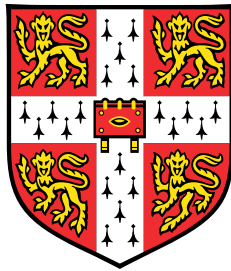


Three Essays on Economic Inequality



Gustavo Nicolas Paez Salamanca

Faculty of Economics
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

Selwyn College

November 2018

Abstract

Title: Three Essays on Economic Inequality

Author: Gustavo Nicolas Paez Salamanca

This PhD dissertation studies how market structures and economic incentives transform heterogeneity at agent levels into unequal economic outcomes.

The first chapter studies the economic incentives that lead a country to specialise its production in specific segments of a supply chain, and how these incentives transform heterogeneity at the productivity level into wage differences between countries. This chapter presents an innovative framework that incorporates production networks to the Ricardian trade model. It describes the price formation mechanism that occurs along supply chains and how it induces countries to focus on the production of specific goods. Moreover, the model highlights the role of the network structure in the determination of prices, and uses it to explain how changes in the productivity of a country have consequences in the production decisions and wages of the other countries that produce goods in the supply chain.

The second chapter studies the effects that the heterogeneity of income flows has over the implementation of collective agreements. Collective agreements are the primary mechanism by which communities cope with market failures. However, the lack of enforcement mechanisms generates coordination challenges. This chapter presents a theoretical framework that studies how inequality among individuals affects the participation incentives of the individuals and explains why agreements that balance the rent-seeking behaviour of wealthy individuals with the redistribution interests of the poor reduce the adverse effects of heterogeneity, and can even use it to create more robust agreements.

The third chapter studies heterogeneity at the level of academic journals. This chapter models the interaction between authors and journals as a platform market and uses this model to explain how general interest journals compete against field-specific journals. The model provides new insights into the way in which general interest journals link the different publication incentives of journals across fields. The theoretical results explain why general interest journals tend to attract higher quality publications and how changes in the publication capacity of a journal, or the volume of research in a field, can affect the quality of ideas published in both field-specific and general interest journals. Finally, this chapter applies the previous theoretical results to understand how the Top 5 journals in economics obtained their central role, and how their influence has changed between 1980 and the present.

I dedicate this work to God.

God allowed me to take the PhD challenge and gave me the tools I needed for it.

God provided me with the ability to speak math and to translate it to and from reality.

God enhanced my life with my parents, my brother, and all my family who are always there for me, and who always advise me and give me unconditional support for any decision I take.

God encouraged my spirit by filling it with friends, tireless companions, who gave me their energy and effort every time I was running out of my own, and who reminded me of the colour of my name when the path was blurry.

God illuminated my path, by letting me know excellent scholars, who taught me that, in an academic debate, the winners are those who are able to learn from the position of the other and use it to improve their own thoughts; no great researcher can be a one trick dog, and by learning from others, our own ideas flourish.

Thanks to God I could conclude this work, and I'm eager to see what is the next challenge.

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. Collaborative work is limited to Chapter 3, which is co-authored with Pr. Sanjeev Goyal (University of Cambridge), Dr. Marco van der Leij (University of Amsterdam), and Dr. Lorenzo Ductor (Middlesex University).

This work is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text.

Finally, this dissertation contains fewer than 60,000 words including appendices, bibliography, footnotes, tables and equations.

Gustavo Nicolas Paez Salamanca

November 2018

Acknowledgements

I would like to thank the people whose advice and time had a direct effect in this work.

Firstly, I would like to begin by giving thanks to the Gates Foundation for granting me with the resources, via the Gates Cambridge scholarship, that allowed me to dedicate myself entirely to the PhD. Beyond financial support, the Gates community has provided me with both emotional and intellectual support all these years. The community of students associated with the scholarship was always ready to help and contribute ideas in all forms. They are truly amazing people.

Secondly, I want to give special thanks for all the time that Zeina Hasna, Ruoyun Hui, Saba Sharma, Joanna Ferreira, and Jörg Kalbfuss invested in this dissertation. The text of this thesis is intelligible thanks to their amazing editing skills.

Thirdly, I thank the Faculty of Economics, both the administrative and the academic staff, as well as the academics that I met there over these years. In the administrative staff, I want to especially thank Silvana Dean for always being there, ready to solve any question or problem, so that the students didn't need to worry about anything admin-related. In the same vein, I also want to thank Craig Peacock and the IT team. Every minute they invested in solving our computer issues saved us hours of stress. Finally, I want to give thanks to Simon Frost and the team at the Marshall library. I was always really impressed by the speed at which they find any book I needed.

I am very grateful to the academic staff and academic visitors at the faculty, for always being ready to listen to my ideas and, based on their own expertise, share with me excellent advice on how to improve my work. During these years, I received valuable advice from Sir James Mirrlees, Alexey Onatskiy, Chryssi Giannitsarou, Donald Robertson, Melvyn Weeks, Meredith Crowley, Robert Rowthorn, Sriya Iyer, Vasco Carvalho, Kaivan Munshi, Dilip Mookherjee, Fernando Vega-Redondo, Francis Bloch, Kai Konrad, Marcin Dziubiński, Daniel Goya, John Spray, Juan Block, Leonnie Bauman, and Sihua Ding. Additionally, I want to give thanks to those academics who gave me friendly advice during the Network reading group, the Max Planck Workshop in Advances in Political Economy and Public

Economics, the Third Annual Conference on Network Science in Economics, and the International Sustainable Development Research Society where the first two chapters were presented. I'm sorry that I can't name all of them, but I did my best to incorporate all their kind and generous advice into this work.

Fourthly, I want to highlight my most sincere gratitude towards some academics that were key for this dissertation. First, I want to mention Matthew Elliot and Flavio Toxvaerd for their continuously positive and encouraging attitude towards each of the different chapters presented in this dissertation. Every time I talked to them, I got multiple ideas on how to improve my work, and how to challenge myself and my models to reach new limits. Next, I want to recognize the fantastic support that each chapter received from excellent academics who helped me to conceptualise the problems that I wanted to address, showed me ways to approach them, and most important, taught me how to learn from the process. Regarding this, I want to express my most profound gratitude towards Tiago Cavalcanti and Ha-Joon Chang for all their advice with the first chapter; Sir Partha Dasgupta and Shailaja Fennell for all their guidance with the second chapter; and to Marco van der Leij and Lorenzo Ductor for all their support with the third chapter.

Finally, I want to give special thanks to my supervisor, Prof. Sanjeev Goyal. During these years I learned a lot from him as a teacher, advisor, and co-author. From each of these perspectives, I learned about economic research as a diversion and as a profession. He was always aware of my progress, seeking to help me finishing the thesis in a proper manner. In particular, when my thoughts deviated and put my dissertation at risk, he dedicated an enormous amount of time to help me re-focus the ideas, and with his words of wisdom, he encouraged me to adjust the path towards the dissertation. It is beyond doubt that he played a central role in this work, and thanks to him and his encouragement, I was able to finish it within three years.

Table of Contents

List of Figures	xiii
List of Tables	xv
1 On Global Supply Chains and Inequality	1
1.1 Introduction	1
1.2 Literature Review	3
1.3 Supply Chain Model	6
1.3.1 General Environment	6
1.3.2 General Equilibrium	10
1.3.3 Acyclical Structures and Constant Productivities	13
1.3.4 Linked Economies and Cascade Effects	23
1.4 Conclusions	31
Appendices to Chapter 1	33
Appendix 1.A Math Notation	33
Appendix 1.B Proofs	34
2 On the Role of Unequal Treatment in Collective Agreements	49
2.1 Introduction	49
2.2 Literature Review	51
2.3 Collective Agreements	53
2.3.1 General Environment	53
2.3.2 Club Good Agreements	55
2.3.3 Common Pool Agreements	57
2.4 Efficiency vs. Robustness	59
2.4.1 On Efficiency	59

2.4.2	On Robustness	60
2.5	CGAs vs. CPAs	69
2.6	Discussion and Future Work	73
2.7	Conclusions	76
Appendices to Chapter 2		77
Appendix 2.A	Math Notation	77
Appendix 2.B	Proofs	77
3	On the Influence of Economic Journals	97
3.1	Introduction	97
3.2	Trends in Journal Influence	100
3.2.1	Data Sources	100
3.2.2	Journal Influence	102
3.3	Publication Platforms Model	110
3.3.1	General Environment	110
3.3.2	Equilibrium Analysis	114
3.4	Comparative Statics	123
3.4.1	Top 5 and Tier 2 Journals	123
3.4.2	Top 5 and Top Field Journals	130
3.5	Conclusions	133
Appendices to Chapter 3		135
Appendix 3.A	Journal List	135
Appendix 3.B	Robustness Analysis	137
3.B.1	Influence Using Article Influence Score	137
3.B.2	Capacity Effects on Influence	140
Appendix 3.C	Math Notation	141
Appendix 3.D	Proofs	141
References		163

List of Figures

1.2.1 Production network: fragmentation models	3
1.3.1 Production network, Example 1.1	10
1.3.2 Production network: definitions	15
1.3.3 Economic structure, Example 1.4	20
1.3.4 Changes in productivity, Example 1.4	22
1.3.5 Economic structure, Example 1.5	24
1.3.6 Changes in productivity, Example 1.5	26
1.3.7 Economic structure, Example 1.6	29
1.3.8 Changes in productivity, Example 1.6	31
2.4.1 Robust rules	64
2.4.2 Expected contribution	64
2.4.3 Minimum δ per individual	65
2.4.4 Total welfare and minimum δ for the community	66
2.5.1 Future trade-off	71
2.5.2 Expected utility	72
3.2.1 Influence gap	104
3.2.2 Influence gap of Top Field journals	105
3.2.3 Journals' citation network before 2000	107
3.2.4 Journals' citation network after 2000	108
3.3.1 Authors decisions	115
3.3.2 Thresholds and authors benefits	116
3.3.3 Influence and influence ratio	117
3.3.4 Changes in general interest journal capacity (κ_g)	120
3.3.5 Changes in field 2 size (n_2)	121
3.3.6 Changes in readership (α)	122

3.4.1 Articles per year	126
3.4.2 Citations per year	127
3.4.3 Evolution of articles and authors	128
3.4.4 Citation evolution for 1980 and 1990	129
3.B.1 Influence gap using article influence	138
3.B.2 Influence gap of Top Field journals using article influence	139

List of Tables

1.3.1 Basic equilibrium, Example 1.3	17
1.3.2 Technological equilibrium, Example 1.3	18
1.3.3 Equilibrium variables, Example 1.4	21
1.3.4 Profit per technique, Example 1.4	21
1.3.5 Equilibrium variables, Example 1.5	25
1.3.6 Profit per technique, Example 1.5	25
1.3.7 Equilibrium variables, Example 1.6	30
1.3.8 Profit per technique, Example 1.6	30
3.2.1 Top journal groups	101
3.2.2 Evolution of references	109
3.B.1 Stationarity tests	140
3.B.2 Cointegration tests	141

Chapter 1

On Global Supply Chains and Inequality

1.1 Introduction

Global supply chains are central to understanding production choices and income differences across countries. Over the last few decades, new technologies have reduced the transport and coordination costs of international trade, allowing firms to both obtain inputs from and allocate production to geographically distant regions. Due to supply chains, market interdependencies are increasing between regions over time. Therefore, the price formation mechanisms within a country depend on the prices at which other countries are willing to sell inputs as well as the prices at which other countries are willing to buy goods. The objective of this paper is to develop an original framework that explains the price formation mechanism by which countries decide which goods to produce in a given supply chain and use it to understand the emergence of wage inequality between countries.

In contrast to the rich empirical literature on global supply chains, theoretical explanations are still insufficient to understand the interactions between inequality and the network structure of a supply chain. On the one hand, most of the research associated with trade use models with very simplified supply chains; these models often reduce supply chains to certain factors, some intermediate goods and a final good that aggregates the previous goods. On the other hand, the literature that theoretically models production networks currently focuses on the spread of shocks across different sectors of the economy without including an international trade component. A few authors, such as Costinot et al. (2012) and Baldwin and Venables (2013), have bridged the gap between trade and production networks. However, these studies represent supply chains using two specific types of input-output matrices, and their policy implications are limited to these settings.

Moving beyond the previous models, this paper extends a Ricardian international trade framework to incorporate a wide range of production networks. It describes how countries with heterogeneous levels of overall productivity decide which goods to produce given the size of their respective labour forces. The model accurately describes the economic incentives that allocate countries to different parts of the supply chain and explains how the network structure determines the equilibrium prices and wages. Finally, it specifies how changes in a country's productivity induce income inequality by changing the equilibrium wages along the supply chain. The model simultaneously solves two theoretical puzzles that are at the intersection of trade and production structures for which previous models offered limited solutions.

First, it explains why countries with the lowest productivity levels usually produce raw materials, whereas countries with the highest productivity levels do not necessarily produce the final goods. Due to their linear structure, models such as the one developed by Costinot et al. (2012) are not able to explain this non-linear effect in the allocation of production. The model presented in this paper explains what model assumptions sort countries by productivity and how countries allocate themselves when these assumptions do not hold. In particular, it describes how location decisions depend on the revenue per worker of each good. Ergo, in cases where final goods are labour intensive, countries with high productivity prefer to produce intermediate goods that require less labour per unit of output.

Second, this paper explains why productivity shocks can affect countries that are downstream, but it also studies under which cases there are also upstream effects, and in which cases there are no effects on the wages or production decisions of other countries. Not all productivity shocks are equal, and understanding which shocks reduce inequality is crucial for international trade policies. This paper characterises productivity shock spillovers and provides new insights into the mechanisms by which shocks flow across the supply chains. In particular, it argues that shocks propagate if they affect the price ranges of the goods that the countries produce. Depending on the network structure, the changes in the productivity of a country can affect only the prices of downstream goods, or affect also prices of upstream goods. In the latter case, that shock will also have consequences over the production decisions and wages of upstream countries.

The paper proceeds as follows. Section 1.2 embeds this work in the broader literature on international trade and supply chains. A concise account of the onset of the relevant research is provided, and recurring principles and premises across models are described. The results and limitations of closely related works are also discussed to reveal the gap that this study attempts to fill. Section 1.3 presents the model. First, the basic setup is introduced, and the

necessary notation is established. Second, the properties of the competitive equilibrium are characterised, and the resulting income distributions across countries are discussed. Third, the implications of acyclic supply chains and constant within-country productivities across sectors are explored to elucidate the effects of idiosyncratic productivity shocks on the chain as a whole. Last, the assumptions required to render the model comparable to existing work are explained, and the plausibility and implications of said restrictions are discussed. Finally, Section 1.4 concludes.

1.2 Literature Review

As the globalisation of production has progressed over recent decades, several economists have recognised its relevance to international trade and built theoretical models to explain the offshoring and outsourcing decisions of firms (Antras and Helpman, 2004). Articles building on this framework frequently assume that the world consists of two generic areas, a Global North and a Global South. From the perspective of the North, the South possesses cheaper resources that are accessible via offshoring or outsourcing. While offshoring incurs transport and coordination costs (Grossman and Helpman, 2002; Grossman and Rossi-Hansberg, 2008), outsourcing reduces the benefits and poses coordination problems (Kohler, 2004).

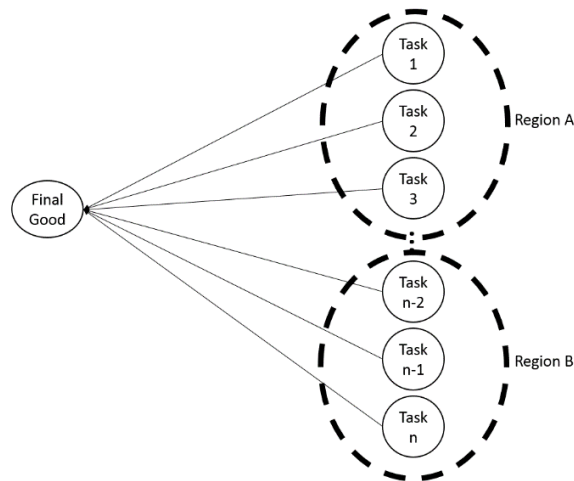


Fig. 1.2.1 Production network: fragmentation models

Based on Grossman and Rossi-Hansberg (2012), Figure 1.2.1 represents the basic production network underlying these models. The final good requires n tasks, and the firm decides which tasks to offshore/outsource and which to produce locally. Thus, this framework limits global supply chains to the analysis of two regions and a production structure in which all

tasks are performed in parallel. Baldwin and Venables (2013) extended previous works to allow for sequential production structures in which task i is the input of task $i + 1$. However, Markusen (2006) illustrated how, under the two-region approach, conclusions related to production decisions depend entirely on the characteristics associated with the regions *a priori* (e.g. only the North can produce the final product, or only the high-skilled sector has location externalities). Although such simplifications are useful to study firms' decisions, they provide a limited understanding of income differences at the country level and production allocation decisions, as most of them are already predetermined by the model. Moreover, the role of the supply chain in these models is minimalistic. Therefore, the effect of the structure of the supply chain on price formation along the chain cannot be studied. Finally, some of these models presume market failures and externalities that make them unable to distinguish between the effects of supply under perfect equilibrium and the effects of market power on trade.

Costinot et al. (2012) developed a different approach to describe the international fragmentation of production processes. Based on Kremer (1993) O-Ring theory, the authors model production as a sequential process that only generates value if it proceeds without failure. The model thereby assumes an exogenous probability of production failure for each country. The authors equate these probabilities to the productivity of the country: given the same amount of inputs, some countries can produce a smaller or a larger amount of high-quality output. Their model concluded that low-performance countries focus their labour on the early stages of production. As a consequence, their wages are systematically lower than those in high-performance countries. Regarding the price formation mechanism along the supply chain, the model infers that even without market imperfections, sequential production networks increase economic inequality in the world. Finally, productivity shocks only have downstream spillovers that reduce the income of countries with higher productivity and reduce the dispersion of wages.

Costinot et al. (2012) provide first-order insights into the relationship between production structures, productivity, and income. However, their model has two main limitations. First, their conclusions rely on a stylized linear production network, and therefore, most of the economic implications only apply to this specific production structure. Second, the model assumes a continuum of goods, which means that there is no single good produced by two countries at the same time. These constraints lead to predictions that are inconsistent with empirical observations. In particular, the implications of the model collide with two key empirical facts. On the one hand, productivity shocks can also substantially affect upstream producers. On the other hand, the most productive countries are not necessarily final good

producers, which contrasts with the main conclusion of their model. Due to the limitations of both the Costinot et al. (2012) framework and the foregoing fragmentation models, this paper takes a step back and proposes an alternative bridge between production networks and trade.

In the post-Second World War period, production networks were a central topic in economics. After Leontief (1953) formalised the interactions between intermediate goods using input-output matrices, several authors used the concept to measure how to promote development via production linkages. However, the popularity of the topic waned thereafter. Only recently, authors such as Acemoglu et al. (2012) have recovered this idea to demonstrate how shocks to particular firms can have first-order impacts on the overall economy, depending on the respective firm's location in the broader production network. Their model uses the input-output matrix of an economy to explain how different arrangements along the production structure induce a ripple effect that promulgates idiosyncratic shocks. This approach is more suitable for this paper because it couples an input-output structure with a perfect competition framework. This characteristic allows the effects of production structures on prices to be separated from the effects of market failures. However, Acemoglu et al. (2012) did not incorporate trade between countries. Instead, perfect factor mobility across different sectors was assumed. Therefore, perfect competition implies price equalisation; in particular, there is no variation in wages. Hence, to understand income inequality, it is necessary to restrict labour mobility between countries. Thus, this approach needs to be complemented with a trade theory that includes the limited mobility of factors.

Many trade theorists have considered it sufficient to study final goods to understand trade dynamics (Markusen, 2006). However, the main trade models are very sensitive to the inclusion of production structures. For instance, Ethier (1984) indicated that the principal results of Heckscher-Ohlin-Vanek (HOV) do not hold in economies with more than two goods and two factors. As another example, Samuelson (2001) demonstrated how the inclusion of intermediate inputs has significant effects on the production possibility frontier of the Ricardian model. Whereas the main classical results in the trade literature originate from these two models, their limitations do not allow them to explain supply chain mechanisms.

Recently, Shiozawa (2007) extended the Ricardian model to accommodate input-output relations and proved that specialisation occurs under general production networks. This adaptation offers a new path to connect trade and supply chains because it provides the mathematics required to treat both phenomena with similar tools. Therefore, this paper adopts Shiozawa (2007) methods to study price formation in supply chains and merges them with a network analysis that characterises the role of production structures in the dispersion of productivity shocks and the emergence of income inequality between different

countries. Using this approach, this paper overcomes the limitations of the previous studies and examines supply chain price formation for a broader range of production structures.

1.3 Supply Chain Model

The goal of this section is to develop a theoretic framework that integrates both general global supply chains and international trade. The first subsection thereby introduces the basic elements, such as the decision problem of individual countries in the global economy and the notational conventions used throughout the chapter. The second subsection solves for the general equilibrium of the model and discusses the relationship between specialisation and income distribution in a supply chain. The third subsection focuses on the price formation mechanism and the flow of productivity shocks in the economy. Lastly, in the fourth subsection, the paper's framework is compared to existing models by making explicit the assumptions that render them comparable and the model implications of these assumptions.

1.3.1 General Environment

The world economy consists of $c = \{1, 2, \dots, C\}$ indexed countries. Each country has access to a fixed number of workers. The positive vector $q \in \mathbb{R}^C$ represents the total number of workers in the economy, where q_c is the number of workers in country c . Finally, the world economy, in which all the countries participate, produces $m = \{1, 2, \dots, M\}$ indexed goods, where good 1 is known as the final good.

The workers, who are also consumers, maximise their utility solely by consuming the final good. They have a unit of labour that they inelastically supply to firms in exchange for their wages. Finally, workers are immobile between countries but perfectly mobile among industries within their own country.

A production function, or production technique, describes the way in which a firm transforms a bundle of inputs and a unit of labour into a certain quantity of output. Each country possesses a specific production technique for each of the M different goods. These techniques must satisfy the following conditions:¹

Assumption A1 *Each technique is Leontief and produces a single good.*

Assumption A2 *The final good is not an input of any other good. Furthermore, to produce one unit of the final good, all the other goods are required directly or indirectly.*

¹The first three conditions are adapted from Shiozawa (2007).

Assumption A3 *In autarky, each country is capable of producing the final good.*

Assumption A4 *Each country requires the same amount of inputs to transform one unit of labour into a given output. However, different countries obtain different quantities of output.*

Firms are profit maximisers and work under perfect competition. Each firm produces a single good based on the technique available to it in the country. Therefore, a firm that produces a particular good in a given country is wholly identified with the technique that produces that good in that country. This direct link allows the terms *firm* and *technique* to be used interchangeably in the context of this model.

Assumptions (A1) to (A4) specify the types of production techniques that this paper examines. For this reason, the remainder of this subsection formalises each assumption and explains their scope and limitations. After the four assumptions are presented, Example 1.1 illustrates how these building blocks represent supply chains.

Assumption A1:

The fact that the production techniques are Leontief is not a substantial restriction.² Moreover, it provides a clear intuition of a supply chain: firms work under protocols in which predetermined inputs are required to obtain a desired output. In this way, a technique is represented by a vector $\tau \in \mathbb{R}^M$, where $\tau_i > 0$ (≤ 0) denotes that τ produces (requires) $|\tau_i|$ units of good i as an output (input) for each unit of labour used. The fact that each technique produces only one good implies that there is a unique i such that $\tau_i > 0$. Finally, to identify each production technique, two auxiliary functions G and C are defined. If τ is the technique used by country c to produce m , then $C(\tau) = c$ and $G(\tau) = m$. When specificity is needed, $\tau(m, c)$ denotes the technique that c uses to manufacture m .

Given the previous notation, let Ξ be the set of all techniques available in the economy, and $\gamma \subseteq \Xi$. $A(\gamma) \in \mathbb{R}^{|\gamma| \times M}$ is defined as a matrix whose rows represent each of the techniques of γ . For this framework, matrix notation is useful to measure production quantities.³ For example, let x_i be the number of workers that the countries employ for each $A_i \in \gamma$. Then, $A(\gamma)^T x$ is a vector in which the i -th component is equivalent to the net production of good i .

Similarly, $I(\gamma) \in \mathbb{R}^{|\gamma| \times C}$ is defined such that $I(\gamma)_i = e(C(A(\gamma)_i))^T$, where $e(i)$ is the canonical vector that has a 1 at entry i and 0 in the rest of the entries. Hence, $I(\gamma)$ represents

²Baldwin and Robert-Nicoud (2014) showed that the Leontief model is equivalent to a first order Taylor approximation of a homothetic production function.

³This paper bases its notations and proofs on linear algebra. Appendix 1.A introduces the main notation used throughout the paper.

a matrix in which all the components are 0 except for the one associated with the country that produces each technique. In this case, $I(\gamma)^T x$ is a vector in which the i -th component is the total labour that country i employs across all the different techniques that it produces.

Assumption A2:

The core characteristic of a supply chain is that goods are used to produce goods. However, because the final good is a consumption goods, it is not used as an intermediate input of any other good. This assumption can incorporate economies with multiple final goods. As Trefler and Zhu (2010) proved, in a scenario in which all consumers have homothetic and identical taste and there is a given a set of prices, each person will invest fixed proportions of each of the final goods. Under this result, assumption (A2) is equivalent to claiming that there is a final good that requires fixed proportions of the consumption goods as inputs, where the labour required to produce it represents the utility for leisure.

Assumption A3:

Underlying perfect competition in trade is the idea that countries can be self-sufficient, but they can achieve better results with trade. For that reason, in autarky, each country can produce the final good. This implies that if a country has spare workers, it can produce positive quantities of each good without having a net loss of any good. Formally, this means that for each country c , there exists $\gamma \subseteq \Xi$ where $|\gamma| = M$ such that $C(\gamma) = \{c\}$ and $\exists x \geq \mathbb{O}_{C \times 1}$, for which $A(\gamma)^T x \gg \mathbb{O}_{M \times 1}$, where $\mathbb{O}$ denotes a 0-vector.⁴

Assumption A4:

Consistent with Acemoglu and Zilibotti (2001), countries acquire the same knowledge about the inputs needed for a given production. However, they differ in how much output they can produce. In other words, given the same inputs, countries obtain more or less output depending on how productive they are in that sector. Formally, for all goods m , and countries c, c' , it holds that $\tau(m, c)_{-m} - \tau(m, c')_{-m} = \mathbb{O}_{(M-1) \times 1}$.

Input-output matrices that define the production structures can be represented as a network. In this case, there is a link from good n to good m (denoted as $n \rightarrow m$) if there is a technique τ such that $\tau_m > 0 > \tau_n$. Moreover, the network is weighted via the amount of inputs required for one unit of labour. Assumption (A4) implies that these weights do not depend on the

⁴This condition is the primary building block for Arrow (1951) seminal paper on general equilibrium in input-output structures.

country that produces the technique. Furthermore, in the network representation, assumption (A2) implies that for all $i \in \{2, \dots, M\}$, there is a path between i and 1. Finally, it also implies that the node representing good 1 (i.e. the final good) has outdegree 0.

Example 1.1 clarifies the abovementioned notation and illustrates how they represent supply chains.

Example 1.1 (Production structures)

Consider an economy of four goods (1, 2, 3, and 4) and two countries (c and c'). Both countries are endowed with one unit of labour. The four production techniques available to country c form the set γ and are represented by the following matrices:

$$A(\gamma) = \begin{bmatrix} 1 & -2 & 0 & -2 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$I(\gamma) = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix}$$

In this example, c requires one unit of labour and two units of goods 2 and 4 to produce one unit of good 1, as represented in the first row of $A(\gamma)$. $I(\gamma)$ has two columns; the first corresponds to c , and the second corresponds to c' . Given that all techniques in γ belong to country c , $I(\gamma)$ only has 1's in the first column. Finally, each of the four techniques is Leontief and produces a single good, which is consistent with assumptions (A1) and (A2).

Country c can allocate its unit of labour to produce the final good. For example, it can schedule its workers in the following way: first, produce $4/9$ of good 3 and use it to produce $2/9$ of good 2; then produce $2/9$ of good 4 and use it, with the $2/9$ of good 2, to produce $1/9$ of good 1. In the end, all intermediate goods have net production 0, and the final good has positive production, which is consistent with assumption (A3). In matrix notation, this is equivalent to:

$$A(\gamma)^T x = \begin{bmatrix} 1 & -2 & 0 & -2 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}^T \begin{bmatrix} 1/9 \\ 2/9 \\ 4/9 \\ 2/9 \end{bmatrix} = \begin{bmatrix} 1/9 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$I(\gamma)^T x = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix}^T \begin{bmatrix} 1/9 \\ 2/9 \\ 4/9 \\ 2/9 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Now, contrast the production techniques of country c against the techniques of country c' described by the set γ' :

$$A(\gamma') = \begin{bmatrix} 1.5 & -2 & 0 & -2 \\ 0 & 1.2 & -2 & 0 \\ 0 & 0 & 0.8 & 0 \\ 0 & 0 & 0 & 0.4 \end{bmatrix}$$

This new matrix also fulfils assumptions (A1), (A2), and (A3). Moreover, any two techniques that produce the same good in countries c and c' only differ by the amount of output, as required by assumption (A4). In this case, c is more productive than c' in the production of $\{1, 2\}$, and c' is more productive than c in the production of $\{3, 4\}$.

Finally, Figure 1.3.1 represents the corresponding production network that is derived from the input-output matrix. It is important to highlight that, due to assumption (A4), its representation is independent of the country.

