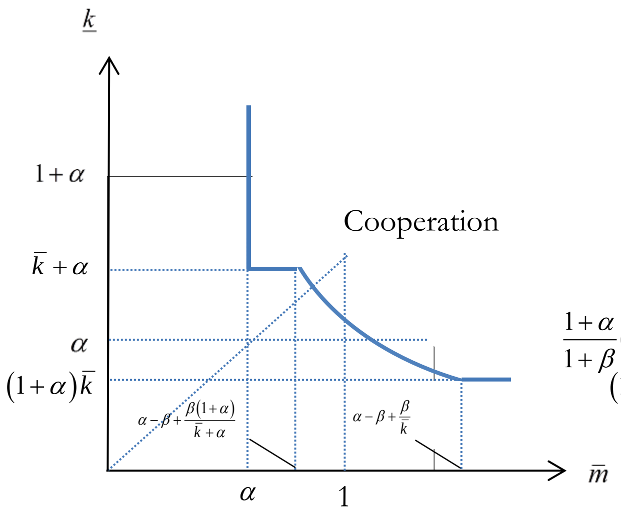


Figure 3-4: $\alpha \leq \beta$

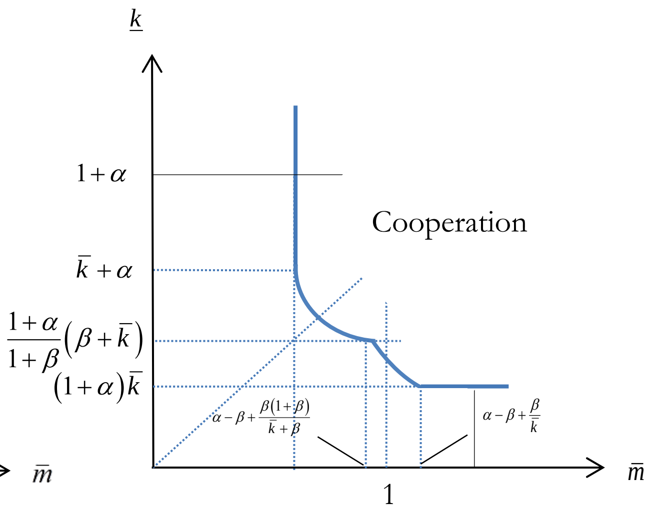


Figure 3-5: $\alpha > \beta$

3.7. Conclusion

In this paper, we analyze an evolutionary model where individuals are able to signal that they internalized a particular social norm, namely a norm for mutual cooperation. This preference was embedded in a Prisoners' Dilemma. In section 3.6, we present a theorem that states necessary and sufficient conditions for full or partial cooperation to be prevalent in a stable equilibrium. These conditions refer to the difference in signaling cost between the cooperative and the opportunistic type, the extent of the cooperative norm and the model parameters of the PD, i.e. the temptation to defect and the sucker's payoff. We obtain several interesting results. First, it is true that the exact size of the behavioral parameter measuring the internal bias in favor of mutual cooperation is not important for the behavioral consequence for each individual. However, when it comes to the presence of stable equilibria characterized by partial cooperation its size and its relation to the incentive to defect do become relevant. More precisely, the stronger the inner motive to cooperate is, the less restrictive are the conditions on the spread in signaling cost. Second, for cooperative agents to coexist with defecting agents in a stable equilibrium, it is not necessary that the signaling technology fully cancels the incentive to defect. Since this would be necessary for many corresponding results that are based on some sort of involuntary redistribution (e.g. punishment), our approach may explain cooperation in more cases than the latter approaches. Furthermore, the range of signalling cost for the low-type individuals allowing for partial or full cooperation is weakly increasing in the strength of the social norm for mutual cooperation. Finally, the set of pairs of signalling cost for the defective type and level of cooperative norm allowing for partial or full cooperation is strictly increasing in signalling cost for the high type and strictly decreasing in the sucker's payoff and the incentive to defect on cooperation.

We achieved these results by analyzing the evolution of norms concerning cooperation in the PD with one of the most general class of dynamics considered in evolutionary game theory, namely the class of payoff-monotone dynamics. Existing literature has already demonstrated that signaling may point a way out of a social dilemma where mechanisms as reputation, reciprocity or assortative matching are absent or fail to work sufficiently well. Yet only a few approaches incorporate a formal model. The novelty of our approach is the derivation of the full set of behavioral equilibria, i.e. all separating, pooling and semi-pooling equilibria of the signaling-extended PD. This would be only a technical note if it did not induce a richer set of equilibria concerning the distribution of an internalized norm that can stabilize cooperation. In particular, it is worthwhile to observe the existence of an inner equilibrium, i.e. an equilibrium where norm bearers and non-bearers coexist, that is stabilized by the interplay of a separating, a semi-pooling and a pooling equilibrium of the evolutionary signaling game. It is exactly this interplay that stabilizes the share of norm bearers and dissolves the necessity to introduce evolutionary forces into the dynamics of norm adoption beyond payoff monotonicity that are frequency based[22].

Since cooperative equilibria exist when agents may signal their cooperative attitude, large societies aiming for more cooperation are not completely limited to the reduction of anonymity in social interaction (and hence, giving up some of the advantages of large societies) or the use of formal institutions. Politicians may also try to provide hard-to-falsify signals of internal motives to cooperate in areas where interaction is rather anonymous. Then, informal institutions may spontaneously and easily evolve even in large unstructured interaction environments. Even if politics cannot alter the underlying incentives of the social dilemma to the extent that the dilemma aspect would indeed vanish, partial reduction of the incentive to defect or partial insurance for the suckers' payoff may be sufficient to allow for cooperation to evolve. The share of norm bearers in our model is driven by evolutionary forces that are beyond the scope of any policy measure. However, politics might have some leverage on how strong the internal sanctions are that support the norm once it is internalized. Hence, strengthening the internalized norms will also increase the chance for cooperation.

If we argue that it is foremost the spontaneous institutions that repel defection in large unstructured societies, then these insights lead us to argue that concepts of institutions should not require that all individuals adhere to the behavior prescribed by the spontaneous institution. Instead, a definition of institutions should allow for a substantial share of the population to deviate from its rule. We add a theoretical basis to this insight, which seems obvious from an empirical point of view.

We have not modeled the interplay of different PD situations in a society. Without going into any detail here, we conjecture from our signaling model that cooperation in one PD may serve as a

[22] Gintis et al. (2001) show in one of the few formal evolutionary signaling models that a stable separating equilibrium may exist. However, under general payoff monotonicity, this equilibrium would cease to exist since their type that corresponds to our high-types face an evolutionary advantage. As a consequence, their share of the population would increase and eventually exceed the threshold beyond which the separating equilibrium breaks down.

signal to have the internal cooperation in order to fare better in another PD. The temptation to defect in the first game would be the cost to falsely signal having the internal motivation to cooperate. Hence, the interplay between different PD situations does not allow for scaling up: temptation in the first game cannot be larger than in the second game, or cooperation there cannot be complete. Further research is needed on the details of the interplay between different PD games in an unstructured society.

The analysis for a more general norm than the one we considered is left open to future research. We believe that the size of the parameter measuring the strength of the internalized norm is not driven by evolutionary forces, since no fitness payoff differences depend on it. However, the size of the parameter does determine the range in which cooperative equilibria exist. Hence, if two separate populations with different levels of the internalized norms are considered, the one with the higher value is more likely to evolve towards a cooperative state. If in the course of time, both populations start interacting with each other, a cooperative population might induce cooperation in a defective population and vice versa. To analyze such an environment may be relevant for studying migrational effects on cooperation.

4. The evolution of inequality aversion in a simplified game of life.

4.1. Abstract

The increasing prominence of other-regarding preferences as an explanation for empirical and experimental findings calls for a rationalization of such preferences from an evolutionary perspective. The sensitivity of study results on the evolution of preferences with respect to the considered environment calls for an evolutionary approach that considers a compound environment, which comprises at best all relevant classes of environments. This paper attempts to address these two issues. I suggest a 2x2 simplified game of life that comprises a dilemma involving a coordination and distribution problem. An analysis of the separate environments makes strong predictions with respect to the advantageousness of inequality aversion. In particular, the global advantage in the dilemma and the global disadvantage in the problem of distribution are surprising. As expected, the simplified game of life gives rise to a greater variety in potential equilibrium distributions of preferences. In particular, the strong predictions for the single environments are put into perspective. Surprisingly, the expected stabilization of inner equilibria occurs only if the problem of coordination shows the same feature.

Keywords: inequality aversion – evolution

JEL Classifications: C72, C73

4.2. Introduction

At the latest with the seminal work of Fehr and Schmidt (1999) and Bolton and Ockenfels (2000) an other-regarding preference in the form of inequality aversion has become a prominent explanation for many empirical and experimental findings which departure from the prediction of standard economic theory. The increasing importance calls for a rationalization for such preferences, otherwise it may be regarded as a rather ad-hoc adjustment of preferences to explain empirical results. As Güth and Napel (2006) point out such preferences should in particular be compatible with the physical necessity to strive and compete for material rewards in an environment characterized by the scarcity in resources. In other words such preferences ought to be rationalizable from an evolutionary point of view.

Analyzing the evolution of preferences offers a unifying framework for traditional microeconomic analysis concerned with forward looking agents with fixed preferences on the one hand. And on the other, it incorporates evolutionary biology focusing on the interplay of the social or biological environment and the success of certain behavioral strategies in that environment. In the past the evolution of preferences has been studied in highly artificial single-game environments (e.g. Huck and Oechssler 1999; Koçkesen et al. 2000a, 2000b and Sethi and Somanathan 2001). As a consequence, these studies were inconclusive in explaining the presence of certain preferences, because the behavior induced by a certain preference might be advantageous in one environment, but disadvantageous in another. The agents' imperfect mental model of the world requires at least some link between the intrinsic motivations in different environments. Given this restriction, agents will be limited in the possibility to develop game-specific or role-specific preferences. Hence, the decentralized results for the single environments need to be combined to a centralized picture in order to explain the success or failure of behavioral determinants such as inequality aversion, reciprocity and truthfulness in the complex social and biological environment that comprises seemingly endlessly many of those small worlds, the 'game of life' (Güth and Napel 2006). I therefore in this paper address as a first aim the rationalizability of a preference for equality in an environment that contains the major classes of games constituting the game of life.

More recently, some attempts were made to analyze the evolution of preferences in more complex environments. Güth and Napel (2006) analyze how the personal characteristic of inequality aversion evolves in a setting containing two well-studied and characteristic games: the Ultimatum game and the Dictator game. Poulsen and Poulsen (2006) study the evolution of other-regarding preferences in an environment that comprises a simultaneous and a sequential Prisoners' Dilemma. Their analysis illustrates that the study of evolution of preferences in a compound strategic environment yields more interesting and intuitive results than game-specific analysis. However the considered environments are not meant to and indeed aren't even rough approximations of a game of life.

A prerequisite for the analysis of the evolution of preferences in the game of life is the structuring of the infinite set of potential games, which is the second aim of the paper. There is evidence that human behavior is not game-specific, but acts of men are similar in entire, quite general classes of games (see Yamagishi et al. 2013; Ashraf et al. 2006; Blanco et al. 2011; Chaudhuri and Gangadharan 2007 and Slonim and Garbarino 2008). This raises hope that the overwhelming complexity of the real world might be reducible to these classes when the evolution of preferences is considered. Many authors implicitly or explicitly share and express the view point that there are two fundamentally different societal problems (see e.g. Sugden 1986; Milgrom et al. 1990), problems of coordination and social dilemmas. Apart from these two classes, Schotter (1981), Ullmann-Margalit (1977) and others share the view that there is (at least) a third type of social problem, one of redistributive nature. A problem of distribution is characterized by unequal payoffs in equilibrium. The notion of a game of life I suggest will comprise these three classes of games.

As a first step to achieve the eager first goal I restrict in this paper to the class of 2x2 games. 2x2 games are omnipresent as they serve as the workhorses in applied game theory and their simplicity is their power as they combine remarkable diversity with minimal machinery. The eight numbers that represent such a game yield a class of 144 problems of remarkable richness and complexity (Robinson and Goforth 2005). Besides, the analysis will reveal that the 2x2 case is representative in uncovering the major forces that in their interplay will determine the distribution of inequality aversion in the population. Furthermore the purpose of the paper is to conduct an analysis for a world that is in some sense complete, i.e. to consider an environment that contains representatives of all classes present in my classification. In other words the focus of the paper in terms of generality is on completeness within a certain world of games (2x2 games) rather than on the world of games as such (e.g. all finite games). I consider this as a first step to explore the effects of considering a complete world, although restricted in size. I thus refine the first question in asking for the rationalizability of inequality aversion in what I will refer to as the 'simplified game of life'. With respect to the second goal although definitions are given for the 2x2 case the classification of games readily translates to all finite normal-form games.

The remainder of the paper proceeds as follows. In Section 4.3 the precise definitions for the games which are comprised in the simplified game of life will be given. The evolution of a preference for equality in material outcomes for each of the single-game environments is studied in Section 4.4. Thereafter the environment of the simplified game of life is considered in Section 4.5. Before I conclude in Section 4.7, I discuss the robustness of the results in Section 4.6.

4.3. Definition of terms

The informal classification of games given above in terms of issues of coordination, dilemma or redistribution is based on equilibrium considerations and will therefore depend on the equilibrium concept applied. The relevance of the chosen equilibrium concept stems not only from its consequences for the classification of games, but also from its implications for the study of evolution of preferences. A particular preference may influence the set of equilibrium outcomes that differs in accordance with the applied equilibrium concept, differently. As one of the standard solution concept, I will apply the notion of Nash equilibrium. The implication of applying different concepts is discussed in section 4.6. In the following I will give the formal definitions for three social problems for the 2x2 case. Note that the definitions in 4.3.1 and 4.3.2 readily extent to any finite normal-form game with N players.

4.3.1. Dilemma and Problem of coordination

In Milgrom et al. 1990 the complex institutional structure that facilitates agreements among US Congressmen is mentioned. The purpose of those institutions is either to facilitate coordination (Banks and Calvert 1989) or to prevent renegotiation on agreements (Weingast and Marshall 1988). If the conditions of the renegotiation-proofness principle (Hart and Tirole 1988) are violated the presence of renegotiation can restrict the set of achievable outcomes and might prevent the achievement of a Pareto-superior outcome. This aspect of prevention of Pareto-improvement is suggestive of what I have in mind talking about a dilemma. It is a non-cooperative strategic interaction between multiple agents with the property that there exists an outcome that is considered as advantageous by all agents but cannot be supported on purely egoistic grounds in the sense that once agreed upon a certain collective behavior some agents have an incentive to deviate from the implied behavior. In other words this superior outcome is not supported as equilibrium. Coordination problems are characterized by a non-dilemma situation with multiple equilibria.

Let $\gamma\left(A^1,A^2\right)$ denote a generic 2x2 game with strategy spaces $S^1=S^2=\{0,1\}\equiv S$ and payoffs $A^1=\left(a_{ij}^1\right)$ and $A^2=\left(a_{ij}^2\right)$, $(i,j)\in S\times S$ for player 1 and 2 respectively. Let ΔS represent the mixed extension of S. Finally I write the expected payoff of player 1 for a pair of mixed strategy as $\pi^1\left(s^1,s^2\right)=\sum_{i=0}^{1}\sigma_i^1 a_{(i,j)}^1$, $s^n=\left(\sigma_0^n,\sigma_1^n\right)\in\Delta S$ and $\pi^2\left(s^1,s^2\right)$ accordingly. The set of (pure) Nash equilibria of $\gamma\left(A^1,A^2\right)$ is denoted by $NE(\gamma)\left(NE^{pure}(\gamma)\right)$. For symmetric games we have $A^1=\left(A^2\right)^T\equiv A$ and I simply write $\gamma(A)$. Let $d_{(i,j)}\equiv\left|a_{(i,j)}^1-a_{(i,j)}^2\right|$ measure the absolute level of inequality if player one (two) plays $i(j)$. Finally, let $AP_{(i,j)} \equiv \frac{a^1_{(i,j)} + a^2_{(i,j)}}{2}$ denote the average payoff if player one (two) plays $i(j)$. Note that for symmetric games $d_{(i,j)} = d_{(j,i)}$ and $AP_{(i,j)} = AP_{(j,i)}$.

Definition A game $\gamma(A^1, A^2)$ is a *Dilemma* if

$$\exists(s^1,s^2) \in \Delta S^2 : \pi^n(s^1,s^2) > \pi^n\left((s^1,s^2)^*\right),\ n \in \{1,2\}, \forall (s^1,s^2)^* \in NE(\gamma)$$

In words, a game constitutes a dilemma if there exists a strategy profile such that the implied payoffs strictly Pareto-dominate the payoffs associated with the set of all Nash equilibria. Alternatively, in some sense on the other end of the spectrum, one could define a dilemma if there exists a Pareto-improvement for at least one Nash equilibrium. In the latter case a dilemma is present whenever from the perspective of a particular equilibrium there is a non-equilibrium Pareto-improvement. In contrast to such a definition, mine declares game to be a dilemma only if this holds for all equilibria, i.e. prior to the equilibrium selection. I consider the ex-ante viewpoint as more appropriate as it makes the classification of games and the analysis of the evolution of preferences less sensitive to assumptions regarding equilibrium selection. Furthermore, in the more general class of finite normal-form games the majority of games would constitute a social dilemma following the alternative definition. Consider for instance a game with Pareto-ranked equilibria and an inferior equilibrium being Pareto-dominated by some non-equilibrium outcome. A classification as a social dilemma appears unintuitive as the problem for this society is rather to coordinate on a Pareto-superior equilibrium. As problem of coordination are complementary to dilemmas and are characterized by the presence of multiple equilibria I define them as follows.

Definition A game $\gamma(A^1, A^2)$ is a *problem of coordination* if $|NE(\gamma)| > 1$ and there exists no non-equilibrium outcome which Pareto-dominates all of these equilibria.

Before I turn to problems of redistribution being asymmetric in nature I briefly want to elaborate on the structure of all symmetric 2x2 games. Note that all symmetric 2x2 games which neither constitute a dilemma nor a problem of coordination are exactly those with a unique equilibrium which is not Pareto-dominated by some non-equilibrium outcome. In the world of symmetric games such situations appear rather unproblematic since no dilemma, no coordination, and—as we will see—no problem of distribution is present. In other words, the set of symmetric games can be partitioned into three classes, dilemmas, problem of coordination s and unproblematic situations. In a symmetric world considering dilemmas and problem of coordination is thus in some sense complete as only unproblematic situations are excluded.

4.3.2. Problems of distribution

Before I give a precise definition of a "problem of distribution", it is necessary to clarify the intuition of such problems informally. First of all, any plausible definition of distributional concern is related to a notion of asymmetry in payoffs. Again, one could take an *ex-ante* or an *ex-post* point-of-view. With an *ex-post* point-of-view, a game would constitute a problem of distribution if the equilibrium played by the individuals shows asymmetric payoffs. From an *ex-*

ante perspective a game would constitute a problem of distribution if all equilibria would show asymmetric payoffs, all in favor of the same player. To illustrate the difference consider two situations. In the first strong individuals play against weak ones and all equilibria are characterized by higher payoffs for the stronger. Such a situation will not only by chance lead to asymmetric payoffs but it will do so systematically. In the second two identical individuals play a game with multiple equilibria, some of them favoring one individual, some favoring the other. In the latter case the game will only occasionally lead to asymmetries and whereas in the former case the game implies systematic asymmetries. It is more convincing, and in line with the corresponding decision with respect to the definition of social dilemmas, to take the *ex-ante* point-of-view.

Definition A game $\gamma\left(A^1, A^2\right)$ is a *problem of distribution* if $\left|NE(\gamma)\right| > 1$ and $\exists n \in \{1,2\}: \pi^n\left(\left(s^1, s^2\right)^*\right) > \pi^{-n}\left(\left(s^1, s^2\right)^*\right),\ \forall \left(s^1, s^2\right)^* \in NE(\gamma)$.

The qualification in the definition for $\gamma\left(A^1, A^2\right)$ to have multiple equilibria is made for simplicity only. I will refer to those individuals (dis)favored in the problem of distribution as (low) high types.

4.3.3. Inequality aversion

In Sections 4.4 and 4.5 on the evolution of inequality aversion, I will make use of the standard evolutionary model, which is concerned with a large population. This population is structured by personal characteristics and by the way individuals are matched. With respect to the former there are two sources of heterogeneity among individuals. The population is on the one hand divided into two subpopulations that correspond to the two different roles assigned in the problem of distribution. On the other hand there is heterogeneity with respect to the evaluation of payoff distributions, i.e. agents show different levels of inequality aversion. Inequality aversion is modeled as follows. I will apply the definition suggested by (Fehr and Schmidt 1999) which in a 2x2 setting amounts to $u^n_{(i,j)} = a^n_{(i,j)} - \sigma^n \max\left\{a^{-n}_{(i,j)} - a^n_{(i,j)}, 0\right\} - \omega^n \max\left\{a^n_{(i,j)} - a^{-n}_{(i,j)}, 0\right\},\ \sigma^n, \omega^n \in [0,1]$, i.e. σ^n and ω^n measure the degree of aversion of player n to inequality which disfavors or, respectively, favors him. I make the simplifying assumption that $\sigma^n = \omega^n \equiv \theta^n$. The qualitative implication of a relaxation of this assumption is discussed in section 4.6. Hence, inequality aversion is parameterized by the one dimensional space $[0,1]$. At time t agents' preference regarding equality in material payoffs is distributed over $[0,1]$ according to the distribution function F^t_H and F^t_L for high types and low types, respectively. Initially, the density functions corresponding to F^t_H and F^t_L are assumed to have full support. I will drop the superscript t to represent equilibrium distributions, i.e. $F_{H,L} = \lim_{t\to\infty} F^t_{H,L}$.

4.3.4. The simplified game of life

As I will elaborate more deeply in the subsequent analysis, inequality aversion transforms the game $\gamma\left(A^1, A^2\right)$ into the game $\gamma\left(U^1, U^2\right)$. The latter and the former may well differ in the set of Nash equilibria. To ease reading and interpretation, I will make use of the following definitions.

Definition I say that an equilibrium (i, j) in the game $\gamma\left(A^1,A^2\right)$ is *contested* by player 1(2) if $\exists u^1_{(\neg i,j)} > u^1_{(i,j)} \left(u^2_{(i,\neg j)} > u^2_{(i,j)}\right)$, i.e. strategy *i(j)* loses its property of being a best response to strategy *j(i)* in the game $\gamma\left(U^1,U^2\right)$. An equilibrium in the game $\gamma\left(A^1,A^2\right)$ is *contestable*, if it may be contested by at least one player I say that the strategy pair (i, j) is *stabilizable* if it is an equilibrium of $\gamma\left(U^1,U^2\right)$ for some levels of θ^1 and θ^2.

Note that if an equilibrium is contested by some player it is contested by any player who shows a weakly higher degree of inequality aversion.

To simplify the analysis of the simplified game of life I will restrict the included games in a way which ensures that in the game $\gamma\left(U^1,U^2\right)$ no situation with a unique mixed Nash equilibrium will occur. Since a unique mixed Nash equilibrium in particular arises if a player who contests all pure Nash equilibria is matched with a purely selfish player, I make the following definition.

Definition A game $\gamma\left(A^1,A^2\right)$ is called *strict* if there is no player who can contest all equilibria.

The term "strict" as defined in here parallels the concept of strict equilibrium since a player will not be able to contest all equilibria if at least one equilibrium is sufficiently strict for him, i.e. the material loss from unilateral deviations is sufficiently high. Note that if inequality aversion has a leverage on strict games it will do so for games that are not strict. Note further that in general finite normal-form games this condition will be satisfied in the majority of the cases. Allowing the play of mixed equilibria has interesting consequences on the sharpness of the prediction regarding the stable distributions of preferences though. This will be outlined in section 4.6.

Based on the classifications of social problems in section 4.3.1-4.3.2 and the definition given above, I am now able to define an environment that comprises all these classes.

Definition The *simplified game of life* is a game that comprises a symmetric dilemma, a strict symmetric problem of coordination and a strict problem of distribution.

The qualification for the dilemma and the problem of coordination to be symmetric is made in order to isolate the effects that the asymmetry of the problem of distribution implies. Strictness of an arbitrary game $\gamma\left(A^1,A^2\right)$ either implies the existence of multiple equilibria or the unique equilibrium in mixed or pure strategies is not contestable. If the unique equilibrium is in mixed strategy then no non-equilibrium outcome can be stabilized without contesting the mixed equilibrium, hence for strict games no evolutionary pressure that favors or disfavor a preference for equality will emerge. If the unique equilibrium is realized in pure strategies then a bilateral deviation could be stabilized by inequality-averse players. Essentially this is the only case that is excluded by the assumption of multiplicity of equilibria in the definition of a problem of distribution. Strictness for the dilemma is not required as this class of games will not show mixed play. In 4.3.1 I argued that a classification into dilemmas and problems of coordination is in some sense complete by partitioning the set of all symmetric games with their complement reflecting rather unproblematic situations. Requiring strictness for problems of coordination limits to some extent this completeness. Problems of coordination that are not strict are thereby excluded from analysis. The only difference between strict problems of coordination and such problems that are not strict is that in the latter in a match of an individual with a very high degree of inequality aversion and an individual with a very low degree of inequality aversion a unique Nash equilibrium in mixed strategy exists. I refer the reader again to the discussion in Section 4.6.

4.3.5. Evolutionary framework

Before I can start with the evolutionary analysis, the analytical framework needs to be set up. In what follows I will state the assumptions I make with respect to informational aspects, the matching process, evolutionary dynamics and the applied stability concept.

I assume that agents can mutually observe their attitude towards unequal payoff distributions. This assumption could be weakened to an awareness of the inequality aversion in a positive fraction of interactions, the availability of sufficiently accurate signals or sufficiently cheap screening technologies (see Güth 1995; Sethi and Somanathan 2001; Güth et al. 2003). With respect to matching consider the following procedure. First a random draw selects among the three types of games that constitute the simplified game of life. In case of a dilemma or a problem of coordination individuals from the total population are randomly matched into pairs playing the selected game. Thereby each pair has the same probability in each short period of time. The interaction in the problem of distribution will be modeled as a 2-population model (see e.g. Weibull 1997), i.e. individuals interact across populations but not within. Again, each pairing has the same probability, relative size of the subpopulations of high and low types matters for expected payoffs though. If for instance the subpopulation of low types is ten times as large as the subpopulation for high types then any high type will play ten times as often as a low type. This will however only amplify the advantage or disadvantage of high types over low types. For notational simplicity I may thus assume that the two subpopulations are equal in size. Payoffs given by A^1 and A^2 represent the material payoffs of the stage game that will be decisive with respect to evolutionary success.

Whereas the belonging to one of the subpopulations due to role assignment in the problem of distribution is exogenous and common knowledge, the distribution of inequality-averse individuals in each of the two subpopulations is endogenous. Since inequality aversion reflects a particular evaluation of material payoffs, I will apply the indirect evolutionary approach pioneered by (Güth and Yaari 1992)[23], i.e. preferences determine behavior and behavior in turn determines fitness. Fitness measured by material payoffs will determine the evolution of F^t. The evolutionary process is modeled by payoff monotone selection dynamics[24] (see e.g. Weibull 1997). With respect to stability I will apply the concept of asymptotic stability (see. e.g. Samuelson 1997 for definitions). An asymptotically stable equilibrium will be reconstituted as

[23] The indirect evolutionary approach has been applied in various strategic settings (ultimatum game, Huck and Oechssler 1999) or to analyze the evolutionary stability of altruistic preferences (Bester and Güth 1998) or of altruistic and spiteful preferences (Possajennikov 2000).

[24] There are other forces than evolutionary selection shaping individual preferences. Bisin and Verdier (2001) for instance study intergenerational cultural transmission mechanisms.

time approaches infinity after a small but – in terms of the composition of mutation-strategies – arbitrary perturbation.

Since I am concerned with games that allow for multiple equilibria an assumption with respect to equilibrium selection needs to be made. An appropriate equilibrium selection criterion should not *a priori* favor or disfavor a preference for equality with respect to evolutionary success. I therefore assume that if $\gamma\left(U^1,U^2\right)$ has multiple pure-strategy Nash equilibria, then players randomize over all pure-strategy Nash equilibria with equal probability. It turns out that symmetry of the probability distribution over the set of pure Nash equilibria is necessary and sufficient for the selection criteria to satisfy the requirement to be neutral with respect to the evolutionary advantageousness of inequality aversion for all games considered (see discussion in 4.6.1). Since 2x2 games show at most two pure Nash equilibria symmetry amounts to uniformity. To clarify, it is not the players who randomize over strategies of different pure Nash equilibria independently, but pairs of players randomize jointly over the set of pure Nash equilibria. If for instance $\gamma\left(U^1,U^2\right)$ has two pure Nash equilibria then a given pair of players will play each of the two with probability one-half. To put it differently, individuals are assumed to play the correlated equilibrium that is the linear combination with equal weights of the two correlated equilibria that correspond to the two pure Nash equilibria (see 4.6.2).

Let $\Gamma,\left(\Gamma_{sym}\right)$ denote the set of (symmetric) 2x2 games, $\Gamma^{\circ},\left(\Gamma^{\circ}_{sym}\right)$ the set of (symmetric) 2x2 games with neither weakly nor strictly dominated strategies. Games with weakly dominated strategies can be treated as the limiting case of games in Γ°. More precisely as $\Gamma\subset\mathbb{R}^8,\left(\Gamma_{sym}\subset\mathbb{R}^4\right)$ the subset of $\Gamma,\left(\Gamma_{sym}\right)$ containing no weakly dominated strategies is dense in $\Gamma,\left(\Gamma_{sym}\right)$ according to the Euclidean norm. Since the critical level of inequality aversion are continuous in the parameters of a game $\gamma\left(A^1,A^2\right)$, the results for any game with weakly dominated strategies are a limit case of games in $\Gamma^{\circ},\left(\Gamma^{\circ}_{sym}\right)$[25]. With this technical note in mind I can concentrate on games with no weakly dominated strategies.

In the next section I will study the evolution of the trait of inequality aversion in each of the three games separately. In section 4.5 I will contrast those results with the analysis in the compound environment of the simplified game of life.

4.4. Inequality aversion in the separate environments

Symmetric dilemma Note that for symmetric games there is always an equilibrium in pure strategies. Furthermore games with multiple equilibria are free of the dilemma property. To see this note that a dilemma requires the existence of a non-equilibrium outcome that Pareto-dominates all Nash equilibria. That is, such a pair off payoffs must yield higher payoffs than in

[25] More precisely the mapping $\Phi:\Gamma\to\mathbb{R}$ which assigns to any game the critical value $\theta^{D,C,R}$ (see Section 4.4) is continuous.

any Nash equilibrium. Consider a symmetric games with two pure Nash-equilibria. A necessary condition for such a game to constitute a social dilemma would be that there is an outcome in pure strategies that gives each player more than the maximum of the two Nash equilibria in pure strategies. But the existence of such an outcome violates the Nash-equilibrium property in the first place, because in 2x2 games this implies the existence of an alternative reply with higher payoffs than in equilibrium. Hence a symmetric social dilemma must be in the set $\Gamma_{sym} \setminus \Gamma_{sym}^{\circ}$, the set of games with weakly or strictly dominated strategies. If a player has a strictly dominant strategy then by symmetry his opponent has the same strictly dominant strategy. As a unilateral deviation from equilibrium can never lead to a strict Pareto-improvement, only the symmetric non-equilibrium outcome realized by bilateral deviation can yield strictly higher payoffs for both players. That is in 2x2 games a symmetric dilemma corresponds to the classical Prisoners' Dilemma. Lemma 4-1 summarizes this insight. All proofs are given in Appendix C.

Lemma 4-1 Let $\gamma(A) \in \Gamma_{sym}$. $\gamma(A)$ constitutes a dilemma if and only if $\gamma(A)$ is strictly dominance-solvable by the unique symmetric Nash equilibrium (i^*, i^*) and $AP_{(i^*,i^*)} < AP_{(\neg i^*,\neg i^*)}$.

Dominance solvability implies that the only stabilizable outcome is the symmetric non-equilibrium outcome. Lemma 4-2 states the conditions on the required degree of inequality aversion for this to be the case.

Lemma 4-2 Let $\gamma(A) \in \Gamma_{sym}$ be a social dilemma and (i^*, i^*) be its unique equilibrium. Then the only pair of strategies $(i, j) \neq (i^*, i^*)$ that is stabilizable is $(\neg i^*, \neg i^*)$. $(\neg i^*, \neg i^*)$ is stabilizable if and only if $\theta^1, \theta^2 \geq \theta^D \equiv \dfrac{a_{(\neg i^*, i^*)} - a_{(\neg i^*, \neg i^*)}}{d_{(i^*, \neg i^*)}} \in [0,1]$.

Lemma 4-2 states that whenever two sufficiently inequality-averse players interact, the symmetric non-equilibrium outcome in $\gamma(A)$ constitutes an equilibrium in $\gamma(U)$. The threshold θ^D has a straight forward economic meaning. Since $a_{(\neg i^*, i^*)} - a_{(\neg i^*, \neg i^*)}$ measures the material gain of deviating from the non-equilibrium pair of strategies $(\neg i^*, \neg i^*)$ and $d_{(i^*, \neg i^*)}$ measures the implied loss in equality induced by such a deviation, θ^D measures the material price per unit of equality gained. Sufficient inequality aversion therefore translates into a sufficient willingness to pay for equality. Given the characterization of social dilemmas in Lemma 4-1 and the characterization of stabilizable strategy profiles in $\gamma(U)$ in Lemma 4-2, Proposition 4-1 characterizes the stable distributions of inequality aversion.

Proposition 4-1 Let $\gamma(A) \in \Gamma_{sym}$ be a social dilemma. If $(\neg s_1^*, \neg s_2^*)$ is stabilizable, then there exists a $\theta^D \in [0,1]$, such that the globally stable equilibrium is $F(\theta^D) = 0$,[26] furthermore the

[26] As in the dilemma and problem of coordination players role is symmetric, I will in the corresponding subsections drop the subscripts reflecting types in the problem of redistribution.

material advantage of sufficiently inequality-averse individuals is increasing in the share of individuals with $\theta \geq \theta^D$, i.e. $\operatorname{sgn}\left(\Pi^{\theta \geq \theta^D} - \Pi^{\theta < \theta^D}\right)_{\left(1-F\left(\theta^D\right)\right)} = 1$, where $\left(\Pi^{\theta \geq \theta^D} - \Pi^{\theta < \theta^D}\right)_{\left(1-F\left(\theta^D\right)\right)}$ denotes the derivative w.r.t. $1 - F\left(\theta^D\right)$, the share of inequality-averse individuals. Otherwise the share of inequality-averse individuals is determined by initial conditions and random shift.

In the following paragraphs I will give the intuition behind this result. The potential for an evolutionary advantage of inequality-averse individuals stems from that fact that a pair of sufficiently inequality-averse players will be able to transform the social dilemma into a coordination game. In a symmetric dilemma "sufficiently high inequality aversion" translates to $\theta \geq \theta^D$, where θ^D measures the ratio of the material incentive to deviate from the Pareto-superior outcome and the potential loss in equality stemming from such a deviation. In other words, if both players have an aversion against inequality larger than θ^D, the material gain of an deviation from the symmetric diagonal outcome is more than compensated for in utility terms by the loss in equality. Thereby a match of two such individuals transforms the dilemma into a problem of coordination. By definition of the dilemma the stabilized outcome yields Pareto-superior payoffs which benefits inequality-averse individuals as they randomize over all pure Nash equilibria.

Symmetric problem of coordination In games within the set of $\mathring{\Gamma}_{sym}$ which show multiple pure-strategy Nash equilibria either the two diagonal symmetric payoff-pairs or the two off-diagonal asymmetric payoff-pairs constitute the Nash equilibrium payoffs.

Lemma 4-3 Let $\gamma(A) \in \Gamma_{sym}$. $\gamma(A)$ constitutes a problem of coordination if and only if (1) $NE^{pure}\left(\gamma(A)\right) = \left\{(i,i)\right\}$ or (2) $NE^{pure}\left(\gamma(A)\right) = \left\{(i,j) \middle| i \neq j\right\}$.

Before I will characterize the stable distribution of inequality aversion in Proposition 4-2, I will define a threshold θ^C which is the equivalent to θ^D in the symmetric dilemma. However, in the symmetric coordination game each of the off-diagonal equilibria of $\gamma(A)$ may be contestable for both players, hence θ^C will be the minimum of the two ratios measuring the material price per unit of equality gained for player one and two. These prices may differ as equilibria in $\gamma(A)$ are asymmetric and players face different incentives to deviate. Formally, $\theta^C \equiv \min_{(i,j)\in NE^{pure}(\gamma(A))} \left\{ \frac{a_{(i,j)} - a_{(i,i)}}{d} \right\}$, where $d \equiv d_{(i,j)}\Big|_{(i,j)\in NE^{pure}(\gamma(A))}$.

Let AP measure the average payoff of the equilibria in $\gamma(A)$, i.e. $AP \equiv AP_{(i,j)}\Big|_{(i,j)\in NE^{pure}(\gamma(A))}$. For ease of readability I will refer to individuals with $\theta \geq \theta^C$ $\left(\theta < \theta^C\right)$ as inequality-averse individuals and selfish players respectively.

Proposition 4-2 Let $\gamma(A) \in \Gamma_{sym}$ be a strict problem of coordination. Then:

If $NE^{pure}(\gamma(A)) = \{(i,i)\}$ or $NE^{pure}(\gamma(A)) = \{(i,j) | i \neq j\}$ and none of the material equilibria is contestable then the share of inequality-averse individuals in the population is determined by initial conditions and random shift.

If equilibria are contestable then:

1. if the destabilized equilibrium is materially favorable for inequality-averse individuals then the globally stable equilibrium is characterized by $F(\theta^C) = 1$, furthermore $\operatorname{sgn}\left(\Pi^{\theta \geq \theta^C} - \Pi^{\theta < \theta^C}\right)_{(1-F(\theta^C))} \in \{-1,0,1\}$.
2. if the destabilized equilibrium is materially favorable for selfish individuals then the globally stable equilibrium is characterized by $F(\theta^C) = \theta^C \dfrac{d}{AP - AP_{(i,i)}}$, furthermore $\operatorname{sgn}\left(\Pi^{\theta \geq \theta^C} - \Pi^{\theta < \theta^C}\right)_{(1-F(\theta^C))} = -1$.

where $AP_{(i,i)}$ is the average payoff of the outcome that is stabilizable by two sufficiently inequality-averse individuals.

In case (1) of Lemma 4-3 the material equilibria are not contestable as any deviation from symmetric material payoffs not only reduces material payoff but also increases inequality. As a consequence no evolutionary pressure will emerge favoring or disfavoring inequality aversion. In case (2) this is not necessarily true. With respect to utility a deviation from materially asymmetric payoffs associated with a gain in equality might outweigh the material loss from deviation. Proposition 4-2 reveals that in strict problem of coordination s a strong preference for equality is weakly disadvantageous from an evolutionary point of view. Intuitively, strictness of the problem of coordination excludes the possibility of both equilibria being destabilized. If one equilibrium is contestable then it is destabilized by sufficiently inequality-averse agents. If the destabilized equilibrium is materially favorable for inequality-averse individuals then not only they suffer from deviating from material equilibrium, but lose relative to more selfish individuals as that equilibrium is destabilized where they gain more than selfish players. As a consequence individuals with a strong preference for equality face an evolutionary disadvantage and will become extinct. If the reverse is true, i.e. the destabilized equilibrium is favorable for selfish players then the disadvantage from unilaterally deviating from material equilibria is partially compensated by no longer playing a disadvantageous equilibrium and thereby increasing average payoffs. However, this effect diminishes as the share of sufficiently inequality-averse agents increases. This stabilizes a distribution of preference where selfish and inequality-averse (relative to θ^C) individuals coexist.

Problem of distribution Lemma 4-4 below characterizes problems of redistribution and differentiates two cases which will become relevant in the course of the argument.

Lemma 4-4 Let $\gamma\left(A^1,A^2\right)\in\Gamma$. $\gamma\left(A^1,A^2\right)$ constitutes a strict problem of distribution if and only if all Nash equilibria favor the same individual and:

(1): $\gamma\left(A^1,A^2\right)$ has multiple equilibria which are not Pareto-ranked.

(2): $\gamma\left(A^1,A^2\right)$ has multiple equilibria which are Pareto-ranked.

Let $\widehat{\theta}^R_{H,L}\left(\breve{\theta}^R_{H,L}\right)$ denote the thresholds for the high and low type respectively such that the more (less) equal material equilibrium is destabilized, which requires of course that the considered equilibrium may be contested by the player. To give a formal definition requires a lot of complicated notation and is not very insightful. The formal definition is given in Appendix C for the representative case in the proof of Proposition 4-3. The economic meaning of the thresholds is the same as for the thresholds in the problem of coordination or the dilemma. The critical values for the inequality aversion of players relate the material incentive and the gain in equality induced by an unilateral deviation from an equilibrium, i.e. they measure the price of deviation per unit equality gained. Let $\theta^R_{H,L}=\min\left\{\widehat{\theta}^R_{H,L},\breve{\theta}^R_{H,L}\right\}$, i.e. the type-contingent threshold $\theta^R_{H,L}$ plays the same role as θ^D and θ^C in the symmetric dilemma and the symmetric problem of coordination respectively, i.e. if the degree of inequality aversion for at least one player exceeds $\theta^R_{H,L}$ then at least one of the equilibria of $\gamma\left(A^1,A^2\right)$ loses its equilibrium property in $\gamma\left(U^1,U^2\right)$.

Proposition 4-3 Let $\gamma\left(A^1,A^2\right)$ constitute a strict problem of distribution.

1. If one of the material equilibria is contestable by low types, the unique globally stable equilibrium distribution is characterized by a homomorphic population with only inequality-averse individuals. $F_L\left(\theta^R_L\right)=0,\ \operatorname{sgn}\left(\Pi_L^{\theta\geq\theta^R_L}-\Pi_L^{\theta<\theta^R_L}\right)_{\left(1-F_H\left(\theta^R_H\right)\right)}\in\{-1,0,1\}$.
2. If one of the material equilibria is contestable by high types, with one exception the globally stable equilibrium distribution is characterized by
$$F_H\left(\theta^R_H\right)=1,\ \operatorname{sgn}\left(\Pi_H^{\theta\geq\theta^R_L}-\Pi_H^{\theta<\theta^R_L}\right)_{\left(1-F_L\left(\theta^R_L\right)\right)}=1.$$
The exception arises in case of two Pareto-ranked equilibria (case (2) of Lemma 4-4) with the Pareto-inferior equilibrium being contestable for both types. In that case the globally stable equilibrium distribution is characterized by
$$F_H\left(\theta^R_H\right)=0,\ \operatorname{sgn}\left(\Pi_H^{\theta\geq\theta^R_L}-\Pi_H^{\theta<\theta^R_L}\right)_{\left(1-F_L\left(\theta^R_L\right)\right)}=-1.$$
Otherwise the distribution is determined by initial conditions and random shift.

Note that the payoff differences for both types depend on the share of inequality-averse individuals in the subpopulation of the other types as there is no interaction within subpopulation, but only across them.

To see the intuition behind Proposition 4-3 I first elaborate on case (1) of Lemma 4-4. In case (1) one of the pure strategy equilibria shows strictly less inequality. Hence the more (less) unequally distributed equilibrium is preferred by the high (low) type. It turns out that for the high type the less unequally distributed equilibrium is never contestable. As a consequence the condition on the problem of distribution to be strict is essentially a condition with respect to the payoffs of the low type.

I first consider the case where the more equally distributed equilibrium is also not destabilized by the low type as in the first case of Proposition 4-3. If on the one hand the more unequal equilibrium is destabilized by both players, then the more equally distributed equilibrium will become the unique equilibrium. In that case such high types will with certainty play the less favorable equilibrium of $\gamma\left(A^1, A^2\right)$ and face an evolutionary disadvantage. Furthermore the extent of the disadvantage for the high types increases with the share of sufficiently inequality-averse low types since more and more often they will end up in playing the relative unfavorable equilibrium. The reverse argument applies for the low types. If on the other hand the more unequal equilibrium is only destabilized by the high type the same argument applies for the high types but the disadvantage is now independent of the share of inequality-averse low types as their best response behavior is not altered by inequality aversion.

I second consider the case where the more equal distributed equilibrium is destabilized by the low type as in the second case of Proposition 4-3. If high types destabilize the more unequally distributed equilibrium then this will result in an evolutionary disadvantage as the relatively less favorable equilibrium will be selected. As no player can destabilize all equilibria inequality-averse low types will face an evolutionary disadvantage as they destabilize the relative favorable of the two pure Nash equilibria in $\gamma\left(A^1, A^2\right)$. In all other cases the distribution of the preference parameter is undetermined. The major difference between case (1) and (2) of Lemma 4-4 responsible for the deviations in equilibrium distribution stems from the following fact. In case (1) of Lemma 4-4 the less unequally distributed equilibrium which is relative less favorable for the high type was not contestable. In case (2) however it is the Pareto-superior equilibrium which is not contestable. In this difference lies the potential for an evolutionary advantage of inequality-averse individuals among high types.

Before I turn to the analysis in the simplified game of life I briefly summarize the results obtained so far (see also Figure 4-1). The analysis in the separate environments revealed that if inequality aversion has a leverage on the set of equilibria played then inequality aversion enjoys a global evolutionary advantage over more selfish preferences in a dilemma. In the class of problems of coordination inequality aversion surprisingly faces a weak evolutionary disadvantage in the sense that at most a stable inner equilibrium exists where relative inequality-averse and relative selfish players coexist, in all other cases relative inequality-averse players will eventually disappear. In the problem of distribution evolutionary selection dynamics will always favor the preference for equality among the disfavored individuals. Among the individuals favored by the problem of distribution in all cases except for one inequality aversion will eventually disappear.

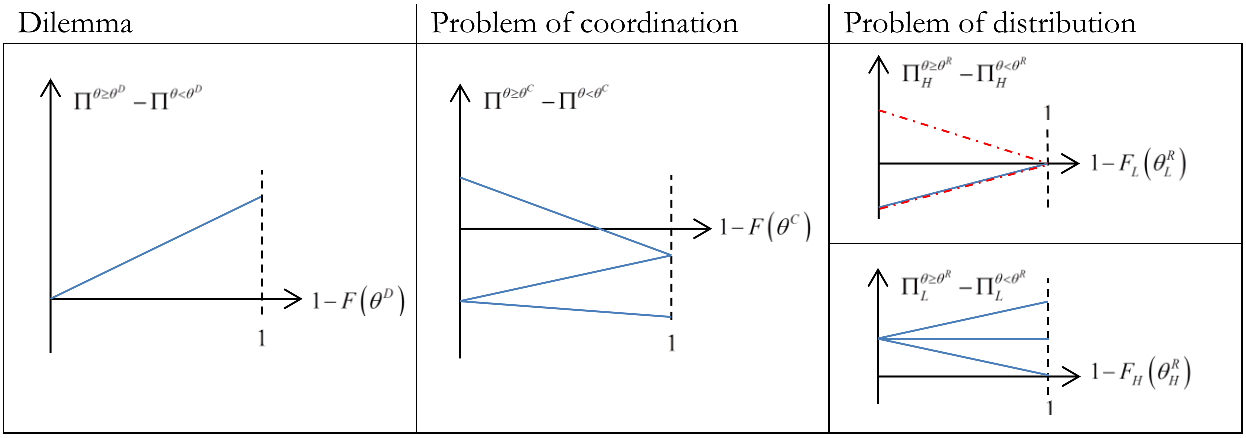


Figure 4-1: Differences in material payoffs in the games constituting the simplified game of life. For high types in the problem of distribution the blue lines correspond to case A, the red dash-dotted lines to case (2) of Lemma 4-4.

Note that the three characteristics: the slope, the intercept and having a root in the open unit interval gives rise to eight different loci of the linear payoff differences[27]. Remarkably, the analysis so far predicts that for a single environment at most three of them are needed to describe the differences in payoffs between inequality-averse and selfish individuals (see Figure 4-1).

4.5. Evolution of inequality aversion in the 2x2 simplified game of life

In this section we will analyze the interplay of the different types of interaction present in the simplified game of life. For ease of exposition I will assume that the thresholds of the single environments coincide, i.e. $\theta^D = \theta^C = \theta^R_{H,L} \equiv \theta^{crit}$. The profit for an individual in the simplified game of life is simply the weighted average of the profits earned in the single environments[28], i.e.:

$$
\begin{aligned}
\Pi^{S,\theta\geq\theta^{crit}}_{H,L} &= \mu\, \Pi^{D,\theta\geq\theta^{crit}}\left(F^t\left(\theta^{crit}\right)\right) + \nu\, \Pi^{C,\theta\geq\theta^{crit}}\left(F^t\left(\theta^{crit}\right)\right) + (1-\mu-\nu)\Pi^{R,\theta\geq\theta^{crit}}_{H,L}\left(F^t_{L,H}\left(\theta^{crit}\right)\right) \\
\Pi^{S,\theta<\theta^{crit}}_{H,L} &= \mu\, \Pi^{D,\theta<\theta^{crit}}\left(F^t\left(\theta^{crit}\right)\right) + \nu\, \Pi^{C,\theta<\theta^{crit}}\left(F^t\left(\theta^{crit}\right)\right) + (1-\mu-\nu)\Pi^{R,\theta<\theta^{crit}}_{H,L}\left(F^t_{L,H}\left(\theta^{crit}\right)\right)
\end{aligned} \tag{4.1}
$$

Hence payoff differences are given by[29]:

$$
\begin{aligned}
&\Pi^{S,\theta\geq\theta^{crit}}_{H,L} - \Pi^{S,\theta<\theta^{crit}}_{H,L} = \\
&\mu\left(\underbrace{\Pi^{D,\theta\geq\theta^{crit}} - \Pi^{D,\theta<\theta^{crit}}}_{\geq 0}\right) + \nu\left(\underbrace{\Pi^{C,\theta\geq\theta^{crit}} - \Pi^{C,\theta<\theta^{crit}}}_{\leq 0}\right) + (1-\mu-\nu)\left(\underbrace{\Pi^{R,\theta\geq\theta^{crit}}_{H,L} - \Pi^{R,\theta<\theta^{crit}}_{H,L}}_{H:\leq 0^*,\, L:\geq 0}\right)
\end{aligned} \tag{4.2}
$$

Let $d\Pi$ denote the difference in payoffs between relatively inequality-averse and selfish players. Equations in can now be expressed in a more compact way as:

---

[27] The eight cases refer to a positive or negative function that is increasing or decreasing or a function with a root in the open unit interval that shows negative or positive slope.

[28] D – dilemma; C – problem of coordination ; R – problem of distribution; S – simplified game of life.

[29] The asterisk in equation and refers to the exception in case (2) of Lemma 4-4 in which also among high types inequality-averse individuals enjoy an evolutionary advantage.

$$d\Pi_{H,L}^{S}=\mu\left(\underbrace{d\Pi^{D}\left(1-F_{H}^{t}-F_{L}^{t}\right)}_{\geq 0}\right)+\nu\left(\underbrace{d\Pi^{C}\left(1-F_{H}^{t}-F_{L}^{t}\right)}_{\leq 0}\right)+\left(1-\mu-\nu\right)\left(\underbrace{d\Pi_{H,L}^{R}\left(1-F_{L,H}^{t}\right)}_{H:\leq 0^{*},\ L:\geq 0}\right) \tag{4.3}$$

Note that whereas the differences in the dilemma and the problem of coordination depend on the total share of inequality-averse individuals in population the according difference in payoffs for the problem of distribution depends only on the share in the subpopulation of the opposite type. Making use of the linearity of the payoffs differences I can write (4.3) in the following way:

$$\begin{aligned} d\Pi_{H}^{S} &= \mu\beta^{D}\left(1-F_{H}^{t}-F_{L}^{t}\right)+\nu\left(\alpha^{C}+\beta^{C}\left(1-F_{H}^{t}-F_{L}^{t}\right)\right)+\left(1-\mu-\nu\right)\left(\alpha_{H}^{R}+\beta_{H}^{R}\left(1-F_{L}^{t}\right)\right) \\ d\Pi_{L}^{S} &= \mu\beta^{D}\left(1-F_{H}^{t}-F_{L}^{t}\right)+\nu\left(\alpha^{C}+\beta^{C}\left(1-F_{H}^{t}-F_{L}^{t}\right)\right)+\left(1-\mu-\nu\right)\left(\alpha_{L}^{R}+\beta_{L}^{R}\left(1-F_{H}^{t}\right)\right) \\ &= d\Pi_{H}^{S}+\left(1-\mu-\nu\right)\underbrace{\left(\alpha_{L}^{R}+\beta_{L}^{R}\left(1-F_{H}^{t}\right)-\alpha_{H}^{R}-\beta_{H}^{R}\left(1-F_{L}^{t}\right)\right)}_{\geq 0^{*}} \end{aligned} \tag{4.4}$$

, where $\alpha^{D}=0,\alpha^{C},\alpha_{H,L}^{R}$ and $\beta^{D},\beta^{C},\beta_{H,L}^{R}$ denote intercepts and slops of $d\Pi^{D},d\Pi^{C},d\Pi_{H,L}^{R}$ respectively.

If in the case with two Pareto-ranked equilibria (case (2) of Lemma 4-4) the Pareto-inferior equilibrium is destabilized by the low type, then inequality-averse players are favored also among high types. In that case if the problem of coordination is not played too often or involved differences in payoffs are comparably small inequality-averse players in both sub-populations face an evolutionary advantage. In other words, the globally stable equilibrium distribution will be characterized by $F_{H,L}\left(\theta^{crit}\right)=0$, i.e. population will consist of inequality-averse individuals only. I therefore concentrate in the following on the non-exceptional cases with a problem of distribution being accompanied with a global disadvantage of inequality-averse players among high types. Note that in that case $\beta_{H}^{R}=-\alpha_{H}^{R}$ (see Figure 4-1). Note further that since low types and high types earn the same profits in the dilemma and the problem of coordination a positive payoff difference for high types implies a positive difference for low types (see (4.4)). This has the immediate consequence that a locally stable equilibrium characterized by $F_{H}\left(\theta^{crit}\right)=0, F_{L}\left(\theta^{crit}\right)=1$, i.e. an equilibrium with only inequality-averse high types and only selfish low types does not exist in the simplified game of life.

The following theorem characterizes the equilibria that may emerge in the simplified game of life for the predominant case of a problem of distribution which is disadvantageous for inequality-averse high types. For ease of readability I abbreviate $F_{H}\left(\theta^{crit}\right)=F_{H},F_{L}\left(\theta^{crit}\right)=F_{L}$.

Theorem Let $\theta^D = \theta^C = \theta^R_{H,L} \equiv \theta^{crit}$ and $d\Pi^R_H \le 0$ then the set of equilibrium distributions of a preference for equality is characterized by:

$$F_L = 0, F_H = \begin{cases} 0 & \alpha^C > -\dfrac{\mu\beta^D + \nu\beta^C}{\nu} \\ 1 + \dfrac{\nu\alpha^C}{\mu\beta^D + \nu\beta^C} & 0 \le \alpha^C \le -\dfrac{\mu\beta^D + \nu\beta^C}{\nu} \end{cases}$$

$$F_H = 1,\ F_L = \begin{cases} 0 & -\dfrac{1-\mu-\nu}{\nu}\alpha^R_L \le \alpha^C \le 0 \\ 1 - \dfrac{\nu\alpha^C}{\mu\beta^D + \nu\beta^C} + \dfrac{1-\mu-\nu}{\mu\beta^D + \nu\beta^C}\alpha^R_L & -\dfrac{1-\mu-\nu}{\nu}\alpha^R_L + \dfrac{\mu\beta^D + \nu\beta^C}{\nu} < \alpha^C < -\dfrac{1-\mu-\nu}{\nu}\alpha^R_L \\ 1 & \alpha^C \le -\dfrac{1-\mu-\nu}{\nu}\alpha^R_L + \dfrac{\mu\beta^D + \nu\beta^C}{\nu} \end{cases}$$

Figure 4-2 illustrates the set of equilibria graphically. Only if the advantage of inequality-averse individuals increases or the disadvantage decreases in the share of inequality-averse individuals when the dilemma and the problem of coordination are considered alone multiple equilibria may arise ($\mu\beta^D + \nu\beta^C > 0$). Inner equilibria with relative inequality-averse and selfish players in coexistence may only arise if the reverse is true. In such inner equilibria only in one of the subpopulation that correspond to the role assignment in the problem of redistribution inequality-averse and selfish players may coexist.

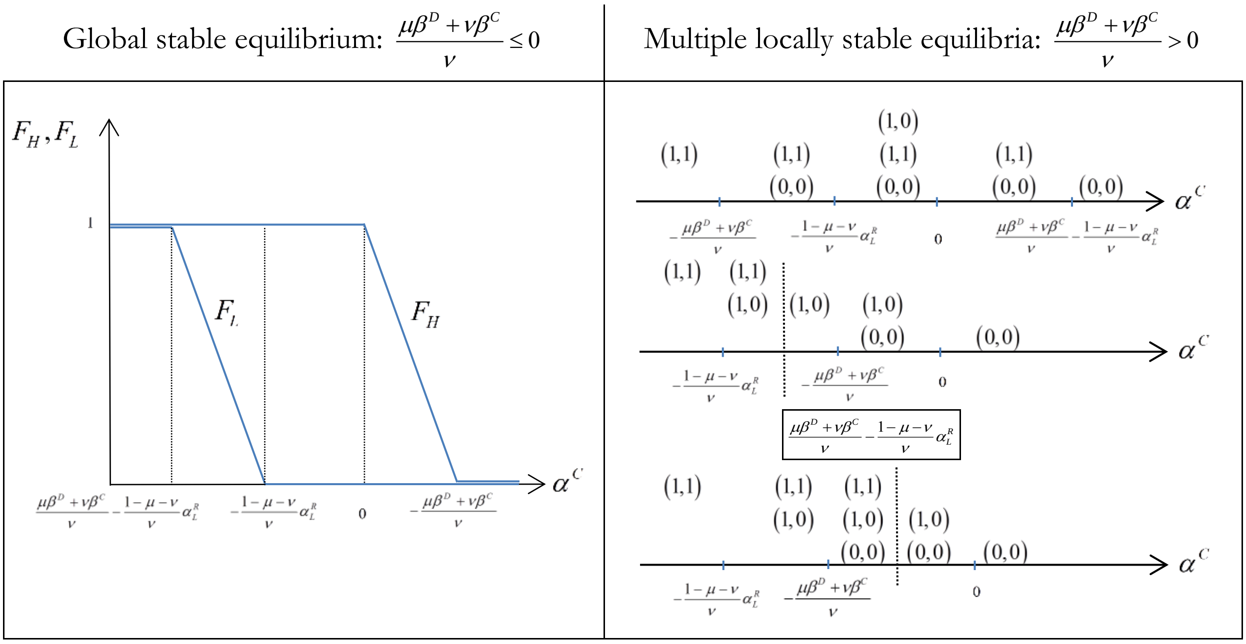


Figure 4-2: set of equilibria, (F_H, F_L) for the right column.

As a consequence of the analysis in the single environments one should expect a tradeoff between the advantageous dilemma environment and the disadvantageous distributional problem with respect for the evolution of the preference for equality. We saw in Proposition 4-3 that in almost all cases a preference for equality above θ^{crit} cannot be sustained in equilibrium among high types in the redistribution problem. A particularly interesting question is therefore whether there is a stable equilibrium with a positive share of inequality-averse players among high types if the simplified game of life is considered, i.e. $F_H\left(\theta^{crit}\right)\in(0,1)$. The theorem reveals that this is indeed the case. However an inner equilibrium with relatively inequality-averse and relatively selfish players can only emerge in the case where the problem of coordination on its own would stabilized such a distribution of preferences ($0\leq\alpha^C$, see Figure 4-2). If in a population of only inequality-averse players selfish individuals would on average face an evolutionary disadvantage when only the dilemma and the problem of coordination are considered ($\mu\beta^D+\nu\left(\alpha^C+\beta^C\right)>0$) then inequality aversion will be advantageous for high and low types and a stable equilibrium with $F_L=0, F_H=0$ exists. In all other cases the inequality-averse high types are deemed to extinction also in the simplified game of life. If the problem of coordination is not too disadvantageous for inequality-averse individuals then the advantageousness for the dilemma and the problem of distribution carries over to the simplified game of life and a stable equilibrium with only inequality-averse players exists. At an intermediate level of disadvantageousness both inequality-averse and selfish players will coexist in the subpopulation of low types. Finally, if the disadvantage in the problem of coordination dominates then in both sub-populations only selfishness may be part of a stable distribution of preferences.

In summary, on the one hand the simplified game of life as expected gives rise to a greater variety in potential equilibrium distributions of preferences. In particular the surprisingly strong predictions for the single environments are put into perspective. The global advantage of inequality-averse players in the dilemma and the global disadvantage for inequality-averse high types in almost all cases become subject to some qualification. On the other hand the expected stabilization of inner equilibria in which relatively inequality-averse individuals and relatively selfish individuals coexist occurs if and only if the single environments show the same feature. The reason for this is that advantageousness in the dilemma is increasing and the disadvantageousness in the problem of distribution is decreasing in the share of inequality-averse players. For an inner equilibrium a decreasing advantageousness that eventually turns into a disadvantage is required though.

4.6. Discussion

In this section I want to discuss the robustness of the results with regard to several issues. These issues consider the core assumptions of the paper: the equilibrium-selection criteria, the equilibrium concept, the strictness property, and the model of inequality aversion.

4.6.1. Equilibrium selection

I now turn to the assumption concerning equilibrium selection that agents jointly randomize over the set of pure Nash equilibria with equal weight. I claimed in section 4.3.5 that, lacking a general theory of equilibrium selection, the requirement on the selection criteria to *a priori* be neutral with respect to the evolutionary success of inequality aversion amounts to a symmetric probability distribution over the set of equilibria. This requirement stems from the fact that I am solely interested in the evolutionary forces that follow from the impact of a particular preference on the set of Nash equilibria and not in forces that are based on some selection bias. A symmetric probability distribution implies neutrality, because in that case any two matches of pairs of individuals with potentially different degrees of inequality aversion will earn the same expected material payoff as long as the set of pure Nash equilibria coincide. Symmetry is thus sufficient for neutrality. To see necessity, consider the following numerical example of a problem of coordination. Table 4-1 below presents the material payoffs of $\gamma\left(A^1, A^2\right)$ and their evaluation.

	0	1
0	$3, 3$	$2, 4$
1	$4, 2$	$0, 0$

Table 4-1:Payoffs in $\gamma\left(A^1, A^2\right)$

	0	1
0	$3, 3$	$2-2\theta_1, 4-2\theta_2$
1	$4-2\theta_1, 2-2\theta_2$	$0, 0$

Payoffs in $\gamma\left(U^1, U^2\right)$.

In a match of two individuals with inequality aversion $0<\theta_1<\theta_2<\frac{1}{2}$, i.e. preferences of player two shows a higher degree of inequality aversion, the set of pure Nash equilibria of $\gamma\left(A^1, A^2\right)$ and $\gamma\left(U^1, U^2\right)$ coincide. Any asymmetric probability distribution over the set $\{(0,1),(1,0)\}$ will (dis)favor the relatively inequality-averse player if a (smaller) larger weight is put on $(0,1)$. Thus an asymmetric distribution gives an evolutionary advantage or disadvantage to the relatively inequality-averse player, but is not neutral.

Next to concerns of some economists about the play of mixed strategies, the assumption that the randomization is over the set of pure Nash equilibria and thus excludes the equilibrium in mixed strategies from the support of the probability distribution was made for simplicity. Interestingly this parallels the application of a coarser equilibrium concept than the notion of Nash equilibrium. In that case, average payoffs may only change if this subset of Nash equilibria changes. The role of the play of mixed equilibria will be analyzed when the assumption for the problems of coordination and distribution to be strict is discussed in section 4.6.3. The results of section 4.6.3 and of the discussion of correlated equilibria in 4.6.2 indicate that the finer the equilibrium concept, the more sensitive the equilibrium set to changes in preference parameter and thus the higher the precision of the prediction characterizing the equilibrium distribution of inequality aversion.

As mentioned before, the assumption of a uniform randomization over the set of pure Nash equilibria is equivalent to a play of the correlated equilibrium that assigns equal weights to each of the pure Nash equilibria. In other words, if multiple pure Nash equilibria exist individuals play a particular correlated equilibrium. The implications of considering not one but the whole set of correlated equilibria is discussed in the next section.

4.6.2. Equilibrium concept

As pointed out before, not only is the notion of the equilibrium decisive with respect to the classification of a game into a dilemma, a problem of coordination or a problem of distribution, but it also plays a role in the evolutionary analysis. The reason why a preference for equality may be advantageous or disadvantageous from an evolutionary perspective lies in its leverage on the equilibrium set. Changes in the set of equilibria may change average payoffs and thereby generate evolutionary pressure. How the set of equilibria is altered by transforming the underlying game in material payoffs by preferences evaluating these payoffs may depend on the applied notion of equilibrium. Most of applied game theory takes the Nash equilibrium as its reference point and deals with finer or coarser equilibrium concepts relative to the Nash-concept. Let me illustrate the effects for a concrete alternative, that of correlated equilibria[30]. This concept not only enlarges the set of equilibria and thereby enlarges the class of problems of coordination but it also increases the set of achievable payoffs generated by correlated strategies. To make the argument precise I will restate some definitions and results of Calvó-Armengol (2006) who studies the set of correlated equilibria for 2x2 games. For $\gamma\left(A^1,A^2\right)\in\Gamma^\circ$ define $\alpha_1=\left|a_{00}^1-a_{10}^1\right|/\left|a_{11}^1-a_{01}^1\right|$ and $\alpha_2=\left|a_{00}^2-a_{01}^2\right|/\left|a_{11}^2-a_{10}^2\right|$. In the absence of neither weakly nor strictly dominated strategies α_1 and α_2 are well defined and strictly positive. The defined values give rise to three different types of games:

	0	1
0	α_1,α_2	$0,0$
1	$0,0$	$1,1$

$\gamma_I\left(\alpha_1,\alpha_2\right)$: coordination

	0	1
0	$-\alpha_1,-\alpha_2$	$0,0$
1	$0,0$	$-1,-1$

$\gamma_{II}\left(\alpha_1,\alpha_2\right)$: anti-coordination

	0	1
0	$-\alpha_1,\alpha_2$	$0,0$
1	$0,0$	$-1,1$

$\gamma_{III}\left(\alpha_1,\alpha_2\right)$: competitive

Table 4-2: Classification of 2x2 games by Calvó-Armengol (2006)

Lemma 4-5 (Calvó-Armengol 2006, Lemma 1) Let $\gamma\left(A^1,A^2\right)\in\Gamma^\circ$. Then, for the set of correlated equilibria (*CE*) of $\gamma\left(A^1,A^2\right)\in\Gamma^\circ$ holds:

$$CE\left(\gamma\left(A^1,A^2\right)\right)=CE\left(\gamma_l\left(\alpha,\beta\right)\right),\text{ for some } l\in\{I,II,III\}.$$

The restated result of Calvó-Armengol (2006) proves that the set Γ° of 2x2 games can be partitioned into three equivalence classes for the set of correlated equilibrium strategies. It is easily verified that $CE\left(\gamma_{III}\left(\alpha_1,\alpha_2\right)\right)=NE\left(\gamma_{III}\left(\alpha_1,\alpha_2\right)\right)$, i.e. the sets of correlated equilibria and Nash equilibria coincide and the set consist of a single point in Δ_3, the 3-dimensional simplex of $\mathbb{R}^4$. Let

	0	1
0	μ_{00}	μ_{01}
1	μ_{10}	μ_{11}

Table 4-3: The canonical representation of a correlated strategy.

[30] There is plenty of theoretical and empirical support for the relevance of the concept of correlated equilibrium. Aumann (1974) shows that a particular definition of Bayesian rationality generates outcomes identical to the set of correlated equilibria. This result was extended by Brandenburger and Dekel (1987) and Tan and da Costa Werlang, Sérgio Ribeiro (1988). Nyarko (1994) showed that Bayesian learning leads to correlated equilibria in normal form games. In an evolutionary context Lenzo and Sarver (2006) establish the correlated equilibrium as a natural solution concept. In particular they show that every interior stationary state, Lyapunov stable state, or limit of an interior solution is equivalent to a correlated equilibrium. This result is generalized by Koch (2008). They show for boundedly rational agents that a set of signal contingent strategies is asymptotically stable only if it represents a strict correlated equilibrium.

be the representation of a correlated strategy $\mu = (\mu_{00}, \mu_{11}, \mu_{10}, \mu_{01}) \in \Delta_3$.

Lemma 4-6 (Calvó-Armengol 2006, Lemma 2) $\mu \in CE(\gamma_I(\alpha_1, \alpha_2))$ if and only if $\tau(\mu) \in CE\left(\gamma_{II}\left(\alpha_1, \frac{1}{\alpha_2}\right)\right)$, where $\tau(x) = (x_3, x_4, x_1, x_2)$ for $(x_1, x_2, x_3, x_4) \in \mathbb{R}^4$.

Lemma 4-6 reveals that the class of coordination games and the class of anti-coordination games are isomorphic to one another. It thus suffices to characterize the set of correlated equilibria for one class. I will restate the result for the class of coordination games. A game $\gamma_I(\alpha_1, \alpha_2)$ of that class has three Nash equilibria and two correlated equilibria, the probability measures of which are given in Table 4-4.

μ	μ_{00}	μ_{11}	μ_{10}	μ_{01}	
$\mu_C^*(\alpha,\beta)$	1	0	0	0	
$\mu_D^*(\alpha,\beta)$	0	1	0	0	
$\mu_E^*(\alpha,\beta)$	$\frac{1}{(1+\alpha_1)(1+\alpha_2)}$	$\frac{\alpha_1\alpha_2}{(1+\alpha_1)(1+\alpha_2)}$	$\frac{\alpha_2}{(1+\alpha_1)(1+\alpha_2)}$	$\frac{\alpha_1}{(1+\alpha_1)(1+\alpha_2)}$	(4.5)
$\mu_F^*(\alpha,\beta)$	$\frac{1}{1+\alpha_2+\alpha_1\alpha_2}$	$\frac{\alpha_1\alpha_2}{1+\alpha_2+\alpha_1\alpha_2}$	$\frac{\alpha_2}{1+\alpha_2+\alpha_1\alpha_2}$	0	
$\mu_G^*(\alpha,\beta)$	$\frac{1}{1+\alpha_1+\alpha_1\alpha_2}$	$\frac{\alpha_1\alpha_2}{1+\alpha_1+\alpha_1\alpha_2}$	0	$\frac{\alpha_1}{1+\alpha_1+\alpha_1\alpha_2}$	

Table 4-4: Probability measures for correlated equilibria and Nash equilibria for a game $\gamma_I(\alpha_1, \alpha_2)$.

Proposition 4-4 relates the 5 vertices given in Table 4-4 and the set of correlated equilibria.

Proposition 4-4 (Calvó-Armengol 2006, Proposition 1) $CE(\gamma_I(\alpha_1, \alpha_2))$ is a polytope of Δ_3 with five vertices given in Table 4-4.

Note that for symmetric games the class of competitive games is empty. By a similar argument as given in 4.4 symmetric coordination and anti-coordination games are free of the dilemma property. Hence a symmetric social dilemma must be in the set $\Gamma_{sym} \setminus \Gamma_{sym}^{\circ}$, the set of games with weakly or strictly dominated strategies. If a player has a strictly dominant strategy then by symmetry his opponent has the same strictly dominant strategy. Two cases can be distinguished. The first one corresponds to the classical Prisoners' Dilemma, i.e. one of the diagonal outcomes is the unique correlated equilibrium payoff being Pareto-inferior to the other diagonal outcome. The second is given by payoffs where the equilibrium payoff is equal or even Pareto-superior to the non-equilibrium diagonal outcome, but where a correlation of out-of-diagonal outcomes yields higher payoffs for both players. In other words with respect to symmetric games both classes, that of dilemma and that of coordination grow whereas the class of what I referred to as unproblematic situations shrinks. Note that any strictly dominated strategy cannot be played with strictly positive probability in any correlated equilibrium of a finite game. Hence Lemma 4-2 also holds if the concept of correlated equilibria is applied. In particular the definition of the critical threshold for the required inequality aversion carries over. If two sufficiently inequality-averse players are matched they play a game γ_I (see Table 4-2). Given the assumption that player randomize over the set of all equilibria, the derivation of expected material payoffs for the set of correlated equilibria is more involved. Due to the linearity of the inner product, the calculation of the expected payoff amounts to the determination of the center of mass of P, the polytope of Proposition 4-4. Equation (4.6) states this property formally, where $\pi = \left(a_{00}, a_{11}, a_{10}, a_{01}\right)$ denotes the payoffs associated with the payoff matrix A of the game.

$$E\pi = \int_{\mu \in P} \frac{1}{Vol(P)} \langle \pi, \mu \rangle \, d\mu = \left\langle \pi, \frac{1}{Vol(P)} \int_{\mu \in P} \mu \, d\mu \right\rangle = \left\langle \pi, \mu^{CM} \right\rangle \tag{4.6}$$

The following Lemma presents the center of mass for the polytope P.

Lemma 4-7 Let $\gamma_I\left(\alpha_1, \alpha_2\right) \in \Gamma^\circ$. Then the center of mass μ^{CM} of P is given by:

$$\begin{pmatrix} \mu_{00}^{CM} \\ \mu_{10}^{CM} \\ \mu_{01}^{CM} \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 1 + \frac{1}{1+\alpha_1+\alpha_1\alpha_2} + \frac{1}{1+\alpha_2+\alpha_1\alpha_2} + \frac{1}{(1+\alpha_1)(1+\alpha_2)} - \frac{2}{2+\alpha_1+\alpha_2+2\alpha_1\alpha_2} \\ \alpha_2 \left(\frac{1}{1+\alpha_2+\alpha_1\alpha_2} + \frac{1}{(1+\alpha_1)(1+\alpha_2)} - \frac{1}{2+\alpha_1+\alpha_2+2\alpha_1\alpha_2} \right) \\ \alpha_1 \left(\frac{1}{1+\alpha_1+\alpha_1\alpha_2} + \frac{1}{(1+\alpha_1)(1+\alpha_2)} - \frac{1}{2+\alpha_1+\alpha_2+2\alpha_1\alpha_2} \right) \end{pmatrix} \tag{4.7}$$

Given the center of mass, I can now compare expected profits in matches of individuals with potentially different degrees of inequality aversion.

Proposition 4-5 Let $\gamma\left(A^1, A^2\right) \in \Gamma^\circ$. Then:

$$E\pi_1 - E\pi_2 > 0 \Leftrightarrow \left(\mu_{10}^{CM} - \mu_{01}^{CM}\right)\left(a_{10} - a_{01}\right) > 0 \Leftrightarrow \left(\alpha_2 - \alpha_1\right)\left(a_{10} - a_{01}\right) > 0 \tag{4.8}$$

The first equivalence follows from symmetry on the diagonal of the bimatrix representing $\gamma\left(A^1, A^2\right)$, i.e. only the weights of off-diagonal outcomes may account for a difference in expected material payoffs. Without loss of generality, I will focus on the case where $\left(a_{10} - a_{01}\right) > 0$. In that case player 1 earns more than his opponent if and only if relatively more weight is put on outcome $\left(1, 0\right)$ that favors player 1. The second equivalence may be less obvious, but concerning this matter the vertices of F and G presented in Table 4-4 are already quite suggestive. Equation (4.7) reveals that this property of the weights μ_{10} and μ_{01} for vertices F and G respectively carry over to the center of mass.

As already mentioned, if the concept of correlated equilibrium is applied there is another type of dominance solvable game next to the Prisoner's Dilemma that constitutes a social dilemma. It turns out that the qualitative results for the Prisoner's Dilemma type do not change, but gain in precision in the sense that the equilibrium distribution of the preference-parameter measuring the degree of inequality aversion can be characterized more precisely. This gain in precision stems from the fact that two individuals who are sufficiently inequality-averse to transform the

Prisoner's Dilemma into a coordination-game no longer earn the same expected payoff when the concept of correlated equilibrium is applied (see Equation (4.8)). However, in the other case where the equilibrium payoff is equal or even Pareto-superior to the non-equilibrium diagonal outcome, but where a correlation of out-of-diagonal outcomes yields higher payoffs for both players, results change significantly. In particular it is the case that inequality-averse players face a global evolutionary disadvantage. I will refer to this case as the non-PD-case.

Proposition 4-6 Let $\gamma(A) \in \Gamma_{sym}$ be a social dilemma. If $\left(\neg s_1^*, \neg s_2^*\right)$ is stabilizable, then there exists a $\theta^D \in [0,1]$, such that the globally stable equilibrium in case of the Prisoners' Dilemma is characterized by $\theta = \theta^D$ for all individuals in the population. In the non-PD-case the globally stable equilibrium is characterized by $F\left(\theta^D\right) = 1$.

In other words, in case of the PD a precise value of inequality aversion is selected by evolutionary forces. This value corresponds to the lowest value that suffices to transform the dilemma into a coordination-game. In the non-PD-case stabilization of the material non-equilibrium outcome implies an evolutionary disadvantage of inequality aversion, i.e. the reverse result. The intuition behind this is, that it is relatively advantageous for inequality-averse individuals if a relatively low weight is put on the disadvantageous one of the two off-diagonal outcome which on average earns higher profits then the unique PD-outcome. In other word it pays off to be relatively opportunistic among the inequality-averse players, because than more weight is put on the off-diagonal outcome which is relatively advantageous. Consequently, while more successful players are selected by evolution less weight is put on the off-diagonals, ultimately leading to a randomization among the two diagonals. This randomization is advantageous in the PD and disadvantageous in the non-PD-case.

The analysis of the class of social dilemmas reveals that results may change when a different concept of equilibrium is applied. With respect to generalizability of the results for the Nash equilibrium concept the preliminary results are ambiguous. On the one hand, the results for the Nash equilibrium carry over to the correlated equilibrium in case of the Prisoners' Dilemma. Interestingly, I obtained a huge gain in precision with respect to the prediction of the stable distribution of preferences. Whereas in the Nash case the distribution could be characterized up a threshold, this threshold was picked as the unique equilibrium value in case of correlated equilibria. On the other hand, a new case which constitutes a dilemma in case of correlated equilibria but not under Nash equilibria emerges. In this case the reverse result with respect to the evolutionary advantageousness of inequality aversion was obtained. Thus, the chosen equilibrium concept appears to have some impact on the results. A detailed analysis for all classes of games is left for future research. However, the effect on the precision of prediction regarding equilibrium distribution of preference observed when applying the notion of correlated equilibria will to some extent also be present when the play of mixed strategies is allowed. This role of randomized play points to the assumption for the problems of coordination and distribution respectively to be strict, which is discussed in the next section.

4.6.3. Strictness

For agents to apply mixed strategies, players need to be indifferent between the involved pure strategies. The involved probabilities equalize the expected payoffs of the pure strategies and are sensitive to marginal changes in payoffs whereas best responses in a strict Nash equilibrium change only in a discrete manner. As a consequence the set of equilibria is more elastic with respect to changes in the preference parameter when mixed equilibria are considered. To see this, consider a symmetric problem of coordination. As mentioned in Section 4.3.4 the requirement for such a problem to be strict excludes the case where two players transform the problem of coordination into a game with a unique mixed equilibrium given their degree of inequality aversion. This is the case if and only if an inequality-averse player who destabilizes both material equilibria is matched with a selfish player who destabilizes none of the material equilibria (see proof of Proposition 4-2 in Appendix B). The unique mixed equilibrium is given by $\mu_E^*(\alpha_1,\alpha_2)$ for a game $\gamma_I(\alpha_1,\alpha_2)$ (see Table 4-4 and Lemma 4-6 for a game $\gamma_{II}(\alpha_1,\alpha_2)$). Note that a game of type $\gamma_I(\alpha_1,\alpha_2)$ is always strict since any degree of inequality aversion will increase the strictness of the two symmetric equilibria on the diagonal of $\gamma_I(\alpha_1,\alpha_2)$. Hence I concentrate on non-strict problems of coordination of type $\gamma_{II}(\alpha_1,\alpha_2)$, i.e. I analyze anti-coordination games such that both pure Nash equilibria are contestable by each player individually. If both equilibria are contestable by one player, by symmetry they are also contestable by the opponent.

Proposition 4-7 Let $\gamma(A) \in \Gamma_{sym}$ be a problem of coordination such that both equilibria are contestable by one player.

(1) If the equilibria are less strict for those player who are favored in the equilibria then no additional stable equilibria arise. In particular there is no stable distribution of preferences that assigns a positive share to players by whom both equilibria are contested.

(2) If the equilibria are less strict for those player who are disfavored in the equilibria then additional stable equilibria arise. In particular there may be a stable distribution of preferences only with players by whom both equilibria are contested. Furthermore there may be a stable distribution of preferences where player who contest none of the equilibria and players who contest both equilibria coexist. No stable equilibrium distributions exist with all three types of players, those who contest none equilibrium, those who contest one equilibrium and those who contest both equilibria.

In case (1) of Proposition 4-7 giving up strictness has no consequences with respect to the characterization of the stable distribution of preferences. However, in case (2) the results presented in Proposition 4-2 experience two qualifications. First, there is a minor qualification with respect to the existence of an inner equilibrium where opportunistic and inequality-averse individuals coexist. In a non-strict problem of coordination there may also be a stable equilibrium with highly inequality-averse players who so far were excluded from analysis and opportunistic players. Second, and this is a major qualification, the result implied by Proposition 4-2 that inequality-averse individuals may at most partially be present in equilibrium is put into perspective. In case (2) of Proposition 4-7 there may be a stable equilibrium with only (highly) inequality-averse individuals. However, it still holds for medium inequality-averse individuals, i.e. player who contest one equilibrium, that they may at most partially be present in equilibrium. Thus, the assumption for problems of coordination to be strict implies that the evolutionary success of inequality aversion is underestimated. This transfers to the simplified game of life and introduces another case how inequality aversion could be stabilized among high types in the problem of distribution.

I will now discuss the consequences of relaxing the restriction for problems of distribution to be strict. In particular I am interested in whether the strong prediction of an evolutionary disadvantage for inequality-averse high types carries over to non-strict problems of distribution. Proposition 4-3 revealed that with one exception the distribution of inequality aversion among high types is characterized by $F_H\left(\theta_H^R\right)=1$, i.e. only relatively opportunistic players are present in equilibrium. This exception occurs if the two pure Nash equilibria are Pareto-ranked. If equilibria are not ranked then the distribution always shows the property of an evolutionary disadvantage of inequality aversion among high types. In any case, the prerequisite was that one of the material equilibria is contestable by high types. Given up strictness now allows both equilibria to be contestable by the same player. However, only low types may contest both equilibria since for high types in any case at most one equilibrium is contestable. If equilibria are not Pareto-ranked it is the more equally distributed equilibrium that is not contestable, if equilibria are ranked it is the Pareto-superior equilibrium.

Proposition 4-8 Let $\gamma\left(A^1,A^2\right)$ constitute a non-strict problem of distribution, such that the pure Nash equilibria are not Pareto-ranked. Then the globally stable equilibrium distribution is characterized by $F_H\left(\theta_H^R\right)=1$.

Proposition 4-8 shows that the disadvantage of inequality-averse high types transfers to non-strict problems of redistribution if equilibria are not Pareto-ranked. However, next to the two cases distinguished in Lemma 4-4 there is a third class of games that may constitute a problem of distribution if strictness is given up, namely that of a competitive game (see Table 4-2) with the unique Nash equilibrium being in mixed strategies. This case and the one with Pareto-ranked equilibria are left for future research.

4.6.4. Modelling inequality aversion

Finally, dropping the assumption that individuals care about favorable and unfavorable inequality in the same way has interesting consequences. In what follows I will elaborate on the consequences of a more complex model of inequality aversion proposed by Fehr and Schmidt (1999)[31], i.e. $u_{(i,j)}^n=a_{(i,j)}^n-\sigma^n\max\left\{a_{(i,j)}^{-n}-a_{(i,j)}^n,0\right\}-\omega^n\max\left\{a_{(i,j)}^n-a_{(i,j)}^{-n},0\right\},\ \sigma^n,\omega^n\in[0,1]$. Thus, individuals preference for equality is no longer characterized by the single parameter θ, but by a pair (σ,ω). Consider again the symmetric prisoners' dilemma. In this game the Pareto-superior outcome can be stabilized by sufficiently inequality-averse players as they devaluate the material gain from defecting on a cooperative opponent due to the induced inequality generated by such a defection. Hence, in case of a symmetric dilemma not inequality aversion *per se* but aversion against favorable outcomes is required to support cooperation. With respect to problems of coordination two cases were distinguished in Proposition 4-2. In the first case, the destabilized

[31] Note that the concept of inequality aversion of Bolton and Ockenfels (2000) implies symmetry, but it is left for further research whether this notion will change qualitative results of the evolutionary analysis.

equilibrium is materially favorable for inequality-averse players. Hence, in that case an aversion against favorable inequality is decisive. In the second case the reverse holds, i.e. the destabilized equilibrium is materially favorable for selfish individuals. Hence in that case an aversion against unfavorable inequality becomes relevant. For problems of distribution there is no such clear assignment for the thresholds of Proposition 4-3. To see this, consider the example given in Table 4-5 which belongs to the first case in Proposition 4-3. In the example the column player is the high type and the row player is the low type. Furthermore the game presented in Table 4-5 has two pure non-pareto-ranked Nash equilibria on the diagonal. I consider the case where none of the equilibria is contestable by high types and the $(0,0)$ is contestable by low types.

	0	1
0	$A-\theta^2\lvert A-a\rvert$ $a-\theta^1\lvert A-a\rvert$	$B-\theta^2\lvert B-b\rvert$ $b-\theta^1\lvert B-b\rvert$
1	$C-\theta^2\lvert C-c\rvert$ $c-\theta^1\lvert C-c\rvert$	$D-\theta^2\lvert D-d\rvert$ $d-\theta^1\lvert D-d\rvert$

Table 4-5: $A>B,\ D>C,\ a>c,\ d>b,\ a<d<D<A$

The $(0,0)$-equilibrium is contested by a low type if and only if:

$$a-\theta^1\left|A-a\right|<c-\theta^1\left|C-c\right| \tag{4.9}$$

The example implies that $A-a>0$, but no relation for $C-c$. If the outcome of playing $(1,0)$ also favors high types, i.e. $C-c>0$ then (4.9) becomes

$$a-\sigma^1\left(A-a\right)<c-\sigma^1\left(C-c\right) \tag{4.10}$$

This suggests that if high types are favored no matter which strategies are played, then the threshold θ_L^R in Proposition 4-3 refers to inequality aversion concerning favorable outcomes.

If however the reverse is true, i.e. $C-c<0$ then (4.9) becomes

$$a-\sigma^1\left(A-a\right)<c-\omega^1\left(c-C\right) \tag{4.11}$$

In this case both parameters become relevant and no clear assignment to the thresholds in Proposition 4-3 is possible. Rewriting (4.11) as

$$\sigma^1>\frac{a-c}{A-a}+\frac{c-C}{A-a}\omega^1 \tag{4.12}$$

reveals that the threshold θ_L^R needs to be substituted by a linear condition, described by (4.12) when considered as equality, which separates the two dimensional parameter space characterizing the preference for equality by (σ,ω)-pairs. Thus, individuals with (σ,ω) located above (below) that line can(not) contest the equilibrium.

A similar argument applies to high types. If the example above is changed in a way such that the $(0,0)$-equilibrium becomes contestable for the high type it again depends on the sign of the difference in payoffs of the outcome which is realized if a high type is sufficiently inequality-averse such that $(0,0)$ is indeed destabilized, i.e. on the sign of $B-b$. If $B-b>0$ then $(0,0)$ is destabilized by a high type if and only if $A-\omega^2(A-a)<B-\omega^2(B-b)$. Thus, for the example θ_H^R in Proposition 4-3 refers to inequality aversion concerning favorable outcomes. As for low types, if the reverse holds, i.e. $B-b<0$, then both parameters become relevant and θ_H^R needs to be substituted by a linear condition in the fashion of (4.13).

In summary the thresholds of inequality aversion I derived for a dilemma (θ^D, see Proposition 4-1) and a problem of coordination (θ^C, see Proposition 4-2) still remain valid but will under the more complex model of inequality aversion refer to the parameter measuring aversion regarding favorable or unfavorable inequality. Thus the precision of the prediction increases as in the more complex model statements will refer not only to the level but also to the type of inequality aversion. The example for a problem of redistribution indicates that in the (σ,ω)-model of inequality aversion the thresholds θ_L^R and θ_H^R may transfer to thresholds regarding σ for low types and ω for high types or need to be replaced by a linear condition on the relation of σ and ω. For problems of redistribution a detailed analysis is left for future research. Proposition 4-9 below summarizes these insights formally.

Proposition 4-9 Let $\gamma(A)\in\Gamma_{sym}$ be a social dilemma, then $\theta^D=\omega^D$. Let $\gamma(A)\in\Gamma_{sym}$ be a strict problem of coordination. If equilibria are contestable then:

1. if the destabilized equilibrium is materially favorable for inequality-averse individuals then $\theta^D=\omega^D$.
2. if the destabilized equilibrium is materially favorable for selfish individuals then $\theta^C=\sigma^C$.

In summary, with respect to the assumption of an uniform distribution over the set of all pure Nash equilibria, it turned out that neutrality of the distribution concerning the evolutionary success of inequality aversion implies symmetry and symmetry implies uniformity if 2x2 games are considered. With respect to generalizability of the results for the Nash equilibrium concept (Proposition 4-1-Proposition 4-3) the preliminary results (Proposition 4-6) are ambiguous and further research is needed to fully understand the sensitivity of the results regarding the coarseness of the applied equilibrium concept relative to the Nash equilibrium. Concerning the assumption on the problem of coordination to be strict, the degree of disadvantageous of inequality aversion (Proposition 4-2) is put into perspective as in a non-strict problem of coordination there may exist a stable equilibrium with only inequality-avers players. However, this requires that the equilibria are less strict for those players who are disfavored in the equilibria (see case (2) in Proposition 4-7). If the reverse is true though, no additional equilibria arise if the assumption of strictness is relaxed. Proposition 4-8 proofs that the strong prediction of an evolutionary disadvantage for inequality-averse high types also holds for non-strict problems of distribution if equilibria of $\gamma(A^1,A^2)$ are not Pareto-ranked (Proposition 4-8). Finally, if a model of preferences is applied that distinguishes between aversion against favorable and unfavorable inequality, then the results of Proposition 4-1 (dilemma) and Proposition 4-2 (problem of coordination) carry over. However, the parameter measuring inequality aversion in the simplified model is replaced by either the parameter for aversion against favorable or by the one for unfavorable inequality. For problems of distribution the discussion in 4.6.4 suggests that the thresholds of Proposition 4-3 are either replaced by a threshold referring to aversion against favorable inequality or by a linear constraint relating the two parameters of the alternative model of inequality aversion.

4.7. Conclusion

The purpose of the paper was twofold. Following the argument for a requirement to analyze the evolution of preference in an environment that comprises at best all relevant classes of games individuals engage in, I have suggested a particular notion of a simplified game of life. The simplified game of life as I have defined it comprises three classes of games: a symmetric dilemma, a symmetric and strict problem of coordination and a strict problem of distribution. Second I have analyzed the evolution of a particular type of other-regarding preference namely that of inequality aversion in the 2x2 simplified game of life.

The analysis in the separate environments revealed that if inequality aversion has a leverage on the set of equilibria played, then inequality aversion enjoys a global evolutionary advantage over more selfish preferences in a dilemma. In the class of problems of coordination inequality aversion surprisingly faces a weak evolutionary disadvantage in the sense that at most a stable inner equilibrium exists where relative inequality-averse and relative selfish players coexist, in all other cases relatively inequality-averse players will eventually disappear. In the problem of distribution a preference for equality will always be favored by evolutionary selection dynamics among those individuals disfavored by the problem. For those individuals favored in the problem of distribution in all cases up to one inequality aversion will eventually disappear. I consider these predictions in the light of the considered generality as rather strong. Furthermore, due to the exemplary variations of assumptions discussed in Section 4.6 these predictions appear quite robust. Note that among the eight loci a linear function can take in the unit interval one is selected for a dilemma, three for problem of coordination, two for high types in the problem of distribution and three for low types.

The simplified game of life that comprises all three types of interaction, on the one hand as expected gives rise to a greater variety in potential equilibrium distributions of preferences. In particular the surprisingly strong predictions for the single environments are put into perspective. The global advantage of inequality-averse players in the dilemma and the global disadvantage for inequality-averse high types in almost all cases experiences significant qualification. In particular whenever the interplay of the dilemma and the problem of distribution allows for a locally stable equilibrium with only inequality-averse players then this transfers to the simplified game of life, i.e. inequality aversion may also be present among high types. On the other hand the expected stabilization of inner equilibria in which relatively inequality-averse individuals and relatively selfish individuals coexist occurs if and only if the problem of coordination shows the same feature, i.e. the coexistence of both types. The reason for this is that advantageousness in the dilemma is increasing and the disadvantageousness in the problem of distribution is decreasing in the share of inequality-averse players. For an inner equilibrium a decreasing advantageousness that eventually turns into a disadvantage is required though.

The contribution of the paper is threefold. First, the different results in the single-game environments and in the simplified game of life again underpin the necessity to carefully select the relevant game environment in any study of the evolution of preferences. Otherwise any negative or positive results with respect to the rationalization of a particular preference may only point to a potential evolutionary force, which however may not be decisive if all relevant environments are considered. Second, the paper contributes methodologically to the field of evolutionary economics by making a precise suggestion of an evolutionary framework for the study of the evolution of preferences. Third, the paper gives an evolutionary rationale for the presence of inequality aversion within the compound environment of the simplified game of life.

References

Angeletos, G.-M., D. Laibson, A. Repetto, J. Tobacman, S. Weinberg. 2001. The hyperbolic consumption model: Calibration, simulation, and empirical evaluation. *Journal of Economic Perspectives* **15**(3) 47–68.

Aoki, M. 2000. *Information, Corporate Governance and Institutional Diversity: Competitiveness in Japan, the USA, and the Transitional Economies*. Oxford University Press.

Arrow, K. J. 1951. *Social choice and individual values*. Yale University Press.

Aumann, R. J. 1974. Subjectivity and correlation in randomized strategies. *Journal of mathematical Economics* **1**(1) 67–96.

Ashraf, N., I. Bohnet, N. Piankov. 2006. Decomposing trust and trustworthiness. *Experimental Economics* **9**(3) 193–208.

Axelrod, R. 1984. *The Evolution of Cooperation*. New York: Basic Book.

Axelrod, R., W. D. Hamilton. 1981. The evolution of cooperation. *Science* **211**(4489) 1390–1396.

Banks, J., R. Calvert. 1989. Equilibria in Coordination Games. MS. University of Rochester.

Benabou, R., J. Tirole. 2006. Belief in a just world and redistributive politics. *The Quarterly Journal of Economics* **121**(2) 699–746.

Bendor, J., P. Swistak. 1998. Evolutionary equilibria: Characterization theorems and their implications. *Theory and decision* **45**(2) 99–159.

Bergsten, G. S. 1985. On the role of social norms in a market economy. *Public Choice* **45**(2) 113–137.

Bester, H., W. Güth. 1998. Is altruism evolutionarily stable? *Journal of Economic Behavior & Organization* **34**(2) 193–209.

Binmore, K. 1998. *Game Theory and the Social Contract - Volume II: Just Playing*. MIT Press.

Bisin, A., T. Verdier. 2001. The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory* **97**(2) 298-319.

Blanco, M., D. Engelmann, H. T. Normann. 2011. A within-subject analysis of other-regarding preferences. *Games and Economic Behavior* **72**(2) 321–338.

Bohnet, I., B.S. Frey, S. Huck. 2001. More Order with Less Law: On Contract Enforcement, Trust, and Crowding. *The American Political Science Review* **95**(1) 131–144.

Bolton, G. E., A. Ockenfels. 2000. ERC: A theory of equity, reciprocity, and competition. *American Economic Review* **90**(1) 166–193.

Bowles, S. 2000. Economic institutions as ecological niches. *Behavioral and Brain Sciences* **23**(1) 148–149.

Bowles, S., H. Gintis. 2000. Walrasian economics in retrospect. *Quarterly Journal of Economics* **115**(4) 1411–1439.

Boyd, R., P.J. Richerson. 1985. *Culture and the evolutionary process*. Chicago: University of Chicago Press.

Brandenburger, A., E. Dekel. 1987. Rationalizability and correlated equilibria. *Econometrica: Journal of the Econometric Society* **55** 1391–1402.

Brekke, K.A., S. Kverndokk, K. Nyborg. 2003. An economic model of moral motivation. *Journal of Public Economics* **87**(9) 1967–1983.

Brenner, T. 1999. *Modelling learning in economics*. Edward Elgar Publishing.

Calvert, R. L. 1995. Rational actors, equilibrium, and social institutions. In *Explaining social institutions*, edited by Jack Knight and Itai Sened. University of Michigan Press. pp. 57–95.

Calvo-Armengol, A. 2006. The Set of Correlated Equilibria of 2x2 Games. Mimeo.

Camerer, C., G. Loewenstein. 2004. *Behavioral economics: Past, present, future.* Princeton University Press.

Chaudhuri, A., L. Gangadharan. 2007. An experimental analysis of trust and trustworthiness. *Southern Economic Journal* **73**(4) 959–985.

Cialdini, R.B., N.J. Goldstein. 2004. Social influence: compliance and conformity. *Annual Review of Psychology* **55** 591–621.

Cooper, R., D. V. DeJong, R. Forsythe, T. W. Ross. 1996. Cooperation without reputation: experimental evidence from prisoner's dilemma games. *Games and Economic Behavior* **12**(2) 187–218.

Cunha, J. M., N. Augenblick. 2014. Competition and Cooperation in a Public Goods Game: A Field Experiment.

Ek, K., P. Söderholm. 2008. Norms and economic motivation in the Swedish green electricity market. *Ecological Economics* **68**(1) 169–182.

Ellison, G. 1994. Cooperation in the prisoner's dilemma with anonymous random matching. *The Review of Economic Studies* **61**(3) 567–588.

Fehr, E., E. Kirchler, A. Weichbold, S. Gächter. 1998. When social norms overpower competition: Gift exchange in experimental labor markets. *Journal of Labor Economics* **16**(2) 324–351.

Fehr, E., K. M. Schmidt. 1999. A theory of fairness, competition, and cooperation. *The quarterly journal of economics* **114**(3) 817–868.

Fehr, E., S. Gächter. 2000. Do Incentive Contracts Crowd Out Voluntary Cooperation?. Institute for Empirical Research in Economics, University of Zürich, Working Paper No. 34.

Frey, B.S., R. Jegen. 2001. Motivation Crowding Theory. *Journal of Economic Surveys* **15**(5) 589–611.

Frey, B., S. Meier. 2004. Social comparisons and pro-social behavior: testing 'conditional cooperation' in a field experiment. *The American Economic Review* **94**(5) 1717-1722.

Frisch, R. 1926. Sur un probleme d'economie pure. *Norsk Matematisk Forenings Skrifter* **I**(16) 1–40.

Fudenberg, D., D. G. Rand, A. Dreber. 2012. Slow to anger and fast to forgive: cooperation in an uncertain world. *The American Economic Review* **102**(2) 720–749.

Fudenberg, D., E. Maskin. 1986. The folk theorem in repeated games with discounting or with incomplete information. *Econometrica* **54** 533–554.

Gintis, H. 2007. A framework for the unification of the behavioral sciences. *Behaviorial and Brain Sciences* **30**(1) 1–16.

Gintis, H., E. A. Smith, S. Bowles. 2001. Costly signaling and cooperation. *Journal of Theoretical Biology* **213**(1) 103–119.

Gneezy, U. A. Rustichini. 2000. "A Fine is a Price". *Journal of Legal Studies* **29**(1) 1–18.

Gneezy, U., A. Rustichini. 2004. Incentives, punishment and behavior. In *Advances in behavioral economics*, edited by C. F. Camerer, G. Lowenstein and M. Rabin. Princeton University Press. pp. 572–589.

Greif, A. 1989. Reputation and coalitions in medieval trade: evidence on the Maghribi traders. *The journal of economic history* **49**(4) 857–882.

Greif, A. 1997. Self-enforcing political system and economic growth: Late medieval Genoa. *Mimeo. Stanford University.*

Greif, A. 1998. Genoa and the Maghribi traders: Historical and comparative institutional analysis. *Book manuscript.*

Güth, W. 1995. An evolutionary approach to explaining cooperative behavior by reciprocal incentives. *International Journal of Game Theory* **24**(4) 323–344.

Güth, W., A. Ockenfels. 2005. The coevolution of morality and legal institutions: an indirect evolutionary approach. *Journal of Institutional Economics* **1**(2) 155–174.

Güth, W., C. Schmidt, M. Sutter. 2003. Fairness in the mail and opportunism in the internet: A newspaper experiment on ultimatum bargaining. *German Economic Review* **4**(2) 243–265.

Güth, W., H. Kliemt, B. Peleg. 2000. Co-evolution of Preferences and Information in Simple Games of Trust. *German Economic Review* **1**(1) 83–110.

Güth, W., H. Kliemt. 1994. Competition or Co-operation: on the evolutionary economics of trust, exploration and moral attitutes. *Metroeconomics* **45**(2) 155–187.

Güth, W., M. Yaari. 1992. An evolutionary approach to explain reciprocal behavior in a simple strategic game. *U. Witt. Explaining Process and Change–Approaches to Evolutionary Economics. Ann Arbor* 23–34.

Güth, W., S. Napel. 2006. Inequality Aversion in a Variety of Games-An Indirect Evolutionary Analysis*. *The Economic Journal* **116**(514) 1037–1056.

Haken, H. 1977. Synergetics. An Introduction. Nonequilibrium Phase Trasitions and Self-organization in Physics, Chemistry, and Biology. Berlin: Springer.

Hamilton, W. D. 1964a. The genetical evolution of social behavior. I. *Journal of Theoretical Biology* **7**(1) 1–16.

Hamilton, W. D. 1964b. The genetical evolution of social behaviour. II. *Journal of Theoretical Biology* **7**(1) 17–52.

Harrington, J. E. 1987. Finite rationalizability and cooperation in the finitely repeated Prisoners' Dilemma. *Economics Letters* **23**(3) 233–237.

Harsanyi, J. C. 1967. Games with Incomplete Information Played by" Bayesian" Players, I-III. Part I. The Basic Model. *Management Science* **14**(3) 159–182.

Harsanyi, J. C. 1968a. Games with incomplete information played by 'Bayesian'players, part III. The basic probability distribution of the game. *Management Science* **14**(7) 486–502.

Harsanyi, J. C. 1968b. Games with Incomplete Information Played by "Bayesian" Players Part II. Bayesian Equilibrium Points. *Management Science* **14**(5) 320–334.

Hart, O. D., J. Tirole. 1988. Contract renegotiation and Coasian dynamics. *The Review of Economic Studies* **55**(4) 509–540.

Heap, B., J. Kent. 2000. Towards sustainable consumption: a European perspective. Royal Society.

Hirshleifer, D., E. Rasmusen. 1989. Cooperation in a repeated prisoners' dilemma with ostracism. *Journal of Economic Behavior & Organization* **12**(1) 87–106.

Hoffman, E., K. McCabe, K. Shachat, V. Smith. 1994. Preferences, property rights, and anonymity in bargaining games. *Games and Economic Behavior* **7**(3) 346–380.

Hogarth, R. M., H. J. Einhorn. 1992. Order effects in belief updating: The belief-adjustment model. *Cognitive psychology* **24**(1) 1–55.

Hong, H., M. Kacperczyk. 2009. The price of sin: The effects of social norms on markets. *Journal of Financial Economics* **93**(1) 15–36.

Huck, S., D. Kübler, J. Weibull. 2012. Social norms and economic incentives in firms. *Journal of Economic Behavior and Organization* **83**(2) 173–185.

Huck, S., J. Oechssler. 1999. The indirect evolutionary approach to explaining fair allocations. *Games and Economic Behavior* **28**(1) 13–24.

Hull, D. L. 2001. *Science and selection: Essays on biological evolution and the philosophy of science.* Cambridge University Press.

Hurwicz, L. 1993. Toward a framework for analyzing institutions and institutional change. In *Markets and democracy: Participation, accountability and efficiency*, edited by S. Bowles, H. Gintis and B. Gustafsson. Cambridge University Press. pp. 51–67.

Hurwicz, L. 1996. Institutions as families of games forms*. *Japanese Economic Review* **47**(2) 113–132.

Janssen, M. C. W., J. Gorter, van de Meerendonk, Sjoerd. 1997. Cooperation in a modified version of the finitely repeated prisoners' dilemma game. *Journal of Economic Behavior & Organization* **32**(4) 613–619.

Janssen, M.A., W. Jager. 2001. Fashions, habits and changing preferences: Simulation of psychological factors affecting market dynamics. *Journal of Economic Psychology* **22**(6) 745–772.

Johnson, S. 2004. Gender norms in financial markets: evidence from Kenya. *World Development* **32**(8) 1355–1374.

Kahneman, D., S. Frederick. 2002. Representativeness revisited: Attribute substitution in intuitive judgment. In *Heuristics of intuitive judgment: Extensions and applications*, edited by T.Gilovich, D. Griffin and D. Kahneman. Cambridge University Press.

Kandori, M. 1992. Social norms and community enforcement. *The Review of Economic Studies* **59**(1) 63–80.

Kim, A.M. 2007. North versus South: The Impact of Social Norms in the Market Pricing of Private Property Rights in Vietnam. *World Development* **35**(12) 2079–2095.

Koch, L. P. 2008. *Evolution and correlated equilibrium,* Bonn econ discussion papers.

Koçkesen, L., E. A. Ok, R. Sethi. 2000a. Evolution of interdependent preferences in aggregative games. *Games and Economic Behavior* **31**(2) 303–310.

Koçkesen, L., E. A. Ok, R. Sethi. 2000b. The strategic advantage of negatively interdependent preferences. *Journal of Economic Theory* **92**(2) 274–299.

Kreps, D. M., P. Milgrom, J. Roberts, R. Wilson. 1982. Rational cooperation in the finitely repeated prisoners' dilemma. *Journal of economic theory* **27**(2) 245–252.

Kurz, M. 1977. Altruistic equilibrium. Bela Balassa and Richard Nelson, Economic Progress, Private Values and Public Policy 177–200.

Leimar, O., P. Hammerstein. 2001. Evolution of cooperation through indirect reciprocity. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **268**(1468) 745–753.

Lenzo, J., T. Sarver. 2006. Correlated equilibrium in evolutionary models with subpopulations. *Games and Economic Behavior* **56**(2) 271–284.

Lindbeck, A., S. Nyberg, J. Weibull. 1999. Social norms and economic incentives in the welfare state. *The Quarterly Journal of Economics* **114**(1) 1–35.

Maynard Smith, J., G. R. Price. 1973. The Logic of Animal Conflict. *Nature* **246** 15–18.

MiD – Mobilität in Deutschland 2008. Ergebnisbericht; Struktur – Aufkommen – Emission – Trends. infas; DLR. Bundesministerium für Verkehr, Bau und Stadtentwicklung (Hg.).

Milgrom, P. R., D. C. North, B. R. Weingast. 1990. The role of institutions in the revival of trade: The law merchant, private judges, and the champagne fairs. *Economics & Politics* **2**(1) 1–23.

Nelson, R. R. 1995. Recent Evolutionary Theorizing About Economic Change. *Journal of Economic Literature* (33) 49–90.

Nelson, R. R., S. G. Winter. 1982. *An evolutionary theory of economic change.* Harvard University Press.

Neyman, A. 1985. Bounded complexity justifies cooperation in the finitely repeated prisoners' dilemma. *Economics Letters* **19**(3) 227–229.

North, D. C. 1990. *Institutions, institutional change and economic performance.* Cambridge University Press.

Nowak, M. A., K. Sigmund. 1998. Evolution of indirect reciprocity by image scoring. *Nature* **393**(6685) 573–577.

Nyarko, Y. 1994. Bayesian learning leads to correlated equilibria in normal form games. *Economic Theory* **4**(6) 821–841.

Nyborg, K., R.B. Howarth, K.A. Brekke. 2006. Green consumers and public policy: On socially contingent moral motivation. *Resource and energy economics* **28**(4) 351–366.

Okuguchi, K., F. Szidarovszky, K. Okuguchi. 1990. *The theory of oligopoly with multi-product firms.* Berlin: Springer.

Platteau, J. 1994. Behind the market stage where real societies exist. Part II: the role of moral norms. *The Journal of Development Studies* **30** 753–817.

Possajennikov, A. 2000. On the evolutionary stability of altruistic and spiteful preferences. *Journal of Economic Behavior & Organization* **42**(1) 125–129.

Poulsen, A., O. Poulsen. 2006. Endogenous preferences and social-dilemma institutions. *Journal of Institutional and Theoretical Economics* **162**(4) 627–660.

Robinson, D., D. Goforth. 2005. *The Topology of the 2× 2 Games.* New York: Routledge.

Roth, A. Bargaining experiments, in: Kagel J., Roth A., The Handbook of Experimental Economics, 1995. Princeton: Princeton University Press.

Rubinstein, A. 1979. Equilibrium in supergames with the overtaking criterion. *Journal of economic theory* **21**(1) 1–9.

Rutherford, M. 2001. Institutional economics: then and now. *Journal of Economic Perspectives* **15**(3) 173–194.

Samuelson, L. 1997. Evolutionary games and equilibrium selection. Cambridge: MIT Press.

Samuelson, P.A. 1947. *Foundations of Economic Analysis*. Cambridge: Harvard Univ. Press.

Schotter, A. 1981. *The economic theory of social institutions*. New York: Cambridge University Press.

Sethi, R. 1996. Evolutionary stability and social norms. *Journal of Economic Behavior & Organization* **29**(1) 113–140.

Sethi, R., E. Somanathan. 2001. Preference evolution and reciprocity. *Journal of Economic Theory* **97**(2) 273–297.

Simon, H. A. 1976. From Substantive to procedural rationality. In *25 Years of Economic Theory,* edited by T. J. Kastelain et al. pp. 65–86.

Slonim, R., E. Garbarino. 2008. Increases in trust and altruism from partner selection: Experimental evidence. *Experimental Economics* **11**(2) 134–153.

Smith, E. A., R. Bliege Bird. 2000. Turtle hunting and tombstone opening: public generosity as costly signaling. *Evolution and Human Behavior* **21**(4) 245–261.

Soler, M. 2012. Costly signaling, ritual and cooperation: evidence from Candomblé, an Afro-Brazilian religion. *Evolution and Human Behavior* **33**(4) 346–356.

Spence, M. 1973. Job market signaling. *The Quarterly Journal of Economics* **87**(3) 355–374.

Sugden, R. 1986. The economics of rights, co-operation and welfare. Oxford: Basil Blackwell.

Sugden, R. 1989. Spontaneous order. *The Journal of Economic Perspectives* **3**(4) 85–97.

Tan, T. C.-C., da Costa Werlang, Sérgio Ribeiro. 1988. The Bayesian foundations of solution concepts of games. *Journal of Economic Theory* **45**(2) 370–391.

Taylor, M. 1976. *Anarchy and cooperation*. London: Wiley.

Trivers, R. L. 1971. The evolution of reciprocal altruism. *Quarterly review of biology* **46**(1) 35–57.

Ullmann-Margalit, E. 1977. *The emergence of norms.* Oxford: Clarendon Press.

United Nations. 2002. Plan of Implementation of the World Summit on Sustainable Development. In *Report of the World Summit on Sustainable Development*, UN Document A/CONF.199/20*, New York: UN.

Wedekind, C., M. Milinski. 2000. Cooperation through image scoring in humans. *Science* **288**(5467) 850–852.

Weibull, J. W. 1997. *Evolutionary game theory.* Cambridge: MIT Press.

Weidlich, W. and Haag, G. 1983. *Concepts and Models of a Quantitative Sociology: The Dynamics of Interacting Populations*. Heidelberg and New York: Springer.

Weingast, B. R., W. J. Marshall. 1988. The industrial organization of Congress; or, why legislatures, like firms, are not organized as markets. *The Journal of Political Economy* **96**(1) 132–163.

Witt, U. 1986. Evolution and stability of cooperation without enforceable contracts. *Kyklos* **39**(2) 245–266.

Witt, U. 2003. *The evolving economy: essays on the evolutionary approach to economics*. Edward Elgar Publishing.

Wright, J. 1999. Altruism as a signal: Zahavi's alternative to kin selection and reciprocity. *Journal of Avian Biology* **30**(1) 108–115.

Yamagishi, T., N. Mifune, Y. Li, M. Shinada, H. Hashimoto, Y. Horita, A. Miura, K. Inukai, S. Tanida, T. Kiyonari. 2013. Is behavioral pro-sociality game-specific? Pro-social preference and expectations of pro-sociality. *Organizational Behavior and Human Decision Processes* **120**(2) 260–271.

A. Appendix to Chapter 2

Proof of Proposition 2-1: The demand system (2.3) in vector notion is given by: $\begin{pmatrix} p^e \\ p^g \end{pmatrix} = A\begin{pmatrix} X^e \\ X^g \end{pmatrix} + b$. According to Okuguchi and Szidarovszky (1990, p.34), given the linear structure of the model, negative definiteness of $A + A^T$ is sufficient for uniqueness of the Cournot equilibrium. Eigenvalues of $A + A^T$ are given by $-\frac{1}{\kappa - \lambda}\left(1 \pm \sqrt{1 - \frac{4\lambda(\kappa - \lambda)}{(\kappa + \lambda)^2}}\right)$ and negative by inspection. QED

Proof of Lemma 2-1: $\frac{\partial m^*}{\partial q} \gtreqless 0 \Leftrightarrow \frac{\chi_a^e - \chi_h^e}{\sqrt{k\kappa}} \gtreqless 0$. QED

Equilibrium prices of system (2.3) are given by:

$$p^e = \frac{1}{\kappa^2 - \lambda^2}\left(\left(q\chi_a^e + (1-q)\chi_h^e\right)\left(\frac{\kappa^2 - \lambda^2}{(1+m)\kappa} + \frac{\lambda^2}{(n+1)\kappa}\right) + \frac{\lambda}{n+1}\left(q\chi_a^g + (1-q)\chi_h^g\right)\right) + c^e\frac{m}{m+1} + c^g\frac{(n-m)\lambda}{(n+1)(m+1)\kappa}$$

$$p^g = \frac{1}{n+1}\frac{\left(q\chi_a^g + (1-q)\chi_h^g\right)\kappa + \left(q\chi_a^e + (1-q)\chi_h^e\right)\lambda}{\kappa^2 - \lambda^2} + \frac{n}{n+1}c^g$$

Proof of Lemma 2-2: The proof is given in the paper. QED

Derivation of vertices of MES:

The lower left vertex (point A) is given by $\begin{pmatrix} \Delta^e \\ \Delta^g \end{pmatrix}^A = \begin{pmatrix} -\theta^e + \frac{(1-CB)\alpha\delta_h}{(1+n)(1-\alpha)\delta_a}\left(n\theta^g - \lambda\theta^e\right) \\ -\theta^g + \frac{(1-CB)\alpha\delta_h\lambda}{n(1+n)(1-\alpha)\delta_a}\left(n\theta^g - \lambda\theta^e\right) \end{pmatrix}$, because $q^{Max}\left(\left(\Delta^e,\Delta^g\right)^A\right) = 1$ and $\dot{q}\left(q^{Max} = 1\right) = -\frac{X^g}{X^e}(1)$. Hence, $\left(\Delta^e,\Delta^g\right)^A$ is a solution to $X^g(1) = 0$ and $\dot{q}\left(q^{Max} = 1\right) = 0$.

The upper left vertex (point B) is derived by similar conditions, $\dot{q}\left(q^B\right) = 0$, $\dot{q}'\left(q^B\right) = 0$ and $\tilde{X}^g(1) = 0$. The first two conditions are reduced to:

$\frac{\alpha}{1-\alpha}(lq+k)Z + (1-q)Z^2 = q,\ \frac{\alpha}{1-\alpha}l - Z + (1-q)Z' - \frac{1}{Z} + q\frac{Z'}{Z^2} = 0$, where $lq + k = (1+CB)\sigma_a - q\left((1+CB)\sigma_a + (1-CB)\sigma_h\right)$. After some algebra, it turns out that

$q^B = \frac{\theta^e - \frac{\alpha}{1-\alpha}\frac{(n+1)\left(l\theta^e - k\Delta^e\right)}{n\theta^g - \lambda\theta^e}\theta^e}{\Delta^e + 2\theta^e - \frac{\alpha}{1-\alpha}\frac{(n+1)\left(l\theta^e - k\Delta^e\right)}{n\theta^g - \lambda\theta^e}\Delta^e}$. Again, to obtain a relation between Δ^e and Δ^g, we plug this value into $\dot{q}\left(q^{B}\right)=0$. This gives us the third vertex:

$$\begin{pmatrix}\Delta^{e}\\ \Delta^{g}\end{pmatrix}^{B}=\left(\begin{array}{l}\dfrac{1}{(1+CB)^{2}(1+n)^{2}\alpha^{2}\delta_{a}^{2}(-2+\delta_{h})}\\ \left(\begin{array}{l}-(1+CB)(1+n)^{2}\alpha^{2}\delta_{a}\delta_{h}(-1+\delta_{a}+CB(1+\delta_{a}-\delta_{h})+\delta_{h})\theta^{e}\\ +(1+n)\delta_{a}((1+n)(-2+\alpha(4+\alpha(-2+2(1+CB)^{2}\delta_{a}+\delta_{h}-CB^{2}\delta_{h})))\theta^{e}\\ +(1+CB)n(-1+\alpha)\alpha\theta^{g}-(1+CB)(-1+\alpha)\alpha\lambda\theta^{e})+\sqrt{\Upsilon}\end{array}\right)\\ \dfrac{\lambda}{n}\left(\Delta^{e}\right)^{B}+\dfrac{\lambda}{n}\theta^{e}\end{array}\right)$$

$$\begin{aligned}\Upsilon\equiv{}&((1+n)^{2}\delta_{a}^{2}(((1+n)(2+\alpha(-4+\alpha(2+(-1+CB^{2})\delta_{h})))\theta^{e}-(1+CB)n(-1+\alpha)\alpha\theta^{g}+(1+CB)(-1+\alpha)\alpha\theta^{e}\lambda)^{2}\\ &+(1+CB)^{2}\alpha^{2}\delta_{h}^{2}(-n\theta^{g}+\lambda\theta^{e}+\alpha((-1+CB)(1+n)\theta^{e}+n\theta^{g}-\lambda\theta^{e}))^{2}-\\ &2(1+CB)\alpha^{2}\delta_{h}((1+CB)n^{2}(-1+\alpha)^{2}\left(\theta^{g}\right)^{2}+(1+CB)n(-1+\alpha)\theta^{e}\theta^{g}(-(-1+CB)(1+n)\alpha(-3+\delta_{h})+2\lambda-2\alpha\lambda)\\ &+\left(\theta^{e}\right)^{2}((-1+CB)(1+n)^{2}(2(-1+\delta_{h})+\alpha(4-4\delta_{h}+\alpha(-2+\delta_{h}+CB^{2}\delta_{h})))\\ &+(-1+CB^{2})(1+n)(-1+\alpha)\alpha(-3+\delta_{h})\lambda+(1+CB)(-1+\alpha)^{2}\lambda^{2})))\end{aligned}$$

We can solve for the upper right vertex (point C) only for $\alpha=0$. It is given by the intersection of $\Delta^{e,Max}\left(\Delta^{g}\right)$ and $\Delta^{e,Min}\left(\Delta^{g}\right)$, characterized by solutions to $\dot{q}\left(q^{Max}\right)=0$ and $\dot{q}\left(q^{Min}\right)=0$. It follows that such a $\left(\Delta^{e},\Delta^{g}\right)$ pair is given by three conditions $\dot{q}\left(q^{C}\right)=0$, $\dot{q}'\left(q^{C}\right)=0$ and $\dot{q}''\left(q^{C}\right)=0$. We can solve for q^{IP} explicitly: $q^{C}=\dfrac{n\theta^{g}-\lambda\theta^{e}}{3\left(n\Delta^{g}-\Delta^{e}\lambda\right)+4\left(n\theta^{g}-\lambda\theta^{e}\right)}$. However, $\left(\Delta^{e},\Delta^{g}\right)^{C}$ cannot be solved analytically.

q^{C} is derived by rewriting $\dot{q}\left(q^{C}\right)=0$, $\dot{q}'\left(q^{C}\right)=0$ and $\dot{q}''\left(q^{C}\right)=0$ as $Z^{2}(q)=\dfrac{q}{1-q}$, $(1-q)Z'-Z-\dfrac{1}{Z}+q\dfrac{Z'}{Z^{2}}=0$ and $-2Z'+2\dfrac{Z'}{Z}+(1-q)Z''+q\dfrac{Z''}{Z}-2q\dfrac{(Z')^{2}}{Z^{3}}=0$, where $Z=\dfrac{\tilde{X}^{e}}{\tilde{X}^{g}}$. The definition of Z implies the following relation between Z and its second derivative: $Z''=-2Z'\dfrac{\tilde{X}'^{g}}{\tilde{X}^{g}}$.

<u>Proof of Lemma 2-3:</u>

The situation where MES is empty corresponds to the case point A, B and C are equal, i.e. where $q^{Max}=q^{Min}\equiv q^{IP}=1$, $\dot{q}(1)=0$ and $X^{g}(1)=0$. The latter two conditions provide a solution for Δ^{g} as a function of α: $\Delta^{g}(\alpha)=\dfrac{\alpha\lambda\delta_{h}(1-CB)\left(n\theta^{g}-\lambda\theta^{e}\right)}{n(n+1)(1-\alpha)\delta_{a}}-\theta^{g}$. The first condition amounts to a condition for α as a function of Δ^{g}:

$$\Delta\alpha\left(\Delta^{g}\right)=$$

$$\frac{2n(n+1)^{2}\delta_{a}(\Delta^{g}+\theta^{g})\left(n\Delta^{g}+\tau\right)+\delta_{h}\lambda^{2}\tau^{2}}{2n(n+1)^{2}\delta_{a}(\Delta^{g}+\theta^{g})\left(n\Delta^{g}+\tau\right)+\delta_{h}\lambda^{2}\tau^{2}+(1+n)\lambda\tau\left(n\left((1+CB)\delta_{a}+2(1-CB)\delta_{h}\right)(\Delta^{g}+\theta^{g})-(1-CB)\delta_{h}\theta^{e}\lambda\right)}$$

, where $\tau=\left(n\theta^{g}-\lambda\theta^{e}\right)$. Solving these two equations for α yields the critical value stated in the lemma. QED

Proof of Lemma 2-4:

For $\alpha = 0$, the approximation strategy is described in the paper. We therefore present only the general solution for the tangent point D:

$$\left(\Delta^e\right)^D = \frac{1}{4(1+n)^2(1+\alpha(-2+CB^2\alpha))\delta_a}\left(-\Omega+\sqrt{\frac{8(1+n)^2(1+\alpha(-2+CB^2\alpha))\delta_a\theta^e}{q^3(1-\alpha)\delta_h}\Psi+\Omega^2}\right)$$

$$\left(\Delta^g\right)^D = \frac{1}{2nq^2(1-\alpha)\delta_h}\left(\begin{array}{l} -q^2\left(\Delta^e\right)^D\left(\begin{array}{l}(1+n)\alpha\left((1+CB)\delta_a+(1-CB)\delta_h\right)\\ -2(1-\alpha)\delta_h\lambda\end{array}\right)\\ +q\left(\begin{array}{l}-2\delta_h\tau+\alpha((1+CB)(1+n)\delta_a(\left(\Delta^e\right)^D-\theta^e)\\ +\delta_h((-1+CB)(1+n)\theta^e+2\tau))+(1+n)\left(\Delta^e\right)^D\Sigma\end{array}\right)\\ +(1+n)\theta^e\left((1+CB)\alpha\delta_a+\Sigma\right)\end{array}\right)$$

, where

$$\tau \equiv \left(n\theta^g - \lambda\theta^e\right)$$

$$\Sigma \equiv \sqrt{(1+CB)^2(-1+q)^2\alpha^2\delta a^2 - 2(-1+q)q(2+\alpha(-4+\alpha+CB^2\alpha))\delta a\delta h + (-1+CB)^2 q^2\alpha^2\delta h^2}$$

$$\begin{aligned}\Psi \equiv\ & (1+n)(2q\alpha\delta_a\delta_h(-3(1+n)(-3+2q)\theta^e + 2(1+CB)n(-1+q)\theta^g - 2(1+CB)(-1+q)\theta^e\lambda) + \\ & \alpha^3(-2q\delta_a\delta_h((1+n)(-1+CB^2(-2+q)+q)\theta^e - (1+CB)(1+CB^2)n(-1+q)\theta^g + (1+CB)(1+CB^2)(-1+q)\theta^e\lambda) + \\ & (1+CB)^2(-1+q)^2\delta_a^{\,2}(-(1+CB)n\theta^g + \theta^e(1+n+\lambda+CB\lambda)) + (-1+CB)^2q^2\delta_h^{\,2}(-(1+CB)n\theta^g + \theta^e(1+n+\lambda+CB\lambda))) + \\ & \alpha(-(1+CB)(1+n)(-1+q)\delta_a\theta^e + q\delta_h(-(-1+CB)(1+n)\theta^e - 4n\theta^g + 4\theta^e\lambda))\Sigma + 2q\delta_h((1+n)(-3+2q)\delta_a\theta^e + (n\theta^g - \theta^e\lambda)\Sigma) - \\ & \alpha^2((1+CB)^2(1+n)(-1+q)^2\delta_a^{\,2}\theta^e + \delta_a(2q\delta_h(-(1+n)(-7+CB^2(-2+q)+5q)\theta^e + 4(1+CB)n(-1+q)\theta^g - 4(1+CB)(-1+q)\theta^e\lambda) - \\ & (1+CB)(-1+q)(-(1+CB)n\theta^g + \theta^e(1+n+\lambda+CB\lambda))\Sigma) + q\delta_h((-1+CB)^2(1+n)q\delta_h\theta^e + (-(1+CB^2)n\theta^g + \theta^e(1+n+\lambda+CB(-1-n+CB\lambda)))\Sigma)))\end{aligned}$$

$$\Omega \equiv \frac{1}{q^2(-1+\alpha)\delta_h}\left(\Psi + 2(1+n)^2q(-1+\alpha)(1+\alpha(-2+CB^2\alpha))\delta_a\delta_h\theta^e\right)$$

To find such a point, we apply the following approach: First, we express two of the conditions for the inflection point C $\dot{q}(q)=0, \dot{q}'(q)=0$ in terms of $\Delta^g(q)$. With these two conditions, we can solve for Δ^e. However, we still have to find a q that will be greater than q^{IP} and independent of Δ^e and Δ^g.

For the general case $\alpha \neq 0$, we again choose q such that we can be sure that it will correspond to a point on the graph of $\Delta^{e,Max}\left(\Delta^g\right)$. This can be achieved by choosing $\left(\Delta^g\right)^B$ as a lower bound for Δ^g and $-\theta^e$ as a lower bound for Δ^e.

$$\underline{q} = q^{IP}\left(\Delta^g = \left(\Delta^g\right)^{\mathrm{B}}, \Delta^e = -\theta^e\right) =$$

$$\frac{\left(\begin{array}{l}\delta_h\left[(1-CB)\alpha(1+n)n\left(\Delta^g\theta^e+\theta^g\Delta^e\right)+2n(1-\alpha)\left(n\Delta^g-\lambda\Delta^e\right)\left(n\theta^g-\theta^e\lambda\right)-2\lambda\left(\left((1-CB)(1+n)\alpha\Delta^e\right)\theta^e\right)\right]\\ +(1+n)\delta_a\left[(1+CB)n\alpha\left(\Delta^g\theta^e+\theta^g\Delta^e\right)-\left(\Delta^e\right)^2((1+n)(1-\alpha)-(1+CB)\alpha\lambda)+\Delta^e\left(n(2\theta^e-\alpha((1+CB)\Delta^g+2\theta^e))+2\theta^e(1-\alpha-(1+CB)\alpha\lambda)\right)\right]\end{array}\right)}{3((1+n)\delta_a\Delta^e((1+n)(-1+\alpha)\Delta^e-(1+CB)n\alpha\Delta^g+(1+CB)\alpha\Delta^e\lambda)-\delta_h(n\Delta^g-\Delta^e\lambda)(n\Delta^g-\Delta^e\lambda+\alpha((1-CB)(1+n)\Delta^e+n\Delta^g-\Delta^e\lambda)))}$$

This gives us a lower bound for the maxima that correspond to $\Delta^{e,Max}\left(\Delta^g\right)$ independent of Δ^e and Δ^g. We can then calculate the slope at point D: $\frac{d\Delta^e}{d\Delta^e} = \frac{\left(\Delta^e\right)^{\mathrm{D}}\underline{q}+\theta^e}{\left(\Delta^g\right)^{\mathrm{D}}\underline{q}+\theta^g}$. QED

Proof of Lemma 2-5: Inserting equilibrium quantities given in (2.13) into (2.24) and reformulation yields equation (2.25). QED

Proof of Lemma 2-6:

(1.) At the discontinuities we have $m^* = m^{eq}$ and thus $\dot{q}(q) = \dot{\hat{q}}(q)$. Otherwise, $m^* > m^{eq}$ implies $\tilde{X}^e > \hat{X}^e$ and $\tilde{X}^g < \hat{X}^g$ due to $\frac{dX^{e*}}{dm} > 0$ and $\frac{dX^{g*}}{dm} < 0$. Hence, $\dot{q}(q) < \dot{\hat{q}}(q)$ for all q in the intervals of continuity. For the second part of the claim, note that we can write $\dot{q} = \dot{q}\left(\frac{X^e}{X^g}(m(q),q),q\right)$ and thus, $\frac{d\dot{q}}{dq} = \frac{\partial \dot{q}}{\partial \frac{X^e}{X^g}}\left(\frac{\partial \frac{X^e}{X^g}}{\partial m}\frac{dm}{dq} + \frac{\partial \frac{X^e}{X^g}}{\partial q}\right) + \frac{\partial \dot{q}}{\partial q}$. Since $\frac{dm^*}{dq} = \frac{\Delta^e}{\sqrt{k\kappa}}$ and $\frac{dm^{eq}}{dq} = 0$ for all q in the intervals of continuity, and the other terms in $\frac{d\dot{q}}{dq}$ are the same for the discontinuous version of $\dot{q}$ and its continuous approximation $\dot{\hat{q}}$, the observation $\frac{\partial \dot{q}}{\partial \frac{X^e}{X^g}} = (1-\alpha)\left((1-q)\sigma_a + \frac{q\sigma_h}{\left(\frac{X^e}{X^g}\right)^2}\right) > 0$ implies the second claim of the lemma.

(2.) The distance between two discontinuities is a natural multiple of $\sqrt{k\kappa}/\Delta^e$ because $\frac{dm^*}{dq} = \frac{\Delta^e}{\sqrt{k\kappa}}$ and thus m^* reaches the next integer at this frequency. Finally, for $\Delta^e > 0$, the lower limit at the discontinuities is obviously smaller than the upper limit and size of the "jumps" of $\hat{X}^e$ at the discontinuity q_i, which is given by $\left(\frac{\hat{m}}{\hat{m}+1} - \frac{\hat{m}-1}{\hat{m}}\right)\left(\chi_h^e + \Delta^e q_i - \kappa c^e + \lambda c^g\right) = \frac{\sqrt{k\kappa}}{\hat{m}}$, where $\hat{m} = m^*(q_i)$. Since m^* grows in q, the size of the "jumps" declines in q. For $\Delta^e < 0$, exactly the opposite is true. QED

Derivation of the partial effects on the critical value $\alpha^{crit.}$:

$$\alpha^{crit.} = \frac{(1-CB)(n+1)\theta^e + 2\left(n\theta^g - \lambda\theta^e\right) - \sqrt{\left((1-CB)(n+1)\theta^e\right)^2 + 4\left(n\theta^g - \lambda\theta^e\right)^2\left(1-CB^2\right)}}{2\left(CB^2\left(n\theta^g - \lambda\theta^e\right) + (n+1)(1-CB)\theta^e\right)}$$

$$\underset{y=2\left(n\theta^g-\lambda\theta^e\right)>0}{\overset{x=(1-CB)(n+1)\theta^e>0}{=}} \frac{x + y - \sqrt{x^2 + y^2\left(1-CB^2\right)}}{CB^2 y + 2x}$$

$$1.claim:\frac{\partial \alpha^{crit.}}{\partial \theta^{g}}>0:$$

$$\frac{\partial \alpha^{crit.}}{\partial \theta^{g}}=\frac{\left(y'-\frac{2yy'\left(1-CB^{2}\right)}{2\sqrt{x^{2}+y^{2}\left(1-CB^{2}\right)}}\right)\left(CB^{2}y+2x\right)-\left(x+y-\sqrt{x^{2}+y^{2}\left(1-CB^{2}\right)}\right)\left(CB^{2}y'\right)}{\left(CB^{2}y+2x\right)^{2}}$$

$$=\frac{y'\left(CB^{2}y+2x\right)-\frac{2yy'\left(1-CB^{2}\right)\left(CB^{2}y+2x\right)}{2\sqrt{x^{2}+y^{2}\left(1-CB^{2}\right)}}-xCB^{2}y'-yCB^{2}y'+CB^{2}y'\sqrt{x^{2}+y^{2}\left(1-CB^{2}\right)}}{\left(CB^{2}y+2x\right)^{2}}$$

$$=\frac{\left(2-CB^{2}\right)xy'-\frac{2yy'\left(1-CB^{2}\right)\left(CB^{2}y+2x\right)}{2\sqrt{x^{2}+y^{2}\left(1-CB^{2}\right)}}+CB^{2}y'\sqrt{x^{2}+y^{2}\left(1-CB^{2}\right)}}{\left(CB^{2}y+2x\right)^{2}}>0\Leftrightarrow$$

$$\left(2-CB^{2}\right)xy'2\sqrt{x^{2}+y^{2}\left(1-CB^{2}\right)}-2yy'\left(1-CB^{2}\right)\left(CB^{2}y+2x\right)+CB^{2}y'2\left(x^{2}+y^{2}\left(1-CB^{2}\right)\right)>0\overset{y'>0}{\Leftrightarrow}$$

$$\left(2-CB^{2}\right)x\sqrt{x^{2}+y^{2}\left(1-CB^{2}\right)}-y\left(1-CB^{2}\right)\left(CB^{2}y+2x\right)+CB^{2}\left(x^{2}+y^{2}\left(1-CB^{2}\right)\right)>0\overset{x>0}{\Leftrightarrow}$$

$$\left(2-CB^{2}\right)\sqrt{x^{2}+y^{2}\left(1-CB^{2}\right)}-2y\left(1-CB^{2}\right)+xCB^{2}>0\Leftrightarrow$$

$$\left(2-CB^{2}\right)\sqrt{x^{2}+y^{2}\left(1-CB^{2}\right)}>2y\left(1-CB^{2}\right)-xCB^{2}\Leftarrow$$

$$\left(2-CB^{2}\right)^{2}\left(x^{2}+y^{2}\left(1-CB^{2}\right)\right)>\left(2y\left(1-CB^{2}\right)-xCB^{2}\right)^{2}\Leftrightarrow$$

$$\left(2-CB^{2}\right)^{2}\left(x^{2}+y^{2}\left(1-CB^{2}\right)\right)-\left(2y\left(1-CB^{2}\right)-xCB^{2}\right)^{2}=\left(1-CB^{2}\right)\left(2x+yCB^{2}\right)^{2}>0$$

$$2.claim: \frac{\partial \alpha^{crit.}}{\partial CB} > 0$$

$$\frac{\partial \alpha^{crit.}}{\partial CB} = \frac{\left(x' - \frac{2xx' - 2CBy^2}{2\sqrt{x^2 + y^2\left(1 - CB^2\right)}}\right)\left(CB^2 y + 2x\right) - \left(x + y - \sqrt{x^2 + y^2\left(1 - CB^2\right)}\right)\left(2CBy + 2x'\right)}{\left(CB^2 y + 2x\right)^2}$$

$$= \frac{x'\left(CB^2 y + 2x\right) - \frac{2xx' - 2CBy^2}{2\sqrt{x^2 + y^2\left(1 - CB^2\right)}}\left(CB^2 y + 2x\right) - \left(x + y - \sqrt{x^2 + y^2\left(1 - CB^2\right)}\right)\left(2CBy + 2x'\right)}{\left(CB^2 y + 2x\right)^2} > 0$$

$$\Leftrightarrow$$

$$x'\left(CB^2 y + 2x\right) - \frac{2xx' - 2CBy^2}{2\sqrt{x^2 + y^2\left(1 - CB^2\right)}}\left(CB^2 y + 2x\right) - \left(2CBy + 2x'\right)\left(x + y\right) + \left(2CBy + 2x'\right)\sqrt{x^2 + y^2\left(1 - CB^2\right)} > 0$$

$$\Leftrightarrow$$

$$-\frac{2xx' - 2CBy^2}{2\sqrt{x^2 + y^2\left(1 - CB^2\right)}}\left(CB^2 y + 2x\right) - \left(2CBy\right)\left(x + y\right) - \left(2 - CB^2\right)x'y + \left(2CBy + 2x'\right)\sqrt{x^2 + y^2\left(1 - CB^2\right)} > 0$$

$$\overset{*2\sqrt{x^2+y^2\left(1-CB^2\right)}}{\Leftrightarrow}$$

$$-\left(2xx' - 2CBy^2\right)\left(CB^2 y + 2x\right) - \left(2CBy\left(x + y\right) + \left(2 - CB^2\right)x'y\right)2\sqrt{x^2 + y^2\left(1 - CB^2\right)} + \left(2CBy + 2x'\right)2\left(x^2 + y^2\left(1 - CB^2\right)\right) > 0$$

$$\Leftrightarrow$$

$$-2xx'\left(CB^2 y + 2x\right) + 2CBy^2\left(CB^2 y + 2x\right) + \left(2CBy + 2x'\right)2x^2 + \left(2CBy + 2x'\right)2y^2\left(1 - CB^2\right)$$
$$-\left(2CBy\left(x + y\right) + \left(2 - CB^2\right)x'y\right)2\sqrt{x^2 + y^2\left(1 - CB^2\right)} > 0$$

$$\Leftrightarrow$$

$$-2xx'CB^2 y + 2CBy^2\left(CB^2 y + 2x\right) + 2CBy2x^2 + \left(2CBy + 2x'\right)2y^2 - CB^2\left(2CBy + 2x'\right)2y^2$$
$$-\left(2CBy\left(x + y\right) + \left(2 - CB^2\right)x'y\right)2\sqrt{x^2 + y^2\left(1 - CB^2\right)} > 0$$

$$\overset{:2y,\, y>0}{\Leftrightarrow}$$

$$-xx'CB^2 + CBy\left(CB^2 y + 2x\right) + CB2x^2 + \left(2CBy + 2x'\right)y - CB^2\left(2CBy + 2x'\right)y$$
$$-\left(2CB\left(x + y\right) + \left(2 - CB^2\right)x'\right)\sqrt{x^2 + y^2\left(1 - CB^2\right)} > 0$$

$$\Leftrightarrow$$

$$-xx'CB^2 + CBy\left(CB^2 y + 2x\right) + CB2x^2 + 2CBy^2 - CB^2 2CBy^2 + 2x'y\left(\underbrace{1 - CB^2}_{=(1-CB)(1+CB)}\right)$$
$$-\left(2CB\left(x + y\right) + \left(2 - CB^2\right)x'\right)\sqrt{x^2 + y^2\left(1 - CB^2\right)} > 0$$

$$\overset{x' = -\frac{x}{1-CB}}{\Leftrightarrow}$$

$$-xx'CB^2 + CBy\left(CB^2 y + 2x\right) + CB2x^2 + 2CBy^2 - CB^2 2CBy^2 - 2xy\left(1 + CB\right)$$
$$-\left(2CB\left(x + y\right) + \left(2 - CB^2\right)x'\right)\sqrt{x^2 + y^2\left(1 - CB^2\right)} > 0$$

$$\Leftrightarrow$$

$$\underbrace{-xx'CB^2}_{>0,\, x'<0} + CBy^2\underbrace{\left(2 - CB^2\right)}_{>0} - 2xy + 2x^2 CB - \left(2CB\left(x + y\right) + \left(2 - CB^2\right)x'\right)\sqrt{x^2 + y^2\left(1 - CB^2\right)} > 0$$

In the next step, we first rearrange the term on the left-hand side of the last inequality and second, we distinguish two cases to establish the strict positivity of the term.

$$-xx'CB^2 + CBy^2\left(2-CB^2\right) - 2xy + 2x^2CB - \left(2CB\left(x+y\right) + \left(2-CB^2\right)x'\right)\sqrt{x^2+y^2\left(1-CB^2\right)} \overset{x'=-\frac{x}{1-CB}}{=}$$

$$x^2\frac{CB^2}{1-CB} + CBy^2\left(2-CB^2\right) - 2xy + 2x^2CB - \left(2CB\left(x+y\right) - \left(2-CB^2\right)\frac{x}{1-CB}\right)\sqrt{x^2+y^2\left(1-CB^2\right)} \overset{*(1-CB)}{=}$$

$$x^2CB^2 + CBy^2\left(2-CB^2\right)\left(1-CB\right) - 2xy\left(1-CB\right) + 2x^2CB\left(1-CB\right) - \left(2CB\left(1-CB\right)\left(x+y\right) - \left(2-CB^2\right)x\right)\sqrt{x^2+y^2\left(1-CB^2\right)} =$$

$$x^2CB\left(2-CB\right) + y^2CB\left(2-CB^2\right)\left(1-CB\right) - 2xy\left(1-CB\right) - \left(2CB\left(1-CB\right)\left(x+y\right) - \left(2-CB^2\right)x\right)\sqrt{x^2+y^2\left(1-CB^2\right)} =$$

$$x^2CB\left(2-CB\right) + y^2CB\left(2-CB^2\right)\left(1-CB\right) - 2xy\left(1-CB\right) - \left(2CB\left(1-CB\right)y - \left(2+CB^2-2CB\right)x\right)\sqrt{x^2+y^2\left(1-CB^2\right)}$$

<u>1.case:</u> $\left(2CB\left(1-CB\right)y - \left(2+CB^2-2CB\right)x\right) > 0$

$$x^2CB\left(2-CB\right) + y^2CB\left(2-CB^2\right)\left(1-CB\right) - 2xy\left(1-CB\right) - \left(2CB\left(1-CB\right)y - \left(2+CB^2-2CB\right)x\right)\sqrt{x^2+y^2\left(1-CB^2\right)}$$

$$\overset{\sqrt{x^2+y^2\left(1-CB^2\right)}<x+y\sqrt{\left(1-CB^2\right)}}{>}$$

$$x^2CB\left(2-CB\right) + y^2CB\left(2-CB^2\right)\left(1-CB\right) - 2xy\left(1-CB\right) - \left(2CB\left(1-CB\right)y - \left(2+CB^2-2CB\right)x\right)\left(x+y\sqrt{\left(1-CB^2\right)}\right) > 0$$

$$\overset{:(1-CB)}{\Leftrightarrow}$$

$$x^2\left(\frac{CB\left(2-CB^2\right)}{1-CB} + \frac{\left(2+CB^2-2CB\right)}{1-CB}\right) + CBy^2\left(2-CB^2-2\sqrt{1-CB^2}\right) + 2xy\left(-1-CB+\left(2+CB^2-2CB\right)\frac{\sqrt{\left(1-CB^2\right)}}{2\left(1-CB\right)}\right) =$$

$$= x^2\left(\frac{2}{1-CB}\right) + CBy^2\underbrace{\left(2-CB^2-2\sqrt{1-CB^2}\right)}_{\geq 0} + 2xy\underbrace{\left(-1-CB+\frac{1}{2\left(1-CB\right)}\left(1+\left(1-CB\right)^2\right)\right)}_{\geq 0} > 0 \Rightarrow \text{claim}$$

<u>2.case:</u> $\left(2CB\left(1-CB\right)y - \left(2+CB^2-2CB\right)x\right) < 0$

$$x^2CB\left(2-CB\right) + y^2CB\left(2-CB^2\right)\left(1-CB\right) - 2xy\left(1-CB\right) - \left(2CB\left(1-CB\right)y - \left(2+CB^2-2CB\right)x\right)\sqrt{x^2+y^2\left(1-CB^2\right)} =$$

$$x^2CB\left(2-CB\right) + y^2CB\left(2-CB^2\right)\left(1-CB\right) - 2xy\left(1-CB\right) + \underbrace{\left(\left(2+CB^2-2CB\right)x - 2CB\left(1-CB\right)y\right)}_{>0}\sqrt{x^2+y^2\left(1-CB^2\right)}$$

$$\overset{\sqrt{x^2+y^2\left(1-CB^2\right)}>y\sqrt{\left(1-CB^2\right)}}{>}$$

$$x^2CB\left(2-CB\right) + y^2CB\left(2-CB^2\right)\left(1-CB\right) - 2xy\left(1-CB\right) + \left(\left(2+CB^2-2CB\right)x - 2CB\left(1-CB\right)y\right)y\sqrt{\left(1-CB^2\right)} =$$

$$x^2CB\left(2-CB\right) + y^2\left(CB\left(2-CB^2\right)\left(1-CB\right) - 2CB\left(1-CB\right)\sqrt{\left(1-CB^2\right)}\right) - 2xy\left(1-CB\right) + xy\left(2+CB^2-2CB\right)\sqrt{\left(1-CB^2\right)} =$$

$$x^2CB\left(2-CB\right) + \left(1-CB\right)CBy^2\underbrace{\left(2-CB^2-2\sqrt{\left(1-CB^2\right)}\right)}_{\geq 0} + xy\underbrace{\left(\left(2+CB^2-2CB\right)\sqrt{\left(1-CB^2\right)} - 2\left(1-CB\right)\right)}_{\geq 0} > 0 \Rightarrow \text{claim}$$

B. Appendix to Chapter 3

B.1 Proofs

Proof of Lemma 3-1: Full cooperation can only be achieved with only high-types present in the population, i.e. $\lambda = 1$. There are only two equilibria which support cooperation among high-types that are supported at $\lambda = 1$ under certain conditions and potentially exhibit a fitness advantage for high-types (necessary for local stability), the separating cooperative equilibrium and the high pooling cooperative equilibrium. With respect to the former, the support condition amounts to $\frac{\underline{k}}{1+\alpha} \geq 1 \Leftrightarrow \underline{k} \geq 1+\alpha$, and the fitness condition to $\bar{k} < 1$ (see Table 3-3). With respect to the latter, the support condition amoutns to $\underline{k} < 1+\alpha$, and the fitness condition to $(\beta-\alpha)-\beta+\underline{k}-\bar{k} > 0 \Leftrightarrow \underline{k}-\bar{k} > \alpha$. If $\underline{k}-\bar{k} = \alpha$ stability requires a strict positive difference in fitness payoffs for high-types for λ close to 1, i.e. $\beta-\alpha < 0$. QED

Proof of Lemma 3-2: The first pair of inequalities $\frac{\beta}{\bar{m}-\alpha+\beta}\bar{m} < \underline{k}-\bar{k} < \alpha$ arises from the condition of the root ($1-\frac{\underline{k}-\bar{k}-\alpha}{\beta-\alpha}$) of the fitness difference for the high pooling cooperative equilibrium to lie in the support of this equilibrium, i.e. $\max\left\{\frac{\underline{k}}{1+\alpha}, \frac{\beta}{(\beta+\bar{m}-\alpha)}\right\} < 1-\frac{\underline{k}-\bar{k}-\alpha}{\beta-\alpha} < 1$. Stability requires a negative slope of the fitness difference function, i.e. $\beta-\alpha$. Let us first consider $\frac{\underline{k}}{1+\alpha} \leq \frac{\beta}{(\beta+\bar{m}-\alpha)}$. In this case, the within-support condition amounts to $\frac{\beta}{\bar{m}-\alpha+\beta} < 1-\frac{\underline{k}-\bar{k}-\alpha}{\beta-\alpha} < 1$, rearranging yields $\frac{\beta}{\bar{m}-\alpha+\beta}\bar{m} < \underline{k}-\bar{k} < \alpha$. If on the other hand $\frac{\underline{k}}{1+\alpha} > \frac{\beta}{(\beta+\bar{m}-\alpha)}$, the within-support condition amounts to $\frac{\underline{k}}{1+\alpha} < 1-\frac{\underline{k}-\bar{k}-\alpha}{\beta-\alpha} < 1$, rearranging yields $\beta-\frac{\underline{k}}{1+\alpha}(\beta-\alpha) < \underline{k}-\bar{k} < \alpha$. Summarizing gives us $\frac{\beta}{\bar{m}-\alpha+\beta}\bar{m} < \underline{k}-\bar{k} < \alpha$ and $\beta-\frac{\underline{k}}{1+\alpha}(\beta-\alpha) < \underline{k}-\bar{k} \Leftrightarrow \beta < \frac{1+\beta}{1+\alpha}\underline{k}-\bar{k}$. Note that $\frac{\beta}{\bar{m}-\alpha+\beta}\bar{m} < \underline{k}-\bar{k} < \alpha$ implies that $\beta-\alpha < 0$, because $\frac{\beta}{\bar{m}-\alpha+\beta}\bar{m} < \alpha \Leftrightarrow \bar{m}(\beta-\alpha) < \alpha(\beta-\alpha) \overset{\bar{m}>\alpha}{\Leftrightarrow} \beta-\alpha < 0$ QED

Proof of Lemma 3-3: For an inner λ-stable equilibrium to exist at $\lambda = \frac{\underline{k}}{1+\alpha}$, we need (1) the connectivness of the supports of the involved equilibria, (2) a fitness advantage for high-types to the left of $\lambda = \frac{\underline{k}}{1+\alpha}$, (3) a fitness disadvantage for high types to the right of $\lambda = \frac{\underline{k}}{1+\alpha}$ and finally (4) for being an inner equilibrium $\lambda = \frac{\underline{k}}{1+\alpha} \in (0,1)$. (1) gives us $\frac{\beta}{\bar{m}-\alpha+\beta} \leq \frac{\underline{k}}{1+\alpha}$, (2) yields $\frac{\underline{k}}{1+\alpha}-\bar{k} > 0$, (3) amounts to $\frac{\underline{k}}{1+\alpha}(\beta-\alpha)-\beta+\underline{k}-\bar{k} < 0 \Leftrightarrow \frac{\underline{k}}{1+\alpha} < \frac{\bar{k}+\beta}{1+\beta}$, if high-types and

low-types fare equally well at $\lambda = \frac{\underline{k}}{1+\alpha}$, then for stability, high-types need to earn strictly less to the right of $\lambda = \frac{\underline{k}}{1+\alpha}$. In essense, if $\frac{\underline{k}}{1+\alpha} = \frac{\overline{k}+\beta}{1+\beta}$, then $\beta - \alpha < 0$, (4) is equivalent to $\underline{k} < 1+\alpha$.
(1) and (3) are equivalent to

$$\frac{\beta}{\overline{m}-\alpha+\beta} \leq \frac{\underline{k}}{1+\alpha} \underset{\substack{\text{if = then} \\ \alpha>\beta}}{\leq} \frac{\overline{k}+\beta}{1+\beta} \quad (*)$$

(2) and (4) are equivalent to

$$\overline{k} < \frac{\underline{k}}{1+\alpha} < 1 \quad (**)$$

Note that (2) and (3) imply (4), hence what remains is:

$$\frac{\beta}{\overline{m}-\alpha+\beta} \leq \frac{\underline{k}}{1+\alpha} \underset{\substack{\text{if = then} \\ \alpha>\beta}}{\leq} \frac{\overline{k}+\beta}{1+\beta} \text{ and } \overline{k} < \frac{\underline{k}}{1+\alpha}$$ QED

B.2 Stable semi-pooling p-equilibria

	Equilibrium	Support	Conditions for existence	Payoff differentials (superscript "f" indicates difference in fitness payoffs)
$CC\bar{m}$ $CD\underline{m}$ $\underline{m}$	$p_{CD\underline{m}} = \frac{\bar{k} + (1-\lambda)\beta}{\lambda(1+\bar{m})}$ $p_{CC\bar{m}} = 1 - \frac{\bar{k} + (1-\lambda)\beta}{\lambda(1+\bar{m})}$ $p_{\underline{m}} = 1$	1. $\frac{\beta}{(\bar{m}-\alpha+\beta)} < \frac{\bar{k}}{(1+\alpha)} < \frac{1+\bar{m}}{(1+\alpha)}$: $\frac{\bar{k}+\beta}{(1+\bar{m}+\beta)} < \lambda < 1$	2. $\frac{\bar{k}}{(1+\alpha)} \le \frac{\beta}{(\bar{m}-\alpha+\beta)}$: $1 - \frac{(\bar{m}-\alpha)}{\beta}\frac{\bar{k}}{(1+\alpha)} < \lambda < 1$	$\Pi_{\bar{m}}(CD,\underline{m}) - \Pi_{\underline{m}}(\underline{m}) = \lambda(p_{CC\bar{m}})(1+\bar{m}) - \lambda p_{CC\bar{m}}(1+\alpha)$ $= \lambda(\bar{m}-\alpha)p_{CC\bar{m}}$ $= \lambda(\bar{m}-\alpha)\left(1 - \frac{\bar{k}+(1-\lambda)\beta}{\lambda(1+\bar{m})}\right) > 0$ $\left(\Pi_{\bar{m}}(CD,\underline{m}) - \Pi_{\underline{m}}(\underline{m})\right)^f = -\alpha\lambda\left(1 - \frac{\bar{k}+(1-\lambda)\beta}{\lambda(1+\bar{m})}\right) < 0$
$DC\bar{m}$ $CD\underline{m}$ $\underline{m}$ $\bar{m}$	$p_{CD\underline{m}} = \frac{1}{2}\left[1 + \frac{\underline{k}}{\lambda(1+\alpha)}\right]$ $p_{DC\bar{m}} = \frac{1}{2}\left[1 - \frac{\underline{k}}{\lambda(1+\alpha)}\right]$ $p_{\underline{m}} = \frac{1}{2}\left[1 + \frac{1}{(1-\lambda)\beta}\left[\frac{(1+\bar{m})}{(1+\alpha)}\underline{k} - \bar{k}\right]\right]$ $p_{\bar{m}} = \frac{1}{2}\left[1 - \frac{1}{(1-\lambda)\beta}\left[\frac{(1+\bar{m})}{(1+\alpha)}\underline{k} - \bar{k}\right]\right]$	$0 < \frac{\beta + (\underline{k} - \bar{k})}{(\bar{m}-\alpha+\beta)} < \lambda$ $< 1 - \frac{1}{\beta}\left[\frac{(1+\bar{m})}{(1+a)}\underline{k} - \bar{k}\right] < 1$	$\beta(1+\alpha) >$ $\frac{(\bar{m}-\alpha+\beta)}{(\bar{m}-\alpha)}\left[(1+\bar{m})\underline{k} - (1+\alpha)\bar{k}\right]$ $+ \frac{\underline{k}-\bar{k}}{\bar{m}-\alpha}\beta(1+\alpha)$	$\Pi_{\bar{m}}(CD,\underline{m}) - \Pi_{\underline{m}}(\underline{m})$ $= \lambda p_{DC\bar{m}}(1+\bar{m}) + (1-\lambda)\left[p_{\bar{m}}(-\beta)\right] - \lambda p_{DC\bar{m}}(1+\alpha)$ $= \lambda(\bar{m}-\alpha)p_{DC\bar{m}} - \beta(1-\lambda)p_{\bar{m}}$ $\left(\Pi_{\bar{m}}(CD,\underline{m}) - \Pi_{\underline{m}}(\underline{m})\right)^f = -\alpha\lambda - \beta(1-\lambda)p_{\bar{m}} < 0$
$DC\bar{m}$ $CD\underline{m}$ $\underline{m}$	$p_{CD\underline{m}} = \frac{1}{2}\left[1 + \frac{\bar{k}+(1-\lambda)\beta}{\lambda(1+\bar{m})}\right]$ $p_{DC\bar{m}} = \frac{1}{2}\left[1 - \frac{\bar{k}+(1-\lambda)\beta}{\lambda(1+\bar{m})}\right]$ $p_{\bar{m}} = 0$	$\lambda > \max\left\{\frac{\bar{k}+\beta}{(1+\bar{m}+\beta)}, 1 - \frac{1}{\beta(1+\alpha)}\left((1+\bar{m})\underline{k} - (1+\alpha)\bar{k}\right), 1 - \frac{\bar{m}-\alpha}{(\bar{m}-\alpha+\beta)(1+\bar{m}) + (1+\alpha)\beta}(1+\bar{m}+\bar{k})\right\}$		$\Pi_{\bar{m}}(CD,\underline{m}) - \Pi_{\underline{m}}(\underline{m}) = \lambda p_{DC\bar{m}}(1+\bar{m}) - \lambda p_{DC\bar{m}}(1+\alpha)$ $= \lambda(\bar{m}-\alpha)p_{DC\bar{m}} > 0$ $\left(\Pi_{\bar{m}}(CD,\underline{m}) - \Pi_{\underline{m}}(\underline{m})\right)^f = -\alpha\lambda < 0$
$CD\bar{m}$ $\bar{m}$ $\underline{m}$	$p_{CD\bar{m}} = 1$	$\lambda = \frac{\underline{k}}{(1+\alpha)}$	1. $\underline{k} < (1+\alpha)$ 2. $p_{\bar{m}} < \frac{\lambda(\bar{m}-\alpha)}{(1-\lambda)\beta}$	$\Pi_{\bar{m}}(CD,\bar{m}) - \Pi_{\underline{m}}(\underline{m}) = \frac{\underline{k}}{(1+\alpha)}(1+\bar{m}) - \bar{k} - \beta(1-\lambda)p_{\bar{m}}$ $\left(\Pi_{\bar{m}}(CD,\bar{m}) - \Pi_{\underline{m}}(\underline{m})\right)^f = \frac{\underline{k}}{(1+\alpha)} - \bar{k} - \beta(1-\lambda)p_{\bar{m}}$

B.3 Stability of p-equilibria

In this appendix we are concerned with stability of the p-equilibria derived in Appendix A. Before we start we will make some comments on the dynamics and on our way of argument. The dynamics among high types is described by eight differential equation, seven being independent, one for each of the eight signal-dependent strategies $CC\bar{m}$, $CD\bar{m}$, $DC\bar{m}$, $DD\bar{m}$ and $CC\underline{m}$, $CD\underline{m}$, $DC\underline{m}$, $DD\underline{m}$. Due to the dominance of defection among low types there are only two differential equations, one being independent, reflecting the signaling behavior. Since it is assumed that the evolution of the share of high types is comparably slow to the evolution of the shares adopting the different strategies these seven independent differential equations for high types and the one for low types give rise to a coupled system of differential equation which itself consists of the two aforementioned systems (high and low types). We are confronted with two types of equilibria, on the one hand equilibrium points, i.e. an equilibrium that specifies a precise level for each share $p_{XYm}, m, \in \{\underline{m}, \bar{m}\}, X, Y \in \{C, D\}$ among high types and $p_m, m \in \{\underline{m}, \bar{m}\}$ among low types. There are on the other hand equilibrium set, i.e. non-singleton subsets of $\mathbb{R}^{10}$. We apply the notion of asymptotical stability as a stability concept. An equilibrium point is a fix point $\vec{p}_f$ of the dynamical system $\dot{p}(t) = F(\vec{p}(t))$ and is said to be asymptotically stable if it meets two conditions. First it needs to be Lyapunow-stable, i.e. $\forall \varepsilon > 0, \exists \delta > 0 : \|\vec{p}(0) - \vec{p}_f\| < \delta \Rightarrow \|\vec{p}(t) - \vec{p}_f\| < \varepsilon, \forall t \geq 0, \forall \vec{p}(t)$ being a trajectory, second it needs to be an attractor, i.e. $\exists \delta > 0$: any trajectory $\vec{p}(t)$ with $\|\vec{p}(0) - \vec{p}_f\| < \delta$ then $\|\vec{p}(t) - \vec{p}_f\| \xrightarrow[t \to \infty]{} 0$. The definitions for an equilibrium set are accordingly (see e.g. Samuelson 1997). To proof stability or instability of an equilibrium we will rely on phase diagrams. We will proof instability by arguing that the system cannot be Lyapunow-stable. In case of an equilibrium point in the interior of the support of the equilibrium the involved strategies earn strictly higher payoffs then non-equilibrium strategies. Small perturbation will not alter this property. Payoff monotone dynamics will decrease the share of the non-equilibrium strategies. Hence for analyzing the stability properties in that case it suffices to consider the involved equilibrium strategies and whether the dynamics will reestablish the equilibrium values given a small perturbation. At the boundaries of the support of an equilibrium point a non-equilibrium strategy will earn the same profits as the equilibrium strategies. In that case these strategies needs to be included in the analysis. However with respect to all other strategies the previous argument still applies.