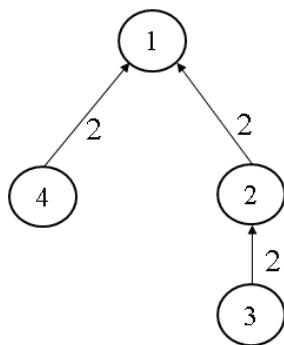


Fig. 1.3.1 Production network, Example 1.1

■

1.3.2 General Equilibrium

This subsection explores the price mechanism that emerges from competitive markets under general equilibrium. Let p, w denote the vector of prices and wages in the economy, where

p_m represents the price of good m , and w_c represents the wage of c .⁵ Hence, the profit vector of all the techniques is calculated via $A(\Xi)p - I(\Xi)w$. Using this notation, Definition 1.1 states the general equilibrium conditions of the economic problem:

Definition 1.1 (Equilibrium) *Under general equilibrium, there is a vector $x \in \mathbb{R}^{|\Xi|}$ (where x_i is the number of workers assigned to $A(\Xi)_i$), a vector of prices $p \in \mathbb{R}^M$ and a vector of wages $w \in \mathbb{R}^C$ such that:*

$$\text{Non-positive profits for any firm} \quad A(\Xi)p - I(\Xi)w \leq \mathbb{0}_{|\Xi| \times 1} \quad (1.1)$$

$$\text{Zero profit for active firms} \quad x_i > 0 \Rightarrow A(\Xi)_i p - A(\Xi)_i w = 0 \quad (1.2)$$

$$\text{Market clears for each intermediate good} \quad A(\Xi)_{.-1}^T x = \mathbb{0}_{(M-1) \times 1} \quad (1.3)$$

$$\text{Market clears for final good} \quad x^T A(\Xi) p = q^T w \quad (1.4)$$

$$\text{Labour market clears} \quad I(\Xi)^T x = q \quad (1.5)$$

The previous five conditions define the general equilibrium as a situation where markets clear, the firms that produce goods have no profits, and those that are not producing goods have non-positive profits if they decide to produce. Definition 1.1 displays the general equilibrium as a system of linear equations that Theorem 1.1 uses to translate the economic problem to a linear optimisation program.⁶

Theorem 1.1 (General equilibrium) *Define $R = \begin{bmatrix} -A(\Xi)_{.-1}^T \\ I(\Xi)^T \end{bmatrix}$, and consider the following linear optimisation program:*

$$\Omega = \max_{x \in \mathbb{R}_+^C} A(\Xi)_{.-1}^T x \quad \text{subject to} \quad Rx \leq \begin{bmatrix} \mathbb{0}_{(M-1) \times 1} \\ q \end{bmatrix} \quad \text{and} \quad x \geq \mathbb{0}_{|\Xi| \times 1}$$

The optimisation problem is feasible, and x is a solution of the linear program with p (shadow prices of the first $M - 1$ constraints) and w (shadow prices of the last C constraints) if and only if x , $\begin{bmatrix} 1 & p^T \end{bmatrix}^T$, and w define the workers allocation, prices, and wages of the general equilibrium.⁷

Theorem 1.1 provides a bijection between supply chains under perfect competition and linear programming. Given that the optimisation is feasible, this theorem implies that the

⁵This paper uses the words *income* and *wage* interchangeably.

⁶All the proofs for this chapter are presented in Appendix 1.B

⁷The "only if" requires prices to be normalized by $p_1 = 1$.

equilibrium exists. In addition to making explicit that the outcome of the economy is efficient, this theorem provides new analytic elements to study supply chains.

Central to this bijection is the fact that the solution to the linear program determines the set of basic constraints of the problem. These are the constraints that are binding in the optimum. On the other hand, restrictions are associated with the profit of a firm. Hence, the bijection implies that there is a set of techniques (hereafter known as basic techniques) that contains only those firms that are active in equilibrium (i.e. those that are represented by restrictions that bind). Furthermore, Proposition 1.1 shows how income emerges and is defined by these basic techniques.

Proposition 1.1 *In equilibrium,*

- (a) *The wage of a country is equivalent to the marginal product of its workers in the final good.*
- (b) *Conditional on the set of basic techniques, wage is independent of the population size.*

Theorem 1.1 and Proposition 1.1 present three characteristics of the economy. First, the equilibrium allocates workers in such a way that maximises the world's production of the final good given the labour constraints. Second, the income of the workers of a country equals its marginal productivity in the final good. This result extends the neoclassical idea that the price of a factor is equal to its marginal product to supply chains, where it is relevant to analyse the marginal product with respect to the final good rather than the marginal product of an intermediate input. Finally, Proposition 1.1 explicitly shows how the production structure, via the basic techniques, determines wages.

Theorem 1.2 provides a deeper understanding of the characteristics of the basic techniques and their relationship with income ordering.

Theorem 1.2 (Income ranking and interdependency) *Consider an equilibrium where γ is a set of basic techniques and $\{\tau(m, c), \tau(m', c')\} \subseteq \gamma$. Then,*

- (a) $\tau(m, c')_m > \tau(m, c)_m \Rightarrow w_{c'} > w_c$.
- (b) $\tau(m, c')_m > \tau(m, c)_m \Rightarrow 1 \leq \frac{p_{m'}(\tau(m', c')_{m'} - \tau(m', c)_{m'})}{p_m(\tau(m, c')_m - \tau(m, c)_m)}$.
- (c) $\tau(m, c') \in \gamma \Rightarrow w_{c'} = w_c + p_m(\tau(m, c')_m - \tau(m, c)_m)$.

The first part of Theorem 1.2 implies that, at equilibrium, the country with the i -th highest income is at least the i -th most productive country for each of the goods that it produces.

Hence, underlying the allocation of production is a mechanism of comparative advantage in which countries specialise in a given set of products. Along these lines, the second part of the theorem redefines comparative advantage for the context of supply chains. It states that countries allocate their workers to those sectors that have the highest revenue differences with respect to the other countries.⁸

The last part of the theorem describes the interdependency of workers' incomes across the production structure. The wage differences among countries that manufacture the same goods depend on the prices of these common products. For example, consider an equilibrium where γ is a set of basic techniques and $\{\tau(m, c), \tau(m, c')\} \subseteq \gamma$. If there is a shock in the economy where $\tau(m, c)$, $\tau(m, c')$ are still basic techniques, but p_m increased, Theorem 1.2 implies that the wage gap between c and c' countries widens.

In summary, this model formalises four key characteristics of supply chains. First, under perfect competition, supply chains generate equilibrium solutions that maximise the world's production of the final good. Second, the wage of a country is equivalent to the marginal contribution of its workers to the final good. Additionally, prices are independent of small demographic changes. Third, the ranking of productivity of the goods that a country manufactures is strongly related to its income level. Furthermore, a comparative advantage mechanism explains why countries do not produce the goods at which they are the most productive but prefer those with the highest revenue difference. Finally, goods whose production is shared by multiple countries directly determine the wage differences among them.

1.3.3 Acyclical Structures and Constant Productivities

The previous subsection discussed the most general characteristics of supply chains. This subsection incorporates two additional constraints that accentuate the price formation mechanism that takes place along supply chains. These assumptions are as follows:

Assumption A5 *The production network is a directed acyclical graph.*

Assumption A6 *The productivity of a country is constant between sectors.*

Assumption (A5) excludes production networks that contain cycles. Due to the framework presented in subsection 1.3.1, this paper considers that cycles can be omitted. There are two

⁸The determination of a generalized comparative advantage measure for production structures has been a challenge due to the multidimensionality of the comparison across countries and goods (Deardorff, 2005), still Theorem 1.2 defines it in such a way that is suitable for supply chains.

reasons for this. First, the goods described in this model are those in which the inputs are incorporated into the output. For example, buttons, yarn, and cloth are required to produce a jacket. All three inputs are embedded in the output. A sewing machine can certainly be used, but it is not in the same category as the previous inputs: while a button used in a jacket is already part of it and cannot be used again without destroying the final good, a sewing machine can be used to produce several jackets. More than being an input, the machine reflects a productivity improvement. On the one hand, the inputs required for a jacket are the same independent of whether it is sewed by hand or with a machine. On the other hand, having machines allows for more jackets to be produced within an hour of work and wastes less input. For these reasons, the machine is not an input, but it is part of the productivity of the technique. Based on this reasoning, when the concept of input focuses on incorporated goods, acyclic processes are a good representation of the reality of production. The other reason to validate assumption (A5) is that it allows upstream and downstream goods to be distinguished from each other in relation to the production structure. This view focuses the model in the relationship between the early and late production stages, and has been used by authors such as Baldwin and Venables (2013) and Ostrovsky (2008) to explain the underlying mechanisms of shocks that spread across the providers and consumers of an intermediate good.

Under assumption (A5), the network characteristics of the production structure become more relevant to the supply chain framework. In particular, three new concepts emerge. First, good m is a raw material if it does not require any input other than labour, i.e. it has 0 in-degree. Second, the production structure of m , denoted by $Prod(m)$, is defined by the subgraph formed by all the goods with a path leading to m , including m itself. Consistently, the set of nodes of $Prod(m)$ are denoted as $Nodes(Prod(m))$. From the previous definition, two production structures are equal if their subgraphs are isomorphic. Third, the raw distance of a good is defined as $rd(m) := \max\{d(m, n) : n \in Nodes(Prod(m))\}$ where $d(m, n)$ is the unweighted directed length of the path between m and n .⁹

Figure 1.3.2 illustrates the definitions above with an example.

While assumption (A5) focuses on the network structure, assumption (A6) focuses on the differences between countries. It simplifies the productivity of a country by assuming homogeneity between sectors. Recall that the indicator of productivity in good m is $\tau(m, c)_m$ because it represents, with one unit of labour and a given quantity of inputs, how much output the country can produce. In the seminal O-ring model (Kremer, 1993), this parameter represented the amount of good quality output. Later, Costinot et al. (2012) associated

⁹All the relevant paths are well defined as only members of $Prod(m)$ are considered.

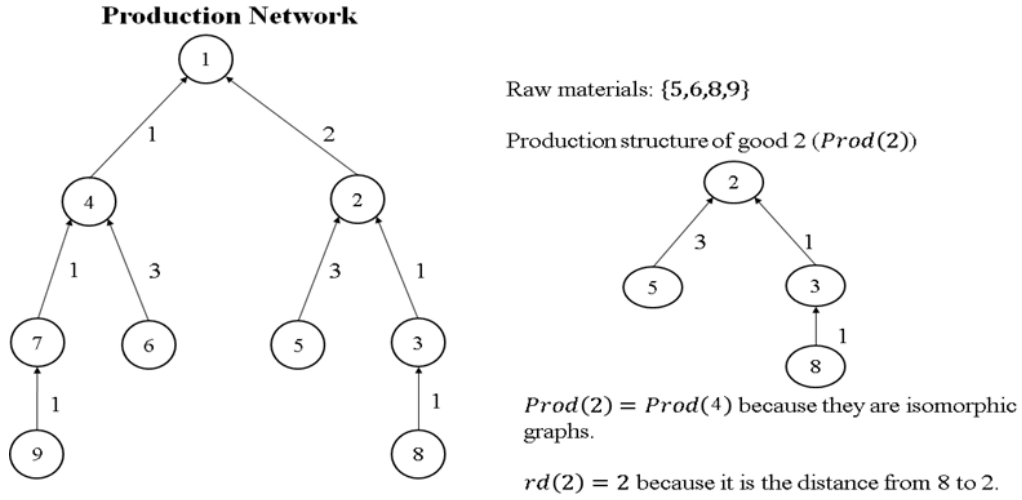


Fig. 1.3.2 Production network: definitions

an equivalent measure with the total factor productivity. These authors assumed that each country has only one productivity parameter to identify the effect of the supply chain without confounding its effect with the heterogeneity of productivity within the sectors of a country. This section follows this idea, which formally implies that for all c, m, m' , it holds that $\tau(m, c)_m = \tau(m', c)_{m'} = k_c$.

Without loss of generality, under assumption (A6), countries' indices are sorted so that productivity is decreasing in a country's index (i.e. country 1 is the most productive, while country C is the least productive). Theorem 1.2 implies that if $k_c < k_{c'}$, then $w_c > w_{c'}$. Thus, this assumption implicitly predetermines the wage to be sorted along with the productivity of countries. For this reason, in the remaining results, the research focus shifts from the wage formation and wage ranking of countries to the mechanisms that determine the production allocation and to how changes in productivity affect the income gaps between workers of different countries.

Assumption (A6) also allows normalization of the production techniques. Assume that under ideal production conditions (following Kremer's interpretation), one unit of labour produces one unit of output (i.e. a unit of merchandise is defined in terms of the amount obtained by one unit of labour under ideal technologies). Thus, for every country c , $1 \geq k_c$, where k_c can be interpreted as the percentage of output that the country achieves relative to the ideal situation. To better illustrate this normalization, Example 1.3 displays the implications of the previous assumptions and provides the key insights of the next set of results.

Example 1.2 (Jacket production; part 1)

Consider a supply chain for the manufacture of jackets: wool is needed to produce cloth, and cloth is required to create jackets. Under the best available technologies, one hour of labour shearing sheep can produce $x\text{kg}$ of wool. Hence, a unit of wool weighs $x\text{kg}$. Country c 's production protocols and know-how are such that its wool production technique can only obtain $0.95 \times x\text{kg}$ of wool from a unit of labour. Thus, $k_c = 0.95$.

The second stage uses spinning machines to transform wool into cloth. Under ideal conditions, one hour of labour transforms $10x\text{kg}$ of wool into $y\text{m}^2$ of cloth. Hence, the ideal production technique requires ten units of wool and one unit of labour to produce one unit ($y\text{m}^2$) of cloth. Assumption (A6) implies that on average, c obtains $0.95y\text{m}^2$. Finally, the technology of weaving is very labour intensive, and in one hour, a worker can only weave $\frac{y}{10}\text{m}^2$. In this case, the ideal production technique for one hour of labour produces 1 unit of jackets from 0.1 units of cloth.

If country c is in autarky and its income is normalized to 1, the price of a unit of wool is $\frac{1}{0.95} \approx 1.05$, the price of a unit of cloth is $\frac{1}{0.95}(10 \times \frac{1}{0.95} + 1) \approx 12.13$, and the price of a unit of jackets is $\frac{1}{0.95}(0.1 \times \frac{1}{0.95}(10 \times \frac{1}{0.95} + 1) + 1) \approx 2.33$. The price of a jacket is less than the price of cloth due to normalization: The amount of cloth needed to produce a jacket costs less than the jacket (otherwise, the technique would not be profitable), yet the amount of cloth produced by one unit of labour has a higher value than the amount of jackets produced with one unit of labour. Finally, if the world's technology improves (e.g. a very good sewing machine is invented) and now $2y\text{m}^2$ of cloth can be processed with one unit, the price of a new unit of jackets increases to $\frac{1}{0.95}(2 \times \frac{1}{0.95}(10 \times \frac{1}{0.95} + 1) + 1) \approx 26.60$.

■

Given assumptions (A1) to (A6), the model describes two fundamental trade mechanisms: specialisation patterns and structure-price equalisation.

Theorem 1.3 (Prices and structures) *Let p, w be the price and income vectors of an equilibrium. Let γ be a set of basic techniques and $\{\tau(m, c), \tau(m', c')\} \subseteq \gamma$. Then,*

- (a) *Structure-Price equalisation: If $\text{Prod}(m) = \text{Prod}(m')$, then $p_m = p_{m'}$.*
- (b) *Price specialisation: If $c' < c$, then $p_{m'} \geq p_m$.*

Theorem 1.3 reveals the main economic mechanisms that determine prices and specialisation in a supply chain. The first result is analogous to the factor price equalisation theorem

(Samuelson, 1948). Theorem 1.3 part (a) proves that isomorphic structures have the same prices, independent of the production country. The underlying mechanism comes from the same notion in Samuelson's theorem. The value of a good depends on its inputs and factors. Then, if two equivalent production structures incorporate the same amount of labour, the price must reflect this fact.

Theorem 1.3 part (b) explains that countries allocate themselves along price ranges. However, specialisation in price ranges is not equivalent to specialisation along the supply chain. For Costinot et al. (2012), countries sort themselves by their productivity along the supply chain. Thus, low-income countries always produce raw materials, and developed countries produce the final good. Nevertheless, there are many examples, such as the maquiladora industry in Mexico, where highly technological countries produce intermediate goods, and the final goods are assembled in countries with lower productivity levels (Bergin et al., 2009).

This model overcomes the previous inconsistency between theory and evidence by explaining that countries organise themselves along price ranges; independent if the goods are downstream or upstream. Hence, if the most productive countries realise that the volume of an intermediate good that they can generate per unit of work generates more revenue than the volume of the final good that can be produced per unit of work, then these countries will produce the former good.

Example 1.3 (Jacket production; part 2)

Consider the jacket production scenario in the previous example, but in this case, there are three countries, each with 100 workers and with respective productivity levels of 95%, 90%, and 85%. In the first scenario, the ideal technology can process ten units of wool for a unit of cloth and 1/10-th of a unit of cloth for a unit of jackets.

In this case, the equilibrium is given by Table 1.3.1.

Table 1.3.1 Basic equilibrium, Example 1.3

Goods	Prices	Workers assigned to each technique		
		Country 1 (95%)	Country 2 (90%)	Country 3 (85%)
Jacket	2.70	86	43	0
Cloth	13.64	14	0	0
Wool	1.18	0	57	100
Workers		100	100	100
Income		1.19	1.06	1.00

Consistent with Theorem 1.3, country 1 works on the products with the highest prices (cloth and jackets); country 2 works on the two products with the lowest prices (jackets and wool); and country 3 produces only the raw material, which has the lowest price.

In this case, the most productive country is not only producing the final good but also producing cloth. In contrast, country 2 produces both the final and the raw material but has no production in the intermediate stage. This ordering illustrates how, in a basic example, countries are not sorting themselves along the chain, as Costinot et al. (2012) would suggest. In contrast, and aligned with empirical observations, price specialisation exhibits patterns that can easily represent processes such as maquiladoras.

Consider a change in the world's technology that improves sewing machines, and now, two units of cloth can be processed within one hour of labour into a unit of jackets. In this case, the new equilibrium is represented in Table 1.3.2.

Table 1.3.2 Technological equilibrium, Example 1.3

Goods	Prices	Workers assigned to each technique		
		Country 1 (95%)	Country 2 (90%)	Country 3 (85%)
Jacket	29.72	11	0	0
Cloth	13.56	24	0	0
Wool	1.18	65	100	100
Workers		100	100	100
Income		1.12	1.06	1.00

The new technology changes the specialisation schedule as well as the prices. In particular, consistent with Theorem 1.3, country 1 now produces all the goods, while countries 2 and 3 move towards the production of the raw material. Hence, changes in offshoring and outsourcing decisions across a supply chain are also affected by new technologies.

■

As noted in Example 1.3, Theorem 1.3 part (b) provides new insights into two trends in global supply chains. First, it explains the mechanism by which developed countries (the most productive ones) focus on high added-value products (Timmer et al., 2014). As mentioned previously, comparative advantage incentivises countries to produce those goods in which they have the highest revenue differences. Due to assumption (A6), this is equivalent to claiming that highly productive countries focus their workforce on those products whose price is the highest per working unit. Second, the specialisation idea also helps to understand offshoring phenomena. Baldwin and Venables (2013) documented that countries decide to

offshore or locally produce goods depending on the costs and market failures. As detailed in Example 1.3, Theorem 1.3 part (b) complements the previous explanation by describing how a technological shock that changes the comparative advantage of the countries can reallocate production between countries.

The following corollaries typify the division of labour between countries and provide the first comparative static results when the productivity of one country changes.

Corollary 1.1 *Let p, w be the price and income vectors of an equilibrium characterised by the set of basic techniques γ .*

- (a) *Let $c > c'' > c'$ such that $p^T \tau(m, c) - w_c = p^T \tau(m, c') - w_{c'} = 0$. If there exists a good n such that $\tau(n, c'') \in \gamma$, then $p_n = p_m$.*
- (b) *Let $p^T \tau(m', c) - w_c = p^T \tau(m, c) - w_c = 0$, where $p_m \leq p_{m'}$. If there exists a good n such that $p_m \leq p_n \leq p_{m'}$, then $p^T \tau(n, c) - w_c = 0$.*
- (c) *Assume that c exists such that, for all $\tau(m, c) \in \gamma$, it holds that $p_m = \alpha > 0$. If an increase in k_c does not change the basic techniques in the equilibrium, then the shock only affects w_c .*

Corollary 1.2 *Let p, w be the price and income vectors of an equilibrium characterised by the set of basic techniques γ . Then, there exists $C + 1$ prices, $P_C, \leq P_{C-1} \leq, \dots, \leq P_0$ where $P_0 = 1$, and $P_C = 0$, such that $p^T \tau(m, c) - w_c = 0 \Leftrightarrow p_m \in [P_C, P_{C-1}]$.*

Corollary 1.1 and Corollary 1.2 develop a deeper understanding of Theorem 1.3 part (b) and describe how countries specialise in price ranges that overlap in at most one value. A direct implication is that the poorest country (C) always has to produce raw materials, as they always have the lowest prices in the chain. This result explains the empirical regularity described by authors such as Felipe et al. (2012), where low-income countries strongly focus on the production of raw materials and commodities.

In contrast to Costinot et al. (2012), Corollary 1.1 shows that cases exist where productivity shocks in particular countries do not induce a ripple effect. For this situation to occur, the countries must sell all their goods at the same price. Whereas this situation seems very unlikely, it can represent the situation for many countries that have specialised their economies in a unique sector.

Complementing the previous results, Proposition 1.2 describes changes in wages and prices where multiple equilibria exist.

Proposition 1.2 Consider an equilibrium with a unique distribution of labour x but multiple sets of basic techniques. Let γ, γ' be two sets of basic techniques, and p, w and p', w' its respective price and income vectors. Then:

- (a) If for all c it holds that $\{p_m : \tau'(m, c) \in \gamma'\} \subseteq \{p_n : \tau(n, c) \in \gamma\}$, then $p = p'$ and $w = w'$.
- (b) If there is a country c such that $\{p_m : \tau'(m, c) \in \gamma'\} \neq \{p_n : \tau(n, c) \in \gamma\}$, $\gamma \setminus \gamma' = \{\tau(m, c)\}$, and $\gamma' \setminus \gamma = \{\tau'(m', c')\}$, then there exists $\beta \in \mathbb{R}^C$, such that for every country, c , $w'_c = w_c + \beta_c$ and $\beta^T q = 0$.

Whereas Proposition 1.2 is a technical result from the simplex algorithm, it provides one of the core results of comparative statics in supply chains. It proves that if the portfolio of a country changes (reflected by a change in basic techniques), then it is possible that the wage vector does not change. This occurs when the price specialisation in both equilibria has the same range, as defined in Corollary 1.2. However, if this is not the case, the income distribution of all countries in the economy will change so that the average wage remains constant (mean-preserving spread). Example 1.4 shows that Proposition 1.2 together with Corollary 1.1 describe how shocks in production networks have non-linear effects on the income relations between countries. For some countries, there are complementarities, and for others, the increase in productivity affects the others as if they were substitutes.

Example 1.4 (General linear production)

Consider an economy where $C = 8$ and $M = 12$. The production network and the productivity parameters are described in Figure 1.3.3.

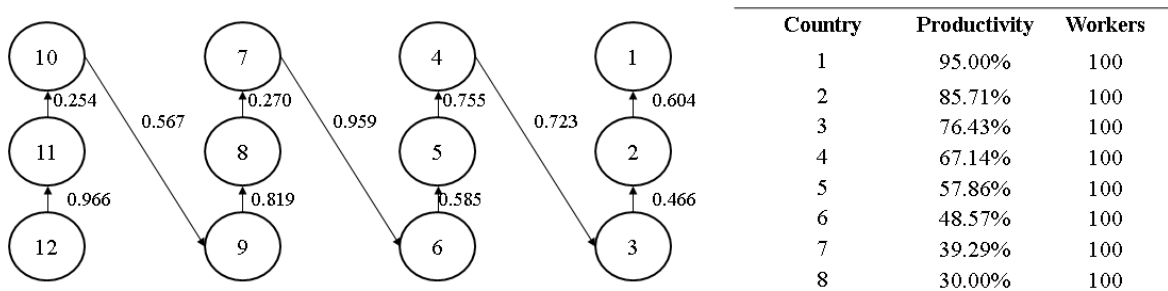


Fig. 1.3.3 Economic structure, Example 1.4

This economy consists of a linear production network. The productivity of countries is assigned values between 30% and 95%. Each country has 100 workers. The numbers in the

production network are based on ideal techniques. For example, to build one unit of good 1, under ideal conditions, 0.6 units of good 2 and one unit of labour are needed. Tables 1.3.3 and 1.3.4 present the equilibrium characteristics.

Table 1.3.3 Equilibrium variables, Example 1.4

M	P	Workers assigned to each technique							
		Countries							
		1	2	3	4	5	6	7	8
1	74.39	52	70	0	0	0	0	0	0
2	70.11	0	10	78	10	0	0	0	0
3	83.10	48	0	0	0	0	0	0	0
4	70.11	0	20	22	0	0	0	0	0
5	51.31	0	0	0	48	0	0	0	0
6	44.58	0	0	0	42	0	0	0	0
7	22.50	0	0	0	0	68	1	0	0
8	31.98	0	0	0	0	32	0	0	0
9	17.22	0	0	0	0	0	55	0	0
10	10.69	0	0	0	0	0	3	76	0
11	11.37	0	0	0	0	0	41	0	0
12	3.33	0	0	0	0	0	0	24	100
Workers		100	100	100	100	100	100	100	100
Income		28.29	21.38	14.87	8.36	4.39	2.30	1.31	1.00

Table 1.3.4 Profit per technique, Example 1.4

M	Countries							
	1	2	3	4	5	6	7	8
1	0.00	0.00	-0.40	-0.79	-3.73	-8.55	-14.47	-21.06
2	-0.40	0.00	0.00	0.00	-2.54	-6.96	-12.48	-18.68
3	0.00	-0.81	-2.02	-3.22	-6.95	-12.60	-19.32	-26.73
4	-0.40	0.00	0.00	0.00	-2.54	-6.96	-12.48	-18.68
5	-5.64	-3.49	-1.75	0.00	-0.80	-3.47	-7.24	-11.70
6	-7.51	-4.74	-2.37	0.00	-0.17	-2.22	-5.37	-9.20
7	-15.54	-10.72	-6.30	-1.88	0.00	0.00	-1.10	-2.88
8	-12.02	-8.08	-4.54	-1.00	0.00	-0.88	-2.86	-5.52
9	-17.99	-12.68	-7.77	-2.86	-0.49	0.00	-0.61	-1.90
10	-21.02	-15.11	-9.59	-4.07	-1.10	0.00	0.00	-0.68
11	-20.71	-14.86	-9.40	-3.95	-1.03	0.00	-0.06	-0.81
12	-25.12	-18.52	-12.32	-6.12	-2.46	-0.68	0.00	0.00

Following Theorem 1.3 part (b), countries organise their production by price and not by their sequence in the network. In particular, country 1 produces goods 1 and 3, but it does

not manufacture good 2. Furthermore, and consistent with Corollary 1.1, countries 2, 3, and 4 produce the same good, and the price range of country 3 is a single element. Additionally, as Corollary 1.2 implies, country 8 manufactures the raw material in equilibrium.

Now consider two scenarios. In scenario 1, the productivity of country 3 changes from 68% to 85%. In scenario 2, the productivity of country 4 changes from 58% to 76%. Figure 1.3.4 presents the corresponding results.

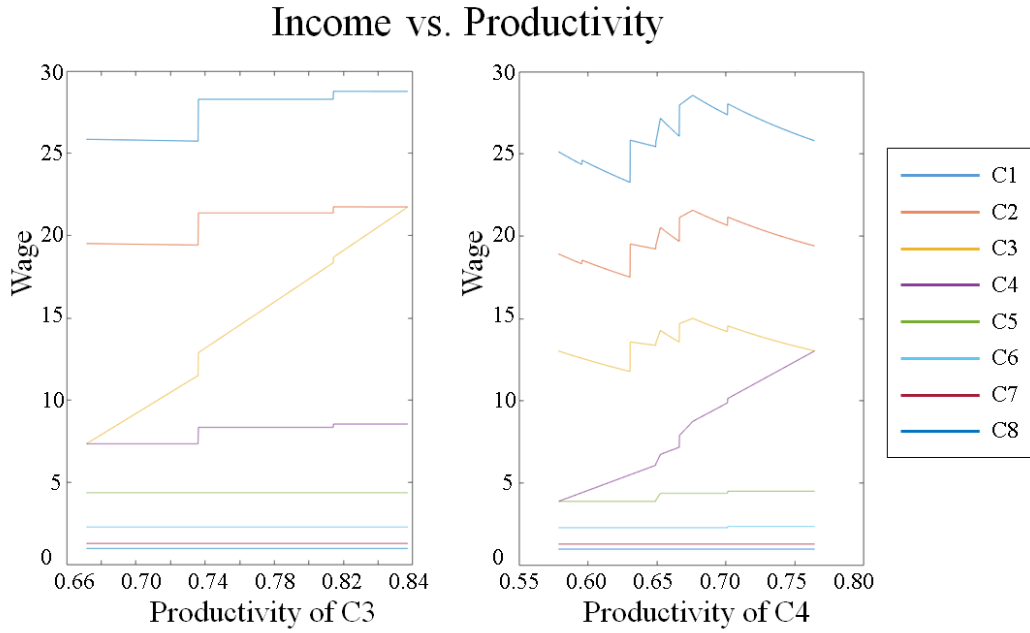


Fig. 1.3.4 Changes in productivity, Example 1.4

Scenario 1 illustrates Corollary 1.1. Given that country 3 produces only two goods at the same price, small changes in its productivity do not affect the income of other countries. As the shock increases, there is a discrete change in basic techniques, as Proposition 1.2 predicts.¹⁰ Scenario 2 illustrates that even in a simple scenario, a change in productivity can produce either a decrease in the spread of wages associated with lower inequality (around $k_c = 0.62$), an increase in the spread of wages associated with greater inequality (around $k_c = 0.74$), or some intermediate effect (around $k_c = 0.65$). Hence, the production structures induce non-linear effects between countries, as previously discussed. The last observation from this example highlights a particular characteristic of country 4. It is the only country with no product priced at P_4 as constructed in Corollary 1.2. Therefore, the countries can

¹⁰For comparison purposes, Figures 1.3.4, 1.3.6, and 1.3.8 standardize $w_C = 1$. However, the mean spread in Proposition 1.2 applies when the standardization is done with $p_1 = 1$. This is the reason why it cannot be visualized from the graph.

be partitioned into two groups: $A = \{1, 2, 3, 4\}$, $B = \{5, 6, 7, 8\}$ such that there is no good produced by the countries of both groups at the same time. As the next subsection describes, this characteristic determines whether it is possible to describe how shocks spread along the production network.

■

By adding assumptions (A5) and (A6), the current model clarifies the price formation and specialisation mechanisms. In particular, it highlights four of their characteristics. First, the price of a product depends solely on its production structure and is independent of the producer. Second, countries organise themselves to produce within ranges of output prices. Third, the least productive country always manufactures raw materials. However, the order of production along the supply chain depends on the available techniques. Fourth, productivity shocks in a country do not necessarily induce changes in the income of other countries. Nevertheless, infinitesimal shocks can produce discrete jumps in the income distribution when they induce a change in the basic techniques.

1.3.4 Linked Economies and Cascade Effects

This last subsection defines a special case of the model that makes it comparable with the current research that models trade in linear supply chains. This is a two-step comparison. First, it adds an assumption that is common to most of the theoretical models of trade and supply chains and explores its implications for price formation and inequality mechanisms. Second, it describes what additional constraints are used in the models that define supply chains as linear input-output structures and explains the implications of these restrictions.

A very common assumption in the literature that bridges trade and production networks is that labour intensity decreases from upstream to downstream. Assumption (A7) formalises this concept.