In the first cell of the first row in the following tables equilibrium strategies can be found. In second cell the precise values or set conditions are given. The third cell contains the support of the equilibrium. The last cell may state some additional conditions concerning the existence of the considered equilibrium.

1. Separating Equilibria

$CD\bar{m}$ $\underline{m}$	$p_{CD\bar{m}}=1,\ p_{\underline{m}}=1$	$\frac{\bar{k}}{1+\bar{m}}\le\lambda\le\frac{\underline{k}}{1+\alpha}$	$\bar{k}<1+\bar{m}$

This equilibrium is certainly stable in the interior range $\frac{\bar{k}}{1+\bar{m}}<\lambda<\frac{\underline{k}}{1+\alpha}$ since all inequalities hold strict. At the upper boundary $\lambda=\frac{\underline{k}}{1+\alpha}$ low types are indifferent between sending the signal and not sending. Any deviation from $p_{\underline{m}}=1$ would need an decrease in $p_{CD\bar{m}}$ to reestablish $p_{\underline{m}}=1$. However, for small such changes $CD\bar{m}$ is still the dominant strategy for high types, hence $p_{CD\bar{m}}=1$ persists and there is no force bringing back $p_{\underline{m}}=1$. Hence at $\lambda=\frac{\underline{k}}{1+\alpha}$ this equilibrium is not stable. At the lower boundary $\lambda=\frac{\bar{k}}{1+\bar{m}}$ high types are indifferent between $CD\bar{m}$ and $DD\underline{m}$:
$\Pi_{\bar{m}}(CD,\bar{m})=\lambda\left[(p_{CD\bar{m}})(1+\bar{m})+(p_{DD\bar{m}})(-\beta)\right]-\bar{k}$ Consider any small increase by random shift in $p_{DD\bar{m}}$, this will lower profits for $CD\bar{m}$ and leave profits $\Pi_{\bar{m}}(DD,\underline{m})=0$
for $DD\underline{m}$ unchanged, hence equilibrium will not be restored. In other words at $\lambda=\frac{\bar{k}}{1+\bar{m}}$ this equilibrium is not stable.

$CD\bar{m}$ $DD\bar{m}$ $\underline{m}$	$p_{CD\bar{m}}=\frac{\beta}{\bar{m}+\beta-\alpha},\ p_{DD\bar{m}}=\frac{\bar{m}-\alpha}{\bar{m}+\beta-\alpha},\ p_{\underline{m}}=1$	$\frac{(\bar{m}-\alpha+\beta)}{\beta}\frac{\bar{k}}{(1+\alpha)}\le\lambda\le\frac{(\bar{m}-\alpha+\beta)}{\beta}\frac{\underline{k}}{(1+\alpha)}$	$\frac{\bar{k}}{(1+\alpha)}<\frac{\beta}{(\bar{m}-\alpha+\beta)}$

$\Pi_{\bar{m}}(CD,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})\ge 0\Leftrightarrow$

$\lambda\left[p_{CD\bar{m}}(\bar{m}-\alpha)+p_{DD\bar{m}}(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]\ge 0$

$\Leftrightarrow\lambda p_{CD\bar{m}}(\bar{m}-\alpha+\beta)-\lambda\beta+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]\ge 0$

$\Leftrightarrow p_{CD\bar{m}}\ge\frac{\beta-(1-\lambda)\beta p_{\underline{m}}}{\lambda(\bar{m}-\alpha+\beta)}$

$\Pi_{\underline{m}}(\underline{m})-\Pi_{\underline{m}}(\bar{m})\ge 0\Leftrightarrow$

$-\lambda p_{CD\bar{m}}(1+\alpha)+\underline{k}\ge 0\Leftrightarrow p_{CD\bar{m}}\le\frac{\underline{k}}{\lambda(1+\alpha)}$,

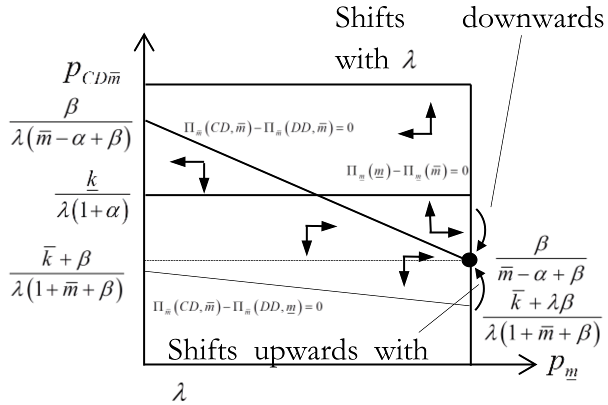
note that for the support of that equilibrium

$\frac{\underline{k}}{\lambda(1+\alpha)}>\frac{\bar{k}+\lambda\beta-(1-\lambda)\beta p_{\underline{m}}}{\lambda(1+\bar{m}+\beta)}$ holds.

Hence we obtain the following phase diagram:

Note: $\lambda\le\frac{(\bar{m}-\alpha+\beta)}{\beta}\frac{\underline{k}}{(1+\alpha)}\Rightarrow\frac{\beta}{(\bar{m}-\alpha+\beta)}\le\frac{\underline{k}}{\lambda(1+\alpha)}$

$\frac{\bar{k}+\lambda\beta}{\lambda(1+\bar{m}+\beta)}\le\frac{\beta}{\bar{m}-\alpha+\beta}\Leftrightarrow\frac{\bar{k}}{\lambda(1+\alpha)}\le\frac{\beta}{\bar{m}-\alpha+\beta}$



As the diagram clearly indicates, this equilibrium is unstable for all λ in the support.

2. Pooling Equilibria

$CC\underline{m}, DC\underline{m}$ $\underline{m}$	$p_{CC\underline{m}} + p_{DC\underline{m}} = 1$	$\frac{\beta}{(\beta+\bar{m}-\alpha)} \le \lambda$	

This set of equilibria is stable for $\lambda > \frac{\beta}{(\beta+\bar{m}-\alpha)}$ since all inequalities hold strictly, i.e. for any small perturbation the equilibrium strategies earn strictly more than any other strategy. Note that not necessarily the pre-perturbation shares are reestablished, but that the sum of their shares equals unity.

At the boundary $\lambda = \frac{\beta}{(\beta+\bar{m}-\alpha)}$ agents become indifferent between $CC_{\underline{m}} / DC_{\underline{m}}$ and $CD_{\underline{m}} / DD_{\underline{m}}$.

Low types still strictly prefer not to signal.

$$\Pi_{\bar{m}}(CC,\underline{m}) - \Pi_{\bar{m}}(CD,\underline{m}) \ge 0 \Leftrightarrow \lambda\left[\left(p_{CC\underline{m}} + p_{DC\underline{m}}\right)(\bar{m}-\alpha) + \left(p_{CD\underline{m}} + p_{DD\underline{m}}\right)(-\beta)\right] - \beta(1-\lambda) \ge 0 \Leftrightarrow p_{CC\underline{m}} \ge \frac{\beta}{\lambda(\bar{m}-\alpha+\beta)} - p_{DC\underline{m}}$$

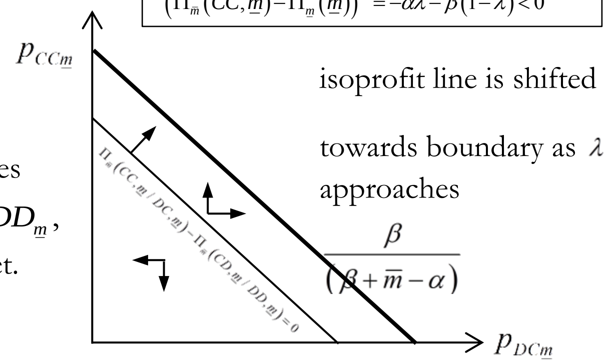
Note that at $\lambda = \frac{\beta}{(\beta+\bar{m}-\alpha)}$ a perturbation from $CC_{\underline{m}}$ towards $DD_{\underline{m}}$ decreases the payoffs for the equilibrium strategies strictly more than for the $DD_{\underline{m}}$-strategy and decreases profits for all other strategies weakly more, i.e. those strategies still earn strictly less than $DD_{\underline{m}}$, and the share of $DD_{\underline{m}}$ increases. Hence there is no force reestablishing the equilibrium set.

As the diagram clearly indicates this equilibrium set is stable for $\frac{\beta}{(\beta+\bar{m}-\alpha)} < \lambda$.

Difference in fitness payoffs:

$\Pi_{\bar{m}}(CC,\underline{m}) - \Pi_{\underline{m}}(\underline{m}) = \lambda(\bar{m}-\alpha) - \beta(1-\lambda)$

$\left(\Pi_{\bar{m}}(CC,\underline{m}) - \Pi_{\underline{m}}(\underline{m})\right)^f = -\alpha\lambda - \beta(1-\lambda) < 0$



$CD\underline{m}, DD\underline{m}$ $\underline{m}$	$p_{CD\underline{m}} + p_{DD\underline{m}} = 1,\ p_{CD\underline{m}} \le \frac{1}{\lambda}\min\left\{\frac{\bar{k}+\beta}{1+\bar{m}+\beta}, \frac{\bar{k}}{1+\alpha}\right\}$	$0 < \lambda < 1$	

$\Pi_{\bar{m}}(CD,\underline{m}) - \Pi_{\bar{m}}(DD,\underline{m}) = -\beta(1-\lambda)\, p_{\bar{m}} \le 0$

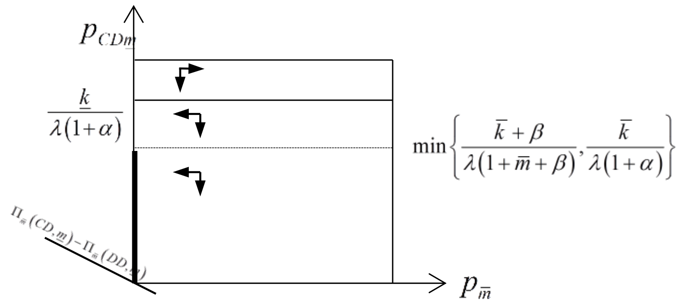
$\Pi_{\underline{m}}(\underline{m}) - \Pi_{\underline{m}}(\bar{m}) = \underline{k} - \lambda(1+\alpha)\, p_{CD\underline{m}} \ge 0 \Leftrightarrow p_{CD\underline{m}} \le \frac{\underline{k}}{\lambda(1+\alpha)}$

Hence we obtain the following phase diagram:

Difference in fitness payoffs:

$\Pi_{\bar{m}}(CD,\underline{m}) - \Pi_{\underline{m}}(\underline{m}) = 0$

$\left(\Pi_{\bar{m}}(CD,\underline{m}) - \Pi_{\underline{m}}(\underline{m})\right)^f = 0$



As the diagram clearly indicates, this equilibrium set is stable for all λ in the support.

$CC\underline{m}, DC\underline{m}, CD\underline{m}, DD\underline{m}$ $\underline{m}$	$p_{CD\underline{m}}+p_{DD\underline{m}}+p_{CC\underline{m}}+p_{DC\underline{m}}=1$	$\frac{\beta}{(\bar{m}-\alpha+\beta)}<\lambda<1$ $\left(p_{CC\underline{m}}+p_{DC\underline{m}}\right)=\frac{\beta}{\lambda(\bar{m}-\alpha+\beta)}$ $\left(p_{CD\underline{m}}-p_{DC\underline{m}}\right)\leq\frac{\bar{k}}{\lambda(1+\bar{m}+\beta)}$	

$\Pi_{\bar{m}}(CC,\underline{m})=\lambda\left(p_{CC\underline{m}}+p_{DC\underline{m}}\right)(1+\bar{m}+\beta)-\beta$

$\Pi_{\bar{m}}(CD,\underline{m})=\lambda\left(p_{CC\underline{m}}+p_{DC\underline{m}}\right)(1+\alpha)$

$\Pi_{\bar{m}}(DC,\underline{m})=\lambda\left(p_{CC\underline{m}}+p_{DC\underline{m}}\right)(1+\bar{m}+\beta)-\beta$

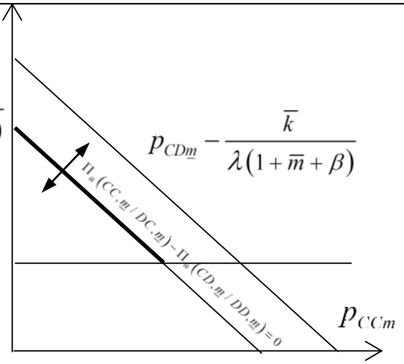
$\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left(p_{CC\underline{m}}+p_{DC\underline{m}}\right)(1+\alpha)$

$\Pi_{\bar{m}}(CC,\underline{m})-\Pi_{\bar{m}}(CD,\underline{m})=\Pi_{\bar{m}}(DC,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})$

$=\lambda\left(p_{CC\underline{m}}+p_{DC\underline{m}}\right)(\bar{m}-\alpha+\beta)-\beta\geq 0 \Leftrightarrow p_{DC\underline{m}}\geq\frac{\beta}{\lambda(\bar{m}-\alpha+\beta)}-p_{CC\underline{m}}$

all other differences vanish.

Differences depend only on two shares. As the diagram clearly indicates this equilibrium set is unstable.



$CC\bar{m}, CD\bar{m}$ $\bar{m}$	$p_{CC\bar{m}}+p_{CD\bar{m}}=1$	$\lambda\geq\max\left\{\frac{\underline{k}}{1+\alpha},\frac{\beta}{(\beta+\bar{m}-\alpha)}\right\}$	$\underline{k}<1+\alpha$ $\quad p_{CD\bar{m}}\geq\frac{\underline{k}}{\lambda(1+\alpha)}$

$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CD,\bar{m})=-\beta(1-\lambda)p_{\underline{m}}\leq 0$

$\Pi_{\underline{m}}(\underline{m})-\Pi_{\underline{m}}(\bar{m})=\underline{k}-\lambda(1+\alpha)p_{CD\bar{m}}\geq 0\Leftrightarrow p_{CD\bar{m}}\leq\frac{\underline{k}}{\lambda(1+\alpha)}$

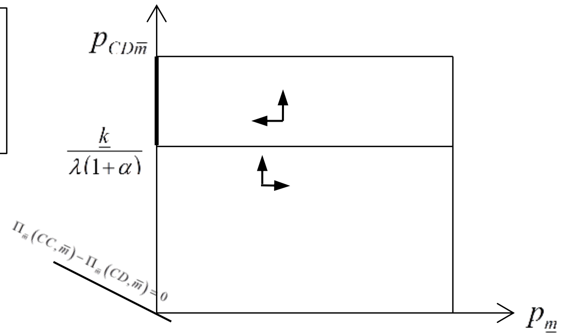
Difference in fitness payoffs:

$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\underline{m}}(\bar{m})=\lambda(\bar{m}-\alpha+\beta)-\beta+\underline{k}-\bar{k}$

$\left(\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\underline{m}}(\bar{m})\right)^{f}=\lambda(\beta-\alpha)-\beta+\underline{k}-\bar{k}$

Hence we obtain the following phase diagram:

Note that at $p_{CD\bar{m}}=\frac{\underline{k}}{\lambda(1+\alpha)}$ low types are indifferent between signaling and no signaling. As soon as low types start not to signal $CD\bar{m}$ earns strictly higher payoffs than $CC\bar{m}$ such that the incentive for low types to signal will be restored. As the diagram clearly indicates, this equilibrium set is stable.



$CC\bar{m}, CD\bar{m}, DC\bar{m}, DD\bar{m}$ $\bar{m}$	$p_{CC\bar{m}}+p_{CD\bar{m}}+p_{DC\bar{m}}+p_{DD\bar{m}}=1$	$\lambda\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)=\frac{\beta}{(\bar{m}-\alpha+\beta)}$, $\lambda\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)\geq\frac{\underline{k}}{1+\alpha}$	$\underline{k}<1+\alpha$

$\Pi_{\bar{m}}(CC,\bar{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)(1+\bar{m})+\left(p_{DC\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)+p_{\bar{m}}(-\beta)\right]-\bar{k}$

$\Pi_{\bar{m}}(CD,\bar{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)(1+\bar{m})+\left(p_{DC\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]-\bar{k}$

$\Pi_{\bar{m}}(DC,\bar{m})=\lambda\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)(1+\alpha)+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]-\bar{k}$

$\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)(1+\alpha)-\bar{k}$

Hence we obtain the following condensed phase diagram:

As the diagram clearly indicates this equilibrium is unstable.

let $\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)=x, \left(p_{DC\bar{m}}+p_{DD\bar{m}}\right)=y=1-x$

$\Pi_{\bar{m}}(CC,\bar{m})=\lambda\left[x(1+\bar{m}+\beta)\right]-\beta-\bar{k}$

$\Pi_{\bar{m}}(CD,\bar{m})=\lambda\left[x(1+\bar{m}+\beta)\right]-\beta-\bar{k}$

$\Pi_{\bar{m}}(DC,\bar{m})=\lambda x(1+\alpha)-\bar{k}$

$\Pi_{\bar{m}}(DD,\bar{m})=\lambda x(1+\alpha)-\bar{k}$

3. **Semi-Pooling Equilibria** (The denotation of the following equilibria in the first column refers to the corresponding subsection in App. A.)

<table>
<tr>
<td rowspan="2">2.2.1.1.</td>
<td>$CC\bar{m}$
$CD\bar{m}$
$DC\bar{m}$
$DD\bar{m}$
$CD\underline{m}$
$DD\underline{m}$
$\underline{m}$</td>
<td>$p_{CC\bar{m}} = -\frac{\bar{k}}{\lambda(1+\alpha)} + \frac{\beta}{\lambda(\bar{m}-\alpha+\beta)} - p_{DC\bar{m}}$
$p_{CD\bar{m}} = p_{DC\bar{m}},$
$p_{DD\bar{m}} = -\frac{\bar{k}}{\lambda(1+\alpha)}\frac{\bar{m}-\alpha}{\beta} - \frac{\beta}{\lambda(\bar{m}-\alpha+\beta)} + \frac{1}{\lambda} - p_{DC\bar{m}}$
$p_{CD\underline{m}} = \frac{\bar{k}}{\lambda(1+\alpha)}$
$p_{DD\underline{m}} = \frac{\bar{k}}{\lambda(1+\alpha)}\frac{\bar{m}-\alpha}{\beta} - \frac{1-\lambda}{\lambda}$</td>
<td>$\lambda \geq 1 - \frac{\bar{k}}{(1+\alpha)}\frac{\bar{m}-\alpha}{\beta}$</td>
<td>1. $\frac{\bar{k}}{1+\alpha} < \frac{\beta}{\bar{m}-\alpha+\beta}$
2.
$\beta < (\bar{m}-\alpha)$: $p_{DC\bar{m}} \leq \frac{1}{\lambda}\left(\frac{\beta}{\bar{m}-\alpha+\beta} - \frac{\bar{k}}{1+\alpha}\right)$
$\beta \geq (\bar{m}-\alpha)$: $p_{DC\bar{m}} \leq \frac{1}{\lambda}\frac{\bar{m}-\alpha}{\beta}\left(\frac{\beta}{\bar{m}-\alpha+\beta} - \frac{\bar{k}}{1+\alpha}\right)$</td>
</tr>
<tr>
<td colspan="3">Note that the payoffs for non-signaling high types is independent of their own share. However payoffs for all other behavioral strategies strictly increase in the share $p_{CD\underline{m}}$ and weakly decrease in $p_{DD\underline{m}}$. Hence if the set is perturbed such that the equilibrium level for $p_{CD\underline{m}}$ is exceeded than there is no force bringing it back to that level. Hence this set of equilibria is unstable.</td>
<td>$\Pi_{\bar{m}}(CC,\bar{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}} + p_{CD\underline{m}}\right)(1+\bar{m}) + \left(p_{DC\bar{m}} + p_{DD\bar{m}} + p_{DD\underline{m}}\right)(-\beta)\right] - \beta(1-\lambda) - \bar{k}$
$\Pi_{\bar{m}}(CD,\bar{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(1+\bar{m}) + \left(p_{CD\underline{m}}\right)(1+\alpha) + \left(p_{DC\bar{m}} + p_{DD\bar{m}}\right)(-\beta)\right] - \bar{k}$
$\Pi_{\bar{m}}(DC,\bar{m}) = \lambda\left[\left(p_{CD\underline{m}}\right)(1+\bar{m}) + \left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(1+\alpha) + \left(p_{DD\underline{m}}\right)(-\beta)\right] - \beta(1-\lambda) - \bar{k}$
$\Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}} + p_{CD\underline{m}}\right)(1+\alpha)\right] - \bar{k}$

$\Pi_{\bar{m}}(CD,\underline{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(1+\bar{m}) + \left(p_{CD\bar{m}} + p_{DD\bar{m}}\right)(-\beta)\right]$
$\Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(1+\alpha)\right]$</td>
</tr>
</table>

<table>
<tr>
<td>2.2.1.2.1.</td>
<td>$CC\bar{m}$
$CD\bar{m}$
$DC\bar{m}$
$DD\bar{m}$
$CD\underline{m}$
$\underline{m}$
$\bar{m}$</td>
<td>$p_{CC\bar{m}} = \frac{(1+\alpha)(\beta+\bar{k})-(1+\beta+\bar{m})\underline{k}}{\lambda(1+\alpha)(\bar{m}-\alpha+\beta)} - p_{CD\bar{m}}$
$p_{DC\bar{m}} = p_{CD\bar{m}} + \frac{\underline{k}-\bar{k}}{\lambda(\bar{m}-\alpha+\beta)}$
$p_{DD\bar{m}} = 1 - \frac{\beta+\underline{k}-\bar{k}}{\lambda(\bar{m}-\alpha+\beta)} - p_{CD\bar{m}}$
$p_{CD\underline{m}} = \frac{\underline{k}}{\lambda(1+\alpha)} + \frac{\underline{k}-\bar{k}}{\lambda(\bar{m}-\alpha+\beta)} = \frac{(1+\beta+\bar{m})\underline{k}-(1+\alpha)\bar{k}}{\lambda(1+\alpha)(\bar{m}-\alpha+\beta)}$
$p_{\underline{m}} = \frac{1}{1-\lambda}\frac{\bar{m}-\alpha}{\beta}\frac{(1+\beta+\bar{m})\underline{k}-(1+\alpha)\bar{k}}{(1+\alpha)(\bar{m}-\alpha+\beta)}$</td>
<td>$\frac{\beta+\underline{k}-\bar{k}}{(\bar{m}-\alpha+\beta)} < \lambda$
$\leq 1 - \frac{\bar{m}-\alpha}{\beta}\frac{(1+\beta+\bar{m})\underline{k}-(1+\alpha)\bar{k}}{(1+\alpha)(\bar{m}-\alpha+\beta)}$</td>
<td>1.
$(1+\alpha)\beta > \frac{\bar{m}-\alpha+\beta}{(\bar{m}-\alpha)}\left((1+\bar{m})\underline{k}-(1+\alpha)\bar{k}\right)$
2.
$0 \leq p_{CD\bar{m}} \leq \min\left\{\underbrace{1-\frac{\beta+\underline{k}-\bar{k}}{\lambda(\bar{m}-\alpha+\beta)}}_{x}, \underbrace{\frac{\beta+\bar{k}-\frac{1+\beta+\bar{m}}{1+\alpha}\underline{k}}{\lambda(\bar{m}-\alpha+\beta)}}_{y}\right\}$
$x \geq y \Leftrightarrow \lambda \geq \frac{2\beta}{(\bar{m}-\alpha+\beta)} - \frac{\underline{k}}{1+\alpha}\left(\overset{(3)b)}{>}\frac{\beta+\underline{k}-\bar{k}}{(\bar{m}-\alpha+\beta)}\right)$</td>
</tr>
</table>

Note that the payoff for non-signaling high types is independent of their own share. However payoffs for all other behavioral strategies strictly increase in the share $p_{CD\underline{m}}$ and weakly decrease in $p_{DD\bar{m}}$. Consider a perturbation such that the equilibrium level for $p_{CD\underline{m}}$ is exceeded and $p_{DD\bar{m}}$ decreases. Payoffs for signaling high types strictly increase and there is no force bringing $p_{CD\underline{m}}$ back to that level. Hence this set of equilibria is unstable.

$$\Pi_{\bar{m}}(CC,\bar{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}} + p_{CD\underline{m}}\right)(1+\bar{m}) + \left(p_{DC\bar{m}} + p_{DD\bar{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\underline{m}}(-\beta) + p_{\bar{m}}(-\beta)\right] - \bar{k}$$
$$\Pi_{\bar{m}}(CD,\bar{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(1+\bar{m}) + p_{CD\underline{m}}(1+\alpha) + \left(p_{DC\bar{m}} + p_{DD\bar{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\bar{m}}(-\beta)\right] - \bar{k}$$
$$\Pi_{\bar{m}}(DC,\bar{m}) = \lambda\left[p_{CD\underline{m}}(1+\bar{m}) + \left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(1+\alpha)\right] + (1-\lambda)\left[p_{\underline{m}}(-\beta)\right] - \bar{k}$$
$$\Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}} + p_{CD\underline{m}}\right)(1+\alpha)\right] - \bar{k}$$
$$\Pi_{\bar{m}}(CD,\underline{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(1+\bar{m}) + \left(p_{CD\bar{m}} + p_{DD\bar{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\bar{m}}(-\beta)\right]$$
$$\Pi_{\underline{m}}(\underline{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(1+\alpha)\right]$$
$$\Pi_{\underline{m}}(\bar{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}} + p_{CD\underline{m}}\right)(1+\alpha)\right] - \underline{k}$$

2.2.1.2.2.

$CC\bar{m}$
$CD\bar{m}$
$DC\bar{m}$
$DD\bar{m}$
$CD\underline{m}$
$\underline{m}$

$$p_{CC\bar{m}} = \left(1 - p_{CD\underline{m}}\right)\frac{\beta}{(\bar{m}-\alpha+\beta)} - p_{CD\bar{m}}$$
$$p_{DC\bar{m}} = p_{CD\bar{m}} - \frac{\bar{k}}{\lambda(1+\bar{m}+\beta)} + \frac{(1+\alpha)}{(1+\bar{m}+\beta)}p_{CD\underline{m}}$$
$$p_{DD\bar{m}} = \left(1 - p_{CD\underline{m}}\right)\frac{\bar{m}-\alpha}{(\bar{m}-\alpha+\beta)} - p_{DC\bar{m}}$$
$$p_{CD\underline{m}} = \frac{(1-\lambda)\beta}{\lambda(\bar{m}-\alpha)}$$

$$\max\left\{\underbrace{\frac{\beta\left(-\alpha(2+\alpha)+\beta+\bar{m}(2+\bar{m}+\beta)\right)-\bar{k}(\bar{m}-\alpha)(\bar{m}-\alpha+\beta)}{(1+\bar{m})(\bar{m}-\alpha+\beta)^2}}_{z>0},\ \underbrace{1-\frac{(\bar{m}-\alpha)}{\beta(\bar{m}-\alpha+\beta)}\left((1+\bar{m}+\beta)\frac{\underline{k}}{(1+\alpha)}-\bar{k}\right)}_{(*)}\right\}$$
$$<\lambda \underset{(iv)}{<} 1-\frac{\bar{k}}{(1+\alpha)}\frac{(\bar{m}-\alpha)}{\beta}$$

1. $\frac{\bar{k}}{(1+\alpha)} < \frac{\beta}{(\bar{m}-\alpha+\beta)}$
2. $0 \le p_{CD\bar{m}} \le$
$$\min\left\{\underbrace{\left(1-\frac{(1-\lambda)\beta}{\lambda(\bar{m}-\alpha)}\right)\frac{\beta}{(\bar{m}-\alpha+\beta)}}_{y},\ \underbrace{\left(1-\frac{(1-\lambda)\beta}{\lambda(\bar{m}-\alpha)}\right)\frac{\bar{m}-\alpha}{(\bar{m}-\alpha+\beta)}+\frac{\bar{k}}{\lambda(1+\bar{m}+\beta)}-\frac{1+\alpha}{1+\bar{m}+\beta}\frac{(1-\lambda)\beta}{\lambda(\bar{m}-\alpha)}}_{z}\right\}$$

Note that the payoffs for non-signaling high types is independent of their own share. However payoffs for all other behavioral strategies strictly increase in the share $p_{CD\underline{m}}$. Hence if the set is perturbed such that the equilibrium level for $p_{CD\underline{m}}$ is exceeded than there is no force bringing it back to that level. Hence this set of equilibria is unstable.

$$\Pi_{\bar{m}}(CC,\bar{m}) = \lambda\left[(1+\bar{m}) - (1+\bar{m}+\beta)\left(p_{DC\bar{m}} + p_{DD\bar{m}}\right)\right] - \beta(1-\lambda) - \bar{k}$$
$$\Pi_{\bar{m}}(CD,\bar{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(\bar{m}-\alpha) + (1+\alpha) - (1+\alpha+\beta)\left(p_{DC\bar{m}} + p_{DD\bar{m}}\right)\right] - \bar{k}$$
$$\Pi_{\bar{m}}(DC,\bar{m}) = \lambda\left[(1+\bar{m}) - \left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(\bar{m}-\alpha) - (1+\bar{m})\left(p_{DC\bar{m}} + p_{DD\bar{m}}\right)\right] - \beta(1-\lambda) - \bar{k}$$
$$\Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[(1+\alpha) - (1+\alpha)\left(p_{DC\bar{m}} + p_{DD\bar{m}}\right)\right] - \bar{k}$$
$$\Pi_{\bar{m}}(CD,\underline{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(1+\bar{m}) + \left(p_{CD\bar{m}} + p_{DD\bar{m}}\right)(-\beta)\right]$$

$$\Pi_{\bar{m}}(CC,\bar{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}} + p_{CD\underline{m}}\right)(1+\bar{m}) + \left(p_{DC\bar{m}} + p_{DD\bar{m}}\right)(-\beta)\right] - \beta(1-\lambda) - \bar{k}$$
$$\Pi_{\bar{m}}(CD,\bar{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(1+\bar{m}) + p_{CD\underline{m}}(1+\alpha) + \left(p_{DC\bar{m}} + p_{DD\bar{m}}\right)(-\beta)\right] - \bar{k}$$
$$\Pi_{\bar{m}}(DC,\bar{m}) = \lambda\left[p_{CD\underline{m}}(1+\bar{m}) + \left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(1+\alpha)\right] - \beta(1-\lambda) - \bar{k}$$
$$\Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}} + p_{CD\underline{m}}\right)(1+\alpha)\right] - \bar{k}$$
$$\Pi_{\bar{m}}(CD,\underline{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(1+\bar{m}) + \left(p_{CD\bar{m}} + p_{DD\bar{m}}\right)(-\beta)\right]$$

4.1.3.2.			
	$CC\bar{m}$ $CD\underline{m}$ $\underline{m}$	$p_{CD\underline{m}} = \frac{\bar{k}+(1-\lambda)\beta}{\lambda(1+\bar{m})}$ $p_{CC\bar{m}} = 1-\frac{\bar{k}+(1-\lambda)\beta}{\lambda(1+\bar{m})}$ $p_{\underline{m}} = 1$	1. $\frac{\bar{k}+\beta}{(1+\bar{m}+\beta)} < \lambda < 1$: $\frac{\beta}{(\bar{m}-\alpha+\beta)} < \frac{\bar{k}}{(1+\alpha)}\left(< \frac{1+\bar{m}}{(1+\alpha)}\right)$ 2. $1-\frac{(\bar{m}-\alpha)}{\beta}\frac{\bar{k}}{(1+\alpha)} < \lambda < 1$: $\frac{\bar{k}}{(1+\alpha)} \le \frac{\beta}{(\bar{m}-\alpha+\beta)}$

$\Pi_{\bar{m}}(CC,\bar{m}) = \lambda\left(p_{CC\bar{m}}+p_{CD\underline{m}}\right)(1+\bar{m})-\beta(1-\lambda)-\bar{k}$

$\Pi_{\bar{m}}(CD,\underline{m}) = \lambda\left(p_{CC\bar{m}}\right)(1+\bar{m})+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]$

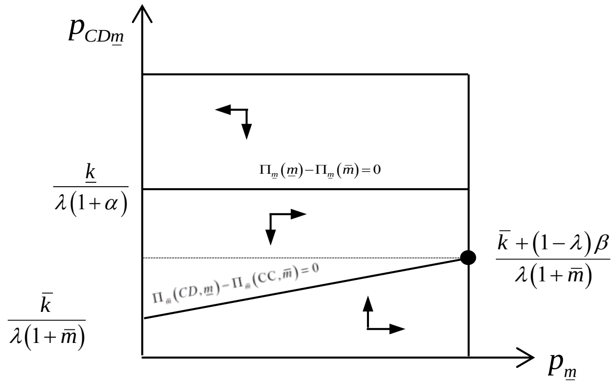
$\Pi_{\underline{m}}(\underline{m}) = \lambda p_{CC\bar{m}}(1+\alpha)$

$\Pi_{\underline{m}}(\bar{m}) = \lambda\left(p_{CC\bar{m}}+p_{CD\underline{m}}\right)(1+\alpha)-\underline{k}$

$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CD,\underline{m}) \ge 0 \Leftrightarrow \lambda p_{CD\underline{m}}(1+\bar{m})-\beta(1-\lambda)-\bar{k}-(1-\lambda)\left[p_{\bar{m}}(-\beta)\right] \ge 0 \Leftrightarrow p_{CD\underline{m}} \ge \frac{\beta(1-\lambda)+\bar{k}+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]}{\lambda(1+\bar{m})} = \frac{\bar{k}+\beta(1-\lambda)p_{\underline{m}}}{\lambda(1+\bar{m})}$

$\Pi_{\underline{m}}(\underline{m})-\Pi_{\underline{m}}(\bar{m}) \ge 0 \Leftrightarrow \lambda p_{CC\bar{m}}(1+\alpha)-\lambda\left(p_{CC\bar{m}}+p_{CD\underline{m}}\right)(1+\alpha)+\underline{k} \ge 0 \Leftrightarrow p_{CD\underline{m}} \le \frac{\underline{k}}{\lambda(1+\alpha)}$, note that $\frac{\underline{k}}{(1+\alpha)} > \frac{\bar{k}+\beta(1-\lambda)}{(1+\bar{m})}$ due to the condition (see support) $\lambda > 1-\frac{(\bar{m}-\alpha)\bar{k}}{\beta(1+\alpha)}$.

Hence we obtain the following phase diagram:



Note that the lower bound in the support condition implies $\left(\frac{\underline{k}}{\lambda(1+\alpha)} >\right)\frac{\bar{k}}{\lambda(1+\alpha)} > \frac{\bar{k}+(1-\lambda)\beta}{\lambda(1+\bar{m})}$

Differences in fitness payoffs:

$\Pi_{\bar{m}}(CD,\underline{m})-\Pi_{\underline{m}}(\underline{m}) = \lambda\left(p_{CC\bar{m}}\right)(1+\bar{m})-\lambda p_{CC\bar{m}}(1+\alpha) = \lambda(\bar{m}-\alpha)p_{CC\bar{m}} = \lambda(\bar{m}-\alpha)\left(1-\frac{\bar{k}+(1-\lambda)\beta}{\lambda(1+\bar{m})}\right) > 0$

$\left(\Pi_{\bar{m}}(CD,\underline{m})-\Pi_{\underline{m}}(\underline{m})\right)^{f} = -\alpha\lambda\left(1-\frac{\bar{k}+(1-\lambda)\beta}{\lambda(1+\bar{m})}\right) < 0$

As the diagram clearly indicates this equilibrium is stable.

4.2.3.2.

$DC\bar{m}$ $CD\underline{m}$ $\underline{m}$ $\bar{m}$	$p_{CD\underline{m}} = \frac{1}{2}\left[1 + \frac{\underline{k}}{\lambda(1+\alpha)}\right]$ $p_{DC\bar{m}} = \frac{1}{2}\left[1 - \frac{\underline{k}}{\lambda(1+\alpha)}\right]$ $p_{\underline{m}} = \frac{1}{2}\left[1 + \frac{1}{(1-\lambda)\beta}\left[\frac{(1+\bar{m})}{(1+\alpha)}\underline{k} - \bar{k}\right]\right]$ $p_{\bar{m}} = \frac{1}{2}\left[1 - \frac{1}{(1-\lambda)\beta}\left[\frac{(1+\bar{m})}{(1+\alpha)}\underline{k} - \bar{k}\right]\right]$	$0 < \frac{\beta + (\underline{k} - \bar{k})}{(\bar{m} - a + b)} < \lambda < 1 - \frac{1}{\beta}\left[\frac{(1+\bar{m})}{(1+a)}\underline{k} - \bar{k}\right] < 1$	$\beta(1+\alpha) > \frac{(\bar{m} - \alpha + \beta)}{(\bar{m} - \alpha)}\left[(1+\bar{m})\underline{k} - (1+\alpha)\bar{k}\right] + \frac{\underline{k} - \bar{k}}{\bar{m} - \alpha}\beta(1+\alpha)$

$\Pi_{\bar{m}}(DC,\bar{m}) = \lambda p_{CD\underline{m}}(1+\bar{m}) + (1-\lambda)\left[p_{\underline{m}}(-\beta)\right] - \bar{k}$

$\Pi_{\bar{m}}(CD,\underline{m}) = \lambda p_{DC\bar{m}}(1+\bar{m}) + (1-\lambda)\left[p_{\bar{m}}(-\beta)\right]$

$\Pi_{\underline{m}}(\underline{m}) = \lambda p_{DC\bar{m}}(1+\alpha)$

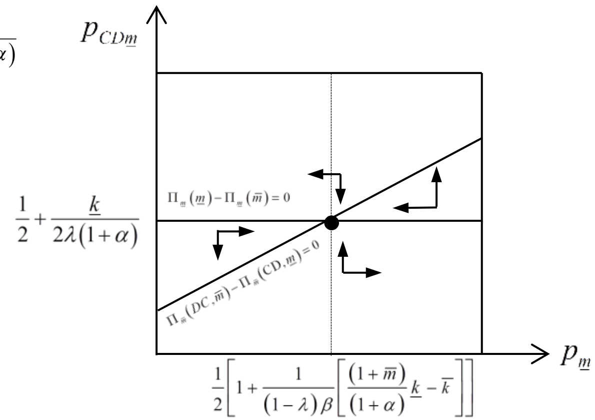
$\Pi_{\underline{m}}(\bar{m}) = \lambda p_{CD\underline{m}}(1+\alpha) - \underline{k}$

$\Pi_{\bar{m}}(DC,\bar{m}) - \Pi_{\bar{m}}(CD,\underline{m}) \geq 0 \Leftrightarrow \lambda p_{CD\underline{m}}(1+\bar{m}) + (1-\lambda)\left[p_{\underline{m}}(-\beta)\right] - \bar{k} - \lambda p_{DC\bar{m}}(1+\bar{m}) - (1-\lambda)\left[p_{\bar{m}}(-\beta)\right] \geq 0 \Leftrightarrow \lambda p_{CD\underline{m}}(1+\bar{m}) - \beta(1-\lambda)p_{\underline{m}} - \bar{k} - \lambda(1-p_{CD\underline{m}})(1+\bar{m}) + \beta(1-\lambda)(1-p_{\underline{m}}) \geq 0 \Leftrightarrow$

$2\lambda p_{CD\underline{m}}(1+\bar{m}) - 2\beta(1-\lambda)p_{\underline{m}} - \bar{k} - \lambda(1+\bar{m}) + \beta(1-\lambda) \geq 0 \Leftrightarrow p_{CD\underline{m}} \geq \frac{2\beta(1-\lambda)p_{\underline{m}} + \bar{k} + \lambda(1+\bar{m}) - \beta(1-\lambda)}{2\lambda(1+\bar{m})} = \frac{1}{2} + \frac{\bar{k} - \beta(1-\lambda) + 2\beta(1-\lambda)p_{\underline{m}}}{2\lambda(1+\bar{m})}$

$\Pi_{\underline{m}}(\underline{m}) - \Pi_{\underline{m}}(\bar{m}) \geq 0 \Leftrightarrow \lambda p_{DC\bar{m}}(1+\alpha) - \lambda p_{CD\underline{m}}(1+\alpha) + \underline{k} \geq 0 \Leftrightarrow \lambda(1+\alpha) - 2\lambda p_{CD\underline{m}}(1+\alpha) + \underline{k} \geq 0 \Leftrightarrow p_{CD\underline{m}} \leq \frac{\lambda(1+\alpha) + \underline{k}}{2\lambda(1+\alpha)} = \frac{1}{2} + \frac{\underline{k}}{2\lambda(1+\alpha)}$

Hence we obtain the following phase diagram:



Differences in fitness payoffs:

$\Pi_{\bar{m}}(CD,\underline{m}) - \Pi_{\underline{m}}(\underline{m}) = \lambda p_{DC\bar{m}}(1+\bar{m}) + (1-\lambda)\left[p_{\bar{m}}(-\beta)\right] - \lambda p_{DC\bar{m}}(1+\alpha) = \lambda(\bar{m} - \alpha)p_{DC\bar{m}} - \beta(1-\lambda)p_{\bar{m}}$

$\left(\Pi_{\bar{m}}(CD,\underline{m}) - \Pi_{\underline{m}}(\underline{m})\right)^{f} = -\alpha\lambda - \beta(1-\lambda)p_{\bar{m}} < 0$

As the diagram clearly indicates this equilibrium is stable.

4.2.3.3.

$DC\bar{m}$ $CD\underline{m}$ $\underline{m}$	$p_{CD\underline{m}} = \frac{1}{2}\left[1+\frac{\bar{k}+(1-\lambda)\beta}{\lambda(1+\bar{m})}\right]$ $p_{DC\bar{m}} = \frac{1}{2}\left[1-\frac{\bar{k}+(1-\lambda)\beta}{\lambda(1+\bar{m})}\right]$ $p_{\bar{m}} = 0$	$\lambda > \max\left\{\frac{\bar{k}+\beta}{(1+\bar{m}+\beta)}, 1-\frac{1}{\beta(1+\alpha)}\left((1+\bar{m})\underline{k}-(1+\alpha)\bar{k}\right), 1-\frac{\bar{m}-\alpha}{(\bar{m}-\alpha+\beta)(1+\bar{m})+(1+\alpha)\beta}(1+\bar{m}+\bar{k})\right\}$

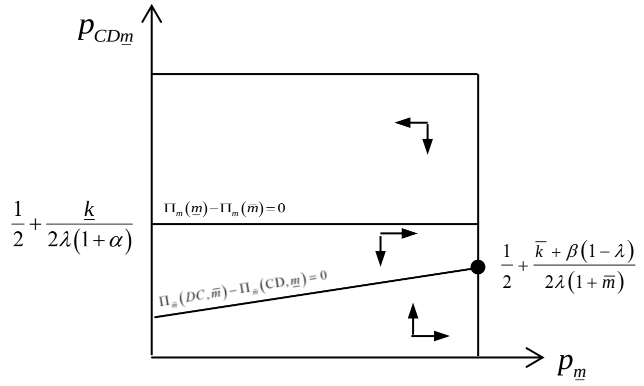
as above:

$$\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(CD,\underline{m})\geq 0 \Leftrightarrow p_{CD\underline{m}} \geq \frac{1}{2}+\frac{\bar{k}-\beta(1-\lambda)+2\beta(1-\lambda)p_{\underline{m}}}{2\lambda(1+\bar{m})}$$

$$\Pi_{\underline{m}}(\underline{m})-\Pi_{\underline{m}}(\bar{m})\geq 0 \Leftrightarrow p_{CD\underline{m}} \leq \frac{\lambda(1+\alpha)+\underline{k}}{2\lambda(1+\alpha)} = \frac{1}{2}+\frac{\underline{k}}{2\lambda(1+\alpha)}$$

Note that $\frac{1}{2}+\frac{\underline{k}}{2\lambda(1+\alpha)} > \frac{1}{2}+\frac{\bar{k}+\beta(1-\lambda)}{2\lambda(1+\bar{m})}$ since this is equivalent to the second argument in the support condition.

Hence we obtain the following phase diagram:



Differences in fitness payoffs:

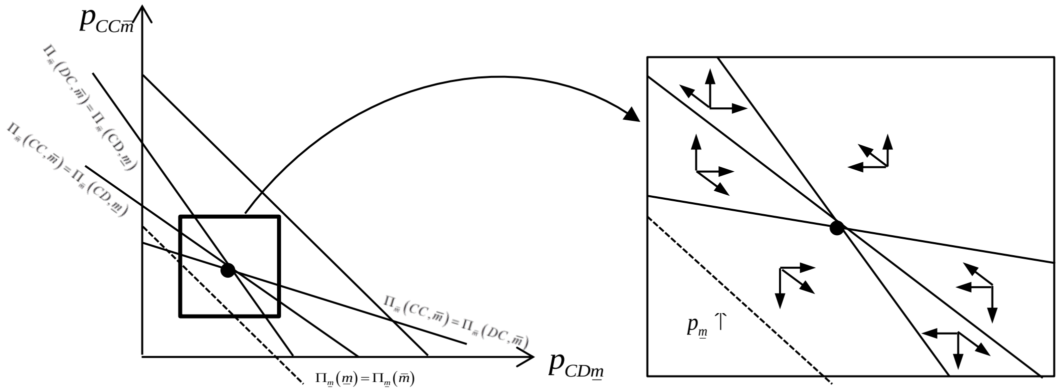
$$\Pi_{\bar{m}}(CD,\underline{m})-\Pi_{\underline{m}}(\underline{m}) = \lambda p_{DC\bar{m}}(1+\bar{m})-\lambda p_{DC\bar{m}}(1+\alpha) = \lambda(\bar{m}-\alpha)p_{DC\bar{m}} > 0$$

$$\left(\Pi_{\bar{m}}(CD,\underline{m})-\Pi_{\underline{m}}(\underline{m})\right)^{f} = -\alpha\lambda < 0$$

As the diagram clearly indicates this equilibrium is stable.

$$\Pi_{\bar{m}}(CD,\underline{m})-\Pi_{\underline{m}}(\underline{m}) = \lambda p_{DC\bar{m}}(1+\bar{m})-\lambda p_{DC\bar{m}}(1+\alpha) = \lambda(\bar{m}-\alpha)p_{DC\bar{m}} > 0$$

$$\left(\Pi_{\bar{m}}(CD,\underline{m})-\Pi_{\underline{m}}(\underline{m})\right)^{f} = -\alpha\lambda < 0$$

<table>
<tr>
<td rowspan="2">4.3.1.1.</td>
<td>$DC\bar{m}$
$CC\bar{m}$
$CD\underline{m}$
$\bar{m}$</td>
<td>$p_{DC\bar{m}} = \frac{(1+\bar{m})\left(\lambda(\bar{m}-\alpha+\beta)-\beta\right)-\bar{k}(\bar{m}-\alpha)}{\lambda\left(2(1+\bar{m})(\bar{m}-\alpha+\beta)-\beta(1+\alpha)\right)}$
$p_{CC\bar{m}} = \frac{\beta\left(2(1+\bar{m})+\beta-\lambda(1+\bar{m}+\beta)-\bar{k}\right)}{\lambda\left(2(1+\bar{m})(\bar{m}-\alpha+\beta)-\beta(1+\alpha)\right)}$
$p_{CD\underline{m}} = \frac{\bar{k}(\bar{m}-\alpha+\beta)+(1+\bar{m}+\beta)\left(\lambda(\bar{m}-\alpha+\beta)-\beta\right)}{\lambda\left(2(1+\bar{m})(\bar{m}-\alpha+\beta)-\beta(1+\alpha)\right)}$</td>
<td>$\frac{2}{\beta}\left(\underline{k}\frac{(1+\bar{m})}{(1+\alpha)}-\bar{k}\right)+\frac{\beta-(\underline{k}-\bar{k})}{(\bar{m}-\alpha+\beta)}<\lambda<1$</td>
</tr>
<tr>
<td colspan="3">We obtain the following phase diagram (derivation below):


As the diagram clearly indicates this equilibrium is unstable.</td>
</tr>
</table>

4.3.1.2.	$DC\bar{m}$ $CC\bar{m}$ $CD\underline{m}$ $\bar{m}$ $\underline{m}$	$\left\{\begin{array}{l} p_{DC\bar{m}}=\frac{1}{2}\left(1-\frac{\underline{k}-\bar{k}-\beta}{\lambda(\bar{m}-\alpha+\beta)}\right) \\ p_{CC\bar{m}}=\frac{\beta+\bar{k}-\frac{1+\bar{m}+\beta}{1+\alpha}\underline{k}}{\lambda(\bar{m}-\alpha+\beta)} \\ p_{CD\underline{m}}=\frac{1}{2}\left(1+\frac{\left(2\frac{(\bar{m}-\alpha+\beta)}{1+\alpha}+1\right)\underline{k}-\bar{k}-\beta}{\lambda(\bar{m}-\alpha+\beta)}\right) \\ p_{\underline{m}}=\frac{1}{\beta(1-\lambda)}\left(\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right)-\frac{1}{2(1-\lambda)(\bar{m}-\alpha+\beta)}(\underline{k}-\bar{k})-\frac{\lambda(\bar{m}-\alpha+\beta)-\beta}{2(1-\lambda)(\bar{m}-\alpha+\beta)} \end{array}\right\}$	1. $\frac{\beta+\underline{k}-\bar{k}}{(\bar{m}-\alpha+\beta)}<\lambda<\min\left\{\begin{array}{l}\frac{\beta+\bar{k}-\underline{k}}{(\bar{m}-\alpha+\beta)}+\frac{2}{\beta}\left(\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right), \\ 2-\left(\frac{\beta+\bar{k}-\underline{k}}{(\bar{m}-\alpha+\beta)}+\frac{2}{\beta}\left(\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right)\right)\end{array}\right\}$ 2. $\underline{k}-\bar{k}<\bar{m}-\alpha$ 3. $\underline{k}<1+\alpha$ 4. $(1+\alpha)\beta>(1+\bar{m}+\beta)\underline{k}-(1+\alpha)\bar{k}$

$\Pi_{\bar{m}}(CC,\bar{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{CD\underline{m}}\right)(1+\bar{m})+p_{DC\bar{m}}(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)+p_{\bar{m}}(-\beta)\right]-\bar{k}$

$\Pi_{\bar{m}}(DC,\bar{m})=\lambda\left[p_{CD\underline{m}}(1+\bar{m})+p_{CC\bar{m}}(1+\alpha)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]-\bar{k}$

$\Pi_{\bar{m}}(CD,\underline{m})=\lambda\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)(1+\bar{m})+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]$

$\Pi_{\underline{m}}(\underline{m})=\lambda\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)(1+\alpha)$

$\Pi_{\underline{m}}(\bar{m})=\lambda\left(p_{CC\bar{m}}+p_{CD\underline{m}}\right)(1+\alpha)-\underline{k}$

$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(DC,\bar{m})\geq 0\Leftrightarrow\lambda\left[\left(p_{CC\bar{m}}+p_{CD\underline{m}}\right)(1+\bar{m})+p_{DC\bar{m}}(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)+p_{\bar{m}}(-\beta)\right]-\bar{k}-\lambda\left[p_{CD\underline{m}}(1+\bar{m})+p_{CC\bar{m}}(1+\alpha)\right]-(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]+\bar{k}\geq 0\Leftrightarrow$

$\lambda\left[\left(p_{CC\bar{m}}\right)(\bar{m}-\alpha)+p_{DC\bar{m}}(-\beta)\right]\geq(1-\lambda)\beta p_{\bar{m}}\Leftrightarrow p_{CC\bar{m}}\geq\frac{(1-\lambda)\beta p_{\bar{m}}+\lambda\beta}{\lambda(\bar{m}-\alpha+\beta)}-\frac{\beta}{(\bar{m}-\alpha+\beta)}p_{CD\underline{m}}$

$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CD,\underline{m})\geq 0\Leftrightarrow\lambda\left[\left(p_{CC\bar{m}}+p_{CD\underline{m}}\right)(1+\bar{m})+p_{DC\bar{m}}(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)+p_{\bar{m}}(-\beta)\right]-\bar{k}-\lambda\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)(1+\bar{m})-(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]\geq 0\Leftrightarrow$

$\lambda\left[\left(p_{CD\underline{m}}\right)(1+\bar{m})-\left(1-p_{CD\underline{m}}-p_{CC\bar{m}}\right)(1+\bar{m}+\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]-\bar{k}\geq 0\Leftrightarrow\lambda\left[\left(p_{CD\underline{m}}\right)\left(2(1+\bar{m})+\beta\right)+p_{CC\bar{m}}(1+\bar{m}+\beta)\right]\geq(1-\lambda)\beta p_{\underline{m}}+\bar{k}+(1+\bar{m}+\beta)\lambda$

$\Leftrightarrow p_{CC\bar{m}}\geq 1+\frac{(1-\lambda)\beta p_{\underline{m}}+\bar{k}}{\lambda(1+\bar{m}+\beta)}-\frac{2(1+\bar{m})+\beta}{(1+\bar{m}+\beta)}p_{CD\underline{m}}$

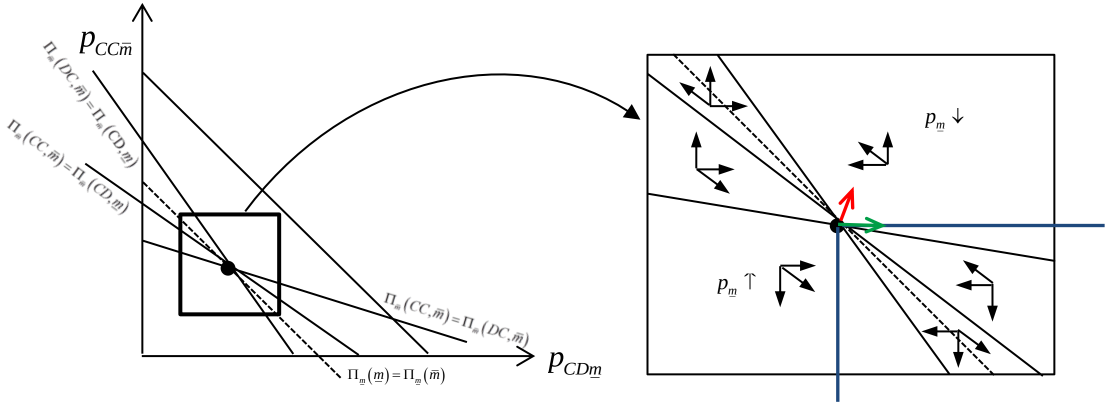
$\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(CD,\underline{m})\geq 0\Leftrightarrow\lambda\left[p_{CD\underline{m}}(1+\bar{m})+p_{CC\bar{m}}(1+\alpha)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]-\bar{k}-\lambda\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)(1+\bar{m})-(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]\geq 0\Leftrightarrow$

$\lambda\left[2p_{CD\underline{m}}(1+\bar{m})+p_{CC\bar{m}}(1+\alpha)\right]-\lambda(1+\bar{m})-\bar{k}+2(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]+(1-\lambda)\beta\geq 0\Leftrightarrow p_{CC\bar{m}}\geq\frac{\lambda(1+\bar{m})+\bar{k}+2(1-\lambda)\beta p_{\underline{m}}-(1-\lambda)\beta}{\lambda(1+\alpha)}-\frac{2(1+\bar{m})}{(1+\alpha)}p_{CD\underline{m}}$

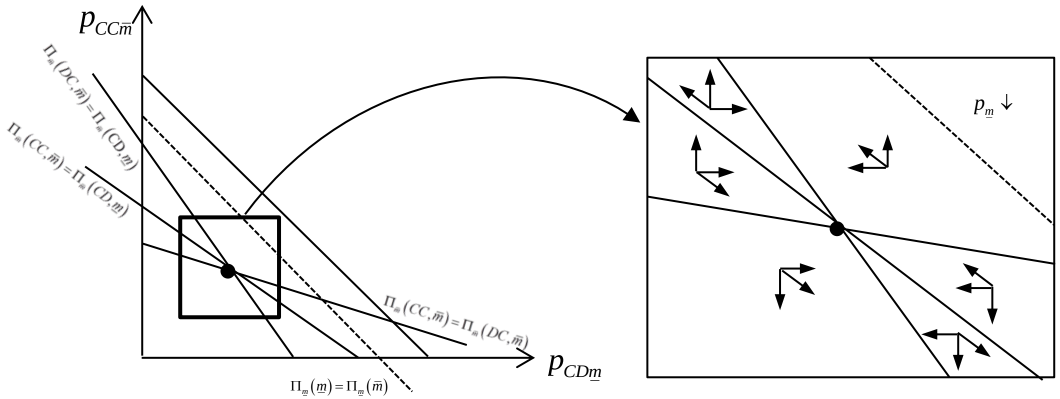
$\Pi_{\underline{m}}(\underline{m})-\Pi_{\underline{m}}(\bar{m})\geq 0\Leftrightarrow\lambda\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)(1+\alpha)-\lambda\left(p_{CC\bar{m}}+p_{CD\underline{m}}\right)(1+\alpha)+\underline{k}\geq 0\Leftrightarrow\lambda\left(1-p_{CD\underline{m}}\right)(1+\alpha)-\lambda\left(p_{CC\bar{m}}+p_{CD\underline{m}}\right)(1+\alpha)+\underline{k}\geq 0\Leftrightarrow p_{CC\bar{m}}\leq 1+\frac{\underline{k}}{\lambda(1+\alpha)}-2p_{CD\underline{m}}$

Note: $-\frac{2(1+\bar{m})}{(1+\alpha)}<-2<-\frac{2(1+\bar{m})+\beta}{(1+\bar{m}+\beta)}<-\frac{\beta}{(\bar{m}-\alpha+\beta)}$

Hence we obtain the following phase diagram:



We saw for the previously analyzed equilibrium with strict inequality with respect to the signaling decision among low types that this equilibrium is unstable. The question here is whether the indeterminacy of the low types in equilibrium could have a stabilizing effect. It turns out that it doesn't. The reason is that if $p_{\underline{m}}$ decreases the new intersection of the iso-profit lines (not an equilibrium) lies in the fourth quadrant relative to the equilibrium point. We will argue for the most favorite scenario that an adjustment in $p_{\underline{m}}$ will not stabilize the equilibrium. Consider therefor a perturbation of the type indicated by the red arrow. For such an perturbation equilibrium will not be restored in the absence of an adjusting $p_{\underline{m}}$. What kind of adjustment is most favorable with respect to stabilization? The instability can only be circumvented if the induced shift of the intersection point of iso-profit lines and thereby a shift of regions with the depicted dynamics would bring the pertubated point into a region with dynamics point at the equilibrium. Most favorable is a strong and fast movement to right at the boundary of the fourth quadrant (indicated by the green arrow). It is important to note that even this most favorable movement cannot induce the pertubated point to be pushed into a region to the left of the dotted line (unaltered by changes in $p_{\underline{m}}$), because than $p_{\underline{m}}$ would start to increase again. And if this is assumed to be fast and strong, than the iso-profit lines will be shifted back towards its equilibrium locations. In other words the only thing that can happen is that the population state pointed at by the red arrow is find itself in the area between the $\Pi_{\bar{m}}(DC,\bar{m}) = \Pi_{\bar{m}}(\mathrm{CD},\underline{m})$-isoline and the dotted line. However this will not lead to a reestablishment of equilibrium but to further movement away. The same argument applies to the second diagram.

<table>
<tr>
<td rowspan="2">4.3.1.3.</td>
<td>$DC\bar{m}$
$CC\bar{m}$
$CD\underline{m}$
$\underline{m}$</td>
<td>$p_{DC\bar{m}}=\frac{(\bar{m}-\alpha)\left((1+\bar{m}+\beta)\lambda-\beta-\bar{k}\right)}{\lambda\left(2(1+\bar{m})(\bar{m}-\alpha+\beta)-\beta(1+\alpha)\right)}$
$p_{CC\bar{m}}=\frac{\beta\left((1+\bar{m}+\beta)\lambda-\beta-\bar{k}\right)}{\lambda\left(2(1+\bar{m})(\bar{m}-\alpha+\beta)-\beta(1+\alpha)\right)}$
$p_{CD\underline{m}}=\frac{\bar{k}(\bar{m}-\alpha+\beta)+(1+\bar{m}+\beta)\lambda(\bar{m}-\alpha)+(\bar{m}-\alpha+\beta)(1-\lambda)\beta}{\lambda\left(2(1+\bar{m})(\bar{m}-\alpha+\beta)-\beta(1+\alpha)\right)}$</td>
<td>$\lambda>\max\left\{\underbrace{\frac{\beta+\bar{k}}{(1+\bar{m}+\beta)}}_{x},\underbrace{2-\left(\frac{\beta+\bar{k}-\underline{k}}{(\bar{m}-\alpha+\beta)}+\frac{2}{\beta}\left(\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right)\right)}_{y}\right\}$
note:
$y>x\Leftrightarrow 2-\left(\frac{\beta+\bar{k}-\underline{k}}{(\bar{m}-\alpha+\beta)}+\frac{2}{\beta}\left(\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right)\right)-\frac{\beta+\bar{k}}{(1+\bar{m}+\beta)}>0$
$\Leftrightarrow (1+\alpha)\beta>(1+\bar{m}+\beta)\underline{k}-(1+\alpha)\bar{k} \quad \wedge \quad \underline{k}<1+\alpha$</td>
<td>$\bar{k}<1+\bar{m}$</td>
</tr>
<tr>
<td colspan="4">We obtain the following phase diagram (derivation above):

As the diagram clearly indicates this equilibrium is unstable.</td>
</tr>
</table>

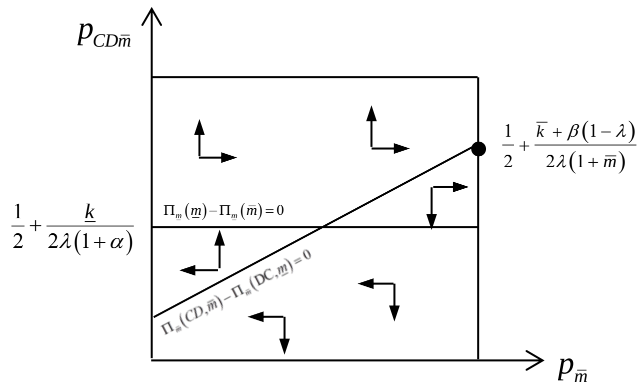
4.4.1.1.	$CD\bar{m}$ $DC\underline{m}$ $\bar{m}$	$p_{DC\underline{m}} = \frac{1}{2} - \frac{\bar{k} + \beta(1-\lambda)}{2\lambda(1+\bar{m})}$ $p_{CD\bar{m}} = \frac{1}{2} + \frac{\bar{k} + \beta(1-\lambda)}{2\lambda(1+\bar{m})}$	$0 \overset{2.}{<} \frac{1+\alpha}{(\bar{m}-\alpha+\beta)(1+\bar{m})+(1+\alpha)\beta}\left(\beta\left(1+\frac{1+\bar{m}}{1+\alpha}\right)-\frac{\bar{k}}{1+\alpha}(\bar{m}-\alpha)\right) < \lambda$ $< 1 - \frac{1}{\beta}\left(\underline{k}\frac{1+\bar{m}}{1+\alpha} - \bar{k}\right) < 1$	1. $(1+\alpha)\beta > (\bar{m}-\alpha+\beta)\bar{k}$ 2. $\underline{k} < \frac{(1+\alpha)\left(\beta(\bar{m}-\alpha)+\bar{k}(\bar{m}-\alpha+2\beta)\right)}{(\bar{m}-\alpha+\beta)(1+\bar{m})+(1+\alpha)\beta}$

as below:

$$\Pi_{\bar{m}}(CD,\bar{m}) - \Pi_{\bar{m}}(DC,\underline{m}) \geq 0 \Leftrightarrow \lambda p_{CD\bar{m}}(1+\bar{m}) + (1-\lambda)\left[p_{\bar{m}}(-\beta)\right] - \bar{k} - \lambda p_{DC\underline{m}}(1+\bar{m}) - (1-\lambda)\left[p_{\underline{m}}(-\beta)\right] \geq 0 \Leftrightarrow p_{CD\bar{m}} \geq \frac{2\beta(1-\lambda)p_{\bar{m}} + \bar{k} + \lambda(1+\bar{m}) - \beta(1-\lambda)}{2\lambda(1+\bar{m})} = \frac{1}{2} + \frac{\bar{k} - \beta(1-\lambda) + 2\beta(1-\lambda)p_{\bar{m}}}{2\lambda(1+\bar{m})}$$

$$\Pi_{\underline{m}}(\underline{m}) - \Pi_{\underline{m}}(\bar{m}) \geq 0 \Leftrightarrow \lambda p_{DC\underline{m}}(1+\alpha) - \lambda p_{CD\bar{m}}(1+\alpha) + \underline{k} \geq 0 \Leftrightarrow p_{CD\bar{m}} \leq \frac{\lambda(1+\alpha)+\underline{k}}{2\lambda(1+\alpha)} = \frac{1}{2} + \frac{\underline{k}}{2\lambda(1+\alpha)}$$

Hence we obtain the following phase diagram:



As the diagram clearly indicates this equilibrium is unstable. Note $\frac{1}{2} + \frac{\bar{k}+\beta(1-\lambda)}{2\lambda(1+\bar{m})} > \frac{1}{2} + \frac{\underline{k}}{2\lambda(1+\alpha)} \Leftrightarrow \frac{\bar{k}+\beta(1-\lambda)}{(1+\bar{m})} > \frac{\underline{k}}{(1+\alpha)} \Leftrightarrow \lambda < 1 - \frac{1}{\beta}\left(\underline{k}\frac{1+\bar{m}}{1+\alpha} - \bar{k}\right)$, what is equivalent to the upper bound of the support.

4.4.1.2.

$CD\bar{m}$ $DC\underline{m}$ $\bar{m}$ $\underline{m}$	$p_{CD\bar{m}} = \frac{1}{2} + \frac{\underline{k}}{2\lambda(1+\alpha)}$ $p_{DC\underline{m}} = \frac{1}{2} - \frac{\underline{k}}{2\lambda(1+\alpha)}$ $p_{\bar{m}} = \frac{1}{2} + \frac{1}{2\beta(1-\lambda)}\left[\frac{(1+\bar{m})}{(1+\alpha)}\underline{k} - \bar{k}\right]$	$\frac{\beta+\underline{k}-\bar{k}}{(\bar{m}-\alpha+\beta)} < \lambda < 1 - \frac{1}{\beta}\left(\frac{(1+\bar{m})}{(1+\alpha)}\underline{k} - \bar{k}\right)$	1. $(1+\alpha)\beta > (\bar{m}-\alpha+\beta)\bar{k}$ 2. $\underline{k} < \frac{(1+\alpha)(\beta(\bar{m}-\alpha)+\bar{k}(\bar{m}-\alpha+2\beta))}{(\bar{m}-\alpha+\beta)(1+\bar{m})+(1+\alpha)\beta}$ 3. $\bar{k} < 1+\alpha$

$\Pi_{\bar{m}}(CD,\bar{m}) = \lambda p_{CD\bar{m}}(1+\bar{m}) + (1-\lambda)\left[p_{\bar{m}}(-\beta)\right] - \bar{k}$

$\Pi_{\bar{m}}(DC,\underline{m}) = \lambda p_{DC\underline{m}}(1+\bar{m}) + (1-\lambda)\left[p_{\underline{m}}(-\beta)\right]$

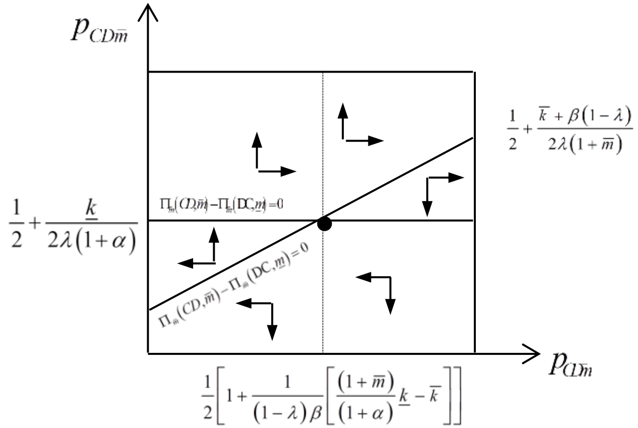
$\Pi_{\underline{m}}(\underline{m}) = \lambda p_{DC\underline{m}}(1+\alpha)$

$\Pi_{\underline{m}}(\bar{m}) = \lambda p_{CD\bar{m}}(1+\alpha) - \underline{k}$

$$\Pi_{\bar{m}}(CD,\bar{m}) - \Pi_{\bar{m}}(DC,\underline{m}) \geq 0 \Leftrightarrow \lambda p_{CD\bar{m}}(1+\bar{m}) + (1-\lambda)\left[p_{\bar{m}}(-\beta)\right] - \bar{k} - \lambda p_{DC\underline{m}}(1+\bar{m}) - (1-\lambda)\left[p_{\underline{m}}(-\beta)\right] \geq 0 \Leftrightarrow p_{CD\bar{m}} \geq \frac{2\beta(1-\lambda)p_{\bar{m}} + \bar{k} + \lambda(1+\bar{m}) - \beta(1-\lambda)}{2\lambda(1+\bar{m})} = \frac{1}{2} + \frac{\bar{k} - \beta(1-\lambda) + 2\beta(1-\lambda)p_{\bar{m}}}{2\lambda(1+\bar{m})}$$

$$\Pi_{\underline{m}}(\underline{m}) - \Pi_{\underline{m}}(\bar{m}) \geq 0 \Leftrightarrow \lambda p_{DC\underline{m}}(1+\alpha) - \lambda p_{CD\bar{m}}(1+\alpha) + \underline{k} \geq 0 \Leftrightarrow p_{CD\bar{m}} \leq \frac{\lambda(1+\alpha)+\underline{k}}{2\lambda(1+\alpha)} = \frac{1}{2} + \frac{\underline{k}}{2\lambda(1+\alpha)}$$

Hence we obtain the following phase diagram:



Note:

$$\frac{1}{2} + \frac{\bar{k} + \beta(1-\lambda)}{2\lambda(1+\bar{m})} > \frac{1}{2} + \frac{\underline{k}}{2\lambda(1+\alpha)}$$

$$\Leftrightarrow \frac{\bar{k} + \beta(1-\lambda)}{(1+\bar{m})} > \frac{\underline{k}}{(1+\alpha)} \Leftrightarrow \lambda < 1 - \frac{1}{\beta}\left(\underline{k}\frac{1+\bar{m}}{1+\alpha} - \bar{k}\right)$$

, what is equivalent to the upper bound of the support.
As the diagram clearly indicates this equilibrium is unstable.

4.4.1.3.				
	$CD\bar{m}$ $DC\underline{m}$ $\underline{m}$	$p_{CD\bar{m}} = \frac{1}{2}\left(1+\frac{\bar{k}-\beta(1-\lambda)}{\lambda(1+\bar{m})}\right)$ $p_{DC\underline{m}} = \frac{1}{2}\left(1-\frac{\bar{k}-\beta(1-\lambda)}{\lambda(1+\bar{m})}\right)$	$\frac{\bar{k}(\bar{m}-\alpha)+\beta(2+\alpha+\bar{m})}{(1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha)} < \lambda < 1$	1. $\bar{k} < (1+\bar{m})$

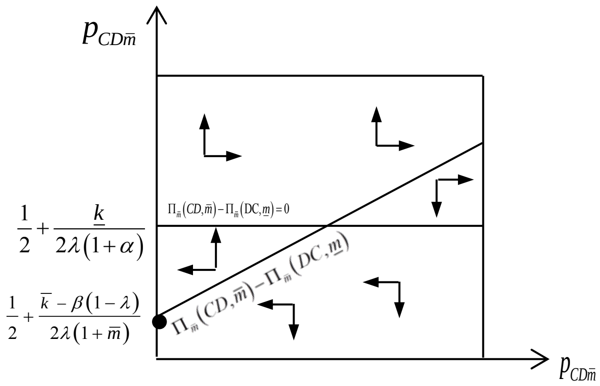
as above:

$\Pi_{\bar{m}}(CD,\bar{m})-\Pi_{\bar{m}}(DC,\underline{m}) \geq 0 \Leftrightarrow \lambda p_{CD\bar{m}}(1+\bar{m})+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]-\bar{k}-\lambda p_{DC\underline{m}}(1+\bar{m})-(1-\lambda)\left[p_{\underline{m}}(-\beta)\right] \geq 0$

$\Leftrightarrow p_{CD\bar{m}} \geq \frac{2\beta(1-\lambda)p_{\bar{m}}+\bar{k}+\lambda(1+\bar{m})-\beta(1-\lambda)}{2\lambda(1+\bar{m})} = \frac{1}{2}+\frac{\bar{k}-\beta(1-\lambda)+2\beta(1-\lambda)p_{\bar{m}}}{2\lambda(1+\bar{m})}$

$\Pi_{\underline{m}}(\underline{m})-\Pi_{\underline{m}}(\bar{m}) \geq 0 \Leftrightarrow \lambda p_{DC\underline{m}}(1+\alpha)-\lambda p_{CD\bar{m}}(1+\alpha)+\underline{k} \geq 0 \Leftrightarrow p_{CD\bar{m}} \leq \frac{\lambda(1+\alpha)+\underline{k}}{2\lambda(1+\alpha)} = \frac{1}{2}+\frac{\underline{k}}{2\lambda(1+\alpha)}$

Hence we obtain the following phase diagram:



Note: $\frac{1}{2}+\frac{\bar{k}-\beta(1-\lambda)}{2\lambda(1+\bar{m})} < \frac{1}{2}+\frac{\underline{k}}{2\lambda(1+\alpha)}$

As the diagram clearly indicates this equilibrium is unstable.

4.4.2.	$CD\bar{m}$ $DD\underline{m}$ $\underline{m}$	$p_{CD\bar{m}} = \dfrac{\bar{k}}{\lambda(1+\bar{m})}$ $p_{DD\underline{m}} = 1 - \dfrac{\bar{k}}{\lambda(1+\bar{m})}$	$\lambda > \dfrac{\bar{k}}{(1+\bar{m})}$	1. $\bar{k} < (1+\bar{m})$

$\Pi_{\bar{m}}(CD,\bar{m}) = \lambda p_{CD\bar{m}}(1+\bar{m}) + (1-\lambda)\left[p_{\bar{m}}(-\beta)\right] - \bar{k}$

$\Pi_{\bar{m}}(DD,\underline{m}) = 0$

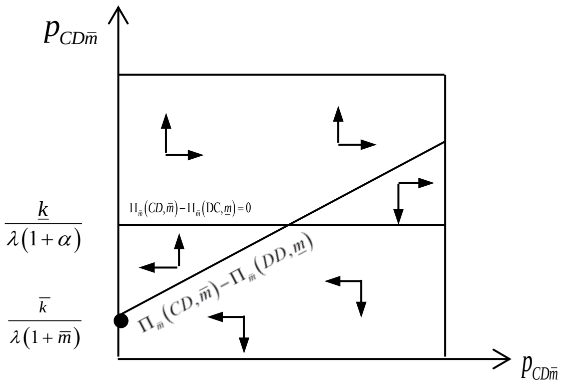
$\Pi_{\underline{m}}(\underline{m}) = 0$

$\Pi_{\underline{m}}(\bar{m}) = \lambda p_{CD\bar{m}}(1+\alpha) - \underline{k}$

$\Pi_{\bar{m}}(CD,\bar{m}) - \Pi_{\bar{m}}(DD,\underline{m}) \geq 0 \Leftrightarrow \lambda p_{CD\bar{m}}(1+\bar{m}) + (1-\lambda)\left[p_{\bar{m}}(-\beta)\right] - \bar{k} \geq 0 \Leftrightarrow p_{CD\bar{m}} \geq \dfrac{\bar{k} + \beta(1-\lambda)p_{\bar{m}}}{\lambda(1+\bar{m})}$

$\Pi_{\underline{m}}(\underline{m}) - \Pi_{\underline{m}}(\bar{m}) \geq 0 \Leftrightarrow -\lambda p_{CD\bar{m}}(1+\alpha) + \underline{k} \geq 0 \Leftrightarrow p_{CD\bar{m}} \leq \dfrac{\underline{k}}{\lambda(1+\alpha)}$

Hence we obtain the following phase diagram:



As the diagram clearly indicates this equilibrium is unstable.