Assumption A7 *Increasing standardisation of processes:* $\forall i : \tau_i < 0 \Rightarrow |\tau_i| \geq 1$.

Whereas the assumption is restrictive, it has been used in seminal papers on trade, such as Grossman and Rossi-Hansberg (2008), to focus their results on production structures when complexity increases along the supply chain. The following results use assumptions (A1) to (A7) to characterise the market.

Theorem 1.4 (Price ordering) *Let p, w be the price and income vectors of an equilibrium characterised by the set of basic techniques γ .*

(a) $m \rightarrow m' \Rightarrow p_{m'} > p_m$.

(b) *If $\{\tau(m, c), \tau(m', c')\} \subseteq \gamma$ and there is a path from m to m' , then $c \geq c'$.*

Theorem 1.4 shows that under assumption (A7), countries organise themselves along the supply chain in what Costinot et al. (2012) referred to as vertical specialisation. Moreover, Theorem 1.4 proves that this result is not constrained to linear supply chains and holds in any acyclical production network where assumption (A7) holds. The next example illustrates the way in which countries allocate their production under this assumption.

Example 1.5 (Non-linear production)

Consider an economy where $C = 11$ and $M = 16$. The production network and the productivity parameter are described in Figure 1.3.5.

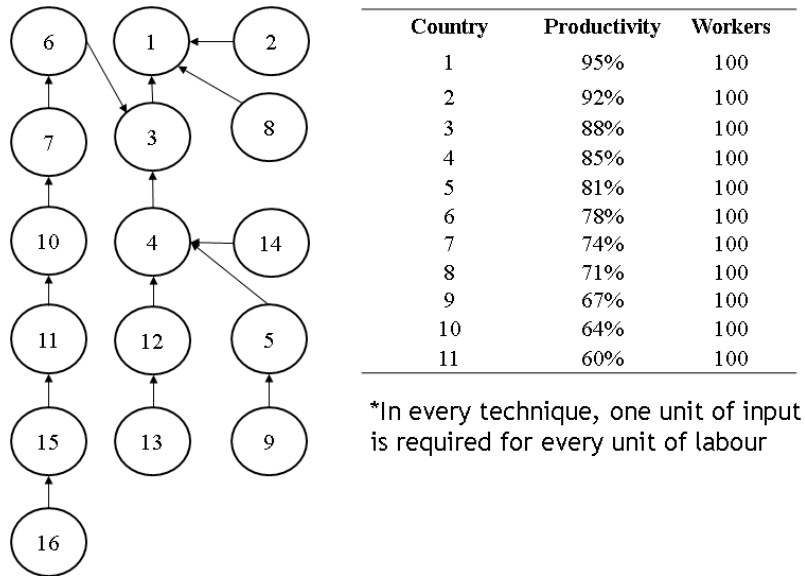


Fig. 1.3.5 Economic structure, Example 1.5

This economy has a rich production structure, including both sequential and parallel tasks. The productivity of the countries is evenly assigned between 60% and 95%. Additionally, each country has 100 workers. Here, every technique requires one unit of each input and one unit of labour. For example, to produce one unit of good 1 under ideal conditions, one unit of goods 2, 3, and 8 are needed along with one unit of labour. Tables 1.3.5 and 1.3.6 present the equilibrium characteristics.

Table 1.3.5 Equilibrium variables, Example 1.5

M	P	Workers assigned to each technique										
		Countries										
		1	2	3	4	5	6	7	8	9	10	11
1	41.48	42	0	0	0	0	0	0	0	0	0	0
2	1.66	0	0	0	0	0	14	0	4	0	45	0
3	33.27	45	0	0	0	0	0	0	0	0	0	0
4	12.57	0	39	11	0	0	0	0	0	0	0	0
5	3.80	0	0	0	0	61	0	0	0	0	0	0
6	16.21	13	36	0	0	0	0	0	0	0	0	0
7	12.57	0	25	28	0	0	0	0	0	0	0	0
8	1.66	0	0	0	0	0	0	0	0	63	0	0
9	1.66	0	0	0	0	0	0	34	0	0	0	59
10	9.25	0	0	61	0	0	0	0	0	0	0	0
11	6.32	0	0	0	72	0	0	0	0	0	0	0
12	3.80	0	0	0	28	32	0	0	0	0	0	0
13	1.66	0	0	0	0	0	0	0	0	0	55	41
14	1.66	0	0	0	0	0	0	66	0	0	0	0
15	3.80	0	0	0	0	7	86	0	0	0	0	0
16	1.66	0	0	0	0	0	0	0	96	37	0	0
Workers		100	100	100	100	100	100	100	100	100	100	100
Income		2.82	2.26	1.82	1.55	1.42	1.29	1.23	1.17	1.12	1.06	1.00

Table 1.3.6 Profit per technique, Example 1.5

M	Countries										
	1	2	3	4	5	6	7	8	9	10	11
1	0.00	-0.88	-1.88	-3.05	-4.36	-5.66	-7.05	-8.43	-9.81	-11.19	-12.57
2	-1.25	-0.74	-0.36	-0.15	-0.07	0.00	0.00	0.00	0.00	0.00	0.00
3	0.00	-0.59	-1.31	-2.19	-3.22	-4.24	-5.34	-5.34	-7.53	-8.63	-9.72
4	-0.13	0.00	0.00	-0.16	-0.47	-0.77	-1.15	-1.15	-1.91	-2.29	-2.67
5	-0.88	-0.45	-0.14	0.00	0.00	0.00	-0.07	-0.07	-0.22	-0.30	-0.37
6	0.00	0.00	-0.13	-0.42	-0.85	-1.28	-1.78	-1.78	-2.79	-3.30	-3.80
7	-0.13	0.00	0.00	-0.16	-0.47	-0.77	-1.15	-1.15	-1.91	-2.29	-2.67
8	-1.25	-0.74	-0.36	-0.15	-0.07	0.00	0.00	0.00	0.00	0.00	0.00
9	-1.25	-0.74	-0.36	-0.15	-0.07	0.00	0.00	0.00	0.00	0.00	0.00
10	-0.36	-0.12	0.00	-0.05	-0.24	-0.43	-0.69	-0.69	-1.22	-1.48	-1.74
11	-0.61	-0.27	-0.05	0.00	-0.09	-0.18	-0.34	-0.34	-0.66	-0.82	-0.98
12	-0.88	-0.45	-0.14	0.00	0.00	0.00	-0.07	-0.07	-0.22	-0.30	-0.37
13	-1.25	-0.74	-0.36	-0.15	-0.07	0.00	0.00	0.00	0.00	0.00	0.00
14	-1.25	-0.74	-0.36	-0.15	-0.07	0.00	0.00	0.00	0.00	0.00	0.00
15	-0.88	-0.45	-0.14	0.00	0.00	0.00	-0.07	-0.15	-0.22	-0.30	-0.37
16	-1.25	-0.74	-0.36	-0.15	-0.07	0.00	0.00	0.00	0.00	0.00	0.00

According to Theorem 1.3, goods 5, 12, and 15 have the same production structure, ergo, the same price at equilibrium. Additionally, according to Theorem 1.4, in each path from a raw material to the final good, all the participating countries are organised by their productivity, e.g. $\tau(9, 11)$, $\tau(5, 5)$, $\tau(4, 2)$, $\tau(9, 1)$, and $\tau(1, 1)$.

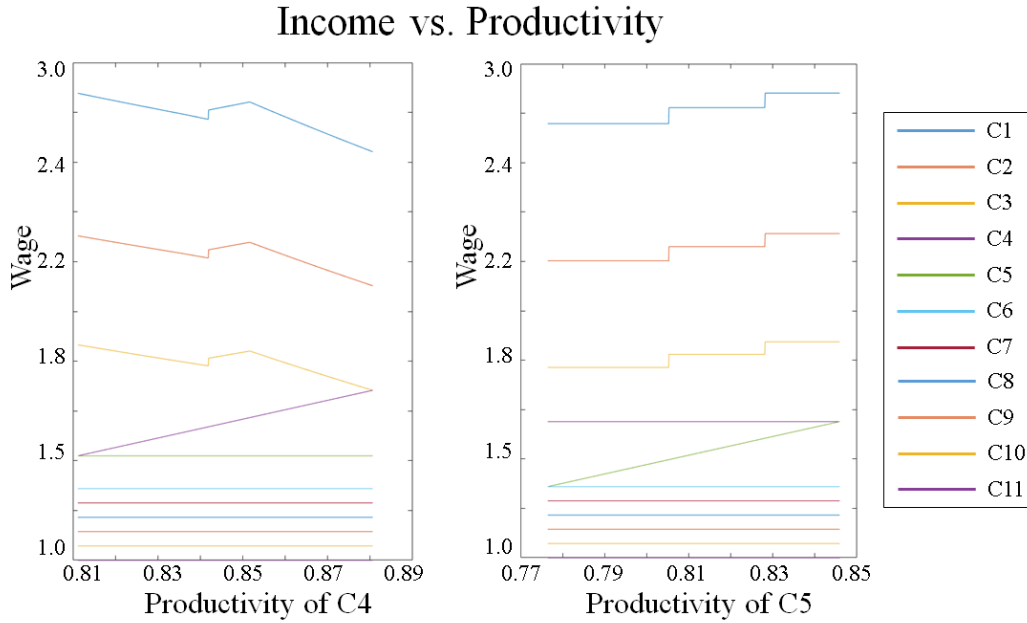


Fig. 1.3.6 Changes in productivity, Example 1.5

Consider two scenarios. In scenario 1, the productivity of country 4 changes from 81% to 88%. Meanwhile, in scenario 2, the productivity of country 5 changes from 78% to 85%. Figure 1.3.6 presents the results.

Scenario 2 reflects Corollary 1.1. Given that country 5 only produces goods 5 and 12, which have the same price, small shocks in its productivity do not affect the income of other countries. Then, as the productivity change increases, a change in the basic set of techniques arises, represented as discrete jumps. In contrast, scenario 1 shows that even under assumption (A7), a change in productivity can decrease (around $k_c = 0.8.3$) or increase (around $k_c = 0.85$) the spread of income. This highlights the emergence of income inequality in response to changes in productivity are structure-dependent. Furthermore, as in the previous example, no product is priced at P_4 . This means that the countries in this economy can be partitioned into two groups, $A = \{1, 2, 3, 4\}$, $B = \{5, 6, 7, 8, 9, 10\}$, where there is no common production between the groups.

■

Examples 1.4 and 1.5 suggest that non-linear spillovers are related to the fact that countries can be partitioned into two groups. To better understand this role, Definition 1.2 formalises this concept, and the rest of this section explains how this condition determines that positive productivity shocks generate only downstream spillovers and reduce the income spread between countries.

Definition 1.2 (Linked economy) *Given a set of basic techniques, an equilibrium generates a linked economy if p, w satisfy that for each country c , a good m exists, and a country $c' \neq c$ such that $p^T \tau(m, c') - w_{c'} = p^T \tau(m, c) - w_c = 0$.*

As the previous examples showed, linked economies are not the general case. However, under assumption (A7), there is an empirical motivation. Given that the countries are vertically organised along any path of the supply chain, when a country improves its productivity, it stops producing some inputs and focuses more on goods with higher value. However, the country is still capable of producing the discontinued inputs competitively, and sometimes residual production continues. Hence, for that good, both the new and the former producers will have 0 profit, which by construction is a linked economy. This characteristic coincides with the empirical observations of Timmer et al. (2014) about production reallocation between countries, and therefore, it is relevant to study them as a special case.

Corollary 1.3 *Consider an economy that satisfies assumptions (A1) to (A7) in which $\forall m < M$, $m + 1$ and labour are the only inputs for the production of m . Then, in equilibrium, any basic set of techniques describes a linked economy.*

Corollary 1.3 shows that the discrete version of the model presented by Costinot et al. (2012) is a linked economy. Hence, the next theorem shows how the price formation mechanisms described by these authors can be extended to any other linked economies.

Definition 1.3 (Price-adjusting algorithm) *Let γ be a basic set of techniques, $w_C = \alpha > 0$, and $s = C$.*

1. $\forall m$ such that $\tau(m, s) \in \gamma$, and $\forall n$ such that $n \rightarrow m$ and p_n has been defined, set

$$p_m = \frac{1}{k_s} (w_s + \sum_{n \rightarrow m} |A(\tau(m, s))_n| p_n)$$

Note: If m is a raw material, $p_m = \frac{1}{k_s} w_s$.

2. Iterate step 1 until $\forall m$ such that $\tau(m, s) \in \gamma$, p_m has been defined.

3. Let $P_s = \max\{p_m : \tau(m, s) \in \gamma\}$, $w_{s-1} = w_s + P_s(k_{s-1} - k_s)$.
4. Iterate from step 1 replacing s for $s - 1$.
5. Iterate step 4 until $s = 1$.

Theorem 1.5 Consider an economy that satisfies assumptions (A1) to (A7).

- (a) The price-adjusting algorithm is well defined.
- (b) If, in equilibrium, the set of basic techniques γ defines a linked economy, then the price-adjusting algorithm determines the prices and wages of the economy. In particular, P_c corresponds with the notation of Corollary 1.2.

The price-adjusting algorithm is a cornerstone in linked economies. Theorem 1.5 presents it as a constructive algorithm to find the prices and income vectors of an economy. It implicitly proves that productivity shocks in a country will only affect downstream countries and products.¹¹ Therefore, similar to the results of Acemoglu et al. (2012) and Costinot et al. (2012), in linked economies, productivity shocks flow downstream.

Proposition 1.3 Consider a linked economy defined by the basic set of techniques γ . Consider an increase in the productivity of country $c \neq C$ such that $k'_c = \beta k_c$ in which the basic techniques are the same in both equilibria. Let, p, w and p', w' be the variables associated with the equilibrium before and after the change, where both are normalized to $w_C = 1$. Then:

- (a) If there are m, n such that $\{\tau(m, c), \tau(n, c)\} \subseteq \gamma$ and $p_m \neq p_n$, then

- i $w'_{c'} = w_{c'}$ if $c' > c$.
- ii $w'_c = w_c + (1 - \beta)k_c$.
- iii $c' < c'' < c \Rightarrow w'_{c'} < w_{c'}$ and $|w_{c'} < w'_{c'}| > |w'_{c''} < w_{c''}|$.

- (b) If the productivity shock occurs in country C , $c' < c'' < C$, then $w'_{c'} < w_{c'}$ and $|w'_{c'} < w_{c'}| > |w'_{c''} < w_{c''}|$.

Complementing Theorem 1.5, Proposition 1.3 details how productivity shocks spill over the chain. A positive productivity shock in a country that produces goods with at least two different prices (otherwise Corollary 1.1 applies) will reduce income inequality (measured

¹¹For the algorithm, the equilibrium is standardized with $w_C = \alpha$ instead of $p_1 = 1$.

by the range of the income distribution) as long as the basic techniques do not change. However, if the techniques change, Corollary 1.1 and Theorem 1.5 describe how this induces a mean-preserving spread in which the income of all the countries that are upstream of the first basic technique to change moves in the same direction and proportion.

Example 1.6 (Simplified linear production)

Consider a linear production where for each unit of output, only one unit of input is needed. This is shown in Figure 1.3.7.

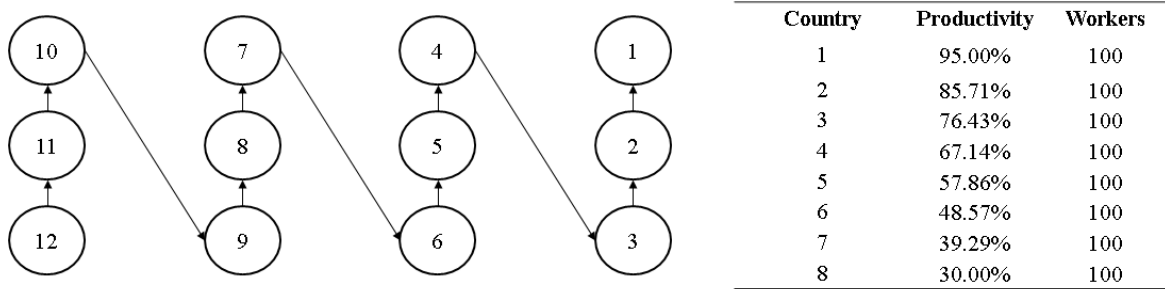


Fig. 1.3.7 Economic structure, Example 1.6

From Corollary 1.3, this equilibrium is always described by a linked economy. Tables 1.3.7 and 1.3.8 specify the equilibrium of this economy.

Consider two scenarios. In scenario 1, the productivity of country 2 changes from 68% to 85%. In scenario 2, the productivity of country 4 changes from 58% to 76%. Figure 1.3.8 presents the results.

Scenario 1 and scenario 2 are consistent with Proposition 1.3. Because the economy is linked, productivity changes that do not change the set of basic techniques do not reduce the range of income among countries. However, as Corollary 1.1 states, a shock that changes the basic techniques, such as the one that occurs in scenario 1 (k_2 around 0.8), can also increase the range of income among countries. In particular, this last observation demonstrates that considering a discrete framework provides cases in which the effects of a productivity change in inequality are qualitatively the opposite of those obtained in a continuous framework.

■

To summarise this last subsection, when the model includes assumption (A7), it restricts the price formation mechanism and determines two economic patterns. First, countries organise themselves so that more productive countries produce downstream products. Second, an increase in the productivity of any country other than country 1 reduces the range of wages (i.e. reduces the inequality) among countries.

Table 1.3.7 Equilibrium variables, Example 1.6

M	P	Workers assigned to each technique							
		Countries							
		1	2	3	4	5	6	7	8
1	210.28	20	0	0	0	0	0	0	0
2	181.64	21	0	0	0	0	0	0	0
3	154.44	22	0	0	0	0	0	0	0
4	128.59	23	0	0	0	0	0	0	0
5	104.04	14	11	0	0	0	0	0	0
6	80.72	0	30	0	0	0	0	0	0
7	60.73	0	35	0	0	0	0	0	0
8	43.59	0	24	18	0	0	0	0	0
9	28.90	0	0	55	0	0	0	0	0
10	17.67	0	0	27	52	0	0	0	0
11	9.09	0	0	0	48	81	0	0	0
12	3.33	0	0	0	0	19	100	100	100
Workers		100	100	100	100	100	100	100	100
Income		18.12	8.46	4.41	2.77	1.93	1.62	1.31	1.00

Table 1.3.8 Profit per technique, Example 1.6

M	Countries							
	1	2	3	4	5	6	7	8
1	0.00	-9.86	-25.34	-43.23	-61.91	-81.12	-100.34	-119.55
2	0.00	-7.21	-20.02	-35.25	-51.27	-67.83	-84.39	-100.94
3	0.00	-4.68	-14.97	-27.67	-41.17	-55.20	-69.23	-83.26
4	0.00	-2.28	-10.17	-20.47	-31.57	-43.20	-54.83	-66.46
5	0.00	0.00	-5.61	-13.63	-22.45	-31.80	-41.15	-50.50
6	-2.17	0.00	-3.45	-9.30	-15.95	-23.14	-30.32	-37.51
7	-4.02	0.00	-1.59	-5.59	-10.38	-15.71	-21.04	-26.37
8	-5.61	0.00	0.00	-2.41	-5.61	-9.35	-13.09	-16.82
9	-8.34	-1.36	0.00	-1.04	-2.88	-5.26	-7.63	-10.00
10	-10.43	-2.41	0.00	0.00	-0.80	-2.13	-3.46	-4.79
11	-12.82	0.00	-0.80	0.00	0.00	-0.54	-1.07	-1.61
12	-14.96	-5.60	-1.87	-0.54	0.00	0.00	0.00	0.00

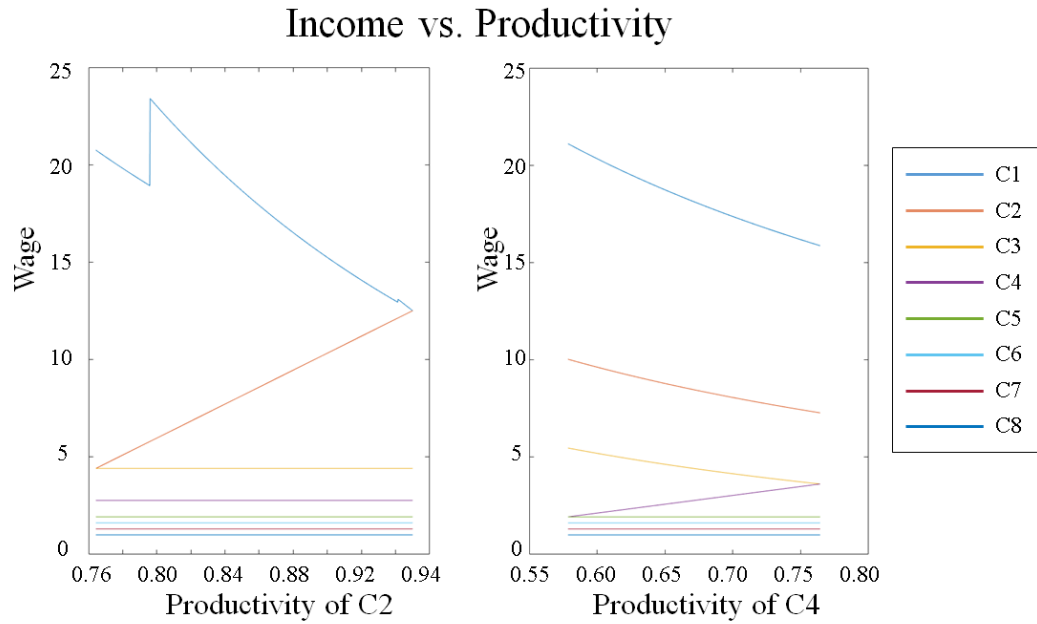


Fig. 1.3.8 Changes in productivity, Example 1.6

1.4 Conclusions

Supply chains are a central element in today's economy. However, theoretical research that bridges production networks and trade has been very limited. This paper takes a step back from the current modelling strategies, and by extending a Ricardian framework, it offers a new way to understand the interactions between these two fields.

This model relies on six assumptions that recreate an environment that is consistent with reality and that highlights the main price formation mechanism in a supply chain. Moreover, this paper shows that the current modelling strategies in the literature can be viewed as instantiations of this model.

Regarding the main outcomes of the model, it described the mechanisms that determine the allocation of countries along supply chains and how the production structure translates productivity differences into inequality at the income level of the workers in each country. Furthermore, it explains how changes in the productivity of a country affect all the other countries via production linkages and highlights the cases in which these shocks can increase and decrease inequality. Finally, as the model is set up under a perfect competition paradigm, it contributes to disentangling the effects of market dynamics and power relations in economies where production is ruled by global supply chains.

Appendix to Chapter 1

Appendix 1.A Math Notation

This section defines the technical notation that the paper uses in the main text and in the proofs. The notation used in the proofs of Section 1.B, that is not mentioned in this section, was explicitly defined in the main text.

Vectors

Let $x, y \in \mathbb{R}^n$:

- All vectors are represented as column vectors.
- x_i is the i -th component x .
- x_{-i} is a vector of $\mathbb{R}^{n-1}$ obtained by removing the i -th component x .
- $x \ll y \Leftrightarrow \forall i : x_i < y_i$.
- $x < y \Leftrightarrow (\forall i : x_i \leq y_i \wedge \exists j : x_j < y_j)$.
- $x \leq y \Leftrightarrow (x < y \vee x = y)$.
- $e(i)$ is the i -th canonical vector. I.e. $e(i)_i = 1$ and $\forall j \neq i, e(i)_j = 0$.
- $\mathbb{1}$ is a vector with all its components equal to 1.
- $\mathbb{0}$ is a vector with all its components equal to 0.

Matrices

Let $A \in \mathbb{R}^{n \times m}$:

- A^T is A transposed.

- A_{ij} is the element at row i and column j , $A_{i\cdot}$ is the i -th row of a matrix, and $A_{\cdot j}$ is the j -th column.
- $A_{-i\cdot}$ is a matrix in $\mathbb{R}^{(n-1) \times m}$ formed by removing the i -th row of A . In a similar way, $A_{\cdot -j}$ is defined as $(A_{-j\cdot}^T)^T$.
- $\mathcal{I}$ denotes an identity matrix (of the appropriate dimensions).
- Let $B \in \mathbb{R}^{n \times k}$, then $C = \begin{bmatrix} A & B \end{bmatrix} \in \mathbb{R}^{n \times (m+k)}$ denotes a matrix that has $m+k$ columns, the first m coincide with the columns of A , while the last k coincide with the columns of B . Analogous, if $B \in \mathbb{R}^{k \times m}$, then $\begin{bmatrix} A \\ B \end{bmatrix} \in \mathbb{R}^{(n+k) \times m}$ is defined as $\begin{bmatrix} A^T & B^T \end{bmatrix}^T$.

Appendix 1.B Proofs

Lemmas:

Lemma 1.1 *Let Ξ be the set of production techniques in the economy. Define the production set as $P(q) := \{y | \exists x \in \mathbb{R}^{|\Xi|} : x \geq \mathbb{O}_{|\Xi| \times 1} \wedge y = A(\Xi)^T x \wedge q \geq I(\Xi)^T x\}$ and the efficient frontier as $E(q) := \{y \in P(q) | \neg \exists y' \in P(q) : y' > y\}$.*

Properties of $P(q)$:

(a) *Strict superset of $\{\mathbb{O}_{M \times 1}\}$.*

(b) *Convex.*

(c) *Compact.*

Properties of $E(q)$:

(d) *If $y \in E(q)$, then there exists $x \geq \mathbb{O}_{|\Xi| \times 1}$ such that $A(\Xi)^T x = y$ and $I(\Xi)^T x = q$.*

(e) *$\exists y \in E(q)$ such that $\forall x \in E(q), y_1 \geq x_1$ and $y_{-1} = \mathbb{O}_{(M-1) \times 1}$.*

Proof:

Part a:

Define the set of techniques available in country c as $\xi_c := \{\tau(1, c), \tau(2, c), \dots, \tau(M, c)\}$; by definition, $\bigcup_{c=1}^C \xi_c = \Xi$. Assumption (A3) implies that for every country c , there exists $x_c > \mathbb{O}_{M \times 1}$ such that $A(\xi_c)^T x_c \gg \mathbb{O}_{C \times 1}$.

Based on the previous observation, define $x^T := \begin{bmatrix} \frac{q_1}{x_1^T \mathbb{I}} x_1^T & \frac{q_2}{x_2^T \mathbb{I}} x_2^T & \dots & \frac{q_C}{x_C^T \mathbb{I}} x_C^T \end{bmatrix}$, and arrange the rows of $A(\Xi)$ such that:

$$A(\Xi) = \begin{bmatrix} A(\xi_1) \\ \vdots \\ A(\xi_C) \end{bmatrix}$$

Hence, x satisfies the following properties:

$$x \geq \mathbb{O}_{|\Xi| \times 1}$$

$$I(\Xi)^T x = q$$

$$A(\Xi)^T x \gg \mathbb{O}_{M \times 1}$$

Therefore,

$$A(\Xi)^T x \in P(q)$$

Finally, $\mathbb{O}_{M \times 1} \in P(q)$ because $x = \mathbb{O}_{|\Xi| \times 1}$ also satisfies the conditions of the set. Thus, $\mathbb{O}_{M \times 1}$ is a strict subset of $P(q)$.

Part b:

If $y, y' \in P(q)$, then there exist $x, x' \geq \mathbb{O}_{|\Xi| \times 1}$ such that

$$A(\Xi)^T \begin{bmatrix} x & x' \end{bmatrix} = \begin{bmatrix} y & y' \end{bmatrix} \geq \begin{bmatrix} \mathbb{O}_{M \times 1} & \mathbb{O}_{M \times 1} \end{bmatrix}$$

Thus, for all $\alpha \in [0, 1]$, $\alpha x + (1 - \alpha)x'$ satisfies the following properties:

- $\alpha x + (1 - \alpha)x' \geq \mathbb{O}_{|\Xi| \times 1}$
- $A(\Xi)^T (\alpha x + (1 - \alpha)x') = \alpha y + (1 - \alpha)y' \geq \mathbb{O}_{M \times 1}$
- $I(\Xi)^T (\alpha x + (1 - \alpha)x') = \alpha I(\Xi)^T x + (1 - \alpha)I(\Xi)^T x' \leq \alpha q + (1 - \alpha)q = q$

Hence, $\alpha y + (1 - \alpha)y' \in P(q)$

Part c:

Let $Q = q^T \mathbb{I}$. $P(q)$ is bounded by a hypercube with extremes placed at the origin ($\mathbb{O}_{M \times 1}$) and at $Q \times \begin{bmatrix} \max_{c \in \{1, \dots, C\}} \tau(1, c)_1 & \dots & \max_{c \in \{1, \dots, C\}} \tau(M, c)_M \end{bmatrix}$. Hence, to show that $P(q)$ is compact, it is enough to prove that the set is closed.

Let $\{y^i\}_{i=1}^\infty$ a sequence of elements in $P(q)$ such that $y^i \rightarrow y$. Then, there exists a sequence of $\{x^i\}_{i=1}^\infty$ such that $A(\Xi)^T x^i = y^i$. Due to the linearity of the problem, $x^i \rightarrow x$,

where $A(\Xi)^T x = y$. Furthermore, for all i , $x^i \geq \mathbb{O}_{|\Xi| \times 1}$ and $I(\Xi)^T x^i \leq q$. Then, it holds that $x \geq \mathbb{O}_{|\Xi| \times 1}$ and $I(\Xi)^T x \leq q$. Hence, $y \in P(q)$. Ergo, $P(q)$ is closed.

Part d:

If $y \in P(q)$, then there exists $x \geq \mathbb{O}_{|\Xi| \times 1}$ such that $A(\Xi)^T x = y$ and $I(\Xi)^T x \leq q$. Assume that $I(\Xi)^T x < q$. Therefore, there exists $t > \mathbb{O}_{|\Xi| \times 1}$ such that $I(\Xi)^T x + t = q$. Thus, $\exists t_c > 0$ which implies that country c has idle workers. However, due to assumption (A3), c can use the spare workers to create a positive quantity of a set of goods without any net negative production. When this is done, the overall production is greater than y . Hence, $y \notin E(q)$. Contradiction.

Part e:

From parts (a), (b), and (c), and Weierstrass theorem, there exists $y \in E(q)$ such that for all $z \in E(q)$, $y_1 \geq z_1$. By construction, there exists an $x \in \mathbb{R}^{|\Xi|}$ such that $A(\Xi)^T x = y$.

Assume that $y_{-1} \neq \mathbb{O}_{(M-1) \times 1}$. Then, there exists an element $i > 1$ such that $y_i > 0$, which implies the existence of a technique j in country c that produces good i and employs $x_j > 0$ workers.