4.6.1.1.	$CD\bar{m}$ $DD\bar{m}$ $DC\underline{m}$ $\bar{m}$	$p_{DC\underline{m}} = \frac{1}{\lambda}\left[\frac{\beta(1+\alpha)-\bar{k}(\bar{m}-\alpha+\beta)}{\beta(1+\alpha)+(1+\bar{m})(\bar{m}-\alpha+\beta)}\right]$ $p_{CD\bar{m}} = \frac{1}{\lambda}\left[\frac{\beta(1+\bar{m}+\bar{k})}{\beta(1+\alpha)+(1+\bar{m})(\bar{m}-\alpha+\beta)}\right]$ $p_{DD\bar{m}} = 1+\frac{1}{\lambda}\left[\frac{\bar{k}(\bar{m}-\alpha)-\beta(2+\bar{m}+\alpha)}{\beta(1+\alpha)+(1+\bar{m})(\bar{m}-\alpha+\beta)}\right]$	$\frac{\bar{k}(\bar{m}-\alpha)+\beta(2+\alpha+\bar{m})}{(1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha)} < \lambda < 1$	1. $\beta(1+\alpha) > \bar{k}(\bar{m}-\alpha+\beta)$ 2. $\underline{k} < \frac{(1+\alpha)(\beta(\bar{m}-\alpha)+\bar{k}(\bar{m}-\alpha+2\beta))}{(1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha)}$

$\Pi_{\bar{m}}(CD,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})\geq 0 \Leftrightarrow \lambda\left[p_{CD\bar{m}}(1+\bar{m})+p_{DD\bar{m}}(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]-\lambda p_{CD\bar{m}}(1+\alpha)\geq 0 \Leftrightarrow \lambda p_{CD\bar{m}}(\bar{m}-\alpha)+\lambda(1-p_{CD\bar{m}}-p_{DC\underline{m}})(-\beta)+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]\geq 0 \Leftrightarrow$

$$\lambda p_{CD\bar{m}}(\bar{m}-\alpha+\beta)-\lambda\beta+\lambda\beta p_{DC\underline{m}}+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]\geq 0 \Leftrightarrow p_{CD\bar{m}}\geq\frac{\lambda\beta-\lambda\beta p_{DC\underline{m}}+(1-\lambda)\beta p_{\bar{m}}}{\lambda(\bar{m}-\alpha+\beta)}=\frac{\lambda\beta+(1-\lambda)\beta p_{\bar{m}}}{\lambda(\bar{m}-\alpha+\beta)}-\frac{\beta}{(\bar{m}-\alpha+\beta)}p_{DC\underline{m}}\overset{p_{\bar{m}}=1}{=}\frac{\beta}{\lambda(\bar{m}-\alpha+\beta)}-\frac{\beta}{(\bar{m}-\alpha+\beta)}p_{DC\underline{m}}$$

$\Pi_{\bar{m}}(CD,\bar{m})-\Pi_{\bar{m}}(DC,\underline{m})\geq 0 \Leftrightarrow \lambda\left[p_{CD\bar{m}}(1+\bar{m})+p_{DD\bar{m}}(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]-\bar{k}-\lambda p_{DC\underline{m}}(1+\bar{m})-(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]\geq 0 \Leftrightarrow$

$$\lambda\left[p_{CD\bar{m}}(1+\bar{m})+(1-p_{CD\bar{m}}-p_{DC\underline{m}})(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]-\bar{k}-\lambda p_{DC\underline{m}}(1+\bar{m})-(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]\geq 0 \Leftrightarrow \lambda p_{CD\bar{m}}(1+\bar{m}+\beta)-\lambda p_{DC\underline{m}}(1+\bar{m}-\beta)-\lambda\beta-\beta(1-\lambda)p_{\bar{m}}-\bar{k}+\beta(1-\lambda)p_{\underline{m}}\geq 0 \Leftrightarrow$$

$$\Leftrightarrow p_{CD\bar{m}}\geq\frac{\lambda p_{DC\underline{m}}(1+\bar{m}-\beta)+2\lambda\beta+2\beta(1-\lambda)p_{\bar{m}}+\bar{k}-\beta}{\lambda(1+\bar{m}+\beta)}=\frac{2\lambda\beta+2\beta(1-\lambda)p_{\bar{m}}+\bar{k}-\beta}{\lambda(1+\bar{m}+\beta)}+\frac{(1+\bar{m}-\beta)}{(1+\bar{m}+\beta)}p_{DC\underline{m}}\overset{p_{\bar{m}}=1}{=}\frac{\beta+\bar{k}}{\lambda(1+\bar{m}+\beta)}+\frac{(1+\bar{m}-\beta)}{(1+\bar{m}+\beta)}p_{DC\underline{m}}$$

$\Pi_{\bar{m}}(DD,\bar{m})-\Pi_{\bar{m}}(DC,\underline{m})\geq 0 \Leftrightarrow \lambda p_{CD\bar{m}}(1+\alpha)-\bar{k}-\lambda p_{DC\underline{m}}(1+\bar{m})-(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]\geq 0 \Leftrightarrow p_{CD\bar{m}}\geq\frac{\bar{k}-\beta(1-\lambda)p_{\underline{m}}}{\lambda(1+\alpha)}+\frac{(1+\bar{m})}{(1+\alpha)}p_{DC\underline{m}}\overset{p_{\bar{m}}=1}{=}\frac{\bar{k}}{\lambda(1+\alpha)}+\frac{(1+\bar{m})}{(1+\alpha)}p_{DC\underline{m}}$

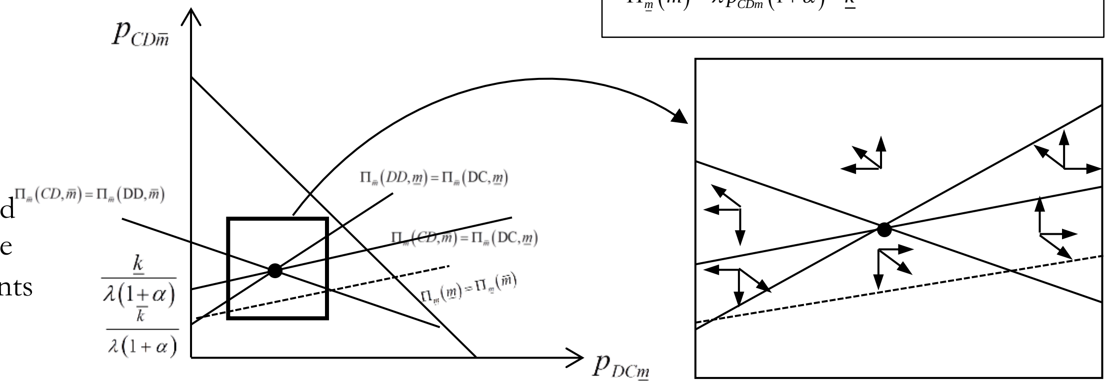
$\Pi_{\underline{m}}(\underline{m})-\Pi_{\underline{m}}(\bar{m})\geq 0 \Leftrightarrow \lambda p_{DC\underline{m}}(1+\alpha)-\lambda p_{CD\bar{m}}(1+\alpha)+\underline{k}\geq 0 \Leftrightarrow p_{CD\bar{m}}\leq\frac{\underline{k}}{\lambda(1+\alpha)}+p_{DC\underline{m}}$

$\Pi_{\bar{m}}(CD,\bar{m})$
$=\lambda\left[p_{CD\bar{m}}(1+\bar{m})+p_{DD\bar{m}}(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]-\bar{k}$
$\Pi_{\bar{m}}(DD,\bar{m})=\lambda p_{CD\bar{m}}(1+\alpha)-\bar{k}$
$\Pi_{\bar{m}}(DC,\underline{m})=\lambda p_{DC\underline{m}}(1+\bar{m})+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]$
$\Pi_{\underline{m}}(\underline{m})=\lambda p_{DC\underline{m}}(1+\alpha)$
$\Pi_{\underline{m}}(\bar{m})=\lambda p_{CD\bar{m}}(1+\alpha)-\underline{k}$

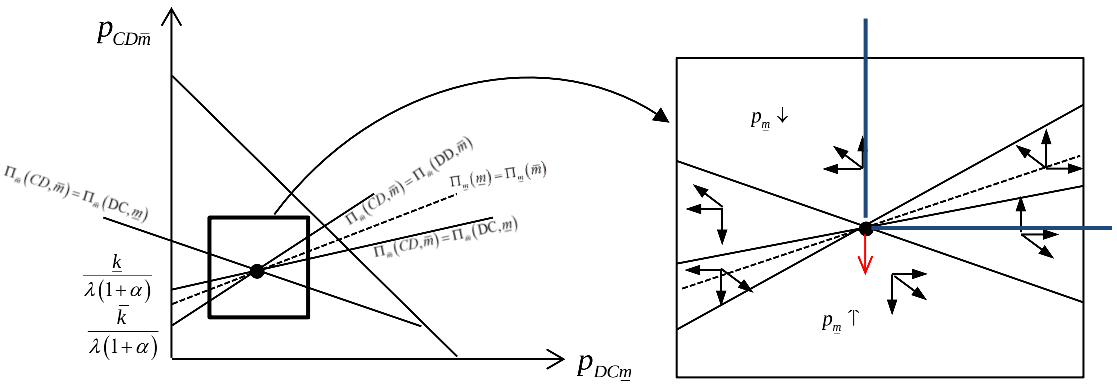
Note that $-\frac{\beta}{(\bar{m}-\alpha+\beta)}<\frac{(1+\bar{m}-\beta)}{(1+\bar{m}+\beta)}<1<\frac{(1+\bar{m})}{(1+\alpha)}$.

Hence we obtain the following phase diagram:

As the diagram clearly indicates this equilibrium is unstable. If in the course of the dynamics the dotted line is crossed, low types start to prefer not to signal since signaling is not often enough rewarded by $CD\bar{m}$-player and too often punished by $DC\underline{m}$-player. The induced decline in the share of signaling low types will shift all three constraints downwards.



<table>
<tr>
<td rowspan="2">4.6.1.2.</td>
<td>$CD\bar{m}$
$DD\bar{m}$
$DC\underline{m}$
$\bar{m}$
$\underline{m}$</td>
<td>$p_{DC\underline{m}} = \frac{\beta(1+\alpha)+(1-\bar{m}-\beta+2\alpha)\underline{k}-(1+\alpha)\bar{k}}{2(1+\alpha)(\bar{m}-\alpha+\beta)}$
$p_{CD\bar{m}} = \frac{\beta(1+\alpha)+(1+\bar{m}+\beta)\underline{k}-(1+\alpha)\bar{k}}{2(1+\alpha)(\bar{m}-\alpha+\beta)}$
$p_{DD\bar{m}} = \left[\frac{-\beta+\lambda(\bar{m}-\alpha+\beta)+\bar{k}-\underline{k}}{\lambda(\bar{m}-\alpha+\beta)}\right],$
$p_{\underline{m}} = -\frac{(1+\alpha)(-\beta(\bar{m}-\alpha)+(\alpha-2\beta-\bar{m})\bar{k})+((1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha))\underline{k}}{2(1+\alpha)(1-\lambda)(\bar{m}-\alpha+\beta)}$
$p_{\bar{m}} = 1-p_{\underline{m}}$</td>
<td>See Appendix B.4</td>
<td></td>
</tr>
<tr>
<td colspan="4">Hence we obtain the following phase diagram:



We will apply here the same logic as in the case for the equilibrium 4.3.1.2. Consider a perturbation that pushes the population state in the lower triangular region (red arrow). Given that low types will strictly prefer not to signal, which in turn shifts the intersection point of the iso-profit lines into the first quadrant relative to the equilibrium. As the picture clearly indicates this will not help to stabilize the equilibrium.</td>
</tr>
</table>

4.6.1.3.	$CD\bar{m}$ $DD\bar{m}$ $DC\underline{m}$ $\underline{m}$	$p_{CD\bar{m}} = \frac{\beta\left(\lambda\left(1+\bar{m}+\beta\right)+\bar{k}-\beta\right)}{\lambda\left(\left(1+\bar{m}\right)\left(\bar{m}-\alpha+\beta\right)+\beta\left(1+\alpha\right)\right)}$ $p_{DC\underline{m}} = \frac{\beta\left(\left(1-\lambda\right)\left(\bar{m}-\alpha+\beta\right)+\lambda\left(1+\alpha\right)\right)-\left(\bar{m}-\alpha+\beta\right)\bar{k}}{\lambda\left(\left(1+\bar{m}\right)\left(\bar{m}-\alpha+\beta\right)+\beta\left(1+\alpha\right)\right)}$ $p_{DD\bar{m}} = \frac{\left(\bar{m}-\alpha\right)\left(\lambda\left(1+\bar{m}+\beta\right)+\bar{k}-\beta\right)}{\lambda\left(\left(1+\bar{m}\right)\left(\bar{m}-\alpha+\beta\right)+\beta\left(1+\alpha\right)\right)}$	$\frac{\left(1+\alpha\right)\beta\left(\bar{m}-\alpha+2\beta\right)+\left(\bar{m}-\alpha\right)\left(\bar{m}-\alpha+\beta\right)\bar{k}}{2\beta\left(1+\alpha\right)\left(\bar{m}-\alpha+\beta\right)} < \lambda <$ $\frac{\left(1+\alpha\right)\beta\left(\bar{m}-\alpha+2\beta\right)+\left(\left(1+\bar{m}\right)\left(\bar{m}-\alpha+\beta\right)+\beta\left(1+\alpha\right)\right)\underline{k}-\left(\left(1+\alpha\right)\left(\bar{m}-\alpha+\beta\right)+\beta\left(1+\alpha\right)\right)\bar{k}}{2\beta\left(1+\alpha\right)\left(\bar{m}-\alpha+\beta\right)}$	$\beta\left(1+\alpha\right) >$ $\frac{\left(\left(1+\bar{m}\right)\left(\bar{m}-\alpha+\beta\right)+\beta\left(1+\alpha\right)\right)}{\left(\bar{m}-\alpha\right)}\underline{k}$ $-\frac{\left(\left(1+\alpha\right)\left(\bar{m}-\alpha+\beta\right)+\beta\left(1+\alpha\right)\right)}{\left(\bar{m}-\alpha\right)}\bar{k}$

as above

$$\Pi_{\bar{m}}\left(CD,\bar{m}\right)-\Pi_{\bar{m}}\left(DD,\bar{m}\right)\geq 0 \Leftrightarrow p_{CD\bar{m}} \geq \frac{\lambda\beta-\lambda\beta p_{DC\underline{m}}+\left(1-\lambda\right)\beta p_{\bar{m}}}{\lambda\left(\bar{m}-\alpha+\beta\right)} = \frac{\lambda\beta+\left(1-\lambda\right)\beta p_{\bar{m}}}{\lambda\left(\bar{m}-\alpha+\beta\right)} - \frac{\beta}{\left(\bar{m}-\alpha+\beta\right)}p_{DC\underline{m}} \overset{p_{\bar{m}}=0}{=} \frac{\lambda\beta}{\lambda\left(\bar{m}-\alpha+\beta\right)} - \frac{\beta}{\left(\bar{m}-\alpha+\beta\right)}p_{DC\underline{m}}$$

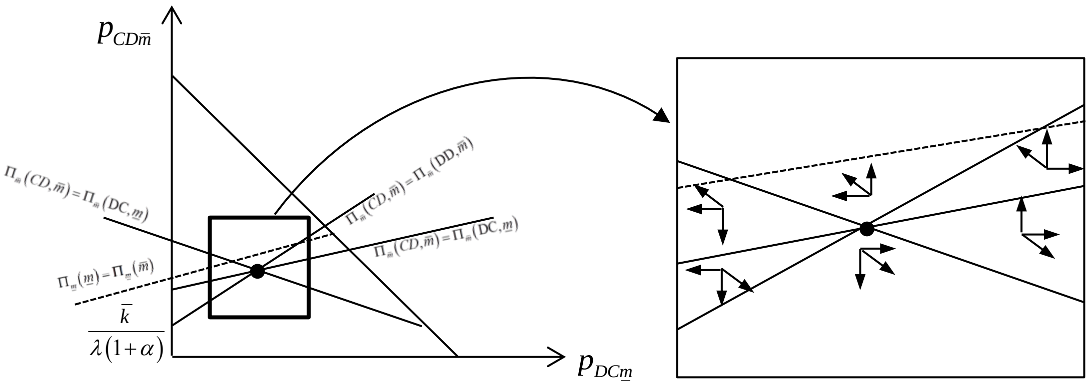
$$\Pi_{\bar{m}}\left(CD,\bar{m}\right)-\Pi_{\bar{m}}\left(DC,\underline{m}\right)\geq 0 \Leftrightarrow p_{CD\bar{m}} \geq \frac{\lambda p_{DC\underline{m}}\left(1+\bar{m}-\beta\right)+2\lambda\beta+2\beta\left(1-\lambda\right)p_{\bar{m}}+\bar{k}-\beta}{\lambda\left(1+\bar{m}+\beta\right)} = \frac{2\lambda\beta+2\beta\left(1-\lambda\right)p_{\bar{m}}+\bar{k}-\beta}{\lambda\left(1+\bar{m}+\beta\right)} + \frac{\left(1+\bar{m}-\beta\right)}{\left(1+\bar{m}+\beta\right)}p_{DC\underline{m}} \overset{p_{\bar{m}}=0}{=} \frac{2\lambda\beta-\beta+\bar{k}}{\lambda\left(1+\bar{m}+\beta\right)} + \frac{\left(1+\bar{m}-\beta\right)}{\left(1+\bar{m}+\beta\right)}p_{DC\underline{m}}$$

$$\Pi_{\bar{m}}\left(DD,\bar{m}\right)-\Pi_{\bar{m}}\left(DC,\underline{m}\right)\geq 0 \Leftrightarrow p_{CD\bar{m}} \geq \frac{\bar{k}-\beta\left(1-\lambda\right)p_{\underline{m}}}{\lambda\left(1+\alpha\right)} + \frac{\left(1+\bar{m}\right)}{\left(1+\alpha\right)}p_{DC\underline{m}} \overset{p_{\bar{m}}=0}{=} \frac{\bar{k}-\beta\left(1-\lambda\right)}{\lambda\left(1+\alpha\right)} + \frac{\left(1+\bar{m}\right)}{\left(1+\alpha\right)}p_{DC\underline{m}}$$

$$\Pi_{\underline{m}}\left(\underline{m}\right)-\Pi_{\underline{m}}\left(\bar{m}\right)\geq 0 \Leftrightarrow \lambda p_{DC\underline{m}}\left(1+\alpha\right)-\lambda p_{CD\bar{m}}\left(1+\alpha\right)+\underline{k}\geq 0 \Leftrightarrow p_{CD\bar{m}} \leq \frac{\underline{k}}{\lambda\left(1+\alpha\right)}+p_{DC\underline{m}}$$

Note that $-\frac{\beta}{\left(\bar{m}-\alpha+\beta\right)} < \frac{\left(1+\bar{m}-\beta\right)}{\left(1+\bar{m}+\beta\right)} < 1 < \frac{\left(1+\bar{m}\right)}{\left(1+\alpha\right)}$.

In comparison to 4.6.1.1. the three lines corresponding to equal profits among the equilibrium strategies shifts such that the equilibrium lies below the dotted line, which is constant with respect to changes in the share of signaling low types.
Hence we obtain the following phase diagram:



As the diagram clearly indicates this equilibrium is unstable. If in the course of the dynamics the dotted line is crossed, low types start to prefer to signal since signaling is often enough rewarded by $CD\bar{m}$-player and not too often punished by $DC\underline{m}$-player. The induced incline in the share of signaling low types will shift all three constraints upwards.

4.6.2.	$CD\bar{m}$ $DD\bar{m}$ $DD\underline{m}$ $\underline{m}$	$p_{CD\bar{m}}=\frac{\underline{k}}{\lambda(1+\alpha)}$ $p_{DD\bar{m}}=\frac{(\bar{m}-\alpha)}{\beta}\frac{\underline{k}}{\lambda(1+\alpha)}$ $p_{DD\underline{m}}=1-\frac{(\bar{m}-\alpha+\beta)}{\beta}\frac{\underline{k}}{\lambda(1+\alpha)}$	$\frac{\bar{k}(\bar{m}-\alpha+\beta)}{(1+\alpha)\beta}<\lambda<1$	$\beta(1+\alpha)>(\bar{m}-\alpha+\beta)\bar{k}$

$\Pi_{\bar{m}}(CD,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})\ge 0 \Leftrightarrow \lambda\left[p_{CD\bar{m}}(1+\bar{m})+p_{DD\bar{m}}(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]-\bar{k}-\lambda p_{CD\bar{m}}(1+\alpha)+\bar{k}\ge 0 \Leftrightarrow \lambda\left[p_{CD\bar{m}}(\bar{m}-\alpha)+(1-p_{CD\bar{m}}-p_{DD\underline{m}})(-\beta)\right]\ge(1-\lambda)\beta p_{\bar{m}} \Leftrightarrow$

$$p_{CD\bar{m}}\ge\frac{(1-\lambda)\beta p_{\bar{m}}+\lambda(1-p_{DD\underline{m}})\beta}{\lambda(\bar{m}-\alpha+\beta)}=\frac{(1-\lambda)\beta p_{\bar{m}}+\lambda\beta}{\lambda(\bar{m}-\alpha+\beta)}-\frac{\beta}{(\bar{m}-\alpha+\beta)}p_{DD\underline{m}}\overset{p_{\bar{m}}=0}{=}\frac{\beta}{(\bar{m}-\alpha+\beta)}-\frac{\beta}{(\bar{m}-\alpha+\beta)}p_{DD\underline{m}}$$

$\Pi_{\bar{m}}(CD,\bar{m})-\Pi_{\bar{m}}(DD,\underline{m})\ge 0 \Leftrightarrow \lambda\left[p_{CD\bar{m}}(1+\bar{m})+p_{DD\bar{m}}(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]-\bar{k}\ge 0 \Leftrightarrow \lambda\left[p_{CD\bar{m}}(1+\bar{m})+(1-p_{CD\bar{m}}-p_{DD\underline{m}})(-\beta)\right]\ge\bar{k}+(1-\lambda)\beta p_{\bar{m}} \Leftrightarrow$

$$p_{CD\bar{m}}\ge\frac{\bar{k}+(1-\lambda)\beta p_{\bar{m}}+\lambda\beta}{\lambda(1+\bar{m}+\beta)}-\frac{\beta}{(1+\bar{m}+\beta)}p_{DD\underline{m}}\overset{p_{\bar{m}}=0}{=}\frac{\bar{k}+\lambda\beta}{\lambda(1+\bar{m}+\beta)}-\frac{\beta}{(1+\bar{m}+\beta)}p_{DD\underline{m}}$$

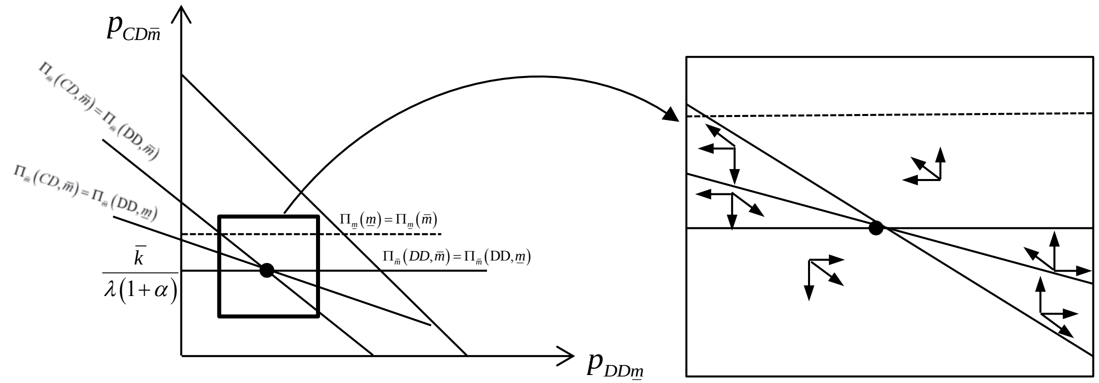
$\Pi_{\bar{m}}(DD,\bar{m})-\Pi_{\bar{m}}(DD,\underline{m})\ge 0 \Leftrightarrow \lambda p_{CD\bar{m}}(1+\alpha)-\bar{k}\ge 0 \Leftrightarrow p_{CD\bar{m}}\ge\frac{\bar{k}}{\lambda(1+\alpha)}$

$\Pi_{\underline{m}}(\underline{m})-\Pi_{\underline{m}}(\bar{m})\ge 0 \Leftrightarrow -\lambda p_{CD\bar{m}}(1+\alpha)+\underline{k}\ge 0 \Leftrightarrow p_{CD\bar{m}}\le\frac{\underline{k}}{\lambda(1+\alpha)}$

Note: $-\frac{\beta}{(1+\bar{m}+\beta)}>-\frac{\beta}{(\bar{m}-\alpha+\beta)}$

$$\Pi_{\bar{m}}(CD,\bar{m})=\lambda\left[p_{CD\bar{m}}(1+\bar{m})+p_{DD\bar{m}}(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]-\bar{k}$$
$$\Pi_{\bar{m}}(DD,\bar{m})=\lambda p_{CD\bar{m}}(1+\alpha)-\bar{k}$$
$$\Pi_{\bar{m}}(DD,\underline{m})=0$$
$$\Pi_{\underline{m}}(\underline{m})=0$$
$$\Pi_{\underline{m}}(\bar{m})=\lambda p_{CD\bar{m}}(1+\alpha)-\underline{k}$$

Hence we obtain the following phase diagram:



Since the phase diagram is ambiguous with respect to stability we will study the Eigenvalues of the linearized system.

	$\begin{matrix} p & q & \\ CD,\bar{m} & DD,\bar{m} & DD,\underline{m} \\ \varepsilon & & -\varepsilon \end{matrix}$: $\Delta\Pi_{\bar{m}}\left(CD,\bar{m}\right)=\lambda\left(1+\bar{m}\right)\varepsilon$ $\Delta\Pi_{\bar{m}}\left(DD,\bar{m}\right)=\lambda\left(1+\alpha\right)\varepsilon$ $\Delta\Pi_{\bar{m}}\left(DD,\underline{m}\right)=0$ hence $\mathrm{p}_{CD,\bar{m}}\uparrow$ and $\mathrm{p}_{DD,\underline{m}}\downarrow\Rightarrow f_p^p>0$ $\begin{matrix} p & q & \\ CD,\bar{m} & DD,\bar{m} & DD,\underline{m} \\ & \varepsilon & -\varepsilon \end{matrix}$: $\Delta\Pi_{\bar{m}}\left(CD,\bar{m}\right)=-\lambda\left(\beta\right)\varepsilon$ $\Delta\Pi_{\bar{m}}\left(DD,\bar{m}\right)=0$ $\Delta\Pi_{\bar{m}}\left(DD,\underline{m}\right)=0$ hence $\mathrm{p}_{CD,\bar{m}}\downarrow \overset{\text{payoff monotonicity}}{\underset{\Pi_{\bar{m}}(DD,\bar{m})=\Pi_{\bar{m}}(DD,\underline{m})\Leftrightarrow\dot{p}_{DD,\bar{m}}=\dot{p}_{DD,\underline{m}}}{\Rightarrow}} \mathrm{p}_{DD,\bar{m}}=\mathrm{p}_{DD,\underline{m}}\uparrow\Rightarrow f_q^q>0$ hence $\begin{pmatrix} f_p^p & f_q^p \\ f_p^q & f_q^q \end{pmatrix}=\begin{pmatrix} >0 & \\ & >0 \end{pmatrix}$ Hence at least one of the Eigenvalues is strictly positive and therefore this equilibrium is unstable.

4.7.1.1	$CD\bar{m}$ $\bar{m}$ $\underline{m}$	$p_{CD\bar{m}}=1$	$\lambda=\dfrac{\underline{k}}{\left(1+\alpha\right)}$	1. $\underline{k}<\left(1+\alpha\right)$ 2. $p_{\bar{m}}<\dfrac{\lambda\left(\bar{m}-\alpha\right)}{\left(1-\lambda\right)\beta}$, note that 2. is only binding if: $\dfrac{\lambda\left(\bar{m}-\alpha\right)}{\left(1-\lambda\right)\beta}<1\Leftrightarrow\lambda<\dfrac{\beta}{\left(\bar{m}-\alpha+\beta\right)}\Leftrightarrow\dfrac{\underline{k}}{\left(1+\alpha\right)}<\dfrac{\beta}{\left(\bar{m}-\alpha+\beta\right)}$

$\Pi_{\bar{m}}\left(CD,\bar{m}\right)=\lambda p_{CD\bar{m}}\left(1+\bar{m}\right)+\left(1-\lambda\right)\left[p_{\bar{m}}\left(-\beta\right)\right]-\bar{k}$

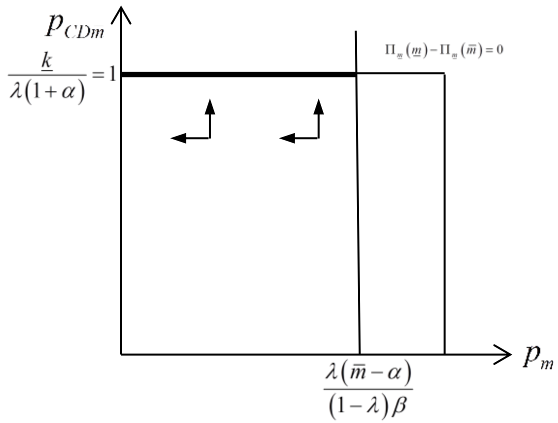
$\Pi_{\underline{m}}\left(\underline{m}\right)=0$

$\Pi_{\underline{m}}\left(\bar{m}\right)=\lambda p_{CD\bar{m}}\left(1+\alpha\right)-\underline{k}$

Hence we obtain the following phase diagram:

$\Pi_{\bar{m}}\left(CD,\bar{m}\right)-\Pi_{\underline{m}}\left(\underline{m}\right)=\dfrac{\underline{k}}{\left(1+\alpha\right)}\left(1+\bar{m}\right)-\bar{k}+\left(1-\lambda\right)\left[p_{\bar{m}}\left(-\beta\right)\right]$

$\left(\Pi_{\bar{m}}\left(CD,\bar{m}\right)-\Pi_{\underline{m}}\left(\underline{m}\right)\right)^{f}=\dfrac{\underline{k}}{\left(1+\alpha\right)}-\bar{k}+\left(1-\lambda\right)\left[p_{\bar{m}}\left(-\beta\right)\right]$



As the diagram clearly indicates this equilibrium set is stable.

Any perturbation induces a drift towards the separating equilibrium $p_{CD\bar{m}}=1, p_{\bar{m}}=0$.

4.7.1.2.	$CD\bar{m}$ $DD\bar{m}$ $\bar{m}$ $\underline{m}$	$p_{CD\bar{m}} = \frac{\underline{k}}{\lambda(1+\alpha)}$ $p_{DD\bar{m}} = 1 - \frac{\underline{k}}{\lambda(1+\alpha)}$ $p_{\bar{m}} = \frac{(\bar{m}-\alpha+\beta)\underline{k} - \lambda(1+\alpha)\beta}{(1-\lambda)(1+\alpha)\beta}$ $p_{\underline{m}} = \frac{(1+\alpha)\beta - (\bar{m}-\alpha+\beta)\underline{k}}{(1-\lambda)(1+\alpha)\beta}$	$\frac{\underline{k}}{(1+\alpha)} < \lambda < \frac{(\bar{m}-\alpha+\beta)}{\beta}\frac{\underline{k}}{(1+\alpha)}$	1. $(1+\alpha)\beta > (\bar{m}-\alpha+\beta)\underline{k}$ 2. $\underline{k} < (1+\alpha)$

$\Pi_{\bar{m}}(CD,\bar{m}) = \lambda\left[p_{CD\bar{m}}(1+\bar{m}) + p_{DD\bar{m}}(-\beta)\right] + (1-\lambda)\left[p_{\bar{m}}(-\beta)\right] - \bar{k}$

$\Pi_{\bar{m}}(DD,\bar{m}) = \lambda p_{CD\bar{m}}(1+\alpha) - \bar{k}$

$\Pi_{\underline{m}}(\underline{m}) = 0$

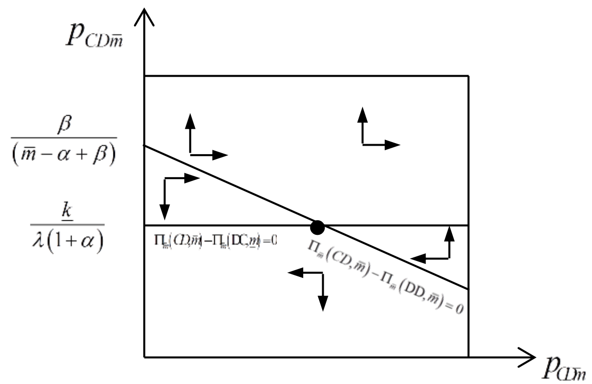
$\Pi_{\underline{m}}(\bar{m}) = \lambda p_{CD\bar{m}}(1+\alpha) - \underline{k}$

$\Pi_{\bar{m}}(CD,\bar{m}) - \Pi_{\bar{m}}(DD,\bar{m}) \geq 0 \Leftrightarrow \lambda\left[p_{CD\bar{m}}(1+\bar{m}) + p_{DD\bar{m}}(-\beta)\right] + (1-\lambda)\left[p_{\bar{m}}(-\beta)\right] - \bar{k} - \lambda p_{CD\bar{m}}(1+\alpha) + \bar{k} \geq 0 \Leftrightarrow \lambda\left[p_{CD\bar{m}}(\bar{m}-\alpha) + p_{DD\bar{m}}(-\beta)\right] + (1-\lambda)\left[p_{\bar{m}}(-\beta)\right] \geq 0 \Leftrightarrow$

$$p_{CD\bar{m}} \geq \frac{\lambda\beta - (1-\lambda)\beta p_{\bar{m}}}{\lambda(\bar{m}-\alpha+\beta)}$$

$\Pi_{\underline{m}}(\underline{m}) - \Pi_{\underline{m}}(\bar{m}) \geq 0 \Leftrightarrow p_{CD\bar{m}} \leq \frac{\underline{k}}{\lambda(1+\alpha)}$

Hence we obtain the following phase diagram:



As the diagram clearly indicates this equilibrium is unstable.

B.4 Derivation of p-equilibria

SEPARATING-EQUILIBRIA

1. High types signal, low types don't (**HSE** - high separating equilibrium)

$$p_{XY\underline{m}} = 0, p_{\underline{m}} = 1 \text{ (only high types signal) } 0 < \lambda < 1, \bar{m} > \alpha$$

high types:

$$\Pi_{\bar{m}}(CC,\bar{m}) = \lambda\left[(p_{CC\bar{m}} + p_{CD\bar{m}})(1+\bar{m}) + (0)(1+\alpha) + (0)(0) + (p_{DC\bar{m}} + p_{DD\bar{m}})(-\beta)\right] + (1-\lambda)(-\beta) - \bar{k} <$$

$$\Pi_{\bar{m}}(CD,\bar{m}) = \lambda\left[(p_{CC\bar{m}} + p_{CD\bar{m}})(1+\bar{m}) + (0)(1+\alpha) + (0)(0) + (p_{DC\bar{m}} + p_{DD\bar{m}})(-\beta)\right] + (1-\lambda)0 - \bar{k}$$

$$\Pi_{\bar{m}}(DC,\bar{m}) = \lambda\left[(0)(1+\bar{m}) + (p_{CC\bar{m}} + p_{CD\bar{m}})(1+\alpha) + (p_{DC\bar{m}} + p_{DD\bar{m}})(0) + (0)(-\beta)\right] + (1-\lambda)(-\beta) - \bar{k} <$$

$$\Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[(0)(1+\bar{m}) + (p_{CC\bar{m}} + p_{CD\bar{m}})(1+\alpha) + (p_{DC\bar{m}} + p_{DD\bar{m}})(0) + (0)(-\beta)\right] + (1-\lambda)0 - \bar{k}$$

$$\Pi_{\bar{m}}(CC,\underline{m}) = \lambda\left[(p_{CC\bar{m}} + p_{DC\bar{m}})(1+\bar{m}) + (0)(1+\alpha) + (0)(0) + (p_{CD\bar{m}} + p_{DD\bar{m}})(-\beta)\right] + (1-\lambda)(-\beta) <$$

$$\Pi_{\bar{m}}(CD,\underline{m}) = \lambda\left[(p_{CC\bar{m}} + p_{DC\bar{m}})(1+\bar{m}) + (0)(1+\alpha) + (0)(0) + (p_{CD\bar{m}} + p_{DD\bar{m}})(-\beta)\right] + (1-\lambda)0$$

$$\Pi_{\bar{m}}(DC,\underline{m}) = \lambda\left[(0)(1+\bar{m}) + (p_{CC\bar{m}} + p_{DC\bar{m}})(1+\alpha) + (p_{CD\bar{m}} + p_{DD\bar{m}})(0) + (0)(-\beta)\right] + (1-\lambda)(-\beta) <$$

$$\Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[(0)(1+\bar{m}) + (p_{CC\bar{m}} + p_{DC\bar{m}})(1+\alpha) + (p_{CD\bar{m}} + p_{DD\bar{m}})(0) + (0)(-\beta)\right] + (1-\lambda)0 \qquad \Rightarrow$$

$$\Pi_{\bar{m}}(CD,\bar{m}) = \lambda\left[(p_{CD\bar{m}})(1+\bar{m}) + (0)(1+\alpha) + (0)(0) + (p_{DD\bar{m}})(-\beta)\right] + (1-\lambda)0 - \bar{k} = \lambda\left[(p_{CD\bar{m}})(1+\bar{m}) + (p_{DD\bar{m}})(-\beta)\right] - \bar{k}$$

$$\Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[(0)(1+\bar{m}) + (p_{CD\bar{m}})(1+\alpha) + (p_{DD\bar{m}})(0) + (0)(-\beta)\right] + (1-\lambda)0 - \bar{k} = \lambda\left[(p_{CD\bar{m}})(1+\alpha)\right] - \bar{k}$$

$$\Pi_{\bar{m}}(CD,\underline{m}) = \lambda\left[(0)(1+\bar{m}) + (0)(1+\alpha) + (0)(0) + (p_{CD\bar{m}} + p_{DD\bar{m}})(-\beta)\right] + (1-\lambda)0 = \lambda\left[(p_{CD\bar{m}} + p_{DD\bar{m}})(-\beta)\right]$$

$$\Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[(0)(1+\bar{m}) + (0)(1+\alpha) + (p_{CD\bar{m}} + p_{DD\bar{m}})(0) + (0)(-\beta)\right] + (1-\lambda)0 = 0$$

low types: $\Pi_{\underline{m}}(\bar{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(1+\alpha) + \left(p_{DC\bar{m}} + p_{DD\bar{m}}\right)0\right] + (1-\lambda)0 - \underline{k}$ $\Pi_{\underline{m}}(\underline{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(1+\alpha) + \left(p_{CD\bar{m}} + p_{DD\bar{m}}\right)0\right] + (1-\lambda)0$ $\overset{iteration}{\underset{hightypes}{\Rightarrow}}$ $\Pi_{\underline{m}}(\bar{m}) = \lambda\left[\left(p_{CD\bar{m}}\right)(1+\alpha) + \left(p_{DD\bar{m}}\right)0\right] - \underline{k} = \lambda\left[\left(p_{CD\bar{m}}\right)(1+\alpha)\right] - \underline{k}$ $\Pi_{\underline{m}}(\underline{m}) = 0$ $\Rightarrow$	high types: $\Pi_{\bar{m}}(CD,\bar{m}) = \lambda\left[\left(p_{CD\bar{m}}\right)(1+\bar{m}) + \left(p_{DD\bar{m}}\right)(-\beta)\right] - \bar{k}$ $\Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[\left(p_{CD\bar{m}}\right)(1+\alpha)\right] - \bar{k}$ $\Pi_{\bar{m}}(CD,\underline{m}) = \lambda\left[\left(p_{CD\bar{m}} + p_{DD\bar{m}}\right)(-\beta)\right] = -\lambda\beta <$ $\Pi_{\bar{m}}(DD,\underline{m}) = 0$ low types: $\Pi_{\underline{m}}(\bar{m}) = \lambda\left[\left(p_{CD\bar{m}}\right)(1+\alpha)\right] - \underline{k}$ $\Pi_{\underline{m}}(\underline{m}) = 0$

For a separating equilibrium where high types send the signal and low types don't there are only two not dominated strategies left, $CD,\bar{m}$ and $DD,\bar{m}$, i.e. $p_{CD\bar{m}} + p_{DD\bar{m}} = 1$, hence $CD,\underline{m}$ is strictly dominated by $DD,\underline{m}$ in such an equilibrium.

1.1. Let's first analyze the case $p_{CD\bar{m}} = 1,\ p_{\underline{m}} = 1$:

$CD_{\bar{m}} \succ DD_{\bar{m}}$, always satisfied (i)

$$CD_{\bar{m}} \succ DD_{\underline{m}} \Leftrightarrow \lambda \geq \frac{\bar{k}}{1+\bar{m}} \qquad \text{(ii)}$$

$$\underline{m} \succ \bar{m} \Leftrightarrow \lambda \leq \frac{\underline{k}}{1+\alpha}. \qquad \text{(iii)}$$

Note that since $\underline{k} > \bar{k},\ \bar{m} > \alpha$ the lambda support for this equilibrium is not empty. Intuitively there need to be sufficiently many CD players such that it is worthwhile to signal for high types, but not too many to deter low types from also signaling.

1.2. Let's now analyze the case $p_{DD\bar{m}} = 1,\ p_{\underline{m}} = 1$:

$DD,\underline{m}$ strictly dominates all other strategies, hence such an equilibrium cannot exist.

1.3. Finally, $p_{CD\bar{m}} + p_{DD\bar{m}} = 1$ such that both strategies are played:

$$\Pi_{\bar{m}}(CD,\bar{m}) = \lambda\left[\left(p_{CD\bar{m}}\right)(1+\bar{m}) + \left(p_{DD\bar{m}}\right)(-\beta)\right] - \bar{k} = \lambda\left[\left(p_{CD\bar{m}}\right)(1+\alpha)\right] - \bar{k} = \Pi_{\bar{m}}(DD,\bar{m}) \overset{p_{DD\bar{m}}=1-p_{CD\bar{m}}}{\Leftrightarrow}$$

$$p_{CD\bar{m}} = \frac{\beta}{\bar{m}-\alpha+\beta} \quad \text{(i)}$$

$$\left(CD_{\bar{m}} \approx\right) DD_{\bar{m}} \succ DD_{\underline{m}} \Leftrightarrow p_{CD\bar{m}} \geq \frac{\lambda\beta+\bar{k}}{\lambda\left(1+\bar{m}+\beta\right)} \quad \text{(ii)}$$

$$\underline{m} \succ \bar{m} \Leftrightarrow p_{CD\bar{m}} \leq \frac{\underline{k}}{\lambda\left(1+\alpha\right)}. \quad \text{(iii)}$$

At $p_{CD\bar{m}} = \dfrac{\beta}{\bar{m}+\beta-\alpha}$: (ii) $\dfrac{\beta}{\bar{m}-\alpha+\beta} \geq \dfrac{\lambda\beta+\bar{k}}{\lambda\left(1+\bar{m}+\beta\right)} \Leftrightarrow \lambda \geq \dfrac{\left(\bar{m}-\alpha+\beta\right)}{\beta\left(1+\alpha\right)}\bar{k}$ (ii)’

(iii) $\dfrac{\beta}{\bar{m}+\beta-\alpha} \leq \dfrac{\underline{k}}{\lambda\left(1+\alpha\right)} \Leftrightarrow \lambda \leq \dfrac{\left(\bar{m}-\alpha+\beta\right)}{\beta\left(1+\alpha\right)}\underline{k}$ (iii)‘

To summarize:

equilibrium	Lambda-support	Conditions for Existence
$p_{CD\bar{m}} = 1,\ p_{\underline{m}} = 1$	$\dfrac{\bar{k}}{1+\bar{m}} \leq \lambda \leq \dfrac{\underline{k}}{1+\alpha}$	$\bar{k} < 1+\bar{m}$
$p_{CD\bar{m}} = \dfrac{\beta}{\bar{m}+\beta-\alpha},\ p_{DD\bar{m}} = \dfrac{\bar{m}-\alpha}{\bar{m}+\beta-\alpha},\ p_{\underline{m}} = 1$	$\dfrac{\left(\bar{m}-\alpha+\beta\right)}{\beta\left(1+\alpha\right)}\bar{k} \leq \lambda \leq \dfrac{\left(\bar{m}-\alpha+\beta\right)}{\beta\left(1+\alpha\right)}\underline{k}$	$\bar{k} < \dfrac{\beta\left(1+\alpha\right)}{\left(\bar{m}-\alpha+\beta\right)}$

2. <u>High types don't signal, low types signal (LSE – low separating equilibrium</u>)

$p_{XY\bar{m}} = 0, p_{\bar{m}} = 1$ (only low types signal) $0 < \lambda < 1, \bar{m} > \alpha$

$$\Pi_{\bar{m}}(CC,\bar{m}) = \lambda\Big[\big(p_{CC\underline{m}} + p_{CD\underline{m}}\big)(1+\bar{m}) + (0)(1+\alpha) + (0)(0) + \big(p_{DC\underline{m}} + p_{DD\underline{m}}\big)(-\beta)\Big] + (1-\lambda)(-\beta) - \bar{k} < \Pi_{\bar{m}}(DC,\bar{m})$$

$$\Pi_{\bar{m}}(CD,\bar{m}) = \lambda\Big[(0)(1+\bar{m}) + \big(p_{CC\underline{m}} + p_{CD\underline{m}}\big)(1+\alpha) + \big(p_{DC\underline{m}} + p_{DD\underline{m}}\big)(0) + (0)(-\beta)\Big] + (1-\lambda)(-\beta) - \bar{k} < \Pi_{\bar{m}}(DD,\bar{m})$$

$$\Pi_{\bar{m}}(DC,\bar{m}) = \lambda\Big[\big(p_{CC\underline{m}} + p_{CD\underline{m}}\big)(1+\bar{m}) + (0)(1+\alpha) + (0)(0) + \big(p_{DC\underline{m}} + p_{DD\underline{m}}\big)(-\beta)\Big] + (1-\lambda)0 - \bar{k}$$

$$\Pi_{\bar{m}}(DD,\bar{m}) = \lambda\Big[(0)(1+\bar{m}) + \big(p_{CC\underline{m}} + p_{CD\underline{m}}\big)(1+\alpha) + \big(p_{DC\underline{m}} + p_{DD\underline{m}}\big)(0) + (0)(-\beta)\Big] + (1-\lambda)0 - \bar{k}$$

$$\Pi_{\bar{m}}(CC,\underline{m}) = \lambda\Big[\big(p_{CC\underline{m}} + p_{DC\underline{m}}\big)(1+\bar{m}) + (0)(1+\alpha) + (0)(0) + \big(p_{CD\underline{m}} + p_{DD\underline{m}}\big)(-\beta)\Big] + (1-\lambda)(-\beta) < \Pi_{\bar{m}}(DC,\underline{m})$$

$$\Pi_{\bar{m}}(CD,\underline{m}) = \lambda\Big[(0)(1+\bar{m}) + \big(p_{CC\underline{m}} + p_{DC\underline{m}}\big)(1+\alpha) + \big(p_{CD\underline{m}} + p_{DD\underline{m}}\big)(0) + (0)(-\beta)\Big] + (1-\lambda)(-\beta) < \Pi_{\bar{m}}(DD,\underline{m})$$

$$\Pi_{\bar{m}}(DC,\underline{m}) = \lambda\Big[\big(p_{CC\underline{m}} + p_{DC\underline{m}}\big)(1+\bar{m}) + (0)(1+\alpha) + (0)(0) + \big(p_{CD\underline{m}} + p_{DD\underline{m}}\big)(-\beta)\Big] + (1-\lambda)0$$

$$\Pi_{\bar{m}}(DD,\underline{m}) = \lambda\Big[(0)(1+\bar{m}) + \big(p_{CC\underline{m}} + p_{DC\underline{m}}\big)(1+\alpha) + \big(p_{CD\underline{m}} + p_{DD\underline{m}}\big)(0) + (0)(-\beta)\Big] + (1-\lambda)0$$

$$\Pi_{\underline{m}}(\bar{m}) = \lambda\Big[\big(p_{CC\underline{m}} + p_{CD\underline{m}}\big)(1+\alpha) + \big(p_{DC\underline{m}} + p_{DD\underline{m}}\big)0\Big] + (1-\lambda)0 - \underline{k}$$

$$\Pi_{\underline{m}}(\underline{m}) = \lambda\Big[\big(p_{CC\underline{m}} + p_{DC\underline{m}}\big)(1+\alpha) + \big(p_{CD\underline{m}} + p_{DD\underline{m}}\big)0\Big] + (1-\lambda)0 \qquad \Rightarrow$$

$$\Pi_{\underline{m}}(\bar{m}) = \lambda\Big[0(1+\alpha) + \big(p_{DC\underline{m}} + p_{DD\underline{m}}\big)0\Big] + (1-\lambda)0 - \underline{k} = -\underline{k} <$$

$$\Pi_{\underline{m}}(\underline{m}) = \lambda\Big[\big(p_{DC\underline{m}}\big)(1+\alpha) + \big(p_{DD\underline{m}}\big)0\Big] + (1-\lambda)0 = \lambda\Big[\big(p_{DC\underline{m}}\big)(1+\alpha)\Big]$$

$\bar{m}$ is strictly dominated for low types implying, that such an equilibrium cannot exist.

POOLING_EQUILIBRIA

1. High types and low types don't signal (NSPE – no-signal pooling equilibrium)

$$p_{XY\bar{m}} = 0, p_{\underline{m}} = 1 \text{ (all signal } \underline{m}\text{, i.e. nobody signals) } 0<\lambda<1, \bar{m}>\alpha$$

$$\Pi_{\bar{m}}(CC,\bar{m}) = \lambda\left[\left(p_{CC\underline{m}} + p_{CD\underline{m}}\right)(1+\bar{m}) + (0)(1+\alpha) + (0)(0) + \left(p_{DC\underline{m}} + p_{DD\underline{m}}\right)(-\beta)\right] + (1-\lambda)(-\beta) - \bar{k} = \lambda\left[\left(p_{CC\underline{m}} + p_{CD\underline{m}}\right)(1+\bar{m}) + \left(p_{DC\underline{m}} + p_{DD\underline{m}}\right)(-\beta)\right] + (1-\lambda)(-\beta) - \bar{k}$$

$$= \lambda\left[\left(p_{CC\underline{m}} + p_{CD\underline{m}}\right)(1+\bar{m}+\beta)\right] - \beta - \bar{k}$$

$$\Pi_{\bar{m}}(CD,\bar{m}) = \lambda\left[(0)(1+\bar{m}) + \left(p_{CC\underline{m}} + p_{CD\underline{m}}\right)(1+\alpha) + \left(p_{DC\underline{m}} + p_{DD\underline{m}}\right)(0) + (0)(-\beta)\right] + (1-\lambda)0 - \bar{k} = \lambda\left[\left(p_{CC\underline{m}} + p_{CD\underline{m}}\right)(1+\alpha)\right] - \bar{k}$$

$$\Pi_{\bar{m}}(DC,\bar{m}) = \lambda\left[\left(p_{CC\underline{m}} + p_{CD\underline{m}}\right)(1+\bar{m}) + (0)(1+\alpha) + (0)(0) + \left(p_{DC\underline{m}} + p_{DD\underline{m}}\right)(-\beta)\right] + (1-\lambda)(-\beta) - \bar{k} = \lambda\left[\left(p_{CC\underline{m}} + p_{CD\underline{m}}\right)(1+\bar{m}) + \left(p_{DC\underline{m}} + p_{DD\underline{m}}\right)(-\beta)\right] + (1-\lambda)(-\beta) - \bar{k}$$

$$= \lambda\left[\left(p_{CC\underline{m}} + p_{CD\underline{m}}\right)(1+\bar{m}+\beta)\right] - \beta - \bar{k}$$

$$\Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[(0)(1+\bar{m}) + \left(p_{CC\underline{m}} + p_{CD\underline{m}}\right)(1+\alpha) + \left(p_{DC\underline{m}} + p_{DD\underline{m}}\right)(0) + (0)(-\beta)\right] + (1-\lambda)0 - \bar{k} = \lambda\left[\left(p_{CC\underline{m}} + p_{CD\underline{m}}\right)(1+\alpha)\right] - \bar{k}$$

$$\Pi_{\bar{m}}(CC,\underline{m}) = \lambda\left[\left(p_{CC\underline{m}} + p_{DC\underline{m}}\right)(1+\bar{m}) + (0)(1+\alpha) + (0)(0) + \left(p_{CD\underline{m}} + p_{DD\underline{m}}\right)(-\beta)\right] + (1-\lambda)(-\beta) = \lambda\left[\left(p_{CC\underline{m}} + p_{DC\underline{m}}\right)(1+\bar{m}) + \left(p_{CD\underline{m}} + p_{DD\underline{m}}\right)(-\beta)\right] + (1-\lambda)(-\beta)$$

$$= \lambda\left[\left(p_{CC\underline{m}} + p_{DC\underline{m}}\right)(1+\bar{m}+\beta)\right] - \beta$$

$$\Pi_{\bar{m}}(CD,\underline{m}) = \lambda\left[(0)(1+\bar{m}) + \left(p_{CC\underline{m}} + p_{DC\underline{m}}\right)(1+\alpha) + \left(p_{CD\underline{m}} + p_{DD\underline{m}}\right)(0) + (0)(-\beta)\right] + (1-\lambda)0 = \lambda\left[\left(p_{CC\underline{m}} + p_{DC\underline{m}}\right)(1+\alpha)\right]$$

$$\Pi_{\bar{m}}(DC,\underline{m}) = \lambda\left[\left(p_{CC\underline{m}} + p_{DC\underline{m}}\right)(1+\bar{m}) + (0)(1+\alpha) + (0)(0) + \left(p_{CD\underline{m}} + p_{DD\underline{m}}\right)(-\beta)\right] + (1-\lambda)(-\beta) = \lambda\left[\left(p_{CC\underline{m}} + p_{DC\underline{m}}\right)(1+\bar{m}) + \left(p_{CD\underline{m}} + p_{DD\underline{m}}\right)(-\beta)\right] + (1-\lambda)(-\beta)$$

$$= \lambda\left[\left(p_{CC\underline{m}} + p_{DC\underline{m}}\right)(1+\bar{m}+\beta)\right] - \beta$$

$$\Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[(0)(1+\bar{m}) + \left(p_{CC\underline{m}} + p_{DC\underline{m}}\right)(1+\alpha) + \left(p_{CD\underline{m}} + p_{DD\underline{m}}\right)(0) + (0)(-\beta)\right] + (1-\lambda)0 = \lambda\left[\left(p_{CC\underline{m}} + p_{DC\underline{m}}\right)(1+\alpha)\right]$$

$$\Pi_{\underline{m}}(\bar{m}) = \lambda\left[\left(p_{CC\underline{m}} + p_{CD\underline{m}}\right)(1+\alpha) + \left(p_{DC\underline{m}} + p_{DD\underline{m}}\right)0\right] + (1-\lambda)0 - \underline{k} = \lambda\left[\left(p_{CC\underline{m}} + p_{CD\underline{m}}\right)(1+\alpha)\right] - \underline{k}$$

$$\Pi_{\underline{m}}(\underline{m}) = \lambda\left[\left(p_{CC\underline{m}} + p_{DC\underline{m}}\right)(1+\alpha) + \left(p_{CD\underline{m}} + p_{DD\underline{m}}\right)0\right] + (1-\lambda)0 = \lambda\left[\left(p_{CC\underline{m}} + p_{DC\underline{m}}\right)(1+\alpha)\right] \Rightarrow$$

$\Pi_{\bar{m}}(CC,\bar{m})=\lambda\left[\left(p_{CC\underline{m}}+p_{CD\underline{m}}\right)(1+\bar{m}+\beta)\right]-\beta-\bar{k}$ $\Pi_{\bar{m}}(CD,\bar{m})=\lambda\left[\left(p_{CC\underline{m}}+p_{CD\underline{m}}\right)(1+\alpha)\right]-\bar{k}$ $\Pi_{\bar{m}}(DC,\bar{m})=\lambda\left[\left(p_{CC\underline{m}}+p_{CD\underline{m}}\right)(1+\bar{m}+\beta)\right]-\beta-\bar{k}$ $\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left[\left(p_{CC\underline{m}}+p_{CD\underline{m}}\right)(1+\alpha)\right]-\bar{k}$	$\Pi_{\bar{m}}(CC,\underline{m})=\lambda\left[\left(p_{CC\underline{m}}+p_{DC\underline{m}}\right)(1+\bar{m}+\beta)\right]-\beta$ $\Pi_{\bar{m}}(CD,\underline{m})=\lambda\left[\left(p_{CC\underline{m}}+p_{DC\underline{m}}\right)(1+\alpha)\right]$ $\Pi_{\bar{m}}(DC,\underline{m})=\lambda\left[\left(p_{CC\underline{m}}+p_{DC\underline{m}}\right)(1+\bar{m}+\beta)\right]-\beta$ $\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[\left(p_{CC\underline{m}}+p_{DC\underline{m}}\right)(1+\alpha)\right]$	$\Pi_{\underline{m}}(\bar{m})=\lambda\left[\left(p_{CC\underline{m}}+p_{CD\underline{m}}\right)(1+\alpha)\right]-\underline{k}$ $\Pi_{\underline{m}}(\underline{m})=\lambda\left[\left(p_{CC\underline{m}}+p_{DC\underline{m}}\right)(1+\alpha)\right]$

Note that in a pooling equilibrium where nobody sends the signal, CC and DC (CD and DD) will always earn the same profits irrespective of the chosen signal and the particular composition. Since those pairs are indistinguishable we only have to consider the following cases:

1.1. Consider first the case $p_{CC\underline{m}}+p_{DC\underline{m}}=1,\left(p_{CD\underline{m}}+p_{DD\underline{m}}=0\right)$:

$$CC_{\underline{m}}/DC_{\underline{m}}\overset{s}{\succ}CC_{\bar{m}}/DC_{\bar{m}},\ CD_{\underline{m}}/DD_{\underline{m}}\overset{s}{\succ}CD_{\bar{m}}/DD_{\bar{m}}\ \text{and}\ \underline{m}\overset{s}{\succ}\bar{m}\ \text{since}\ p_{CD\underline{m}}=0$$

$CC_{\underline{m}}/DC_{\underline{m}}\succ CD_{\underline{m}}/DD_{\underline{m}}\Leftrightarrow\lambda\geq\dfrac{\beta}{(\bar{m}-\alpha+\beta)}$, since $CC_{\underline{m}}/DC_{\underline{m}}\succ CD_{\underline{m}}/DD_{\underline{m}}\succ CD_{\bar{m}}/DD_{\bar{m}}$, $\lambda\geq\dfrac{\beta}{(\beta+\bar{m}-\alpha)}$ is necessary and sufficient.

1.2. Consider now the case $p_{CD\underline{m}}+p_{DD\underline{m}}=1,\left(p_{CC\underline{m}}+p_{DC\underline{m}}=0\right)$:

$CC_{\underline{m}}/DC_{\underline{m}}\prec CD_{\underline{m}}/DD_{\underline{m}}$, since $p_{CC\underline{m}}+p_{CD\underline{m}}=0$,

$$CD_{\underline{m}}/DD_{\underline{m}}\succ CD_{\bar{m}}/DD_{\bar{m}}\Leftrightarrow\lambda\left(p_{CD\underline{m}}\right)\leq\frac{\bar{k}}{1+\alpha} \quad \text{(i)}$$

$$CD_{\underline{m}}/DD_{\underline{m}}\succ CC_{\bar{m}}/DC_{\bar{m}}\Leftrightarrow\lambda\left(p_{CD\underline{m}}\right)\leq\frac{\beta+\bar{k}}{1+\bar{m}+\beta} \quad \text{(ii)}$$

$$\underline{m} \succ \bar{m} \Leftrightarrow \lambda\left(p_{CD\underline{m}}\right) \leq \frac{\underline{k}}{1+\alpha}. \quad \text{(iii)}$$

Note that $\frac{\underline{k}}{1+\alpha} > \frac{\bar{k}}{1+\alpha} > \frac{\bar{k}}{1+\bar{m}+\beta}$, i.e. (i) implies (iii), hence necessary and sufficient is:

$$\lambda\left(p_{CD\underline{m}}\right) \leq \min\left\{\frac{\bar{k}+\beta}{1+\bar{m}+\beta}, \frac{\bar{k}}{1+\alpha}\right\} \Leftrightarrow p_{CD\underline{m}} \leq \frac{1}{\lambda}\min\left\{\frac{\bar{k}+\beta}{1+\bar{m}+\beta}, \frac{\bar{k}}{1+\alpha}\right\} = \frac{1}{\lambda}\begin{cases} \frac{\bar{k}+\beta}{1+\bar{m}+\beta}, & (1+\alpha)\beta < (\bar{m}-\alpha+\beta)\bar{k} \\ \frac{\bar{k}}{1+\alpha}, & (1+\alpha)\beta > (\bar{m}-\alpha+\beta)\bar{k} \end{cases} \Leftrightarrow$$

$$\lambda\left(p_{CD\underline{m}}\right) \leq \min\left\{\frac{\bar{k}+\beta}{1+\bar{m}+\beta}, \frac{\bar{k}}{1+\alpha}\right\} \Leftrightarrow \lambda \leq \frac{1}{p_{CD\underline{m}}}\min\left\{\frac{\bar{k}+\beta}{1+\bar{m}+\beta}, \frac{\bar{k}}{1+\alpha}\right\} = \frac{1}{p_{CD\underline{m}}}\begin{cases} \frac{\bar{k}+\beta}{1+\bar{m}+\beta}, & (1+\alpha)\beta < (\bar{m}-\alpha+\beta)\bar{k} \\ \frac{\bar{k}}{1+\alpha}, & (1+\alpha)\beta > (\bar{m}-\alpha+\beta)\bar{k} \end{cases}$$

For this constraint to be binding we need to have: $\lambda > \begin{cases} \frac{\bar{k}+\beta}{1+\bar{m}+\beta}, & (1+\alpha)\beta < (\bar{m}-\alpha+\beta)\bar{k} \\ \frac{\bar{k}}{1+\alpha}, & (1+\alpha)\beta > (\bar{m}-\alpha+\beta)\bar{k} \end{cases}$

1.3. Consider finally $p_{CD\underline{m}} + p_{DD\underline{m}} + p_{CC\underline{m}} + p_{DC\underline{m}} = 1$:

i.e. all no-signaling strategies earn the same payoff:

$$\lambda\left(p_{CC\underline{m}} + p_{DC\underline{m}}\right)\left(1+\bar{m}+\beta\right) - \beta = \lambda\left(p_{CC\underline{m}} + p_{DC\underline{m}}\right)\left(1+\alpha\right) \Leftrightarrow \lambda\left(p_{CC\underline{m}} + p_{DC\underline{m}}\right) = \frac{\beta}{\left(\bar{m}-\alpha+\beta\right)}$$

In this case necessary and sufficient is, that any of the no-signaling strategies is better than any signaling strategy, necessary and sufficient for this is:

$$CC_{\underline{m}} / DC_{\underline{m}} \succ CC_{\bar{m}} / DC_{\bar{m}} \Leftrightarrow \lambda\left(p_{CD\underline{m}} - p_{DC\underline{m}}\right) \leq \frac{\bar{k}}{1+\bar{m}+\beta}$$

$$CD_{\underline{m}} / DD_{\underline{m}} \succ CD_{\bar{m}} / DD_{\bar{m}} \Leftrightarrow \lambda\left(p_{CD\underline{m}} - p_{DC\underline{m}}\right) \leq \frac{\bar{k}}{1+\alpha}$$

$$\underline{m} \succ \bar{m} \Leftrightarrow \lambda\left(p_{CD\underline{m}} - p_{DC\underline{m}}\right) \leq \frac{\underline{k}}{1+\alpha}$$

Maximal support: $\lambda\left(p_{CC\underline{m}} + p_{DC\underline{m}}\right) = \frac{\beta}{\left(\bar{m}-\alpha+\beta\right)}$ under the constraint $\lambda\left(p_{CD\underline{m}} - p_{DC\underline{m}}\right) \leq \frac{\bar{k}}{1+\bar{m}+\beta}$, note that the maximal support is given if $p_{CC\underline{m}} + p_{DC\underline{m}}$ has maximal support, i.e. $p_{CC\underline{m}} + p_{DC\underline{m}} \in (0,1)$, consider $p_{DD\underline{m}} = 1-\left(p_{CC\underline{m}} + p_{DC\underline{m}}\right)$ and $p_{CD\underline{m}} = 0$, then the constraint will always be satisfied and we get the support $\frac{\beta}{\left(\bar{m}-\alpha+\beta\right)} < \lambda < 1$

To summarize:

Equilibrium	Lambda-support
$p_{CC\underline{m}} + p_{DC\underline{m}} = 1, \left(p_{CD\underline{m}} + p_{DD\underline{m}} = 0\right), p_{\underline{m}} = 1$	$\lambda \geq \frac{\beta}{\left(\beta+\bar{m}-\alpha\right)}$
$p_{CD\underline{m}} + p_{DD\underline{m}} = 1, \left(p_{CC\underline{m}} + p_{DC\underline{m}} = 0\right), p_{\underline{m}} = 1$	$\lambda\left(p_{CD\underline{m}}\right) \leq \min\left\{\frac{\bar{k}+\beta}{1+\bar{m}+\beta}, \frac{\bar{k}}{1+\alpha}\right\}$
$p_{CD\underline{m}} + p_{DD\underline{m}} + p_{CC\underline{m}} + p_{DC\underline{m}} = 1, p_{\underline{m}} = 1$	$\lambda\left(p_{CC\underline{m}} + p_{DC\underline{m}}\right) = \frac{\beta}{\left(\bar{m}-\alpha+\beta\right)}$, $\lambda\left(p_{CD\underline{m}} - p_{DC\underline{m}}\right) \leq \frac{\bar{k}}{1+\bar{m}+\beta}$

Lambda-support	Equilibrium
$\lambda = \frac{\beta}{\left(\beta + \bar{m} - \alpha\right)}$	$\left\{ p_{CC\underline{m}}, p_{DC\underline{m}}, p_{\underline{m}} \middle| p_{CC\underline{m}} + p_{DC\underline{m}} = 1, \left(p_{CD\underline{m}} + p_{DD\underline{m}} = 0 \right), p_{\underline{m}} = 1 \right\}$
$\lambda > \frac{\beta}{\left(\beta + \bar{m} - \alpha\right)}$	1. $\left\{ \begin{array}{l} p_{CD\underline{m}}, p_{DD\underline{m}}, p_{CC\underline{m}}, p_{DC\underline{m}}, p_{\underline{m}} \middle| p_{CD\underline{m}} + p_{DD\underline{m}} + p_{CC\underline{m}} + p_{DC\underline{m}} = 1, \\ \left(p_{CC\underline{m}} + p_{DC\underline{m}} \right) = \frac{\beta}{\lambda\left(\bar{m} - \alpha + \beta\right)}, \left(p_{CD\underline{m}} - p_{DC\underline{m}} \right) \leq \frac{\bar{k}}{\lambda\left(1 + \bar{m} + \beta\right)}, p_{\underline{m}} = 1 \end{array} \right\}$ 2. $\left\{ p_{CC\underline{m}}, p_{DC\underline{m}}, p_{\underline{m}} \middle| p_{CC\underline{m}} + p_{DC\underline{m}} = 1, \left(p_{CD\underline{m}} + p_{DD\underline{m}} = 0 \right), p_{\underline{m}} = 1 \right\}$
$\lambda \in \left(0,1\right)$	$\left\{ p_{CD\underline{m}}, p_{DD\underline{m}}, p_{\underline{m}} \middle| p_{CD\underline{m}} + p_{DD\underline{m}} = 1, \left(p_{CC\underline{m}} + p_{DC\underline{m}} = 0 \right), p_{\underline{m}} = 1, p_{CD\underline{m}} \leq \frac{1}{\lambda} \min\left\{ \frac{\bar{k} + \beta}{1 + \bar{m} + \beta}, \frac{\bar{k}}{1 + \alpha} \right\} \right\}$, i.e. $p_{CD\underline{m}} \leq \frac{1}{\lambda} \min\left\{ \frac{\bar{k} + \beta}{1 + \bar{m} + \beta}, \frac{\bar{k}}{1 + \alpha} \right\}$ is only binding if $\lambda \geq \max\left\{ \frac{1 + \bar{m} + \beta}{\bar{k} + \beta}, \frac{1 + \alpha}{\bar{k}} \right\}$

2. <u>High types and low types signal (**SPE** – signal pooling equilibrium)</u>

$$p_{XY\underline{m}} = 0, p_{\bar{m}} = 1 \text{ (all signal } \bar{m}\text{)} \; 0 < \lambda < 1, \bar{m} > \alpha$$

$$\Pi_{\bar{m}}(CC,\bar{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(1+\bar{m}) + (0)(1+\alpha) + (0)(0) + \left(p_{DC\bar{m}} + p_{DD\bar{m}}\right)(-\beta)\right] + (1-\lambda)(-\beta) - \bar{k} = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(1+\bar{m}) + \left(p_{DC\bar{m}} + p_{DD\bar{m}}\right)(-\beta)\right] + (1-\lambda)(-\beta) - \bar{k}$$
$$= \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(1+\bar{m}+\beta)\right] - \beta - \bar{k}$$
$$\Pi_{\bar{m}}(CD,\bar{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(1+\bar{m}) + (0)(1+\alpha) + (0)(0) + \left(p_{DC\bar{m}} + p_{DD\bar{m}}\right)(-\beta)\right] + (1-\lambda)(-\beta) - \bar{k} = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(1+\bar{m}) + \left(p_{DC\bar{m}} + p_{DD\bar{m}}\right)(-\beta)\right] + (1-\lambda)(-\beta) - \bar{k}$$
$$= \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(1+\bar{m}+\beta)\right] - \beta - \bar{k}$$
$$\Pi_{\bar{m}}(DC,\bar{m}) = \lambda\left[(0)(1+\bar{m}) + \left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(1+\alpha) + \left(p_{DC\bar{m}} + p_{DD\bar{m}}\right)(0) + (0)(-\beta)\right] + (1-\lambda)0 - \bar{k} = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(1+\alpha)\right] - \bar{k}$$
$$\Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[(0)(1+\bar{m}) + \left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(1+\alpha) + \left(p_{DC\bar{m}} + p_{DD\bar{m}}\right)(0) + (0)(-\beta)\right] + (1-\lambda)0 - \bar{k} = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(1+\alpha)\right] - \bar{k}$$

$$\Pi_{\bar{m}}(CC,\underline{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(1+\bar{m}) + (0)(1+\alpha) + (0)(0) + \left(p_{CD\bar{m}} + p_{DD\bar{m}}\right)(-\beta)\right] + (1-\lambda)(-\beta) = \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(1+\bar{m}) + \left(p_{CD\bar{m}} + p_{DD\bar{m}}\right)(-\beta)\right] + (1-\lambda)(-\beta)$$
$$= \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(1+\bar{m}+\beta)\right] - \beta$$
$$\Pi_{\bar{m}}(CD,\underline{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(1+\bar{m}) + (0)(1+\alpha) + (0)(0) + \left(p_{CD\bar{m}} + p_{DD\bar{m}}\right)(-\beta)\right] + (1-\lambda)(-\beta) = \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(1+\bar{m}) + \left(p_{CD\bar{m}} + p_{DD\bar{m}}\right)(-\beta)\right] + (1-\lambda)(-\beta)$$
$$= \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(1+\bar{m}+\beta)\right] - \beta$$
$$\Pi_{\bar{m}}(DC,\underline{m}) = \lambda\left[(0)(1+\bar{m}) + \left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(1+\alpha) + \left(p_{CD\bar{m}} + p_{DD\bar{m}}\right)(0) + (0)(-\beta)\right] + (1-\lambda)0 = \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(1+\alpha)\right]$$
$$\Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[(0)(1+\bar{m}) + \left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(1+\alpha) + \left(p_{CD\bar{m}} + p_{DD\bar{m}}\right)(0) + (0)(-\beta)\right] + (1-\lambda)0 = \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(1+\alpha)\right]$$

$$\Pi_{\underline{m}}(\bar{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(1+\alpha)\right] - \underline{k}$$
$$\Pi_{\underline{m}}(\underline{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(1+\alpha)\right]$$

$\Pi_{\bar{m}}(CC,\bar{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(1+\bar{m}+\beta)\right] - \beta - \bar{k}$ $\Pi_{\bar{m}}(CD,\bar{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(1+\bar{m}+\beta)\right] - \beta - \bar{k}$ $\Pi_{\bar{m}}(DC,\bar{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(1+\alpha)\right] - \bar{k}$ $\Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(1+\alpha)\right] - \bar{k}$	$\Pi_{\bar{m}}(CC,\underline{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(1+\bar{m}+\beta)\right] - \beta$ $\Pi_{\bar{m}}(CD,\underline{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(1+\bar{m}+\beta)\right] - \beta$ $\Pi_{\bar{m}}(DC,\underline{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(1+\alpha)\right]$ $\Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(1+\alpha)\right]$	$\Pi_{\underline{m}}(\bar{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(1+\alpha)\right] - \underline{k}$ $\Pi_{\underline{m}}(\underline{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(1+\alpha)\right]$

Note that in a pooling equilibrium where everybody sends the signal, CC and CD (DC and DD) will always earn the same profits irrespective of the chosen signal and the particular composition. Since those pairs are indistinguishable we consider the following cases:

1.1. Consider first the case $p_{CC\bar{m}} + p_{CD\bar{m}} = 1, \left(p_{DC\bar{m}} + p_{DD\bar{m}} = 0 \right)$:

$$CC_{\bar{m}} / CD_{\bar{m}} \succ CC_{\underline{m}} / CD_{\underline{m}} \Leftrightarrow \lambda \left(p_{CD\bar{m}} \right) \geq \frac{\bar{k}}{1+\bar{m}+\beta}, \qquad \text{(i)}$$

$$CC_{\bar{m}} / CD_{\bar{m}} \succ DC_{\bar{m}} / DD_{\bar{m}} \Leftrightarrow \lambda \geq \frac{\beta}{\left(\bar{m}-\alpha+\beta\right)} \qquad \text{(ii)}$$

$$CC_{\bar{m}} / CD_{\bar{m}} \succ DC_{\underline{m}} / DD_{\underline{m}} \Leftrightarrow \lambda \geq \frac{\beta+\bar{k}}{\bar{m}-\alpha+\beta+\left(p_{CD\bar{m}}\right)\left(1+\alpha\right)} \qquad \text{(iii)}$$

$$\bar{m} \succ \underline{m} \Leftrightarrow \lambda \left(p_{CD\bar{m}} \right) \geq \frac{\underline{k}}{1+\alpha} \Leftrightarrow \lambda \geq \frac{\underline{k}}{p_{CD\bar{m}}\left(1+\alpha\right)} \qquad \text{(iv)}$$

Note that (iv) implies (i); for (iv) to be satisfied a strictly positive share needs to play CD; note further that (ii) and (iv) imply (iii) because $\frac{\underline{k}}{1+\alpha} > \frac{\bar{k}}{1+\alpha}$ and $\lambda \geq \frac{\underline{k}}{p_{CD\bar{m}}\left(1+\alpha\right)}, \lambda \geq \frac{\beta}{\left(\bar{m}-\alpha+\beta\right)} \Rightarrow \lambda \geq \frac{\beta+\underline{k}}{\bar{m}-\alpha+\beta+\left(p_{CD\bar{m}}\right)\left(1+\alpha\right)} > \frac{\beta+\bar{k}}{\bar{m}-\alpha+\beta+\left(p_{CD\bar{m}}\right)\left(1+\alpha\right)}$ hence for $\lambda \geq \max\left\{ \frac{\underline{k}}{p_{CD\bar{m}}\left(1+\alpha\right)}, \frac{\beta}{\left(\beta+\bar{m}-\alpha\right)} \right\}$ such an equilibrium exists.

1.2. Consider now the case $p_{DC\bar{m}} + p_{DD\bar{m}} = 1, \left(p_{CC\bar{m}} + p_{CD\bar{m}} = 0 \right)$:

This cannot be an equilibrium since low types strictly prefer not to signal.

1.3. Consider finally $p_{CC\bar{m}} + p_{CD\bar{m}} + p_{DC\bar{m}} + p_{DD\bar{m}} = 1$

i.e. all signaling strategies earn the same payoff, i.e. $\lambda\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)=\frac{\beta}{\left(\bar{m}-\alpha+\beta\right)}$. In this case necessary and sufficient is, that any of the signaling strategies is better than any no-signaling strategy, necessary and sufficient for this is:

$$CC_{\bar{m}}/CD_{\bar{m}} \succ CC_{\underline{m}}/CD_{\underline{m}} \Leftrightarrow \lambda\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)\geq\frac{\bar{k}}{1+\bar{m}+\beta}, \qquad \text{(i)}$$

$$DC_{\bar{m}}/DD_{\bar{m}} \succ DC_{\underline{m}}/DD_{\underline{m}} \Leftrightarrow \lambda\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)\geq\frac{\bar{k}}{1+\alpha}, \qquad \text{(ii)}$$

$$\bar{m}\succ\underline{m} \Leftrightarrow \lambda\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)\geq\frac{\underline{k}}{1+\alpha} \qquad \text{(iv)}$$

Note that (ii) implies (i) and (iv) implies (ii), hence for $\lambda\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)\geq\frac{\underline{k}}{1+\alpha}$ such an equilibrium exists.