Therefore, if country c reduces the production of i by $\min\{y_i, x_j A(\Xi)_{ji}\}$ units, the overall production of good 1 does not change. From assumption (A3) these idle workers can be allocated to increase the production of all goods, in particular good 1. Thus, y did not maximize the production of good 1. Contradiction. ■

Lemma 1.2 R , as defined in Theorem 1.1, has full row rank.

Proof:

Step 1: The first $M - 1$ rows are linearly independent.

Define ξ_c as in the proof of Lemma 1.1. Based on the proofs of Arrow (1951), assumption (A3) guarantees that it is not possible to find $y \neq \mathbb{O}_{(M-1) \times 1}$ such that $y^T A(\Xi)_{-1}^T = \mathbb{O}_{|\Xi| \times 1}^T$ and that

$$A(\xi_c) \begin{bmatrix} 0 \\ y \end{bmatrix} = \mathbb{O}_{M \times 1}$$

Therefore $A(\Xi)$ is invertible.

Step 2: The last C rows of R are linearly independent.

Each column in $I(\Xi)$ is a canonical vector. Hence, all the rows of $I(\Xi)^T$ are linearly independent.

Step 3: The first $M - 1$ rows in R are linearly independent to any of the last C rows.

Assume there exists $y \neq \mathbb{O}_{(M-1) \times 1}$ such that $y^T A(\Xi)_{.-1}^T = I(\Xi)_{.c}^T$. In particular, this means that

$$A(\xi_c)_{.-1} y = \mathbb{I}_{M \times 1}$$

From Step 2, $A(\xi_c)$ is invertible. Then, $t = A(\xi)^{-1} \mathbb{I}_{M \times 1} \gg 0$ (Arrow, 1951).

Thus, $\mathbb{I}_{M \times 1} = \sum_{i=1}^M t_i A(\xi_c)_{.i} = \sum_{i=2}^M y_i A(\xi_c)_{.i}$. Therefore, $A(\xi_c)_{.1} = \frac{p_1 - y_1}{p_1} \sum_{i=2}^M A(\xi_c)_{.i}$. This last result contradicts the fact that $A(\xi_c)$ is invertible. Hence, the first $M - 1$ rows in R are linearly independent to any of the last C rows.

Step 4:

From steps 1-3, the $M - 1 + C$ rows of R are linearly independent. Ergo, R is full row rank. ■

Theorems

Proof of Theorem 1.1

From Lemma 1.1 parts (a) and (c), there is a vector x that maximizes the production of the final good. Furthermore, Lemma 1.1 parts (d) and (e) implies that $A(\Xi)_{.-1}^T x = \mathbb{O}_{(M-1) \times 1}$ and $I(\Xi)^T x = q$. Hence, all the constraints of the linear program bind at the optimum.

Karush-Kuhn-Tucker (KKT) conditions state that x solves the linear program if and only if there exist a vector of shadow prices $\lambda^T = [p^T \quad w^T] \gg 0$ and a vector of reduced costs $s \geq 0$ such that:

$$Rx = \begin{bmatrix} \mathbb{O}_{(M-1) \times 1} \\ q \end{bmatrix}$$

$$R^T \lambda + s = A(\Xi)_{.1}$$

$$x \geq \mathbb{O}_{|\Xi| \times 1}$$

$$s \geq \mathbb{O}_{|\Xi| \times 1}$$

$$\lambda \gg \mathbb{O}_{M+C-1 \times 1}$$

$$x' s = 0$$

The fundamental theorem of linear programming states that the solutions of the linear program are (generically) extreme points of the bounded polyhedron created by the constraints, or (non-generically) lie on a face of it. Hence, for any solution, there are at most $M + C - 1$ variables different than 0; in the linear program jargon, these are called the basic

variables. Thus, their correspondent techniques will be known as basic techniques and will be identified with γ .

Given that R is full row rank (Lemma 1.2), its columns can be ordered into two matrices: an invertible squared matrix $B := \begin{bmatrix} -A(\gamma)_{.-1}^T \\ I(\gamma)^T \end{bmatrix}$ and a matrix $N := \begin{bmatrix} -A(\Xi \setminus \gamma)_{.-1}^T \\ I(\Xi \setminus \gamma)^T \end{bmatrix}$.

Based on KKT, given the set of basic variables, x and s satisfy the following constraints:

$$\begin{aligned} x &= \begin{bmatrix} x_B \\ x_N \end{bmatrix}, s = \begin{bmatrix} s_B \\ s_N \end{bmatrix} \\ R &= \begin{bmatrix} B & N \end{bmatrix} \\ \begin{bmatrix} x_B \\ x_N \end{bmatrix} &= \begin{bmatrix} B^{-1} \begin{bmatrix} \mathbb{O}_{(M-1) \times 1} \\ q \end{bmatrix} \\ \mathbb{O}_{(|\Xi| - (C+M-1)) \times 1} \end{bmatrix} \\ \begin{bmatrix} s_B \\ s_N \end{bmatrix} &= \begin{bmatrix} \mathbb{O}_{(M+C-1) \times 1} \\ A(\Xi \setminus \gamma)_{.1} - N^T \lambda \end{bmatrix} \\ \lambda &= (B^T)^{-1} A(\gamma)_{.1} \end{aligned}$$

As it was mentioned before, in the optimization program all the constraints bind in the optimum. Therefore, all intermediate goods and labour market clear (then, by Walras Law, the final good market also clears); i.e. it satisfies conditions (3), (4), and (5) of the equilibrium conditions.

Regarding prices, due to the KKT conditions, $B^T \lambda = A(\gamma)_{.1}$ which implies that,

$$\begin{bmatrix} -A(\gamma)_{.-1} I(\gamma) \end{bmatrix} \begin{bmatrix} p \\ w \end{bmatrix} = A(\gamma)_{.1}$$

Therefore,

$$A(\gamma)p = I(\gamma)w$$

Thus, for all the basic techniques, the profit is 0. By definition, these are a superset of those techniques which are active in the equilibrium. Hence, condition (2) of the equilibrium is satisfied.

Finally, for the remaining techniques

$$N^T \lambda + s_N = A(\Xi \setminus \gamma)_{.1} \Rightarrow A(\Xi \setminus \gamma)p + s_N = I(\Xi \setminus \gamma)w \Rightarrow A(\Xi \setminus \gamma)p \leq I(\Xi \setminus \gamma)w$$

Consequently, all the techniques that are not part of the basic techniques have non-positive profit (which satisfies condition (1) of the equilibrium).

In this way, x , $\begin{bmatrix} 1 & p^T \end{bmatrix}^T$ and w define the vectors of quantities, prices, and wages that satisfy all equilibrium conditions. To prove the other direction of the implication, notice that if x , p , and w satisfy all the equilibrium conditions, they also meet KKT conditions. Ergo, x is a solution of the optimization program. ■

Proof of Theorem 1.2

Part a:

The fact that $\tau(m, c) \in \gamma$ implies that $p^T \tau(m, c) = w_c$ and $p^T \tau(m, c') \leq w_{c'}$. Therefore, $p^T (\tau(m, c') - \tau(m, c)) \leq w_{c'} - w_c$. On the other hand, assumption (A4) implies that $p^T (\tau(m, c') - \tau(m, c)) = p_m (\tau(m, c')_m - \tau(m, c)_m)$. Thus,

$$w_{c'} - w_c \geq p_m (\tau(m, c')_m - \tau(m, c)_m)$$

Part b:

Following the same derivations as part (a), if $p_m (\tau(m, c')_m - \tau(m, c)_m) \leq w_{c'} - w_c$ and $p_{m'} (\tau(m', c')_{m'} - \tau(m', c)_{m'}) \geq w_{c'} - w_c$, then $p_m (\tau(m, c')_m - \tau(m, c)_m) \leq p_{m'} (\tau(m', c')_{m'} - \tau(m', c)_{m'})$.

Thus,

$$\tau(m, c')_m > \tau(m, c)_m \Rightarrow 1 \leq \frac{p_{m'} (\tau(m', c')_{m'} - \tau(m', c)_{m'})}{p_m (\tau(m, c')_m - \tau(m, c)_m)}$$

Part c:

$\{\tau(m, c), \tau(m, c')\} \subseteq \gamma$ implies that $p^T \tau(m, c) = w_c$ and $p^T \tau(m, c') = w_{c'}$. Therefore, $p^T (\tau(m, c') - \tau(m, c)) = w_{c'} - w_c$. Ergo, $w_{c'} - w_c = p_m (\tau(m, c')_m - \tau(m, c)_m)$. ■

Proof of Theorem 1.3

Part a:

Step 1: All raw materials have the same price.

Let m, m' be raw materials such that $\{\tau(m, c), \tau(m', c')\} \subseteq \gamma$ and assume that $p_m > p_{m'}$. From equilibrium conditions (1) and (2), it holds that $k_c p_m - w_c = 0$ and $k_{c'} p_m - w_{c'} \leq 0$. Symmetrically, $k_{c'} p_{m'} - w_{c'} = 0$ and $k_c p_m - w_c \leq 0$. However, the previous equations also imply that, $k_{c'} p_m - w_{c'} > 0$. Contradiction. Hence, both goods have the same prices.

Step 2:

Consider two goods a, b that have the same production structure but different prices. Given that the two networks are isomorphic, $x \in \text{Nodes}(\text{Prod}(a))$ requires the existence of $x' \in \text{Nodes}(\text{Prod}(b))$ such that $\text{Prod}(x) = \text{Prod}(x')$.

Set

$$m \in \underset{x \in \text{Nodes}(\text{Prod}(a))}{\text{argmin}} \{rd(x) | p_x \neq p_{x'}\}$$

and let $m' \in \text{Nodes}(\text{Prod}(b))$ such that $\text{Prod}(m) = \text{Prod}(m')$. Without loss of generality, assume that $p_m > p_{m'}$. Let c, c' be two countries such that, in equilibrium $\{\tau(m, c), \tau(m', c')\} \subseteq \gamma$.

Given that m, m' have isomorphic production structures, for every $n \rightarrow m$, it holds that $rd(n) < rd(m)$. Thus, by construction, all the correspondent inputs of m and m' have the same price. Ergo, in equilibrium, $p^T \tau(m, c) - w_c = 0$ and

$$0 \geq p^T \tau(m, c') - w_{c'} = p^T \tau(m, c) - p_m k_c + p_m k_{c'} - w_{c'}$$

Also,

$$0 = p^T \tau(m', c') - w_{c'} = p^T \tau(m, c) - p_m k_c + p_{m'} k_{c'} - w_{c'} < p^T \tau(m, c) - p_m k_c + p_m k_{c'} - w_{c'}$$

Contradiction. Thus, in equilibrium, isomorphic structures have the same prices.

Part b:

Given that $k_{c'} \geq k_c$, Theorem 1.2 implies that

$$1 \leq \frac{p_{m'}(\tau(m', c')_{m'} - \tau(m', c)_{m'})}{p_m(\tau(m, c')_m - \tau(m, c)_m)} = \frac{p_{m'}(k_{c'} - k_c)}{p_m(k_{c'} - k_c)} = \frac{p_{m'}}{p_m}$$

Ergo, $p_m \leq p_{m'}$.

■

Proof of Theorem 1.4**Part a:**

Let $\tau(m', c') \in \gamma$. Given that $m \rightarrow m'$, then

$$p_{m'} = \frac{1}{k_{c'}}(-p_{-m'}^T \tau(m', c')_{-m'} + w_{c'}) > \frac{p_m |\tau(m', c')_m|}{k_{c'}} \geq p_m$$

Hence, $p_{m'} > p_m$.

Part b:

If $m \rightarrow m'$, then $p_{m'} > p_m$. Due to Theorem 1.3, this observation implies that $p_{m'} > p_m \Rightarrow c \geq c'$. Therefore, if there is a path from m to m' such that $m \rightarrow m_{a_1} \rightarrow \dots \rightarrow m_{a_n} \rightarrow m'$ and for all i , $\tau(m_{a_i}, c_{a_i}) \in \gamma$, then $c \geq c_{a_1} \geq \dots \geq c_{a_n} \geq c'$.

■

Proof of Theorem 1.5**Part a:**

To prove that the algorithm is well defined, it is sufficient to show that it defines w and p . For this purpose, assume there is a set of goods $A = \{m : m \text{ is not priced}\}$ and $m' \in \text{argmin}\{rd(m) : m \in A\}$.

If $\tau(m', C) \in \gamma$, then step 1 of the algorithm and the fact that w_C is given imply that $\exists n \rightarrow m'$ such that n is not priced, which leads to a contradiction because $rd(n) < rd(m')$. Hence, as all the prices of goods produced by C are defined. This implies that w_{C-1} is also defined.

The previous result shows that all goods that C produce are priced. Moreover, notice that if all the prices of the goods produced by $c+1$ are defined, then w_c is also defined. Having these two observations in mind, assume that for all $c' > c$, and for all n such that $\tau(n, c') \in \gamma$, p_n is defined. Therefore if $m \in A$, m has to be produced by a country $c^* \leq c$. Then, the fact that there is a good m' that is not priced implies that $\exists n' \rightarrow m'$ such that n' is not priced or w_c is not defined. But due to the algorithm, w_c is defined. Also, as in the previous case, $rd(n') < rd(m')$ implies that $p_{n'}$ is defined. Ergo, m' has to be defined. This means that $A = \emptyset$. Following the inductive process, the algorithm is well defined.

Part b: Proof by induction

Define w', p' as the equilibrium wages and prices. The normalization implies $w'_C = w_C = \alpha$.

Base Case: Country C :

By construction $w_C = w'_C$.

Let $M = \{m : \tau(m, C) \in \gamma \wedge p_m \neq p'_m\}$ and $m' = \operatorname{argmin}\{rd(m) : m \in M\}$. From step 1 of the algorithm, $p_{m'} = \frac{1}{k_c}(w_c + \sum_{n \rightarrow m'} |A(\tau(m', c))_n| p_n) = \frac{1}{k_c}(w'_c + \sum_{n \rightarrow m'} |A(\tau(m', c))_n| p'_n) = p'_{m'}$. The second equality comes from the fact that $rd(n) < rd(m')$ implies that $p_n = p'_n$. Thus, $M = \emptyset$.

Induction step:

Assume that for all $c' > c$ and for all m such that $\tau(m, c') \in \gamma$, $w_{c'} = w'_{c'}$ and $p_m = p'_m$. Given that the economy is linked, there exists m such that

$$(p')^T \tau(m, c+1) - w'_{c+1} = (p')^T \tau(m, c) - w'_c = 0$$

Therefore,

$$w'_c = w'_{c+1} + p'_m(k_c - k_{c+1}) = w_{c+1} + p_m(k_c - k_{c+1}) = w_c$$

where the second equality is due to induction hypothesis.

Furthermore, let $M = \{m : \tau(m, c) \in \gamma \wedge p_m \neq p'_m\}$ and $m' = \operatorname{argmin}\{rd(m) : m \in M\}$.

From step 1 of the algorithm,

$$p_{m'} = \frac{1}{k_c}(w_c + \sum_{n \rightarrow m'} |A(\tau(m', c))_n| p_n) = \frac{1}{k_c}(w'_c + \sum_{n \rightarrow m'} |A(\tau(m', c))_n| p'_n) = p'_{m'}$$

The second equality comes from the fact that $rd(n) < rd(m')$. Therefore, $M = \emptyset$. Thus, by induction principle, $w = w'$ and $p = p'$.

Finally, by the construction of the algorithm and the fact that it is an equilibrium vector, $p^T \tau(m, c) - w_c = 0 \Leftrightarrow p_m \in [P_c, P_{c-1}]$. Thus, these values coincide with the notation of Corollary 1.2.

■

Propositions

Proof of Proposition 1.1

This proof uses the notation of the proof of Theorem 1.1.

Part a:

Given the optimization program, $\lambda_i = \frac{\partial \Omega}{\partial b_i}$, where $b = \begin{bmatrix} \mathbb{O}_{(M-1) \times 1} \\ q \end{bmatrix}$.

By definition, $\lambda = \begin{bmatrix} p \\ w \end{bmatrix}$. Ergo, $w_c = \frac{\partial \Omega}{\partial q_c}$.

Part b:

By the KKT conditions,

$$\lambda = \begin{bmatrix} p \\ w \end{bmatrix} = \begin{bmatrix} -A(\gamma)_{\cdot 1} & I(\gamma) \end{bmatrix}^{-1} A(\gamma)_{\cdot 1}$$

Hence, given a set of basic techniques, the prices are uniquely determined and are independent of the labour quantities. ■

Proof of Proposition 1.2

Both parts of the proof are based on the simplex algorithm for dynamic programming. Therefore, the vernacular used is associated with the pivoting technique used in the book Applied Linear Programming (Bradley et al., 1977).

Part a:

The fact that $\{p_m : \tau'(m, c) \in \gamma'\} \subseteq \{p_n : \tau(n, c) \in \gamma\}$ implies that for every $\tau'(m, c) \in \gamma'$ exist $\tau(m', c), \tau(m'', c) \in \gamma$ such that $p_{m'} \leq p_m \leq p_{m''}$. From Corollary 1.1, this signifies that $\forall \tau'(m, c) \in \gamma', p^T \tau'(m, c) - w_c = 0$. Then, from Theorem 1.1, p, w are the price and income vectors of γ' .

Part b:

To simplify notation, denote the position of $\tau(m, c)$ as the first vector of the basic techniques γ . Given that optimal worker schedule x is unique, then $x_1 = 0$. Ergo,

$$e(1)^T B^{-1} \begin{bmatrix} \mathbb{O}_{(M-1) \times 1} \\ q \end{bmatrix} = 0$$

Hence, let $z \in \mathbb{R}^C$ be the last C components of $[B^{-1}]_1$. Then $z^T q = 0$.

Also, the reduced costs under γ are defined as

$$s = \begin{bmatrix} 0 \\ A(\Xi)_{\cdot 1N} - N^T \lambda \end{bmatrix} = \begin{bmatrix} A(\gamma) \begin{bmatrix} 1 \\ p \end{bmatrix} - I(\gamma)w \\ A(\Xi \setminus \gamma) \begin{bmatrix} 1 \\ p \end{bmatrix} - I(\Xi \setminus \gamma)w \end{bmatrix}$$

Let $s_{(m',c')} = p^T \tau'(m',c') - w_{c'}$. Due to the steps of the simplex algorithm, when one basic technique is changed (technically known as pivoting), the new reduced costs are

$$s' = s - \frac{s_{(m',c')}}{U_1} [V_1.]^T$$

where

$$U = B^{-1} \begin{bmatrix} A(\tau'(m',c')) \\ I(\tau'(m',c')) \end{bmatrix}, V = B^{-1} R$$

Therefore, the new shadow prices are updated via the following equation:

$$\begin{bmatrix} p' \\ w' \end{bmatrix} = \begin{bmatrix} p \\ w \end{bmatrix} - \frac{s_{(m',c')}}{U_1} [(B^{-1})_1.]^T$$

Hence, $\beta = -\frac{s_{(m',c')}}{U_1} z$. Finally, $\beta^T q = -\frac{s_{(m',c')}}{U_1} z^T q = -\frac{s_{(m',c')}}{U_1} 0 = 0$.

■

Proof of Proposition 1.3

Part a:

Given that the basic techniques are the same, i) is a direct consequence of the pricing algorithm; ii) is as well a consequence of step 3 of the algorithm.

For iii), an inductive process is followed both at the good and at the country level.

Base case: Price updating for country c .

Consider a good m such that $\tau(m,c) \in \gamma$ and $p_m > \max\{p_j : \tau(j,c+1) \in \gamma\}$.

If m is updated in the first iteration of country c , then

$$p'_m = \frac{1}{\beta k_c} (w_c + (1-\beta)k_c + \sum_{n \rightarrow m} |A(\tau(m,c))_n| p_n) = \frac{(1-\beta)k_c}{\beta} + \frac{p_m}{\beta} < p_m$$

Furthermore, for all its inputs $|p'_n - p_n| = 0 < |p'_m - p_m|$.

Induction step: Good level

Assume that for all the prices updated in c , up till the i -th iteration, $j \rightarrow k \Rightarrow |p_k - p'_k| > |p_j - p'_j|$ and $p_k > p'_k$.

Then, if m is updated in the iteration $i+1$,

$$p'_m = \frac{1}{\beta k_c} (w_c + (1-\beta)k_c + \sum_{n \rightarrow m} |A(\tau(m,c))_n| p'_n)$$

$$< \frac{1}{\beta k_c} (w_c + (1 - \beta)k_c + \sum_{n \rightarrow m} |A(\tau(m, c))_n| p_n) = \frac{(1 - \beta)}{\beta} + \frac{p_m}{\beta} < p_m$$

And given that $p_n - p'_n \geq 0$, then

$$p_m - p'_m = \frac{1}{k_c} \left(\sum_{n \rightarrow m} |A(\tau(m, c))_n| (p_n - p'_n) \right) + \frac{\beta - 1}{\beta k_c} ((w_c + k_c) + \sum_{n \rightarrow m} |A(\tau(m, c))_n| p'_n)$$

As all the terms are positive, $|p_m - p'_m| > |p_n - p'_n|$. Then, by induction hypothesis, all the prices updated by c , follow this inequality. Now consider country $c - 1$. From the price adjusting algorithm

$$\begin{aligned} w'_{c-1} - w_{c-1} &= w'_c - w_c + p'_m(k_{c-1} - \beta k_c) - p_m(k_{c-1} - k_c) \\ &= (1 - \beta)k_c + (p'_m - p_m)k_{c-1} + p_mk_c - p'_m\beta k_c < 0 \end{aligned}$$

Induction Step: Country level

Assume that, if k was updated in any iteration between the first iteration of C and the i -th of c' , it holds that $j \rightarrow k \Rightarrow |p_k - p'_k| > |p_j - p'_j|$ and $p'_k < p_k$. Also assume that $w'_{c'} < w_{c'}$.

Let m' be updated in the $i + 1$ iteration of country c' . Then, by induction hypothesis

$$p_{m'} - p'_{m'} = \frac{1}{k_{c'}} (w_c - w'_c + \sum_{n \rightarrow m} |A(\tau(m, c))_n| (p_n - p'_n)) > 0$$

Furthermore, as all the terms are positive, $\forall j \rightarrow m : |p_m - p'_m| > |p_j - p'_j|$. Then, by the induction principle, this condition holds for all the iterations of c' .

From the step 3 of the algorithm, and the induction in prices, $w_{c'-1} - w'_{c'-1} = w_{c'} - w'_{c'} + (P_{C'} - P'_{C'})(k_{c-1} - k_c) > 0$. Thus, by the induction principle, $\forall c' < c, w'_{c'} < w_{c'}$.

Finally,

$$\begin{aligned} (w_{c'-1} - w'_{c'-1}) - (w_{c'} - w'_{c'}) &= (P_{C'} - P'_{C'})(k_{c-1} - k_c) > 0 \\ \Rightarrow |w_{c'-1} - w'_{c'-1}| &> |w_{c'} - w'_{c'}| \end{aligned}$$

Thus, iii) is proven via the double induction in prices and countries.

Part b:

Repeat the proof of part (a) considering the fact that $w_C - w'_C = 0$ from the normalization condition.

■

Corollaries

Proof of Corollary 1.1

Part a:

From Theorem 1.3, price specialisation implies that $p_m \leq p_n$ (regarding c) and $p_n \leq p_m$ (regarding c'). Then, $p_m = p_n$.

Part b:

Assume that $p^T \tau(n, c) - w_c < 0$. Thus, in equilibrium there exists c' such that $p^T \tau(n, c') - w_{c'} = 0$. If $c' > c$, then $p_m(k_{c'} - k_c) \leq w_{c'} - w_c$ and $p_n(k_c - k_{c'}) < w_c - w_{c'}$. Therefore, $p_n(k_c - k_{c'}) < w_c - w_{c'} \leq p_m(k_c - k_{c'})$. Thus, $p_n > p_m$.

But, from Theorem 1.3, $p_n = p_{m'}$. Contradiction (the proof is symmetric if $c' < c$).

Part c:

Define p', w' such that $p = p'$, $w_{-c} = w'_{-c}$, and $w'_c = w_c + \alpha(k_{c'} - k_c) > w_c$. By direct calculation, p', w' satisfy the new equilibrium conditions. Ergo, the only value that changed after the shock was w_c .

■

Proof of Corollary 1.2

Define $P'_c = \min\{p | p^T \tau(m, c) - w_c = 0\}$ and $P''_c = \max\{p_m | p^T \tau(m, c) - w_c = 0\}$. Theorem 1.3 implies that $P''_c \leq P'_{c-1}$. Hence, if $P''_c = P'_{c-1}$, then define P_{c-1} as P'_{c-1} , otherwise define it as $P_{c-1} = \frac{P''_c + P'_{c-1}}{2}$.

From Corollary 1.1,

$$p_m \in [P'_c, P''_c] \Rightarrow p^T \tau(m, c) - w_c = 0$$

and by Theorem 1.3,

$$p_m \in \mathbb{R} \setminus [P'_c, P''_c] \Rightarrow p^T \tau(m, c) - w_c < 0$$

■

Proof of Corollary 1.3

From Theorem 1.4, all the goods in this economy have different prices. Thus, consider $c < c'$ and m such that $\{\tau(m, c), \tau(m, c')\} \subseteq \gamma$. Corollary 1.1 states that, for all $c < c'' < c'$ and for all m' such that $p_{m'} \neq p_m$, it holds that $p^T \tau(m', c'') - w_{c''} < 0$.

Every country has to produce something (to clear its labour market) in the equilibrium. Therefore, the previous observations imply that $\tau(m, c'') \in \gamma$. Furthermore, assumption (A7) and Theorem 1.2 signify that $\neg \exists m, m', c, c'$ such that $\{\tau(m, c), \tau(m, c'), \tau(m', c), \tau(m', c')\} \subseteq \gamma$. Hence, for each c , there can be at most two goods that are produced by other countries: one is shared with $c + 1$ and the other one is shared with $c - 1$. The exception of this rule are countries 1 and C , which only have one good in common with other countries.

Yet, Theorem 1.1 indicates that there are exactly $M + C - 1$ basic techniques. Thus, the previous condition is achieved by every country. Hence, $\forall c, c + 1 \exists! m : \{\tau(m, c), \tau(m, c + 1)\}$.

■

Chapter 2

On the Role of Unequal Treatment in Collective Agreements

2.1 Introduction

Solving collective action problems is crucial to the survival of many communities, therefore cooperation rules frequently emerge in societies to incentivise individuals to produce common goods (Ostrom, 2000). The current literature in this area evaluates such rules according to their capacity to produce collective goods and to maximise the community's overall welfare. However, these evaluations have produced several economic puzzles, and this paper focuses on four of them.

The first puzzle states that individuals are more likely to shirk from collective agreements when commitment is not enforced. However, agreements without commitment are recurrent in many communities. Several authors, such as Dasgupta (2009), have used the theory of repeated games to explain that any agreement that provides better conditions than the outcomes that individuals can obtain in autarky will be supported by communities where people continuously interact and monitor each other.

The second puzzle states that communities do not exhibit the welfare-maximising agreements predicted by the repeated games theory with perfect monitoring. To reduce the gap between the models and reality, authors such as Townsend (1994) have explained that private information creates stronger deviation incentives under welfare-maximising solutions, and therefore, the lack of complete information implies that welfare-maximising agreements are not supported in equilibrium. Nonetheless, even in small communities where perfect informa-

tion is a reasonable assumption and where individuals constantly interact, such agreements still do not maximise the welfare (Bardhan, 2000).

The third puzzle states that individual income heterogeneity can be either a barrier or a facilitator of collective agreements. Heterogeneity can be a barrier by creating different deviation incentives for different individuals. However, heterogeneity can also facilitate collective agreements when the groups that have an interest in the agreements can compromise with the other groups (Dayton-Johnson and Bardhan, 2002). In contrast to the previous puzzles, the models used to understand the relationship between heterogeneity and the viability of collective agreements presume that individuals can commit to cooperate (e.g. Munshi and Rosenzweig (2016)) or that income heterogeneity also implies hierarchical power structures (e.g. Baland and Platteau (1999)).

The last puzzle states that most collective agreements lead to unequal treatment in favour of wealthy groups in society, and at the same time, the agreements manage to reduce the utility differences among their members. To explain this puzzle, authors such as Krueger and Perri (2006), have assumed that individuals are identical, and examine how collective agreements reduce the dispersion of consumable income that individuals obtain.

Thus far, there are multiple theoretical solutions to explain each of the aforementioned puzzles. However, the premises used for each model are different and sometimes incompatible. For example, some puzzles assume the existence of commitment mechanisms, while others assume their absence; some study homogeneous individuals, while others require heterogeneity among them; and some study individuals with the same level of decision power, while others require social structures. Although each solution provides valuable insights, it is important to link these ideas within a unified framework of assumptions to better explain these puzzles. In the real world, the four puzzles discussed can be present in a single community; thus, one set of assumptions about the community should be able to solve all the puzzles. As such, the objective of this paper is to present this unified framework of assumptions.

“Collective action problems” is a broad term that incorporates multiple types of agreements that require cooperation among members of a community. Due to the large variety of topics covered under this label, this paper narrows the scope and focuses on two types of agreements that encompass a significant number of the relevant collective action problems studied in the literature: club good agreements and common pool agreements. In the club good case, individuals dedicate a fraction of their income to the common good, which generates equal benefits for each community member when the good is non-rival or is evenly distributed when the good is rival. In the common pool case, individuals dedicate an equal

fraction of their income to the common good, and the members then divide the benefits in a way that is not necessarily equal. These archetypal agreements are flexible structures that can describe risk-sharing agreements, public good projects, or social production schemes.

This paper represents both agreements as stochastic games played for infinite periods, where agents can stop cooperating in any period. To analyse these games, the current study takes a new approach that focuses on the rules (i.e. game parameters) that support cooperation for the most impatient members of the community (i.e. rules that are robust to the discount factor). These rules have not yet been studied in the literature, but they are central to unifying the previous puzzles into a single framework. On the one hand, such rules can emerge in scenarios where welfare-maximising rules are not an equilibrium of the game. On the other hand, such rules elucidate how collective agreements balance the rent-seeking behaviour of wealthy individuals with the redistribution of the interests of poor individuals in efforts to support cooperative games. Using this methodological approach, the paper explains why rules that minimise deviations are likely to occur in multiple communities and how these rules provide simultaneous answers to the four puzzles as well as providing new insights into the mechanisms that keep communities cooperating in collective agreements.

The structure of the paper is as follows. Section 2.2 positions the current model in the economic development literature. Section 2.3 describes the two types of collective agreements. Section 2.4 theoretically characterises the robust rules in each type of agreement and explains the previous puzzles within this unified framework. Section 2.5 presents a comparative statics analysis of scenarios with heterogeneous communities to explore the economic mechanisms that agreements use to promote cooperation in a community and to contrast welfare-maximising agreements with deviation-minimising agreements. Section 2.6 discusses how reasonable and useful the agreements presented in this paper are to understand how collective agreements work and provides a guideline for future work in this area. Section 2.7 concludes.

2.2 Literature Review

This paper lies at the intersection of two questions at the heart of development economics: 1) which mechanisms allow individuals to participate in collective action problems, and 2) how does inequality affect the existence of collective goods?

Regarding the first question, individuals frequently coordinate to produce goods despite strong incentives to free ride (Coleman, 1988). Dasgupta (2009) suggested five reasons why cooperation occurs: mutual affection, a pro-social disposition, credible threats, external

enforcement, and long-term relationships. The first two reasons provide a trivial solution to the problem: people help others because they care about them. However, this argument applies to specific scenarios and does not explain cooperation amongst selfish agents. The next two reasons, credible threats and external enforcement, provide explanations that include cooperation among selfish individuals; however, they still depend on the existence of institutions (such as law enforcement agents) that are not available in many communities. The last reason explains how the repeated nature of interactions between agents induces commitment, thus allowing cooperation to emerge. This idea appeared in seminal papers such as Thomas and Worrall (1990), Coate and Ravallion (1993), and Kocherlakota (1996), who demonstrated that selfish individuals contribute to risk-sharing agreements without enforcement. These papers show that under the infinite game paradigm, collective agreements are prone to having multiple equilibria, and the authors only focused on the equilibria that composed the Pareto frontier. More recent papers, such as that of Genicot and Ray (2003), extended the literature by refining the set of equilibria to those that are coalition-proof, but they still focus on welfare-maximising solutions.

The second question investigates the effects of heterogeneity on collective agreements. Cardenas (2005) showed that assuming individuals with different utility functions, costs, wealth levels, or outside options leads to a diverse range of conclusions. For that reason, the current study considers that individuals are heterogeneous because the income they receive in each period follows an idiosyncratic distribution. For comparison purposes, individuals are identical in every other aspect.

Under this framework, several authors, such as Alesina and La Ferrara (2000), Bardhan (2000), Bowles and Gintis (2002), and Bardhan and Dayton-Johnson (2002), have argued that the effect of inequality is ambiguous, but in general, it seems to reduce the possibility of cooperative solutions. The theoretical models used to explain these observations either presume that rules treat everyone equally (e.g. La Ferrara (2002)) or that agents are homogenous (e.g. Baland and Platteau (2007)), and in most cases, there is external enforcement so that agents cannot refuse their responsibilities. The main exception is Munshi and Rosenzweig (2016), who built a model similar to the common pool agreement presented in Section 2.3.3, whereby i) heterogeneous agents design a risk-sharing agreement that distributes benefits unevenly among agents, and ii) the unequal treatment provides different incentives to agents. While the authors merged unequal treatment rules and heterogeneous agents in the same framework, they also assumed that the community is capable of enforcing the agreements.

Because the goal of this paper is to solve the four puzzles within a unified framework, the model constructed in Section 2.3 does so by extending the idea of Munshi and Rosenzweig

(2016) to club goods and common pool agreements and embedding it in a repeated game framework. In this way, enforcement mechanisms do not need to be assumed. However, in contrast to the previous literature, this research uses welfare-maximising rules as a benchmark and then elaborates on the concept of robust rules that provide theoretical explanations of the empirical puzzles.

2.3 Collective Agreements

This section models two types of collective action problems in an infinite period game framework. Although the agreements have different mechanisms, they have the same mathematical structure, which facilitates comparisons between them. The first part of this section presents the notation common to both the club good and common pool agreements. The remaining parts of the section explain the specificities of each case as well as the situations in which they are represented.

2.3.1 General Environment

Collective agreements in this paper are structured around four functions: individual utility (u), the income distribution before the collective agreement (f), the collective good production function (g), and the net income allocation after the agreement (Γ).

Regarding the utility function, there are n individuals identified with subindices $i \in \{1, \dots, N\}$ who live for an infinite horizon. The utility function $u(c) : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is common for all agents. Individuals are risk averse (i.e. $u'(c) \geq 0, u''(c) \leq 0$), have a utility of 0 whenever consumption is zero (i.e. $u(0) = 0$), and exhibit satiation (i.e. $\lim_{c \rightarrow \infty} u(c) < \infty$). Finally, individuals have a time discount factor $\delta \in (0, 1)$, which depicts their impatience.

Individuals are heterogeneous in their income realisation. Each individual is identified with an income parameter r_i , such that $0 < r_1 \leq r_2 \leq \dots \leq r_N < \infty$. In each period t , individual i receives a random amount of perishable income Y_{it} (Y_t is a random vector that describes the income of each member of the community). Y_{it} follows the distribution $\frac{1}{r_i} f(\frac{Y_{it}}{r_i})$, where f is a continuous probability density function with support $\mathbb{R}_+$ such that if X follows an $f(x)$ distribution, then $\mathbb{E}[X] = 1$ and $\text{Var}[X] = \sigma^2$. The perishable income distribution is independent across members of the community and independently and identically distributed across time. Finally, y_{it} is the realised income of individual i in period t , and y_t represents the realised income vector of the community in period t .¹

¹In general, the realization of the random variable is denoted with lowercase letters.

The income characterisation has two implications. First, the income distribution of individuals with higher income parameters dominates first-order stochastically those with lower parameters; ergo, higher values of r_i are associated with wealthier people. Second, $\mathbb{E}[Y_{it}] = r_i$ and $\text{Var}[Y_{it}] = r_i^2 \sigma^2$. The first part shows that wealthier individuals have more desirable income distributions, while the second part suggests that individuals are not able to diversify their income sources. Although extreme, the second feature aims to represent small communities with few income sources, whereby rich and poor individuals differ in the amount of income sources but not in their variety. This in turn generates a parsimonious model that avoids diversification strategies that shadow the cooperation mechanisms. To illustrate a scenario in which this feature applies, consider a small agrarian community where farmers have few types of crops. In this case, rich and poor individuals have different amounts of land, but the crops they produce are probably similar, ergo their diversification portfolios are similar despite differences in wealth.

There is a collective good with a production function $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ that is increasing and concave. When there is no investment, the production is 0 (i.e. $g(0) = 0$). The collective action problem emerges because it is not profitable for any single individual to invest (i.e. $\frac{\partial(g(x)-x)}{\partial x} < 0$), but there are gains if everyone in the community contributes (i.e. $\frac{\partial ng(x)-x}{\partial x} \geq 0$). These conditions are summarised as $\frac{1}{n} \leq g'(x) < 1$.

Finally, consider a net income allocation function $\Gamma(y, a, p)_i : \mathbb{R}_+ \times [0, 1]^n \times \Delta^n \rightarrow \mathbb{R}_+^n$ such that

$$\Gamma(y, a, p)_i = y_i - a_i y_i + p_i n g\left(\sum_{j=1}^N a_j y_j\right) \quad (2.1)$$

where Γ describes the net consumable income that each agent has under an agreement type. It is composed of three parts: first, the income realisation of the period (y_i), second, the investment of the individual in the collective good ($a_i y_i$), where a_i represents the percentage of income that the individual contributes, and finally, the income obtained from the collective good, which is a function of the investment of all the members of the community ($p_i n g(\sum_{j=1}^N a_j y_j)$), where p_i represents the share that goes to individual i . When the collective good is equally distributed among the members of the community, the third component becomes $g(\sum_{j=1}^N a_j y_j)$.

The functions $\{u, f, g, \Gamma\}$ are not independent of each other. Indeed, they need to satisfy two assumptions.

Assumption A1 $\mathbb{E}[u(\Gamma(Y, a, p)_i)]$ exists.

Assumption (A1) states that individuals can calculate the expected utility of the agreement.

Assumption A2 $\forall p \in \Delta^n, \forall a, a' \in [0, 1]^n$ such that $\forall a \geq a'$ where $a_i = a_j > a'_i = a'_j$ and $p_i = p_j$, the following inequality holds:

$$|\mathbb{E}[u(\Gamma(Y, a', p)_i) - u(\Gamma(Y, a', p)_j)]| > |\mathbb{E}[u(\Gamma(Y, a, p)_i) - u(\Gamma(Y, a, p)_j)]|$$

Assumption (A2) claims that if two individuals who are investing the same share of their income increase it by the same fraction, then the difference in their expected utility decreases. In other words, if the community contributes more to the common good, the expected utilities of its members becomes more similar. Assumption (A2) holds for any CARA and bounded HARA functions. It is noteworthy that these families of risk-averse utility functions are the most common in the literature on collective agreements. All the proofs presented in this model take as given these two assumptions.

2.3.2 Club Good Agreements

Club good agreements (CGAs) describe a setting in which a common good benefits all the members of society equally, and the community is capable of excluding those who did not participate in its production. As such, individuals invest a fixed fraction of their income and receive the same benefit from the collective good. There are two ways to interpret this fact: either the collective good is non-rival, or the community divides it evenly among its members. Finally, collective goods require a minimum expected level of inputs to be worthy. Therefore, the expected contributions need to be at least k . Naturally, $\sum_{j=1}^N r_j \geq k$ so that the problem is feasible.

CGAs describe a wide range of collective action problems. For example, they can represent a risk-sharing agreement where everyone receives the average of the resources pooled by the community. They can also represent a community tax scheme to provide a common good, such as in La Ferrara (2002). In that case, k represents the expected flow of income that the policy-maker requires for the common good provision to be justified.

Under these definitions, CGAs are represented as a net income allocation function $\Gamma(y, a, \frac{1}{n}\mathbb{I}) : \mathbb{R}_+^n \times [0, 1]^n \rightarrow \mathbb{R}_+^n$ such that

$$\Gamma(y, a, \frac{1}{n}\mathbb{I}) = y_i - a_i y_i + g\left(\sum_{j=1}^N a_j y_j\right) \quad (2.2)$$

The agreement is characterised by $a = (a_1, \dots, a_n)$. Henceforth, $(a, \frac{1}{n}\mathbb{I})$ is known as the allocation rule, where valid rules require that $\sum_{j=1}^N a_j r_j \geq k$.

For illustration purposes, if a community participates in the club good agreement described by $(a, \frac{1}{n}\mathbb{I})$, the utility derived by the i -th individual in period t is represented by equation (2.3).

$$u(\Gamma(y, a, \frac{1}{n}\mathbb{I})_i) \quad (2.3)$$

In contrast, the utility of those individuals that do not participate is $u(y_{it})$. By construction, the investment in the collective good provides positive externalities to the group, but individuals have incentives to reduce their contribution. Hence, the community designs the following mechanism:

At the beginning of each period, individuals observe the income realisation of the members of the community and decide accordingly whether to participate in the collective agreement.² Individuals who contribute in one period can also participate in the agreement in the following period. Those who refuse to invest lose their chance to participate in the collective agreement in that period as well as in subsequent stages. This mechanism induces a stochastic game that repeats itself for infinite periods and therefore has multiple equilibria. This paper focuses on the cooperative equilibrium induced by a grim trigger strategy. The strategy states that in the first period, everyone participates in the collective agreement, and for period $t > 0$, an agent participates if and only if all the agents participated in the previous period.³

²Coleman (1988), Fafchamps (1992), Ostrom (2000), and Bowles and Gintis (2002) highlight that individuals in small communities have high monitoring capacities of the actions of their neighbours, and this is one of the main reasons why they are able to support many collective agreements. Thus, this feature of reality is represented by the assumption.

³This paper presents a grim trigger strategy because it allows the model to focus on the implications of heterogeneity without mixing its effects with the implications of more complex strategies. Naturally, these results can be extended to other strategies, but the lost in parsimony does not compensate the gains in intuition.

The following Bellman equations summarise the one-shot deviation principle:⁴

The utility of individual i when he refuses the agreement $U_R^i(y_t)$:

$$U_R^i(y_t) = u(y_{it}) + \delta \mathbb{E}[U_R^i(Y_{t+1})]$$

where $\mathbb{E}[U_R^i(Y_t)] = \frac{\delta}{1-\delta} \mathbb{E}[u(Y_i)]$.

The utility of individual i when he accepts the agreement $U_A^i(y_t)$:

$$U_A^i(y_t) = u(\Gamma(y_t, a, \frac{1}{n}\mathbb{I})_i) + \delta \mathbb{E}[U_A^i(Y_{t+1})]$$

where $\mathbb{E}[U_A^i(Y_t)] = \frac{\delta}{1-\delta} \mathbb{E}[u(\Gamma(Y, a, \frac{1}{n}\mathbb{I})_i)]$.

The difference between these utilities is formed by two elements:

Present trade-off:

$$PT_i(y_t, a, \frac{1}{n}\mathbb{I}) = u(\Gamma(y_t, a, \frac{1}{n}\mathbb{I})_i) - u(y_{it})$$

Future trade-off:

$$FT_i(a, \frac{1}{n}\mathbb{I}) = \mathbb{E}[u(\Gamma(Y, a, \frac{1}{n}\mathbb{I})_i) - u(Y_i)]$$

The present trade-off quantifies the difference between the utility of participating and the utility of leaving the agreement during the period of the decision. The future trade-off quantifies the difference between the expected utility of participating and the expected utility of leaving the agreement in a generic future period. Hence, it reflects the future implications that the current decision to participate will have on the utility of the individual.

From Lemma 2.2 in Appendix 2.B, $PT_i(y_t, a, \frac{1}{n}\mathbb{I})$ has a lower bound, and from assumption (A1), $FT_i(a, \frac{1}{n}\mathbb{I})$ exists. Thus, as long as $FT_i(a, \frac{1}{n}\mathbb{I})$ is positive, there exists $\delta^* : \forall \delta \geq \delta^*, U_A^i(y_t) - U_R^i(y_t) \geq 0$. In other words, for any $\delta \geq \delta^*$, the individual has no incentive to leave the agreement. Therefore, this strategy is a subgame perfect equilibrium if $FT_i(a, \frac{1}{n}\mathbb{I}) > 0$ and δ is high enough.

2.3.3 Common Pool Agreements

Common pool agreements (CPAs) describe a setting in which the community decides how to allocate a collective good among its members. In each period, individuals first pool an equal fraction (τ) of their income into the production of the collective good, and the community

⁴This methodology is consistent with the works of Thomas and Worrall (1988), Coate and Ravallion (1993), Kocherlakota (1996), and Ligon et al. (2002).

then allocates the benefits among its members. For consistency with the club good agreement assumptions, this setting assumes that $\tau \sum_{j=1}^N r_j \geq k$.

This model generalises Munshi and Rosenzweig (2016) model and encompasses a wide range of collective action problems. It can represent a risk-sharing agreement in which the common good is the sum of the individual incomes as well as other scenarios such as communal crops and shared irrigation systems, similar to Bardhan (2000) and Mukhopadhyay (2008). In contrast to club good agreements, the good produced under a CPA can be unevenly allocated among the population.

To make both arrangements comparable, the production function of the collective good is n times the production function g . Thus, a net income allocation function for a CPA is characterised by $\Gamma(y, \tau\mathbb{I}, p) : \mathbb{R}_+^n \times [0, 1] \times \Delta^n \rightarrow \mathbb{R}_+^n$ such that

$$\Gamma(y, \tau\mathbb{I}, p)_i = y_i - \tau y_i + p_i n g\left(\tau \sum_{j=1}^N y_j\right) \quad (2.4)$$

This agreement is characterised by $p = (p_1, \dots, p_n)$ and τ . Hence, $(\tau\mathbb{I}, p)$ is the allocation rule. This notation allows both agreements to be instances of equation (2.1).

If a community participates in the common pool agreement described by $(\tau\mathbb{I}, p)$, the utility derived by the i -th individual in period t is represented by equation (2.5).

$$u(\Gamma(y, \tau\mathbb{I}, p)_i) \quad (2.5)$$

Analogous to a CGA, the community designs a mechanism to guarantee the participation of its members with a grim trigger strategy. The following Bellman equations describe the one-shot deviation principle:

The utility of individual i if he rejects the agreement $U_R^i(y_t)$:

$$U_R^i(y_t) = u(y_{it}) + \delta \mathbb{E}[U_R^i(Y_{t+1})]$$

where $\mathbb{E}[U_R^i(Y_t)] = \frac{\delta}{1-\delta} \mathbb{E}[u(Y_i)]$.

The utility of individual i if he accepts the agreement $U_A^i(y_t)$:

$$U_A^i(y_t) = u(\Gamma(y_t, \tau\mathbb{I}, p)_i) + \delta \mathbb{E}[U_A^i(Y_{t+1})]$$

where $\mathbb{E}[U_A^i(Y_t)] = \frac{\delta}{1-\delta} \mathbb{E}[\Gamma(Y, \tau\mathbb{I}, p)_i]$.

The difference between these utilities has two elements that have the same interpretation as in the previous case:

Present trade-off:

$$PT_i(y_t, \tau\mathbb{L}, p) = u(\Gamma(y_t, \tau\mathbb{L}, p)_i) - u(y_{it})$$

Future trade-off:

$$FT_i(\tau\mathbb{L}, p) = \mathbb{E}[u(\Gamma(Y, \tau\mathbb{L}, p)_i) - u(Y_i)]$$

As in the previous case, the grim trigger strategy is a subgame perfect equilibrium if $FT_i(\tau\mathbb{L}, p) > 0$ and δ is high enough.

2.4 Efficiency vs. Robustness

The studies by Elbers et al. (2004) and Mulder et al. (2009) showed that inequality in small communities is the norm rather than the exception. This section proves that this asymmetry among individuals creates differences between rules that maximise the aggregate welfare and rules that minimise the deviation incentives of the participants. The first part of the section defines welfare-maximising agreements and uses them as a benchmark scenario to compare rules that minimise deviation incentives. The second part defines robust rules as those that minimise the deviation incentives of impatient individuals and shows how these rules explain the four puzzles that drive this paper.

Definition 2.1 (Admissible rules) *A rule (α, β) is admissible if for all $i \in \{1, \dots, N\}$, it holds that $FT_i(\alpha, \beta) \geq 0$ and $\sum_{j=1}^N \alpha_j r_j \geq k$.*

Definition 2.1 assigns a name to those rules that are a subgame perfect equilibrium for a high enough value of δ and that satisfy the minimum contribution condition. These are the rules that have the potential to become norms, which makes them the focus of this research.

2.4.1 On Efficiency

Club good and common pool agreements have multiple admissible rules. This phenomenon has been well studied by authors such as Fafchamps (1992), Sethi and Somanathan (1996), Ostrom (2000), and Dasgupta (1995). They concluded that the rules chosen by a society depend on historical processes as well as the power relations among its decision-makers. However, the economic literature frequently focuses on welfare-maximising solutions. In this context, maximising the welfare is equivalent to maximising the sum of future trade-off values

(these values are a monotonous transformation of the *ex-ante* benefits of the agreement). For this paper, rules that maximise the welfare are called efficient because they are not Pareto dominated, and when the community is studied as a corporation, they maximise the overall generated benefit.

Definition 2.2 (Efficient rules) *A rule (α, β) is efficient if it is admissible and maximises $\sum_{j=1}^N FT_j(\alpha, \beta)$.*

Based on the previous definition it is possible to obtain the first result on collective agreements.⁵

Theorem 2.1 (Efficient rules) *For both CGA and CPA, $FT_N(\mathbb{I}, \frac{1}{n}\mathbb{I}) > 0$ if and only if $(\mathbb{I}, \frac{1}{n}\mathbb{I})$ is the efficient rule.*

Theorem 2.1 proves that as long as the wealthiest individual (N) finds the rule admissible, an efficient rule treats all community members equally. Furthermore, consistent with the Borch (1962) rule, the marginal utility of each individual is the same, independent of the income draw. Finally, $FT_N(\mathbb{I}, \frac{1}{n}\mathbb{I}) > 0$ is easily achieved when individuals have a high degree of risk aversion or low heterogeneity (low r_i variance). For that reason, and because it provides a clear benchmark for the utilitarian solution, the remainder of this article assumes (unless stated otherwise) that $FT_N(\mathbb{I}, \frac{1}{n}\mathbb{I}) > 0$ and the efficient rule — also called egalitarian rule — is $(\mathbb{I}, \frac{1}{n}\mathbb{I})$ for both CGAs and CPAs.

Theorem 2.1 explains why the literature has focused on equal treatment rules. As long as the income heterogeneity within the population is low enough, welfare-maximising collective agreements are fully specified.

2.4.2 On Robustness

Is it reasonable to choose an egalitarian rule? Due to income heterogeneity, the exit options are different for each participant. Thus, the expected future trade-off in a collective agreement is not the same for everyone, and heterogeneity creates differentiated incentives for people to leave the agreement, i.e. the minimum required changes between individuals. Moreover, there are cases where the community might prefer to design agreements that allow the lowest discount rates. For example, if the community has no information about the discount rate before designing the rules, they might prefer rules that have the highest likelihood to be

⁵All the lemmata as well as all the proofs of the lemmata, theorems, and propositions are presented in Appendix 2.B.

supported in equilibrium. This decision implies that the community is willing to gain stability in exchange for efficiency. Another explanation for why communities implement robust rather than efficient rules is the path dependency of the agreements. Fafchamps (1992) and Sethi and Somanathan (1996) discussed how rules appear randomly, and if individuals are not willing to deviate, these rules might become norms. From this perspective, if there is an unexpected change in the discount rate of some individuals, efficient rules might have a greater chance of disappearing than other rules do. Motivated by the previous discussion, robust rules are defined as rules that can be sustained at equilibrium for the smallest possible δ .⁶

Definition 2.3 (Worst present trade-off) For a rule (α, β) , define the worst present trade-off as $M_i(\alpha, \beta) = \liminf_{y \in \mathbb{R}_+^I} PT_i(y, \alpha, \beta)$.

Definition 2.4 (Robust rules) For a rule (α, β) , let

$$Ds(\alpha, \beta) = \left\{ \delta \mid \forall i M_i(\alpha, \beta) + \frac{\delta}{1-\delta} FT_i(\alpha, \beta) \geq 0 \right\}$$

(α, β) is robust if it is admissible and for all admissible rules: (α', β') , $Ds(\alpha', \beta') \subseteq Ds(\alpha, \beta)$.

Definition 2.3 specifies the scenario in which the present trade-off of an individual is the lowest. Regarding Definition 2.4, if $M_i(\alpha, \beta) + \frac{\delta}{1-\delta} FT_i(\alpha, \beta) < 0$, then individual i will leave the game after some period, and cooperation cannot be a subperfect game equilibrium. Thus, robust rules are equilibrium rules for a large range of δ . Furthermore, robust rules differ from efficient rules because their focus is on preserving the stability of the agreement and not maximising the welfare.

Note that

$$M_i(\alpha, \beta) + \frac{\delta}{1-\delta} FT_i(\alpha, \beta) \geq 0 \Leftrightarrow \delta \geq \frac{1}{1 - \frac{FT_i(\alpha, \beta)}{M_i(\alpha, \beta)}}$$

Hence, a robust rule (α, β) is a solution to the following problem:

$$(\alpha, \beta) = \operatorname{argmax}_{(\alpha', \beta') \in S} \left\{ \min_{i \in \{1, \dots, N\}} \left\{ \frac{FT_i(\alpha', \beta')}{M_i(\alpha', \beta')} \right\} \right\}$$

where the conditions over S depend on the type of agreement.

⁶The definition presented here extends that of Kalai and Stanford (1988).

Proposition 2.1 $\forall f, g, u$:

- (a) *It is possible that there is no egalitarian rule.*
- (b) *In a CGA, if there is an egalitarian rule, then there is a robust rule.*
- (c) *In a CPA, there is a robust rule independent of the existence of egalitarian rules.*

Proposition 2.1 proves that depending on the context (defined by f, g, u), it is sometimes impossible to find egalitarian rules that are subgame perfect equilibria. However, part (b) states that if those rules exist, it is also true that robust rules exist for a club goods agreement. Moreover, part (c) claims that independent of the existence of egalitarian rules, robust rules always exist for common pool agreements. Hence, in every scenario in which an egalitarian rule emerges, it is plausible that a robust rule also emerges. However, there are contexts in which robust rules can emerge but egalitarian rules cannot. Thus, Proposition 2.1 implies that robust rules are more likely to emerge in a society. Finally, these rules are subgame perfect equilibria in which they exhibit cooperation among individuals with limited commitment, therefore satisfying the first puzzle.

To simplify the mathematical notation in the characterisation of robust rules, a new assumption is stated.

Assumption A3 *Let $r_i > r_j$ and define*

$$T(a_i, a_j) = \mathbb{E} \left[\int_0^\infty \int_0^\infty \frac{\partial^2 u(\Gamma(Y, a, \frac{1}{n} \mathbb{I})_i)}{\partial Y_i \partial Y_j} F\left(\frac{y_i}{r_i}\right) F\left(\frac{y_j}{r_j}\right) dy_i dy_j \right]$$

If $a_i r_i \leq a_j r_j$ then $\mathbb{E}[u(Y_i)] - \mathbb{E}[u(Y_j)] \leq T(a_i, a_j) - T(a_j, a_i)$.

Assumption (A3) defines $T(a_i, a_j)$ as the measurement of the expected complementary effect that individual j has on individual i 's investment decision. Then, it states that if a wealthy individual is expected to invest less than a poor individual, then the difference between their utility without agreement is smaller than the difference between their expected complementarities.⁷

⁷Assumption (A3) holds for functions such as constant absolute risk aversion, hyperbolic absolute risk aversion, and constant relative risk aversion, but the proof is case specific. For simplicity, instead of proving it, it is taken as given in Theorem 2.2.

Theorem 2.2 *For a CGA*

- (a) *If $(a, \frac{1}{n}\mathbb{I})$ is robust, then $k = \sum_{i=1}^N a_i r_i$ and if $r_i > r_j$, then $a_i < a_j$ or $a_j = 1$.*
- (b) *Given assumption (A3), if $(a, \frac{1}{n}\mathbb{I})$ is robust, then $r_i > r_j$ implies that $a_i r_i \leq a_j r_j$.*

For a CPA

- (c) *If $(\tau\mathbb{I}, p)$ is robust, then $r_i > r_j$ implies that $p_i > p_j$ and $\tau = k(\sum_{i=1}^N r_i)^{-1}$.*
- (d) *If $(\tau\mathbb{I}, p)$ is robust, then for all i, j , $-\frac{FT_j(\tau\mathbb{I}, p)}{M_j(\tau\mathbb{I}, p)} = -\frac{FT_i(\tau\mathbb{I}, p)}{M_i(\tau\mathbb{I}, p)}$.*

Theorem 2.2 is composed of two parts: the first part describes club good agreements, and the second part describes common pool agreements. Regarding CGAs, part (a) reveals that wealthier individuals contribute lower percentages of their income to the common good, while poor individuals contribute all their income. In contrast, part (b) states that the expected value of the contribution of the rich has to be higher than the contribution of the poor. Regarding CPAs, part (c) states that richer individuals obtain higher shares of the collective good, and part (d) claims that robust agreements balance all the incentives, such that all individuals have a common lowest δ . Finally, parts (a) and (c) express that in both cases, individuals will not contribute as much as they can to the collective good and that the agreements provide preferential treatment to wealthier individuals.

To provide a better understanding of the previous theorems, Example 2.1 constructs efficient and robust agreements for a three-agent scenario.

Example 2.1 (CARA agents; part 1)

There is a community with three individuals. Their utility function is $u(x) = 100(1 - e^{-2x})$, a member of the CARA family. Income follows an exponential distribution with individual parameters $r_1 = 1 - t, r_2 = 1, r_3 = 1 + t$ where $t \in [0, 1)$. Regarding the collective good function, there are two scenarios. In the first scenario, the function represents a risk-sharing agreement defined by $g(x) = x/3$. In the second scenario, there is a positive externality in which the function is $g(x) = x/2$. Finally, $k = 1$. This formulation satisfies assumptions (A1), (A2), and (A3), making the example ideal for illustration purposes.

Figure 2.4.1 shows the different robust agreements for values of $t \in [0, 1)$. Consistent with Theorem 2.2 parts (a) and (c), wealthier individuals have favourable rules. Moreover, in club good agreements, once inequality is high enough, the poorest individuals contribute all their income to the collective good.

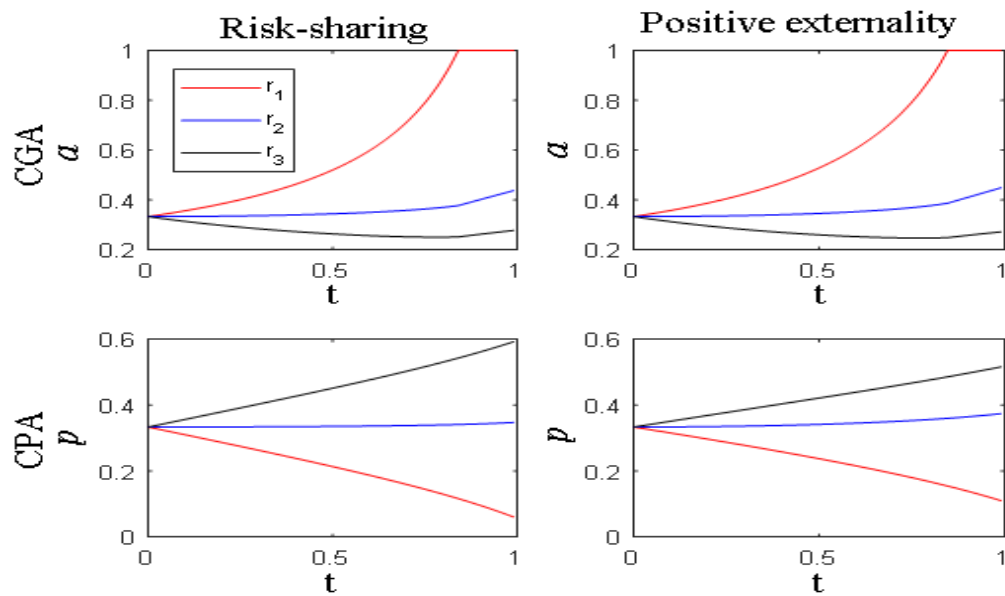


Fig. 2.4.1 Robust rules

Figure 2.4.2 analyses the expected contributions of each person under a club good agreement. As stated in Theorem 2.2 part (b), the wealthiest individual contributes more than the poor individual. Thus, Figure 2.4.1 and 2.4.2 represent the balance between levels and percentages of contributions.