To summarize:

Equilibrium	Lambda-support
$p_{CC\bar{m}}+p_{CD\bar{m}}=1,\left(p_{DC\bar{m}}+p_{DD\bar{m}}=0\right),1$	$\lambda\geq\max\left\{\frac{\underline{k}}{p_{CD\bar{m}}\left(1+\alpha\right)},\frac{\beta}{\left(\bar{m}-\alpha+\beta\right)}\right\}$ Maximal lambda support is given by setting $p_{CD\bar{m}}=1$; the larger lambda the smaller $p_{CD\bar{m}}$ can get. $\lambda\geq\max\left\{\frac{\underline{k}}{1+\alpha},\frac{\beta}{\left(\beta+\bar{m}-\alpha\right)}\right\}$
$p_{CC\bar{m}}+p_{CD\bar{m}}+p_{DC\bar{m}}+p_{DD\bar{m}}=1, p_{\bar{m}}=$	$\lambda\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)=\frac{\beta}{\left(\bar{m}-\alpha+\beta\right)}$, $\lambda\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)\geq\frac{\underline{k}}{1+\alpha}$

Equilibrium	Lambda-support
$\lambda = \frac{\beta}{(\bar{m}-\alpha+\beta)}$	$\{ p_{CC\bar{m}}, p_{DC\bar{m}}, p_{\bar{m}} \mid p_{CC\bar{m}} + p_{CD\bar{m}} = 1, (p_{DC\bar{m}} + p_{DD\bar{m}} = 0), p_{\bar{m}} = 1, p_{CD\bar{m}} \geq \frac{\underline{k}}{\lambda(1+\alpha)} \}$ A necessary condition for $p_{CD\bar{m}} \geq \frac{\underline{k}}{\lambda(1+\alpha)}$ to be feasible is $\lambda \geq \frac{\underline{k}}{(1+\alpha)}$
$\lambda > \frac{\beta}{(\bar{m}-\alpha+\beta)}$ the two equilibria are indeed different	1. $\{ p_{CC\bar{m}}, p_{CD\bar{m}}, p_{DC\bar{m}}, p_{DD\bar{m}}, p_{\bar{m}} \mid p_{CC\bar{m}} + p_{CD\bar{m}} + p_{DC\bar{m}} + p_{DD\bar{m}} = 1, p_{\bar{m}} = 1, (p_{CC\bar{m}} + p_{CD\bar{m}}) = \frac{\beta}{\lambda(\bar{m}-\alpha+\beta)}, (p_{CD\bar{m}}$ for $(p_{CD\bar{m}} - p_{DC\bar{m}}) \geq \frac{\underline{k}}{\lambda(1+\alpha)}$ to be feasible we need $\frac{\beta}{\bar{m}-\alpha+\beta} \geq \frac{\underline{k}}{1+\alpha}$. 2. $p_{CC\bar{m}} + p_{CD\bar{m}} = 1, (p_{DC\bar{m}} + p_{DD\bar{m}} = 0), p_{\bar{m}} = 1 \; p_{CD\bar{m}} \geq \frac{\underline{k}}{\lambda(1+\alpha)}$

SEMI-POOLING-EQUILIBRIA

We will turn to the analysis of equilibria where only parts of high types or low types signal.

Before we start we will have a closer look on the payoffs for various strategies and their differences. This will significantly simplify the analysis. The following table gives the payoffs for each strategy:

$$\Pi_{\bar{m}}(CC,\bar{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}} + p_{CC\underline{m}} + p_{CD\underline{m}}\right)(1+\bar{m}) + \left(p_{DC\bar{m}} + p_{DD\bar{m}} + p_{DC\underline{m}} + p_{DD\underline{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\underline{m}}(-\beta) + p_{\bar{m}}(-\beta)\right] - \bar{k}$$
$$\Pi_{\bar{m}}(CD,\bar{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(1+\bar{m}) + \left(p_{CC\underline{m}} + p_{CD\underline{m}}\right)(1+\alpha) + \left(p_{DC\bar{m}} + p_{DD\bar{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\bar{m}}(-\beta)\right] - \bar{k}$$
$$\Pi_{\bar{m}}(DC,\bar{m}) = \lambda\left[\left(p_{CC\underline{m}} + p_{CD\underline{m}}\right)(1+\bar{m}) + \left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(1+\alpha) + \left(p_{DC\underline{m}} + p_{DD\underline{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\underline{m}}(-\beta)\right] - \bar{k}$$
$$\Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}} + p_{CC\underline{m}} + p_{CD\underline{m}}\right)(1+\alpha)\right] - \bar{k}$$

$$\Pi_{\bar{m}}(CC,\underline{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}} + p_{CC\underline{m}} + p_{DC\underline{m}}\right)(1+\bar{m}) + \left(p_{CD\bar{m}} + p_{DD\bar{m}} + p_{CD\underline{m}} + p_{DD\underline{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\underline{m}}(-\beta) + p_{\bar{m}}(-\beta)\right]$$
$$\Pi_{\bar{m}}(CD,\underline{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(1+\bar{m}) + \left(p_{CC\underline{m}} + p_{DC\underline{m}}\right)(1+\alpha) + \left(p_{CD\bar{m}} + p_{DD\bar{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\bar{m}}(-\beta)\right]$$
$$\Pi_{\bar{m}}(DC,\underline{m}) = \lambda\left[\left(p_{CC\underline{m}} + p_{DC\underline{m}}\right)(1+\bar{m}) + \left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(1+\alpha) + \left(p_{CD\underline{m}} + p_{DD\underline{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\underline{m}}(-\beta)\right]$$
$$\Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}} + p_{CC\underline{m}} + p_{DC\underline{m}}\right)(1+\alpha)\right]$$

$$\Pi_{\underline{m}}(\underline{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}} + p_{CC\underline{m}} + p_{DC\underline{m}}\right)(1+\alpha)\right]$$
$$\Pi_{\underline{m}}(\bar{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}} + p_{CC\underline{m}} + p_{CD\underline{m}}\right)(1+\alpha)\right] - \underline{k}$$

It will be useful to calculate differences among strategies with different behavior but the same signal and among strategies with different signals.

(1) Within-differences:

$$\Pi_{\bar{m}}\left(CC,\bar{m}\right)-\Pi_{\bar{m}}\left(CD,\bar{m}\right)=\lambda\left[\left(p_{CC\underline{m}}+p_{CD\underline{m}}\right)\left(\bar{m}-\alpha\right)+\left(p_{DC\underline{m}}+p_{DD\underline{m}}\right)\left(-\beta\right)\right]+\left(1-\lambda\right)\left[p_{\underline{m}}\left(-\beta\right)\right]$$

$$\Pi_{\bar{m}}\left(CC,\bar{m}\right)-\Pi_{\bar{m}}\left(DC,\bar{m}\right)=\lambda\left[\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)\left(\bar{m}-\alpha\right)+\left(p_{DC\bar{m}}+p_{DD\bar{m}}\right)\left(-\beta\right)\right]+\left(1-\lambda\right)\left[p_{\bar{m}}\left(-\beta\right)\right]$$

$$\Pi_{\bar{m}}\left(CC,\bar{m}\right)-\Pi_{\bar{m}}\left(DD,\bar{m}\right)=\lambda\left[\left(p_{CC\bar{m}}+p_{CD\bar{m}}+p_{CC\underline{m}}+p_{CD\underline{m}}\right)\left(\bar{m}-\alpha\right)+\left(p_{DC\bar{m}}+p_{DD\bar{m}}+p_{DC\underline{m}}+p_{DD\underline{m}}\right)\left(-\beta\right)\right]+\left(1-\lambda\right)\left[p_{\bar{m}}\left(-\beta\right)+p_{\underline{m}}\left(-\beta\right)\right]$$

$$\Pi_{\bar{m}}\left(CD,\bar{m}\right)-\Pi_{\bar{m}}\left(DC,\bar{m}\right)=\lambda\left[\left(p_{CC\bar{m}}+p_{CD\bar{m}}-p_{CC\underline{m}}-p_{CD\underline{m}}\right)\left(\bar{m}-\alpha\right)+\left(p_{DC\bar{m}}+p_{DD\bar{m}}-p_{DC\underline{m}}-p_{DD\underline{m}}\right)\left(-\beta\right)\right]+\left(1-\lambda\right)\left[p_{\bar{m}}\left(-\beta\right)-p_{\underline{m}}\left(-\beta\right)\right]$$

$$\Pi_{\bar{m}}\left(CD,\bar{m}\right)-\Pi_{\bar{m}}\left(DD,\bar{m}\right)=\lambda\left[\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)\left(\bar{m}-\alpha\right)+\left(p_{DC\bar{m}}+p_{DD\bar{m}}\right)\left(-\beta\right)\right]+\left(1-\lambda\right)\left[p_{\bar{m}}\left(-\beta\right)\right]$$

$$\Pi_{\bar{m}}\left(DC,\bar{m}\right)-\Pi_{\bar{m}}\left(DD,\bar{m}\right)=\lambda\left[\left(p_{CC\underline{m}}+p_{CD\underline{m}}\right)\left(\bar{m}-\alpha\right)+\left(p_{DC\underline{m}}+p_{DD\underline{m}}\right)\left(-\beta\right)\right]+\left(1-\lambda\right)\left[p_{\underline{m}}\left(-\beta\right)\right]$$

$$\Pi_{\bar{m}}\left(CC,\underline{m}\right)-\Pi_{\bar{m}}\left(CD,\underline{m}\right)=\lambda\left[\left(p_{CC\underline{m}}+p_{DC\underline{m}}\right)\left(\bar{m}-\alpha\right)+\left(p_{CD\underline{m}}+p_{DD\underline{m}}\right)\left(-\beta\right)\right]+\left(1-\lambda\right)\left[p_{\underline{m}}\left(-\beta\right)\right]$$

$$\Pi_{\bar{m}}\left(CC,\underline{m}\right)-\Pi_{\bar{m}}\left(DC,\underline{m}\right)=\lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)\left(\bar{m}-\alpha\right)+\left(p_{CD\bar{m}}+p_{DD\bar{m}}\right)\left(-\beta\right)\right]+\left(1-\lambda\right)\left[p_{\bar{m}}\left(-\beta\right)\right]$$

$$\Pi_{\bar{m}}\left(CC,\underline{m}\right)-\Pi_{\bar{m}}\left(DD,\underline{m}\right)=\lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}+p_{CC\underline{m}}+p_{DC\underline{m}}\right)\left(\bar{m}-\alpha\right)+\left(p_{CD\bar{m}}+p_{DD\bar{m}}+p_{CD\underline{m}}+p_{DD\underline{m}}\right)\left(-\beta\right)\right]+\left(1-\lambda\right)\left[p_{\bar{m}}\left(-\beta\right)+p_{\underline{m}}\left(-\beta\right)\right]$$

$$\Pi_{\bar{m}}\left(CD,\underline{m}\right)-\Pi_{\bar{m}}\left(DC,\underline{m}\right)=\lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}-p_{CC\underline{m}}-p_{DC\underline{m}}\right)\left(\bar{m}-\alpha\right)+\left(p_{CD\bar{m}}+p_{DD\bar{m}}-p_{CD\underline{m}}-p_{DD\underline{m}}\right)\left(-\beta\right)\right]+\left(1-\lambda\right)\left[p_{\bar{m}}\left(-\beta\right)-p_{\underline{m}}\left(-\beta\right)\right]$$

$$\Pi_{\bar{m}}\left(CD,\underline{m}\right)-\Pi_{\bar{m}}\left(DD,\underline{m}\right)=\lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)\left(\bar{m}-\alpha\right)+\left(p_{CD\bar{m}}+p_{DD\bar{m}}\right)\left(-\beta\right)\right]+\left(1-\lambda\right)\left[p_{\bar{m}}\left(-\beta\right)\right]$$

$$\Pi_{\bar{m}}\left(DC,\underline{m}\right)-\Pi_{\bar{m}}\left(DD,\underline{m}\right)=\lambda\left[\left(p_{CC\underline{m}}+p_{DC\underline{m}}\right)\left(\bar{m}-\alpha\right)+\left(p_{CD\underline{m}}+p_{DD\underline{m}}\right)\left(-\beta\right)\right]+\left(1-\lambda\right)\left[p_{\underline{m}}\left(-\beta\right)\right]$$

(2) Cross-differences:

$$\Pi_{\underline{m}}\left(\bar{m}\right)-\Pi_{\underline{m}}\left(\underline{m}\right)=\lambda\left[\left(p_{CD\bar{m}}+p_{CD\underline{m}}-p_{DC\bar{m}}-p_{DC\underline{m}}\right)\left(1+\alpha\right)\right]-\underline{k}$$

$$\Pi_{\bar{m}}\left(CC,\bar{m}\right)-\Pi_{\bar{m}}\left(CC,\underline{m}\right)=\lambda\left(1+\bar{m}+\beta\right)\left[\left(p_{CD\bar{m}}+p_{CD\underline{m}}-p_{DC\bar{m}}-p_{DC\underline{m}}\right)\right]-\bar{k}$$

$$\Pi_{\bar{m}}\left(CD,\bar{m}\right)-\Pi_{\bar{m}}\left(CD,\underline{m}\right)=\lambda\left[\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)\left(1+\bar{m}+\beta\right)+\left(p_{CD\underline{m}}-p_{DC\underline{m}}\right)\left(1+\alpha\right)\right]-\bar{k}$$

$$\Pi_{\bar{m}}\left(DC,\bar{m}\right)-\Pi_{\bar{m}}\left(DC,\underline{m}\right)=\lambda\left[\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)\left(1+\alpha\right)+\left(p_{CD\underline{m}}-p_{DC\underline{m}}\right)\left(1+\bar{m}+\beta\right)\right]-\bar{k}$$

$$\Pi_{\bar{m}}\left(DD,\bar{m}\right)-\Pi_{\bar{m}}\left(DD,\underline{m}\right)=\lambda\left(1+\alpha\right)\left[\left(p_{CD\bar{m}}+p_{CD\underline{m}}-p_{DC\bar{m}}-p_{DC\underline{m}}\right)\right]-\bar{k}$$

All other differences can be expressed by the within-differences and the four cross differences above.

Observation 1:

$$\Pi_{\bar{m}}\left(CD,\bar{m}\right)-\Pi_{\bar{m}}\left(CD,\underline{m}\right)=\Pi_{\bar{m}}\left(DD,\bar{m}\right)-\Pi_{\bar{m}}\left(DD,\underline{m}\right)+\lambda\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)\left(\bar{m}-\alpha+\beta\right)=\Pi_{\bar{m}}\left(CC,\bar{m}\right)-\Pi_{\bar{m}}\left(CC,\underline{m}\right)-\lambda\left(p_{CD\underline{m}}-p_{DC\underline{m}}\right)\left(\bar{m}-\alpha+\beta\right)$$

(i) $$\Pi_{\bar{m}}\left(DC,\bar{m}\right)-\Pi_{\bar{m}}\left(DC,\underline{m}\right)=\Pi_{\bar{m}}\left(DD,\bar{m}\right)-\Pi_{\bar{m}}\left(DD,\underline{m}\right)+\lambda\left(p_{CD\underline{m}}-p_{DC\underline{m}}\right)\left(\bar{m}-\alpha+\beta\right)=\Pi_{\bar{m}}\left(CC,\bar{m}\right)-\Pi_{\bar{m}}\left(CC,\underline{m}\right)-\lambda\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)\left(\bar{m}-\alpha+\beta\right)$$

$$\Pi_{\bar{m}}\left(CD,\bar{m}\right)-\Pi_{\bar{m}}\left(CD,\underline{m}\right)=\Pi_{\bar{m}}\left(DC,\bar{m}\right)-\Pi_{\bar{m}}\left(DC,\underline{m}\right)+\lambda\left(p_{CD\bar{m}}+p_{DC\underline{m}}-p_{CD\underline{m}}-p_{DC\bar{m}}\right)\left(\bar{m}-\alpha+\beta\right)$$

(ii) $$\Pi_{\underline{m}}\left(\bar{m}\right)-\Pi_{\underline{m}}\left(\underline{m}\right)=\Pi_{\bar{m}}\left(DD,\bar{m}\right)-\Pi_{\bar{m}}\left(DD,\underline{m}\right)+\bar{k}-\underline{k}$$

A consequence of (ii) of observation 1 is that whenever low types are indifferent in an equilibrium between signaling and not signaling, high types strictly prefer to signal over not to signal given unconditional defective behavior. Put differently, if unconditional defection with and without signal is part of an equilibrium, then low types will prefer not to signal in such an equilibrium.

Observation 2:

(i) $\Pi_{\bar{m}}\left(CC,\bar{m}\right)-\Pi_{\bar{m}}\left(CD,\bar{m}\right)=\Pi_{\bar{m}}\left(DC,\bar{m}\right)-\Pi_{\bar{m}}\left(DD,\bar{m}\right)$ note that differences depend only on non-signaling shares

(ii) $\Pi_{\bar{m}}\left(CC,\underline{m}\right)-\Pi_{\bar{m}}\left(CD,\underline{m}\right)=\Pi_{\bar{m}}\left(DC,\underline{m}\right)-\Pi_{\bar{m}}\left(DD,\underline{m}\right)$

(iii) Corollary: $\Pi_{\bar{m}}\left(CC,\bar{m}\right)+\Pi_{\bar{m}}\left(DD,\bar{m}\right)=\Pi_{\bar{m}}\left(CD,\bar{m}\right)+\Pi_{\bar{m}}\left(DC,\bar{m}\right)$; $\Pi_{\bar{m}}\left(CC,\underline{m}\right)+\Pi_{\bar{m}}\left(DD,\underline{m}\right)=\Pi_{\bar{m}}\left(CD,\underline{m}\right)+\Pi_{\bar{m}}\left(DC,\underline{m}\right)$

Implication:
If within the 4 signal or 4 non-signal behaviors 3 strategies earn the same profit then all 4 strategies earn the same profit. Hence, as a first consequence, there are for each of the cases signal/ no signal only three possibilities: either all 4 strategies earn the same payoff, 2 equal profitable strategies earn strictly more than 2 others, or a single strategy earns more than all others.

If we look at the corollary of observation 2 that the sum of profits for unconditional strategies must equal the sum of profits for conditional strategies, then both conditional can only earn the same profits in equilibrium if the two unconditional strategies earn the same profits too, i.e. all 4 strategies earn the same, otherwise the two unconditional (conditional) strategies must be dominated by one conditional (unconditional) strategy. Furthermore this dominating strategy dominates the second condition (unconditional) strategy. Hence either all strategies earn the same profits or a conditional an unconditional strategy earn the same (highest) payoffs or a single conditional/unconditional strategy earns the highest payoff. The following Lemma summarizes.

<u>Lemma</u>: For each signaling strategy (signal/ no-signal) the table below gives all possible behavioral combinations that could be part of an equiilbrium.

	unconditional versus conditional		
1.	CC=CD ; DC=DD	1.1.	**CC=CD=DC=DD**
		1.2.	**CC=CD**>DC=DD
		1.3.	CC=CD<**DC=DD**
2.	**CC**>CD ; **DC**>DD	2.1.	**CC=DC**
		2.2.	**CC**>DC
		2.3.	CC<**DC**
3.	CC<**CD** ; DC<**DD**	3.1.	**CD=DD**
		3.2.	**CD**>DD
		3.3.	CD<**DD**

Table B-1: possible cases for signaling / no signaling

Proof: whenever CC and CD have a strict payoff relation, so do DC and DD, hence either CC/DC and CD/DD have a strict payoff relation or all four strategies earn the same profit. In the former case there are three possible relations among the dominating pair: either the relation is strict, then we have the situation of an unique behavior or they could earn the same payoff. Hence either all behavior earns the same payoff, a pair of conditional and unconditional behavior (CC/DC or CD/DD) earn the highest payoff or any unique behavior earns highest payoff.

If we neglect for a moment that for a given signal all 4 behaviors are part of a semi pooling equilibrium then following the lemma above, the table below gives all possible combinations of strategies in a semipooling equilibrium.

	$CC,\underline{m}$	$DC,\underline{m}$	$CC,\underline{m}/DC,\underline{m}$	$CD,\underline{m}$	$DD,\underline{m}$	$CD,\underline{m}/DD,\underline{m}$
$CC,\overline{m}$	N (2.)		N (2.)			
$DC,\overline{m}$	N (7.)	N (3.)	N (3.)			
$CC,\overline{m}/DC,\overline{m}$	N (2.)	N (3.)	N (3.)		N (6.)	
$CD,\overline{m}$	N (4.)		N (4.)	N (5.)		N (5.)
$DD,\overline{m}$	N (4.)		N (4.)	N (5.)	N (1.)	N (5.)
$CD,\overline{m}/DD,\overline{m}$	N (4.)		N (4.)	N (5.)		N (5.)

Table B-2: N – cannot exist; for colored cells low types don't signal, because either $CD,\overline{m}$ and $CD,\underline{m}$ are not played (blue) or $DD,\underline{m}$ earns highest payoffs (gray) (see 8.-9.)

However, if we have a closer look at the respective differences we can significantly reduce the number of possible combinations.

$$1.\Pi_{\bar{m}}\left(DD,\bar{m}\right)-\Pi_{\bar{m}}\left(DD,\underline{m}\right)=\lambda\left(1+\alpha\right)\left\lfloor\left(p_{CD\bar{m}}+p_{CD\underline{m}}-p_{DC\bar{m}}-p_{DC\underline{m}}\right)\right\rfloor-\bar{k}=-\bar{k}<0$$

$$2.\Pi_{\bar{m}}\left(CC,\bar{m}\right)-\Pi_{\bar{m}}\left(CC,\underline{m}\right)=\lambda\left(1+\bar{m}+\beta\right)\left[\left(p_{CD\bar{m}}+p_{CD\underline{m}}-p_{DC\bar{m}}-p_{DC\underline{m}}\right)\right]-\bar{k}=-\bar{k}<0$$

$$3.\Pi_{\bar{m}}\left(DC,\bar{m}\right)-\Pi_{\bar{m}}\left(DC,\underline{m}\right)=\lambda\left[\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)\left(1+\alpha\right)+\left(p_{CD\underline{m}}-p_{DC\underline{m}}\right)\left(1+\bar{m}+\beta\right)\right]-\bar{k}=\lambda\left[\left(-p_{DC\bar{m}}\right)\left(1+\alpha\right)+\left(-p_{DC\underline{m}}\right)\left(1+\bar{m}+\beta\right)\right]-\bar{k}<0$$

$$4.\Pi_{\bar{m}}\left(CC,\underline{m}\right)-\Pi_{\bar{m}}\left(DC,\underline{m}\right)=\lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)\left(\bar{m}-\alpha\right)+\left(p_{CD\bar{m}}+p_{DD\bar{m}}\right)\left(-\beta\right)\right]+\left(1-\lambda\right)\left[p_{\bar{m}}\left(-\beta\right)\right]\overset{p_{CC\bar{m}}+p_{DC\bar{m}}=0}{<}0$$

$$\Rightarrow:\Pi_{\bar{m}}\left(CC,\underline{m}\right)-\Pi_{\bar{m}}\left(DC,\underline{m}\right)\geq 0\Rightarrow p_{CC\bar{m}}+p_{DC\bar{m}}>0$$

$$5.\Pi_{\bar{m}}\left(CD,\underline{m}\right)-\Pi_{\bar{m}}\left(DD,\underline{m}\right)=\lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)\left(\bar{m}-\alpha\right)+\left(p_{CD\bar{m}}+p_{DD\bar{m}}\right)\left(-\beta\right)\right]+\left(1-\lambda\right)\left[p_{\bar{m}}\left(-\beta\right)\right]\overset{p_{CC\bar{m}}+p_{DC\bar{m}}=0}{<}0$$

$$\Rightarrow:\Pi_{\bar{m}}\left(CD,\underline{m}\right)-\Pi_{\bar{m}}\left(DD,\underline{m}\right)\geq 0\Rightarrow p_{CC\bar{m}}+p_{DC\bar{m}}>0$$

$$6.\Pi_{\bar{m}}\left(DC,\bar{m}\right)-\Pi_{\bar{m}}\left(DD,\underline{m}\right)=\lambda\left[+\left(p_{CD\underline{m}}\right)\left(1+\bar{m}\right)+\left(-p_{DC\bar{m}}\right)\left(1+\alpha\right)+\left(p_{DD\underline{m}}\right)\left(-\beta\right)\right]+\left(1-\lambda\right)\left[p_{\underline{m}}\left(-\beta\right)\right]-\bar{k}\overset{p_{CD\underline{m}}=0}{<}0$$

$$7.\Pi_{\bar{m}}\left(DC,\bar{m}\right)-\Pi_{\bar{m}}\left(CC,\underline{m}\right)=\Pi_{\bar{m}}\left(CC,\bar{m}\right)-\Pi_{\bar{m}}\left(CC,\underline{m}\right)-\left[\Pi_{\bar{m}}\left(CC,\bar{m}\right)-\Pi_{\bar{m}}\left(DC,\bar{m}\right)\right]=\lambda\left(1+\bar{m}+\beta\right)\left[\left(p_{CD\bar{m}}+p_{CD\underline{m}}-p_{DC\bar{m}}-p_{DC\underline{m}}\right)\right]-\bar{k}$$

$$-\left\{\lambda\left[\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)\left(\bar{m}-\alpha\right)+\left(p_{DC\bar{m}}+p_{DD\bar{m}}\right)\left(-\beta\right)\right]+\left(1-\lambda\right)\left[p_{\bar{m}}\left(-\beta\right)\right]\right\}\overset{p_{\bar{m}}=0}{=}$$

$$\lambda\left(1+\bar{m}+\beta\right)\left[\left(-p_{DC\bar{m}}\right)\right]-\bar{k}-\lambda\left[\left(p_{DC\bar{m}}\right)\left(-\beta\right)\right]=-p_{DC\bar{m}}\lambda\left(1+\bar{m}\right)-\bar{k}<0$$

$$8.\Pi_{\underline{m}}\left(\bar{m}\right)-\Pi_{\underline{m}}\left(\underline{m}\right)=\Pi_{\bar{m}}\left(DD,\bar{m}\right)-\Pi_{\bar{m}}\left(DD,\underline{m}\right)+\bar{k}-\underline{k}\overset{\Pi_{\bar{m}}\left(DD,\bar{m}\right)-\Pi_{\bar{m}}\left(DD,\underline{m}\right)\leq 0}{<}0$$

$9.\Pi_{\underline{m}}\left(\bar{m}\right)-\Pi_{\underline{m}}\left(\underline{m}\right)=\lambda\left[\left(p_{CD\bar{m}}+p_{CD\underline{m}}-p_{DC\bar{m}}-p_{DC\underline{m}}\right)\left(1+\alpha\right)\right]-\underline{k}$, hence if neither $CD,\underline{m}$ nor $CD,\bar{m}$ is played then low types strictly prefer not to signal, i.e. $p_{\underline{m}}=1$.

Observation 3:

$\Pi_{\bar{m}}(XY,m)-\Pi_{\bar{m}}(XZ,m)=f(p_{CC\underline{m}},p_{CD\underline{m}},p_{DC\underline{m}},p_{DD\underline{m}})$ $\quad X,Y,Z\in\{C,D\},m\in\{\underline{m},\bar{m}\},Y\neq Z$ and

$\Pi_{\bar{m}}(YX,m)-\Pi_{\bar{m}}(ZX,m)=f(p_{CC\bar{m}},p_{CD\bar{m}},p_{DC\bar{m}},p_{DD\bar{m}})$ $\quad X,Y,Z\in\{C,D\},m\in\{\underline{m},\bar{m}\},Y\neq Z$

Before we turn to the 14 remaining cases of table 2, we check for semi-pooling equilibria that contain all 4 behaviors for at least one signal.

1. **<u>All 8 strategies are played by high types (4 vs. 4)</u>**

Due to $\begin{matrix}\Pi_{\bar{m}}(DD,\bar{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda(1+\alpha)\left[\left(p_{CD\bar{m}}+p_{CD\underline{m}}-p_{DC\bar{m}}-p_{DC\underline{m}}\right)\right]-\bar{k}\\ \Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CC,\underline{m})=\lambda(1+\bar{m}+\beta)\left[\left(p_{CD\bar{m}}+p_{CD\underline{m}}-p_{DC\bar{m}}-p_{DC\underline{m}}\right)\right]-\bar{k}\end{matrix}$ there cannot be an equilibrium such that both equations are satisfied, required for an equilibrium where all strategies earn the same profits.

2. **<u>All four signaling strategies earn same profit, i.e.</u>** $\Pi_{\bar{m}}(CC,\bar{m})=\Pi_{\bar{m}}(CD,\bar{m})=\Pi_{\bar{m}}(DC,\bar{m})=\Pi_{\bar{m}}(DD,\bar{m})$ **(4 versus 2/1)**

2.1. $CC,\bar{m}/DC,\bar{m}/CD,\bar{m}/DD,\bar{m}$ *vs.* $CC,\underline{m}/DC,\underline{m}$, i.e. $\Pi_{\bar{m}}(CC,\underline{m})-\Pi_{\bar{m}}(CD,\underline{m})=\Pi_{\bar{m}}(DC,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})>0$ (*)

2.1.1. $\Pi_{\bar{m}}(CC,\underline{m})\geq\Pi_{\bar{m}}(DC,\underline{m})\left(>\Pi_{\bar{m}}(DD,\underline{m})\right)$

$CC,\bar{m}/CC,\underline{m}$ earn same profits, i.e. $\begin{matrix}\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CC,\underline{m})=\lambda(1+\bar{m}+\beta)\left[\left(p_{CD\bar{m}}+p_{CD\underline{m}}-p_{DC\bar{m}}-p_{DC\underline{m}}\right)\right]-\bar{k}=0\Rightarrow\\ \Pi_{\bar{m}}(DD,\bar{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda(1+\alpha)\left[\left(p_{CD\bar{m}}+p_{CD\underline{m}}-p_{DC\bar{m}}-p_{DC\underline{m}}\right)\right]-\bar{k}<0\end{matrix}$ Hence $DD,\bar{m}$ cannot be part of the equilibrium, cannot earn the same profits as $CC,\bar{m}$. Therefor such an equilibrium cannot exist.

2.1.2. $\Pi_{\bar{m}}(CC,\underline{m})<\Pi_{\bar{m}}(DC,\underline{m})$, i.e. $p_{CC\underline{m}}=p_{CD\underline{m}}=p_{DD\underline{m}}=0$

$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CD,\bar{m})=\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left[\left(p_{DC\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]=0$ is violated if $p_{DC\underline{m}}>0$ which is necessary for a **semi**-pooling equilibrium.

Therefor such an equilibrium cannot exist.

2.2. $CC,\bar{m}/DC,\bar{m}/CD,\bar{m}/DD,\bar{m}$ *vs.* $CD,\underline{m}/DD,\underline{m}$, i.e. $\Pi_{\bar{m}}(CC,\underline{m})-\Pi_{\bar{m}}(CD,\underline{m})=\Pi_{\bar{m}}(DC,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})<0$ $p_{CC\underline{m}}=p_{DC\underline{m}}=0$

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CD,\bar{m})=\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left[\left(p_{CC\underline{m}}+p_{CD\underline{m}}\right)(\bar{m}-\alpha)+\left(p_{DC\underline{m}}+p_{DD\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]=0 \quad \text{(i)}$$

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(DC,\bar{m})=\Pi_{\bar{m}}(CD,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{DC\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]=0 \quad \text{(ii)}$$

$$\Pi_{\bar{m}}(CC,\underline{m})-\Pi_{\bar{m}}(CD,\underline{m})=\Pi_{\bar{m}}(DC,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[\left(p_{CC\underline{m}}+p_{DC\underline{m}}\right)(\bar{m}-\alpha)+\left(p_{CD\underline{m}}+p_{DD\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]<0 \quad \text{(iii)}$$

2.2.1. $\Pi_{\bar{m}}(DD,\underline{m})\le\Pi_{\bar{m}}(CD,\underline{m})$

$$\Pi_{\bar{m}}(CD,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{CD\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]\ge 0 \quad \text{(iv)}$$

$$\Pi_{\bar{m}}(CD,\bar{m})-\Pi_{\bar{m}}(CD,\underline{m})=\Pi_{\bar{m}}(DD,\bar{m})-\Pi_{\bar{m}}(DD,\underline{m})+\lambda\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)(\bar{m}-\alpha+\beta) \quad \text{(v)}$$

$$=\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CC,\underline{m})-\lambda\left(p_{CD\underline{m}}-p_{DC\underline{m}}\right)(\bar{m}-\alpha+\beta)$$

$$=\lambda\left[\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)(1+\bar{m}+\beta)+\left(p_{CD\underline{m}}-p_{DC\underline{m}}\right)(1+\alpha)\right]-\bar{k}=0$$

The last equation implies that $\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)\le 0$ and $\left(p_{CD\underline{m}}-p_{DC\underline{m}}\right)>0$.

$p_{CC\underline{m}}=p_{DC\underline{m}}=0$ →

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CD,\bar{m})=\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left[\left(p_{CD\underline{m}}\right)(\bar{m}-\alpha)+\left(p_{DD\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]=0 \quad \text{(i)}$$

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(DC,\bar{m})=\Pi_{\bar{m}}(CD,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{DC\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]= \quad \text{(ii)}$$

$$\Pi_{\bar{m}}(CC,\underline{m})-\Pi_{\bar{m}}(CD,\underline{m})=\Pi_{\bar{m}}(DC,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[\left(p_{CD\underline{m}}+p_{DD\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]<0 \quad \text{(iii)}$$

$$\Pi_{\bar{m}}(CD,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{CD\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]\ge 0 \quad \text{(iv)}$$

$$\Pi_{\bar{m}}(CD,\bar{m})-\Pi_{\bar{m}}(CD,\underline{m})=\Pi_{\bar{m}}(DD,\bar{m})-\Pi_{\bar{m}}(DD,\underline{m})+\lambda\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)(\bar{m}-\alpha+\beta) \quad \text{(v)}$$
$$=\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CC,\underline{m})-\lambda\left(p_{CD\underline{m}}\right)(\bar{m}-\alpha+\beta)$$
$$=\lambda\left[\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)(1+\bar{m}+\beta)+\left(p_{CD\underline{m}}\right)(1+\alpha)\right]-\bar{k}=0$$

(iii) is always satisfied in a semi-pooling equilibrium

2.2.1.1. Eq. for $\Pi_{\bar{m}}(DD,\underline{m})=\Pi_{\bar{m}}(CD,\underline{m})\overset{Obs.1}{\Rightarrow} p_{\underline{m}}=1$:

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CD,\bar{m})=\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left[\left(p_{CD\underline{m}}\right)(\bar{m}-\alpha)+\left(p_{DD\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]=0 \quad \text{(i)}$$
$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(DC,\bar{m})=\Pi_{\bar{m}}(CD,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{DC\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]= \quad \text{(ii)}$$
$$\Pi_{\bar{m}}(CD,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{CD\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]=0 \quad \text{(iv)}$$
$$\Pi_{\bar{m}}(CD,\bar{m})-\Pi_{\bar{m}}(CD,\underline{m})=\Pi_{\bar{m}}(DD,\bar{m})-\Pi_{\bar{m}}(DD,\underline{m})+\lambda\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)(\bar{m}-\alpha+\beta) \quad \text{(v)}$$
$$=\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CC,\underline{m})-\lambda\left(p_{CD\underline{m}}\right)(\bar{m}-\alpha+\beta)$$
$$=\lambda\left[\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)(1+\bar{m}+\beta)+\left(p_{CD\underline{m}}\right)(1+\alpha)\right]-\bar{k}$$
$$=\lambda(1+\alpha)\left[\left(p_{CD\bar{m}}+p_{CD\underline{m}}-p_{DC\bar{m}}\right)\right]-\bar{k}+\lambda\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)(\bar{m}-\alpha+\beta)=0$$

(v) implies $\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)=0$ then $p_{CD\underline{m}}=\dfrac{\bar{k}}{\lambda(1+\alpha)}$ by (v), $p_{DD\underline{m}}=\dfrac{\bar{k}}{\lambda(1+\alpha)}\dfrac{\bar{m}-\alpha}{\beta}-\dfrac{1-\lambda}{\lambda}$ by (i) and so (ii), (iv) are remaining: given that $\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)=0$ (ii) and (iv) are equivalent., i.e. those two equation amount to one further condition on the shares among high types.

In summary the equilibrium set is given by:

$$\left\{\begin{array}{l}\left(p_{CC\bar{m}}, p_{CD\bar{m}}, p_{DC\bar{m}}, p_{DD\bar{m}}, p_{CD\underline{m}}, p_{DD\underline{m}}, p_{\underline{m}}=1\right) \Big| \\ p_{CC\bar{m}}=-\frac{\bar{k}}{\lambda(1+\alpha)}+\frac{\beta}{\lambda(\bar{m}-\alpha+\beta)}-p_{DC\bar{m}} \\ p_{CD\bar{m}}=p_{DC\bar{m}}, \\ p_{DD\bar{m}}=-\frac{\bar{k}}{\lambda(1+\alpha)}\frac{\bar{m}-\alpha}{\beta}-\frac{\beta}{\lambda(\bar{m}-\alpha+\beta)}+\frac{1}{\lambda}-p_{DC\bar{m}} \\ p_{CD\underline{m}}=\frac{\bar{k}}{\lambda(1+\alpha)} \\ p_{DD\underline{m}}=\frac{\bar{k}}{\lambda(1+\alpha)}\frac{\bar{m}-\alpha}{\beta}-\frac{1-\lambda}{\lambda}\end{array}\right\}$$

Note that the condition $p_{DD\underline{m}} \geq 0 \Leftrightarrow \frac{\bar{k}}{\lambda(1+\alpha)}\frac{\bar{m}-\alpha}{\beta}-\frac{1-\lambda}{\lambda} \geq 0 \Leftrightarrow \lambda \geq 1-\frac{\bar{k}}{(1+\alpha)}\frac{\bar{m}-\alpha}{\beta}$. On the other hand in a semi-pooling equilibrium where high apply both types of signals we must have:

$$p_{CD\underline{m}}+p_{DD\underline{m}}<1 \Leftrightarrow \frac{\bar{k}}{\lambda(1+\alpha)}\frac{\bar{m}-\alpha+\beta}{\beta}-\frac{1-\lambda}{\lambda}<1 \Leftrightarrow \frac{\bar{k}}{(1+\alpha)}\frac{\bar{m}-\alpha+\beta}{\beta}<1$$

All conditions of the type
$p \in [0,1], \sum p<1$ reduce to:

(1): $p_{DD\underline{m}} \geq 0 \Leftrightarrow \frac{\bar{k}}{\lambda(1+\alpha)}\frac{\bar{m}-\alpha}{\beta}-\frac{1-\lambda}{\lambda} \geq 0 \Leftrightarrow \lambda \geq 1-\frac{\bar{k}}{(1+\alpha)}\frac{\bar{m}-\alpha}{\beta}$

(2): $p_{CD\underline{m}} \geq 0$ true

(3): $p_{DD\underline{m}} + p_{CD\underline{m}} < 1 \Leftrightarrow \frac{\bar{k}}{1+\alpha} < \frac{\beta}{\bar{m}-\alpha+\beta}$

(4): $p_{CC\bar{m}}, p_{DD\bar{m}}, p_{CD\bar{m}} \geq 0 \Leftrightarrow 0 \leq p_{DC\bar{m}} \leq \frac{1}{\lambda}\left(\frac{\beta}{\bar{m}-\alpha+\beta}-\frac{\bar{k}}{1+\alpha}\right)\min\left\{\frac{\bar{m}-\alpha}{\beta},1\right\} = \begin{cases} \frac{1}{\lambda}\left(\frac{\beta}{\bar{m}-\alpha+\beta}-\frac{\bar{k}}{1+\alpha}\right) & ,\beta<(\bar{m}-\alpha) \\ \frac{1}{\lambda}\frac{\bar{m}-\alpha}{\beta}\left(\frac{\beta}{\bar{m}-\alpha+\beta}-\frac{\bar{k}}{1+\alpha}\right) & ,\beta>(\bar{m}-\alpha) \end{cases}$

Note:

$a)\ \frac{\beta}{\bar{m}-\alpha+\beta}-\frac{\bar{k}}{1+\alpha}-1+\frac{\bar{k}}{(1+\alpha)}\frac{\bar{m}-\alpha}{\beta} = -\frac{\bar{m}-\alpha}{\bar{m}-\alpha+\beta}+\frac{\bar{k}}{(1+\alpha)}\frac{\bar{m}-\alpha-\beta}{\beta} > 0 \Leftrightarrow \frac{\bar{k}}{(1+\alpha)} > \frac{\beta(\bar{m}-\alpha)}{(\bar{m}-\alpha+\beta)(\bar{m}-\alpha-\beta)} > \frac{\beta}{\bar{m}-\alpha+\beta}$

which violates (3), hence $\frac{1}{\lambda}\left(\frac{\beta}{\bar{m}-\alpha+\beta}-\frac{\bar{k}}{1+\alpha}\right)$ is binding, i.e. <1.

$\frac{\beta(\bar{m}-\alpha)}{(\bar{m}-\alpha+\beta)(\bar{m}-\alpha-\beta)} < \frac{\bar{k}}{(1+\alpha)} < \frac{\beta}{\bar{m}-\alpha+\beta}$ false, since $\frac{(\bar{m}-\alpha)}{(\bar{m}-\alpha-\beta)} > 1$, i.e.

$b)\ \frac{\bar{m}-\alpha}{\beta}\left(\frac{\beta}{\bar{m}-\alpha+\beta}-\frac{\bar{k}}{1+\alpha}\right)-1+\frac{\bar{k}}{(1+\alpha)}\frac{\bar{m}-\alpha}{\beta} = \frac{\bar{m}-\alpha}{\bar{m}-\alpha+\beta}-1<0 \Rightarrow \frac{1}{\lambda}\frac{\bar{m}-\alpha}{\beta}\left(\frac{\beta}{\bar{m}-\alpha+\beta}-\frac{\bar{k}}{1+\alpha}\right)$ is binding, i.e. <1.

Conditions for existence:

1. $\frac{\bar{k}}{1+\alpha} < \frac{\beta}{\bar{m}-\alpha+\beta}$

2. $\lambda \geq 1 - \frac{\bar{k}}{(1+\alpha)} \frac{\bar{m}-\alpha}{\beta}$ (>0, by 1.)

3.

$\beta < (\bar{m}-\alpha)$: $\quad p_{DC\bar{m}} \leq \frac{1}{\lambda}\left(\frac{\beta}{\bar{m}-\alpha+\beta} - \frac{\bar{k}}{1+\alpha}\right)$ (> 0, by 1.)

$\beta \geq (\bar{m}-\alpha)$: $\quad p_{DC\bar{m}} \leq \frac{1}{\lambda}\frac{\bar{m}-\alpha}{\beta}\left(\frac{\beta}{\bar{m}-\alpha+\beta} - \frac{\bar{k}}{1+\alpha}\right)$ (> 0, by 1.)

2.2.1.2. Eq. for $\Pi_{\bar{m}}(DD,\underline{m}) < \Pi_{\bar{m}}(CD,\underline{m})$: $p_{DD\underline{m}} = 0$ plugged into

$$\Pi_{\bar{m}}(CC,\bar{m}) - \Pi_{\bar{m}}(CD,\bar{m}) = \Pi_{\bar{m}}(DC,\bar{m}) - \Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[\left(p_{CD\underline{m}}\right)(\bar{m}-\alpha) + \left(p_{DD\underline{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\underline{m}}(-\beta)\right] = 0 \quad \text{(i)}$$

$$\Pi_{\bar{m}}(CC,\bar{m}) - \Pi_{\bar{m}}(DC,\bar{m}) = \Pi_{\bar{m}}(CD,\bar{m}) - \Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(\bar{m}-\alpha) + \left(p_{DC\bar{m}} + p_{DD\bar{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\bar{m}}(-\beta)\right] = \quad \text{(ii)}$$

$$\Pi_{\bar{m}}(CD,\underline{m}) - \Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(\bar{m}-\alpha) + \left(p_{CD\bar{m}} + p_{DD\bar{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\bar{m}}(-\beta)\right] > 0 \quad \text{(iv)}$$

$$\Pi_{\bar{m}}(CD,\bar{m}) - \Pi_{\bar{m}}(CD,\underline{m}) = \Pi_{\bar{m}}(DD,\bar{m}) - \Pi_{\bar{m}}(DD,\underline{m}) + \lambda\left(p_{CD\bar{m}} - p_{DC\bar{m}}\right)(\bar{m}-\alpha+\beta) \quad \text{(v)}$$

$$= \Pi_{\bar{m}}(CC,\bar{m}) - \Pi_{\bar{m}}(CC,\underline{m}) - \lambda\left(p_{CD\underline{m}}\right)(\bar{m}-\alpha+\beta)$$

$$= \lambda\left[\left(p_{CD\bar{m}} - p_{DC\bar{m}}\right)(1+\bar{m}+\beta) + \left(p_{CD\underline{m}}\right)(1+\alpha)\right] - \bar{k}$$

$$= \lambda(1+\alpha)\left[\left(p_{CD\bar{m}} + p_{CD\underline{m}} - p_{DC\bar{m}}\right)\right] - \bar{k} + \lambda\left(p_{CD\bar{m}} - p_{DC\bar{m}}\right)(\bar{m}-\alpha+\beta) = 0$$

(i) gives us $\lambda\left[\left(p_{CD\underline{m}}\right)(\bar{m}-\alpha)\right] + (1-\lambda)\left[p_{\underline{m}}(-\beta)\right] = 0$ which for a semi-pooling equilibrium ($p_{CD\underline{m}} > 0$) requires $p_{\underline{m}} > 0$ hence $\Pi_{\underline{m}}(\bar{m}) - \Pi_{\underline{m}}(\underline{m}) = \lambda(1+\alpha)\left[\left(p_{CD\bar{m}} + p_{CD\underline{m}} - p_{DC\bar{m}} - p_{DC\underline{m}}\right)\right] - \underline{k} \leq 0$ Note that with $p_{\underline{m}} > 0$ (ii) will always be satisfied.

2.2.1.2.1. $\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})=0$

$$\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})=\lambda(1+\alpha)\left[\left(p_{CD\bar{m}}+p_{CD\underline{m}}-p_{DC\bar{m}}-p_{DC\underline{m}}\right)\right]-\underline{k}=0 \quad (*)$$

$$\lambda(1+\alpha)\left[\left(p_{CD\bar{m}}+p_{CD\underline{m}}-p_{DC\bar{m}}\right)\right]-\underline{k}=0 \Leftrightarrow \lambda(1+\alpha)p_{CD\underline{m}}=\underline{k}-\lambda(1+\alpha)\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right) \quad (*)$$

$$\lambda\left[\left(p_{CD\underline{m}}\right)(\bar{m}-\alpha)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]=0 \quad \text{(i)}$$

$$\lambda\left[\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{DC\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]=0 \quad \text{(ii)}$$

$$\lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{CD\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]>0 \quad \text{(iv)}$$

$$\lambda(1+\alpha)p_{CD\underline{m}}=\bar{k}-\lambda(1+\bar{m}+\beta)\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right) \quad \text{(v)}$$

Hence $p_{DC\bar{m}}-p_{CD\bar{m}}=\dfrac{\underline{k}-\bar{k}}{\lambda(\bar{m}-\alpha+\beta)}$ by (*) and (v) and therefor

$$\lambda(1+\alpha)p_{CD\underline{m}}=\underline{k}-\lambda(1+\alpha)\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)=\underline{k}+\lambda(1+\alpha)\frac{\underline{k}-\bar{k}}{\lambda(\bar{m}-\alpha+\beta)} \Leftrightarrow p_{CD\underline{m}}=\frac{\underline{k}}{\lambda(1+\alpha)}+\frac{\underline{k}-\bar{k}}{\lambda(\bar{m}-\alpha+\beta)} \text{ by } (*)$$

and by (i) $p_{\underline{m}}=\dfrac{\lambda}{1-\lambda}\dfrac{\bar{m}-\alpha}{\beta}\left(\dfrac{\underline{k}}{\lambda(1+\alpha)}+\dfrac{\underline{k}-\bar{k}}{\lambda(\bar{m}-\alpha+\beta)}\right)=\dfrac{1}{1-\lambda}\dfrac{\bar{m}-\alpha}{\beta}\left(\dfrac{\underline{k}}{(1+\alpha)}+\dfrac{\underline{k}-\bar{k}}{(\bar{m}-\alpha+\beta)}\right)$.

Finally, rearrange (i) to: $\lambda\left(p_{CD\underline{m}}\right)(\bar{m}-\alpha)-(1-\lambda)\beta=(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]$ and plug it into (ii): or equivalently

$$p_{CC\bar{m}}+p_{DC\bar{m}}=\frac{\beta}{\lambda(\bar{m}-\alpha+\beta)}-\frac{\underline{k}}{\lambda(1+\alpha)},\ p_{CD\bar{m}}+p_{DD\bar{m}}=1-\frac{\beta}{\lambda(\bar{m}-\alpha+\beta)}-\frac{\underline{k}-\bar{k}}{\lambda(\bar{m}-\alpha+\beta)}=1-\frac{\beta-\bar{k}+\underline{k}}{\lambda(\bar{m}-\alpha+\beta)}$$

Note, that these values imply that (iv) is satisfied :

$$\lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)\left(\bar{m}-\alpha\right)+\left(p_{CD\bar{m}}+p_{DD\bar{m}}\right)\left(-\beta\right)\right]+\left(1-\lambda\right)\left[p_{\bar{m}}\left(-\beta\right)\right]=$$
$$\lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)\left(\bar{m}-\alpha+\beta\right)-\beta+\beta p_{CD\underline{m}}\right]+\lambda p_{CD\underline{m}}\left(\bar{m}-\alpha\right)-\left(1-\lambda\right)\beta=$$
$$\lambda\left(p_{CC\bar{m}}+p_{DC\bar{m}}+p_{CD\underline{m}}\right)\left(\bar{m}-\alpha+\beta\right)-\beta=\underline{k}-\bar{k}>0$$

In summary equilibrium set is given by

$$\left\{\begin{array}{l}\left(p_{CC\bar{m}},p_{CD\bar{m}},p_{DC\bar{m}},p_{DD\bar{m}},p_{CD\underline{m}},p_{\underline{m}}\right)\Big| \\ p_{CC\bar{m}}=\dfrac{\left(1+\alpha\right)\left(\beta+\bar{k}\right)-\left(1+\beta+\bar{m}\right)\underline{k}}{\lambda\left(1+\alpha\right)\left(\bar{m}-\alpha+\beta\right)}-p_{CD\bar{m}} \\ p_{DC\bar{m}}=p_{CD\bar{m}}+\dfrac{\underline{k}-\bar{k}}{\lambda\left(\bar{m}-\alpha+\beta\right)} \\ p_{DD\bar{m}}=1-\dfrac{\beta+\underline{k}-\bar{k}}{\lambda\left(\bar{m}-\alpha+\beta\right)}-p_{CD\bar{m}} \\ p_{CD\underline{m}}=\dfrac{\underline{k}}{\lambda\left(1+\alpha\right)}+\dfrac{\underline{k}-\bar{k}}{\lambda\left(\bar{m}-\alpha+\beta\right)}=\dfrac{\left(1+\beta+\bar{m}\right)\underline{k}-\left(1+\alpha\right)\bar{k}}{\lambda\left(1+\alpha\right)\left(\bar{m}-\alpha+\beta\right)} \\ p_{\underline{m}}=\dfrac{1}{1-\lambda}\dfrac{\bar{m}-\alpha}{\beta}\left(\dfrac{\underline{k}}{\left(1+\alpha\right)}+\dfrac{\underline{k}-\bar{k}}{\left(\bar{m}-\alpha+\beta\right)}\right)=\dfrac{1}{1-\lambda}\dfrac{\bar{m}-\alpha}{\beta}\dfrac{\left(1+\beta+\bar{m}\right)\underline{k}-\left(1+\alpha\right)\bar{k}}{\left(1+\alpha\right)\left(\bar{m}-\alpha+\beta\right)}\end{array}\right\}$$

Note: $p_{CD\bar{m}}+p_{CD\underline{m}}=p_{DC\bar{m}}+\dfrac{\underline{k}}{\lambda\left(1+\alpha\right)}$

$$\Pi_{\underline{m}}=\Pi_{\underline{m}}\left(\underline{m}\right)=\lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)\left(1+\alpha\right)\right]<\lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)\left(1+\alpha\right)\right]+\underline{k}-\bar{k}=\Pi_{\bar{m}}\left(DD,\underline{m}\right)=\Pi_{\bar{m}}$$

<u>Existence</u>:

All conditions of the type

$p \in [0,1], \sum p < 1$ reduce to:

(1): $p_{CD\underline{m}} > 0$ true

(2): $p_{CD\underline{m}} < 1 \Leftrightarrow \frac{\bar{k}}{1+\alpha} + \frac{\underline{k}-\bar{k}}{(\bar{m}-\alpha+\beta)} < \lambda \Leftrightarrow \frac{(1+\beta+\bar{m})\underline{k}-(1+\alpha)\bar{k}}{(\bar{m}-\alpha+\beta)} < \lambda$

(3): $p_{CC\bar{m}}, p_{DD\bar{m}}, p_{DC\bar{m}} \geq 0 \Leftrightarrow 0 \leq \mathrm{p}_{CD\bar{m}} \leq \min\left\{ \underbrace{1-\frac{\beta+\underline{k}-\bar{k}}{\lambda(\bar{m}-\alpha+\beta)}}_{x}, \underbrace{\frac{(1+\alpha)(\beta+\bar{k})-(1+\beta+\bar{m})\underline{k}}{\lambda(1+\alpha)(\bar{m}-\alpha+\beta)}}_{y} \right\}$

$a) x > 0 \Leftrightarrow \lambda > \frac{\beta+\underline{k}-\bar{k}}{(\bar{m}-\alpha+\beta)}$ $\quad x < 1$ true

$b) y > 0 \Leftrightarrow (1+\alpha)(\beta+\bar{k})-(1+\beta+\bar{m})\underline{k} > 0$ $\quad y < 1 \Leftrightarrow \frac{(1+\alpha)(\beta+\bar{k})-(1+\beta+\bar{m})\underline{k}}{(1+\alpha)(\bar{m}-\alpha+\beta)} = \frac{\beta+\bar{k}-\frac{(1+\beta+\bar{m})}{(1+\alpha)}\underline{k}}{(\bar{m}-\alpha+\beta)} < \lambda$ (by 3a))

$c) x \geq y \Leftrightarrow \lambda \geq \frac{2\beta}{(\bar{m}-\alpha+\beta)} - \frac{\underline{k}}{1+\alpha}$

(4): $p_{\underline{m}} \geq 0$ true

(5): $p_{\underline{m}} \leq 1 \Leftrightarrow \lambda \leq 1 - \frac{\bar{m}-\alpha}{\beta}\left(\frac{\underline{k}}{(1+\alpha)} + \frac{\underline{k}-\bar{k}}{(\bar{m}-\alpha+\beta)}\right) = 1 - \frac{\bar{m}-\alpha}{\beta}\frac{(1+\beta+\bar{m})\underline{k}-(1+\alpha)\bar{k}}{(1+\alpha)(\bar{m}-\alpha+\beta)}$ (>0, by 3b))

note: $(1+\alpha)\beta-(\bar{m}-\alpha+\beta)\bar{k} > (1+\beta+\bar{m})(\underline{k}-\bar{k}) > 0$, by 3b) $\Rightarrow \frac{\beta+\underline{k}-\bar{k}}{(\bar{m}-\alpha+\beta)} > \frac{\bar{k}}{1+\alpha} + \frac{\underline{k}-\bar{k}}{(\bar{m}-\alpha+\beta)}$, i.e. (2) is not binding.

1. (3)b): $\left(1+\alpha\right)\beta > \left(1+\beta+\bar{m}\right)\underline{k} - \left(1+\alpha\right)\bar{k}$

2. (3)a) $\wedge$ (2) $\wedge$ (5): $0 < \frac{\beta+\underline{k}-\bar{k}}{\left(\bar{m}-\alpha+\beta\right)} < \lambda \leq 1 - \frac{\bar{m}-\alpha}{\beta}\left(\frac{\underline{k}}{\left(1+\alpha\right)} + \frac{\underline{k}-\bar{k}}{\left(\bar{m}-\alpha+\beta\right)}\right) < 1\left(LHS < RHS, \text{NOT by 3b) see below}\right)$

3. $0 \leq p_{CD\bar{m}} \leq \min\left\{\underbrace{1-\frac{\beta+\underline{k}-\bar{k}}{\lambda\left(\bar{m}-\alpha+\beta\right)}}_{x}, \underbrace{\frac{\beta+\bar{k}-\frac{1+\beta+\bar{m}}{1+\alpha}\underline{k}}{\lambda\left(\bar{m}-\alpha+\beta\right)}}_{y}\right\}$

$$x \geq y \Leftrightarrow \lambda \geq \frac{2\beta}{\left(\bar{m}-\alpha+\beta\right)} - \frac{\underline{k}}{1+\alpha}\left(\overset{(3)b)}{>} \frac{\beta+\underline{k}-\bar{k}}{\left(\bar{m}-\alpha+\beta\right)}\right)$$

$$\frac{\beta+\underline{k}-\bar{k}}{\left(\bar{m}-\alpha+\beta\right)} < 1 - \frac{\bar{m}-\alpha}{\beta}\left(\frac{\underline{k}}{\left(1+\alpha\right)} + \frac{\underline{k}-\bar{k}}{\left(\bar{m}-\alpha+\beta\right)}\right) \Leftrightarrow$$

$$\frac{\bar{m}-\alpha+\beta}{\beta}\frac{\underline{k}-\bar{k}}{\left(\bar{m}-\alpha+\beta\right)} < \frac{\bar{m}-\alpha}{\left(\bar{m}-\alpha+\beta\right)} - \frac{\bar{m}-\alpha}{\beta}\frac{\underline{k}}{\left(1+\alpha\right)} \Leftrightarrow$$

$$\frac{\underline{k}-\bar{k}}{\beta} < \frac{\bar{m}-\alpha}{\left(\bar{m}-\alpha+\beta\right)} - \frac{\bar{m}-\alpha}{\left(1+\alpha\right)}\frac{\underline{k}}{\beta} \Leftrightarrow \frac{1+\bar{m}}{1+\alpha}\frac{\underline{k}}{\beta} - \frac{\bar{k}}{\beta} < \frac{\bar{m}-\alpha}{\left(\bar{m}-\alpha+\beta\right)} \Leftrightarrow$$

$$\left(1+\bar{m}\right)\underline{k} - \left(1+\alpha\right)\bar{k} < \frac{\bar{m}-\alpha}{\left(\bar{m}-\alpha+\beta\right)}\left(1+\alpha\right)\beta \Leftrightarrow \left(1+\alpha\right)\beta > \frac{\bar{m}-\alpha+\beta}{\left(\bar{m}-\alpha\right)}\left(\left(1+\bar{m}\right)\underline{k} - \left(1+\alpha\right)\bar{k}\right)$$

It turns out that this conditions is stronger than 3b), because

$$\frac{\bar{m}-\alpha+\beta}{\left(\bar{m}-\alpha\right)}\left(\left(1+\bar{m}\right)\underline{k} - \left(1+\alpha\right)\bar{k}\right) - \left(\left(1+\beta+\bar{m}\right)\underline{k} - \left(1+\alpha\right)\bar{k}\right) > 0 \Leftrightarrow \underline{k} - \bar{k} > 0$$

Hence we are left with the following conditions for existence:

1. $(1+\alpha)\beta > \dfrac{\bar{m}-\alpha+\beta}{(\bar{m}-\alpha)}\left((1+\bar{m})\underline{k}-(1+\alpha)\bar{k}\right)$

2. $\dfrac{\beta+\underline{k}-\bar{k}}{(\bar{m}-\alpha+\beta)} < \lambda \le 1-\dfrac{\bar{m}-\alpha}{\beta}\left(\dfrac{\underline{k}}{(1+\alpha)}+\dfrac{\underline{k}-\bar{k}}{(\bar{m}-\alpha+\beta)}\right) = 1-\dfrac{\bar{m}-\alpha}{\beta}\dfrac{(1+\beta+\bar{m})\underline{k}-(1+\alpha)\bar{k}}{(1+\alpha)(\bar{m}-\alpha+\beta)}$

3. $0 \le \mathrm{p}_{CD\bar{m}} \le \min\left\{\underbrace{1-\dfrac{\beta+\underline{k}-\bar{k}}{\lambda(\bar{m}-\alpha+\beta)}}_{x}, \underbrace{\dfrac{\beta+\bar{k}-\dfrac{1+\beta+\bar{m}}{1+\alpha}\underline{k}}{\lambda(\bar{m}-\alpha+\beta)}}_{y}\right\}$

$$x \ge y \Leftrightarrow \lambda \ge \frac{2\beta}{(\bar{m}-\alpha+\beta)} - \frac{\underline{k}}{1+\alpha}\left(\overset{(3)\mathrm{b})}{>}\frac{\beta+\underline{k}-\bar{k}}{(\bar{m}-\alpha+\beta)}\right)$$

2.2.1.2.2. $\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})<0 \Rightarrow p_{\underline{m}}=1$

$$\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m}) = \lambda(1+\alpha)\left[\left(p_{CD\bar{m}}+p_{CD\underline{m}}-p_{DC\bar{m}}-p_{DC\underline{m}}\right)\right]-\underline{k}<0 \tag{*}$$

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CD,\bar{m}) = \Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[\left(p_{CD\underline{m}}\right)(\bar{m}-\alpha)+\left(p_{DD\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]=0 \tag{i}$$

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(DC,\bar{m}) = \Pi_{\bar{m}}(CD,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{DC\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]= \tag{ii}$$

$$\Pi_{\bar{m}}(CD,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{CD\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]>0 \tag{iv}$$

$$\Pi_{\bar{m}}(CD,\bar{m})-\Pi_{\bar{m}}(CD,\underline{m}) = \Pi_{\bar{m}}(DD,\bar{m})-\Pi_{\bar{m}}(DD,\underline{m})+\lambda\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)(\bar{m}-\alpha+\beta) \tag{v}$$

$$=\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CC,\underline{m})-\lambda\left(p_{CD\underline{m}}\right)(\bar{m}-\alpha+\beta)$$

$$=\lambda\left[\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)(1+\bar{m}+\beta)+\left(p_{CD\underline{m}}\right)(1+\alpha)\right]-\bar{k}$$

$$=\lambda(1+\alpha)\left[\left(p_{CD\bar{m}}+p_{CD\underline{m}}-p_{DC\bar{m}}\right)\right]-\bar{k}+\lambda\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)(\bar{m}-\alpha+\beta)=0$$

$$\lambda(1+\alpha)\left[\left(p_{CD\bar{m}}+p_{CD\underline{m}}-p_{DC\bar{m}}\right)\right]-\underline{k}<0 \quad (*)$$

$$\lambda\left[\left(p_{CD\underline{m}}\right)(\bar{m}-\alpha)\right]+(1-\lambda)(-\beta)=0 \quad \text{(i)}$$

$$\lambda\left[\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{DC\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]=0 \quad \text{(ii)}$$

$$\lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{CD\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)(-\beta)>0 \quad \text{(iv)}$$

$$\lambda(1+\alpha)\,p_{CD\underline{m}}=\bar{k}-\lambda(1+\bar{m}+\beta)\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right) \quad \text{(v)}$$

Then $p_{CD\underline{m}}=\dfrac{(1-\lambda)\beta}{\lambda(\bar{m}-\alpha)}$ by (i) and

$$p_{CD\bar{m}}-p_{DC\bar{m}}=\frac{\bar{k}}{\lambda(1+\bar{m}+\beta)}-\frac{(1+\alpha)}{(1+\bar{m}+\beta)}p_{CD\underline{m}}=\frac{\bar{k}}{\lambda(1+\bar{m}+\beta)}-\frac{(1-\lambda)}{\lambda}\frac{\beta(1+\alpha)}{(1+\bar{m}+\beta)(\bar{m}-\alpha)}$$ by (v).

Furthermore by (ii):

$$\lambda\left[\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{DC\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]=0\Leftrightarrow\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)(\bar{m}-\alpha+\beta)+\left(1-p_{CD\underline{m}}\right)(-\beta)=0\Leftrightarrow$$

$$\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)=\left(1-p_{CD\underline{m}}\right)\frac{\beta}{(\bar{m}-\alpha+\beta)}\Rightarrow p_{DC\bar{m}}+p_{DD\bar{m}}=1-\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)-p_{CD\underline{m}}=\left(1-p_{CD\underline{m}}\right)\frac{\bar{m}-\alpha}{(\bar{m}-\alpha+\beta)}$$

In summary the equilibrium set is given by:

$$\left\{ \left(p_{CC\bar{m}}, p_{CD\bar{m}}, p_{DC\bar{m}}, p_{DD\bar{m}}, p_{CD\underline{m}}, p_{\underline{m}} = 1 \right) \middle| \begin{array}{l} p_{CC\bar{m}} = \left(1 - p_{CD\underline{m}}\right) \dfrac{\beta}{(\bar{m} - \alpha + \beta)} - p_{CD\bar{m}} \\ p_{DC\bar{m}} = p_{CD\bar{m}} - \dfrac{\bar{k}}{\lambda(1 + \bar{m} + \beta)} + \dfrac{(1+\alpha)}{(1+\bar{m}+\beta)} p_{CD\underline{m}} \\ p_{DD\bar{m}} = \left(1 - p_{CD\underline{m}}\right) \dfrac{\bar{m} - \alpha}{(\bar{m} - \alpha + \beta)} - p_{DC\bar{m}} \\ p_{CD\underline{m}} = \dfrac{(1-\lambda)\beta}{\lambda(\bar{m} - \alpha)} \end{array} \right\}$$

(*) and (iv) remain to be checked:

(iv): $\bar{k}\left(\bar{m}-\alpha\right)-\left(1+\alpha\right)\beta\left(1-\lambda\right)<0 \Leftrightarrow \lambda < 1 - \dfrac{\bar{k}}{(1+\alpha)}\dfrac{(\bar{m}-\alpha)}{\beta}$

(*): $\lambda > 1 - \dfrac{(\bar{m}-\alpha)}{\beta}\left(\dfrac{(1+\beta+\bar{m})\underline{k} - (1+\alpha)\bar{k}}{(1+\alpha)(\bar{m}-\alpha+\beta)}\right)$

note that the lower bound is always smaller than the upper bound due to $\underline{k} > \bar{k}$

All conditions of the type

$p \in [0,1], \sum p < 1$ reduce to:

(1): $p_{CD\underline{m}} > 0$ true

(2): $p_{CD\underline{m}} < 1 \Leftrightarrow \lambda > \frac{\beta}{(\bar{m}-\alpha+\beta)}$

(3): $p_{CC\bar{m}} \geq 0: \frac{\beta}{\lambda(\bar{m}-\alpha+\beta)}\left(1-\frac{1-\lambda}{\bar{m}-\alpha}(\bar{m}-\alpha+\beta)\right) = \frac{\beta}{(\bar{m}-\alpha+\beta)}\left(1-\frac{(1-\lambda)\beta}{\lambda(\bar{m}-\alpha)}\right) > p_{CD\bar{m}}$

$p_{DC\bar{m}} \geq 0: p_{CD\bar{m}} > \frac{\bar{k}(\bar{m}-\alpha)-(1+\alpha)\beta(1-\lambda)}{\lambda(\bar{m}-\alpha)(1+\beta+\bar{m})}$

$p_{DD\bar{m}} \geq 0: p_{CD\bar{m}} < \left(1-\frac{(1-\lambda)\beta}{\lambda(\bar{m}-\alpha)}\right)\frac{\bar{m}-\alpha}{(\bar{m}-\alpha+\beta)} + \frac{\bar{k}}{\lambda(1+\bar{m}+\beta)} - \frac{1+\alpha}{1+\bar{m}+\beta}\frac{(1-\lambda)\beta}{\lambda(\bar{m}-\alpha)}$

$$\underbrace{\frac{\bar{k}}{\lambda(1+\bar{m}+\beta)} - \frac{1+\alpha}{1+\bar{m}+\beta}\frac{(1-\lambda)\beta}{\lambda(\bar{m}-\alpha)}}_{x} \leq p_{CD\bar{m}} \leq \min\left\{\underbrace{\left(1-\frac{(1-\lambda)\beta}{\lambda(\bar{m}-\alpha)}\right)\frac{\beta}{(\bar{m}-\alpha+\beta)}}_{y}, \underbrace{\left(1-\frac{(1-\lambda)\beta}{\lambda(\bar{m}-\alpha)}\right)\frac{\bar{m}-\alpha}{(\bar{m}-\alpha+\beta)} + \frac{\bar{k}}{\lambda(1+\bar{m}+\beta)} - \frac{1+\alpha}{1+\bar{m}+\beta}\frac{(1-\lambda)\beta}{\lambda(\bar{m}-\alpha)}}_{z}\right\}$$

a) $x > 0 \Leftrightarrow \lambda > 1 - \frac{\bar{k}}{(1+\alpha)}\frac{(\bar{m}-\alpha)}{\beta} \overset{(iv)}{\Rightarrow} x < 0$

b) $y > 0 \Leftrightarrow \lambda > \frac{\beta}{(\bar{m}-\alpha+\beta)}$

c) $z > 0 \Leftrightarrow \lambda > \frac{\beta(-\alpha(2+\alpha)+\beta+\bar{m}(2+\bar{m}+\beta)) - \bar{k}(\bar{m}-\alpha)(\bar{m}-\alpha+\beta)}{(1+\bar{m})(\bar{m}-\alpha+\beta)^2}$

$$z>0 \Leftrightarrow \left(\lambda-\frac{(1-\lambda)\beta}{(\bar{m}-\alpha)}\right)\frac{\bar{m}-\alpha}{(\bar{m}-\alpha+\beta)}+\frac{\bar{k}}{(1+\bar{m}+\beta)}-\frac{1+\alpha}{1+\bar{m}+\beta}\frac{(1-\lambda)\beta}{(\bar{m}-\alpha)}>0 \Leftrightarrow$$

$$\left(\lambda-\frac{\beta}{(\bar{m}-\alpha)}+\frac{\lambda\beta}{(\bar{m}-\alpha)}\right)\frac{\bar{m}-\alpha}{(\bar{m}-\alpha+\beta)}+\frac{\bar{k}}{(1+\bar{m}+\beta)}-\frac{1+\alpha}{1+\bar{m}+\beta}\frac{\beta}{(\bar{m}-\alpha)}+\frac{1+\alpha}{1+\bar{m}+\beta}\frac{\lambda\beta}{(\bar{m}-\alpha)}>0 \Leftrightarrow$$

$$\lambda\left(1+\frac{1+\alpha}{1+\bar{m}+\beta}\frac{\beta}{(\bar{m}-\alpha)}\right)-\frac{\beta}{(\bar{m}-\alpha+\beta)}+\frac{\bar{k}}{(1+\bar{m}+\beta)}-\frac{1+\alpha}{1+\bar{m}+\beta}\frac{\beta}{(\bar{m}-\alpha)}>0 \Leftrightarrow$$

$$\lambda\left(\frac{(1+\bar{m})(\bar{m}-\alpha+\beta)}{(1+\bar{m}+\beta)(\bar{m}-\alpha)}\right)>\frac{\beta}{(\bar{m}-\alpha+\beta)}+\frac{1+\alpha}{1+\bar{m}+\beta}\frac{\beta}{(\bar{m}-\alpha)}-\frac{\bar{k}}{(1+\bar{m}+\beta)}$$

Note that for $\frac{\bar{k}}{(1+\alpha)} \geq \frac{\beta}{(\bar{m}-\alpha+\beta)}$ the necessary condition $\frac{\beta}{(\bar{m}-\alpha+\beta)} \overset{3b)}{<} \lambda \overset{(iv)}{<} 1-\frac{\bar{k}}{(1+\alpha)}\frac{(\bar{m}-\alpha)}{\beta}$ gives an nonempty interval if and only if $\frac{\bar{k}}{(1+\alpha)} < \frac{\beta}{(\bar{m}-\alpha+\beta)}$, hence the case $\frac{\bar{k}}{(1+\alpha)} \geq \frac{\beta}{(\bar{m}-\alpha+\beta)}$ can be neglected. It turns out that for the necessary condition $\frac{\bar{k}}{(1+\alpha)} < \frac{\beta}{(\bar{m}-\alpha+\beta)}$, the condition 3c) is stronger than 3b)

Conditions for existence:

1. $\frac{\bar{k}}{(1+\alpha)} < \frac{\beta}{(\bar{m}-\alpha+\beta)}$

2. $0 \le p_{CD\bar{m}} \le \min\left\{ \underbrace{\left(1-\frac{(1-\lambda)\beta}{\lambda(\bar{m}-\alpha)}\right)\frac{\beta}{(\bar{m}-\alpha+\beta)}}_{y}, \underbrace{\left(1-\frac{(1-\lambda)\beta}{\lambda(\bar{m}-\alpha)}\right)\frac{\bar{m}-\alpha}{(\bar{m}-\alpha+\beta)}+\frac{\bar{k}}{\lambda(1+\bar{m}+\beta)}-\frac{1+\alpha}{1+\bar{m}+\beta}\frac{(1-\lambda)\beta}{\lambda(\bar{m}-\alpha)}}_{z} \right\}$

3. $\max\left\{ \underbrace{\frac{\beta\left(-\alpha(2+\alpha)+\beta+\bar{m}(2+\bar{m}+\beta)\right)-\bar{k}(\bar{m}-\alpha)(\bar{m}-\alpha+\beta)}{(1+\bar{m})(\bar{m}-\alpha+\beta)^2}}_{z>0}, \underbrace{1-\frac{(\bar{m}-\alpha)}{\beta(\bar{m}-\alpha+\beta)}\left((1+\bar{m}+\beta)\frac{\underline{k}}{(1+\alpha)}-\bar{k}\right)}_{(*)} \right\} < \lambda \underset{(iv)}{<} 1-\frac{\bar{k}}{(1+\alpha)}\frac{(\bar{m}-\alpha)}{\beta}$

Note that the interval defined by 3. is non-empty due to 1.

2.2.2. $\Pi_{\bar{m}}(DD,\underline{m}) > \Pi_{\bar{m}}(CD,\underline{m})$, i.e. $p_{CC\underline{m}} = p_{CD\underline{m}} = p_{DC\underline{m}} = 0$

$\Pi_{\bar{m}}(DD,\bar{m}) - \Pi_{\bar{m}}(DD,\underline{m}) = \lambda(1+\alpha)\left[\left(p_{CD\bar{m}} + p_{CD\underline{m}} - p_{DC\bar{m}} - p_{DC\underline{m}}\right)\right] - \bar{k} = 0$ is violated.

Therefor such an equilibrium cannot exist.

3. <u>All four non-signaling strategies earn the same payoffs , i.e.</u> $\Pi_{\bar{m}}(CC,\underline{m}) = \Pi_{\bar{m}}(CD,\underline{m}) = \Pi_{\bar{m}}(DC,\underline{m}) = \Pi_{\bar{m}}(DD,\underline{m})$

3.1. $CC,\underline{m}/DC,\underline{m}/CD,\underline{m}/DD,\underline{m}$ *vs.* $CC,\bar{m}/DC,\bar{m}$ i.e. $\Pi_{\bar{m}}(CC,\bar{m}) - \Pi_{\bar{m}}(CD,\bar{m}) = \Pi_{\bar{m}}(DC,\bar{m}) - \Pi_{\bar{m}}(DD,\bar{m}) > 0 \quad p_{CD\bar{m}} = p_{DD\bar{m}} = 0$

$$\Pi_{\underline{m}}(\bar{m}) - \Pi_{\underline{m}}(\underline{m}) = \underbrace{\Pi_{\bar{m}}(DD,\bar{m}) - \Pi_{\bar{m}}(DD,\underline{m})}_{\leq 0} + \bar{k} - \underline{k} < 0 \Rightarrow p_{\bar{m}} = 0$$

$$\Pi_{\bar{m}}(CC,\underline{m}) - \Pi_{\bar{m}}(DC,\underline{m}) = \Pi_{\bar{m}}(CD,\underline{m}) - \Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[(p_{CC\bar{m}} + p_{DC\bar{m}})(\bar{m}-\alpha) + (p_{CD\bar{m}} + p_{DD\bar{m}})(-\beta)\right] + (1-\lambda)\left[p_{\bar{m}}(-\beta)\right] = 0$$

(i)

(i) becomes: $\Pi_{\bar{m}}(CC,\underline{m}) - \Pi_{\bar{m}}(DC,\underline{m}) = \Pi_{\bar{m}}(CD,\underline{m}) - \Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[(p_{CC\bar{m}} + p_{DC\bar{m}})(\bar{m}-\alpha)\right] = 0$ and implies that $(p_{CC\bar{m}} + p_{DC\bar{m}}) = 0$ however a **semi**-pooling equilibrium requires strict positivity for at least one of the shares.

Therefor such an equilibrium cannot exist.

3.2. $CC,\underline{m}/DC,\underline{m}/CD,\underline{m}/DD,\underline{m}$ *vs.* $CD,\bar{m}/DD,\bar{m}$, i.e. $\Pi_{\bar{m}}(CC,\bar{m}) - \Pi_{\bar{m}}(CD,\bar{m}) = \Pi_{\bar{m}}(DC,\bar{m}) - \Pi_{\bar{m}}(DD,\bar{m}) < 0 \quad p_{CC\bar{m}} = p_{DC\bar{m}} = 0$

3.2.1. $\Pi_{\bar{m}}(DD,\bar{m}) > \Pi_{\bar{m}}(CD,\bar{m})$, i.e. $p_{CC\bar{m}} = p_{DC\bar{m}} = p_{CD\bar{m}} = 0$

Then $\Pi_{\bar{m}}(CD,\bar{m}) - \Pi_{\bar{m}}(CD,\underline{m}) = \underbrace{\Pi_{\bar{m}}(DD,\bar{m}) - \Pi_{\bar{m}}(DD,\underline{m})}_{=0} + \lambda(p_{CD\bar{m}} - p_{DC\bar{m}})(\bar{m}-\alpha+\beta) = 0$ violates $\Pi_{\bar{m}}(CD,\underline{m}) = \Pi_{\bar{m}}(DD,\bar{m}) > \Pi_{\bar{m}}(CD,\bar{m})$

3.2.2. $\Pi_{\bar{m}}(DD,\bar{m}) \leq \Pi_{\bar{m}}(CD,\bar{m})$, i.e. $p_{CC\bar{m}} = p_{DC\bar{m}} = 0$

$$\Pi_{\underline{m}}(\bar{m}) - \Pi_{\underline{m}}(\underline{m}) = \underbrace{\Pi_{\bar{m}}(DD,\bar{m}) - \Pi_{\bar{m}}(DD,\underline{m})}_{\leq 0} + \bar{k} - \underline{k} < 0 \Rightarrow p_{\bar{m}} = 0$$

$$\Pi_{\bar{m}}\left(CC,\underline{m}\right)-\Pi_{\bar{m}}\left(CD,\underline{m}\right)=\Pi_{\bar{m}}\left(DC,\underline{m}\right)-\Pi_{\bar{m}}\left(DD,\underline{m}\right)=\lambda\left[\left(p_{CC\underline{m}}+p_{DC\underline{m}}\right)\left(\bar{m}-\alpha\right)+\left(p_{CD\underline{m}}+p_{DD\underline{m}}\right)\left(-\beta\right)\right]+\left(1-\lambda\right)\left[p_{\underline{m}}\left(-\beta\right)\right]= \quad \text{(i)}$$

$$\Pi_{\bar{m}}\left(CC,\underline{m}\right)-\Pi_{\bar{m}}\left(DC,\underline{m}\right)=\Pi_{\bar{m}}\left(CD,\underline{m}\right)-\Pi_{\bar{m}}\left(DD,\underline{m}\right)=\lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)\left(\bar{m}-\alpha\right)+\left(p_{CD\bar{m}}+p_{DD\bar{m}}\right)\left(-\beta\right)\right]+\left(1-\lambda\right)\left[p_{\bar{m}}\left(-\beta\right)\right]= \quad \text{(ii)}$$

$$\Pi_{\bar{m}}\left(DD,\bar{m}\right)-\Pi_{\bar{m}}\left(DD,\underline{m}\right)=\lambda\left(1+\alpha\right)\left[\left(p_{CD\bar{m}}+p_{CD\underline{m}}-p_{DC\bar{m}}-p_{DC\underline{m}}\right)\right]-\bar{k}\leq 0 \quad \text{(iii)}$$

$$\Pi_{\bar{m}}\left(CC,\bar{m}\right)-\Pi_{\bar{m}}\left(CD,\bar{m}\right)=\Pi_{\bar{m}}\left(DC,\bar{m}\right)-\Pi_{\bar{m}}\left(DD,\bar{m}\right)=\lambda\left[\left(p_{CC\underline{m}}+p_{CD\underline{m}}\right)\left(\bar{m}-\alpha\right)+\left(p_{DC\underline{m}}+p_{DD\underline{m}}\right)\left(-\beta\right)\right]+\left(1-\lambda\right)\left[p_{\underline{m}}\left(-\beta\right)\right]\leq \quad \text{(iv)}$$

$$\Pi_{\bar{m}}\left(CD,\bar{m}\right)-\Pi_{\bar{m}}\left(CD,\underline{m}\right)=\Pi_{\bar{m}}\left(DD,\bar{m}\right)-\Pi_{\bar{m}}\left(DD,\underline{m}\right)+\lambda\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)\left(\bar{m}-\alpha+\beta\right)= \quad \text{(v)}$$

$$\Pi_{\bar{m}}\left(CC,\bar{m}\right)-\Pi_{\bar{m}}\left(CC,\underline{m}\right)-\lambda\left(p_{CD\underline{m}}-p_{DC\underline{m}}\right)\left(\bar{m}-\alpha+\beta\right)=0$$

a) If (iii) holds with equality then the last equality implies that $p_{CD\bar{m}}-p_{DC\bar{m}}=0\Rightarrow p_{CD\bar{m}}=0$ then (ii) becomes: $\Pi_{\bar{m}}\left(CC,\underline{m}\right)-\Pi_{\bar{m}}\left(DC,\underline{m}\right)=\Pi_{\bar{m}}\left(CD,\underline{m}\right)-\Pi_{\bar{m}}\left(DD,\underline{m}\right)=\lambda\left[\left(p_{DD\bar{m}}\right)\left(-\beta\right)\right]=0\Rightarrow p_{DD\bar{m}}=0$ Hence such a **semi**-pooling equilibrium cannot exist.

b) If (iii) holds as a strict inequality then $p_{DD\bar{m}}=0$ and (ii) becomes $\Pi_{\bar{m}}\left(CC,\underline{m}\right)-\Pi_{\bar{m}}\left(DC,\underline{m}\right)=\lambda\left[\left(p_{CD\bar{m}}\right)\left(-\beta\right)\right]=0$ which holds only for $p_{CD\bar{m}}=0$, i.e. in a pooling but not **semi**-pooling equilibrium.