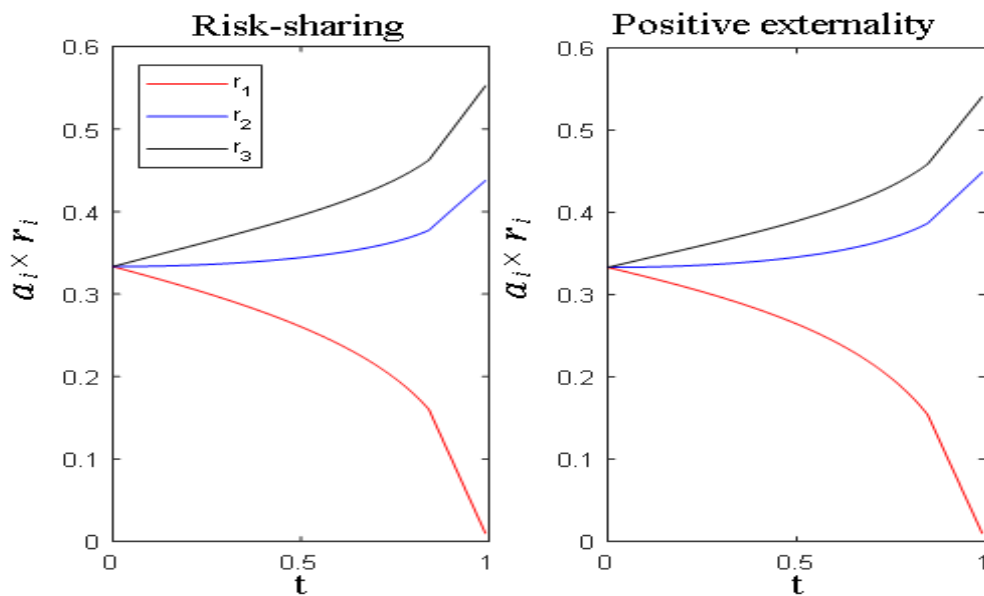


Fig. 2.4.2 Expected contribution

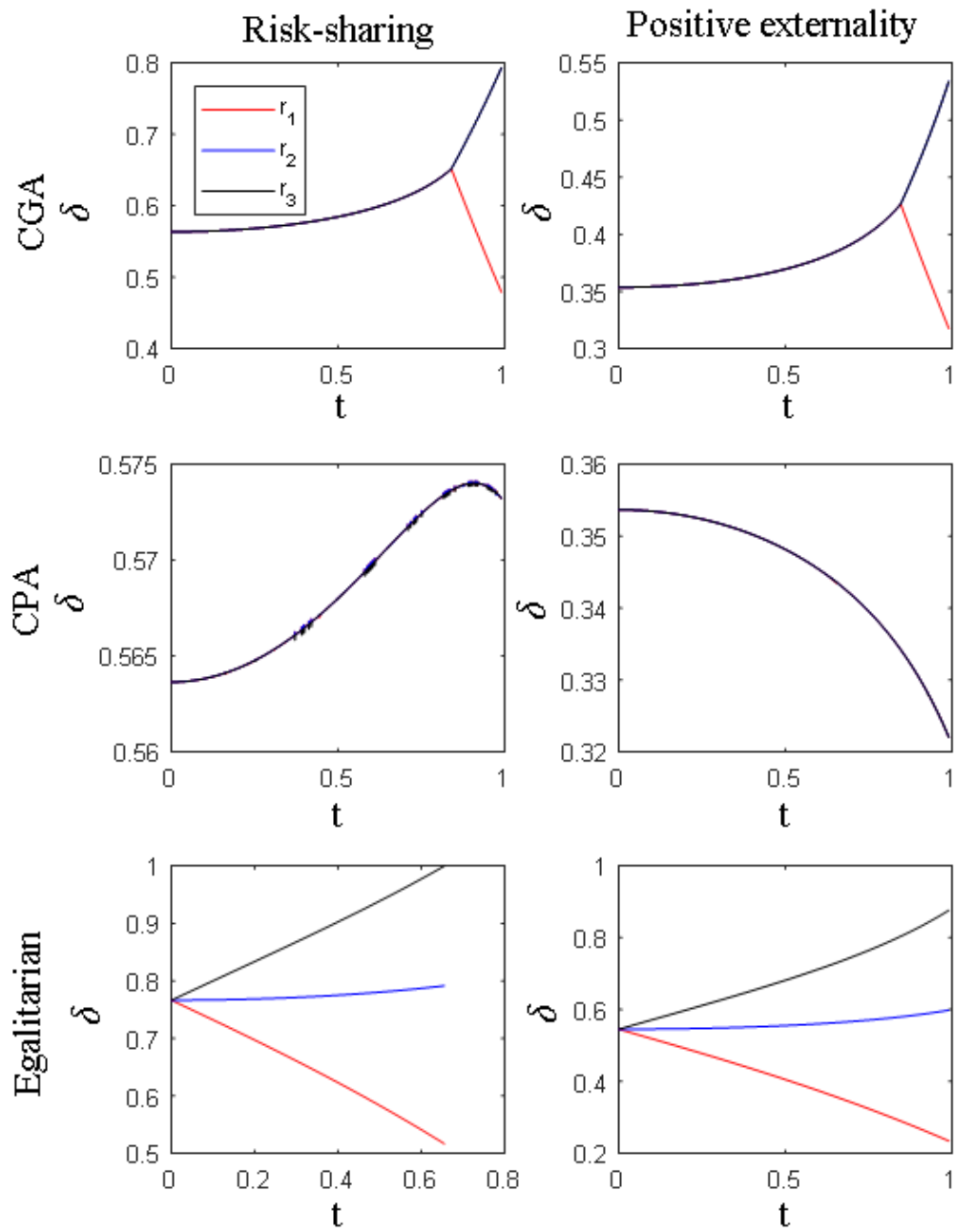
Fig. 2.4.3 Minimum δ per individual

Figure 2.4.3 focuses on the minimum δ that each individual requires. Following Theorem 2.2 part (d), common pool agreements are able to balance the incentives of all the individuals in order for them to have the same δ . Club good agreements do the same until poor individuals contribute all their income to the collective good. From this point onwards, the deviation incentives of other agents grow steeper as inequality increases, while the deviation incentives of the poorest individuals decrease. Finally, for an egalitarian agreement, rich individuals will have higher incentives to deviate, and in the risk-sharing case, the richest individual reaches $\delta = 1$ at around $t = 0.65$, which implies that egalitarian agreements are not feasible beyond this level of inequality.

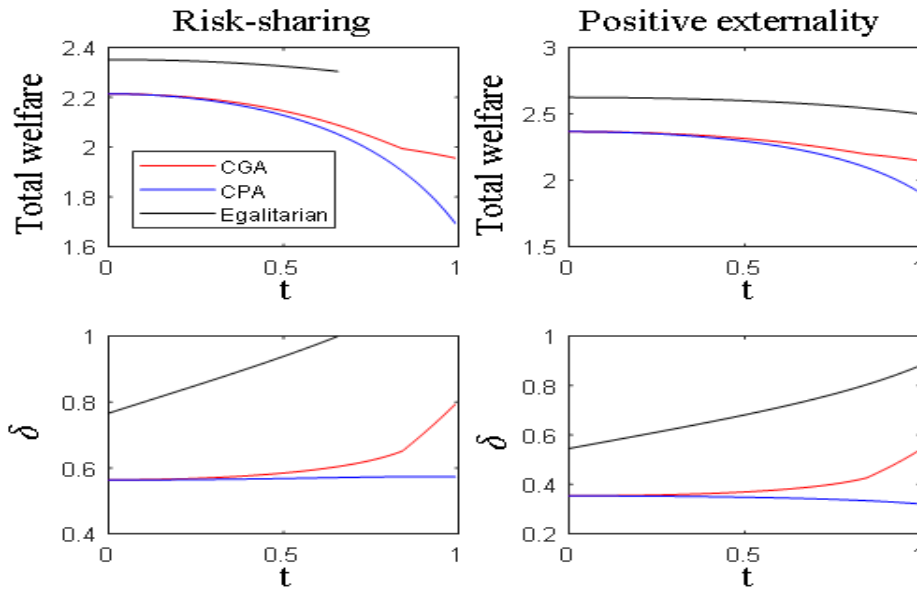


Fig. 2.4.4 Total welfare and minimum δ for the community

Finally, Figure 2.4.4 visualises Theorem 2.1 and Proposition 2.1. For both cases, the egalitarian rule produces higher total welfare for the community. However, as soon as the richest individual finds that the agreement not admissible, the egalitarian agreement stops being an option for the community. In contrast, both CPAs and CGAs are still valid options.

■

Example 2.1 displays the implications of the theorems at the abstract level. However, these results also provide new insights into the puzzles described in the introduction. First, under risk-sharing agreements, the community might avoid full risk sharing even with perfect monitoring. This result complements the theoretical framework developed by Kocherlakota

(1996) and Ligon et al. (2002) as well as the empirical observations of Townsend (1994), Udry (1994), and Fafchamps and Lund (2003). These authors explained that communities do not pool as much risk as expected, and they explained the difference via moral hazard and asymmetric information. They claimed that agents have limited information, and therefore it is not prudent to share all their income, as other agents can take advantage of the different signals that each agent receives. Complementing that approach, Theorem 2.2 proves that if an agreement was chosen for its robustness, full risk-sharing is not an option; by default, income correlates with consumption. Furthermore, Theorem 2.2 suggests that this correlation is stronger for people with higher income. Thus, this model explains the second puzzle, which asks why, in cases where egalitarian agreements are plausible, people might still follow norms that are not welfare maximising and in which their consumption is correlated with their income.

Second, people may support social norms in a community that preserves inequality. Scott (1977) explained that solidarity within peasants emerges from subsistence needs and not from optimal production solutions. Therefore, some communities are willing to accept inequality as long as it improves their living standards. Theorem 2.2 claims that when the priority is having the participation of all members of the community, inequality is the norm. In this case, richer individuals contribute to the collective good with a lower fraction of their income, and consequently, their consumption is structurally higher.

In general, robustness implies tolerance for rules that favour the richer people in society. In accordance with this conclusion, several authors such as Wade (1987), Baland and Platteau (1999), and Elbers et al. (2004) empirically found that wealthy agents are usually engaged in the creation and maintenance of collective agreements. These authors claimed that wealth differences implied bargaining power differences, and therefore, the wealthy members of a community can preserve unequal treatment in their favour. However, Theorem 2.2 provides an alternative yet complementary explanation by highlighting that even if every member has the same bargaining power, differential treatment emerges by changing an efficient-oriented agreement into a robust-oriented agreement. Thus, related to the third and fourth puzzles, unequal treatment does not need to be a consequence of power relations, but it can emerge from the objective of the community.

Third, as illustrated in Figure 2.4.1, if inequality is high enough, the poorest individual will dedicate all his income to the community. Hence, contrary to Olson (1971) theory, this model shows that there are agreements where poor people are entirely devoted to the project, and the incentives are structured in such a way that they have no deviation incentives. Indeed,

as Figure 2.4.3 shows, wealthier individuals will have more incentives to deviate than poorer individuals.

Finally, regarding the role of inequality as a barrier to or facilitator of collective agreements (i.e. the third puzzle), Theorem 2.2 and Example 2.1 provide two insights. First, robust agreements are supported by less patient individuals. Hence, inequality helps to reduce people's deviation incentives. Second, as Figure 2.4.3 explains, under egalitarian agreements, inequality increases deviation incentives. However, under common pool agreements, this effect is non-linear, and sometimes, higher inequality can reduce the minimum δ required by the community. In this way, the current model explains the paradox of the third puzzle by describing the impatience level required by the type of agreement that the community wants to implement. Thus, some level of inequality can provide the conditions for more robust agreements in which the community can participate. However, depending on the conditions of the problem, an increase in heterogeneity can also increase δ .

While robust rules provide explanations for the first three puzzles, it is not clear why a community will actively look for a policy that creates structural inequality in the consumable income of individuals. One way to answer this question is that the community may not have been able to decide. Most likely, the rule appeared, and its repetition made it a norm. Another way to approach this problem is to realise that inequality at the utility level is more critical for policy-making decisions than is inequality at the income level. Hence, the remaining portion of this section describes the performance of robust rules under the utility approach.

Definition 2.5 (Equality promotion policies) *Let EU_i be the expected utility of individual i . Consider a policy P such that the expected utility after implementation is EU'_i . Then, P promotes equality if $\forall i, j$:*

$$\frac{\max\{EU_i, EU_j\}}{\min\{EU_i, EU_j\}} \geq \frac{\max\{EU'_i, EU'_j\}}{\min\{EU'_i, EU'_j\}}$$

Although there are many definitions of inequality, Definition 2.5 provides a widely accepted measure of a reduction in inequality that is useful to explain the effects of robust rules on inequality at the utility level.

Proposition 2.2 *In CGAs and CPAs, robust agreements are policies that promote equality.*

It is evident that efficient agreements promote the highest equality levels, as members of the community have the same expected utility. Proposition 2.2 shows that even though robust agreements preserve structural inequality at the income level, they also promote equality at

the utility level. Furthermore, because they are an equilibrium of the game, they are better than the state of no cooperation. Hence, these rules also improve the living conditions of each member. Thus, robustness is an alternative explanation for the fact that even with heterogeneous agents and limited commitment contracts, income inequality is higher than consumption inequality (Krueger and Perri, 2006). Furthermore, this result lends theoretical support to one of the main observations by James Scott about the peasants in South-East Asia: “Village egalitarianism in this sense is not radical; it claims that all should have a place, a living, not that all should be equal” (Scott, 1977, p.74).

Together with Theorem 2.2, Proposition 2.2 reveals the central mechanism that induces stability among these agreements. Robust rules manage to balance rent-seeking behaviour from the rich (i.e. their unequal treatment) with redistribution incentives from the poor (i.e. the inequality reduction). This result is capable of explaining the empirical observations of Marchiori (2014), who suggested that collective agreements are affected by two forces: effort augmenting and effort mixing. The effort-augmenting force is associated with agreements that have differentiated rules that favour some agents in society, and the effort-mixing force is associated with the need for balance in agreements. Thus, Theorem 2.2 and Proposition 2.2 conceptualise these forces.

This section illustrated how robust rules produce an alternative explanation for the fourth puzzle. In addition, robust rules are equilibria of the game in more scenarios than efficient rules are. Moreover, these rules can appear in communities where decision-makers have no information about the δ of their members or under scenarios that δ can suffer a shock. Hence, it is reasonable to argue that robust rules are likely to become norms in a community.

Theoretical models are not sufficient to learn about the generation of norms in a community. History-dependent events can create an infinite number of rules. Nevertheless, due to their favourable characteristics, robust agreements are likely to emerge in a society. In conclusion, this section explained within a unified framework how the robust approach provides insights into the four puzzles, what the mechanisms are underlying these rules, and why these rules are observed in the field.

2.5 CGAs vs. CPAs

Sometimes the community cannot decide on the type of agreement it wants to adopt. For example, even if a good has a rival nature, it does not mean that it is socially acceptable for some people to receive higher benefits than others. However, assuming that both agreements

are possible, understanding their differences can provide a better understanding of the mechanisms underlying them.

The criteria to compare the agreements are changes in the deviation incentives, changes in the overall welfare, and changes in the individual welfare. Proposition 2.3 presents these comparisons.

Proposition 2.3 *Consider an infinitesimal mean-spread deviation of a homogeneous population where $\forall i : r_i = c + \Delta r_i$ and $\sum_{i=1}^N \Delta r_i = 0$, for a fixed $c > 0$. Then:*

(a) *CPAs are subgame perfect equilibria for lower δ than CGAs if and only if:*

$$FT_i\left(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I}\right)\left(\frac{dM_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})}{da}\right) \\ \leq M_i\left(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I}\right)\mathbb{E}[Y_i u'((1 - \frac{k}{cn})Y_i + g(\frac{k}{cn}\sum_{j=1}^N Y_j))(ng'(\frac{k}{cn}\sum_{j=1}^N Y_j) - 1)].$$

(b) *The overall welfare of CPAs does not change, while for CGAs, the change is non-negative.*

(c) *In CPAs, the change in the future expected trade-off of all the members with $r_i \geq c$ is non-positive, and for all members with $r_i \leq c$, the change is non-negative. In CGAs, the change of the future expected trade-off is non-positive for all members with $r_i \geq c$ and some members with $r_i < c$.*

(d) *In CPAs, the change in the expected utility of all members with $r_i \geq c$ is non-negative and for all the members with $r_i \leq c$, the change is non-positive. In CGAs, the change in the expected utility is non-positive for all members with $r_i \leq c$ and some members with $r_i > c$.*

(e) *In risk-sharing scenarios (i.e. $g(x) = x/n$) both models have the same behaviour for parts (c) and (d).*

For a better understanding of the mechanisms underlying Proposition 2.3, Example 2.1 is continued, and each graph is described in terms of the theoretical result.

Example 2.2 (CARA agents; part 2)

This example is the continuation of Example 2.1. Figure 2.4.4 shows that robust rules for club good agreements provide higher total welfare but are bounded by higher δ . This result

is explained by Proposition 2.3 parts (a) and (b). Part (a) explains that in many conditions, and in particular for risk-sharing agreements ($ng'(\frac{k}{cn}\sum_{j=1}^N Y_j) - 1 = 0$), club goods require higher values of δ (i.e. are less robust), while part (b) shows that their total welfare is higher.

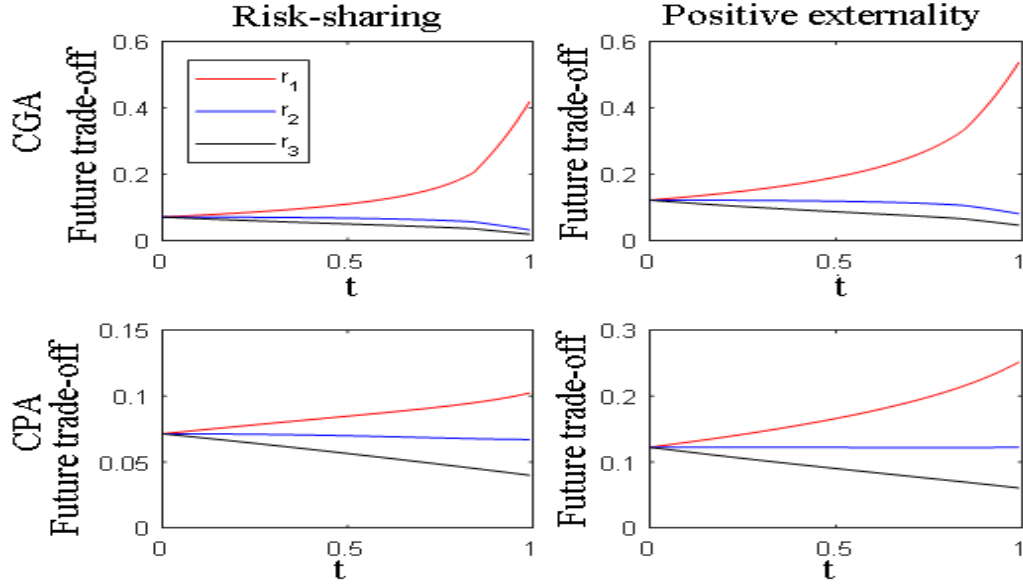


Fig. 2.5.1 Future trade-off

Complementing this analysis, Figures 2.5.1 and 2.5.2 represent Proposition 2.3 part (c) and (d), respectively. Figure 2.5.1 shows that as inequality increases, the future trade-off improves for the poorest individual, while it deteriorates for the richest individual. In a seemingly contradictory result, Figure 2.5.2 reveals that the expected utility declines for the poorest individual and increases for the richest. This paradox is the essence of the rent-seeking vs. re-distribution trade-off that the robust rules balance.

These four results describe the differences between the incentive mechanisms underlying the two types of collective action problems. Club good agreements create *ex-ante* unequal treatment (contribution side), while common pool agreements create *ex-post* unequal treatment (distribution side). This difference allows common pool agreements to be more effective at adjusting the incentives of each member of the community and a lower δ to be obtained. By contrast, club good agreements set the contribution levels according to the incentives of each individual; thus, individual utilities are less penalized by the contributions. Both agreements depict that the poorest individual is compensated with a better future trade-off and that the richest individual is compensated with a better expected utility. Therefore, poor people participate in the agreement because it is better than staying in autarky, and this benefit increases as inequality increases (due to re-distribution benefits). For rich people, as they

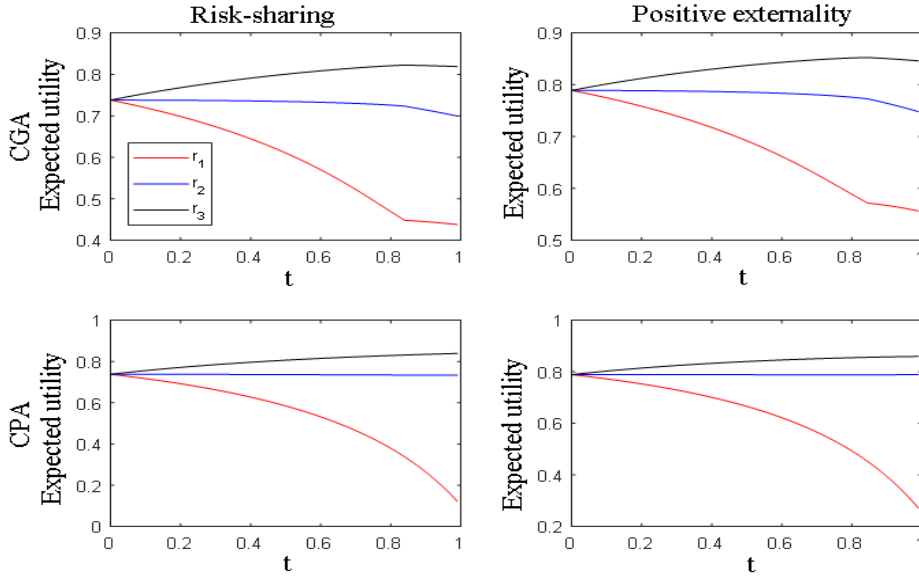


Fig. 2.5.2 Expected utility

become richer, their future trade-off is less compelling, but the agreements offer them higher expected utility to incentivize their participation (rent-seeking incentives). Common pool agreements divide rich and poor depending on their difference from the average individual. Instead, club good agreements provide fewer incentives to average individuals, except for the risk-sharing case, as Proposition 2.3 part (e) claims.

■

Proposition 2.3 highlights the incentive structure that robust rules use to promote participation. For example, for low inequality levels, robust rules preserve the total welfare as agreements become more robust. Regarding the third puzzle, the previous idea explains that inequality can be beneficial for agreements because it reduces the deviation incentives (i.e. reduces the δ required to coordinate) while still providing the same welfare level. However, as heterogeneity increases, this result becomes weaker as non-linear effects appear.

In conclusion, Section 2.4 and Section 2.5 showed that robust rules provide new insights that can explain the four puzzles stated in the introduction and why such rules are likely to be present in communities. Finally, there is no absolute advantage between common pool and club good agreements. Deciding between the two is equivalent to deciding between robustness and efficiency. Under general conditions, club good agreements generate higher total welfare but are less robust than common pool agreements.

2.6 Discussion and Future Work

This section discusses three topics related to the use of robust agreements to describe real scenarios: i) the prevalence of robust agreements in society, ii) the veracity of group coordination in network contexts, and iii) the effect of the market structure on the stability of agreements. This section then concludes with suggestions for future extensions of the current model.

Are robust agreements common in society?

The previous sections proved that robust rules explain several empirical observations. Given that these agreements require lower patience levels at equilibrium, they should be more common than egalitarian agreements: heterogeneous agents are more likely to deviate from agreements that treat everyone equally (Abramitzky, 2008). Nonetheless, as stated in the third puzzle, heterogeneity sometimes works as a barrier to the creation of agreements.

While robust agreements explain the different scenarios that appear in reality, the theoretical construction is not able to determine whether the barrier effect is stronger than the facilitator effect without a specific context. However, experimental and field work provide some additional insights into the previous puzzle. In a review of experiments in this field, Cardenas (2005) showed that it was challenging for individuals to postulate differential treatment rules. As such, unless they emerge naturally, they are not the first option for policy-makers. Additionally, Baland and Platteau (2007) and Ruttan (2008) provided some examples of successful unequal treatment agreements. However, the authors emphasised that rules are community-specific and that their implementation is unlikely to be transferable to other societies. In theory, while differential treatment agreements should be regularly observed, it is still essential to understand the context of each community to determine whether such treatment is plausible for its members.

Is group coordination a realistic assumption in network contexts?

Both club good and common pool agreements are mechanisms by which a community coordinates towards a common goal. This assumption is realistic when society is working towards a common good (e.g. an irrigation system (Bardhan, 2000), a common field (Wade, 1987), or a common forest (Naidu, 2009)). However, authors such as Bates (1990), Fafchamps and Lund (2003), Bramoullé and Kranton (2007), and Bloch et al. (2008) have suggested that in cases with risk-sharing and other resource-pooling mechanisms, the relationship is not

with the community but between individuals. Therefore, the shares that individuals receive depend on their position in the network. Nonetheless, authors such as Jackson et al. (2012) and Ambrus et al. (2014) have explained that both in theory and in the field, the most stable networks have members that have strong regular connections with all the other members of the same cluster. Therefore, under models such as that of Bramoullé and Kranton (2007), where the agents cannot distinguish who receives the benefit, a regular network implies that all individuals receive the same share of the common good, i.e. CGAs.

The previous argument supports the fact that club good agreements are more common than common pool agreements. Furthermore, the expected even distribution of benefits in such network agreements explains why different communities tolerate people that hide their “share-able income”. For example, Bates (1990), Fafchamps (1992), and Di Falco and Bulte (2011) documented cases in which some community members found ways to reduce the resources they were pooling, such as buying non-shareable assets, placing only a portion of their income in the pool, or hiding a proportion of their income. Implicitly, these practices are equivalent to a robust rule in club good agreements. Hence, even in agreements ruled by approximately regular network interactions, it is possible to observe club good agreements in the field. In contrast, common pool agreements require community coordination and do not seem to be feasible in network contexts.

Are robust agreements stable vis-à-vis markets?

The rise of markets plays a central role in the collective agreements literature. Munshi and Rosenzweig (2007) gathered empirical evidence on how new market structures increase individuals’ outside options, thus making them more prone to deviate. Continuing on this topic, Gagnon and Goyal (2017) studied how the nature of the interaction between the market and community agreements (whether a social agreement complements or substitutes for market effects) can affect their stability. From the lens of the “substitution effect”, individuals who benefit less from the community are prone to leave. In another analysis, Banerjee and Newman (1998) argued that in heterogeneous communities, both the richest and the poorest individuals are more likely to leave the network and take higher risks in the market because markets improve the outside option for the richest and offer new opportunities for the poorest.

As these models suggest, markets increase the outside options for some agents within the community. If a shock is unexpected, the effect of the market represents a decrease in the future trade-off or a decrease in the worst-case present trade-off. Theorem 2.2 implies that robust agreements resist these shocks better than any other rule. Alternatively, if the

community anticipates a shock generated by the market and can renegotiate the agreement, the new agreement is likely to improve the treatment of agents who have better outside options. Hence, heterogeneous communities that manage to preserve their collective goods are expected to exhibit unequal treatment after a market shock. While the relationship between market interactions and communities might not be causal, it does reflect a survivorship bias associated with a preference for robustness in the community.

Future work

In the search for a parsimonious model, the current paper strongly relied on three premises: independent income distributions, simple rules, and perfect information among community members.

The independence among individuals in the distribution of income suggests that income follows a Poisson-like process across individuals. This assumption is realistic for multiple cases. For example, in agricultural villages where the land is homogeneous and the amount of grain obtained is proportional to the amount of rain that the crop receives, the Poisson process of rainfall guarantees the independence of production between individuals. Thus, this assumption can be reasonable for spatially closed communities, such as those observed in Fafchamps and Gubert (2007). Another way to obtain independence among individuals occurs when the people in the agreement are strategically chosen to have independent sources of income. Luke and Munshi (2006) presented an example of this situation. In their study, communities encourage exogamous marriages so that the family linkages are not correlated. Hence, individual independence is a valid assumption for many scenarios. In contrast, time independence is a stronger assumption. For example, if the income of a hunter-gatherer community is the amount of game collected, the number of animals around the village is likely to present a serial correlation. Thus, it is appropriate to extend the current work to cases where income is serially correlated because such cases might offer different insights regarding the income stream behaviour and history-dependent agreements in a heterogeneous agent context.

Regarding simple rules, because it is already difficult for individuals to construct unequal treatment rules under experimental conditions, more sophisticated rules will be very unlikely to emerge in reality (Kalai and Stanford, 1988). However, as Kocherlakota (1996) and Ligon et al. (2002) showed, history and state-contingent agreements can improve the efficiency of agreements. Hence, it is worth studying the robustness of agreements under more elaborate

rules that can provide better insights into history and state-contingent contracts and how they are affected by a robustness approach.

Finally, with respect to perfect monitoring, it is plausible that everyone in a small community is aware of the realities of other members. However, as the population size increases, this assumption becomes less likely. In that case, the network perspective mentioned previously becomes more relevant. Even if the network is regular, the flow of information might only be local, as in Bloch et al. (2008). Therefore, future work should include the network component to help explain how the incentives of community members to monitor each other allow agreements to exist.

2.7 Conclusions

The economic theory of community agreements has been heavily influenced by a utilitarian approach in which ideal policies maximise the overall welfare of the community. However, these are not the only rules that can emerge in communities. The results revealed in the previous sections present a case for policies that follow a Rawlsian approach in which communities develop rules that reduce the deviation incentives for individuals who are more likely to deviate from collective agreements. This change of perspective provides a unified framework to explain some of the central puzzles that link inequality and collective agreements.

Robust agreements are not only theoretical constructions: they do exist in many communities. However, in contrast to other mechanisms such as equally distributed full risk-sharing arrangements, robust agreements are difficult to generalise because they are context-specific. Nevertheless, the primary results from this paper can be extended to guide policies that balance the trade-off between stability and efficient agreements.

Appendix to Chapter 2

Appendix 2.A Math Notation

This section defines the technical notation that the paper uses in the main text and in the proofs. The notation used in the proofs of Section 2.B, that is not mentioned in this section, was explicitly defined in the main text.

- $\mathbb{R}_+^n$ is an n -dimensional vector where each component is in $[0, \infty)$.
- $\mathbb{O}$ is a zero vector of appropriate size.
- $\mathbb{I}$ is a one vector of appropriate size.
- Δ^n is an n -dimensional simplex.
- $\mathbb{E}$ is the expected value function.
- If A is a matrix, A^T is its transpose.

Appendix 2.B Proofs

Lemmas:

Lemma 2.1 *If $r_i > r_j$, then Y_i first orders stochastic dominates Y_j .*

Proof:

The cumulative probability function of Y_i and Y_j are $F(\frac{y_i}{r_i})$ and $F(\frac{y_j}{r_j})$. As F is increasing, $\forall x \in [0, \infty)$, $F(\frac{x}{r_i}) \leq F(\frac{x}{r_j})$. Hence, Y_i first orders stochastic dominates Y_j .

■

Lemma 2.2 $PT_i(y_t, \alpha, \beta)$ has a lower bound.

Proof:

By definition, u is bounded. Let M be a constant such that $\forall x : M > u(x)$. Then, for every individual i , it holds that $PT_i(y_t, \alpha, \beta) > -M$. ■

Lemma 2.3 Let $h_1(x)$ be an increasing convex function such that $h_1(x) = 0 \Leftrightarrow x = 0$ and let $h_2(x)$ be a positive concave function such that $h_2(0) = 0$. Then, $\frac{h_2(x)}{h_1(x)}$ is decreasing for $x > 0$.

Proof:

From the conditions of h_1 , for all $\lambda \in (0, 1)$, $\lambda h_1(x) \geq h_1(\lambda x) \geq 0$, and from the conditions of $h_2(x)$, for all $\lambda \in (0, 1)$, $0 \leq \lambda h_2(x) \leq h_2(x)$. Thus, $\frac{h_2(\lambda x)}{h_1(\lambda x)} \geq \frac{h_2(x)}{h_1(x)}$. ■

Lemma 2.4 Let $h(x, y)$ be a continuous function and define $h_x(x, y) = \frac{dh(x, y)}{dx}$ and $h_y(x, y) = \frac{dh(x, y)}{dy}$. Then,

$$\frac{d\mathbb{E}[h(Y_i, r_i)]}{dr_i} = \frac{\mathbb{E}[Y_i h_x(Y_i, r_i)]}{r_i} + \mathbb{E}[h_y(Y_i, r_i)]$$

Proof:

$$\begin{aligned} \frac{d\mathbb{E}[h(Y_i, r_i)]}{dr_i} &= \int_0^\infty \frac{dh(y_i, r_i)}{dr_i} \frac{1}{r_i} f\left(\frac{y_i}{r_i}\right) dy_i \\ &= \int_0^\infty h_y(y_i, r_i) \frac{1}{r_i} f\left(\frac{y_i}{r_i}\right) dy_i + \int_0^\infty h(y_i, r_i) \frac{d\left(\frac{1}{r_i} f\left(\frac{y_i}{r_i}\right)\right)}{dr_i} dy_i \\ &= \mathbb{E}[h_y(Y_i, r_i)] + \lim_{T \rightarrow \infty} h(y_i, r_i) \frac{dF\left(\frac{y_i}{r_i}\right)}{dr_i} \Big|_{y_i=0}^T - \int_0^\infty h_x(y_i, r_i) \frac{dF\left(\frac{y_i}{r_i}\right)}{dr_i} dy_i \\ &= \mathbb{E}[h_y(Y_i, r_i)] + \frac{1}{r_i} \int_0^\infty y_i h_x(y_i, r_i) \frac{1}{r_i} f\left(\frac{y_i}{r_i}\right) dy_i = \frac{\mathbb{E}[Y_i h_x(Y_i, r_i)]}{r_i} + \mathbb{E}[h_y(Y_i, r_i)] \end{aligned}$$
■

Theorems:**Proof of Theorem 2.1:**

The first direction of the proof: “if the efficient rule for the CGAs and CPAs is $(\mathbb{I}, \frac{1}{n}\mathbb{I}) \Rightarrow FT_N(\mathbb{I}, \frac{1}{n}\mathbb{I}) > 0$ ” is trivial: The counterreciprocal states that if $FT_N(\mathbb{I}, \frac{1}{n}\mathbb{I}) \leq 0$, then the rule $(\mathbb{I}, \frac{1}{n}\mathbb{I})$ is not admissible because the wealthiest individual has incentives to deviate. Hence, the rule cannot be efficient.

The proof of the other direction is independently done for each model.

CGA:

Step 1: Define $\mathcal{F}(a) = \sum_{i=1}^N FT_i(a, \frac{1}{n}\mathbb{I})$ and prove that $\mathcal{F}(a)$ is concave.

This proof requires two parts. The first part proves that for all y_t , and for every individual i , $\mathcal{C}_i(a, y) = u(\Gamma(y, a, \frac{1}{n}\mathbb{I})_i) - u(y_i)$ is concave. Once this is proven, the second part states that $\mathcal{F}(a) = \sum_{i=1}^N \mathbb{E}[\mathcal{C}_i(a)]$ is concave because the expected value preserves concavity on a .

Without loss of generality, the first part of the proof is done for $i = 1$. The proof for the other values of i is identical.

Let y_{-1} be a vector of $(n-1) \times 1$ constructed by removing the first element of y . Then,

$$H(-\mathcal{C}_1(a, y)) = \begin{bmatrix} A_{1 \times 1} & C_{1 \times (n-1)} \\ C^T & B_{(n-1) \times (n-1)} \end{bmatrix}$$

where H is the Hessian function, and

$$A = - \left(u'' \left(\Gamma(y, a, \frac{1}{n}\mathbb{I})_1 \right) \left(g' \left(\sum_{j=1}^N a_j y_j \right) - 1 \right)^2 + g'' \left(\sum_{j=1}^N a_j y_j \right) u' \left(\Gamma(y, a, \frac{1}{n}\mathbb{I})_1 \right) \right) y_1^2$$

$$B = - \left(u'' \left(\Gamma(y, a, \frac{1}{n}\mathbb{I})_1 \right) \left(g' \left(\sum_{j=1}^N a_j y_j \right) \right)^2 + u' \left(\Gamma(y, a, \frac{1}{n}\mathbb{I})_1 \right) g'' \left(\sum_{j=1}^N a_j y_j \right) \right) y_{-1} (y_{-1})^T$$

$$C_{1 \times j} =$$

$$- y_1 y_{j+1} \left(u'' \left(\Gamma(y, a, \frac{1}{n}\mathbb{I})_1 \right) \left(g' \left(\sum_{j=1}^N a_j y_j \right) - 1 \right) g' \left(\sum_{j=1}^N a_j y_j \right) + g'' \left(\sum_{j=1}^N a_j y_j \right) u' \left(\Gamma(y, a, \frac{1}{n}\mathbb{I})_1 \right) \right)$$

Given that u and g are concave, $A \geq 0$ and B is positive semidefinite. Finally, using row operations,

$$\det(H(-\mathcal{C}_1(a, y))) = \det \begin{bmatrix} D_{1 \times 1} & \mathbb{O}_{1 \times (n-1)} \\ C^T & B_{(n-1) \times (n-1)} \end{bmatrix}$$

where

$$D = - \frac{u''\left(\Gamma(y, a, \frac{1}{n}\mathbb{I})_1\right) g''(\sum_{j=1}^N a_j y_j) u'\left(\Gamma(y, a, \frac{1}{n}\mathbb{I})_1\right) \left(2g'(\sum_{j=1}^N a_j y_j) - 1\right)^2 y_1^2}{u''\left(\Gamma(y, a, \frac{1}{n}\mathbb{I})_1\right) \left(g'(\sum_{j=1}^N a_j y_j)\right)^2 + u'\left(\Gamma(y, a, \frac{1}{n}\mathbb{I})_1\right) g''(\sum_{j=1}^N a_j y_j)} \geq 0.$$

Thus, as B and D are positive semidefinite, then $H(-\mathcal{C}_1(a, y_t))$ is positive semidefinite. Ergo, $\mathcal{C}_1(a, y)$ is concave in a . Thus, $\mathcal{F}(a)$ is concave.

Step 2:

Notice that, $FT_N(a, \frac{1}{n}\mathbb{I}) > 0$ and, because Y_N first order stochastically dominates Y_i , $\mathbb{E}[u(Y_N) - u(Y_i)] \geq 0$. Thus, $FT_i(a, \frac{1}{n}\mathbb{I}) = FT_N(a, \frac{1}{n}\mathbb{I}) + \mathbb{E}[u(Y_N) - u(Y_i)] > 0$.

Consider the first order derivatives of $\mathcal{F}(a)$ evaluated in $a = \mathbb{I}$.

$$\begin{aligned} \frac{\partial \mathcal{F}(\mathbb{I})}{\partial a_i} &= \mathbb{E} \left[Y_i \left(\sum_{j=1}^N \left(u' \left(g \left(\sum_{j=1}^N Y_j \right) \right) g' \left(\sum_{j=1}^N Y_j \right) \right) - u' \left(g \left(\sum_{j=1}^N Y_j \right) \right) \right) \right] \\ &= \mathbb{E} \left[Y_i u' \left(g \left(\sum_{j=1}^N Y_j \right) \right) \left(n g' \left(\sum_{j=1}^N Y_j \right) - 1 \right) \right] \geq 0 \end{aligned}$$

From step 1, $\mathcal{F}(a)$ is concave. Moreover, $a = \mathbb{I}$ satisfies all the Karush-Kuhn-Tucker (KKT) conditions (i.e. $\sum_{i=1}^N a_i r_i \geq k$, $\forall i$ $0 \leq a_i \leq 1$ $FT_i(a, \frac{1}{n}\mathbb{I}) > 0$, and $\frac{\partial \mathcal{F}(\mathbb{I})}{\partial a_i} \geq 0$). Then, $\mathcal{F}(a)$ is a global maximum. Ergo, $(\mathbb{I}, \frac{1}{n}\mathbb{I})$ is efficient.

CPA:

Let $\tau = 1$ and consider $p \neq \frac{1}{n}\mathbb{I}$ and a permutation $\pi : \{1, \dots, N\} \rightarrow \{1, \dots, N\}$ such that for all $i > j$, $p_{\pi(i)} \geq p_{\pi(j)}$. Hence, for all $j < N$ it is true that $\sum_{i=1}^j p_{\pi(i)} < \frac{j}{n}$ and $\sum_{i=1}^N p_{\pi(i)} = 1$.

Given that u is concave, Karamanta inequality states that for all y

$$\begin{aligned} \sum_{i=1}^N u\left(p_i n g\left(\sum_{j=1}^N y_j\right)\right) - \sum_{i=1}^N u\left(\frac{1}{n} n g\left(\sum_{j=1}^N y_j\right)\right) &< 0 \\ \Rightarrow \mathbb{E}\left[\sum_{i=1}^N u\left(p_i n g\left(\sum_{j=1}^N Y_j\right)\right) - \sum_{i=1}^N u\left(\frac{1}{n} n g\left(\sum_{j=1}^N Y_j\right)\right)\right] &< 0 \\ \Rightarrow \sum_{i=1}^N FT_i(\mathbb{I}, p) &< \sum_{i=1}^N FT_i(\mathbb{I}, \frac{1}{n}\mathbb{I}) \end{aligned}$$

Finally, note that, as in the CGA, $FT_i(\mathbb{I}, \frac{1}{n}\mathbb{I}) = FT_N(\mathbb{I}, \frac{1}{n}\mathbb{I}) + \mathbb{E}[u(Y_N) - u(Y_i)] > 0$. Thus, $\frac{1}{n}\mathbb{I}$ both maximizes the objective function and satisfies KKT conditions. Ergo, $(\mathbb{I}, \frac{1}{n}\mathbb{I})$ is efficient. ■

Proof of Theorem 2.2:

This proof relies in the characteristics of $M_i(\alpha, \beta)$. Therefore, these characteristics are presented and then the theorem is proven.

By the envelope theorem,

$$\begin{aligned} \frac{dM_i(\alpha, \beta)}{d\alpha_i} &= y_i^* u' \left(y_i^* - \alpha_i y_i^* + \beta_i n g(\alpha_i y_i^*) \right) \left(\beta_i n g'(\alpha_i y_i^*) - 1 \right) < 0 \\ \frac{d^2 M_i(\alpha, \beta)}{d\alpha_i^2} &= (y_i^*)^2 \left(u'' \left(y_i^* - \alpha_i y_i^* + \beta_i n g(\alpha_i y_i^*) \right) \left(\beta_i n g'(\alpha_i y_i^*) - 1 \right)^2 \right) \\ &\quad + (y_i^*)^2 \left(u' \left(y_i^* - \alpha_i y_i^* + \beta_i n g(\alpha_i y_i^*) \right) \beta_i n g''(\alpha_i y_i^*) \right) < 0 \end{aligned}$$

where y_i^* is the income that minimises $PT(y, \alpha, \beta)$. Furthermore, is clear that $M_i(\alpha, \beta) \leq 0$, and $M_i(\alpha, \mathbb{O}) = 0$.

From the definition of $M_i(\alpha, \beta)$, if $\alpha_i = \alpha_j$ and $\beta_i = \beta_j$ then $M_i(\alpha, \beta) = M_j(\alpha, \beta)$. Moreover, if α is parametrized as an increasing linear function of t ,

$$\begin{aligned} \frac{d^2 FT_i(\alpha, \beta)}{dt^2} &= \mathbb{E} \left[u' \left(Y_i - \alpha_i Y_i + \beta \text{ing} \left(\sum_{k=1}^N \alpha_k Y_k \right) \right) \beta \text{ing}'' \left(\sum_{k=1}^N \alpha_k Y_k \right) \left(\sum_{k=1}^N Y_k \frac{d\alpha_k}{dt} \right)^2 \right] \\ &+ \mathbb{E} \left[u'' \left(Y_i - \alpha_i Y_i + \beta \text{ing} \left(\sum_{k=1}^N \alpha_k Y_k \right) \right) \left(\beta \text{ing}' \left(\sum_{k=1}^N \alpha_k Y_k \right) \left(\sum_{k=1}^N Y_k \frac{d\alpha_k}{dt} \right) - Y_i \frac{d\alpha_i}{dt} \right)^2 \right] \leq 0 \end{aligned}$$

From the definition of M , it is also clear that $FT_i(\mathbb{O}, \beta) = 0$. Using these results, the proof of the theorem goes as follows.

Part a:

Assume that $r_i > r_j$ and $a_i = a_j < 1$. From the smoothing condition,

$$\begin{aligned} \left| \mathbb{E} \left[u \left(\Gamma(Y, a, \frac{1}{n} \mathbb{I})_i \right) - u \left(\Gamma(Y, a, \frac{1}{n} \mathbb{I})_j \right) \right] \right| &< \left| \mathbb{E} \left[u \left(\Gamma(Y, \mathbb{O}, \frac{1}{n} \mathbb{I})_i \right) - u \left(\Gamma(Y, \mathbb{O}, \frac{1}{n} \mathbb{I})_j \right) \right] \right| \\ &= \left| \mathbb{E}[u(Y_i)] - \mathbb{E}[u(Y_j)] \right| \end{aligned}$$

Given that Y_i first order stochastically dominates Y_j , then $\mathbb{E}[u(Y_i)] > \mathbb{E}[u(Y_j)]$. Therefore,

$$\begin{aligned} FT_i(a, \frac{1}{n} \mathbb{I}) - FT_j(a, \frac{1}{n} \mathbb{I}) &= \\ \mathbb{E} \left[u \left(\Gamma(Y, a, \frac{1}{n} \mathbb{I})_i \right) - u \left(\Gamma(Y, a, \frac{1}{n} \mathbb{I})_j \right) \right] - \mathbb{E}[u(Y_i)] + \mathbb{E}[u(Y_j)] &< 0 \end{aligned}$$

Thus, $\min_{k \in \{1, \dots, N\}} \left\{ -\frac{FT_k(a, \frac{1}{n} \mathbb{I})}{M_k(a, \frac{1}{n} \mathbb{I})} \right\} \neq -\frac{FT_j(a, \frac{1}{n} \mathbb{I})}{M_j(a, \frac{1}{n} \mathbb{I})}$, and given that $FT_j(a, \frac{1}{n} \mathbb{I}) > FT_i(a, \frac{1}{n} \mathbb{I})$, $a_j < 1$, and $-M_i(a, \frac{1}{n} \mathbb{I}) = -M_j(a, \frac{1}{n} \mathbb{I}) > 0$, a marginal increase in a_j will increase the utility of all the others without decreasing the utility of $FT_j(a, \frac{1}{n} \mathbb{I})$ enough for it to become the minimum or 0. Thus $(a, \frac{1}{n} \mathbb{I})$ was not a robust rule.

For the case where $a_j \neq a_i$, assume that $a_j < a_i$. Notice that $-\frac{FT_j(a, \frac{1}{n} \mathbb{I})}{M_j(a, \frac{1}{n} \mathbb{I})}$ is decreasing in a . Hence, consider two rules $(a, \frac{1}{n} \mathbb{I})$, and $(a', \frac{1}{n} \mathbb{I})$, and where for all $k \neq i$ it holds that $a'_k = a_k$ and $a'_i = a_j$. In that case,

$$-\frac{FT_i(a, \frac{1}{n} \mathbb{I})}{M_i(a, \frac{1}{n} \mathbb{I})} < -\frac{FT_i(a', \frac{1}{n} \mathbb{I})}{M_i(a, \frac{1}{n} \mathbb{I})} < -\frac{FT_j(a, \frac{1}{n} \mathbb{I})}{M_j(a, \frac{1}{n} \mathbb{I})}$$

Henceforth, reducing the contribution of j improves the ratio for all the other individuals and does not reduce the minimum ratio. Hence, $(a, \frac{1}{n} \mathbb{I})$ cannot be robust.

Finally, consider a robust rule $(a, \frac{1}{n}\mathbb{I})$ and assume that $k < \sum_{k=1}^N \alpha_k r_k$. Define $\alpha_i(t) = a_i t$. From the proof of Theorem 2.1, for all i , $FT_i(a, \frac{1}{n}\mathbb{I}) > 0$ and for the range $t = [0, 1]$, $FT_i(\alpha(t), \frac{1}{n}\mathbb{I})$ is positive, concave, and it is only 0 when $t = 0$. Moreover, based on the results at the beginning of this proof, $-M_i(\alpha(t), \frac{1}{n}\mathbb{I})$ is increasing and convex, being 0 only when $t = 0$. Hence, from Lemma 2.3, $-\frac{FT_i(\alpha(t), \frac{1}{n}\mathbb{I})}{M_i(\alpha(t), \frac{1}{n}\mathbb{I})}$ is decreasing in t . Thence, $(a, \frac{1}{n}\mathbb{I})$ cannot be a robust rule because the rule $(\alpha(t^*), \frac{1}{n}\mathbb{I})$ where t^* is such that $k = \sum_{k=1}^N \alpha_k(t^*) r_k$ has a higher value of $-\frac{FT_i(\alpha(t^*), \frac{1}{n}\mathbb{I})}{M_i(\alpha(t^*), \frac{1}{n}\mathbb{I})}$. Contradiction.

Part b:

By replicating the last part of the proof of part (a) it is enough to prove that if $(a, \frac{1}{n}\mathbb{I})$ is robust then it is not possible that $a_i r_i = a_j r_j$. As in the previous case, this is proven by contradiction.

Step 1:

Given that $M_j(a, \frac{1}{n}\mathbb{I}) > M_i(a, \frac{1}{n}\mathbb{I})$, it is sufficient to prove that $FT_i(a, \frac{1}{n}\mathbb{I}) \geq FT_j(a, \frac{1}{n}\mathbb{I})$ to conclude that $\min_{k \in \{1, \dots, N\}} \left\{ -\frac{FT_k(a, \frac{1}{n}\mathbb{I})}{M_k(a, \frac{1}{n}\mathbb{I})} \right\} \neq -\frac{FT_i(a, \frac{1}{n}\mathbb{I})}{M_i(a, \frac{1}{n}\mathbb{I})}$.

Once this is proven, and given that $a_i < a_j \leq 1$, a marginal decrease in a_i increases $\min_{k \in \{1, \dots, N\}} \left\{ -\frac{FT_k(a, \frac{1}{n}\mathbb{I})}{M_k(a, \frac{1}{n}\mathbb{I})} \right\}$, proving that a was not robust.

Step 2: $FT_j(a, \frac{1}{n}\mathbb{I}) \leq FT_i(a, \frac{1}{n}\mathbb{I})$

Integrating by parts,

$$FT_i(a, \frac{1}{n}\mathbb{I}) = \lim_{y_i \rightarrow \infty, y_j \rightarrow \infty} \mathbb{E} \left[(u(\Gamma(Y, a, \frac{1}{n}\mathbb{I}))) \middle| Y_i = y_i, Y_j = y_j \right] + T(a_i, a_j) - \mathbb{E} [u(Y_i)]$$

Ergo, $FT_j(a, \frac{1}{n}\mathbb{I}) - FT_i(a, \frac{1}{n}\mathbb{I}) = T(a_j, a_i) - T(a_i, a_j) + \mathbb{E}[u(Y_i)] - \mathbb{E}[u(Y_j)] \leq 0$. Hence, due to step 1, $(a, \frac{1}{n}\mathbb{I})$ is not robust.

Part c:

Step 1: $\tau = k(\sum_{k=1}^N r_k)^{-1}$

By contradiction, assume a robust rule $(\tau\mathbb{I}, p)$ and $k < \tau \sum_{j=1}^N r_j$. Analogous to part (a), for every individual i , $FT_i(\tau\mathbb{I}, p) > 0$, and for the range $\tau = [0, 1]$, $FT_i(\tau\mathbb{I}, p)$ is concave and positive where only $FT_i(0\mathbb{I}, p) = 0$. Also, $-M_i(\tau\mathbb{I}, p)$ is increasing and convex, where it is only 0 when $\tau = 0$. Hence, from Lemma 2.3, $-\frac{FT_i(\tau\mathbb{I}, p)}{M_i(\tau\mathbb{I}, p)}$ is decreasing in τ . Thence, $(\tau\mathbb{I}, p)$ cannot be a robust rule because the rule $(\tau'\mathbb{I}, \frac{1}{n}\mathbb{I})$ where $\tau' = k(\sum_{k=1}^N r_k)^{-1}$ has a higher value of $-\frac{FT_i(\tau'\mathbb{I}, p)}{M_i(\tau'\mathbb{I}, p)}$.

Step 2:

If p is such that $p_i = 0$, $FT_i(\tau\mathbb{I}, p) < 0$. Therefore, p is not robust.

Step 3:

Let $(\tau\mathbb{I}, p)$ be a robust rule. Assume that $r_i > r_j$ and $p_i = p_j$. From the smoothing condition,

$$\begin{aligned} \left| \mathbb{E} \left[u \left(\Gamma(Y, \tau\mathbb{I}, p)_i \right) - u \left(\Gamma(Y, \tau\mathbb{I}, p)_j \right) \right] \right| &< \left| \mathbb{E} \left[u \left(\Gamma(Y, \mathbb{O}, p)_i \right) - u \left(\Gamma(Y, \mathbb{O}, p)_j \right) \right] \right| \\ &= \left| \mathbb{E} \left[u(Y_i) \right] - \mathbb{E} \left[u(Y_j) \right] \right| \end{aligned}$$

Given that Y_i first order stochastically dominates Y_j , then $\mathbb{E}[u(Y_i)] > \mathbb{E}[u(Y_j)]$. Therefore,

$$\begin{aligned} FT_i(\tau\mathbb{I}, p) - FT_j(\tau\mathbb{I}, p) &= \\ \mathbb{E} \left[u \left(\Gamma(Y, \tau\mathbb{I}, p)_i \right) - u \left(\Gamma(Y, \tau\mathbb{I}, p)_j \right) \right] - \mathbb{E} \left[u(Y_i) \right] + \mathbb{E} \left[u(Y_j) \right] &< 0 \end{aligned}$$

Thus,

$$\min_{k \in \{1, \dots, N\}} \left\{ -\frac{FT_k(\tau\mathbb{I}, p)}{M_k(\tau\mathbb{I}, p)} \right\} \neq -\frac{FT_j(\tau\mathbb{I}, p)}{M_j(\tau\mathbb{I}, p)}$$

To understand the previous result, consider that $FT_j(\tau\mathbb{I}, p) > FT_i(\tau\mathbb{I}, p)$ and $-M_i(\tau\mathbb{I}, p) = -M_j(\tau\mathbb{I}, p) > 0$. Moreover $p_j > 0$. Therefore, a marginal decrease in p_j that is equivalent to an increase in the share (p_i) of all the other individuals will increase their utility without decreasing the utility of j enough to make $FT_j(\tau\mathbb{I}, p)$ the minimum. Ergo, this change increases the $\min_{k \in \{1, \dots, N\}} \left\{ -\frac{FT_k(\tau\mathbb{I}, p)}{M_k(\tau\mathbb{I}, p)} \right\}$ and is still an admissible rule. Thus, the original rule was not robust.

Part d:

Assume the existence of individuals i, j such that $-\frac{FT_i(\tau\mathbb{I}, p)}{M_i(\tau\mathbb{I}, p)} < -\frac{FT_j(\tau\mathbb{I}, p)}{M_j(\tau\mathbb{I}, p)}$. From step 2 of part (c), $p_j > 0$. Hence, a marginal decrease in p_j compensated with an increase of p_i will increase the utility of all the others without decreasing the utility of j enough to make $FT_j(\tau\mathbb{I}, p)$ the minimum. Ergo, this change increases the minimum and is still an admissible rule. Thus, the original rule cannot be robust. ■

Propositions:

Proof of Proposition 2.1:

Part a:

For part (a) it is enough to provide an example of a situation where $(\mathbb{I}, \frac{1}{n}\mathbb{I})$ is not admissible.

Let $n = 2$, $u(x) = 1 - e^{-\rho x}$, $g(x) = \frac{1}{2}x$, and $f(x) = 1 - e^{-x}$, i.e. a community of two constant absolute risk-averse agents whose income follows an exponential distribution and the common good is a risk-sharing function. Then, for $i \in \{1, 2\}$

$$FT_i(\mathbb{I}, \frac{1}{n}\mathbb{I}) \propto \frac{1}{1 + r_i \rho} - \left(\frac{1}{1 + \frac{r_1}{2} \rho} \right) \left(\frac{1}{1 + \frac{r_2}{2} \rho} \right)$$

Thus, if $\frac{2(r_2 - r_1)}{r_1 r_2} > \rho$, then the second individual has a negative future trade-off. Ergo, the rule is not admissible.

Part b:

If $(\mathbb{I}, \frac{1}{n}\mathbb{I})$ is admissible, then $Ds(\mathbb{I}, \frac{1}{n}\mathbb{I}) \neq \emptyset$. Given that the set of rules $(a, \frac{1}{n}\mathbb{I})$ is defined in a compact $(a \in [0, 1])$ such that $\sum_{i=1}^N a_i r_i \geq k$, by Weierstrass Theorem, there exists $(a', \frac{1}{n}\mathbb{I})$ admissible such that $Ds(a, \frac{1}{n}\mathbb{I}) \subseteq Ds(a', \frac{1}{n}\mathbb{I})$ for all $a \in [0, 1]$.

Part c:

Let $\tau = 1$ and $p_i = \frac{r_i}{\sum_{j=1}^N r_j}$, and define $v(x) = u(r_i x)$ and $X_i = \frac{Y_i}{r_i}$. By construction $\forall i: X_i \sim iid f(x)$. Moreover,

$$\begin{aligned} \mathbb{E} \left[u \left(p_i n g \left(\sum_{j=1}^N Y_j \right) \right) \right] &\geq \mathbb{E} \left[u \left(p_i \sum_{j=1}^N Y_j \right) \right] = \mathbb{E} \left[v \left(\frac{\sum_{j=1}^N r_j X_j}{\sum_{j=1}^N r_j} \right) \right] \\ &> \mathbb{E} \left[\frac{\sum_{j=1}^N r_j v(X_j)}{\sum_{j=1}^N r_j} \right] = \mathbb{E} [v_i(X_i)] = \mathbb{E} [u(Y_i)] \end{aligned}$$

The first inequality comes from the definition of $g(x)$, and the second comes from Jensen's Inequality. The previous result implies that for every individual i , $FT_i(\mathbb{I}, p) > 0$. As there is an admissible rule, and given that all the functions are continuous and τ and p are defined in compacts ($\tau \in [k(\sum_{k=1}^N r_k)^{-1}, 1]$ and $p \in \Delta^n$), Weierstrass Theorem states the existence of $(\tau' \mathbb{I}, p')$ admissible such that $Ds(\tau \mathbb{I}, p) \subseteq Ds(\tau' \mathbb{I}, p')$ for all $\tau \in [k(\sum_{k=1}^N r_k)^{-1}, 1]$ and $p \in \Delta^n$. ■

Proof of Proposition 2.2:

CGA:

Let $(a, \frac{1}{n}\mathbb{I})$ be a robust agreement in a smoothing context.

Step 1: $\forall i \neq j$ either $-\frac{FT_i(a, \frac{1}{n}\mathbb{I})}{M_i(a, \frac{1}{n}\mathbb{I})} = -\frac{FT_j(a, \frac{1}{n}\mathbb{I})}{M_j(a, \frac{1}{n}\mathbb{I})}$ or $a_{k'} = 1$ for every $k' < \argmax_{k \in \{i, j\}} \left\{ -\frac{FT_k(a, \frac{1}{n}\mathbb{I})}{M_k(a, \frac{1}{n}\mathbb{I})} \right\}$.

Without loss of generality, if $-\frac{FT_i(a, \frac{1}{n}\mathbb{I})}{M_i(a, \frac{1}{n}\mathbb{I})} > -\frac{FT_j(a, \frac{1}{n}\mathbb{I})}{M_j(a, \frac{1}{n}\mathbb{I})}$ and $a_i \neq 1$, then a marginal decrease in a_i increases $\min_{k \in \{1, \dots, N\}} \left\{ -\frac{FT_k(a, \frac{1}{n}\mathbb{I})}{M_k(a, \frac{1}{n}\mathbb{I})} \right\}$. Thus, $(a, \frac{1}{n}\mathbb{I})$ was not robust.

If $a_i = 1$, then mimicking the proof of part (c) of Theorem 2.2 it follows that for all $j < i$, and unless $a_j = 1$, if $a_i \geq a_j$, then $(a, \frac{1}{n}\mathbb{I})$ is not robust.

Step 2: $FT_i(a, \frac{1}{n}\mathbb{I}) \leq FT_j(a, \frac{1}{n}\mathbb{I})$ for all $i > j$.

Form Theorem 2.2 part (a), if $a_i = a_j = 1$, then $FT_i(a, \frac{1}{n}\mathbb{I}) \leq FT_j(a, \frac{1}{n}\mathbb{I})$.

If $a_i < a_j$, then $-\frac{FT_i(a, \frac{1}{n}\mathbb{I})}{M_i(a, \frac{1}{n}\mathbb{I})} = -\frac{FT_j(a, \frac{1}{n}\mathbb{I})}{M_j(a, \frac{1}{n}\mathbb{I})}$. Hence, $M_j(a, \frac{1}{n}\mathbb{I}) \leq M_i(a, \frac{1}{n}\mathbb{I}) < 0$ implies that $FT_i(a, \frac{1}{n}\mathbb{I}) \leq FT_j(a, \frac{1}{n}\mathbb{I})$.