4. 1-2 strategies versus 1-2 strategies

In general there are 36 possible matchings: CC, DC, CC and DC, CD, DD, CD and DD , six for each signaling strategy, however as summarized in Table B-3 we excluded 21 of them; in the following we consider the remaining 14 cases:

	$CC,\underline{m}$	$DC,\underline{m}$	$CC,\underline{m}/DC,\underline{m}$	$CD,\underline{m}$	$DD,\underline{m}$	$CD,\underline{m}/DD,\underline{m}$
$CC,\bar{m}$	N (2.)	N (4.1.1)	N (2.)	(4.1.3.)	N (4.1.2.)	N (4.1.2.)
$DC,\bar{m}$	N (7.)	N (3.)	N (3.)	(4.2.3.2./4.2.3.3.)	N (4.2.1.)	N (4.2.2.)
$CC,\bar{m}/DC,\bar{m}$	N (2.)	N (3.)	N (3.)	(4.3.1.)	N (6.)	N (4.3.2.)
$CD,\bar{m}$	N (4.)	(4.4.1.)	N (4.)	N (5.)	(4.4.2.)	N (5.)
$DD,\bar{m}$	N (4.)	N (4.5.)	N (4.)	N (5.)	N (1.)	N (5.)
$CD,\bar{m}/DD,\bar{m}$	N (4.)	(4.6.1.)	N (4.)	N (5.)	(4.6.2.)	N (5.)

Table B-3: Overview of subcases; N: non-existence of the considered equilibrium; number in parenthesis either refers to the list of payoff differences below Table B-2 or subsection dealing with the corresponding case.

4.1. $CC,\bar{m}$

4.1.1. and $DC,\underline{m}$ ($p_{\bar{m}}=0$)

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(DC,\underline{m})=\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(DC,\underline{m})-\left[\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(CC,\bar{m})\right]= \tag{i}$$

$$\lambda\left[\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)(1+\alpha)+\left(p_{CD\underline{m}}-p_{DC\underline{m}}\right)(1+\bar{m}+\beta)\right]-\bar{k}+\lambda\left[\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{DC\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]$$

$$=\lambda\left[\left(p_{CD\bar{m}}\right)(1+\bar{m})-p_{DC\bar{m}}(1+\alpha+\beta)+\left(p_{CC\bar{m}}\right)(\bar{m}-\alpha)-p_{DD\bar{m}}\beta+\left(p_{CD\underline{m}}-p_{DC\underline{m}}\right)(1+\bar{m}+\beta)\right]-\bar{k}+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]=0$$

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(DC,\bar{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{DC\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]>0 \tag{ii}$$

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CD,\bar{m})=\lambda\left[\left(p_{CC\underline{m}}+p_{CD\underline{m}}\right)(\bar{m}-\alpha)+\left(p_{DC\underline{m}}+p_{DD\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]>0 \tag{iii}$$

$$\Pi_{\bar{m}}(CC,\underline{m})-\Pi_{\bar{m}}(DC,\underline{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{CD\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]<0 \tag{iv}$$

$$\Pi_{\bar{m}}(DC,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[\left(p_{CC\underline{m}}+p_{DC\underline{m}}\right)(\bar{m}-\alpha)+\left(p_{CD\underline{m}}+p_{DD\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]>0 \tag{v}$$

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(DC,\underline{m})=\lambda\left[p_{CC\bar{m}}(\bar{m}-\alpha)-p_{DC\underline{m}}(1+\bar{m}+\beta)\right]-\bar{k}=0 \quad \text{(i)}$$
$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(DC,\bar{m})=\lambda\left[\left(p_{CC\bar{m}}\right)(\bar{m}-\alpha)\right]>0 \quad \text{(ii)}$$
$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CD,\bar{m})=\lambda\left[\left(p_{DC\underline{m}}\right)(-\beta)\right]-\beta(1-\lambda)>0 \quad \text{(iii)}$$
$$\Pi_{\bar{m}}(CC,\underline{m})-\Pi_{\bar{m}}(DC,\underline{m})=\lambda\left[\left(p_{CC\bar{m}}\right)(\bar{m}-\alpha)\right]<0 \quad \text{(iv)}$$
$$\Pi_{\bar{m}}(DC,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[\left(p_{DC\underline{m}}\right)(\bar{m}-\alpha)\right]-\beta(1-\lambda)>0 \quad \text{(v)}$$

By (iii) such a semi-pooling equilibrium cannot exist.

4.1.2. and $DD,\underline{m}$ or $CD,\underline{m}/DD,\underline{m}$ ($p_{\bar{m}}=0$)

$$\Pi_{\bar{m}}(CD,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{CD\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]\leq 0 \quad \text{(i)}$$
$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(DC,\bar{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{DC\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]>0 \quad \text{(ii)}$$
$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CD,\bar{m})=\lambda\left[\left(p_{CC\underline{m}}+p_{CD\underline{m}}\right)(\bar{m}-\alpha)+\left(p_{DC\underline{m}}+p_{DD\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]>0 \quad \text{(iii)}$$
$$\Pi_{\bar{m}}(CC,\underline{m})-\Pi_{\bar{m}}(CD,\underline{m})=\lambda\left[\left(p_{CC\underline{m}}+p_{DC\underline{m}}\right)(\bar{m}-\alpha)+\left(p_{CD\underline{m}}+p_{DD\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]<0 \quad \text{(iv)}$$
$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(DD,\underline{m})=0 \quad \text{(v)}$$

$\rightarrow$

$$\Pi_{\bar{m}}(CD,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[\left(p_{CC\bar{m}}\right)(\bar{m}-\alpha)\right]\leq 0 \quad \text{(i)}$$
$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(DC,\bar{m})=\lambda\left[\left(p_{CC\bar{m}}\right)(\bar{m}-\alpha)\right]>0 \quad \text{(ii)}$$
$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CD,\bar{m})=\lambda\left[\left(p_{CD\underline{m}}\right)(\bar{m}-\alpha)+\left(p_{DD\underline{m}}\right)(-\beta)\right]-\beta(1-\lambda)>0 \quad \text{(iii)}$$
$$\Pi_{\bar{m}}(CC,\underline{m})-\Pi_{\bar{m}}(CD,\underline{m})=\lambda\left[\left(p_{CC\underline{m}}\right)(\bar{m}-\alpha)+\left(p_{DD\underline{m}}\right)(-\beta)\right]-\beta(1-\lambda)<0 \quad \text{(iv)}$$
$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(DD,\underline{m})=0 \quad \text{(v)}$$

By (i) such a **semi**-pooling equilibrium cannot exist.

4.1.3. and $CD, \underline{m}$

$$\Pi_{\bar{m}}(CD,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{CD\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]>0 \quad \text{(i)}$$

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(DC,\bar{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{DC\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]>0 \quad \text{(ii)}$$

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CD,\bar{m})=\lambda\left[\left(p_{CC\underline{m}}+p_{CD\underline{m}}\right)(\bar{m}-\alpha)+\left(p_{DC\underline{m}}+p_{DD\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]>0 \quad \text{(iii)}$$

$$\Pi_{\bar{m}}(CC,\underline{m})-\Pi_{\bar{m}}(CD,\underline{m})=\lambda\left[\left(p_{CC\underline{m}}+p_{DC\underline{m}}\right)(\bar{m}-\alpha)+\left(p_{CD\underline{m}}+p_{DD\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]<0 \quad \text{(iv)}$$

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CD,\underline{m}) \quad \text{(v)}$$

$$=\lambda\left[\left(p_{CD\underline{m}}\right)(1+\bar{m})+\left(p_{CC\underline{m}}\right)(\bar{m}-\alpha)+\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)(1+\bar{m}+\beta)-p_{DC\underline{m}}(1+\alpha+\beta)-p_{DD\underline{m}}\beta\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]-\bar{k}$$

$$=\lambda\left[\left(p_{CD\underline{m}}\right)(1+\bar{m})+\left(-p_{DC\bar{m}}\right)(1+\bar{m}+\beta)-p_{DD\underline{m}}\beta\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]-\bar{k}$$

$$=\lambda\left[\left(p_{CD\underline{m}}\right)(1+\bar{m})\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]-\bar{k}=0$$

→

$$\Pi_{\bar{m}}(CD,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[\left(p_{CC\bar{m}}\right)(\bar{m}-\alpha)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]>0 \quad \text{(i)}$$

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(DC,\bar{m})=\lambda\left[\left(p_{CC\bar{m}}\right)(\bar{m}-\alpha)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]>0 \quad \text{(ii)}$$

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CD,\bar{m})=\lambda\left[\left(p_{CD\underline{m}}\right)(\bar{m}-\alpha)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]>0 \quad \text{(iii)}$$

$$\Pi_{\bar{m}}(CC,\underline{m})-\Pi_{\bar{m}}(CD,\underline{m})=\lambda\left[\left(p_{CD\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]<0 \quad \text{(iv)}$$

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CD,\underline{m})=\lambda\left[\left(p_{CD\underline{m}}\right)(1+\bar{m})\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]-\bar{k}=0 \quad \text{(v)}$$

The sum of (ii) and (iii) implies: $\lambda(\bar{m}-\alpha+\beta)>\beta$

$$\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})=\lambda\left[\left(p_{CD\bar{m}}+p_{CD\underline{m}}-p_{DC\bar{m}}-p_{DC\underline{m}}\right)(1+\alpha)\right]-\underline{k}=\lambda p_{CD\underline{m}}(1+\alpha)-\underline{k} \quad \text{(vi)}$$

4.1.3.1. $\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})>0 \Leftrightarrow p_{CD\underline{m}} > \frac{\underline{k}}{\lambda(1+\alpha)} (\Rightarrow p_{\underline{m}}=0)$

Then $p_{CD\underline{m}} = \frac{\bar{k}}{\lambda(1+\bar{m})}$, $p_{CC\bar{m}} = 1-\frac{\bar{k}}{\lambda(1+\bar{m})}$ by (v) and (iii) is satisfied and $\frac{(1-\lambda)\beta}{\lambda(\bar{m}-\alpha)} < p_{CC\bar{m}}$ by (ii)

Since $\frac{\bar{k}}{\lambda(1+\bar{m})} < \frac{\underline{k}}{\lambda(1+\alpha)}$ such an equilibrium cannot exist.

4.1.3.2. $\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})<0 \Leftrightarrow p_{CD\underline{m}} < \frac{\underline{k}}{\lambda(1+\alpha)} (\Rightarrow p_{\underline{m}}=1)$

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CD,\bar{m}) = \lambda\left[\left(p_{CD\underline{m}}\right)(\bar{m}-\alpha)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]>0 \quad \text{(iii)}$$

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CD,\underline{m}) = \lambda\left[\left(p_{CD\underline{m}}\right)(1+\bar{m})\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]-\bar{k}=0 \quad \text{(v)}$$

$$\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m}) = \lambda\left[\left(p_{CD\bar{m}}+p_{CD\underline{m}}-p_{DC\bar{m}}-p_{DC\underline{m}}\right)(1+\alpha)\right]-\underline{k} = \lambda p_{CD\underline{m}}(1+\alpha)-\underline{k}<0 \quad \text{(vi)}$$

Then $p_{CD\underline{m}} = \frac{\bar{k}+(1-\lambda)\beta}{\lambda(1+\bar{m})}$, $p_{CC\bar{m}} = 1-\frac{\bar{k}+(1-\lambda)\beta}{\lambda(1+\bar{m})}$ by (v) and (ii) is satisfied.

Conditions (iii) and (vi) need to be checked.

(ii): $\lambda > \frac{\bar{k}+\beta}{(1+\bar{m}+\beta)}$

(iii): $\lambda > 1-\frac{(\bar{m}-\alpha)}{\beta}\frac{\bar{k}}{(1+\alpha)}$

(vi): $\lambda > 1-\frac{1}{\beta}\left(\frac{1+\bar{m}}{(1+\alpha)}\underline{k}-\bar{k}\right)$

It turns out, that (iii) implies (vi), hence we are left with:

$$\text{(ii)}\wedge\text{(iii)}: \lambda > \max\left\{\frac{\bar{k}+\beta}{(1+\bar{m}+\beta)}, 1-\frac{(\bar{m}-\alpha)}{\beta}\frac{\bar{k}}{(1+\alpha)}\right\}$$

$$= \begin{cases} \frac{\bar{k}+\beta}{(1+\bar{m}+\beta)} & , \frac{\bar{k}}{(1+\alpha)} > \frac{\beta}{(\bar{m}-\alpha+\beta)} \\ 1-\frac{(\bar{m}-\alpha)}{\beta}\frac{\bar{k}}{(1+\alpha)} & , \frac{\bar{k}}{(1+\alpha)} \leq \frac{\beta}{(\bar{m}-\alpha+\beta)} \end{cases}$$

Finally all condition of the type:

$p \in [0,1], \sum p<1$ reduce to:

(1): $p_{CC\bar{m}} < 1$ true

(2): $p_{CC\bar{m}} > 0 \Leftrightarrow \lambda > \frac{\beta+\bar{k}}{1+\bar{m}+\beta}$

, hence $\max\left\{1-\frac{(\bar{m}-\alpha)}{\beta}\frac{\bar{k}}{(1+\alpha)}, \frac{\beta+\bar{k}}{1+\bar{m}+\beta}\right\} < \lambda < 1$, hence we need $\frac{\beta+\bar{k}}{1+\bar{m}+\beta} < 1 \Leftrightarrow \bar{k} < 1+\bar{m}$

To summarize: $\left\{ p_{\underline{m}} = 1, p_{CD\underline{m}} = \frac{\bar{k} + (1-\lambda)\beta}{\lambda(1+\bar{m})}, p_{CC\bar{m}} = 1 - \frac{\bar{k} + (1-\lambda)\beta}{\lambda(1+\bar{m})} \right\}$

Conditions for existence:

1. $(1+\alpha)\frac{\beta}{(\bar{m}-\alpha+\beta)} < \bar{k} < 1+\bar{m}$: $\frac{\bar{k}+\beta}{(1+\bar{m}+\beta)} < \lambda < 1$

2. $\bar{k} \le (1+\alpha)\frac{\beta}{(\bar{m}-\alpha+\beta)}$: $1 - \frac{(\bar{m}-\alpha)}{\beta}\frac{\bar{k}}{(1+\alpha)} < \lambda < 1$

4.1.3.3. $\Pi_{\underline{m}}(\bar{m}) - \Pi_{\underline{m}}(\underline{m}) = 0$

then $p_{CD\underline{m}} = \frac{\underline{k}}{\lambda(1+\alpha)}, p_{CC\bar{m}} = 1 - \frac{\underline{k}}{\lambda(1+\alpha)}$ by (vi) and shares for low types are given by (v): $\frac{\frac{(1+\bar{m})}{(1+\alpha)}\underline{k} - \bar{k}}{(1-\lambda)\beta} = p_{\underline{m}}$

furthermore by (ii) and (iii): $\frac{(1-\lambda)p_{\bar{m}}\beta}{\lambda(\bar{m}-\alpha)} < p_{CC\bar{m}} < 1 - \frac{(1-\lambda)p_{\underline{m}}\beta}{\lambda(\bar{m}-\alpha)}$

$$\frac{(1-\lambda)\beta-\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}}{\lambda(\bar{m}-\alpha)}<1-\frac{\underline{k}}{\lambda(1+\alpha)}<1-\frac{\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}}{\lambda(\bar{m}-\alpha)}\Leftrightarrow$$

$$\frac{(1-\lambda)\beta-\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}}{(\bar{m}-\alpha)}<\lambda-\frac{\underline{k}}{(1+\alpha)}<\lambda-\frac{\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}}{(\bar{m}-\alpha)}\Leftrightarrow$$

$$(1-\lambda)\beta-\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}-\lambda(\bar{m}-\alpha)<-\frac{(\bar{m}-\alpha)}{(1+\alpha)}\underline{k}<-\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}+\bar{k}\Leftrightarrow$$

$$\beta-\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}-\lambda(\bar{m}-\alpha+\beta)+\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}<\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\frac{(\bar{m}-\alpha)}{(1+\alpha)}\underline{k}<+\bar{k}\Leftrightarrow$$

$$\beta-\bar{k}-\lambda(\bar{m}-\alpha+\beta)<\underline{k}<\bar{k}$$

The last inequality is violated; hence such an equilibrium cannot exist.

4.2. $DC,\bar{m}$

4.2.1. and $DD,\underline{m}$ ($p_{\bar{m}}=0$)

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(DC,\bar{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{DC\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]<0 \tag{i}$$

$$\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left[\left(p_{CC\underline{m}}+p_{CD\underline{m}}\right)(\bar{m}-\alpha)+\left(p_{DC\underline{m}}+p_{DD\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]>0 \tag{ii}$$

$$\Pi_{\bar{m}}(CC,\underline{m})-\Pi_{\bar{m}}(CD,\underline{m})=\Pi_{\bar{m}}(DC,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[\left(p_{CC\underline{m}}+p_{DC\underline{m}}\right)(\bar{m}-\alpha)+\left(p_{CD\underline{m}}+p_{DD\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]<0 \tag{iii}$$

$$\Pi_{\bar{m}}(CD,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{CD\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]<0 \tag{iv}$$

$$\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(DD,\underline{m})=0 \tag{v}$$

$$\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left[\left(p_{DD\underline{m}}\right)(-\beta)\right]-\beta(1-\lambda)>0 \quad \text{(ii)}$$

B

By (ii) such a semi-pooling equilibrium cannot exist.

4.2.2. and $CD,\underline{m}/DD,\underline{m}$

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(DC,\bar{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{DC\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]<0 \quad \text{(i)}$$

$$\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left[\left(p_{CC\underline{m}}+p_{CD\underline{m}}\right)(\bar{m}-\alpha)+\left(p_{DC\underline{m}}+p_{DD\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]>0 \quad \text{(ii)}$$

$$\Pi_{\bar{m}}(CC,\underline{m})-\Pi_{\bar{m}}(CD,\underline{m})=\Pi_{\bar{m}}(DC,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[\left(p_{CC\underline{m}}+p_{DC\underline{m}}\right)(\bar{m}-\alpha)+\left(p_{CD\underline{m}}+p_{DD\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]<0 \quad \text{(iii)}$$

$$\Pi_{\bar{m}}(CD,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{CD\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]=0 \quad \text{(iv)}$$

$$\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(DD,\underline{m})=0 \quad \text{(v)}$$

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(DC,\bar{m})=\lambda p_{DC\bar{m}}(-\beta)+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]<0 \quad \text{(i)}$$

$$\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left[\left(p_{CD\underline{m}}\right)(\bar{m}-\alpha)+\left(p_{DD\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]>0 \quad \text{(ii)}$$

$$\Pi_{\bar{m}}(CC,\underline{m})-\Pi_{\bar{m}}(CD,\underline{m})=\Pi_{\bar{m}}(DC,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[\left(p_{CD\underline{m}}+p_{DD\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]<0 \quad \text{(iii)}$$

$$\Pi_{\bar{m}}(CD,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda p_{DC\bar{m}}(\bar{m}-\alpha)+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]=0 \quad \text{(iv)}$$

$$\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(DD,\underline{m})=\Pi_{\bar{m}}(DD,\bar{m})-\Pi_{\bar{m}}(DD,\underline{m})-\left[\Pi_{\bar{m}}(DD,\bar{m})-\Pi_{\bar{m}}(DC,\bar{m})\right]=\lambda(1+\alpha)\left[\left(p_{CD\bar{m}}+p_{CD\underline{m}}-p_{DC\bar{m}}-p_{DC\underline{m}}\right)\right]-\bar{k} \quad \text{(v)}$$

$$+\lambda\left[\left(p_{CC\underline{m}}+p_{CD\underline{m}}\right)(\bar{m}-\alpha)+\left(p_{DC\underline{m}}+p_{DD\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]=0$$

(iv) requires $p_{\bar{m}}>0$, i.e.

$$\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})=\lambda\left[\left(p_{CD\bar{m}}+p_{CD\underline{m}}-p_{DC\bar{m}}-p_{DC\underline{m}}\right)(1+\alpha)\right]-\underline{k}=\lambda\left[\left(p_{CD\underline{m}}-p_{DC\bar{m}}\right)(1+\alpha)\right]-\underline{k}\geq 0 \quad (*)$$

(i)and (iii) are always satisfied, note that (*) and (v) violate (ii):

$\lambda\left[\left(p_{CD\underline{m}} - p_{DC\bar{m}}\right)(1+\alpha)\right] \geq \underline{k}$ by (*) and $\underbrace{\lambda(1+\alpha)\left[\left(p_{CD\underline{m}} - p_{DC\bar{m}}\right)\right] - \bar{k}}_{>0} + \lambda\left[\left(p_{CD\underline{m}}\right)(\bar{m}-\alpha) + \left(p_{DD\underline{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\underline{m}}(-\beta)\right] = 0$ by (v) ,

hence $\lambda\left[\left(p_{CD\underline{m}}\right)(\bar{m}-\alpha) + \left(p_{DD\underline{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\underline{m}}(-\beta)\right] < 0$

Hence such a semi-pooling equilibrium cannot exist.

4.2.3. and $CD, \underline{m}$

$$\Pi_{\bar{m}}(CC,\bar{m}) - \Pi_{\bar{m}}(DC,\bar{m}) = \lambda\left[\left(p_{DC\bar{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\bar{m}}(-\beta)\right] < 0 \quad \text{(i)}$$

$$\Pi_{\bar{m}}(DC,\bar{m}) - \Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[\left(p_{CD\underline{m}}\right)(\bar{m}-\alpha) + \left(p_{DD\underline{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\underline{m}}(-\beta)\right] > 0 \quad \text{(ii)}$$

$$\Pi_{\bar{m}}(CC,\underline{m}) - \Pi_{\bar{m}}(CD,\underline{m}) = \Pi_{\bar{m}}(DC,\underline{m}) - \Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[\left(p_{CD\underline{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\underline{m}}(-\beta)\right] < 0 \quad \text{(iii)}$$

$$\Pi_{\bar{m}}(CD,\underline{m}) - \Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[\left(p_{DC\bar{m}}\right)(\bar{m}-\alpha)\right] + (1-\lambda)\left[p_{\bar{m}}(-\beta)\right] > 0 \quad \text{(iv)}$$

$$\Pi_{\bar{m}}(DC,\bar{m}) - \Pi_{\bar{m}}(CD,\underline{m}) = \Pi_{\bar{m}}(CD,\bar{m}) - \Pi_{\bar{m}}(CD,\underline{m}) - \left[\Pi_{\bar{m}}(CD,\bar{m}) - \Pi_{\bar{m}}(DC,\bar{m})\right] = 0 \quad \text{(v)}$$

(i)and (iii) are always satisfied

$$\Pi_{\bar{m}}(DC,\bar{m}) - \Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[\left(p_{CD\underline{m}}\right)(\bar{m}-\alpha)\right] + (1-\lambda)\left[p_{\underline{m}}(-\beta)\right] > 0 \quad \text{(ii)}$$

$$\Pi_{\bar{m}}(CD,\underline{m}) - \Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[\left(p_{DC\bar{m}}\right)(\bar{m}-\alpha)\right] + (1-\lambda)\left[p_{\bar{m}}(-\beta)\right] > 0 \quad \text{(iv)}$$

$$\Pi_{\bar{m}}(DC,\bar{m}) - \Pi_{\bar{m}}(CD,\underline{m}) = \lambda\left[\left(p_{CD\underline{m}} - p_{DC\bar{m}}\right)(1+\bar{m})\right] - \bar{k} - (1-\lambda)\left[p_{\bar{m}}(-\beta) - p_{\underline{m}}(-\beta)\right] = 0 \quad \text{(v)}$$

$$\Pi_{\underline{m}}(\bar{m}) - \Pi_{\underline{m}}(\underline{m}) = \lambda\left[\left(p_{CD\underline{m}} - p_{DC\bar{m}}\right)(1+\alpha)\right] - \underline{k} \quad \text{(*)}$$

4.2.3.1. $\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})>0$

$$\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left[\left(p_{CD\underline{m}}\right)(\bar{m}-\alpha)\right]>0 \quad \text{(ii)}$$
$$\Pi_{\bar{m}}(CD,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[\left(p_{DC\bar{m}}\right)(\bar{m}-\alpha)\right]+(1-\lambda)(-\beta)>0 \quad \text{(iv)}$$
$$\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(CD,\underline{m})=\lambda\left[\left(p_{CD\underline{m}}-p_{DC\bar{m}}\right)(1+\bar{m})\right]-\bar{k}+\beta(1-\lambda)=0 \quad \text{(v)}$$
$$\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})=\lambda\left[\left(p_{CD\underline{m}}-p_{DC\bar{m}}\right)(1+\alpha)\right]-\underline{k}>0\Rightarrow p_{\bar{m}}=1 \quad (*)$$

(ii) is satisfied; (*) violates (v) because

$$\lambda\left[\left(p_{CD\underline{m}}-p_{DC\bar{m}}\right)(1+\bar{m})\right]-\bar{k}+\beta(1-\lambda)=0\Rightarrow\lambda\left[\left(p_{CD\underline{m}}-p_{DC\bar{m}}\right)(1+\bar{m})\right]-\bar{k}<0\Rightarrow\lambda\left[\left(p_{CD\underline{m}}-p_{DC\bar{m}}\right)(1+\alpha)\right]-\underline{k}<0$$

Hence such a semi-pooling equilibrium cannot exist.

4.2.3.2. $\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})=0$

$$\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left[\left(p_{CD\underline{m}}\right)(\bar{m}-\alpha)\right]-\beta(1-\lambda)\,p_{\underline{m}}>0 \quad \text{(ii)}$$
$$\Pi_{\bar{m}}(CD,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[\left(p_{DC\bar{m}}\right)(\bar{m}-\alpha)\right]-\beta(1-\lambda)\,p_{\bar{m}}>0 \quad \text{(iv)}$$
$$\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(CD,\underline{m})=\lambda\left[\left(p_{CD\underline{m}}-p_{DC\bar{m}}\right)(1+\bar{m})\right]-\bar{k}-(1-\lambda)\left[p_{\bar{m}}(-\beta)-p_{\underline{m}}(-\beta)\right]=0 \quad \text{(v)}$$
$$\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})=\lambda\left[\left(p_{CD\underline{m}}-p_{DC\bar{m}}\right)(1+\alpha)\right]-\underline{k}=0\Leftrightarrow p_{CD\underline{m}}-p_{DC\bar{m}}=\frac{\underline{k}}{\lambda(1+\alpha)} \quad (*)$$

$p_{CD\underline{m}}-p_{DC\bar{m}}=\dfrac{k}{\lambda(1+\alpha)}\Rightarrow p_{CD\underline{m}}=\dfrac{1}{2}\left[1+\dfrac{k}{\lambda(1+\alpha)}\right]$, $p_{DC\bar{m}}=\dfrac{1}{2}\left[1-\dfrac{k}{\lambda(1+\alpha)}\right]$ by (*), if we plug in these values in (iv) and (ii)

$$p_{\bar{m}}<\frac{\lambda(\bar{m}-\alpha)}{\beta(1-\lambda)}p_{DC\bar{m}}=\frac{\lambda(\bar{m}-\alpha)}{\beta(1-\lambda)}\frac{1}{2}\left[1-\frac{k}{\lambda(1+\alpha)}\right] \quad \text{(iv)}$$

$$p_{\underline{m}} < \frac{\lambda\left(\bar{m}-\alpha\right)}{\beta\left(1-\lambda\right)} p_{CD\underline{m}} = \frac{\lambda\left(\bar{m}-\alpha\right)}{\beta\left(1-\lambda\right)} \frac{1}{2}\left[1+\frac{\underline{k}}{\lambda\left(1+\alpha\right)}\right] \qquad \text{(ii)}$$

Furthermore: $\frac{\left(1+\bar{m}\right)}{\left(1+\alpha\right)}\underline{k}-\bar{k}=\left(1-\lambda\right)\beta\left[p_{\underline{m}}-p_{\bar{m}}\right] \Leftrightarrow p_{\underline{m}}=\frac{1}{2}\left[1+\frac{1}{\left(1-\lambda\right)\beta}\left[\frac{\left(1+\bar{m}\right)}{\left(1+\alpha\right)}\underline{k}-\bar{k}\right]\right], p_{\bar{m}}=\frac{1}{2}\left[1-\frac{1}{\left(1-\lambda\right)\beta}\left[\frac{\left(1+\bar{m}\right)}{\left(1+\alpha\right)}\underline{k}-\bar{k}\right]\right]$ by (v)

w.r.t. (iv):

$$\frac{1}{2}\left[1-\frac{1}{\left(1-\lambda\right)\beta}\left[\frac{\left(1+\bar{m}\right)}{\left(1+\alpha\right)}\underline{k}-\bar{k}\right]\right] < \frac{\lambda\left(\bar{m}-\alpha\right)}{\beta\left(1-\lambda\right)}\frac{1}{2}\left[1-\frac{\underline{k}}{\lambda\left(1+\alpha\right)}\right] \Leftrightarrow$$

$$1-\frac{1}{\left(1-\lambda\right)\beta}\left[\frac{\left(1+\bar{m}\right)}{\left(1+\alpha\right)}\underline{k}-\bar{k}\right] < \frac{\lambda\left(\bar{m}-\alpha\right)}{\beta\left(1-\lambda\right)}\left[1-\frac{\underline{k}}{\lambda\left(1+\alpha\right)}\right] \Leftrightarrow$$

$$1-\frac{\lambda\left(\bar{m}-\alpha\right)}{\beta\left(1-\lambda\right)} < \frac{1}{\left(1-\lambda\right)\beta}\left[\frac{\left(1+\bar{m}\right)}{\left(1+\alpha\right)}\underline{k}-\bar{k}\right]+\frac{\lambda\left(\bar{m}-\alpha\right)}{\beta\left(1-\lambda\right)}\frac{\underline{k}}{\lambda\left(1+\alpha\right)} \Leftrightarrow$$

$$\beta-\lambda\left(\bar{m}-\alpha+\beta\right) < \frac{\left(1+\bar{m}\right)}{\left(1+\alpha\right)}\underline{k}-\bar{k}+\left(\bar{m}-\alpha\right)\frac{\underline{k}}{\left(1+\alpha\right)} \Leftrightarrow$$

$$\beta-\lambda\left(\bar{m}-\alpha+\beta\right) < \underline{k}-\bar{k}$$

w.r.t. (ii):

$$\frac{1}{2}\left[1+\frac{1}{(1-\lambda)\beta}\left[\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right]\right]<\frac{\lambda(\bar{m}-\alpha)}{\beta(1-\lambda)}\frac{1}{2}\left[1+\frac{\underline{k}}{\lambda(1+\alpha)}\right]\Leftrightarrow$$

$$\left[1+\frac{1}{(1-\lambda)\beta}\left[\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right]\right]<\frac{\lambda(\bar{m}-\alpha)}{\beta(1-\lambda)}+\frac{\lambda(\bar{m}-\alpha)}{\beta(1-\lambda)}\frac{\underline{k}}{\lambda(1+\alpha)}\Leftrightarrow$$

$$\left[(1-\lambda)\beta+\left[\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right]\right]<\lambda(\bar{m}-\alpha)+(\bar{m}-\alpha)\frac{\underline{k}}{(1+\alpha)}\Leftrightarrow$$

$$\beta-\lambda(\bar{m}-\alpha+\beta)<-\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}+\bar{k}+(\bar{m}-\alpha)\frac{\underline{k}}{(1+\alpha)}\Leftrightarrow$$

$$\beta-\lambda(\bar{m}-\alpha+\beta)<-(\underline{k}-\bar{k})\Leftrightarrow\frac{\beta+(\underline{k}-\bar{k})}{(\bar{m}-\alpha+\beta)}<\lambda$$

We observe that (ii) is binding.

Finally all condition of the type

$p \in [0,1], \sum p < 1$ reduce to:

(1): $p_{CD\underline{m}} > 0$ true

(2): $p_{CD\underline{m}} < 1 \Leftrightarrow \dfrac{\underline{k}}{(1+\alpha)} < \lambda$

(3): $p_{\underline{m}} > 0$ true

(4): $p_{\underline{m}} < 1 \Leftrightarrow \lambda < 1 - \dfrac{1}{\beta}\left[\dfrac{(1+\bar{m})}{(1+\alpha)}\underline{k} - \bar{k}\right]$

,i.e. $0 < \dfrac{\underline{k}}{(1+\alpha)} < \lambda < 1 - \dfrac{1}{\beta}\left[\dfrac{(1+\bar{m})}{(1+\alpha)}\underline{k} - \bar{k}\right] < 1$

note: $\dfrac{\underline{k}}{(1+\alpha)} < 1 - \dfrac{1}{\beta}\left[\dfrac{(1+\bar{m})}{(1+\alpha)}\underline{k} - \bar{k}\right] \Leftrightarrow \dfrac{(1+\bar{m}+\beta)}{\beta(1+\alpha)}\underline{k} < 1 + \dfrac{\bar{k}}{\beta} \Leftrightarrow (1+\alpha)\beta > (1+\bar{m}+\beta)\underline{k} - (1+\alpha)\bar{k}$,

To summarize:

$$\left\{ p_{CD\underline{m}} = \frac{1}{2}\left[1 + \frac{\underline{k}}{\lambda(1+\alpha)}\right], p_{DC\bar{m}} = \frac{1}{2}\left[1 - \frac{\underline{k}}{\lambda(1+\alpha)}\right], p_{\underline{m}} = \frac{1}{2}\left[1 + \frac{1}{(1-\lambda)\beta}\left[\frac{(1+\bar{m})}{(1+\alpha)}\underline{k} - \bar{k}\right]\right], p_{\bar{m}} = \frac{1}{2}\left[1 - \frac{1}{(1-\lambda)\beta}\left[\frac{(1+\bar{m})}{(1+\alpha)}\underline{k} - \bar{k}\right]\right]\right\}$$

Conditions:

$$(1):0<\max\left\{\frac{\underline{k}}{(1+a)},\frac{\beta+(\underline{k}-\bar{k})}{(\bar{m}\text{-}a+b)}\right\}<\lambda<1-\frac{1}{\beta}\left[\frac{(1+\bar{m})}{(1+a)}\underline{k}-\bar{k}\right]<1$$

$$note:\ a)\ \frac{\underline{k}}{(1+\alpha)}<1-\frac{1}{\beta}\left[\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right]\Leftrightarrow\frac{(1+\bar{m}+\beta)}{\beta(1+\alpha)}\underline{k}<1+\frac{\bar{k}}{\beta}\Leftrightarrow(1+\alpha)\beta>(1+\bar{m}+\beta)\underline{k}-(1+\alpha)\bar{k}$$

$$b)\ \frac{\beta+(\underline{k}-\bar{k})}{(\bar{m}-\alpha+\beta)}<1-\frac{1}{\beta}\left[\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right]\Leftrightarrow\frac{(\underline{k}-\bar{k})}{(\bar{m}-\alpha+\beta)}<\frac{\bar{m}-\alpha}{(\bar{m}-\alpha+\beta)}-\frac{1}{\beta}\left[\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right]\Leftrightarrow\underline{k}-\bar{k}<\bar{m}-\alpha-\frac{(\bar{m}-\alpha+\beta)}{\beta(1+\alpha)}\left[(1+\bar{m})\underline{k}-(1+\alpha)\bar{k}\right]$$

$$\Leftrightarrow\frac{\underline{k}-\bar{k}}{\bar{m}-\alpha}\beta(1+\alpha)<\beta(1+\alpha)-\frac{(\bar{m}-\alpha+\beta)}{(\bar{m}-\alpha)}\left[(1+\bar{m})\underline{k}-(1+\alpha)\bar{k}\right]\Leftrightarrow\beta(1+\alpha)>\frac{(\bar{m}-\alpha+\beta)}{(\bar{m}-\alpha)}\left[(1+\bar{m})\underline{k}-(1+\alpha)\bar{k}\right]+\frac{\underline{k}-\bar{k}}{\bar{m}-\alpha}\beta(1+\alpha)$$

$$\Leftrightarrow\beta(1+\alpha)>\frac{(\bar{m}-\alpha+\beta)}{(\bar{m}-\alpha)}\left[(1+\bar{m})\underline{k}-(1+\alpha)\bar{k}\right]+\frac{\underline{k}-\bar{k}}{\bar{m}-\alpha}\beta(1+\alpha)>\frac{(\bar{m}-\alpha+\beta)}{(\bar{m}-\alpha)}\left[(1+\bar{m})\underline{k}-(1+\alpha)\bar{k}\right]-\frac{\underline{k}-\bar{k}}{\bar{m}-\alpha}\beta(1+\alpha)$$

$$=\frac{(\bar{m}-\alpha+\beta)(1+\bar{m})-\beta(1+\alpha)}{(\bar{m}-\alpha)}\underline{k}-\frac{(\bar{m}-\alpha)(1+\alpha)}{(\bar{m}-\alpha)}\bar{k}=\frac{(\bar{m}-\alpha)(1+\bar{m}+\beta)}{(\bar{m}-\alpha)}\underline{k}-\frac{(\bar{m}-\alpha)(1+\alpha)}{(\bar{m}-\alpha)}=\bar{k}\,(1+\bar{m}+\beta)\underline{k}-(1+\alpha)\bar{k}$$

,i.e. the necessary condition for a non-empty interval a) is weaker than the b)

$$c)\ \frac{\underline{k}}{(1+\alpha)}<\frac{\beta-(\underline{k}-\bar{k})}{(\bar{m}-\alpha+\beta)}\left(<\frac{\beta+(\underline{k}-\bar{k})}{(\bar{m}-\alpha+\beta)}\right)\Leftrightarrow(\bar{m}-\alpha+\beta)\underline{k}<(1+\alpha)\beta-(1+\alpha)(\underline{k}-\bar{k})$$

$$\Leftrightarrow(1+\alpha)\beta>(\bar{m}-\alpha+\beta)\underline{k}+(1+\alpha)(\underline{k}-\bar{k})=(\bar{m}-\alpha+\beta)\underline{k}+(1+\alpha)(\underline{k}-\bar{k})=(1+\bar{m}+\beta)\underline{k}-(1+\alpha)\bar{k}$$

,i.e. the necessary condition for a non-empty interval a) implies that $\frac{\underline{k}}{(1+\alpha)}<\frac{\beta+(\underline{k}-\bar{k})}{(\bar{m}-\alpha+\beta)}$

Conditions for existence:

$$(1):0<\frac{\beta+(\underline{k}-\bar{k})}{(\bar{m}\text{-}a+b)}<\lambda<1-\frac{1}{\beta}\left[\frac{(1+\bar{m})}{(1+a)}\underline{k}-\bar{k}\right]<1$$

$$(2):\beta(1+\alpha)>\frac{(\bar{m}-\alpha+\beta)}{(\bar{m}-\alpha)}\left[(1+\bar{m})\underline{k}-(1+\alpha)\bar{k}\right]+\frac{\underline{k}-\bar{k}}{\bar{m}-\alpha}\beta(1+\alpha)$$

4.2.3.3. $\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})<0$

$$\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left[\left(p_{CD\underline{m}}\right)(\bar{m}-\alpha)\right]-\beta(1-\lambda)>0 \quad \text{(ii)}$$

$$\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(CD,\underline{m})=\lambda\left[\left(p_{CD\underline{m}}-p_{DC\bar{m}}\right)(1+\bar{m})\right]-\bar{k}-(1-\lambda)\beta=0 \quad \text{(iv)}$$

$$\Pi_{\bar{m}}(CD,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[\left(p_{DC\bar{m}}\right)(\bar{m}-\alpha)\right]>0 \quad \text{(v)}$$

$$\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})=\lambda\left[\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)(1+\alpha)\right]-\underline{k}<0\Rightarrow p_{\bar{m}}=0 \quad (*)$$

(v) is satisfied; $p_{CD\underline{m}}=\frac{1}{2}\left[1+\frac{\bar{k}+(1-\lambda)\beta}{\lambda(1+\bar{m})}\right]$, $p_{DC\bar{m}}=\frac{1}{2}\left[1-\frac{\bar{k}+(1-\lambda)\beta}{\lambda(1+\bar{m})}\right]$ by (iv),

$$\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left[\left(p_{CD\underline{m}}\right)(\bar{m}-\alpha)\right]-\beta(1-\lambda)\,p_{\underline{m}}>0 \quad \text{(ii)}$$

$$\Pi_{\bar{m}}(CD,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[\left(p_{DC\bar{m}}\right)(\bar{m}-\alpha)\right]-\beta(1-\lambda)\,p_{\bar{m}}>0 \quad \text{(iv)}$$

$$\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(CD,\underline{m})=\lambda\left[\left(p_{CD\underline{m}}-p_{DC\bar{m}}\right)(1+\bar{m})\right]-\bar{k}-(1-\lambda)\left[p_{\bar{m}}(-\beta)-p_{\underline{m}}(-\beta)\right]=0 \quad \text{(v)}$$

$$\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})=\lambda\left[\left(p_{CD\underline{m}}-p_{DC\bar{m}}\right)(1+\alpha)\right]-\underline{k}=0\Leftrightarrow p_{CD\underline{m}}-p_{DC\bar{m}}=\frac{\underline{k}}{\lambda(1+\alpha)} \quad (*)$$

(iv): $\lambda > \dfrac{\bar{k}+\beta}{(1+\bar{m}+\beta)}$

(*): $\lambda > \dfrac{\bar{k}+\beta-\dfrac{(1+\bar{m})}{(1+\alpha)}\underline{k}}{\beta} = 1-\dfrac{1}{\beta(1+\alpha)}\left((1+\bar{m})\underline{k}-(1+\alpha)\bar{k}\right)$

(ii): $\lambda > \dfrac{\beta(2+\alpha+\bar{m})-(\bar{m}-\alpha)\bar{k}}{(\bar{m}-\alpha+\beta)(1+\bar{m})+(1+\alpha)\beta} = \dfrac{\beta(1+\alpha)+\beta(1+\bar{m})+(\bar{m}-\alpha)(1+\bar{m})-(\bar{m}-\alpha)(1+\bar{m})-(\bar{m}-\alpha)\bar{k}}{(\bar{m}-\alpha+\beta)(1+\bar{m})+(1+\alpha)\beta}$

$= 1-\dfrac{(\bar{m}-\alpha)}{(\bar{m}-\alpha+\beta)(1+\bar{m})+(1+\alpha)\beta}(1+\bar{m}+\bar{k})$

note:

$1.(iv)-(*): \dfrac{\bar{k}+\beta}{(1+\bar{m}+\beta)}-1+\dfrac{1}{\beta(1+\alpha)}\left((1+\bar{m})\underline{k}-(1+\alpha)\bar{k}\right)>0 \Leftrightarrow \dfrac{\bar{k}}{(1+\bar{m}+\beta)}-\dfrac{1+\bar{m}}{(1+\bar{m}+\beta)}+\dfrac{(1+\bar{m})\underline{k}}{\beta(1+\alpha)}-\dfrac{\bar{k}}{\beta}>0$

$\Leftrightarrow -\dfrac{(1+\bar{m})}{(1+\bar{m}+\beta)\beta}\bar{k}-\dfrac{1+\bar{m}}{(1+\bar{m}+\beta)}+\dfrac{(1+\bar{m})}{\beta(1+\alpha)}\underline{k}>0 \Leftrightarrow \underline{k}-\dfrac{(1+\alpha)}{(1+\bar{m}+\beta)}\bar{k}>\dfrac{\beta(1+\alpha)}{(1+\bar{m}+\beta)} \Leftrightarrow (1+\bar{m}+\beta)\underline{k}-(1+\alpha)\bar{k}>\beta(1+\alpha)$

$2.(iv)-(ii): \dfrac{\bar{k}+\beta}{(1+\bar{m}+\beta)}-1+\dfrac{(\bar{m}-\alpha)}{(\bar{m}-\alpha+\beta)(1+\bar{m})+(1+\alpha)\beta}(1+\bar{m}+\bar{k})>0 \Leftrightarrow$

$\dfrac{\bar{k}}{(1+\bar{m}+\beta)}-\dfrac{1+\bar{m}}{(1+\bar{m}+\beta)}+\dfrac{(\bar{m}-\alpha)(1+\bar{m})}{(\bar{m}-\alpha+\beta)(1+\bar{m})+(1+\alpha)\beta}+\dfrac{(\bar{m}-\alpha)}{(\bar{m}-\alpha+\beta)(1+\bar{m})+(1+\alpha)\beta}\bar{k}>0 \Leftrightarrow$

$\left(\dfrac{1}{(1+\bar{m}+\beta)}+\dfrac{(\bar{m}-\alpha)}{(\bar{m}-\alpha+\beta)(1+\bar{m})+(1+\alpha)\beta}\right)\bar{k} > \left(\dfrac{1}{(1+\bar{m}+\beta)}-\dfrac{(\bar{m}-\alpha)}{(\bar{m}-\alpha+\beta)(1+\bar{m})+(1+\alpha)\beta}\right)(1+\bar{m}) \Leftrightarrow$

$\dfrac{\bar{k}}{(1+\bar{m})} > \dfrac{(\bar{m}-\alpha+\beta)(1+\bar{m})+(1+\alpha)\beta-(1+\bar{m}+\beta)(\bar{m}-\alpha)}{(\bar{m}-\alpha+\beta)(1+\bar{m})+(1+\alpha)\beta+(1+\bar{m}+\beta)(\bar{m}-\alpha)} = \dfrac{(1+\alpha)\beta}{(1+\bar{m})(\bar{m}-\alpha+\beta)} \Leftrightarrow \dfrac{\bar{k}}{(1+\alpha)} > \dfrac{\beta}{(\bar{m}-\alpha+\beta)} \Leftrightarrow (\bar{m}-\alpha+\beta)\bar{k} > (1+\alpha)\beta$

$$3.(*)-(ii):1-\frac{1}{\beta\left(1+\alpha\right)}\left(\left(1+\bar{m}\right)\underline{k}-\left(1+\alpha\right)\bar{k}\right)-1+\frac{\left(\bar{m}-\alpha\right)}{\left(\bar{m}-\alpha+\beta\right)\left(1+\bar{m}\right)+\left(1+\alpha\right)\beta}\left(1+\bar{m}+\bar{k}\right)>0\Leftrightarrow$$

$$\frac{\left(\bar{m}-\alpha\right)}{\left(\bar{m}-\alpha+\beta\right)\left(1+\bar{m}\right)+\left(1+\alpha\right)\beta}\left(1+\bar{m}+\bar{k}\right)>\frac{1}{\beta\left(1+\alpha\right)}\left(\left(1+\bar{m}\right)\underline{k}-\left(1+\alpha\right)\bar{k}\right)\Leftrightarrow$$

$$\frac{\left(\bar{m}-\alpha\right)\left(1+\bar{m}\right)}{\left(\bar{m}-\alpha+\beta\right)\left(1+\bar{m}\right)+\left(1+\alpha\right)\beta}>\frac{\left(1+\bar{m}\right)}{\beta\left(1+\alpha\right)}\underline{k}-\frac{\bar{k}}{\beta}-\frac{\left(\bar{m}-\alpha\right)}{\left(\bar{m}-\alpha+\beta\right)\left(1+\bar{m}\right)+\left(1+\alpha\right)\beta}\bar{k}\Leftrightarrow$$

$$1>\frac{\left(\bar{m}-\alpha+\beta\right)\left(1+\bar{m}\right)+\left(1+\alpha\right)\beta}{\beta\left(1+\alpha\right)\left(\bar{m}-\alpha\right)}\underline{k}-\frac{\left(\bar{m}-\alpha+\beta\right)\left(1+\bar{m}\right)+\left(1+\alpha\right)\beta}{\beta\left(\bar{m}-\alpha\right)\left(1+\bar{m}\right)}\bar{k}-\frac{1}{\left(1+\bar{m}\right)}\bar{k}\Leftrightarrow$$

$$1>\frac{\left(\bar{m}-\alpha+\beta\right)\left(1+\bar{m}\right)+\left(1+\alpha\right)\beta}{\beta\left(1+\alpha\right)\left(\bar{m}-\alpha\right)}\underline{k}-\frac{\left(\bar{m}-\alpha+2\beta\right)}{\beta\left(\bar{m}-\alpha\right)}\bar{k}\Leftrightarrow\beta\left(1+\alpha\right)>\frac{\left(\bar{m}-\alpha+\beta\right)\left(1+\bar{m}\right)+\left(1+\alpha\right)\beta}{\left(\bar{m}-\alpha\right)}\underline{k}-\frac{\left(\bar{m}-\alpha+\beta\right)\left(1+\alpha\right)+\left(1+\alpha\right)\beta}{\left(\bar{m}-\alpha\right)}\bar{k}$$

$$\Leftrightarrow\beta\left(1+\alpha\right)>\frac{\left(\bar{m}-\alpha+\beta\right)\left(1+\bar{m}\right)+\left(1+\alpha\right)\beta}{\left(\bar{m}-\alpha\right)}\underline{k}-\frac{\left(\bar{m}-\alpha+\beta\right)\left(1+\alpha\right)+\left(1+\alpha\right)\beta}{\left(\bar{m}-\alpha\right)}\bar{k}$$

$$\left(1+\bar{m}+\beta\right)\underline{k}-\left(1+\alpha\right)\bar{k}>\left(\bar{m}-\alpha+\beta\right)\bar{k}\Rightarrow\left((iv)-(ii)>0\Rightarrow(iv)-(*)>0\right)$$

$$\left(1+\alpha\right)\beta<\left(\bar{m}-\alpha+\beta\right)\bar{k}:(iv)$$

$$\left(\bar{m}-\alpha+\beta\right)\bar{k}<\left(1+\alpha\right)\beta<\left(1+\bar{m}+\beta\right)\underline{k}-\left(1+\alpha\right)\bar{k}:(ii)$$

$$\left(1+\bar{m}+\beta\right)\underline{k}-\left(1+\alpha\right)\bar{k}<\left(1+\alpha\right)\beta:(*)oder(ii)$$

$$\beta(1+\alpha) > (1+\bar{m}+\beta)\underline{k} - (1+\alpha)\bar{k}$$

$$\beta(1+\alpha) > \frac{(\bar{m}-\alpha+\beta)(1+\bar{m})+(1+\alpha)\beta}{(\bar{m}-\alpha)}\underline{k} - \frac{(\bar{m}-\alpha+\beta)(1+\alpha)+(1+\alpha)\beta}{(\bar{m}-\alpha)}\bar{k}$$

$$(1+\bar{m}+\beta)\underline{k} - (1+\alpha)\bar{k} - \frac{(\bar{m}-\alpha+\beta)(1+\bar{m})+(1+\alpha)\beta}{(\bar{m}-\alpha)}\underline{k} + \frac{(\bar{m}-\alpha+\beta)(1+\alpha)+(1+\alpha)\beta}{(\bar{m}-\alpha)}\bar{k}$$

$$= \frac{(1+\bar{m}+\beta)(\bar{m}-\alpha)-(\bar{m}-\alpha+\beta)(1+\bar{m})-(1+\alpha)\beta}{(\bar{m}-\alpha)}\underline{k} + \frac{(\bar{m}-\alpha+\beta)(1+\alpha)+(1+\alpha)\beta-(1+\alpha)(\bar{m}-\alpha)}{(\bar{m}-\alpha)}\bar{k}$$

$$= \frac{-2(1+\alpha)\beta}{(\bar{m}-\alpha)}\underline{k} + \frac{2(1+\alpha)\beta}{(\bar{m}-\alpha)}\bar{k} = \frac{2(1+\alpha)\beta}{(\bar{m}-\alpha)}(\bar{k}-\underline{k}) < 0 \Rightarrow$$

$$(1+\alpha)\beta < (\bar{m}-\alpha+\beta)\bar{k} : (iv)$$

$$(\bar{m}-\alpha+\beta)\bar{k} < (1+\alpha)\beta < (1+\bar{m}+\beta)\underline{k} - (1+\alpha)\bar{k} : (ii)$$

$$(1+\bar{m}+\beta)\underline{k} - (1+\alpha)\bar{k} < (1+\alpha)\beta < \frac{(\bar{m}-\alpha+\beta)(1+\bar{m})+(1+\alpha)\beta}{(\bar{m}-\alpha)}\underline{k} - \frac{(\bar{m}-\alpha+\beta)(1+\alpha)+(1+\alpha)\beta}{(\bar{m}-\alpha)}\bar{k} : (ii)$$

$$\frac{(\bar{m}-\alpha+\beta)(1+\bar{m})+(1+\alpha)\beta}{(\bar{m}-\alpha)}\underline{k} - \frac{(\bar{m}-\alpha+\beta)(1+\alpha)+(1+\alpha)\beta}{(\bar{m}-\alpha)}\bar{k} < (1+\alpha)\beta : (*)$$

Finally all condition of the type

$p \in [0,1], \sum p < 1$ reduce to:

(1): $p_{CD\underline{m}} > 0$ true ,

(2): $p_{CD\underline{m}} < 1 \Leftrightarrow \lambda > \dfrac{\bar{k}+\beta}{(1+\bar{m}+\beta)}$

To summarize:

$$\left\{ p_{CD\underline{m}} = \frac{1}{2}\left[1+\frac{\bar{k}+(1-\lambda)\beta}{\lambda(1+\bar{m})}\right],\ p_{DC\bar{m}} = \frac{1}{2}\left[1-\frac{\bar{k}+(1-\lambda)\beta}{\lambda(1+\bar{m})}\right],\ p_{\bar{m}} = 0 \right\}$$

Conditions for Existence:

$$\lambda > \max\left\{ \frac{\bar{k}+\beta}{(1+\bar{m}+\beta)},\ 1-\frac{1}{\beta(1+\alpha)}\left((1+\bar{m})\underline{k}-(1+\alpha)\bar{k}\right),\ 1-\frac{\bar{m}-\alpha}{(\bar{m}-\alpha+\beta)(1+\bar{m})+(1+\alpha)\beta}(1+\bar{m}+\bar{k}) \right\}$$

$$= \begin{cases} \dfrac{\bar{k}+\beta}{(1+\bar{m}+\beta)}, & (1+\alpha)\beta < (\bar{m}-\alpha+\beta)\bar{k} \\ 1-\dfrac{(\bar{m}-\alpha)}{(\bar{m}-\alpha+\beta)(1+\bar{m})+(1+\alpha)\beta}(1+\bar{m}+\bar{k}), & (\bar{m}-\alpha+\beta)\bar{k} < (1+\alpha)\beta < \dfrac{(\bar{m}-\alpha+\beta)(1+\bar{m})+(1+\alpha)\beta}{(\bar{m}-\alpha)}\underline{k}-\dfrac{(\bar{m}-\alpha+\beta)(1+\alpha)+(1+\alpha)\beta}{(\bar{m}-\alpha)}\bar{k} \\ 1-\dfrac{1}{\beta(1+\alpha)}\left((1+\bar{m})\underline{k}-(1+\alpha)\bar{k}\right), & \dfrac{(\bar{m}-\alpha+\beta)(1+\bar{m})+(1+\alpha)\beta}{(\bar{m}-\alpha)}\underline{k}-\dfrac{(\bar{m}-\alpha+\beta)(1+\alpha)+(1+\alpha)\beta}{(\bar{m}-\alpha)}\bar{k} < (1+\alpha)\beta \end{cases}$$

4.3. $CC,\bar{m}\,/\,DC,\bar{m}$

4.3.1. $CD,\underline{m}$

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(DC,\bar{m}) = \lambda\left[(p_{CC\bar{m}}+p_{CD\bar{m}})(\bar{m}-\alpha)+(p_{DC\bar{m}}+p_{DD\bar{m}})(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]=0 \tag{i}$$

$$\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[(p_{CC\underline{m}}+p_{CD\underline{m}})(\bar{m}-\alpha)+(p_{DC\underline{m}}+p_{DD\underline{m}})(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]>0 \tag{ii}$$

$$\Pi_{\bar{m}}(CD,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[(p_{CC\bar{m}}+p_{DC\bar{m}})(\bar{m}-\alpha)+(p_{CD\bar{m}}+p_{DD\bar{m}})(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]>0 \tag{iii}$$

$$\Pi_{\bar{m}}(CC,\underline{m})-\Pi_{\bar{m}}(CD,\underline{m}) = \Pi_{\bar{m}}(DC,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[(p_{CC\underline{m}}+p_{DC\underline{m}})(\bar{m}-\alpha)+(p_{CD\underline{m}}+p_{DD\underline{m}})(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]<0 \tag{iv}$$

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CD,\underline{m}) = \Pi_{\bar{m}}(CD,\bar{m})-\Pi_{\bar{m}}(CD,\underline{m})-\left[\Pi_{\bar{m}}(CD,\bar{m})-\Pi_{\bar{m}}(CC,\bar{m})\right]= \tag{v}$$

$$\lambda\left[(p_{CD\bar{m}}-p_{DC\bar{m}})(1+\bar{m}+\beta)+(p_{CD\underline{m}}-p_{DC\underline{m}})(1+\alpha)\right]-\bar{k}+\lambda\left[(p_{CC\underline{m}}+p_{CD\underline{m}})(\bar{m}-\alpha)+(p_{DC\underline{m}}+p_{DD\underline{m}})(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]=0$$

$$\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m}) = \lambda\left[(p_{CD\bar{m}}+p_{CD\underline{m}}-p_{DC\bar{m}}-p_{DC\underline{m}})(1+\alpha)\right]-\underline{k} \tag{*}$$

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(DC,\bar{m})=\lambda\left[(p_{CC\bar{m}})(\bar{m}-\alpha)+(p_{DC\bar{m}})(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]=0 \quad \text{(i)}$$

$$\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left[(p_{CD\underline{m}})(\bar{m}-\alpha)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]>0 \quad \text{(ii)}$$

$$\Pi_{\bar{m}}(CD,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[(p_{CC\bar{m}}+p_{DC\bar{m}})(\bar{m}-\alpha)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]>0 \quad \text{(iii)}$$

$$\Pi_{\bar{m}}(CC,\underline{m})-\Pi_{\bar{m}}(CD,\underline{m})=\Pi_{\bar{m}}(DC,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[(p_{CD\underline{m}})(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]<0 \quad \text{(iv)}$$

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CD,\underline{m})=\lambda\left[(-p_{DC\bar{m}})(1+\bar{m}+\beta)+(p_{CD\underline{m}})(1+\alpha)\right]-\bar{k}+\lambda\left[(p_{CD\underline{m}})(\bar{m}-\alpha)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]= \quad \text{(v)}$$

$$\lambda\left[(-p_{DC\bar{m}})(1+\bar{m}+\beta)+(p_{CD\underline{m}})(1+\bar{m})\right]-\bar{k}+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]=$$

$$\lambda\left[(p_{CD\underline{m}}-p_{DC\bar{m}})(1+\bar{m})-p_{DC\bar{m}}\beta\right]-\bar{k}+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]=0$$

$$\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})=\lambda\left[(p_{CD\bar{m}}+p_{CD\underline{m}}-p_{DC\bar{m}}-p_{DC\underline{m}})(1+\alpha)\right]-\underline{k} \quad (*)$$

(iv) is always satisfied in a semi-pooling equilibrium, (i) implies (iii) if DC is played by strictly positive share.

4.3.1.1. $\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})>0$

$$\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})=\lambda\left[(p_{CD\underline{m}}-p_{DC\bar{m}})(1+\alpha)\right]-\underline{k}>0\Rightarrow p_{\bar{m}}=1 \quad (*)$$

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(DC,\bar{m})=\lambda\left[(p_{CC\bar{m}})(\bar{m}-\alpha)+(p_{DC\bar{m}})(-\beta)\right]-(1-\lambda)\beta=0 \quad \text{(i)}$$

$$\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left[(p_{CD\underline{m}})(\bar{m}-\alpha)\right]>0 \quad \text{(ii)}$$

$$\Pi_{\bar{m}}(CD,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[(p_{CC\bar{m}}+p_{DC\bar{m}})(\bar{m}-\alpha)\right]-\beta>0 \quad \text{(iii)}$$

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CD,\underline{m})=\lambda\left[(p_{CD\underline{m}}-p_{DC\bar{m}})(1+\bar{m})-p_{DC\bar{m}}\beta\right]-\bar{k}=0 \quad \text{(v)}$$

(*) and (v) imply : $\lambda\left[\frac{\underline{k}}{\lambda(1+\alpha)}(1+\bar{m})-p_{DC\bar{m}}\beta\right]-\bar{k}<0 \Leftrightarrow \frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}<\lambda\beta p_{DC\bar{m}}$

(i) and (v) give:

$$p_{DC\bar{m}}=\frac{(1+\bar{m})(\lambda(\bar{m}-\alpha+\beta)-\beta)-\bar{k}(\bar{m}-\alpha)}{\lambda(2(1+\bar{m})(\bar{m}-\alpha+\beta)-\beta(1+\alpha))}$$

$$p_{CC\bar{m}}=\frac{\beta(2(1+\bar{m})+\beta-\lambda(1+\bar{m}+\beta)-\bar{k})}{\lambda(2(1+\bar{m})(\bar{m}-\alpha+\beta)-\beta(1+\alpha))}$$

$$p_{CD\underline{m}}=\frac{\bar{k}(\bar{m}-\alpha+\beta)+(1+\bar{m}+\beta)(\lambda(\bar{m}-\alpha+\beta)-\beta)}{\lambda(2(1+\bar{m})(\bar{m}-\alpha+\beta)-\beta(1+\alpha))}$$

Given this solution for the respective shares (*) and (iii) need to be checked for.

$$(iii):-(\bar{m}-\alpha+\beta)\frac{(1+\bar{m})(\lambda(\bar{m}-\alpha+\beta)-\beta)-\bar{k}(\bar{m}-\alpha)}{(2(1+\bar{m})(\bar{m}-\alpha+\beta)-\beta(1+\alpha))}=-\lambda(\bar{m}-\alpha+\beta)p_{DC\bar{m}}<0$$

$$(*):\frac{(1+\alpha)(\bar{k}(2\bar{m}-2\alpha+\beta)+\beta(\lambda(\bar{m}-\alpha+\beta)-\beta))}{(2(1+\bar{m})(\bar{m}-\alpha+\beta)-\beta(1+\alpha))}>\underline{k}$$

$\Leftrightarrow$

$$(iii):\frac{\beta}{(\bar{m}-\alpha+\beta)}+\frac{\bar{k}(\bar{m}-\alpha)}{(1+\bar{m})(\bar{m}-\alpha+\beta)}=\frac{\beta(1+\bar{m})+\bar{k}(\bar{m}-\alpha)}{(1+\bar{m})(\bar{m}-\alpha+\beta)}<\lambda<1 \qquad \wedge \qquad \bar{k}<1+\bar{m}$$

$$(*):\frac{(1+\alpha)\left(\bar{k}(2\bar{m}-2\alpha+\beta)+\beta\left(\lambda(\bar{m}-\alpha+\beta)-\beta\right)\right)}{\left(2(1+\bar{m})(\bar{m}-\alpha+\beta)-\beta(1+\alpha)\right)}>\underline{k}$$

$$(1+\alpha)\left(\beta\left(\lambda(\bar{m}-\alpha+\beta)-\beta\right)\right)>\underline{k}\left(2(1+\bar{m})(\bar{m}-\alpha+\beta)-\beta(1+\alpha)\right)-\bar{k}\left(2(1+\alpha)(\bar{m}-\alpha+\beta)-(1+\alpha)\beta\right)$$

$$(1+\alpha)\left(\beta\left(\lambda(\bar{m}-\alpha+\beta)-\beta\right)\right)>2(\bar{m}-\alpha+\beta)\left(\underline{k}(1+\bar{m})-\bar{k}(1+\alpha)\right)-(1+\alpha)\beta\left(\underline{k}-\bar{k}\right)$$

$$\lambda(\bar{m}-\alpha+\beta)-\beta>2\frac{(\bar{m}-\alpha+\beta)}{\beta}\left(\underline{k}\frac{(1+\bar{m})}{(1+\alpha)}-\bar{k}\right)-\left(\underline{k}-\bar{k}\right)$$

$$\lambda>\frac{\beta-\left(\underline{k}-\bar{k}\right)}{(\bar{m}-\alpha+\beta)}+\frac{2}{\beta}\left(\underline{k}\frac{(1+\bar{m})}{(1+\alpha)}-\bar{k}\right)$$

Finally all condition of the type

$p\in[0,1],\sum p<1$ reduce to: $\frac{\beta(1+\bar{m})+\bar{k}(\bar{m}-\alpha)}{(1+\bar{m})(\bar{m}-\alpha+\beta)}<\lambda<1 \qquad \wedge \qquad \bar{k}<1+\bar{m}$ (derived with mathematica)

$$p_{CD\underline{m}}>0:\lambda>\frac{\beta}{(\bar{m}-\alpha+\beta)}-\frac{\bar{k}}{(1+\bar{m}+\beta)}$$

$p_{CD\underline{m}}<1:$

$$\frac{\bar{k}(\bar{m}-\alpha+\beta)+(1+\bar{m}+\beta)\left(\lambda(\bar{m}-\alpha+\beta)-\beta\right)}{\lambda\left(2(1+\bar{m})(\bar{m}-\alpha+\beta)-\beta(1+\alpha)\right)}<1$$

$$\Leftrightarrow \bar{k}(\bar{m}-\alpha+\beta)-\beta(1+\bar{m}+\beta)<\lambda\left(2(1+\bar{m})(\bar{m}-\alpha+\beta)-\beta(1+\alpha)-(1+\bar{m}+\beta)(\bar{m}-\alpha+\beta)\right)$$

$$\Leftrightarrow \bar{k}(\bar{m}-\alpha+\beta)-\beta(1+\bar{m}+\beta)<\lambda\left((1+\bar{m})(\bar{m}-\alpha)-\beta^{2}\right)$$

Note that: $\frac{2}{\beta}\left(\underline{k}\frac{(1+\bar{m})}{(1+\alpha)}-\bar{k}\right)+\frac{\beta-(\underline{k}-\bar{k})}{(\bar{m}-\alpha+\beta)}>\frac{\beta+(\bar{m}-\alpha)\frac{\underline{k}}{(1+\alpha)}}{(\bar{m}-\alpha+\beta)}$, since $\frac{\underline{k}}{(1+\alpha)}\left(1+\bar{m}+\frac{1+\alpha}{2\frac{(\bar{m}-\alpha+\beta)}{\beta}-1}\right)>\bar{k}$.

In summary the equilibrium is given:

$$p_{DC\bar{m}}=\frac{(1+\bar{m})(\lambda(\bar{m}-\alpha+\beta)-\beta)-\bar{k}(\bar{m}-\alpha)}{\lambda(2(1+\bar{m})(\bar{m}-\alpha+\beta)-\beta(1+\alpha))}$$

$$p_{CC\bar{m}}=\frac{\beta(2(1+\bar{m})+\beta-\lambda(\bar{m}-\alpha+\beta)-\bar{k})}{\lambda(2(1+\bar{m})(\bar{m}-\alpha+\beta)-\beta(1+\alpha))}$$

$$p_{CD\underline{m}}=\frac{\bar{k}(\bar{m}-\alpha+\beta)+(1+\bar{m}+\beta)(\lambda(\bar{m}-\alpha+\beta)-\beta)}{\lambda(2(1+\bar{m})(\bar{m}-\alpha+\beta)-\beta(1+\alpha))}$$

Condition for existence: $\frac{2}{\beta}\left(\underline{k}\frac{(1+\bar{m})}{(1+\alpha)}-\bar{k}\right)+\frac{\beta-(\underline{k}-\bar{k})}{(\bar{m}-\alpha+\beta)}<\lambda<1$

4.3.1.2. $\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})=0$

$$\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})=\lambda\left[(p_{CD\underline{m}}-p_{DC\bar{m}})(1+\alpha)\right]-\underline{k}=0\Leftrightarrow 1-2p_{DC\bar{m}}-\frac{\underline{k}}{\lambda(1+\alpha)}=p_{CC\bar{m}} \quad (*)$$

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(DC,\bar{m})=\lambda\left[(p_{CC\bar{m}})(\bar{m}-\alpha)+(p_{DC\bar{m}})(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]=0 \quad \text{(i)}$$

$$\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left[(p_{CD\underline{m}})(\bar{m}-\alpha)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]>0 \quad \text{(ii)}$$

$$\Pi_{\bar{m}}(CD,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[(p_{CC\bar{m}}+p_{DC\bar{m}})(\bar{m}-\alpha)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]>0 \quad \text{(iii)}$$

$$\begin{aligned}
&\Pi_{\bar{m}}\left(CC,\bar{m}\right)-\Pi_{\bar{m}}\left(CD,\underline{m}\right)=\lambda\left[\left(-p_{DC\bar{m}}\right)\left(1+\bar{m}+\beta\right)+\left(p_{CD\underline{m}}\right)\left(1+\alpha\right)\right]-\bar{k}+\lambda\left[\left(p_{CD\underline{m}}\right)\left(\bar{m}-\alpha\right)\right]+\left(1-\lambda\right)\left[p_{\underline{m}}\left(-\beta\right)\right]=\\
&\lambda\left[\left(-p_{DC\bar{m}}\right)\left(1+\bar{m}+\beta\right)+\left(p_{CD\underline{m}}\right)\left(1+\bar{m}\right)\right]-\bar{k}+\left(1-\lambda\right)\left[p_{\underline{m}}\left(-\beta\right)\right]= \qquad \text{(v)}\\
&\lambda\left[\left(p_{CD\underline{m}}-p_{DC\bar{m}}\right)\left(1+\bar{m}\right)-p_{DC\bar{m}}\beta\right]-\bar{k}+\left(1-\lambda\right)\left[p_{\underline{m}}\left(-\beta\right)\right]=0
\end{aligned}$$

To summarize:

$$\left\{\begin{array}{l}
\left(p_{CC\bar{m}},p_{DC\bar{m}},p_{CD\underline{m}},p_{\bar{m}}\right)\Big| \\
p_{DC\bar{m}}=\dfrac{1}{2}\left(1-\dfrac{\underline{k}-\bar{k}-\beta}{\lambda\left(\bar{m}-\alpha+\beta\right)}\right)\\
p_{CC\bar{m}}=\dfrac{\beta+\bar{k}-\dfrac{1+\bar{m}+\beta}{1+\alpha}\underline{k}}{\lambda\left(\bar{m}-\alpha+\beta\right)}\\
p_{CD\underline{m}}=\dfrac{1}{2}\left(1+\dfrac{\left(2\dfrac{\left(\bar{m}-\alpha+\beta\right)}{1+\alpha}+1\right)\underline{k}-\bar{k}-\beta}{\lambda\left(\bar{m}-\alpha+\beta\right)}\right)\\
p_{\underline{m}}=\dfrac{1}{\beta\left(1-\lambda\right)}\left(\dfrac{\left(1+\bar{m}\right)}{\left(1+\alpha\right)}\underline{k}-\bar{k}\right)-\dfrac{1}{2\left(1-\lambda\right)\left(\bar{m}-\alpha+\beta\right)}\left(\underline{k}-\bar{k}\right)-\dfrac{\lambda\left(\bar{m}-\alpha+\beta\right)-\beta}{2\left(1-\lambda\right)\left(\bar{m}-\alpha+\beta\right)}
\end{array}\right\}$$

Given this solution for the respective shares (ii) and (iii) need to be checked for.

$(ii): \lambda\left(\bar{m}-\alpha+\beta\right) > \beta + \underline{k} - \bar{k}$

$(iii): \lambda\left(\bar{m}-\alpha+\beta\right) > \beta - \left(\underline{k} - \bar{k}\right)$, i.e. (ii) implies (iii).

Finally all condition of the type

$p \in [0,1], \sum p < 1$ reduce to:

(high types): $\lambda > \dfrac{\beta+\underline{k}-\bar{k}}{\left(\bar{m}-\alpha+\beta\right)} \quad \wedge \quad (1+\alpha)\beta > (1+\bar{m}+\beta)\underline{k} - (1+\alpha)\bar{k} \quad \wedge \quad \underline{k} < 1+\alpha$

(low types): $\lambda < \min\left\{ \dfrac{\beta+\bar{k}-\underline{k}}{\left(\bar{m}-\alpha+\beta\right)} + \dfrac{2}{\beta}\left(\dfrac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right), 2-\left(\dfrac{\beta+\bar{k}-\underline{k}}{\left(\bar{m}-\alpha+\beta\right)} + \dfrac{2}{\beta}\left(\dfrac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right)\right)\right\}$

note:

$$\frac{\beta+\underline{k}-\bar{k}}{\left(\bar{m}-\alpha+\beta\right)} < 1 \Leftrightarrow \underline{k}-\bar{k} < \bar{m}-\alpha$$

$p_{\underline{m}} > 0$:

$$\frac{1}{\beta(1-\lambda)}\left(\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right) > \frac{1}{2(1-\lambda)\left(\bar{m}-\alpha+\beta\right)}\left(\underline{k}-\bar{k}\right) + \frac{\lambda\left(\bar{m}-\alpha+\beta\right)-\beta}{2(1-\lambda)\left(\bar{m}-\alpha+\beta\right)} \Leftrightarrow$$

$$\frac{2\left(\bar{m}-\alpha+\beta\right)}{\beta}\left(\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right) > \left(\underline{k}-\bar{k}\right) + \lambda\left(\bar{m}-\alpha+\beta\right) - \beta \Leftrightarrow$$

$$\lambda < \frac{\beta+\bar{k}-\underline{k}}{\left(\bar{m}-\alpha+\beta\right)} + \frac{2}{\beta}\left(\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right)$$

$p_{\underline{m}} < 1:$

$$\frac{1}{\beta(1-\lambda)}\left(\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right) < \frac{1}{2(1-\lambda)(\bar{m}-\alpha+\beta)}\left(\underline{k}-\bar{k}\right)+\frac{\lambda(\bar{m}-\alpha+\beta)-\beta}{2(1-\lambda)(\bar{m}-\alpha+\beta)}+1 \Leftrightarrow$$

$$\frac{2(\bar{m}-\alpha+\beta)}{\beta}\left(\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right) < \left(\underline{k}-\bar{k}\right)+\lambda(\bar{m}-\alpha+\beta)-\beta+2(1-\lambda)(\bar{m}-\alpha+\beta) \Leftrightarrow$$

$$\lambda(\bar{m}-\alpha+\beta) < \left(\underline{k}-\bar{k}\right)-\frac{2(\bar{m}-\alpha+\beta)}{\beta}\left(\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right)-\beta+2(\bar{m}-\alpha+\beta) \Leftrightarrow$$

$$\lambda < 2+\frac{\left(\underline{k}-\bar{k}\right)-\beta}{(\bar{m}-\alpha+\beta)}-\frac{2}{\beta}\left(\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right) = 2-\left(\frac{\beta+\bar{k}-\underline{k}}{(\bar{m}-\alpha+\beta)}+\frac{2}{\beta}\left(\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right)\right)$$

LHS<RHS:

$$\frac{\beta+\bar{k}-\underline{k}}{(\bar{m}-\alpha+\beta)}+\frac{2}{\beta}\left(\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right)-2+\left(\frac{\beta+\bar{k}-\underline{k}}{(\bar{m}-\alpha+\beta)}+\frac{2}{\beta}\left(\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right)\right) = 2\frac{\beta+\bar{k}-\underline{k}}{(\bar{m}-\alpha+\beta)}+2\frac{2}{\beta}\left(\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right)-2 > 0 \Leftrightarrow$$

$$\frac{\beta+\bar{k}-\underline{k}}{(\bar{m}-\alpha+\beta)}+\frac{2}{\beta}\left(\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right) > 1 \Leftrightarrow (1+\alpha)\beta\frac{\beta+\bar{k}-\underline{k}}{(\bar{m}-\alpha+\beta)}+2\left((1+\bar{m})\underline{k}-(1+\alpha)\bar{k}\right) > (1+\alpha)\beta$$

from high types:

$$(1+\alpha)\beta > (1+\bar{m}+\beta)\underline{k} - (1+\alpha)\bar{k} \Rightarrow$$

$$(1+\bar{m}+\beta)\underline{k} - (1+\alpha)\bar{k} < (1+\alpha)\beta < (1+\alpha)\beta\frac{\beta+\bar{k}-\underline{k}}{(\bar{m}-\alpha+\beta)} + 2\left((1+\bar{m})\underline{k} - (1+\alpha)\bar{k}\right)$$

$$(1+\bar{m}+\beta)\underline{k} - (1+\alpha)\bar{k} < (1+\alpha)\beta\frac{\beta+\bar{k}-\underline{k}}{(\bar{m}-\alpha+\beta)} + 2\left((1+\bar{m})\underline{k} - (1+\alpha)\bar{k}\right) \Leftrightarrow$$

$$0 < (1+\alpha)\beta\frac{\beta+\bar{k}-\underline{k}}{(\bar{m}-\alpha+\beta)} + (1+\bar{m}-\beta)\underline{k} - (1+\alpha)\bar{k}$$

$$\Leftrightarrow 0 < (1+\alpha)\beta\frac{\beta}{(\bar{m}-\alpha+\beta)} + (1+\alpha)\beta\frac{\bar{k}}{(\bar{m}-\alpha+\beta)} - (1+\alpha)\beta\frac{\underline{k}}{(\bar{m}-\alpha+\beta)} + (1+\bar{m}-\beta)\underline{k} - (1+\alpha)\bar{k}$$

$$\Leftrightarrow 0 < (1+\alpha)\beta\frac{\beta}{(\bar{m}-\alpha+\beta)} - (1+\alpha)\frac{\bar{m}-\alpha}{(\bar{m}-\alpha+\beta)}\bar{k} + \frac{(1+\bar{m})(\bar{m}-\alpha+\beta) - \beta(1+\bar{m}+\beta)}{(\bar{m}-\alpha+\beta)}\underline{k}$$

$$\Leftrightarrow 0 < (1+\alpha)\beta - (1+\alpha)\frac{\bar{m}-\alpha}{\beta}\bar{k} + \left(\frac{\bar{m}-\alpha}{\beta} - \beta(1+\bar{m})\right)\underline{k} \Leftrightarrow 0 < (1+\alpha)\beta - (1+\alpha)\frac{\bar{m}-\alpha}{\beta}\bar{k} + \left(\frac{\bar{m}-\alpha}{\beta} - \beta(1+\bar{m})\right)\underline{k}$$

conditions for existence:

1. $\frac{\beta+\underline{k}-\bar{k}}{(\bar{m}-\alpha+\beta)} < \lambda < \min\left\{\frac{\beta+\bar{k}-\underline{k}}{(\bar{m}-\alpha+\beta)} + \frac{2}{\beta}\left(\frac{(1+\bar{m})}{(1+\alpha)}\underline{k} - \bar{k}\right), 2 - \left(\frac{\beta+\bar{k}-\underline{k}}{(\bar{m}-\alpha+\beta)} + \frac{2}{\beta}\left(\frac{(1+\bar{m})}{(1+\alpha)}\underline{k} - \bar{k}\right)\right)\right\}$
2. $\underline{k} - \bar{k} < \bar{m} - \alpha$
3. $\underline{k} < 1+\alpha$
4. $(1+\alpha)\beta > (1+\bar{m}+\beta)\underline{k} - (1+\alpha)\bar{k}$

4.3.1.3. $\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})<0$

$$\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})=\lambda\left[\left(p_{CD\underline{m}}-p_{DC\bar{m}}\right)(1+\alpha)\right]-\underline{k}<0\Rightarrow p_{\bar{m}}=0 \qquad (*)$$

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(DC,\bar{m})=\lambda\left[\left(p_{CC\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{DC\bar{m}}\right)(-\beta)\right]=0 \qquad \text{(i)}$$

$$\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left[\left(p_{CD\underline{m}}\right)(\bar{m}-\alpha)\right]-(1-\lambda)\beta>0 \qquad \text{(ii)}$$

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CD,\underline{m})=\lambda\left[\left(p_{CD\underline{m}}-p_{DC\bar{m}}\right)(1+\bar{m}+\beta)-p_{DC\bar{m}}\beta\right]-\bar{k}-(1-\lambda)\beta=0 \qquad \text{(v)}$$

Rewrite (i) and (v) to: $p_{CC\bar{m}}=\dfrac{p_{DC\bar{m}}\beta}{\lambda(\bar{m}-\alpha)}$, $p_{CD\underline{m}}=\dfrac{\bar{k}+(1-\lambda)\beta}{\lambda(1+\bar{m}+\beta)}+\left(1+\dfrac{\beta}{(1+\bar{m}+\beta)}\right)p_{DC\bar{m}}$ and plug into: $p_{CD\underline{m}}=1-p_{CC\bar{m}}-p_{DC\bar{m}}$

$$p_{DC\bar{m}}=\frac{(\bar{m}-\alpha)\left((1+\bar{m}+\beta)\lambda-\beta-\bar{k}\right)}{\lambda\left(2(1+\bar{m})(\bar{m}-\alpha+\beta)-\beta(1+\alpha)\right)}$$

EQ: $$p_{CC\bar{m}}=\frac{\beta\left((1+\bar{m}+\beta)\lambda-\beta-\bar{k}\right)}{\lambda\left(2(1+\bar{m})(\bar{m}-\alpha+\beta)-\beta(1+\alpha)\right)}$$

$$p_{CD\underline{m}}=\frac{\bar{k}(\bar{m}-\alpha+\beta)+(1+\bar{m}+\beta)\lambda(\bar{m}-\alpha)+(\bar{m}-\alpha+\beta)(1-\lambda)\beta}{\lambda\left(2(1+\bar{m})(\bar{m}-\alpha+\beta)-\beta(1+\alpha)\right)}$$

Finally, check for (*) and (ii) given those values:

w.r.t. (*):

$$(1+\alpha)\left[\bar{k}\left(2\bar{m}-2\alpha+\beta\right)+\beta\left(\left(2\bar{m}-2\alpha+\beta\right)-\lambda\left(\bar{m}-\alpha+\beta\right)\right)\right]<\left(2\left(1+\bar{m}\right)\left(\bar{m}-\alpha+\beta\right)-\beta\left(1+\alpha\right)\right)\underline{k}\Leftrightarrow$$

$$\bar{k}\left[2\left(1+\alpha\right)\left(\bar{m}-\alpha+\beta\right)-\beta\left(1+\alpha\right)\right]-\left[2\left(1+\bar{m}\right)\left(\bar{m}-\alpha+\beta\right)-\beta\left(1+\alpha\right)\right]\underline{k}+\beta\left(1+\alpha\right)\left[2\left(\bar{m}-\alpha+\beta\right)-\beta\right]<\lambda\beta\left(1+\alpha\right)\left(\bar{m}-\alpha+\beta\right)\Leftrightarrow$$

$$\frac{1}{\left(\bar{m}-\alpha+\beta\right)}\left(\underline{k}-\bar{k}\right)-\frac{2}{\beta}\left(\frac{\left(1+\bar{m}\right)}{\left(1+\alpha\right)}\underline{k}-\bar{k}\right)+\frac{2\left(\bar{m}-\alpha+\beta\right)-\beta}{\left(\bar{m}-\alpha+\beta\right)}=2-\left(\frac{\beta+\bar{k}-\underline{k}}{\left(\bar{m}-\alpha+\beta\right)}+\frac{2}{\beta}\left(\frac{\left(1+\bar{m}\right)}{\left(1+\alpha\right)}\underline{k}-\bar{k}\right)\right)<\lambda$$

w.r.t. (ii):

$$\lambda>\frac{\beta+\bar{k}}{\left(1+\bar{m}+\beta\right)}\wedge\bar{k}<1+\bar{m}$$

To summarize:

$$p_{DC\bar{m}}=\frac{\left(\bar{m}-\alpha\right)\left(\left(1+\bar{m}+\beta\right)\lambda-\beta-\bar{k}\right)}{\lambda\left(2\left(1+\bar{m}\right)\left(\bar{m}-\alpha+\beta\right)-\beta\left(1+\alpha\right)\right)}$$

$$p_{CC\bar{m}}=\frac{\beta\left(\left(1+\bar{m}+\beta\right)\lambda-\beta-\bar{k}\right)}{\lambda\left(2\left(1+\bar{m}\right)\left(\bar{m}-\alpha+\beta\right)-\beta\left(1+\alpha\right)\right)}$$

$$p_{CD\underline{m}}=\frac{\bar{k}\left(\bar{m}-\alpha+\beta\right)+\left(1+\bar{m}+\beta\right)\lambda\left(\bar{m}-\alpha\right)+\left(\bar{m}-\alpha+\beta\right)\left(1-\lambda\right)\beta}{\lambda\left(2\left(1+\bar{m}\right)\left(\bar{m}-\alpha+\beta\right)-\beta\left(1+\alpha\right)\right)}$$

Finally all condition of the type

$p\in[0,1],\sum p<1$ reduce to: $\lambda>\dfrac{\beta+\bar{k}}{\left(1+\bar{m}+\beta\right)}\wedge\bar{k}<1+\bar{m}$

Hence we are left with:

$$\lambda > \max\left\{ \underbrace{\frac{\beta+\bar{k}}{(1+\bar{m}+\beta)}}_{x}, \underbrace{2-\left(\frac{\beta+\bar{k}-\underline{k}}{(\bar{m}-\alpha+\beta)}+\frac{2}{\beta}\left(\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right)\right)}_{y} \right\} \wedge \bar{k} < 1+\bar{m}$$

note:

$$y > x \Leftrightarrow 2-\left(\frac{\beta+\bar{k}-\underline{k}}{(\bar{m}-\alpha+\beta)}+\frac{2}{\beta}\left(\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right)\right)-\frac{\beta+\bar{k}}{(1+\bar{m}+\beta)} > 0 \Leftrightarrow (1+\alpha)\beta > (1+\bar{m}+\beta)\underline{k}-(1+\alpha)\bar{k} \quad \wedge \quad \underline{k} < 1+\alpha$$

Conditions for existence:

1. $\bar{k} < 1+\bar{m}$

2. $\lambda > \max\left\{ \underbrace{\frac{\beta+\bar{k}}{(1+\bar{m}+\beta)}}_{x}, \underbrace{2-\left(\frac{\beta+\bar{k}-\underline{k}}{(\bar{m}-\alpha+\beta)}+\frac{2}{\beta}\left(\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right)\right)}_{y} \right\}$

$$= \begin{cases} 2-\left(\frac{\beta+\bar{k}-\underline{k}}{(\bar{m}-\alpha+\beta)}+\frac{2}{\beta}\left(\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right)\right), & (1+\alpha)\beta > (1+\bar{m}+\beta)\underline{k}-(1+\alpha)\bar{k} \quad \wedge \quad \underline{k} < 1+\alpha \\ \frac{\beta+\bar{k}}{(1+\bar{m}+\beta)}, & \text{else} \end{cases}$$

note:

$$y > x \Leftrightarrow 2-\left(\frac{\beta+\bar{k}-\underline{k}}{(\bar{m}-\alpha+\beta)}+\frac{2}{\beta}\left(\frac{(1+\bar{m})}{(1+\alpha)}\underline{k}-\bar{k}\right)\right)-\frac{\beta+\bar{k}}{(1+\bar{m}+\beta)} > 0 \Leftrightarrow (1+\alpha)\beta > (1+\bar{m}+\beta)\underline{k}-(1+\alpha)\bar{k} \quad \wedge \quad \underline{k} < 1+\alpha$$

4.3.2. and $CD,\underline{m}/DD,\underline{m} \Rightarrow p_{\underline{m}} = 1$

$$\Pi_{\bar{m}}(CC,\bar{m}) - \Pi_{\bar{m}}(DC,\bar{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(\bar{m}-\alpha) + \left(p_{DC\bar{m}} + p_{DD\bar{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\bar{m}}(-\beta)\right] = 0 \quad \text{(i)}$$

$$\Pi_{\bar{m}}(DC,\bar{m}) - \Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[\left(p_{CC\underline{m}} + p_{CD\underline{m}}\right)(\bar{m}-\alpha) + \left(p_{DC\underline{m}} + p_{DD\underline{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\underline{m}}(-\beta)\right] > 0 \quad \text{(ii)}$$

$$\Pi_{\bar{m}}(CD,\underline{m}) - \Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(\bar{m}-\alpha) + \left(p_{CD\bar{m}} + p_{DD\bar{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\bar{m}}(-\beta)\right] = 0 \quad \text{(iii)}$$

$$\Pi_{\bar{m}}(CC,\underline{m}) - \Pi_{\bar{m}}(CD,\underline{m}) = \Pi_{\bar{m}}(DC,\underline{m}) - \Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[\left(p_{CC\underline{m}} + p_{DC\underline{m}}\right)(\bar{m}-\alpha) + \left(p_{CD\underline{m}} + p_{DD\underline{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\underline{m}}(-\beta)\right] < 0 \quad \text{(iv)}$$

$$\Pi_{\bar{m}}(CC,\bar{m}) - \Pi_{\bar{m}}(CD,\underline{m}) = \Pi_{\bar{m}}(CD,\bar{m}) - \Pi_{\bar{m}}(CD,\underline{m}) - \left[\Pi_{\bar{m}}(CD,\bar{m}) - \Pi_{\bar{m}}(CC,\bar{m})\right] =$$
$$\lambda\left[\left(p_{CD\bar{m}} - p_{DC\bar{m}}\right)(1+\bar{m}+\beta) + \left(p_{CD\underline{m}} - p_{DC\underline{m}}\right)(1+\alpha)\right] - \bar{k} + \lambda\left[\left(p_{CC\underline{m}} + p_{CD\underline{m}}\right)(\bar{m}-\alpha) + \left(p_{DC\underline{m}} + p_{DD\underline{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\underline{m}}(-\beta)\right] = 0$$

$$\Pi_{\underline{m}}(\bar{m}) - \Pi_{\underline{m}}(\underline{m}) = \lambda\left[\left(p_{CD\bar{m}} + p_{CD\underline{m}} - p_{DC\bar{m}} - p_{DC\underline{m}}\right)(1+\alpha)\right] - \underline{k}$$

$$\Pi_{\bar{m}}(CC,\bar{m}) - \Pi_{\bar{m}}(DC,\bar{m}) = \lambda\left[\left(p_{CC\bar{m}}\right)(\bar{m}-\alpha) + \left(p_{DC\bar{m}}\right)(-\beta)\right] = 0 \quad \text{(i)}$$

$$\Pi_{\bar{m}}(DC,\bar{m}) - \Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[\left(p_{CD\underline{m}}\right)(\bar{m}-\alpha) + \left(p_{DD\underline{m}}\right)(-\beta)\right] - (1-\lambda)\beta > 0 \quad \text{(ii)}$$

$$\Pi_{\bar{m}}(CD,\underline{m}) - \Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(\bar{m}-\alpha)\right] = 0 \quad \text{(iii)}$$

$$\Pi_{\bar{m}}(CC,\underline{m}) - \Pi_{\bar{m}}(CD,\underline{m}) = \Pi_{\bar{m}}(DC,\underline{m}) - \Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[\left(p_{CD\underline{m}} + p_{DD\underline{m}}\right)(-\beta)\right] - (1-\lambda)\beta < 0 \quad \text{(iv)}$$

$$\Pi_{\bar{m}}(CC,\bar{m}) - \Pi_{\bar{m}}(CD,\underline{m}) = \lambda\left[\left(-p_{DC\bar{m}}\right)(1+\bar{m}+\beta) + \left(p_{CD\underline{m}}\right)(1+\alpha)\right] - \bar{k} + \lambda\left[\left(p_{CD\underline{m}}\right)(\bar{m}-\alpha) + \left(p_{DD\underline{m}}\right)(-\beta)\right] - (1-\lambda)\beta = 0$$

By (iii) such a **semi**-pooling equilibrium cannot exist.

4.4. $CD,\bar{m}$

4.4.1. and $DC,\underline{m}$

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CD,\bar{m})=\lambda\left[\left(p_{CC\underline{m}}+p_{CD\underline{m}}\right)(\bar{m}-\alpha)+\left(p_{DC\underline{m}}+p_{DD\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]<0 \quad \text{(i)}$$

$$\Pi_{\bar{m}}(CD,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{DC\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]>0 \quad \text{(ii)}$$

$$\Pi_{\bar{m}}(DC,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[\left(p_{CC\underline{m}}+p_{DC\underline{m}}\right)(\bar{m}-\alpha)+\left(p_{CD\underline{m}}+p_{DD\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]>0 \quad \text{(iii)}$$

$$\Pi_{\bar{m}}(CC,\underline{m})-\Pi_{\bar{m}}(DC,\underline{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{CD\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]<0 \quad \text{(iv)}$$

$$\Pi_{\bar{m}}(CD,\bar{m})-\Pi_{\bar{m}}(DC,\underline{m})=\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(DC,\underline{m})-\left[\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(CD,\bar{m})\right]=0 \quad \text{(v)}$$

$$\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})=\lambda\left[\left(p_{CD\bar{m}}+p_{CD\underline{m}}-p_{DC\bar{m}}-p_{DC\underline{m}}\right)(1+\alpha)\right]-\underline{k}$$

$$\Pi_{\bar{m}}(CC,\bar{m})-\Pi_{\bar{m}}(CD,\bar{m})=\lambda\left[\left(p_{DC\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]<0 \quad \text{(i)}$$

$$\Pi_{\bar{m}}(CD,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left[\left(p_{CD\bar{m}}\right)(\bar{m}-\alpha)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]>0 \quad \text{(ii)}$$

$$\Pi_{\bar{m}}(DC,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[\left(p_{DC\underline{m}}\right)(\bar{m}-\alpha)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]>0 \quad \text{(iii)}$$

$$\Pi_{\bar{m}}(CC,\underline{m})-\Pi_{\bar{m}}(DC,\underline{m})=\lambda\left[\left(p_{CD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]<0 \quad \text{(iv)}$$

$$\Pi_{\bar{m}}(CD,\bar{m})-\Pi_{\bar{m}}(DC,\underline{m})=\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(DC,\underline{m})-\left[\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(CD,\bar{m})\right]=0$$

$$\lambda\left[\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)(1+\alpha)+\left(p_{CD\underline{m}}-p_{DC\underline{m}}\right)(1+\bar{m}+\beta)\right]-\bar{k}$$

$$+\lambda\left[\left(p_{CC\bar{m}}+p_{CD\bar{m}}-p_{CC\underline{m}}-p_{CD\underline{m}}\right)(\bar{m}-\alpha)+\left(p_{DC\bar{m}}+p_{DD\bar{m}}-p_{DC\underline{m}}-p_{DD\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)-p_{\underline{m}}(-\beta)\right] \quad \text{(v)}$$

$$=\lambda\left[\left(p_{CD\bar{m}}\right)(1+\alpha)+\left(-p_{DC\underline{m}}\right)(1+\bar{m}+\beta)\right]-\bar{k}+\lambda\left[\left(p_{CD\bar{m}}\right)(\bar{m}-\alpha)+\left(-p_{DC\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)-p_{\underline{m}}(-\beta)\right]$$

$$=\lambda\left[\left(p_{CD\bar{m}}-p_{DC\underline{m}}\right)(1+\bar{m})\right]-\bar{k}-\beta(1-\lambda)\left[p_{\bar{m}}-p_{\underline{m}}\right]=0$$

$$\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})=\lambda\left[\left(p_{CD\bar{m}}-p_{DC\underline{m}}\right)(1+\alpha)\right]-\underline{k} \quad \text{(vi)}$$

(i)and (iv) are always satisfied.

4.4.1.1. $\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})>0$

$$\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})=\lambda\left[\left(p_{CD\bar{m}}-p_{DC\underline{m}}\right)(1+\alpha)\right]-\underline{k}>0\Rightarrow p_{\bar{m}}=1 \quad \text{(vi)}$$

$$\Pi_{\bar{m}}(CD,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left[\left(p_{CD\bar{m}}\right)(\bar{m}-\alpha)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]>0 \quad \text{(ii)}$$

$$\Pi_{\bar{m}}(DC,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[\left(p_{DC\underline{m}}\right)(\bar{m}-\alpha)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]>0 \quad \text{(iii)}$$

$$\Pi_{\bar{m}}(CD,\bar{m})-\Pi_{\bar{m}}(DC,\underline{m})=\lambda\left[\left(p_{CD\bar{m}}-p_{DC\underline{m}}\right)(1+\bar{m})\right]-\bar{k}-\beta(1-\lambda)=0 \quad \text{(v)}$$

EQ: $p_{DC\underline{m}}=\frac{1}{2}-\frac{\bar{k}+\beta(1-\lambda)}{2\lambda(1+\bar{m})}$, $p_{CD\bar{m}}=\frac{1}{2}+\frac{\bar{k}+\beta(1-\lambda)}{2\lambda(1+\bar{m})}$

(iii) is always satisfied, (ii) and (vi) need to be checked for::

$$(vi): \lambda < 1 - \frac{1}{\beta}\left(\underline{k}\frac{1+\bar{m}}{1+\alpha} - \bar{k}\right)$$

$$(ii): \lambda > \frac{1+\alpha}{(\bar{m}-\alpha+\beta)(1+\bar{m})+(1+\alpha)\beta}\left(\beta\left(1+\frac{1+\bar{m}}{1+\alpha}\right) - \frac{\bar{k}}{1+\alpha}(\bar{m}-\alpha)\right)$$

$(vi) \wedge (ii) \wedge \lambda \in (0,1)$:

1. $\bar{k} < 1+\alpha$
2. $\frac{1+\alpha}{(\bar{m}-\alpha+\beta)(1+\bar{m})+(1+\alpha)\beta}\left(\beta\left(1+\frac{1+\bar{m}}{1+\alpha}\right) - \frac{\bar{k}}{1+\alpha}(\bar{m}-\alpha)\right) < \lambda < 1 - \frac{1}{\beta}\left(\underline{k}\frac{1+\bar{m}}{1+\alpha} - \bar{k}\right)$
3. $(1+\alpha)\beta > (\bar{m}-\alpha+\beta)\bar{k}\ (\Rightarrow 1.)$
4. $\underline{k} < \frac{(1+\alpha)\left(\beta(\bar{m}-\alpha)+\bar{k}(\bar{m}-\alpha+2\beta)\right)}{-\alpha(1-\beta)+2\beta+\bar{m}(1+\bar{m}-\alpha+\beta)}$

Finally all conditions of the type

$p \in [0,1], \sum p < 1$ reduce to: $\frac{1}{2} + \frac{\bar{k}+\beta(1-\lambda)}{2\lambda(1+\bar{m})}$

$p_{CD\bar{m}} > 0$: true

$p_{CD\bar{m}} < 1 :\Leftrightarrow \lambda > \frac{\beta+\bar{k}}{(1+\bar{m}+\beta)} \wedge \bar{k} < 1+\bar{m}$

Conditions:

1. $$\max\left\{\underbrace{\frac{1+\alpha}{(\bar{m}-\alpha+\beta)(1+\bar{m})+(1+\alpha)\beta}\left(\beta\left(1+\frac{1+\bar{m}}{1+\alpha}\right)-\frac{\bar{k}}{1+\alpha}(\bar{m}-\alpha)\right)}_{x},\underbrace{\frac{\beta+\bar{k}}{(1+\bar{m}+\beta)}}_{y}\right\}<\lambda<\underbrace{1-\frac{1}{\beta}\left(\underline{k}\frac{1+\bar{m}}{1+\alpha}-\bar{k}\right)}_{z}$$
2. $(1+\alpha)\beta>(\bar{m}-\alpha+\beta)\bar{k}$
3. $\underline{k}<\dfrac{(1+\alpha)\left(\beta(\bar{m}-\alpha)+\bar{k}(\bar{m}-\alpha+2\beta)\right)}{-\alpha(1-\beta)+2\beta+\bar{m}(1+\bar{m}-\alpha+\beta)}$

note that 2. implies that $x>y$.

$$x<z:\Leftrightarrow \underline{k}<\frac{(1+\alpha)(\bar{m}-\alpha+2\bar{k})}{(2+\bar{m}+\alpha)} \qquad \wedge \qquad (1+\alpha)\beta>\left(1+\bar{m}+\frac{(1+\bar{m})\beta+(1+\alpha)\beta}{(\bar{m}-\alpha)}\right)\underline{k}-\left(1+\alpha+\frac{2(1+\alpha)\beta}{(\bar{m}-\alpha)}\right)\bar{k}$$

note:

a) $$(\bar{m}-\alpha+\beta)\bar{k}-\left(1+\bar{m}+\frac{(1+\bar{m})\beta+(1+\alpha)\beta}{(\bar{m}-\alpha)}\right)\underline{k}+\left(1+\alpha+\frac{2(1+\alpha)\beta}{(\bar{m}-\alpha)}\right)\bar{k}=$$

$$\left(1+\bar{m}+\frac{(1+\bar{m})\beta+(1+\alpha)\beta}{(\bar{m}-\alpha)}\right)\bar{k}-\left(1+\bar{m}+\frac{(1+\bar{m})\beta+(1+\alpha)\beta}{(\bar{m}-\alpha)}\right)\underline{k}<0$$

b) $$\frac{(1+\alpha)\left(\beta(\bar{m}-\alpha)+\bar{k}(\bar{m}-\alpha+2\beta)\right)}{-\alpha(1-\beta)+2\beta+\bar{m}(1+\bar{m}-\alpha+\beta)}-\frac{(1+\alpha)(\bar{m}-\alpha+2\bar{k})}{(2+\bar{m}+\alpha)}=$$

$$\frac{(1+\alpha)\beta(\bar{m}-\alpha)}{(1+\bar{m})(\bar{m}-\alpha+\beta)+(1+\alpha)\beta}-\frac{(1+\alpha)(\bar{m}-\alpha)}{(2+\bar{m}+\alpha)}+\frac{(1+\alpha)(\bar{m}-\alpha+2\beta)\bar{k}}{(1+\bar{m})(\bar{m}-\alpha+\beta)+(1+\alpha)\beta}-\frac{(1+\alpha)2\bar{k}}{(2+\bar{m}+\alpha)}<0,$$ i.e.

$\underline{k}<\dfrac{(1+\alpha)\left(\beta(\bar{m}-\alpha)+\bar{k}(\bar{m}-\alpha+2\beta)\right)}{-\alpha(1-\beta)+2\beta+\bar{m}(1+\bar{m}-\alpha+\beta)}$ is binding.

To summarize:

EQ: $p_{DC\underline{m}} = \frac{1}{2} - \frac{\bar{k} + \beta(1-\lambda)}{2\lambda(1+\bar{m})}$, $p_{CD\bar{m}} = \frac{1}{2} + \frac{\bar{k} + \beta(1-\lambda)}{2\lambda(1+\bar{m})}$

Conditions for existence:

1. $0 \overset{2.}{<} \frac{1+\alpha}{(\bar{m}-\alpha+\beta)(1+\bar{m})+(1+\alpha)\beta}\left(\beta\left(1+\frac{1+\bar{m}}{1+\alpha}\right) - \frac{\bar{k}}{1+\alpha}(\bar{m}-\alpha)\right) < \lambda < 1 - \frac{1}{\beta}\left(\underline{k}\frac{1+\bar{m}}{1+\alpha} - \bar{k}\right) < 1$
2. $(1+\alpha)\beta > (\bar{m}-\alpha+\beta)\bar{k}$
3. $\underline{k} < \frac{(1+\alpha)\left(\beta(\bar{m}-\alpha)+\bar{k}(\bar{m}-\alpha+2\beta)\right)}{(\bar{m}-\alpha+\beta)(1+\bar{m})+(1+\alpha)\beta}$

4.4.1.2. $\Pi_{\underline{m}}(\bar{m}) - \Pi_{\underline{m}}(\underline{m}) = 0$

$$\Pi_{\underline{m}}(\bar{m}) - \Pi_{\underline{m}}(\underline{m}) = \lambda\left[\left(p_{CD\bar{m}} - p_{DC\underline{m}}\right)(1+\alpha)\right] - \underline{k} = 0 \Leftrightarrow \left(p_{CD\bar{m}} - p_{DC\underline{m}}\right) = \frac{\underline{k}}{\lambda(1+\alpha)} \qquad \text{(vi)}$$

$$\Pi_{\bar{m}}(CD,\bar{m}) - \Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[\left(p_{CD\bar{m}}\right)(\bar{m}-\alpha)\right] + (1-\lambda)\left[p_{\bar{m}}(-\beta)\right] > 0 \qquad \text{(ii)}$$

$$\Pi_{\bar{m}}(DC,\underline{m}) - \Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[\left(p_{DC\underline{m}}\right)(\bar{m}-\alpha)\right] + (1-\lambda)\left[p_{\underline{m}}(-\beta)\right] > 0 \qquad \text{(iii)}$$

(v) becomes:

$$\Pi_{\bar{m}}(CD,\bar{m}) - \Pi_{\bar{m}}(DC,\underline{m}) = \lambda\left[\left(p_{CD\bar{m}} - p_{DC\underline{m}}\right)(1+\bar{m})\right] - \bar{k} - \beta(1-\lambda)\left[p_{\bar{m}} - p_{\underline{m}}\right] = 0 \Leftrightarrow \frac{(1+\bar{m})}{(1+\alpha)}\underline{k} - \bar{k} = \beta(1-\lambda)\left[p_{\bar{m}} - p_{\underline{m}}\right]$$

$$\Leftrightarrow \frac{(1+\bar{m})}{(1+\alpha)}\underline{k} - \bar{k} = \beta(1-\lambda)\left[2p_{\bar{m}} - 1\right] \Leftrightarrow p_{\bar{m}} = \frac{1}{2} + \frac{1}{2\beta(1-\lambda)}\left[\frac{(1+\bar{m})}{(1+\alpha)}\underline{k} - \bar{k}\right]$$

$p_{CD\bar{m}} = \frac{1}{2} + \frac{\underline{k}}{2\lambda(1+\alpha)}, p_{DC\underline{m}} = \frac{1}{2} - \frac{\underline{k}}{2\lambda(1+\alpha)}$ by (vi) ; $p_{\bar{m}} = \frac{1}{2} + \frac{1}{2\beta(1-\lambda)}\left[\frac{(1+\bar{m})}{(1+\alpha)}\underline{k} - \bar{k}\right]$ by (v) and

Given this solution for the respective shares (ii) and (iii) need to be checked for.

$(ii): \lambda(\bar{m} - \alpha + \beta) > \beta + \underline{k} - \bar{k}$

$(iii): \lambda(\bar{m} - \alpha + \beta) > \beta - (\underline{k} - \bar{k})$, i.e. (ii) implies (iii).

$\lambda \in (0,1)$ adds $\underline{k} - \bar{k} < \bar{m} - \alpha$

Finally all conditions of the type

$p \in [0,1], \sum p < 1$ reduce to:

$$\frac{\underline{k}}{1+\alpha} < \lambda < 1 - \frac{1}{\beta}\left(\frac{(1+\bar{m})}{(1+\alpha)}\underline{k} - \bar{k}\right)\left(\Rightarrow \text{nec.:} \frac{\underline{k}}{1+\alpha} < 1\right)$$

$$(1+\alpha)\beta > (1+\bar{m}+\beta)\underline{k} - (1+\alpha)\bar{k}$$

Conditions for existence:

1. $\frac{\beta + \underline{k} - \bar{k}}{(\bar{m} - \alpha + \beta)} < \lambda < 1 - \frac{1}{\beta}\left(\frac{(1+\bar{m})}{(1+\alpha)}\underline{k} - \bar{k}\right)$
2. $(1+\alpha)\beta > (\bar{m} - \alpha + \beta)\bar{k}$
3. $\underline{k} < \frac{(1+\alpha)\left(\beta(\bar{m}-\alpha) + \bar{k}(\bar{m} - \alpha + 2\beta)\right)}{(\bar{m} - \alpha + \beta)(1+\bar{m}) + (1+\alpha)\beta}$
4. $\bar{k} < 1 + \alpha$

4.4.1.3. $\Pi_{\underline{m}}(\bar{m}) - \Pi_{\underline{m}}(\underline{m}) < 0$

$$\Pi_{\underline{m}}(\bar{m}) - \Pi_{\underline{m}}(\underline{m}) = \lambda\left[\left(p_{CD\bar{m}} - p_{DC\underline{m}}\right)(1+\alpha)\right] - \underline{k} < 0 \Leftrightarrow p_{\underline{m}} = 1 \quad \text{(vi)}$$

$$\Pi_{\bar{m}}(CD,\bar{m}) - \Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[\left(p_{CD\bar{m}}\right)(\bar{m}-\alpha)\right] > 0 \quad \text{(ii)}$$

$$\Pi_{\bar{m}}(DC,\underline{m}) - \Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[\left(p_{DC\underline{m}}\right)(\bar{m}-\alpha)\right] - \beta(1-\lambda) > 0 \quad \text{(iii)}$$

$$\Pi_{\bar{m}}(CD,\bar{m}) - \Pi_{\bar{m}}(DC,\underline{m}) = \lambda\left[\left(p_{CD\bar{m}} - p_{DC\underline{m}}\right)(1+\bar{m})\right] - \bar{k} + \beta(1-\lambda) = 0 \quad \text{(v)}$$

$$\text{EQ:} \quad \begin{aligned} p_{CD\bar{m}} &= \frac{1}{2}\left(1 + \frac{\bar{k} - \beta(1-\lambda)}{\lambda(1+\bar{m})}\right) \\ p_{DC\underline{m}} &= \frac{1}{2}\left(1 - \frac{\bar{k} - \beta(1-\lambda)}{\lambda(1+\bar{m})}\right) \end{aligned}$$

(ii) is satisfied, (iii) and (vi) need to be checked for,

$$\text{(vi)}: \lambda < 1 + \frac{1}{\beta}\left(\frac{(1+\bar{m})}{(1+\alpha)}\underline{k} - \bar{k}\right)$$

$$\text{(iii)}: \lambda > \frac{\bar{k}(\bar{m}-\alpha) + \beta(2+\alpha+\bar{m})}{(1+\bar{m})(\bar{m}-\alpha+\beta) + \beta(1+\alpha)}$$

hence adding $\lambda \in (0,1)$: $\dfrac{\bar{k}(\bar{m}-\alpha) + \beta(2+\alpha+\bar{m})}{(1+\bar{m})(\bar{m}-\alpha+\beta) + \beta(1+\alpha)} < \lambda < 1 \qquad \wedge \qquad \bar{k} < (1+\bar{m})$

Finally all conditions of the type:

$p \in [0,1], \sum p < 1$ reduce to:

$$0 < 1 + \frac{\bar{k} - \beta(1-\lambda)}{\lambda(1+\bar{m})} < 2 \Leftrightarrow -\lambda(1+\bar{m}+\beta) < \bar{k} - \beta < \lambda(1+\bar{m}-\beta) \Leftrightarrow$$

$$\lambda(1+\bar{m}-\beta) > \bar{k} - \beta \wedge \lambda(1+\bar{m}+\beta) > \bar{k} - \beta \Leftrightarrow \quad \lambda(1+\bar{m}-\beta) > \bar{k} - \beta \wedge \lambda > \frac{\bar{k} - \beta}{(1+\bar{m}+\beta)}$$

adding $\lambda \in (0,1)$ and $\bar{k} < 1+\bar{m}$ yields $x < \lambda < 1$, $x = \begin{cases} \dfrac{\bar{k} - \beta}{1+\bar{m}-\beta}, \beta \leq \bar{k} \\ \dfrac{\bar{k} - \beta}{1+\bar{m}+\beta}, \beta > \bar{k} \end{cases}$

however, it turns out that $\dfrac{\bar{k}(\bar{m}-\alpha)+\beta(2+\alpha+\bar{m})}{(1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha)} < \lambda$ is binding.

To summarize:

$$p_{CD\bar{m}} = \frac{1}{2}\left(1 + \frac{\bar{k} - \beta(1-\lambda)}{\lambda(1+\bar{m})}\right), \; p_{DC\underline{m}} = \frac{1}{2}\left(1 - \frac{\bar{k} - \beta(1-\lambda)}{\lambda(1+\bar{m})}\right)$$

Conditions for existence:

1. $\dfrac{\bar{k}(\bar{m}-\alpha)+\beta(2+\alpha+\bar{m})}{(1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha)} < \lambda < 1$
2. $\bar{k} < (1+\bar{m})$

4.4.2. and $DD, \underline{m} \Rightarrow p_{\underline{m}} = 1$

$$\Pi_{\bar{m}}(CC,\bar{m}) - \Pi_{\bar{m}}(CD,\bar{m}) = \lambda\left[\left(p_{CC\underline{m}} + p_{CD\underline{m}}\right)(\bar{m}-\alpha) + \left(p_{DC\underline{m}} + p_{DD\underline{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\underline{m}}(-\beta)\right] < 0 \quad \text{(i)}$$

$$\Pi_{\bar{m}}(CD,\bar{m}) - \Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(\bar{m}-\alpha) + \left(p_{DC\bar{m}} + p_{DD\bar{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\bar{m}}(-\beta)\right] > 0 \quad \text{(ii)}$$

$$\Pi_{\bar{m}}(DC,\underline{m}) - \Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[\left(p_{CC\underline{m}} + p_{DC\underline{m}}\right)(\bar{m}-\alpha) + \left(p_{CD\underline{m}} + p_{DD\underline{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\underline{m}}(-\beta)\right] < 0 \quad \text{(iii)}$$

$$\Pi_{\bar{m}}(CD,\underline{m}) - \Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(\bar{m}-\alpha) + \left(p_{CD\bar{m}} + p_{DD\bar{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\bar{m}}(-\beta)\right] < 0 \quad \text{(iv)}$$

$$\Pi_{\bar{m}}(CD,\bar{m}) - \Pi_{\bar{m}}(DD,\underline{m}) = \Pi_{\bar{m}}(DD,\bar{m}) - \Pi_{\bar{m}}(DD,\underline{m}) - \left[\Pi_{\bar{m}}(DD,\bar{m}) - \Pi_{\bar{m}}(CD,\bar{m})\right] = 0 \quad \text{(v)}$$

$$\Pi_{\bar{m}}(CC,\bar{m}) - \Pi_{\bar{m}}(CD,\bar{m}) = \lambda\left[\left(p_{DD\underline{m}}\right)(-\beta)\right] - \beta(1-\lambda) < 0 \quad \text{(i)}$$

$$\Pi_{\bar{m}}(CD,\bar{m}) - \Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[\left(p_{CD\bar{m}}\right)(\bar{m}-\alpha)\right] > 0 \quad \text{(ii)}$$

$$\Pi_{\bar{m}}(DC,\underline{m}) - \Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[\left(p_{DD\underline{m}}\right)(-\beta)\right] - \beta(1-\lambda) < 0 \quad \text{(iii)}$$

$$\Pi_{\bar{m}}(CD,\underline{m}) - \Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[\left(p_{CD\bar{m}}\right)(-\beta)\right] < 0 \quad \text{(iv)}$$

$$\Pi_{\bar{m}}(CD,\bar{m}) - \Pi_{\bar{m}}(DD,\underline{m}) = \Pi_{\bar{m}}(DD,\bar{m}) - \Pi_{\bar{m}}(DD,\underline{m}) - \left[\Pi_{\bar{m}}(DD,\bar{m}) - \Pi_{\bar{m}}(CD,\bar{m})\right] =$$
$$\lambda(1+\alpha)\, p_{CD\bar{m}} - \bar{k} + \lambda\left[\left(p_{CD\bar{m}}\right)(\bar{m}-\alpha)\right] = 0 \Leftrightarrow p_{CD\bar{m}} = \frac{\bar{k}}{\lambda(1+\bar{m})}$$

EQ: $p_{CD\bar{m}} = \dfrac{\bar{k}}{\lambda(1+\bar{m})}$, $p_{DD\underline{m}} = 1 - p_{CD\bar{m}}$

(i),(ii),(iii) and (iv) are satisfied

Finally all conditions of the type $p \in [0,1], \sum p < 1$ reduce to: $\lambda > \dfrac{\bar{k}}{(1+\bar{m})}$

Condition for existence:

1. $\lambda > \dfrac{\bar{k}}{(1+\bar{m})}$
2. $\bar{k} < (1+\bar{m})$

however it turns out that this equilibrium is not stable.

4.5. $DD,\bar{m}$ and $DC,\underline{m}$ $\Rightarrow p_{\underline{m}} = 1$

$$\Pi_{\bar{m}}(DC,\bar{m}) - \Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[\left(p_{CC\underline{m}} + p_{CD\underline{m}}\right)(\bar{m}-\alpha) + \left(p_{DC\underline{m}} + p_{DD\underline{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\underline{m}}(-\beta)\right] < 0 \qquad \text{(i)}$$

$$\Pi_{\bar{m}}(CD,\bar{m}) - \Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{CD\bar{m}}\right)(\bar{m}-\alpha) + \left(p_{DC\bar{m}} + p_{DD\bar{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\bar{m}}(-\beta)\right] < 0 \qquad \text{(ii)}$$

$$\Pi_{\bar{m}}(DC,\underline{m}) - \Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[\left(p_{CC\underline{m}} + p_{DC\underline{m}}\right)(\bar{m}-\alpha) + \left(p_{CD\underline{m}} + p_{DD\underline{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\underline{m}}(-\beta)\right] > 0 \qquad \text{(iii)}$$

$$\Pi_{\bar{m}}(CC,\underline{m}) - \Pi_{\bar{m}}(DC,\underline{m}) = \lambda\left[\left(p_{CC\bar{m}} + p_{DC\bar{m}}\right)(\bar{m}-\alpha) + \left(p_{CD\bar{m}} + p_{DD\bar{m}}\right)(-\beta)\right] + (1-\lambda)\left[p_{\bar{m}}(-\beta)\right] < 0 \qquad \text{(iv)}$$

$$\Pi_{\bar{m}}(DD,\bar{m}) - \Pi_{\bar{m}}(DC,\underline{m}) = \Pi_{\bar{m}}(DC,\bar{m}) - \Pi_{\bar{m}}(DC,\underline{m}) - \left[\Pi_{\bar{m}}(DC,\bar{m}) - \Pi_{\bar{m}}(DD,\bar{m})\right] = 0 \qquad \text{(v)}$$

$$\Pi_{\bar{m}}\left(DC,\bar{m}\right)-\Pi_{\bar{m}}\left(DD,\bar{m}\right)=\lambda\left[\left(p_{DC\underline{m}}\right)\left(-\beta\right)\right]-\beta\left(1-\lambda\right)<0 \qquad \text{(i)}$$

$$\Pi_{\bar{m}}\left(CD,\bar{m}\right)-\Pi_{\bar{m}}\left(DD,\bar{m}\right)=\lambda\left[\left(p_{DD\bar{m}}\right)\left(-\beta\right)\right]<0 \qquad \text{(ii)}$$

$$\Pi_{\bar{m}}\left(DC,\underline{m}\right)-\Pi_{\bar{m}}\left(DD,\underline{m}\right)=\lambda\left[\left(p_{DC\underline{m}}\right)\left(\bar{m}-\alpha\right)\right]-\beta\left(1-\lambda\right)>0 \qquad \text{(iii)}$$

$$\Pi_{\bar{m}}\left(CC,\underline{m}\right)-\Pi_{\bar{m}}\left(DC,\underline{m}\right)=\lambda\left[\left(p_{DD\bar{m}}\right)\left(-\beta\right)\right]<0 \qquad \text{(iv)}$$

$$\Pi_{\bar{m}}\left(DD,\bar{m}\right)-\Pi_{\bar{m}}\left(DC,\underline{m}\right)=\Pi_{\bar{m}}\left(DC,\bar{m}\right)-\Pi_{\bar{m}}\left(DC,\underline{m}\right)-\left[\Pi_{\bar{m}}\left(DC,\bar{m}\right)-\Pi_{\bar{m}}\left(DD,\bar{m}\right)\right]=$$

$$-p_{DC\underline{m}}\lambda\left(1+\bar{m}\right)-\bar{k}+\left(1-\lambda\right)\beta=0\Leftrightarrow\frac{\left(1-\lambda\right)\beta-\bar{k}}{\lambda\left(1+\bar{m}\right)}=p_{DC\underline{m}}$$

$$\Pi_{\bar{m}}\left(DC,\bar{m}\right)-\Pi_{\bar{m}}\left(DC,\underline{m}\right)=\lambda\left[\left(p_{CD\bar{m}}-p_{DC\bar{m}}\right)\left(1+\alpha\right)+\left(p_{CD\underline{m}}-p_{DC\underline{m}}\right)\left(1+\bar{m}+\beta\right)\right]-\bar{k}$$

$$\Pi_{\bar{m}}\left(DC,\bar{m}\right)-\Pi_{\bar{m}}\left(DD,\bar{m}\right)=\lambda\left[\left(p_{DC\underline{m}}\right)\left(-\beta\right)\right]-\beta\left(1-\lambda\right)<0$$

By (v) $p_{DC\underline{m}}=\dfrac{\left(1-\lambda\right)\beta-\bar{k}}{\lambda\left(1+\bar{m}\right)}$, however this is incompatible with $p_{DC\underline{m}}>\dfrac{\left(1-\lambda\right)\beta}{\lambda\left(\bar{m}-\alpha\right)}$ by (iii)

Hence such a semi-pooling equilibrium cannot exist.

4.6. $CD,\bar{m}\,/\,DD,\bar{m}$

4.6.1. and $DC,\underline{m}$

$$\Pi_{\bar{m}}\left(DC,\bar{m}\right)-\Pi_{\bar{m}}\left(DD,\bar{m}\right)=\lambda\left[\left(p_{CC\underline{m}}+p_{CD\underline{m}}\right)\left(\bar{m}-\alpha\right)+\left(p_{DC\underline{m}}+p_{DD\underline{m}}\right)\left(-\beta\right)\right]+\left(1-\lambda\right)\left[p_{\underline{m}}\left(-\beta\right)\right]<0 \qquad \text{(i)}$$

$$\Pi_{\bar{m}}\left(CD,\bar{m}\right)-\Pi_{\bar{m}}\left(DD,\bar{m}\right)=\lambda\left[\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)\left(\bar{m}-\alpha\right)+\left(p_{DC\bar{m}}+p_{DD\bar{m}}\right)\left(-\beta\right)\right]+\left(1-\lambda\right)\left[p_{\bar{m}}\left(-\beta\right)\right]=0 \qquad \text{(ii)}$$

$$\Pi_{\bar{m}}\left(DC,\underline{m}\right)-\Pi_{\bar{m}}\left(DD,\underline{m}\right)=\lambda\left[\left(p_{CC\underline{m}}+p_{DC\underline{m}}\right)\left(\bar{m}-\alpha\right)+\left(p_{CD\underline{m}}+p_{DD\underline{m}}\right)\left(-\beta\right)\right]+\left(1-\lambda\right)\left[p_{\underline{m}}\left(-\beta\right)\right]>0 \qquad \text{(iii)}$$

$$\Pi_{\bar{m}}\left(CC,\underline{m}\right)-\Pi_{\bar{m}}\left(DC,\underline{m}\right)=\lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)\left(\bar{m}-\alpha\right)+\left(p_{CD\bar{m}}+p_{DD\bar{m}}\right)\left(-\beta\right)\right]+\left(1-\lambda\right)\left[p_{\bar{m}}\left(-\beta\right)\right]<0 \qquad \text{(iv)}$$

$$\Pi_{\bar{m}}\left(DD,\bar{m}\right)-\Pi_{\bar{m}}\left(DC,\underline{m}\right)=\Pi_{\bar{m}}\left(DC,\bar{m}\right)-\Pi_{\bar{m}}\left(DC,\underline{m}\right)-\left[\Pi_{\bar{m}}\left(DC,\bar{m}\right)-\Pi_{\bar{m}}\left(DD,\bar{m}\right)\right]=0 \qquad \text{(v)}$$

$$\Pi_{\bar{m}}\left(DC,\bar{m}\right)-\Pi_{\bar{m}}\left(DD,\bar{m}\right)=\lambda\left[\left(p_{DC\underline{m}}\right)\left(-\beta\right)\right]+\left(1-\lambda\right)\left[p_{\underline{m}}\left(-\beta\right)\right]<0 \qquad \text{(i)}$$

$$\Pi_{\bar{m}}\left(CD,\bar{m}\right)-\Pi_{\bar{m}}\left(DD,\bar{m}\right)=\lambda\left[\left(p_{CD\bar{m}}\right)\left(\bar{m}-\alpha\right)+\left(p_{DD\bar{m}}\right)\left(-\beta\right)\right]+\left(1-\lambda\right)\left[p_{\bar{m}}\left(-\beta\right)\right]=0 \qquad \text{(ii)}$$

$$\Pi_{\bar{m}}\left(DC,\underline{m}\right)-\Pi_{\bar{m}}\left(DD,\underline{m}\right)=\lambda\left[\left(p_{DC\underline{m}}\right)\left(\bar{m}-\alpha\right)\right]+\left(1-\lambda\right)\left[p_{\underline{m}}\left(-\beta\right)\right]>0 \qquad \text{(iii)}$$

$$\Pi_{\bar{m}}\left(CC,\underline{m}\right)-\Pi_{\bar{m}}\left(DC,\underline{m}\right)=\lambda\left[\left(p_{CD\bar{m}}+p_{DD\bar{m}}\right)\left(-\beta\right)\right]+\left(1-\lambda\right)\left[p_{\bar{m}}\left(-\beta\right)\right]<0 \qquad \text{(iv)}$$

$$\Pi_{\bar{m}}(DD,\bar{m})-\Pi_{\bar{m}}(DC,\underline{m})=\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(DC,\underline{m})-\left[\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})\right]=$$
$$\lambda\left[(p_{CD\bar{m}}-p_{DC\bar{m}})(1+\alpha)+(p_{CD\underline{m}}-p_{DC\underline{m}})(1+\bar{m}+\beta)\right]-\bar{k}-\lambda\left[(p_{CC\underline{m}}+p_{CD\underline{m}})(\bar{m}-\alpha)+(p_{DC\underline{m}}+p_{DD\underline{m}})(-\beta)\right]-(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]=$$
$$\lambda\left[(p_{CD\bar{m}})(1+\alpha)+(-p_{DC\underline{m}})(1+\bar{m}+\beta)\right]-\bar{k}-\lambda\left[+(p_{DC\underline{m}})(-\beta)\right]+\beta(1-\lambda)\,p_{\underline{m}}= \quad \text{(v)}$$
$$\lambda\left[(p_{CD\bar{m}})(1+\alpha)-p_{DC\underline{m}}(1+\bar{m})\right]-\bar{k}+\beta(1-\lambda)\,p_{\underline{m}}=0$$

$$\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})=\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})=\lambda\left[(p_{CD\bar{m}}-p_{DC\underline{m}})(1+\alpha)\right]-\underline{k}$$

In a semi-pooling equilibrium (i) and (iv) will always be satisfied.

4.6.1.1. $\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})>0$

$$\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})=\lambda\left[(p_{CD\bar{m}}-p_{DC\underline{m}})(1+\alpha)\right]-\underline{k}>0 \Rightarrow p_{\underline{m}}=0 \qquad (*)$$

$$\Pi_{\bar{m}}(CD,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left[(p_{CD\bar{m}})(\bar{m}-\alpha)+(p_{DD\bar{m}})(-\beta)\right]-\beta(1-\lambda)=0 \qquad \text{(ii)}$$

$$\Pi_{\bar{m}}(DC,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[(p_{DC\underline{m}})(\bar{m}-\alpha)\right]>0 \qquad \text{(iii)}$$

$$\Pi_{\bar{m}}(DD,\bar{m})-\Pi_{\bar{m}}(DC,\underline{m})=\lambda\left[(p_{CD\bar{m}})(1+\alpha)-p_{DC\underline{m}}(1+\bar{m})\right]-\bar{k}=0 \qquad \text{(v)}$$

EQ:

$$p_{DC\underline{m}}=\frac{1}{\lambda}\left[\frac{\beta(1+\alpha)-\bar{k}(\bar{m}-\alpha+\beta)}{\beta(1+\alpha)+(1+\bar{m})(\bar{m}-\alpha+\beta)}\right],\ p_{CD\bar{m}}=\frac{\beta}{\lambda(\bar{m}-\alpha+\beta)}-p_{DC\underline{m}}\frac{\beta}{\bar{m}-\alpha+\beta},\ p_{DD\bar{m}}=1-p_{CD\bar{m}}-p_{DC\underline{m}}\Leftrightarrow$$
$$p_{DC\underline{m}}=\frac{1}{\lambda}\left[\frac{\beta(1+\alpha)-\bar{k}(\bar{m}-\alpha+\beta)}{\beta(1+\alpha)+(1+\bar{m})(\bar{m}-\alpha+\beta)}\right],\ p_{CD\bar{m}}=\frac{1}{\lambda}\left[\frac{\beta(1+\bar{m}+\bar{k})}{\beta(1+\alpha)+(1+\bar{m})(\bar{m}-\alpha+\beta)}\right],\ p_{DD\bar{m}}=1+\frac{1}{\lambda}\left[\frac{\bar{k}(\bar{m}-\alpha)-\beta(2+\bar{m}+\alpha)}{\beta(1+\alpha)+(1+\bar{m})(\bar{m}-\alpha+\beta)}\right]$$

(iii) is satisfied, (*) needs to be checked for, however this equilibrium is not stable.

(*) reduces to:

1. $\underline{k} < \dfrac{(1+\alpha)\left(\beta(\bar{m}-\alpha)+\bar{k}(\bar{m}-\alpha+2\beta)\right)}{(1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha)}$
2. $\beta(1+\alpha) > \bar{k}(\bar{m}-\alpha+\beta)$

Finally all conditions of the type

$p \in [0,1], \sum p < 1$ and $0 < \lambda < 1$ reduce to:

1. $\dfrac{\bar{k}(\bar{m}-\alpha)+\beta(2+\alpha+\bar{m})}{(1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha)} < \lambda < 1$
2. $\beta(1+\alpha) > \bar{k}(\bar{m}-\alpha+\beta)$

To summarize

$$p_{DC\underline{m}} = \frac{1}{\lambda}\left[\frac{\beta(1+\alpha)-\bar{k}(\bar{m}-\alpha+\beta)}{\beta(1+\alpha)+(1+\bar{m})(\bar{m}-\alpha+\beta)}\right], p_{CD\bar{m}} = \frac{1}{\lambda}\left[\frac{\beta(1+\bar{m}+\bar{k})}{\beta(1+\alpha)+(1+\bar{m})(\bar{m}-\alpha+\beta)}\right], p_{DD\bar{m}} = 1+\frac{1}{\lambda}\left[\frac{\bar{k}(\bar{m}-\alpha)-\beta(2+\bar{m}+\alpha)}{\beta(1+\alpha)+(1+\bar{m})(\bar{m}-\alpha+\beta)}\right]$$

Conditions for existence:

1. $\dfrac{\bar{k}(\bar{m}-\alpha)+\beta(2+\alpha+\bar{m})}{(1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha)} < \lambda < 1$
2. $\beta(1+\alpha) > \bar{k}(\bar{m}-\alpha+\beta)$
3. $\underline{k} < \dfrac{(1+\alpha)\left(\beta(\bar{m}-\alpha)+\bar{k}(\bar{m}-\alpha+2\beta)\right)}{(1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha)}$

4.6.1.2. $\Pi_{\underline{m}}\left(\bar{m}\right)-\Pi_{\underline{m}}\left(\underline{m}\right)=0$

$$\Pi_{\underline{m}}\left(\bar{m}\right)-\Pi_{\underline{m}}\left(\underline{m}\right)=\lambda\left[\left(p_{CD\bar{m}}-p_{DC\underline{m}}\right)\left(1+\alpha\right)\right]-\underline{k}=0 \qquad (0)$$

$$\Pi_{\bar{m}}\left(CD,\bar{m}\right)-\Pi_{\bar{m}}\left(DD,\bar{m}\right)=\lambda\left[\left(p_{CD\bar{m}}\right)\left(\bar{m}-\alpha\right)+\left(p_{DD\bar{m}}\right)\left(-\beta\right)\right]+\left(1-\lambda\right)\left[p_{\bar{m}}\left(-\beta\right)\right]=0 \qquad (ii)$$

$$\Pi_{\bar{m}}\left(DC,\underline{m}\right)-\Pi_{\bar{m}}\left(DD,\underline{m}\right)=\lambda\left[\left(p_{DC\underline{m}}\right)\left(\bar{m}-\alpha\right)\right]+\left(1-\lambda\right)\left[p_{\underline{m}}\left(-\beta\right)\right]>0 \qquad (iii)$$

$$\Pi_{\bar{m}}\left(DD,\bar{m}\right)-\Pi_{\bar{m}}\left(DC,\underline{m}\right)=\lambda\left[\left(p_{CD\bar{m}}\right)\left(1+\alpha\right)-p_{DC\underline{m}}\left(1+\bar{m}\right)\right]-\bar{k}+\beta\left(1-\lambda\right)p_{\underline{m}}=0 \qquad (v)$$

EQ:

$$p_{DC\underline{m}}=\frac{\beta\left(1+\alpha\right)+\left(1-\bar{m}-\beta+2\alpha\right)\underline{k}-\left(1+\alpha\right)\bar{k}}{2\left(1+\alpha\right)\left(\bar{m}-\alpha+\beta\right)},\ p_{CD\bar{m}}=\frac{\beta\left(1+\alpha\right)+\left(1+\bar{m}+\beta\right)\underline{k}-\left(1+\alpha\right)\bar{k}}{2\left(1+\alpha\right)\left(\bar{m}-\alpha+\beta\right)},\ p_{DD\bar{m}}=\left[\frac{-\beta+\lambda\left(\bar{m}-\alpha+\beta\right)+\bar{k}-\underline{k}}{\lambda\left(\bar{m}-\alpha+\beta\right)}\right],$$

$$p_{\underline{m}}=-\frac{\left(1+\alpha\right)\left(-\beta\left(\bar{m}-\alpha\right)+\left(\alpha-2\beta-\bar{m}\right)\bar{k}\right)+\left(\left(1+\bar{m}\right)\left(\bar{m}-\alpha+\beta\right)+\beta\left(1+\alpha\right)\right)\underline{k}}{2\left(1+\alpha\right)\left(1-\lambda\right)\left(\bar{m}-\alpha+\beta\right)},\ p_{\bar{m}}=1-p_{\underline{m}}$$

Note, that for the equilibrium values (iii) is satisfied:

Finally all conditions of the type

$p\in\left[0,1\right],\sum p<1$ and $0<\lambda<1$ reduce to:

$$-\frac{\left(1+\alpha\right)\left(-\beta\left(\bar{m}-\alpha\right)+\left(\alpha-2\beta-\bar{m}\right)\bar{k}\right)+\left(\left(1+\bar{m}\right)\left(\bar{m}-\alpha+\beta\right)+\beta\left(1+\alpha\right)\right)\underline{k}}{2\left(1+\alpha\right)\left(1-\lambda\right)\left(\bar{m}-\alpha+\beta\right)}>0$$

$p_{\underline{m}} > 0:$

$$-\frac{(1+\alpha)\left(-\beta(\bar{m}-\alpha)+(\alpha-2\beta-\bar{m})\bar{k}\right)+\left((1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha)\right)\underline{k}}{2(1+\alpha)(1-\lambda)(\bar{m}-\alpha+\beta)} > 0 \Leftrightarrow$$

$$(1+\alpha)\left(-\beta(\bar{m}-\alpha)+(\alpha-2\beta-\bar{m})\bar{k}\right)+\left((1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha)\right)\underline{k} < 0 \Leftrightarrow$$

$$\left((1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha)\right)\underline{k}-(1+\alpha)(\bar{m}-\alpha+2\beta)\bar{k} < (1+\alpha)\beta(\bar{m}-\alpha) \Leftrightarrow$$

$$\left((1+\bar{m})\left(1+\frac{\beta}{\bar{m}-\alpha}\right)+(1+\alpha)\frac{\beta}{\bar{m}-\alpha}\right)\underline{k}-(1+\alpha)\left(1+2\frac{\beta}{\bar{m}-\alpha}\right)\bar{k} < (1+\alpha)\beta \Leftrightarrow$$

$$\left((1+\bar{m})+(1+\bar{m})\frac{\beta}{\bar{m}-\alpha}+(1+\alpha)\frac{\beta}{\bar{m}-\alpha}\right)\underline{k}-\left((1+\alpha)+2(1+\alpha)\frac{\beta}{\bar{m}-\alpha}\right)\bar{k} < (1+\alpha)\beta$$

$p_{\underline{m}} < 1:$

$$(1+\alpha)\left(-\beta(\bar{m}-\alpha)+(\alpha-2\beta-\bar{m})\bar{k}\right)+\left((1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha)\right)\underline{k} > 2(1+\alpha)(1-\lambda)(\bar{m}-\alpha+\beta) \Leftrightarrow$$

$$\left((1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha)\right)\underline{k}-(1+\alpha)(\bar{m}-\alpha+2\beta)\bar{k} > (1+\alpha)\beta(\bar{m}-\alpha)+2(1+\alpha)(1-\lambda)(\bar{m}-\alpha+\beta) \Leftrightarrow$$

4.6.1.3. $\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m}) < 0$ (*)

$$\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m}) = \lambda\left[(p_{CD\bar{m}})(1+\alpha)\right]-\underline{k} < 0 \;\Rightarrow p_{\underline{m}} = 1$$

$$\Pi_{\bar{m}}(CD,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[(p_{CD\bar{m}})(\bar{m}-\alpha)+(p_{DD\bar{m}})(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right] = 0 \qquad \text{(ii)}$$

$$\Pi_{\bar{m}}(DC,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[(p_{DC\underline{m}})(\bar{m}-\alpha)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right] > 0 \qquad \text{(iii)}$$

$$\Pi_{\bar{m}}(DD,\bar{m})-\Pi_{\bar{m}}(DC,\underline{m}) = \lambda\left[(p_{CD\bar{m}})(1+\alpha)-p_{DC\underline{m}}(1+\bar{m})\right]-\bar{k}+\beta(1-\lambda)p_{\underline{m}} = 0 \qquad \text{(v)}$$

$$\Pi_{\bar{m}}(CD,\bar{m}) - \Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[(p_{CD\bar{m}})(\bar{m}-\alpha) + (p_{DD\bar{m}})(-\beta)\right] = 0 \quad \text{(ii)}$$

$$\Pi_{\bar{m}}(DC,\underline{m}) - \Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[(p_{DC\underline{m}})(\bar{m}-\alpha)\right] - \beta(1-\lambda) > 0 \quad \text{(iii)}$$

$$\Pi_{\bar{m}}(DD,\bar{m}) - \Pi_{\bar{m}}(DC,\underline{m}) = \lambda\left[(p_{CD\bar{m}})(1+\alpha) - p_{DC\underline{m}}(1+\bar{m})\right] - \bar{k} + \beta(1-\lambda) = 0 \quad \text{(v)}$$

(ii) and (v):

$$\lambda\left[(p_{CD\bar{m}})(\bar{m}-\alpha+\beta) + p_{DC\underline{m}}\beta - \beta\right] = 0$$
$$\lambda\left[(p_{CD\bar{m}})(1+\alpha) - p_{DC\underline{m}}(1+\bar{m})\right] - \bar{k} + \beta(1-\lambda) = 0$$

$$p_{CD\bar{m}} = \frac{\beta - p_{DC\underline{m}}\beta}{\lambda(\bar{m}-\alpha+\beta)}$$

$$\lambda\left[(p_{CD\bar{m}})(1+\alpha) - p_{DC\underline{m}}(1+\bar{m})\right] - \bar{k} + \beta(1-\lambda) = \lambda\left[\frac{\beta - p_{DC\underline{m}}\beta}{\lambda(\bar{m}-\alpha+\beta)}(1+\alpha) - p_{DC\underline{m}}(1+\bar{m})\right] - \bar{k} + \beta(1-\lambda) = 0$$

$$\Rightarrow p_{DC\underline{m}} = \frac{1}{\frac{\beta(1+\alpha)}{(\bar{m}-\alpha+\beta)} + \lambda(1+\bar{m})}\left[\frac{\beta(1+\alpha)}{\lambda(\bar{m}-\alpha+\beta)} - \bar{k} + \beta(1-\lambda)\right]$$

EQ:

$$p_{CD\bar{m}} = \frac{\beta\left(\lambda(1+\bar{m}+\beta) + \bar{k} - \beta\right)}{\lambda\left((1+\bar{m})(\bar{m}-\alpha+\beta) + \beta(1+\alpha)\right)}, \; p_{DC\underline{m}} = \frac{\beta\left((1-\lambda)(\bar{m}-\alpha+\beta) + \lambda(1+\alpha)\right) - (\bar{m}-\alpha+\beta)\bar{k}}{\lambda\left((1+\bar{m})(\bar{m}-\alpha+\beta) + \beta(1+\alpha)\right)}$$

$$p_{DD\bar{m}} = \frac{(\bar{m}-\alpha)\left(\lambda(1+\bar{m}+\beta) + \bar{k} - \beta\right)}{\lambda\left((1+\bar{m})(\bar{m}-\alpha+\beta) + \beta(1+\alpha)\right)}$$

We need to check for (*) and (iii):

(*): $\frac{(1+\alpha)\left((\bar{m}-\alpha+2\beta)\bar{k}+\beta\left(2\lambda(\bar{m}-\alpha+\beta)-(\bar{m}-\alpha+2\beta)\right)\right)-\left((1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha)\right)\underline{k}}{(1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha)}<0 \Leftrightarrow$

$\left((1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha)\right)\underline{k}-(1+\alpha)(\bar{m}-\alpha+2\beta)\bar{k}>(1+\alpha)\beta(\bar{m}-\alpha)+2(1+\alpha)(1-\lambda)(\bar{m}-\alpha+\beta) \Leftrightarrow$

$\lambda<\frac{(1+\alpha)(\beta-\bar{k})(\bar{m}-\alpha+2\beta)+\left((1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha)\right)\underline{k}}{2(1+\alpha)\beta(\bar{m}-\alpha+\beta)}$

$=\frac{(1+\alpha)\beta(\bar{m}-\alpha+2\beta)+\left((1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha)\right)\underline{k}-\left((1+\alpha)(\bar{m}-\alpha+\beta)+\beta(1+\alpha)\right)\bar{k}}{2(1+\alpha)\beta(\bar{m}-\alpha+\beta)}$

Finally all conditions of the type

$p\in[0,1], \sum p<1$ and $0<\lambda<1$ and (iii) reduce to:

1. $0<\frac{(1+\alpha)\beta(\bar{m}-\alpha+2\beta)+(\bar{m}-\alpha)(\bar{m}-\alpha+\beta)\bar{k}}{2\beta(1+\alpha)(\bar{m}-\alpha+\beta)}<\lambda<1$
2. $\beta(1+\alpha)>(\bar{m}-\alpha+\beta)\bar{k}$

adding the upper bound due to (*) we end up with:

1. $\frac{(1+\alpha)\beta(\bar{m}-\alpha+2\beta)+(\bar{m}-\alpha)(\bar{m}-\alpha+\beta)\bar{k}}{2\beta(1+\alpha)(\bar{m}-\alpha+\beta)}<\lambda<\frac{(1+\alpha)\beta(\bar{m}-\alpha+2\beta)+((1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha))\underline{k}-((1+\alpha)(\bar{m}-\alpha+\beta)+\beta(1+\alpha))\bar{k}}{2\beta(1+\alpha)(\bar{m}-\alpha+\beta)}$

2. $\beta(1+\alpha)>(\bar{m}-\alpha+\beta)\bar{k}$

note that LHS<RHS due to $\bar{k}<\underline{k}$

RHS <1: $\frac{(1+\alpha)\beta(\bar{m}-\alpha+2\beta)+((1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha))\underline{k}-((1+\alpha)(\bar{m}-\alpha+\beta)+\beta(1+\alpha))\bar{k}}{2\beta(1+\alpha)(\bar{m}-\alpha+\beta)}<1\Leftrightarrow$

$((1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha))\underline{k}-((1+\alpha)(\bar{m}-\alpha+\beta)+\beta(1+\alpha))\bar{k}<\beta(1+\alpha)(\bar{m}-\alpha)\Leftrightarrow$

$((1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha))\underline{k}-((1+\alpha)(\bar{m}-\alpha+\beta)+\beta(1+\alpha))\bar{k}<\beta(1+\alpha)(\bar{m}-\alpha)\Leftrightarrow$

$$\frac{((1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha))}{(\bar{m}-\alpha)}\underline{k}-\frac{((1+\alpha)(\bar{m}-\alpha+\beta)+\beta(1+\alpha))}{(\bar{m}-\alpha)}\bar{k}<\beta(1+\alpha)\Leftrightarrow$$

note:

$$\frac{((1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha))}{(\bar{m}-\alpha)}\underline{k}-\frac{((1+\alpha)(\bar{m}-\alpha+\beta)+\beta(1+\alpha))}{(\bar{m}-\alpha)}\bar{k}>(\bar{m}-\alpha+\beta)\bar{k}\Leftrightarrow$$

$$\frac{((1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha))}{(\bar{m}-\alpha)}\underline{k}-\frac{((1+\alpha)(\bar{m}-\alpha+\beta)+\beta(1+\alpha))+(\bar{m}-\alpha)(\bar{m}-\alpha+\beta)}{(\bar{m}-\alpha)}\bar{k}>0\Leftrightarrow$$

$\underline{k}-\bar{k}>0$, i.e. $\frac{((1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha))}{(\bar{m}-\alpha)}\underline{k}-\frac{((1+\alpha)(\bar{m}-\alpha+\beta)+\beta(1+\alpha))}{(\bar{m}-\alpha)}\bar{k}<\beta(1+\alpha)$ is binding.

To summarize:

$$p_{CD\bar{m}}=\frac{\beta(\lambda(1+\bar{m}+\beta)+\bar{k}-\beta)}{\lambda((1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha))},\ p_{DC\underline{m}}=\frac{\beta((1-\lambda)(\bar{m}-\alpha+\beta)+\lambda(1+\alpha))-(\bar{m}-\alpha+\beta)\bar{k}}{\lambda((1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha))}$$

$$p_{DD\bar{m}}=\frac{(\bar{m}-\alpha)(\lambda(1+\bar{m}+\beta)+\bar{k}-\beta)}{\lambda((1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha))}$$

Conditions for existence:

1. $\frac{(1+\alpha)\beta(\bar{m}-\alpha+2\beta)+(\bar{m}-\alpha)(\bar{m}-\alpha+\beta)\bar{k}}{2\beta(1+\alpha)(\bar{m}-\alpha+\beta)}<\lambda<\frac{(1+\alpha)\beta(\bar{m}-\alpha+2\beta)+((1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha))\underline{k}-((1+\alpha)(\bar{m}-\alpha+\beta)+\beta(1+\alpha))\bar{k}}{2\beta(1+\alpha)(\bar{m}-\alpha+\beta)}$

2. $\beta(1+\alpha)>\frac{((1+\bar{m})(\bar{m}-\alpha+\beta)+\beta(1+\alpha))}{(\bar{m}-\alpha)}\underline{k}-\frac{((1+\alpha)(\bar{m}-\alpha+\beta)+\beta(1+\alpha))}{(\bar{m}-\alpha)}\bar{k}$

4.6.2. and $DD,\underline{m} \;\Rightarrow p_{\underline{m}}=1$

$$\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left[\left(p_{CC\underline{m}}+p_{CD\underline{m}}\right)(\bar{m}-\alpha)+\left(p_{DC\underline{m}}+p_{DD\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]<0 \qquad \text{(i)}$$

$$\Pi_{\bar{m}}(CD,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{DC\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]=0 \qquad \text{(ii)}$$

$$\Pi_{\bar{m}}(DC,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[\left(p_{CC\underline{m}}+p_{DC\underline{m}}\right)(\bar{m}-\alpha)+\left(p_{CD\underline{m}}+p_{DD\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]<0 \qquad \text{(iii)}$$

$$\Pi_{\bar{m}}(CD,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{CD\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]<0 \qquad \text{(iv)}$$

$$\Pi_{\bar{m}}(DD,\bar{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda(1+\alpha)\left[\left(p_{CD\bar{m}}+p_{CD\underline{m}}-p_{DC\bar{m}}-p_{DC\underline{m}}\right)\right]-\bar{k}=0 \qquad \text{(v)}$$

$$\Pi_{\bar{m}}(DC,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left[\left(p_{DD\underline{m}}\right)(-\beta)\right]-\beta(1-\lambda)<0 \qquad \text{(i)}$$

$$\Pi_{\bar{m}}(CD,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left[\left(p_{CD\bar{m}}\right)(\bar{m}-\alpha)+\left(p_{DD\bar{m}}\right)(-\beta)\right]=0 \qquad \text{(ii)}$$

$$\Pi_{\bar{m}}(DC,\underline{m})-\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[\left(p_{DD\underline{m}}\right)(-\beta)\right]-\beta(1-\lambda)<0 \qquad \text{(iii)}$$

$$\Pi_{\bar{m}}\left(CD,\underline{m}\right)-\Pi_{\bar{m}}\left(DD,\underline{m}\right)=\lambda\left[\left(p_{CD\bar{m}}+p_{DD\bar{m}}\right)\left(-\beta\right)\right]<0 \qquad \text{(iv)}$$

$$\Pi_{\bar{m}}\left(DD,\bar{m}\right)-\Pi_{\bar{m}}\left(DD,\underline{m}\right)=\lambda\left(1+\alpha\right)\left[\left(p_{CD\bar{m}}\right)\right]-\bar{k}=0 \qquad \text{(v)}$$

EQ: $p_{CD\bar{m}}=\dfrac{\underline{k}}{\lambda\left(1+\alpha\right)}, p_{DD\bar{m}}=\dfrac{\left(\bar{m}-\alpha\right)}{\beta}\dfrac{\underline{k}}{\lambda\left(1+\alpha\right)}, p_{DD\underline{m}}=1-\dfrac{\left(\bar{m}-\alpha+\beta\right)}{\beta}\dfrac{\underline{k}}{\lambda\left(1+\alpha\right)}$

(i), (iii) and (iv) are satisfied.

Finally all conditions of the type

$p\in\left[0,1\right],\sum p<1$ and $0<\lambda<1$ reduce to:

1. $0<\dfrac{\bar{k}\left(\bar{m}-\alpha+\beta\right)}{\left(1+\alpha\right)\beta}<\lambda<1$
2. $\beta\left(1+\alpha\right)>\left(\bar{m}-\alpha+\beta\right)\bar{k}$

To summarize:

$$p_{CD\bar{m}}=\frac{\underline{k}}{\lambda\left(1+\alpha\right)}, p_{DD\bar{m}}=\frac{\left(\bar{m}-\alpha\right)}{\beta}\frac{\underline{k}}{\lambda\left(1+\alpha\right)}, p_{DD\underline{m}}=1-\frac{\left(\bar{m}-\alpha+\beta\right)}{\beta}\frac{\underline{k}}{\lambda\left(1+\alpha\right)}$$

Conditions for existence:

1. $\dfrac{\bar{k}\left(\bar{m}-\alpha+\beta\right)}{\left(1+\alpha\right)\beta}<\lambda<1$
2. $\beta\left(1+\alpha\right)>\left(\bar{m}-\alpha+\beta\right)\bar{k}$

Finally we consider semi-pooling equilibria with only low types pooling, i.e. $\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})=\lambda\left[\left(p_{CD\bar{m}}+p_{CD\underline{m}}-p_{DC\bar{m}}-p_{DC\underline{m}}\right)(1+\alpha)\right]-\underline{k}=0$. For this equality to hold we necessarily need $p_{CD\bar{m}}>0$ or $p_{CD\underline{m}}>0$ but not both since this would correspond to a pooling among high types.

4.7. $\Pi_{\underline{m}}(\bar{m})-\Pi_{\underline{m}}(\underline{m})=\lambda\left[\left(p_{CD\bar{m}}+p_{CD\underline{m}}-p_{DC\bar{m}}-p_{DC\underline{m}}\right)(1+\alpha)\right]-\underline{k}=0$

4.7.1. $p_{CD\bar{m}}>0$, i.e. $p_{CC\underline{m}}=p_{CD\underline{m}}=p_{DC\underline{m}}=p_{DD\underline{m}}=0$

$p_{\underline{m}}\in(0,1)$:

$\Pi_{\bar{m}}(CC,\bar{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)(1+\bar{m})+\left(p_{DC\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)+p_{\bar{m}}(-\beta)\right]-\bar{k}<$

$\Pi_{\bar{m}}(CD,\bar{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)(1+\bar{m})+\left(p_{DC\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]-\bar{k}$

$\Pi_{\bar{m}}(DC,\bar{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)(1+\alpha)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]-\bar{k}<$

$\Pi_{\bar{m}}(DD,\bar{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)(1+\alpha)\right]-\bar{k}$

$\Pi_{\bar{m}}(CC,\underline{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)(1+\bar{m})+\left(p_{CD\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)+p_{\bar{m}}(-\beta)\right]<$

$\Pi_{\bar{m}}(CD,\underline{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)(1+\bar{m})+\left(p_{CD\bar{m}}+p_{DD\bar{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]$

$\Pi_{\bar{m}}(DC,\underline{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)(1+\alpha)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]<$

$\Pi_{\bar{m}}(DD,\underline{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)(1+\alpha)\right]$

$\Pi_{\underline{m}}(\underline{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{DC\bar{m}}\right)(1+\alpha)\right]$

$\Pi_{\underline{m}}(\bar{m})=\lambda\left[\left(p_{CC\bar{m}}+p_{CD\bar{m}}\right)(1+\alpha)\right]-\underline{k}$

$$\Rightarrow$$
$$\Pi_{\bar{m}}(CD,\bar{m}) = \lambda\left[(p_{CD\bar{m}})(1+\bar{m})+(p_{DD\bar{m}})(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]-\bar{k}$$
$$\Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[(p_{CD\bar{m}})(1+\alpha)\right]-\bar{k}$$

$$\Pi_{\bar{m}}(CD,\underline{m}) = \lambda\left[(p_{CD\bar{m}}+p_{DD\bar{m}})(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]<0$$
$$\Pi_{\bar{m}}(DD,\underline{m}) = 0$$

$$\Pi_{\underline{m}}(\underline{m}) = 0 = \lambda\left[(p_{CD\bar{m}})(1+\alpha)\right]-\underline{k} = \Pi_{\underline{m}}(\bar{m}) \Leftrightarrow p_{CD\bar{m}} = \frac{\underline{k}}{\lambda(1+\alpha)} \Rightarrow DD,\underline{m} \prec DD,\bar{m}$$

$$\Pi_{\bar{m}}(CD,\bar{m}) = \lambda\left[(p_{CD\bar{m}})(1+\bar{m})+(p_{DD\bar{m}})(-\beta)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]-\bar{k}$$
$$\Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[(p_{CD\bar{m}})(1+\alpha)\right]-\bar{k}$$

$$\Pi_{\underline{m}}(\underline{m}) = 0$$
$$\Pi_{\underline{m}}(\bar{m}) = \lambda\left[(p_{CD\bar{m}})(1+\alpha)\right]-\underline{k}$$

4.7.1.1. $\Pi_{\bar{m}}(CD,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})>0 \Rightarrow p_{CD\bar{m}}=1$

$$\Pi_{\bar{m}}(CD,\bar{m})-\Pi_{\bar{m}}(DD,\bar{m})>0 \Leftrightarrow$$
$$\lambda(\bar{m}-\alpha+\beta) > \lambda\beta+(1-\lambda)\,p_{\bar{m}}\beta \Leftrightarrow p_{\bar{m}} < \frac{\lambda(\bar{m}-\alpha)}{(1-\lambda)\beta}$$
$$\Pi_{\underline{m}}(\underline{m}) = 0 = \lambda(1+\alpha)-\underline{k} = \Pi_{\underline{m}}(\bar{m}) \Leftrightarrow \lambda = \frac{\underline{k}}{(1+\alpha)}$$

To summarize:

$p_{CD\bar{m}} = 1$

Conditions for existence:

1. $\lambda = \frac{\underline{k}}{(1+\alpha)}$
2. $\underline{k} < (1+\alpha)$
3. $p_{\bar{m}} < \frac{\lambda(\bar{m}-\alpha)}{(1-\lambda)\beta}$

note that 3. is only binding if:

$$\frac{\lambda(\bar{m}-\alpha)}{(1-\lambda)\beta} < 1 \Leftrightarrow \lambda < \frac{\beta}{(\bar{m}-\alpha+\beta)} \Leftrightarrow \frac{\underline{k}}{(1+\alpha)} < \frac{\beta}{(\bar{m}-\alpha+\beta)}$$

4.7.1.2. $\Pi_{\bar{m}}(CD,\bar{m}) - \Pi_{\bar{m}}(DD,\bar{m}) = 0$

$$\Pi_{\bar{m}}(CD,\bar{m}) = \Pi_{\bar{m}}(DD,\bar{m}) \Leftrightarrow$$

$$\lambda\left[(p_{CD\bar{m}})(\bar{m}-\alpha+\beta)\right] = \lambda\beta + (1-\lambda)p_{\bar{m}}\beta \Leftrightarrow p_{CD\bar{m}} = \frac{\lambda\beta+(1-\lambda)p_{\bar{m}}\beta}{\lambda(\bar{m}-\alpha+\beta)}$$

$$\Pi_{\underline{m}}(\underline{m}) = 0 = \lambda\left[(p_{CD\bar{m}})(1+\alpha)\right] - \underline{k} = \Pi_{\underline{m}}(\bar{m}) \Leftrightarrow p_{CD\bar{m}} = \frac{\underline{k}}{\lambda(1+\alpha)}$$

$\Rightarrow$

$$\frac{\lambda\beta+(1-\lambda)p_{\bar{m}}\beta}{\lambda(\bar{m}-\alpha+\beta)} = \frac{\underline{k}}{\lambda(1+\alpha)} \Leftrightarrow p_{\bar{m}} = \frac{(\bar{m}-\alpha+\beta)\underline{k}-\lambda(1+\alpha)\beta}{(1-\lambda)(1+\alpha)\beta}$$

EQ: $p_{CD\bar{m}} = \frac{\underline{k}}{\lambda(1+\alpha)}, p_{DD\bar{m}} = 1-\frac{\underline{k}}{\lambda(1+\alpha)}, p_{\bar{m}} = \frac{(\bar{m}-\alpha+\beta)\underline{k}-\lambda(1+\alpha)\beta}{(1-\lambda)(1+\alpha)\beta}, p_{\underline{m}} = \frac{(1+\alpha)\beta-(\bar{m}-\alpha+\beta)\underline{k}}{(1-\lambda)(1+\alpha)\beta}$

Conditions for existence:

1. $\frac{\underline{k}}{(1+\alpha)} < \lambda < \frac{(\bar{m}-\alpha+\beta)}{\beta}\frac{\underline{k}}{(1+\alpha)}$
2. $(1+\alpha)\beta > (\bar{m}-\alpha+\beta)\underline{k}$
3. $\underline{k} < (1+\alpha)$

4.7.2. $p_{CD\underline{m}} > 0$, i.e. $p_{CC\bar{m}} = p_{CD\bar{m}} = p_{DC\bar{m}} = p_{DD\bar{m}} = 0$

$p_{\underline{m}} \in (0,1)$:

$$\Pi_{\bar{m}}(CC,\bar{m}) = \lambda\left[\left(p_{CC\underline{m}}+p_{CD\underline{m}}\right)(1+\bar{m})+\left(p_{DC\underline{m}}+p_{DD\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)+p_{\bar{m}}(-\beta)\right]-\bar{k} <$$
$$\Pi_{\bar{m}}(DC,\bar{m}) = \lambda\left[\left(p_{CC\underline{m}}+p_{CD\underline{m}}\right)(1+\bar{m})+\left(p_{DC\underline{m}}+p_{DD\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]-\bar{k}$$
$$\Pi_{\bar{m}}(CD,\bar{m}) = \lambda\left[\left(p_{CC\underline{m}}+p_{CD\underline{m}}\right)(1+\alpha)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right]-\bar{k} <$$
$$\Pi_{\bar{m}}(DD,\bar{m}) = \lambda\left[\left(p_{CC\underline{m}}+p_{CD\underline{m}}\right)(1+\alpha)\right]-\bar{k}$$

$$\Pi_{\bar{m}}(CC,\underline{m}) = \lambda\left[\left(p_{CC\underline{m}}+p_{DC\underline{m}}\right)(1+\bar{m})+\left(p_{CD\underline{m}}+p_{DD\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)+p_{\bar{m}}(-\beta)\right] <$$
$$\Pi_{\bar{m}}(DC,\underline{m}) = \lambda\left[\left(p_{CC\underline{m}}+p_{DC\underline{m}}\right)(1+\bar{m})+\left(p_{CD\underline{m}}+p_{DD\underline{m}}\right)(-\beta)\right]+(1-\lambda)\left[p_{\underline{m}}(-\beta)\right]$$
$$\Pi_{\bar{m}}(CD,\underline{m}) = \lambda\left[\left(p_{CC\underline{m}}+p_{DC\underline{m}}\right)(1+\alpha)\right]+(1-\lambda)\left[p_{\bar{m}}(-\beta)\right] <$$
$$\Pi_{\bar{m}}(DD,\underline{m}) = \lambda\left[\left(p_{CC\underline{m}}+p_{DC\underline{m}}\right)(1+\alpha)\right]$$

$$\Pi_{\underline{m}}(\underline{m}) = \lambda\left[\left(p_{CC\underline{m}}+p_{DC\underline{m}}\right)(1+\alpha)\right]$$
$$\Pi_{\underline{m}}(\bar{m}) = \lambda\left[\left(p_{CC\underline{m}}+p_{CD\underline{m}}\right)(1+\alpha)\right]-\underline{k}$$

Since $CD,\underline{m}$ is strictly dominated by unconditional defection such an equilibrium cannot exist.

C. Appendix to Chapter 4

Proof of Lemma 4-1: The argument is given in the paper. QED

Proof of Lemma 4-2: By Lemma 4-1 $\gamma(A) \in \Gamma_{sym}$ constitutes a dilemma if and only if it is of the Prisoners' dilemma type. W.l.o.g. let $(1,1)$ be the unique equilibrium. Since unilateral deviation reduces material payoffs and increases inequality $(1,1)$ may not be contested by any player. However $(0,0)$ may be stabilized by two sufficiently inequality-averse players. This is the case if and only if for player one $u^1_{(1,0)} = a^1_{(1,0)} - \theta^1 \left| a^1_{(1,0)} - a^2_{(1,0)} \right| \leq u^1_{(0,0)} = a^1_{(0,0)} \Leftrightarrow \theta^1 \geq \frac{a^1_{(1,0)} - a^1_{(0,0)}}{\left| a^1_{(1,0)} - a^2_{(1,0)} \right|}$ and accordingly for player two. QED

Proof of Proposition 4-1: A symmetric dilemma can be represented by the following matrix $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ showing the payoffs for the column player. I will refer to the strategies as '0' and '1' respectively. Without loss of generality I assume $b > a,\ d > c$, i.e. '1' is the dominant strategy. According to Lemma 4-1 $a > d$ must hold. This implies the following ordering of parameters $b > a > d > c$. Define $\theta^D = \frac{b-a}{b-c} \in (0,1)$. In terms of utility two individuals with inequality aversion θ_1 and θ_2 respectively give rise to the following bimatrix:

	0	1
0	a ; a	$c - \theta_1(b-c)$; $b - \theta_2(b-c)$
1	$b - \theta_2(b-c)$; $c - \theta_1(b-c)$	d ; d

Table C-1: Payoffs in the dilemma $\gamma(U^1, U^2)$.

In the following I will distinguish two cases. The first case corresponds to a match of two players with a degree of inequality aversion above the threshold θ^D. In the second case for at least one player this condition is violated.

(i) $\theta_1, \theta_2 \geq \theta^D$

$(0,0),(1,1)$ are the two pure Nash equilibria over which individuals randomize with equal weight and both gain a material payoff of: $\left(\frac{a+d}{2}\right)$

(ii) $\theta_1 \vee \theta_2 < \theta^D$

'1' remains for at least one agent the dominant strategy. Hence, both individuals will earn: (d)

Note that all individuals with $\theta \geq \theta^D$ earn the same expected payoff $\Pi^{\theta \geq \theta^D} = F(\theta^D)d + (1 - F(\theta^D))\frac{a+d}{2}$, whereas individuals with $\theta < \theta^D$ earn

$\Pi^{\theta<\theta^D} = F(\theta^D)d + (1-F(\theta^D))d = d$. Hence as long as there are some individuals with a degree of inequality aversion above θ^D those players face an evolutionary advantage because $\Pi^{\theta\geq\theta^D} - \Pi^{\theta<\theta^D} = (1-F(\theta^D))\frac{a-d}{2} > 0$. Hence the globally stable distribution of inequality aversion is characterized by $F^\infty(\theta^D) = 0$. The advantage increases with the share of sufficiently inequality-averse players, i.e. $\frac{\partial(\Pi^{\theta\geq\theta^D} - \Pi^{\theta<\theta^D})}{\partial(1-F(\theta^D))} = \frac{a-d}{2} > 0$. QED

Proof of Lemma 4-3: The argument is given in the paper. QED

Proof of Proposition 4-2: A symmetric problem of coordination can be represented by $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$. For a game with the Nash equilibria on the diagonal $a > b,\ d > c$ holds. Hence any degree of inequality aversion leaves the set of pure Nash equilibria unchanged. Hence any match of two players will generate the same payoff, the average of the two pure Nash equilibria. Therefore the distribution of preferences will be determined by initial conditions and random shifts. Hence I shall assume for the Nash equilibria to lie on the off-diagonal, i.e. w.l.o.g. $b > a,\ d < c$. In terms of utility two individuals with inequality aversion θ_1 and θ_2 respectively give rise to bimatrix as depicted in Table C-1.

Define $\theta^C_{(0,1),2} = \theta^C_{(1,0),1} \equiv \frac{b-a}{|b-c|} > 0,\ \theta^C_{(0,1),1} = \theta^C_{(1,0),2} \equiv \frac{c-d}{|b-c|} > 0$. These thresholds represent the ratio of the material incentive to stick to the considered (material) equilibrium and the gain in non-material terms from deviation stemming from an increasing equality. A threshold above one represents a situation where the maximum gain in equality is smaller than the material loss from deviating from (material) equilibrium behavior. In other words no level of inequality aversion can destabilize this equilibrium. If for a player $\theta \geq \theta^C_{(0,1),2}\left(\theta^C_{(1,0),2}\right)$ then for this player the equilibrium $(0,1)((1,0))$ is contestable. In the following (1)-(3) consider the different possible matches according to the relation of the thresholds and the involved players' inequality aversion.

(1) $\theta_1 > \theta^C_{(1,0),1}, \theta_2 > \theta^C_{(0,1),2}$

, i.e. both equilibria are contestable (by different players) and are indeed destabilized. The strategy-tuple $(0,0)$ is stabilized. Now two cases can be distinguished. First '0' has become the dominant strategy for at least one player (a and c) or $(1,1)$ is also stabilized (b).

a) $\theta_1 < \theta^C_{(0,1),1}, \theta_2 < \theta^C_{(1,0),2} \vee \theta^C_{(0,1),1}, \theta^C_{(1,0),2} > 1$

, i.e. $(1,1)$ is not stabilized either because inequality aversion is too weak or the equilibria are not contestable by the considered players. In that case '0' becomes the dominant strategy and the unique Nash equilibrium is given by $(0,0)$. (a,a)

b) $\theta_1 > \theta^C_{(0,1),1}, \theta_2 > \theta^C_{(1,0),2} \wedge \theta^C_{(0,1),1}, \theta^C_{(1,0),2} < 1$

, i.e. $(1,1)$ also becomes an equilibrium. There are now the two pure Nash equilibria $(1,1)$ and $(0,0)$.

$\left(\frac{a+d}{2},\frac{a+d}{2}\right)$

c) $\theta_1 < \theta^C_{(0,1),1} \vee \theta^C_{(0,1),1} > 1, \theta_2 > \theta^C_{(1,0),2} \wedge \theta^C_{(1,0),2} < 1$

, i.e. '0' is the dominant strategy for player one and '0' is the best response for player two. (a,a)

(2) $\theta_1 < \theta^C_{(1,0),1}, \theta_2 < \theta^C_{(0,1),2} \vee \theta^C_{(0,1),2}, \theta^C_{(1,0),1} > 1$

, i.e. $(0,0)$ is not stabilized either because inequality aversion is too weak or the equilibria are not contestable by the considered players.

a) $\theta_1 < \theta^C_{(0,1),1}, \theta_2 < \theta^C_{(1,0),2} \vee \theta^C_{(0,1),1}, \theta^C_{(1,0),2} > 1$

, i.e. $(1,1)$ is not stabilized either because inequality aversion is too weak or the equilibria are not contestable by the considered players. The sets of Nash equilibria of $\gamma(A)$ and $\gamma(U^1,U^2)$ coincide. $\left(\frac{b+c}{2},\frac{b+c}{2}\right)$

b) $\theta_1 > \theta^C_{(0,1),1}, \theta_2 > \theta^C_{(1,0),2} \wedge \theta^C_{(0,1),1}, \theta^C_{(1,0),2} < 1$

, i.e. both material equilibria are contestable and are indeed destabilized. In that case '1' becomes the dominant strategy. (d,d)

(3) w.l.o.g. $\theta_1 < \theta^C_{(1,0),1}$ $\theta_2 > \theta^C_{(0,1),2} \vee \theta^C_{(0,1),2}, \theta^C_{(1,0),1} < 1$ (player 1 is selfish, player 2 is inequality-averse), i.e. one players' inequality aversion makes one equilibrium contestable.

a) $\theta_2 < \theta_{(1,0),2} \vee \theta_{(1,0),2} > 1$

, i.e. this player inequality aversion is either too weak or the remaining equilibrium is not contestable by this player. In that case '0' is the dominant strategy of this player. Two cases can be distinguished for the remaining player.

(i) $\theta_1 < \theta_{(0,1),1} \vee \theta_{(0,1),1} > 1$

, i.e. this opponents' inequality aversion is either too weak to or the remaining equilibrium is not contestable from this perspective. (b,c)

(ii) $\theta_1 > \theta_{(0,1),1} < 1$

, i.e. the remaining equilibrium is also contestable and indeed destabilized. (a,a)

b) $\theta_2 > \theta_{(1,0),2} < 1$

, i.e. this player makes both equilibria contestable and indeed both equilibria are destabilized.

(i) $\theta_1 > \theta_{(0,1),1} < 1$

, i.e. '1' becomes the dominant strategy of this player (d,d)

(ii) $\theta_1 < \theta_{(0,1),1} \vee \theta_{(0,1),1} > 1$

There is a unique mixed equilibrium which is played $\left(\Pi^{mix},\Pi^{mix}\right)$

Note that strictness excludes the cases 1b), 1c), 3b). Table C-2 depicts equilibrium payoffs in the various matches for the case of $\theta^C_{(0,1),2} = \theta^C_{(1,0),1} \equiv \theta^C < 1$, i.e. the case where both equilibria are contestable by different players.

	$\theta_2 < \theta^C$	$\theta^C < \theta_2$
$\theta_1 < \theta^C$	$\left(\frac{b+c}{2}, \frac{b+c}{2}\right)$	(b,c)
$\theta^C < \theta_1$	(c,b)	(a,a)

Table C-2: Equilibrium payoffs according to the degree of inequality aversion of the matched players.

Note that individuals with $\theta \geq \theta^C$ earn the same expected payoff $\Pi^{\theta \geq \theta^C} = F(\theta^C)c + (1 - F(\theta^C))a$, whereas individuals with $\theta < \theta^C$ earn $\Pi^{\theta < \theta^C} = F(\theta^C)\frac{b+c}{2} + (1 - F(\theta^C))b$. Hence $\Pi^{\theta \geq \theta^C} - \Pi^{\theta < \theta^C} = F(\theta^C)\frac{c-b}{2} + (1 - F(\theta^C))(a-b)$. Note that $\left(\Pi^{\theta \geq \theta^C} - \Pi^{\theta < \theta^C}\right)\left(F(\theta^C) = 0\right) = a - b < 0$. If $b > c$, then $\left(\Pi^{\theta \geq \theta^C} - \Pi^{\theta < \theta^C}\right)\left(F(\theta^C) = 1\right) = \frac{c-b}{2} < 0$ and the globally stable equilibrium is characterized by $F(\theta^C) = 1$. Furthermore $\frac{\partial\left(\Pi^{\theta \geq \theta^C} - \Pi^{\theta < \theta^C}\right)}{\partial\left(1 - F(\theta^C)\right)} = a - \frac{b+c}{2}$.

If the reverse holds, i.e. $b < c$ then $\left(\Pi^{\theta \geq \theta^C} - \Pi^{\theta < \theta^C}\right)\left(F(\theta^C) = 1\right) = \frac{c-b}{2} > 0$ and as a consequence there exist a globally stable inner equilibria characterized by $F(\theta^C) = \frac{b-a}{\frac{b+c}{2} - a} = \theta^C \frac{c-b}{\frac{b+c}{2} - a}$. Furthermore $\frac{\partial\left(\Pi^{\theta \geq \theta^C} - \Pi^{\theta < \theta^C}\right)}{\partial\left(1 - F(\theta^C)\right)} = a - \frac{b+c}{2} < 0$. The case $\theta^C_{(0,1),1} = \theta^C_{(1,0),2} < 1$ is analyzed in the analog way $(a \leftrightarrow d, b \leftrightarrow c)$. With the definitions for $AP_{(i,i)}$ and d in the text the claim follows. QED

Proof of Lemma 4-4: The argument is given in the paper. QED

Proof of Proposition 4-3: Let me first consider the case with multiple equilibria which are not Pareto-ranked (case A) with

payoffs given by $A^1 = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $A^2 = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$. W.l.o.g. I will consider a game with Nash equilibria on the diagonal (relabeling the strategies for one player transforms such a game in a game with equilibria on the off-diagonal and vice versa), i.e. $A > B$, $D > C$ and $d > b$, $a > c$. W.l.o.g. let player two be the type who is favored by the problem of distribution, i.e. $A > a$ and $D > d$. The assumption that the two pure Nash equilibria are not Pareto-ranked leaves us with two possibilities, either $a < d < D < A$ or $d < a < A < D$. W.l.o.g. I will assume the first relations to hold. This implies that the equilibrium $(1,1)$ is characterized by a strictly lower degree of inequality. In terms of utility two individuals with inequality aversion θ_1 and θ_2 respectively give rise to the following bimatrix:

	0	1
0	$A-\theta_2\|A-a\|$ $a-\theta_1\|A-a\|$	$B-\theta_2\|B-b\|$ $b-\theta_1\|B-b\|$
1	$C-\theta_2\|C-c\|$ $c-\theta_1\|C-c\|$	$D-\theta_2\|D-d\|$ $d-\theta_1\|D-d\|$

Table C-3: Payoffs in the problem of distribution $\gamma\left(U^1,U^2\right)$.

Note that:

- (i) $D-\theta_2|D-d|\geq d>a>c$
- (ii) $D-\theta_2|D-d|>C-\theta_2|C-c|$, because for $\theta_2=0$ $D>C$ and for $\theta_2=1$ $d>c\geq C-|C-c|$
- (iii) $|A-a|>|D-d|$

Before I analyze the different types of matches, I will define the following thresholds: $\theta^R_{(0,0),2}\equiv\frac{A-B}{|A-a|-|B-b|},\theta^R_{(1,1),2}\equiv\frac{D-C}{|D-d|-|C-c|},\theta^R_{(0,0),1}\equiv\frac{a-c}{|A-a|-|C-c|},\theta^R_{(1,1),1}\equiv\frac{d-b}{|D-d|-|B-b|}$

Note that due to (ii) $\theta^R_{(1,1),2}>1$, i.e. the equilibrium $(1,1)$ is not contestable for player two. Let for all other thresholds $\theta^R_{(0,0),2},\theta^R_{(0,0),1},\theta^R_{(1,1),1}\in(0,1)$, i.e. both equilibria are contestable, $(1,1)$ only by player one, $(0,0)$ by both players.

1. $\theta_2>\theta^R_{(0,0),2}$ ('1' is the dominant strategy for player 2)

(1) $\theta_1<\theta^R_{(1,1),1}$ $\quad(d,D)$

(2) $\theta_1>\theta^R_{(1,1),1}$ $\quad(b,B)$

2. $\theta_2<\theta^R_{(0,0),2}$

a) $\theta_1<\theta^R_{(0,0),1}$

(1) $\theta_1<\theta^R_{(1,1),1}$: $(0,0),(1,1)$ remain both equilibria $\quad\left(\frac{a+d}{2},\frac{A+D}{2}\right)$

(2) $\theta_1>\theta^R_{(1,1),1}$: '0' is the dominant strategy for player 1 $\quad(a,A)$

b) $\theta_1>\theta^R_{(0,0),1}$

(1) $\theta_1<\theta^R_{(1,1),1}$: '1' is the dominant strategy for player 1 $\quad(d,D)$

(2) $\theta_1>\theta^R_{(1,1),1}$: there is a unique mixed equilibrium $\quad\left(\Pi_1^{mixed},\Pi_2^{mixed}\right)$

Note that all other values of threshold can be analyzed via 1. and 2., because for $\theta^R_{(i,i),j}\geq 1$ simply the subcase $\theta>\theta^R_{(i,i),j}$ is left out of the analysis. The same holds for negative values, i.e. $\theta^R_{(i,i),j}<0$ simply the case $\theta>\theta^R_{(i,i),j}$ is left out of the analysis. The last statement may need some clarification. A negative threshold implies that a deviation from an equilibrium (not only decreases the material payoff, but also) increases inequality. In that case for no level of inequality aversion a deviation from equilibrium becomes profitable in utility terms. This is equivalent to a situation where an equilibrium is contestable but inequality aversion is too weak to indeed destabilize the equilibrium, i.e. $\theta > \theta^R_{(i,i),j}$ is left out. Note that strictness of the problem of distribution excludes case 2b (2). Table C-4 depicts equilibrium payoffs in the various matches.

	$0 \le \theta \le \theta^R_{(0,0),2}, \theta^R_{(0,0),2} < 0, \theta^R_{(0,0),2} \ge 1$	$\theta > \theta^R_{(0,0),2}$
$0 \le \theta \le \theta^R_{(0,0),1}$ $\theta^R_{(0,0),1} < 0$	(1) $\left(\frac{a+d}{2}, \frac{A+D}{2}\right)$	(1) (d,D)
$\theta^R_{(0,0),1} \ge 1$	(2) (a,A)	(2) (b,B)
$\theta > \theta^R_{(0,0),1}$	(1) (d,D)	(1) (d,D)

Table C-4: Equilibrium payoffs according to the degree of inequality aversion of the matched players.

$\theta^R_{H,L} = \min\left\{\hat{\theta}^R_{H,L}, \breve{\theta}^R_{H,L}\right\}$, $\theta^R_{(1,1),2} > 1$ and $(1,1)$ being the materially more equal distributed equilibrium imply that $\theta^R_H = \breve{\theta}^R_H = \theta^R_{(0,0),2}$. Furthermore $\theta^R_L = \min\left\{\theta^R_{(0,0),1}, \theta^R_{(1,1),1}\right\}$.

1. $\theta^R_{(0,0),2} \ge 1 \vee \theta^R_{(0,0),2} < 0$ (materially more unequal distributed equilibrium is not contestable for high type)

Obviously all high types will earn the same payoff. Hence the distribution of inequality aversion among high types is determined by initial conditions and random shift. With respect to low types let me first consider the case when the materially more unequal distributed equilibrium is contestable, i.e. $\theta^R_L = \theta^R_{(0,0),1}$. In that case payoffs for low types are given by $\Pi_L^{\theta<\theta^R_L} = \frac{a+d}{2}$ and $\Pi_L^{\theta\ge\theta^R_L} = d$. Hence difference is given by $\Pi_L^{\theta\ge\theta^R_L} - \Pi_L^{\theta<\theta^R_L} = \frac{d-a}{2} > 0$ and the globally stable equilibrium is characterized by $F_L\left(\theta^R_L\right) = 0$. Furthermore $\frac{\partial\left(\Pi_L^{\theta\ge\theta^R_L} - \Pi_L^{\theta<\theta^R_L}\right)}{\partial\left(1-F_H\left(\theta^R_H\right)\right)} = 0$. I will now turn to the case where the materially less unequal equilibrium is contestable for the low type, i.e. $\theta^R_L = \theta^R_{(1,1),1}$. Payoff differences is given by $\Pi_L^{\theta\ge\theta^R_L} - \Pi_L^{\theta<\theta^R_L} = d-a > 0$. Hence the globally stable equilibrium is characterized by $F_L\left(\theta^R_L\right) = 0$. Furthermore $\frac{\partial\left(\Pi_L^{\theta\ge\theta^R_L} - \Pi_L^{\theta<\theta^R_L}\right)}{\partial\left(1-F_H\left(\theta^R_H\right)\right)} = 0$.

Finally if none of the material equilibria is contestable for the low type, the distribution of inequality aversion among low types is determined by initial conditions and random shift.

2. $\theta^R_{(0,0),2} \in (0,1)$

Let me first consider the case when the materially more unequal distributed equilibrium is contestable for the low type, i.e. $\theta^R_L = \theta^R_{(0,0),1}$. In that case payoffs for low types are given by $\Pi_L^{\theta<\theta^R_L} = F_H\left(\theta^R_H\right)\frac{a+d}{2} + \left(1-F_H\left(\theta^R_H\right)\right)d$ and $\Pi_L^{\theta\ge\theta^R_L} = F_H\left(\theta^R_H\right)d + \left(1-F_H\left(\theta^R_H\right)\right)d = d$, for high types $\Pi_H^{\theta<\theta^R_H} = F_L\left(\theta^R_L\right)\frac{A+D}{2} + \left(1-F_L\left(\theta^R_L\right)\right)D$ and $\Pi_H^{\theta\ge\theta^R_H} = F_L\left(\theta^R_L\right)D + \left(1-F_H\left(\theta^R_H\right)\right)D = D$. Hence differences are given by $\Pi_L^{\theta\ge\theta^R_L} - \Pi_L^{\theta<\theta^R_L} = F_H\left(\theta^R_H\right)\frac{d-a}{2} > 0$ and $\Pi_H^{\theta\ge\theta^R_H} - \Pi_H^{\theta<\theta^R_H} = -F_L\left(\theta^R_L\right)\frac{A-D}{2} < 0$. Hence the globally stable equilibrium is characterized by $F_H\left(\theta_H^R\right)=1, F_L\left(\theta_L^R\right)=0$. Furthermore $\frac{\partial\left(\Pi_H^{\theta\geq\theta_H^R}-\Pi_H^{\theta<\theta_H^R}\right)}{\partial\left(1-F_L\left(\theta_L^R\right)\right)}=\frac{A-D}{2}>0$ and $\frac{\partial\left(\Pi_L^{\theta\geq\theta_L^R}-\Pi_L^{\theta<\theta_L^R}\right)}{\partial\left(1-F_H\left(\theta_H^R\right)\right)}=-\frac{d-a}{2}<0$.

I will now turn to the case where the materially less unequal equilibrium is contestable for the low type, i.e. $\theta_L^R=\theta_{(1,1),1}^R$. In that case payoffs are given by $\Pi_L^{\theta<\theta_L^R}=F_H\left(\theta_H^R\right)a+\left(1-F_H\left(\theta_H^R\right)\right)b$ and $\Pi_L^{\theta\geq\theta_L^R}=F_H\left(\theta_H^R\right)d+\left(1-F_H\left(\theta_H^R\right)\right)d=d$, for high types $\Pi_H^{\theta<\theta_H^R}=F_L\left(\theta_L^R\right)A+\left(1-F_L\left(\theta_L^R\right)\right)D$ and $\Pi_H^{\theta\geq\theta_H^R}=F_L\left(\theta_L^R\right)B+\left(1-F_H\left(\theta_H^R\right)\right)D$. Hence differences are given by $\Pi_L^{\theta\geq\theta_L^R}-\Pi_L^{\theta<\theta_L^R}=d-b-F_H\left(\theta_H^R\right)(a-b)$ and $\Pi_H^{\theta\geq\theta_H^R}-\Pi_H^{\theta<\theta_H^R}=F_L\left(\theta_L^R\right)(B-A)\leq 0$. Note that $\left(\Pi_L^{\theta\geq\theta_L^R}-\Pi_L^{\theta<\theta_L^R}\right)\left(F_H\left(\theta_H^R\right)=0\right)=d-b>0$ and $\left(\Pi_L^{\theta\geq\theta_L^R}-\Pi_L^{\theta<\theta_L^R}\right)\left(F_H\left(\theta_H^R\right)=1\right)=d-a>0$. Hence the globally stable equilibrium is given by $F_H\left(\theta_H^R\right)=1, F_L\left(\theta_L^R\right)=0$. Furthermore $\frac{\partial\left(\Pi_H^{\theta\geq\theta_H^R}-\Pi_H^{\theta<\theta_H^R}\right)}{\partial\left(1-F_L\left(\theta_L^R\right)\right)}=A-B>0$ and $\frac{\partial\left(\Pi_L^{\theta\geq\theta_L^R}-\Pi_L^{\theta<\theta_L^R}\right)}{\partial\left(1-F_H\left(\theta_H^R\right)\right)}=a-b$.

Finally if none of the material equilibria is contestable for the low type, the distribution of inequality aversion among low types is determined by initial conditions and random shift. Payoff difference for high types is given by $\Pi_H^{\theta\geq\theta_H^R}-\Pi_H^{\theta<\theta_H^R}=F_L\left(\theta_L^R\right)(B-A)\leq 0$ with $\frac{\partial\left(\Pi_H^{\theta\geq\theta_H^R}-\Pi_H^{\theta<\theta_H^R}\right)}{\partial\left(1-F_L\left(\theta_L^R\right)\right)}=A-B>0$. Hence the globally stable equilibrium is given by $F_H\left(\theta_H^R\right)=1$.

<u>Case (2) of Lemma 4-4:</u>

Let us now turn to case (2) of Lemma 4-4 with two Pareto-ranked equilibria. Given the assumption parallel to case A this leaves us with two possibilities, either $d<a\wedge D<A$ or $a<d\wedge A<D$. For ease of comparability to case (1) of Lemma 4-4 I will w.l.o.g. assume $a<d\wedge A<D$ to hold. Hence the only relation that has changed in comparison to case (1) is the one between parameters A and D. Note that inequalities (i) and (ii) still hold. Again, due to (ii) $\theta_{(1,1),2}^R>1$, i.e. the equilibrium $(1,1)$ is not contestable for player two. That is, in case (2) the Pareto-superior equilibrium is not contestable for high types. The equilibrium analysis is equivalent to case A and equilibrium payoffs correspond to those in Table C-4, their relation to each other may have changed though.

1. $\theta_{(0,0),2}^R\geq 1\vee\theta_{(0,0),2}^R<0$ (Pareto-inferior equilibrium is not contestable for high type)

Parameters A and D are not involved, hence the results are identical to those in case A.

2. $\theta_{(0,0),2}^R\in(0,1)$

Let me first consider the case when the Pareto-inferior equilibrium is contestable for the low type, i.e. $\theta_L^R=\theta_{(0,0),1}^R$. Payoffs are equivalent to case (1). Differences in payoffs among low types is given by $\Pi_L^{\theta\geq\theta_L^R}-\Pi_L^{\theta<\theta_L^R}=F_H\left(\theta_H^R\right)\frac{d-a}{2}>0$ and by $\Pi_H^{\theta\geq\theta_H^R}-\Pi_H^{\theta<\theta_H^R}=-F_L\left(\theta_L^R\right)\frac{A-D}{2}>0$ among high types.

Hence the globally stable equilibrium is given by $F_H\left(\theta_H^R\right)=F_L\left(\theta_L^R\right)=0$. Furthermore $\frac{\partial\left(\Pi_H^{\theta\geq\theta_H^R}-\Pi_H^{\theta<\theta_H^R}\right)}{\partial\left(1-F_L\left(\theta_L^R\right)\right)}=\frac{A-D}{2}<0$ and $\frac{\partial\left(\Pi_L^{\theta\geq\theta_L^R}-\Pi_L^{\theta<\theta_L^R}\right)}{\partial\left(1-F_H\left(\theta_H^R\right)\right)}=-\frac{d-a}{2}<0$.

I will now turn to the case where the Pareto-superior equilibrium is contestable for the low type, i.e. $\theta_L^R=\theta_{(1,1),1}^R$. In that case payoffs are given by $\Pi_L^{\theta<\theta_L^R}=F_H\left(\theta_H^R\right)a+\left(1-F_H\left(\theta_H^R\right)\right)b$ and $\Pi_L^{\theta\geq\theta_L^R}=F_H\left(\theta_H^R\right)d+\left(1-F_H\left(\theta_H^R\right)\right)d=d$, for high types $\Pi_H^{\theta<\theta_H^R}=F_L\left(\theta_L^R\right)A+\left(1-F_L\left(\theta_L^R\right)\right)D$ and $\Pi_H^{\theta\geq\theta_H^R}=F_L\left(\theta_L^R\right)B+\left(1-F_H\left(\theta_H^R\right)\right)D$. Hence differences are given by $\Pi_L^{\theta\geq\theta_L^R}-\Pi_L^{\theta<\theta_L^R}=d-b-F_H\left(\theta_H^R\right)(a-b)$ and $\Pi_H^{\theta\geq\theta_H^R}-\Pi_H^{\theta<\theta_H^R}=F_L\left(\theta_L^R\right)(B-A)\leq 0$. Note that $\left(\Pi_L^{\theta\geq\theta_L^R}-\Pi_L^{\theta<\theta_L^R}\right)\left(F_H\left(\theta_H^R\right)=0\right)=d-b>0$ and $\left(\Pi_L^{\theta\geq\theta_L^R}-\Pi_L^{\theta<\theta_L^R}\right)\left(F_H\left(\theta_H^R\right)=1\right)=d-a>0$. Hence the globally stable equilibrium is given by $F_H\left(\theta_H^R\right)=1, F_L\left(\theta_L^R\right)=0$. Furthermore $\frac{\partial\left(\Pi_H^{\theta\geq\theta_H^R}-\Pi_H^{\theta<\theta_H^R}\right)}{\partial\left(1-F_L\left(\theta_L^R\right)\right)}=A-B>0$ and $\frac{\partial\left(\Pi_L^{\theta\geq\theta_L^R}-\Pi_L^{\theta<\theta_L^R}\right)}{\partial\left(1-F_H\left(\theta_H^R\right)\right)}=a-b$.

Finally if none of the material equilibria is contestable for the low type, the distribution of inequality aversion among low types is determined by initial conditions and random shift. Payoff difference for high types is given by $\Pi_H^{\theta\geq\theta_H^R}-\Pi_H^{\theta<\theta_H^R}=F_L\left(\theta_L^R\right)(B-A)\leq 0$ with $\frac{\partial\left(\Pi_H^{\theta\geq\theta_H^R}-\Pi_H^{\theta<\theta_H^R}\right)}{\partial\left(1-F_L\left(\theta_L^R\right)\right)}=A-B>0$. Hence the globally stable equilibrium is given by $F_H\left(\theta_H^R\right)=1$. QED

<u>Proof of Theorem:</u> Let $\theta^D=\theta^C=\theta_{H,L}^R\equiv\theta^{crit}$ and $d\Pi_H^R\leq 0$. $d\Pi_H^R\leq 0$ implies that $\beta_H^R=-\alpha_H^R$.

Payoff differences are given by

$$\begin{aligned}
&d\Pi_H^S\geq 0\Leftrightarrow\left(1-F_L^t\right)\left(\mu\beta^D+\nu\beta^C-(1-\mu-\nu)\alpha_H^R\right)\geq\mu\beta^D+\nu\beta^C-\nu\alpha^C-(1-\mu-\nu)\alpha_H^R-\left(1-F_H^t\right)\left(\mu\beta^D+\nu\beta^C\right)\\
&d\Pi_L^S\geq 0\Leftrightarrow\left(1-F_L^t\right)\left(\mu\beta^D+\nu\beta^C\right)\geq\mu\beta^D+\nu\beta^C-\nu\alpha^C-(1-\mu-\nu)\alpha_L^R-\left(1-F_H^t\right)\left(\mu\beta^D+\nu\beta^C+(1-\mu-\nu)\beta_L^R\right)
\end{aligned}\quad(*)$$

I will distinguish 3 cases (i) $0<\mu\beta^D+\nu\beta^C<\mu\beta^D+\nu\beta^C-(1-\mu-\nu)\alpha_H^R$, (ii) $\mu\beta^D+\nu\beta^C<\mu\beta^D+\nu\beta^C-(1-\mu-\nu)\alpha_H^R<0$ and (iii) $\mu\beta^D+\nu\beta^C<0<\mu\beta^D+\nu\beta^C-(1-\mu-\nu)\alpha_H^R$.

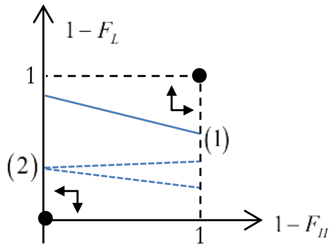
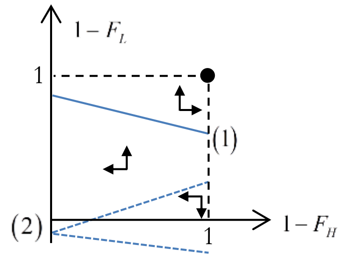
(i) $0<\mu\beta^D+\nu\beta^C<\mu\beta^D+\nu\beta^C-(1-\mu-\nu)\alpha_H^R$

(*) can be written as:

$$\begin{aligned}
&(1):\left(1-F_L^t\right)\geq 1-\frac{\nu\alpha^C}{\mu\beta^D+\nu\beta^C-(1-\mu-\nu)\alpha_H^R}-\frac{\mu\beta^D+\nu\beta^C}{\mu\beta^D+\nu\beta^C-(1-\mu-\nu)\alpha_H^R}\left(1-F_H^t\right)\\
&(2):\left(1-F_L^t\right)\geq 1-\frac{\nu\alpha^C+(1-\mu-\nu)\alpha_L^R}{\mu\beta^D+\nu\beta^C}-\frac{\mu\beta^D+\nu\beta^C+(1-\mu-\nu)\beta_L^R}{\mu\beta^D+\nu\beta^C}\left(1-F_H^t\right)
\end{aligned}$$

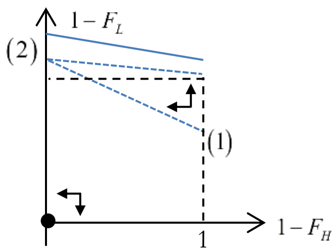
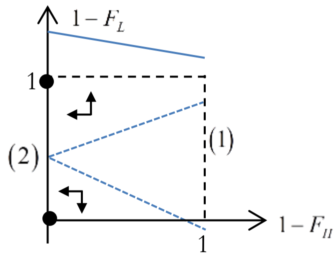
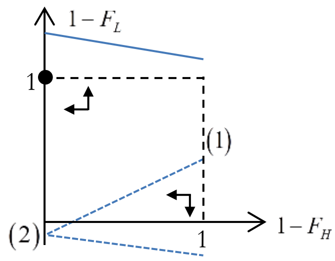
a) $\alpha^C>0$:

It follows that the intercept of (1) in below one and above the intercept of (2). Given the negative slope of (1) essentially 2 cases can be distinguished. The following table depicts the phase diagrams which clearly indicate the stable equilibria. The last row states the precise condition for the case considered.

$F_H = F_L = 0,\ F_H = F_L = 1$	$F_H = F_L = 0$
	
$1-\frac{\nu\alpha^C+(1-\mu-\nu)\alpha_L^R}{\mu\beta^D+\nu\beta^C}>0 \Leftrightarrow$ $\alpha^C<\frac{\mu\beta^D+\nu\beta^C-(1-\mu-\nu)\alpha_L^R}{\nu}$	$1-\frac{\nu\alpha^C+(1-\mu-\nu)\alpha_L^R}{\mu\beta^D+\nu\beta^C}\le 0 \Leftrightarrow$ $\alpha^C\ge\frac{\mu\beta^D+\nu\beta^C-(1-\mu-\nu)\alpha_L^R}{\nu}$

b) $\alpha^C \le 0$:

It follows that the intercept of (1) above one and above the intercept of (2). Given the negative slope of (1) essentially 4 cases can be distinguished. The following table depicts the phase diagrams which clearly indicate the stable equilibria. The last row states the precise condition for the case considered.

$F_H = F_L = 1$	$F_H = 1,\ F_L = 0;\ F_H = F_L = 1$	$F_H = 1,\ F_L = 0$
		
$1-\frac{\nu\alpha^C+(1-\mu-\nu)\alpha_L^R}{\mu\beta^D+\nu\beta^C}>1 \Leftrightarrow$ $\alpha^C<-\frac{1-\mu-\nu}{\nu}\alpha_L^R$	$1-\frac{\nu\alpha^C+(1-\mu-\nu)\alpha_L^R}{\mu\beta^D+\nu\beta^C}\in(0,1)$	$1-\frac{\nu\alpha^C+(1-\mu-\nu)\alpha_L^R}{\mu\beta^D+\nu\beta^C}<0 \Leftrightarrow$ $\alpha^C>\frac{\mu\beta^D+\nu\beta^C-(1-\mu-\nu)\alpha_L^R}{\nu}$

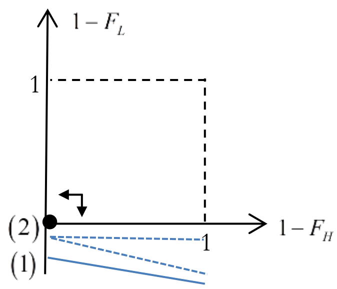
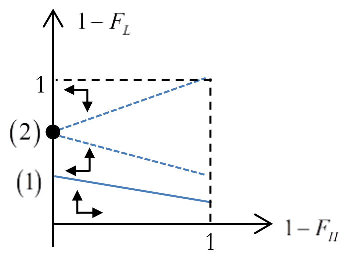
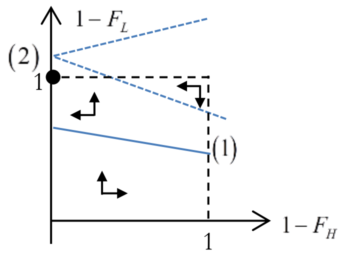
The same three cases emerge if (1) has a value of below one at $1-F_H^t=1$. However in that case an additional locally stable equilibrium arises, that of $F_H = F_L = 0$. The condition for this is $\mu\beta^D+\nu\left(\alpha^C+\beta^C\right)>0$.

(ii) $\mu\beta^D + \nu\beta^C < \mu\beta^D + \nu\beta^C - (1-\mu-\nu)\alpha_H^R < 0$

Note that the slope of (1) is again negative.

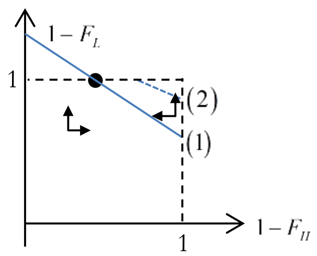
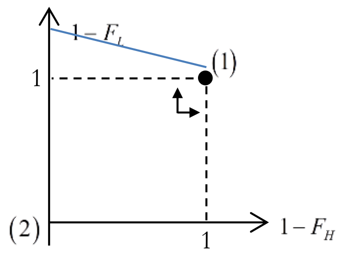
a) $\alpha^C < 0$:

It follows that the intercept of (1) in below one and below the intercept of (2). Given the negative slope of (1) essentially 3 cases can be distinguished. The following table depicts the phase diagrams which clearly indicate the stable equilibria. The last row states the precise condition for the case considered.

<table>
<tr><td>$F_H = F_L = 1$</td><td>$F_H = 1,\ F_L = 1 - \frac{\nu\alpha^C + (1-\mu-\nu)\alpha_L^R}{\mu\beta^D + \nu\beta^C}$</td><td>$F_H = 1,\ F_L = 0$</td></tr>
<tr><td>
</td><td>
</td><td>
</td></tr>
<tr><td>$1 - \frac{\nu\alpha^C + (1-\mu-\nu)\alpha_L^R}{\mu\beta^D + \nu\beta^C} \le 0 \Leftrightarrow$
$\alpha^C \le \frac{\mu\beta^D + \nu\beta^C - (1-\mu-\nu)\alpha_L^R}{\nu}$</td><td>$1 - \frac{\nu\alpha^C + (1-\mu-\nu)\alpha_L^R}{\mu\beta^D + \nu\beta^C} \in (0,1)$</td><td>$1 - \frac{\nu\alpha^C + (1-\mu-\nu)\alpha_L^R}{\mu\beta^D + \nu\beta^C} \ge 1 \Leftrightarrow$
$\alpha^C \ge -\frac{(1-\mu-\nu)}{\nu}\alpha_L^R$</td></tr>
</table>

b) $\alpha^C \ge 0$:

In this case the intercept of (1) is above one. The following two cases can be distinguished.

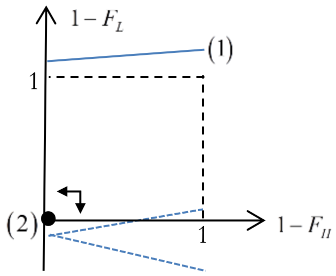
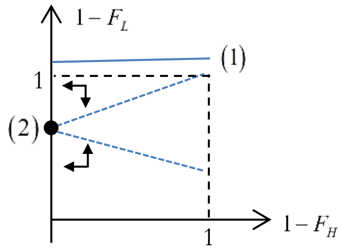
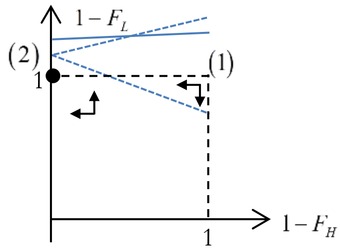
<table>
<tr><td>$F_L = 0,\ F_H = 1 + \frac{\nu\alpha^C}{\mu\beta^D + \nu\beta^C}$</td><td>$F_H = F_L = 0$</td></tr>
<tr><td>
</td><td>
</td></tr>
<tr><td>$1 - \frac{\nu\alpha^C}{\mu\beta^D + \nu\beta^C - (1-\mu-\nu)\alpha_H^R} - \frac{\mu\beta^D + \nu\beta^C}{\mu\beta^D + \nu\beta^C - (1-\mu-\nu)\alpha_H^R} < 1$
$\Leftrightarrow \mu\beta^D + \nu(\alpha^C + \beta^C) < 0$</td><td>$\mu\beta^D + \nu(\alpha^C + \beta^C) \ge 0$</td></tr>
</table>

(iii) $\mu\beta^D + \nu\beta^C < 0 < \mu\beta^D + \nu\beta^C - (1-\mu-\nu)\alpha_H^R$

Note that the intercept of (1) is above one.

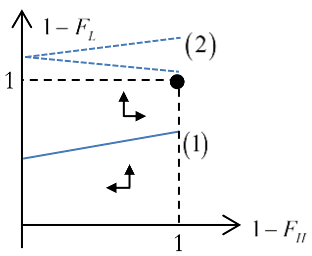
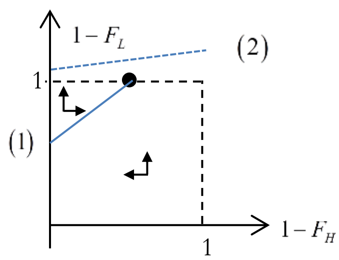
a) $\alpha^C < 0$:

It follows that the slope of (1) is positive. Essentially 3 cases can be distinguished.

$F_H = F_L = 1$	$F_H = 1,\ F_L = 1 - \frac{\nu\alpha^C + (1-\mu-\nu)\alpha_L^R}{\mu\beta^D + \nu\beta^C}$	$F_H = 1,\ F_L = 0$
		
$1 - \frac{\nu\alpha^C + (1-\mu-\nu)\alpha_L^R}{\mu\beta^D + \nu\beta^C} \le 0 \Leftrightarrow$ $\alpha^C \le \frac{\mu\beta^D + \nu\beta^C - (1-\mu-\nu)\alpha_L^R}{\nu}$	$1 - \frac{\nu\alpha^C + (1-\mu-\nu)\alpha_L^R}{\mu\beta^D + \nu\beta^C} \in (0,1)$	$1 - \frac{\nu\alpha^C + (1-\mu-\nu)\alpha_L^R}{\mu\beta^D + \nu\beta^C} \ge 1 \Leftrightarrow$ $\alpha^C \ge -\frac{(1-\mu-\nu)}{\nu}\alpha_L^R$

b) $\alpha^C \ge 0$:

In this case the intercept of (1) is below one whereas the intercept of (2) is above one. The following two cases can be distinguished.

$F_H = F_L = 0$	$F_L = 0,\ F_H = 1 + \frac{\nu\alpha^C}{\mu\beta^D + \nu\beta^C}$
	
$1 - \frac{\nu\alpha^C}{\mu\beta^D + \nu\beta^C - (1-\mu-\nu)\alpha_H^R} - \frac{\mu\beta^D + \nu\beta^C}{\mu\beta^D + \nu\beta^C - (1-\mu-\nu)\alpha_H^R} < 1$ $\Leftrightarrow \mu\beta^D + \nu(\alpha^C + \beta^C) > 0$	$\mu\beta^D + \nu(\alpha^C + \beta^C) \le 0$

QED

Proof of Lemma 4-7: The center of mass of the polytope given by the vertices presented in Table 4-4 can be calculated as the average with relative volume as weights of the centers of mass of the two pyramids DCGE and DCFE, i.e.

$$CM_{P_{CDEFG}} = \frac{V_{P_{DCGE}}}{V_{P_{DCGE}} + V_{P_{DCFE}}} CM_{P_{DCGE}} + \frac{V_{P_{DCFE}}}{V_{P_{DCGE}} + V_{P_{DCFE}}} CM_{P_{DCFE}} .$$

The center of mass for these two pyramids is located on the line segment connecting the center of the (any) triangular base and the top of the pyramids. Some elementary algebra yields

$$V_{P_{DCGE}} = \frac{\alpha_1}{6} \frac{1}{1+\alpha_1+\alpha_1\alpha_2} \frac{\alpha_2}{(1+\alpha_1)(1+\alpha_2)},\ V_{P_{DCFE}} = \frac{\alpha_1}{6} \frac{1}{1+\alpha_2+\alpha_1\alpha_2} \frac{\alpha_2}{(1+\alpha_1)(1+\alpha_2)} .$$

Furthermore the centers of mass of the two pyramids are given by:

$$CM_{P_{DCGE}} = \begin{pmatrix} \mu_{00}^{CM_{P_{DCGE}}} \\ \mu_{10}^{CM_{P_{DCGE}}} \\ \mu_{01}^{CM_{P_{DCGE}}} \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 1 + \frac{1}{1+\alpha_1+\alpha_1\alpha_2} + \frac{1}{(1+\alpha_1)(1+\alpha_2)} \\ \alpha_2 \frac{1}{(1+\alpha_1)(1+\alpha_2)} \\ \alpha_1 \left(\frac{1}{1+\alpha_1+\alpha_1\alpha_2} + \frac{1}{(1+\alpha_1)(1+\alpha_2)} \right) \end{pmatrix}$$

and

$$CM_{P_{DCFE}} = \begin{pmatrix} \mu_{00}^{CM_{P_{DCFE}}} \\ \mu_{10}^{CM_{P_{DCFE}}} \\ \mu_{01}^{CM_{P_{DCFE}}} \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 1 + \frac{1}{1+\alpha_2+\alpha_1\alpha_2} + \frac{1}{(1+\alpha_1)(1+\alpha_2)} \\ \alpha_2 \left(\frac{1}{1+\alpha_2+\alpha_1\alpha_2} + \frac{1}{(1+\alpha_1)(1+\alpha_2)} \right) \\ \alpha_1 \frac{1}{(1+\alpha_1)(1+\alpha_2)} \end{pmatrix}$$

Plugging in values and rearranging terms yield the stated equation. QED

Proof of Proposition 4-5: Again, let the symmetric dilemma be represented by the following matrix $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$. Expected payoffs are then given by: $E\pi_1 = \mu_{00}a + \mu_{01}c + \mu_{10}b + \mu_{11}d$, $E\pi_2 = \mu_{00}a + \mu_{01}b + \mu_{10}a + \mu_{11}d$ and hence the difference by: $E\pi_1 - E\pi_2 = (\mu_{10} - \mu_{01})(b-c)$. Plugging in the values for the center of mass (see Lemma 4-7) yields:

$$E\pi_1 - E\pi_2 = (\alpha_2 - \alpha_1)(b-c) \underbrace{\left(\frac{1+\alpha_1\alpha_2}{(1+\alpha_1+\alpha_1\alpha_2)(1+\alpha_2+\alpha_1\alpha_2)} + \frac{1}{(1+\alpha_1)(1+\alpha_2)} - \frac{1}{(1+\alpha_1+\alpha_1\alpha_2)+(1+\alpha_2+\alpha_1\alpha_2)} \right)}_{>0}$$ QED

Proof of Proposition 4-6: The line of argument in the proof of Proposition 4-1 is still valid, i.e. if and only if two individuals are matched who are sufficiently inequality-averse the set of equilibria changes. In case of the concept of correlated equilibria the vertices of the set are given in Table 4-4. According to Proposition 4-5 for two such individuals the one with the lower degree of inequality-aversion earns higher profits. Furthermore the difference in profits is monotonic decreasing in the difference in the degrees of inequality-aversion. Hence the highest profits is earned by individuals with $\theta = \theta^D$ and the lowest profits such an individual can earn is realized when matched with another individual with $\theta = \theta^D$. Again, let the symmetric dilemma be represented by the following matrix $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$. W.l.o.g. $a > d$, in that case $\theta_1 = \theta_2 = \theta^D$ implies $\alpha_1 = \alpha_2 = 0$ and $\mu_{00}^{CM} = \frac{3}{4}, \mu_{11}^{CM} = \frac{1}{4}, \mu_{01}^{CM} = \mu_{10}^{CM} = 0$ yielding expected payoff $E\pi = \frac{3}{4}a + \frac{1}{4}d > d$ strictly greater than the payoff received by opportunistic individuals. Hence the only stable equilibrium that can emerge is the singular distribution with all agents sharing the same degree of inequality-aversion.

I turn now to the non-PD-case. In that case the results with respect to profits for individuals with $\theta \geq \theta^D$ also hold. No two different values θ_1, θ_2 with $\theta_1, \theta_2 \geq \theta^D$ can be part of an equilibrium, because both individual earn the same profit when matched with an opportunistic opponent and the one with the lower degree of inequality-aversion earns higher profits than the one with the higher value in any match with some other agent with $\theta \geq \theta^D$. Hence only types with $\theta = \theta^D$ could be part of an equilibrium. However the same calculation of expected payoffs as in the PD-case applies, but in the non-PD-case this amount to a disadvantage because w.l.o.g. $b > a,\ d > c,\ \frac{b+c}{2} > d,\ a \leq d$ and thereby $E\pi = \frac{3}{4}a + \frac{1}{4}d \leq d$. Hence the globally stable equilibrium distribution is characterized by $F\left(\theta^D\right) = 1$. QED

<u>Proof of Proposition 4-7:</u> Given the definition of thresholds and the derivation of different equilibria in the proof of Proposition 4-2, I focus herein on the case where one player alone can destabilize both pure Nash equilibria. By symmetry potentially both players can thus destabilize all equilibria individually. Again, since inequality aversion has no leverage on coordination games, I study anti-coordination games. In other words, I am concerned with games represented by a matrix $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ such that $b < a,\ c > d$. Both equilibria being contestable is equivalent to

$$\theta_{(0,1),2}^C, \theta_{(1,0),1}^C < 1,\ \theta_{(0,1),1}^C, \theta_{(1,0),2}^C < 1 \text{ and } \theta_{(0,1),2}^C, \theta_{(1,0),1}^C < 1,\ \theta_{(0,1),1}^C, \theta_{(1,0),2}^C < 1 \Leftrightarrow \begin{cases} c < \left\{a, \frac{b+d}{2}\right\},\ b > c \Rightarrow b > a > c > d \\ b < \left\{d, \frac{a+c}{2}\right\},\ b < c \Rightarrow c > d > b > a \end{cases}.$$

I first study the case $\theta_{(0,1),2}^C = \theta_{(1,0),1}^C < \theta_{(0,1),1}^C = \theta_{(1,0),2}^C \left(\Leftrightarrow \alpha < \beta \Leftrightarrow \frac{b-a}{c-d} < 1 \right)$.

Table C-5 below presents the payoffs depending on the two level of inequality aversion being matched. I will refer to an individual in lowest interval, medium and high interval as A, B and C-types respectively.

	A	B	C
	$\theta_2 < \theta^C_{(0,1),2} < \theta^C_{(1,0),2}$	$\theta^C_{(0,1),2} < \theta_2 < \theta^C_{(1,0),2}$ '0' is dominant str.	$\theta^C_{(0,1),2} < \theta^C_{(1,0),2} < \theta_2$
$\theta_1 < \theta^C_{(1,0),1} < \theta^C_{(0,1),1}$	$\left(\frac{b+c}{2},\frac{b+c}{2}\right)$	(b,c)	$\left(\Pi_1^{mix}(\alpha_1,\alpha_2),\Pi_2^{mix}(\alpha_1,\alpha_2)\right)$
$\theta^C_{(1,0),1} < \theta_1 < \theta^C_{(0,1),1}$	(c,b)	(a,a)	(a,a)
$\theta^C_{(1,0),1} < \theta^C_{(0,1),1} < \theta_1$	$\left(\Pi_1^{mix}(\alpha_1,\alpha_2),\Pi_2^{mix}(\alpha_1,\alpha_2)\right)$	(a,a)	$\left(\frac{a+d}{2},\frac{a+d}{2}\right)$

Table C-5: Payoffs in the various matches.

For the mixed equilibrium: $\alpha_1 = \frac{\left|a-b+\theta_1\left|b-c\right|\right|}{\left|c-\theta_1\left|b-c\right|+d\right|}$, $\alpha_2 = \frac{\left|a-b+\theta_2\left|b-c\right|\right|}{\left|c-\theta_2\left|b-c\right|+d\right|}$, and

$\alpha < \beta \Leftrightarrow \frac{b-a}{c-d} < 1 \Leftrightarrow \theta^C_{(0,1),2} = \theta^C_{(1,0),1} < \theta^C_{(0,1),1} = \theta^C_{(1,0),2}$.

I will first consider the case $b > c$. Note that in that case B-types destabilize the equilibrium that favors them, but not the equilibrium that disfavors them. This suggests an evolutionary disadvantage for B-types. Profits if there exists only a mixed Nash equilibrium, i.e. in a match between A-types and C-types are given by:

$$\Pi_1^{mix}(\alpha_1,\alpha_2) = \frac{1}{(1+\alpha_1)(1+\alpha_2)}(a+\alpha_1 b+\alpha_2 c+\alpha_1\alpha_2 d),\quad \Pi_2^{mix}(\alpha_1,\alpha_2) = \frac{1}{(1+\alpha_1)(1+\alpha_2)}(a+\alpha_1 c+\alpha_2 b+\alpha_1\alpha_2 d)$$

$$\Pi_1^{mix}(\alpha_1,\alpha_2) - \Pi_2^{mix}(\alpha_1,\alpha_2) = \frac{(\alpha_2-\alpha_1)}{(1+\alpha_1)(1+\alpha_2)}(b-c) > 0 \Leftrightarrow \alpha_1 < \alpha_2$$

Consider a match between type A as player one and type C as player two, i.e. player one is opportunistic and player two is highly inequality-averse. In that case

$$\alpha_1 < \alpha_2 \Leftrightarrow \frac{\left|a-b+\theta_1\left|b-c\right|\right|}{\left|c-\theta_1\left|b-c\right|+d\right|} < \frac{\left|a-b+\theta_2\left|b-c\right|\right|}{\left|c-\theta_2\left|b-c\right|+d\right|} \Leftrightarrow \frac{a-b+\theta_1\left|b-c\right|}{d-c+\theta_1\left|b-c\right|} < \frac{a-b+\theta_2\left|b-c\right|}{d-c+\theta_2\left|b-c\right|}$$
$$\Leftrightarrow (a-b)\theta_2\left|b-c\right| + (d-c)\theta_1\left|b-c\right| > (a-b)\theta_1\left|b-c\right| + (d-c)\theta_2\left|b-c\right| \underset{b-a<c-d}{\Leftrightarrow} \theta_1 < \theta_2$$

When I considered a strict and symmetric problem of coordination type C player were simply left out of analysis. Thus I will focus on equilibria with type C players. Note that there can be no B, C equilibrium, because C players would be worse off. For the same reason there cannot be an equilibrium with only C players, since B players could successfully invade.

In an equilibrium with both types A and C present, only players with minimal $\alpha_1 = \alpha_1\left(\theta_1 = \frac{b-a}{b-c}\right) = 0$ among A types and those with minimal $\alpha_2 = \alpha_2(\theta_2 = 1) = \frac{a-c}{d+b-2c}$ can be part of the equilibrium, because $\frac{\partial \Pi_1^{mix}(\alpha_1,\alpha_2)}{\partial \alpha_1} = -\frac{a-c+(b-d)\alpha_2}{(1+\alpha_1)^2(1+\alpha_2)} < 0$ and $\frac{\partial \Pi_2^{mix}(\alpha_1,\alpha_2)}{\partial \alpha_2} = -\frac{a-c+(b-d)\alpha_1}{(1+\alpha_1)(1+\alpha_2)^2} < 0$. Due to $\frac{\partial \alpha_1}{\partial \theta_1} = \frac{b-c}{\left(c-d-\theta_1(b-c)\right)^2}\left(b-a-(c-d)\right) \underset{b-a<c-d}{<} 0$ and

$\frac{\partial \alpha_2}{\partial \theta_2} = \frac{b-c}{\left(c-d-\theta_2(b-c)\right)^2}\left(b-a-(c-d)\right) \underset{b-a<c-d}{<} 0$ minimal α translate into maximal θ. Table C-5 thus simplifies to:

	A	B	C
A	$\left(\frac{b+c}{2},\frac{b+c}{2}\right)$	(b,c)	$\left(\frac{-bc+a(2b-2c+d)}{a+b-3c+d},\frac{-c^2+a(b-c+d)}{a+b-3c+d}\right)$
B	(c,b)	(a,a)	(a,a)
C	$\left(\frac{-c^2+a(b-c+d)}{a+b-3c+d},\frac{-bc+a(2b-2c+d)}{a+b-3c+d}\right)$	(a,a)	$\left(\frac{a+d}{2},\frac{a+d}{2}\right)$

It turns out that type A types earn strictly higher payoffs than type C players, because

$\Pi_2^{mix}\left(\alpha_1=0,\alpha_2=\frac{a-c}{d+b-2c}\right)=\frac{-c^2+a(b-c+d)}{a+b-3c+d}<\frac{b+c}{2}$ and

$\Pi_1^{mix}\left(\alpha_1=0,\alpha_2=\frac{a-c}{d+b-2c}\right)=\frac{-bc+a(2b-2c+d)}{a+b-3c+d}>\frac{a+d}{2}$. Hence such an equilibrium cannot exist.

Intuitively, if $b>c$, then weighting the outcome (0,1) and (1,1) less reduced payoffs for player two. For the lowest weight payoffs for player two are a weighted average of a and c, and therefore higher than c.

Finally, I analyze whether there exists a A,B,C equilibrium. It turns out that for the most profitable type A player an even stronger inequality holds: $\Pi_1^{mix}\left(\alpha_1=0,\alpha_2=\frac{a-c}{d+b-2c}\right)>a$. Hence A-types would earn strictly higher profits than B-types in an A,B,C equilibrium.

Thus no additional equilibria arise.

I now turn to the case when $b<c$. To summarize conditions: $c>d>b>a$, $b-a<c-d$, $b<\frac{a+c}{2}$.

These conditions imply: $\frac{\partial \Pi_2^{mix}(\alpha_1,\alpha_2)}{\partial \alpha_2}=\frac{c-a+(d-b)\alpha_1}{(1+\alpha_1)(1+\alpha_2)^2}>0$ and $\frac{\partial \Pi_1^{mix}(\alpha_1,\alpha_2)}{\partial \alpha_1}=\frac{c-a+(d-b)\alpha_2}{(1+\alpha_1)^2(1+\alpha_2)}>0$.

Again I focus on equilibria with C types being present. I first consider the case with only C-types present in equilibrium. Such a distribution cannot be invaded by B types. The fittest A type is the one with maximal α_1, which transfers to a minimal θ_1. Note that the profit of the fittest A type is independent of the degree of inequality aversion of the C type, $\Pi_1^{mix}\left(\alpha_1=\frac{b-a}{c-d},\alpha_2\right)=\frac{bc-ad}{b-a+c-d}$.

Hence a locally stable equilibrium with only inequality averse players of type C exists if $\frac{bc-ad}{b-a+c-d}<\frac{a+d}{2}$.

I now study whether there is an equilibrium with A and C types present. Again, this demands $\alpha_1 = \frac{b-a}{c-d}, \alpha_2 = \infty$ implying the following profits: $\Pi_1^{mix}\left(\alpha_1 = \frac{b-a}{c-d}, \alpha_2\right) = \frac{bc-ad}{b-a+c-d}$ and $\Pi_2^{mix}\left(\alpha_1 = \frac{b-a}{c-d}, \alpha_2 \to \infty\right) = \frac{-c^2+(a-b)d+cd}{a-b-c+d}$. Thus Table C-5 simplifies to:

	A	**B**	**C**
A	$\left(\frac{b+c}{2}, \frac{b+c}{2}\right)$	(b,c)	$\left(\frac{bc-ad}{b-a+c-d}, \frac{-c^2+(a-b)d+cd}{a-b-c+d}\right)$
B	(c,b)	(a,a)	(a,a)
C	$\left(\frac{-c^2+(a-b)d+cd}{a-b-c+d}, \frac{bc-ad}{b-a+c-d}\right)$	(a,a)	$\left(\frac{a+d}{2}, \frac{a+d}{2}\right)$

Let Π^A and Π^C denote the payoffs of A-types and C-types respectively. Let $F(A)$ denote the share of A-types in equilibrium, then $\Pi^A = F(A)\frac{b+c}{2} + (1-F(A))\frac{bc-ad}{b-a+c-d}$ and $\Pi^C = F(A)\frac{-c^2+(a-b)d+cd}{a-b-c+d} + (1-F(A))\frac{a+d}{2}$.

An A,C equilibrium exists if and only if $\frac{bc-ad}{b-a+c-d} > \frac{a+d}{2}$, because $\frac{b+c}{2} < \frac{-c^2+(a-b)d+cd}{a-b-c+d}$ holds. In that case the equilibrium share of A-types is given by $F(A) = \frac{a^2+2bc-a(b+c)-(b+c)d+d^2}{a^2-b^2+(c-d)^2-2ad+2bd}$. The equilibrium is locally stable if the profits of B are smaller than equilibrium payoffs, given the equilibrium share of A and C-types. Note that a parameterization with $a=0,\ b=\frac{1}{5},\ d=\frac{1}{3},\ c=1$ indeed satisfies all condition, i.e. $c>d>b>a,\ b-a<c-d,\ b<\frac{a+c}{2}$, $\frac{bc-ad}{b-a+c-d} > \frac{a+d}{2}$, $F(A)\in(0,1)$, and $\Pi^B < \Pi^A$, thus such a stable equilibrium indeed exists.

I will finally analyze the existence of an A,B,C equilibrium. Payoffs of the different types are given by: $\Pi^A = F(A)\frac{b+c}{2} + F(B)b + (1-F(A)-F(B))\frac{bc-ad}{b-a+c-d}$,

$\Pi^B = F(A)c + F(B)a + (1-F(A)-F(B))a$, and

$\Pi^C = F(A)\frac{-c^2+(a-b)d+cd}{a-b-c+d} + F(B)a + (1-F(A)-F(A))\frac{a+d}{2}$. The two equations $\Pi^A = \Pi^B$ and $\Pi^B = \Pi^C$ imply the following equilibrium values for $F(A)$ and $F(B)$:

$$F(A) = \frac{2(a-b)(a-d)(a-b-c+d)^2}{\left(\begin{array}{l}2a^4+4b^3c-(b-c)(b+c)(3b+c)d+2(b^2-2bc-c^2)d^2+(b+c)d^3+a^3(-5(b+c)+2d)\\ +a^2(4(b^2+3bc+c^2)-5(b+c)d+2d^2)-a(b^3+b^2(11c-6d)+(c-d)(c^2+3cd-2d^2)+b(3c^2-4cd+3d^2))\end{array}\right)}$$

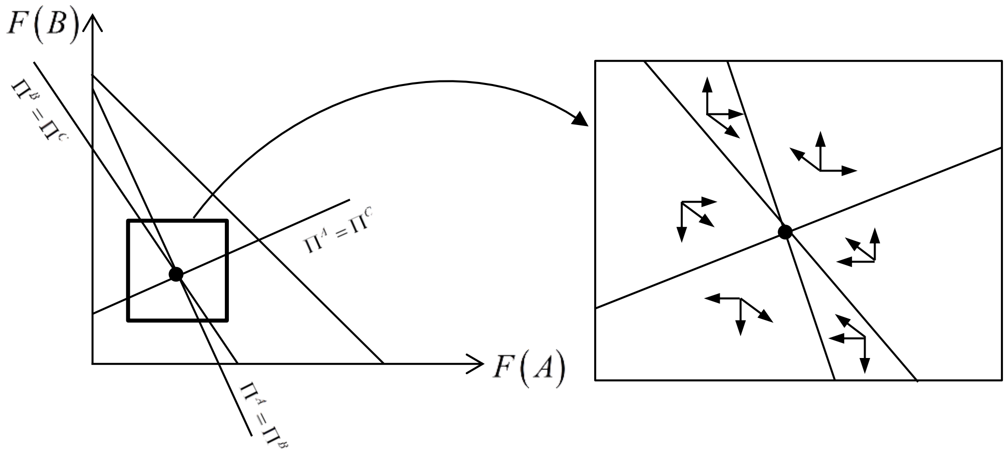
$$F(B) = 1-F(A)\frac{(a^2-a(b+3c-2d)+b(2c-d)+(c-d)d)}{(a-d)(a-b-c+d)}.$$

Given the summarizing conditions of this case $c > d > b > a$, $b - a < c - d$, $b < \frac{a+c}{2}$, it turns out that for the slopes of the three equation the following ordering holds: $\frac{\partial F(B)}{\partial F(A)}^{\Pi^A=\Pi^B} < \frac{\partial F(B)}{\partial F(A)}^{\Pi^B=\Pi^C} < -1 < 0 < \frac{\partial F(B)}{\partial F(A)}^{\Pi^A=\Pi^C}$, where

$$\frac{\partial F(B)}{\partial F(A)}^{\Pi^A=\Pi^B} = \frac{-2a^2 + b^2 + a(b+3c) + c(-c+d) - b(2c+d)}{2(a-b)(b-d)},$$

$$\frac{\partial F(B)}{\partial F(A)}^{\Pi^B=\Pi^C} = \frac{-a^2 + a(b+3c-2d) + b(-2c+d) + d(-c+d)}{(a-d)(a-b-c+d)},$$ and

$$\frac{\partial F(B)}{\partial F(A)}^{\Pi^A=\Pi^C} = \frac{a^2 - b^2 + (c-d)^2 - 2ad + 2bd}{a^2 + 2b^2 - a(3b+c-2d) - bd + (c-d)d}.$$ This gives rise to the following phase diagram.



As the diagram clearly indicates this equilibrium is unstable.

Proof of Proposition 4-8: The set of equilibrium payoffs can be found in the proof of Proposition 4-3. If the two Nash equilibria are not Pareto-ranked then I may w.l.o.g. assume that $a < d < D < A$ (see Table C-3). Two cases with respect to the thresholds for low types may be distinguished. I first consider $\theta^R_{(0,0),1} < \theta^R_{(1,1),1}$

1. $0 < \theta^R_{(0,0),1} < \theta^R_{(1,1),1} < 1$

Table C-6 depicts equilibrium payoffs in the various matches.

	$\theta < \theta^R_{(0,0),2}$	$\theta > \theta^R_{(0,0),2}$
$\theta < \theta^R_{(0,0),1}$	$\left(\frac{a+d}{2}, \frac{A+D}{2}\right)$	(d,D)
$\theta^R_{(0,0),1} < \theta < \theta^R_{(1,1),1}$	(d,D)	(d,D)
$\theta^R_{(1,1),1} < \theta$	$\left(\Pi_1^{mix}(\alpha_1,\alpha_2), \Pi_2^{mix}(\alpha_1,\alpha_2)\right)$	(b,B)

Table C-6: Equilibrium payoffs: $0 < \theta^R_{(0,0),1} < \theta^R_{(1,1),1} < 1$.

I will refer to individuals with $\theta < \theta^R_{(0,0),1}$, $\theta^R_{(0,0),1} < \theta < \theta^R_{(1,1),1}$, and $\theta^R_{(1,1),1} < \theta$ as A types, B types and C types respectively.

There can be no equilibrium with B types only as the more opportunistic A type would earn strictly higher profits as long as some high types are opportunistic. In an equilibrium with A and B types opportunistic high types would earn strictly higher payoffs. I will now consider the case of C types, who give rise to the play of a mixed equilibrium when matched with an opportunistic high type. I will show that $\Pi_2^{mix}(\alpha_1,\alpha_2)>B$, thus in such an equilibrium only opportunistic high types can be present.

Note that $\Pi_2^{mix}(\alpha_1,\alpha_2)>B \Leftrightarrow A+\alpha_1+\alpha_2 C+\alpha_1\alpha_2 D>(1+\alpha_1)(1+\alpha_2)B \Leftrightarrow A-B+\alpha_1\alpha_2(D-B)>\alpha_2(B-C)$

Consider first $B\le b$, then $D>B$ and hence $A-B+\alpha_1\alpha_2(D-B)>\alpha_2(B-C)\Leftarrow A-B>\alpha_2(B-C)$.

Note that $\dfrac{\partial\alpha_2}{\partial\theta_2}=\dfrac{A(c-d)+B(-c+2C+d-2D)-(a+b)(C-D)}{\left(D+C(-1+\theta 2)-(c-d+D)\theta_2\right)^2}$ if $C>c$. This derivative is negative if and only if the numerator is negative which can be written as $-(A-B)(d-c)+(D-C)(a+b-2B)$. This term is negative because $\theta^R_{(0,0),2}<1\Leftrightarrow a+b<2B$. If $C<c$ then $\dfrac{\partial\alpha_2}{\partial\theta_2}=\dfrac{-A(c-2C+d)+B(c+d-2D)-(a+b)(C-D)}{\left(D(-1+\theta_2)-(c+d)\theta_2+C(1+\theta_2)\right)^2}$. This derivative is negative if and only if the numerator is negative which can be written as $-(A-B)(d+c)+2AD-2BD+(D-C)(a+b)$. Note that:

$$
\begin{aligned}
&-(A-B)(d+c)+2AD-2BD+\underbrace{(D-C)}_{>0}(a+b)<0 \underset{\theta^R_{(0,0),2}<1\Leftrightarrow a+b<2B}{\Leftarrow}\\
&-(A-B)(d+c)+2AD-2BD+2B(D-C)<0\Leftrightarrow\\
&\underbrace{(A-B)}_{>0}\left(2C-(d+c)\right)<0
\end{aligned}
$$

This term is negative because $C<c<d$ implies $\left(2C-(d+c)\right)<0$. Thus $\dfrac{\partial\alpha_2}{\partial\theta_2}<0, b\ge B$.

Consider second $B>b$, then $D>B$ because $\theta^R_{(1,1),1}<1\Leftrightarrow 2\underbrace{(d-b)}_{>0}<D-B$. Hence still $\Pi_2^{mix}(\alpha_1,\alpha_2)>B\Leftarrow A-B>\alpha_2(B-C)$ holds. I show that also in this case $\dfrac{\partial\alpha_2}{\partial\theta_2}<0$. $\theta^R_{(0,0),2}<1\Leftrightarrow a+b<2B$. If $C<c$ then $\dfrac{\partial\alpha_2}{\partial\theta_2}=\dfrac{-A(c-2C+d)+B(c-2C+d)-(a-b)(C-D)}{\left(D(-1+\theta_2)-(c+d)\theta_2+C(1+\theta_2)\right)^2}$. This derivative is negative if and only if the numerator is negative which can be written as $-\underbrace{(A-B)}_{>0}\underbrace{(c+d-2C)}_{>0,\ C<c<d}+(a-b)\underbrace{(D-C)}_{>0}$. Note that: $\theta^R_{(0,0),2}<1\Leftrightarrow a-b<0$, hence $\dfrac{\partial\alpha_2}{\partial\theta_2}<0$.

If $C>c$ then $\dfrac{\partial\alpha_2}{\partial\theta_2}=\dfrac{(A-B)(c-d)-(a-b)(C-D)}{\left(D+C(-1+\theta_2)-(c-d+D)\theta_2\right)^2}$. This derivative is negative if and only if the numerator is negative which can be written as $-\underbrace{(A-B)(d-c)}_{>0}+(a-b)\underbrace{(D-C)}_{>0}$. Note that: $\theta^R_{(0,0),2}<1\Leftrightarrow a-b<0$, hence $\dfrac{\partial\alpha_2}{\partial\theta_2}<0$.

In summary, if $0<\theta^R_{(0,0),1},\theta^R_{(1,1),1}<1$ and one equilibrium is contestable for the high type, i.e. $0<\theta^R_{(0,0),2}<1$, then $\dfrac{\partial\alpha_2}{\partial\theta_2}<0$. Note that I did not make use of $\theta^R_{(0,0),1}<\theta^R_{(1,1),1}$. Hence the result also applies for the second case $0<\theta^R_{(1,1),1}<\theta^R_{(0,0),1}<1$ which will be considered next. Hence

$$\Pi_2^{mix}(\alpha_1,\alpha_2)>B \Leftarrow A-B>\alpha_2(B-C) \Leftarrow A-B>\alpha_2^{\max}(B-C) \overset{\frac{\partial\alpha_2}{\partial\theta_2}<0}{=} \alpha_2(\theta_2=0)(B-C)=\frac{A-B}{D-C}(B-C)$$

$$\Leftrightarrow D>B$$

Since the last inequality holds, the claim $\Pi_2^{mix}(\alpha_1,\alpha_2)>B$ is established.

2. $0<\theta^R_{(1,1),1}<\theta^R_{(0,0),1}<1$

In that case Table C-6 becomes:

	$\theta<\theta^R_{(0,0),2}$	$\theta>\theta^R_{(0,0),2}$
$\theta<\theta^R_{(0,0),1}$	$\left(\frac{a+d}{2},\frac{A+D}{2}\right)$	(d,D)
$\theta^R_{(0,0),1}<\theta<\theta^R_{(1,1),1}$	(a,A)	(b,B)
$\theta^R_{(1,1),1}<\theta$	$\left(\Pi_1^{mix}(\alpha_1,\alpha_2),\Pi_2^{mix}(\alpha_1,\alpha_2)\right)$	(b,B)

Table C-7: Equilibrium payoffs: $0<\theta^R_{(1,1),1}<\theta^R_{(0,0),1}<1$.

Since $\Pi_2^{mix}(\alpha_1,\alpha_2)>B$ also holds and since $A>B$ dominance of relative opportunistic players among high types is even strict. Thus, also in this case no inequality-averse individuals can be part of a stable equilibrium. QED

<u>Proof of Proposition 4-9:</u> The argument is given in the paper. QED