Step 3:

From the previous step, if $i > j$, $\frac{FT_i(a, \frac{1}{n}\mathbb{I})}{FT_j(a, \frac{1}{n}\mathbb{I})} \leq 1$; and due to the first order stochastic dominance of the income distributions, $\frac{\mathbb{E}[u(Y_i)]}{\mathbb{E}[u(Y_j)]} \geq 1$. Therefore,

$$\frac{FT_i(a, \frac{1}{n}\mathbb{I})}{FT_j(a, \frac{1}{n}\mathbb{I})} \leq \frac{FT_i(a, \frac{1}{n}\mathbb{I}) + \mathbb{E}[u(Y_i)]}{FT_j(a, \frac{1}{n}\mathbb{I}) + \mathbb{E}[u(Y_j)]} \leq \frac{\mathbb{E}[u(Y_i)]}{\mathbb{E}[u(Y_j)]}$$

where $FT_i(a, \frac{1}{n}\mathbb{I}) + \mathbb{E}[u(Y_i)]$ is the expected utility of individual i after the agreement. Thus, $(a, \frac{1}{n}\mathbb{I})$ promotes equality.

CPA:

Let $(\tau\mathbb{I}, p)$ be a robust agreement.

Step 1: $\forall i \neq j, -\frac{FT_i(\tau\mathbb{I}, p)}{M_i(\tau\mathbb{I}, p)} = -\frac{FT_j(\tau\mathbb{I}, p)}{M_j(\tau\mathbb{I}, p)}$.

Theorem 2.2, part (d).

Step 2: $FT_i(\tau\mathbb{I}, p) \leq FT_j(\tau\mathbb{I}, p)$ for every $i > j$.

Given that $-\frac{FT_i(\tau\mathbb{I}, p)}{M_i(\tau\mathbb{I}, p)} = -\frac{FT_j(\tau\mathbb{I}, p)}{M_j(\tau\mathbb{I}, p)}$, and $M_j(\tau\mathbb{I}, p) \leq M_i(\tau\mathbb{I}, p) < 0$, if $p_i \geq p_j$, then $FT_i(\tau\mathbb{I}, p) \leq FT_j(\tau\mathbb{I}, p)$.

Step 3:

From the previous step, $\frac{FT_i(\tau\mathbb{I}, p)}{FT_j(\tau\mathbb{I}, p)} \leq 1$. Also, from the first order stochastic dominance of the distributions of the income, $\frac{\mathbb{E}[u(Y_i)]}{\mathbb{E}[u(Y_j)]} \geq 1$. Therefore,

$$\frac{FT_i(\tau\mathbb{I}, p)}{FT_j(\tau\mathbb{I}, p)} \leq \frac{FT_i(\tau\mathbb{I}, p) + \mathbb{E}[u(Y_i)]}{FT_j(\tau\mathbb{I}, p) + \mathbb{E}[u(Y_j)]} \leq \frac{\mathbb{E}[u(Y_i)]}{\mathbb{E}[u(Y_j)]}$$

where $FT_i(\tau\mathbb{I}, p) + \mathbb{E}[u(Y_i)]$ is the expected utility of individual i after the agreement. Thus, $(\tau\mathbb{I}, p)$ promotes equality. ■

Proof of Proposition 2.3:

From Lemma 2.4 a Taylor expansion of order 1 in r , centered in $c\mathbb{I}$ follows the next expression:

$$\begin{aligned}
& \Delta F T_i \left(\frac{k}{cn} \mathbb{I}, \frac{1}{n} \mathbb{I} \right) \\
&= \frac{\frac{k}{cn} \sum_{j=1}^N \mathbb{E} \left[Y_i u' \left(\left(1 - \frac{k}{cn} \right) Y_i + g \left(\frac{k}{cn} \sum_{l=1}^N Y_l \right) \right) g' \left(\frac{k}{cn} \sum_{l=1}^N Y_l \right) \right] \Delta r_j}{c} \\
&+ \frac{\mathbb{E} \left[\left(1 - \frac{k}{cn} \right) Y_i u' \left(\left(1 - \frac{k}{cn} \right) Y_i + g \left(\frac{k}{cn} \sum_{l=1}^N Y_l \right) \right) g' \left(\frac{k}{cn} \sum_{l=1}^N Y_l \right) \right] \Delta r_i}{c} - \frac{\mathbb{E} \left[Y_i u' (Y_i) \right] \Delta r_i}{c} \\
&= \frac{\frac{k}{cn} \mathbb{E} \left[Y_i u' \left(\left(1 - \frac{k}{cn} \right) Y_i + g \left(\frac{k}{cn} \sum_{l=1}^N Y_l \right) \right) g' \left(\frac{k}{cn} \sum_{l=1}^N Y_l \right) \right] \sum_{j=1}^N \Delta r_j}{c} \\
&+ \frac{\mathbb{E} \left[\left(1 - \frac{k}{cn} \right) Y_i u' \left(\left(1 - \frac{k}{cn} \right) Y_i + g \left(\frac{k}{cn} \sum_{l=1}^N Y_l \right) \right) g' \left(\frac{k}{cn} \sum_{l=1}^N Y_l \right) \right] \Delta r_i}{c} - \frac{\mathbb{E} \left[Y_i u' (Y_i) \right] \Delta r_i}{c} \\
&= \frac{\mathbb{E} \left[Y_i \left(\left(1 - \frac{k}{cn} \right) u' \left(\left(1 - \frac{k}{cn} \right) Y_i + g \left(\frac{k}{cn} \sum_{l=1}^N Y_l \right) \right) g' \left(\frac{k}{cn} \sum_{l=1}^N Y_l \right) - u' (Y_i) \right) \right] \Delta r_i}{c}
\end{aligned}$$

To simplify notation, during the proof,

$$\mu = \frac{\mathbb{E} \left[Y_i \left(\left(1 - \frac{k}{cn} \right) u' \left(\left(1 - \frac{k}{cn} \right) Y_i + g \left(\frac{k}{cn} \sum_{l=1}^N Y_l \right) \right) g' \left(\frac{k}{cn} \sum_{l=1}^N Y_l \right) - u' (Y_i) \right) \right] \Delta r_i}{c}$$

The previous derivations will be use in the following proofs:

Part a:

Given that in both cases infinitesimal deviations are considered, the relevant functions are approximated by a first order Taylor expansion, centred on $(\frac{k}{cn} \mathbb{I}, \frac{1}{n} \mathbb{I})$ (which is the robust rule of an homogeneous population). For heterogeneous populations, robust agreements are denoted in a way such that for every individual i , $a_i = \frac{k}{cn} + \Delta a_i$ and $p_i = \frac{1}{n} + \Delta p_i$ (τ is not relevant as the shock is a mean-spread). Therefore, a Taylor expansion of order 1 on r, a , and p can be done. Using Lemma 2.4 on the expansion, the following equations are obtained:

CGA:⁸

$$\begin{aligned} \Delta\left(-\frac{FT_i(a, \frac{1}{n}\mathbb{I})}{M_i(a, \frac{1}{n}\mathbb{I})}\right) &= -\frac{\mu}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})}\Delta r_i \\ &- \sum_{j=1}^N \left(\frac{\mathbb{E}\left[Y_j u' \left(\left(1 - \frac{k}{cn}\right)Y_i + g\left(\frac{k}{cn}\sum_{k=1}^N Y_k\right)\right) g'\left(\frac{k}{cn}\sum_{k=1}^N Y_k\right)\right]}{M_j(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})} \Delta a_j \right) \\ &+ \left(\frac{\mathbb{E}\left[Y_j u' \left(\left(1 - \frac{k}{cn}\right)Y_i + g\left(\frac{k}{cn}\sum_{k=1}^N Y_k\right)\right) g'\left(\frac{k}{cn}\sum_{k=1}^N Y_k\right)\right]}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})} + \frac{FT_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})M'_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})^2} \right) \Delta a_i \end{aligned}$$

Because the Taylor expansion is centred in an homogeneous case, all the values that are not denoted with Δ are independent of the individual. Thus,

$$\begin{aligned} \Delta\left(-\frac{FT_i(a, \frac{1}{n}\mathbb{I})}{M_i(a, \frac{1}{n}\mathbb{I})}\right) &= -\frac{\mu}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})}\Delta r_i \\ &- \left(\frac{\mathbb{E}\left[Y_j u' \left(\left(1 - \frac{k}{cn}\right)Y_i + g\left(\frac{k}{cn}\sum_{k=1}^N Y_k\right)\right) g'\left(\frac{k}{cn}\sum_{k=1}^N Y_k\right)\right]}{M_j(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})} \sum_{j=1}^N \Delta a_j \right) \\ &+ \left(\frac{\mathbb{E}\left[Y_j u' \left(\left(1 - \frac{k}{cn}\right)Y_i + g\left(\frac{k}{cn}\sum_{k=1}^N Y_k\right)\right) g'\left(\frac{k}{cn}\sum_{k=1}^N Y_k\right)\right]}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})} + \frac{FT_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})M'_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})^2} \right) \Delta a_i \end{aligned}$$

From Theorem 2.2, for all i, j ,

$$\Delta\left(-\frac{FT_i(a, \frac{1}{n}\mathbb{I})}{M_i(a, \frac{1}{n}\mathbb{I})}\right) = \Delta\left(-\frac{FT_j(a, \frac{1}{n}\mathbb{I})}{M_j(a, \frac{1}{n}\mathbb{I})}\right)$$

Hence,

$$\begin{aligned} \frac{\mu}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})}(\Delta r_i - \Delta r_j) &= \left(\frac{\mathbb{E}\left[Y_j u' \left(\left(1 - \frac{k}{cn}\right)Y_i + g\left(\frac{k}{cn}\sum_{k=1}^N Y_k\right)\right) g'\left(\frac{k}{cn}\sum_{k=1}^N Y_k\right)\right]}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})} \right. \\ &\quad \left. + \frac{FT_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})M'_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})^2} \right) (\Delta a_i - \Delta a_j) \end{aligned}$$

⁸During this proof $M'(\alpha, \beta)$ is a short notation for $\frac{dM}{d\alpha}$ or $\frac{dM}{d\beta}$. The right option is clear by context, and whereas this is an abuse of notation it helps to reduce the size of the equations.

On the other hand,

$$k = \sum_{i=1}^N a_i r_i = \sum_{i=1}^N \left(\frac{k}{cn} + \Delta a_i \right) (c + \Delta r_i) \Rightarrow \sum_{i=1}^N \Delta a_i (c + \Delta r_i) = 0$$

Thus, multiplying the previous result by $(c - \Delta r_j)$ and adding it over all j :

$$\sum_{j=1}^N \left(\frac{\mu}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})} (\Delta r_i - \Delta r_j) (c - \Delta r_j) \right) = \sum_{j=1}^N \left(\frac{\mu}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})} (c \Delta r_i - \Delta r_j^2) \right)$$

and

$$\begin{aligned} & \sum_{j=1}^N \left(\left(\frac{\mathbb{E} \left[Y_j u' \left(\left(1 - \frac{k}{cn} \right) Y_i + g \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right) g' \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right]}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})} \right. \right. \\ & \quad \left. \left. + \frac{FT_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I}) M_i'(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})^2} \right) (\Delta a_i - \Delta a_j) (c - \Delta r_j) \right) \\ &= \left(\frac{\mathbb{E} \left[Y_j u' \left(\left(1 - \frac{k}{cn} \right) Y_i + g \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right) g' \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right]}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})} \right. \\ & \quad \left. + \frac{FT_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I}) M_i'(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})^2} \right) nc \Delta a_i \end{aligned}$$

Therefore,

$$\begin{aligned} \Delta a_i &= \frac{\frac{\mu \Delta r_i}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})} - \sum_{j=1}^N \left(\frac{\mu}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})} \left(\frac{\Delta r_j^2}{nc} \right) \right)}{\left(\frac{\mathbb{E} \left[Y_j u' \left(\left(1 - \frac{k}{cn} \right) Y_i + g \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right) g' \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right]}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})} + \frac{FT_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I}) M_i'(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})^2} \right)} \\ &= \frac{\mu \varsigma \Delta r_i}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})} - \sum_{j=1}^N \left(\frac{\mu \varsigma}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})} \left(\frac{\Delta r_j^2}{nc} \right) \right) \end{aligned}$$

where

$$\begin{aligned} \varsigma^{-1} &= \\ & \left(\frac{\mathbb{E} \left[Y_j u' \left(\left(1 - \frac{k}{cn} \right) Y_i + g \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right) g' \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right]}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})} + \frac{FT_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I}) M_i'(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})^2} \right) < 0 \end{aligned}$$

and

$$\sum_{j=1}^N \Delta a_j = - \sum_{j=1}^N \left(\frac{\mu \varsigma}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})} \left(\frac{\Delta r_j^2}{c} \right) \right)$$

Theorem 2.2 part (a) implies that $\frac{\mu \varsigma}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})} < 0$.

By replacing all these results in $\Delta \left(-\frac{FT_i(a, \frac{1}{n}\mathbb{I})}{M_i(a, \frac{1}{n}\mathbb{I})} \right)$:

$$\begin{aligned} & \Delta \left(-\frac{FT_i(a, \frac{1}{n}\mathbb{I})}{M_i(a, \frac{1}{n}\mathbb{I})} \right) \\ &= -\frac{\mu}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})} \left(\frac{\sum_{j=1}^N \Delta r_j^2}{nc} \right) \left(\frac{\varsigma}{(M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I}))^2} \right) \left(FT_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I}) M'_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I}) \right. \\ & \quad \left. + M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I}) \mathbb{E} \left[Y_i u' \left(\left(1 - \frac{k}{cn} \right) Y_i + g \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right) \left(1 - g' \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right) \right] \right) \leq 0 \end{aligned}$$

if and only if

$$\begin{aligned} & FT_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I}) M'_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I}) \\ & + M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I}) \mathbb{E} \left[Y_i u' \left(\left(1 - \frac{k}{cn} \right) Y_i + g \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right) \left(1 - g' \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right) \right] \leq 0 \end{aligned}$$

CPA:

$$\begin{aligned} & \Delta \left(-\frac{FT_i(\frac{k}{cn}\mathbb{I}, p)}{M_i(\frac{k}{cn}\mathbb{I}, p)} \right) = -\frac{\mu}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})} \Delta r_i \\ & - \left(\frac{\mathbb{E} \left[Y_j u' \left(\left(1 - \frac{k}{cn} \right) Y_i + g \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right) g \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right]}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})} + \frac{FT_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I}) M'_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})^2} \right) \Delta p_i \end{aligned}$$

From Proposition 2.1 $\forall i, j : \Delta \left(-\frac{FT_i(a, \frac{1}{n}\mathbb{I})}{M_i(a, \frac{1}{n}\mathbb{I})} \right) = \Delta \left(-\frac{FT_j(a, \frac{1}{n}\mathbb{I})}{M_j(a, \frac{1}{n}\mathbb{I})} \right)$ and $\sum_{i=1}^N \Delta p = 0$. Then,

$$\begin{aligned}
& \frac{-\mu}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})}(\Delta r_i - \Delta r_j) \\
&= \left(\frac{\mathbb{E}\left[Y_j u' \left(\left(1 - \frac{k}{cn}\right)Y_i + g\left(\frac{k}{cn}\sum_{k=1}^N Y_k\right)\right) g\left(\frac{k}{cn}\sum_{k=1}^N Y_k\right)\right]}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})} \right) (\Delta p_i - \Delta p_j) \\
&+ \left(\frac{FT_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})M'_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})^2} \right) (\Delta p_i - \Delta p_j)
\end{aligned}$$

And, adding over all j ,

$$\begin{aligned}
& \frac{-\mu}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})}\Delta r_i \\
&= \left(\frac{\mathbb{E}\left[Y_j u' \left(\left(1 - \frac{k}{cn}\right)Y_i + g\left(\frac{k}{cn}\sum_{k=1}^N Y_k\right)\right) g\left(\frac{k}{cn}\sum_{k=1}^N Y_k\right)\right]}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})} + \frac{FT_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})M'_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})^2} \right) \Delta p_i
\end{aligned}$$

Thus, $\Delta\left(-\frac{FT_i(\frac{k}{cn}\mathbb{I}, p)}{M_i(\frac{k}{cn}\mathbb{I}, p)}\right) = 0$. When the two cases are compared:

$$\begin{aligned}
& FT_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})M'_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I}) \\
&+ M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})\mathbb{E}\left[Y_i u' \left(\left(1 - \frac{k}{cn}\right)Y_i + g\left(\frac{k}{cn}\sum_{k=1}^N Y_k\right)\right) \left(1 - g'\left(\frac{k}{cn}\sum_{k=1}^N Y_k\right)\right)\right] < 0
\end{aligned}$$

if and only if

$$\min_{i \in \{1, \dots, N\}} \left\{ -\frac{FT_i(\frac{k}{cn}\mathbb{I}, p)}{M_i(\frac{k}{cn}\mathbb{I}, p)} \right\} - \min_{i \in \{1, \dots, N\}} \left\{ -\frac{FT_i(a, \frac{1}{n}\mathbb{I})}{M_i(a, \frac{1}{n}\mathbb{I})} \right\} < 0$$

Part b:

Using the same notation as in the previous case,

CGA:

$$\begin{aligned}
& \sum_{l=1}^N \Delta \left(FT_l(a, \frac{1}{n}\mathbb{I}) + \mathbb{E} \left[u(Y_l) \right] \right) \\
&= n \mathbb{E} \left[Y_i u' \left(\left(1 - \frac{k}{cn} \right) Y_i + g \left(\frac{k}{cn} \sum_{k=1}^N Y_l \right) \right) g' \left(\frac{k}{cn} \sum_{k=1}^N Y_l \right) \right] \sum_{j=1}^N \Delta a_j \\
&- \mathbb{E} \left[u' \left(\left(1 - \frac{k}{cn} \right) Y_i + g \left(\frac{k}{cn} \sum_{k=1}^N Y_l \right) \right) \right] \sum_{j=1}^N \Delta a_j \\
&= \mathbb{E} \left[u' \left(\left(1 - \frac{k}{cn} \right) Y_i + g \left(\frac{k}{cn} \sum_{k=1}^N Y_l \right) \right) \left(n g' \left(\frac{k}{cn} \sum_{k=1}^N Y_l \right) - 1 \right) \right] \sum_{j=1}^N \Delta a_j \geq 0
\end{aligned}$$

CPA:

$$\begin{aligned}
& \sum_{l=1}^N \Delta \left(FT_l \left(\frac{k}{cn} \mathbb{I}, p \right) + \mathbb{E} \left[u(Y_l) \right] \right) \\
&= \left(\frac{\mathbb{E} \left[Y_i u' \left(\left(1 - \frac{k}{cn} \right) Y_i + g \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right) g' \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right]}{\mathbb{E}[Y_i]} \right) \sum_{k=1}^N \Delta p_k = 0
\end{aligned}$$

Part c:

Using the same notation as in the previous case,

CGA:

$$\begin{aligned}
& \Delta FT_i(a, \frac{1}{n}\mathbb{I}) \\
&= \mu \Delta r_i + \mathbb{E} \left[Y_j u' \left(\left(1 - \frac{k}{cn} \right) Y_i + g \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right) g' \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right] \sum_{j=1}^N \Delta a_j \\
&- \mathbb{E} \left[Y_j u' \left(\left(1 - \frac{k}{cn} \right) Y_i + g \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right) g' \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right] \Delta a_i
\end{aligned}$$

Then, replacing the previous results

$$\begin{aligned} \Delta FT_i(a, \frac{1}{n}\mathbb{I}) &= \mu \left(1 - \left(\frac{\mathbb{E} \left[Y_i u' \left(\left(1 - \frac{k}{cn} \right) Y_i + g \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right) \right]}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})} \right) \varsigma \right) \Delta r_i - \\ &\mathbb{E} \left[Y_i u' \left(\left(1 - \frac{k}{cn} \right) Y_i + g \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right) \left(1 - ng' \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right) \right] \varsigma \sum_{j=1}^N \left(\frac{\mu}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})} \left(\frac{\Delta r_j^2}{nc} \right) \right) \end{aligned}$$

From the previous equation, notice that $\left(1 - \left(\frac{\mathbb{E} \left[Y_i u' \left(\left(1 - \frac{k}{cn} \right) Y_i + g \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right) \right]}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})} \right) \varsigma \right) > 0$.

Thus $\Delta FT_i(a, \frac{1}{n}\mathbb{I})$ decreases when Δr_i increases.

Also, $\mathbb{E} \left[Y_i u' \left(\left(1 - \frac{k}{cn} \right) Y_i + g \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right) \left(1 - ng' \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right) \right] \varsigma \leq 0$. Then $\Delta r_i \geq 0$ implies that $\Delta FT_i(a, \frac{1}{n}\mathbb{I}) \leq 0$.

CPA:

$$\begin{aligned} \Delta FT_i(\frac{k}{cn}\mathbb{I}, p) &= \mu \Delta r_i - \mathbb{E} \left[Y_j u' \left(\left(1 - \frac{k}{cn} \right) Y_i + g \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right) g \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right] \Delta p_i \\ &= \mu \left(1 - \left(\frac{\mathbb{E} \left[Y_j u' \left(\left(1 - \frac{k}{cn} \right) Y_i + g \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right) g \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right]}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})} \right) \varsigma \right) \Delta r_i \end{aligned}$$

Finally, $\left(1 - \left(\frac{\mathbb{E} \left[Y_j u' \left(\left(1 - \frac{k}{cn} \right) Y_i + g \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right) g \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right]}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})} \right) \varsigma \right) > 0$. Thus, $\Delta FT_i(\frac{k}{cn}\mathbb{I}, p)$

decreases when Δr_i increases, and $\Delta FT_i(\frac{k}{cn}\mathbb{I}, p)$ has the opposite sign of Δr_i .

Part d:

Using the same notation as in the previous case,

CGA:

$$\begin{aligned} & \Delta\left(FT_i(a, \frac{1}{n}\mathbb{I}) + \mathbb{E}[u(Y_i)]\right) \\ &= \left(\mu\left(1 - \left(\frac{\mathbb{E}\left[Y_i u'\left((1 - \frac{k}{cn})Y_i + g(\frac{k}{cn}\sum_{k=1}^N Y_k)\right)\right]}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})}\right)\varsigma\right) + \frac{\mathbb{E}[Y_i u'(Y_i)]}{c}\right)\Delta r_i \\ & - \mathbb{E}\left[Y_i u'\left((1 - \frac{k}{cn})Y_i + g(\frac{k}{cn}\sum_{k=1}^N Y_k)\right)(1 - ng'(\frac{k}{cn}\sum_{k=1}^N Y_k))\right]\varsigma \sum_{j=1}^N \left(\frac{\mu}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})}\left(\frac{\Delta r_j^2}{nc}\right)\right) \end{aligned}$$

Notice that $\mu + \frac{\mathbb{E}[Y_i u'(Y_i)]}{c} \geq 0$. Hence,

$$\mu\left(1 - \left(\frac{\mathbb{E}\left[Y_i u'\left((1 - \frac{k}{cn})Y_i + g(\frac{k}{cn}\sum_{k=1}^N Y_k)\right)\right]}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})}\right)\varsigma\right) + \frac{\mathbb{E}[Y_i u'(Y_i)]}{c} \geq 0$$

Thus $\Delta\left(FT_i(a, \frac{1}{n}\mathbb{I}) + \mathbb{E}[u(Y_i)]\right)$ increases when Δr_i increases. Moreover, using the same logic as in part (c), if $\Delta r_i < 0$, then $\Delta\left(FT_i(a, \frac{1}{n}\mathbb{I}) + \mathbb{E}[u(Y_i)]\right) < 0$.

CPA:

$$\begin{aligned} & \Delta\left(FT_i(\frac{k}{cn}\mathbb{I}, p) + \mathbb{E}[u(Y_i)]\right) \\ &= \left(\mu\left(1 - \left(\frac{\mathbb{E}\left[Y_j u'\left((1 - \frac{k}{cn})Y_i + g(\frac{k}{cn}\sum_{k=1}^N Y_k)\right)g(\frac{k}{cn}\sum_{k=1}^N Y_k)\right]}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})}\right)\varsigma\right) + \frac{\mathbb{E}[Y_i u'(Y_i)]}{c}\right)\Delta r_i \end{aligned}$$

As in the previous case,

$$\mu\left(1 - \left(\frac{\mathbb{E}\left[Y_j u'\left((1 - \frac{k}{cn})Y_i + g(\frac{k}{cn}\sum_{k=1}^N Y_k)\right)g(\frac{k}{cn}\sum_{k=1}^N Y_k)\right]}{M_i(\frac{k}{cn}\mathbb{I}, \frac{1}{n}\mathbb{I})}\right)\varsigma\right) + \frac{\mathbb{E}[Y_i u'(Y_i)]}{c} > 0$$

Thus, $\left|\Delta\left(FT_i(\frac{k}{cn}\mathbb{I}, p) + \mathbb{E}[u(Y_i)]\right)\right|$ increases when $|\Delta r_i|$ increases. Furthermore,

$$\text{sign}\left(\Delta\left(FT_i(\frac{k}{cn}\mathbb{I}, p) + \mathbb{E}[u(Y_i)]\right)\right) = \text{sign}(\Delta r_i)$$

Part e:

In the case of risk-sharing agreements,

$$\mathbb{E} \left[Y_i u' \left(\left(1 - \frac{k}{cn} \right) Y_i + g \left(\frac{k}{cn} \sum_{k=1}^N Y_k \right) \right) \left(1 - n g' \left(\sum_{k=1}^N Y_k \right) \right) \right] = 0$$

Therefore, the equations in part (c), and (d) conclude that both models have the same behaviour.

■

Chapter 3

On the Influence of Economic Journals

3.1 Introduction

Many economists studying market efficiency have explored whether the organisation of economic research itself is efficient and, if not, how it can be improved. One aspect of the organisation of economic research that has attracted a lot of attention recently is the importance given to publications in top journals, in particular to the so-called "Top 5": *American Economic Review*, *Econometrica*, *Journal of Political Economy*, *Review of Economic Studies*, and *Quarterly Journal of Economics*. At many departments, publishing in Top 5 journals is an important factor, if not a requirement, for tenure or promotion. This issue has raised concerns that the economic profession focuses too much on publishing in these journals. Furthermore, authorships and editorships of the top journals have been dominated by U.S. elite schools, which raises concerns about market power and its implications on the efficiency of research production (Fourcade et al., 2015).¹

The reigning discussion motivates obvious questions, such as: How important are these top journals? Do the Top 5 journals indeed stand out from other top journals? And, has it always been like this, or have the "Top 5" emerged? Even more importantly, is it possible to say something about the causes of the excessive influence of the Top 5 journals?

Regarding the first two questions, one way to analyse them is through citations. While having its measurement drawbacks, citations proxy for influence and are widely measurable and reported. The first step of this paper is then to perform a comprehensive citation analysis,

¹For more insights about these issues, the 2017 AEA Annual Meeting had a panel discussion on this topic: "Publishing and Promotion in Economics: The Curse of the Top Five", which can be viewed at <https://www.aeaweb.org/conference/webcasts/2017>.

considering trends in journal influence of the Top 5 journals in comparison to other important economic journals.

This analysis leads to the following results. First, the Top 5 journals outpaced the other journals in the 1980s and 1990s in terms of journal influence. To illustrate, comparing the Top 5 journals to a set of Tier 2 journals, namely *Economic Journal*, *International Economic Review*, and *Review of Economics and Statistics*, articles in the Top 5 journals received on average 68% more citations than Tier 2 journals in 1980. This influence gap widened, such that in 1999 Top 5 articles received on average 350% more citations than Tier 2 journals. Trends for other top field journals are similar.

However, the second finding is that this widening of the influence gap stopped in 2000, and, since then, it has partially reversed for some journals. For example, in 2017, the average Top 5 article received 100% more citations than an article in a Tier 2 journal, reversing the influence gap to the level of 1983.

Taking a closer look at the rebound after 2000, this paper obtains a surprising third finding: when individual top field journals are considered, the influence measures of these journals have converged. For example, in 2000, the *Journal of Econometrics* received almost three times as many citations as the *Journal of Development Economics*. In contrast, in 2017, the *Journal of Development Economics* has surpassed the *Journal of Econometrics* in the citation rankings, and the *Journal of Development Economics* now receives around 35% more citations than the *Journal of Econometrics*.

The second and third findings are surprising and raise questions about the perceived stability of the outperformance of the Top 5 journals. Moreover, they foment the question of what caused these reversals and convergence in trends.

The second part of this paper considers a theoretical model of a two-sided journal market. In this model, journals act as platforms connecting authors that want to publish their ideas and readers that want to read good ideas. Journals act as screening devices on the quality of an author's ideas, such that authors not only care about publishing in a journal with a large readership but also in a journal that publishes high-quality ideas. Similarly, readers value journals that publish many high-quality ideas, such that journals indeed prefer to publish as many good ideas as possible. However, each journal faces a space constraint, such that it decides to set a quality threshold in order to constrain the number of submissions within its capacity.

This model considers competition between one general interest journal, attracting readers from many fields, and many field journals that specialise in a single field. Upon analysing the model, the first observation is that there are multiple equilibria that differ in the hierarchy of

the journals. In some fields, the general interest journal is the top journal and the field journal is the second, whereas in other fields, the field journal is the top journal and the general interest journal is the second. The existence of multiple equilibria can be understood as a bandwagon effect with respect to quality, that is if the best researchers from a particular field publish in journal *A* then it becomes more attractive to publish in journal *A*, such that journal *A* is able to select to publish the best ideas from the field.

Second, the model shows that the ambiguity in equilibria disappears when the number of authors in the fields and the size of the readership of the general interest journal are large enough. In that case, the benefits of publishing in the general interest journal are so large that the general interest journal is the top journal in all fields.

Third, considering comparative statics, the influence of a journal decreases when the journal is able to publish more articles, and the influence of the field journals converge with an expansion of those fields in which the general interest journal has the highest influence.

This third group of findings suggests potential mechanisms for the observed trends in the citation patterns. Therefore, the final part of the paper returns to the citation data to see if these mechanisms are empirically plausible. Here, the paper indeed finds a negative relation between journal space and journal influence. Moreover, Top 5 journals reduced the number of articles in their journals in the 80s and 90s, whereas the other top journals increased the number of articles in their journals in the same period. This fact mimics the trends observed in the impact factors, which makes changes in journal capacity a plausible explanation to the average changes in journal influence. Similarly, fields that used to have a relatively low impact factor, such as development economics and urban economics, have expanded significantly in recent years, which could explain the convergence of impact factors of *all* fields in the data.

This paper makes the following contributions. It is the first to make a comprehensive comparison between the Top 5 journals and other top journals. This covers and extends previous research, and allows it to obtain new findings regarding journal influence since 2000, and the role of journal space and field sizes in explaining these trends. Concerning previous research, Ellison (2002b) already noted that in the 1980s and 1990s second tier generalised interest journals and top field journals had lost influence relative to Top 5 journals in terms of the impact factor, which estimates the number of citations that the average journal article receives. Card and DellaVigna (2013) performed an extensive citation analysis of the Top 5 journals from the 1970s, and noted an increased number of submissions and lower acceptance rates in the Top 5 journals. However, they did not compare the Top 5 journals to other journals.

Second, this paper makes a contribution to the theoretical literature on the market of academic journals as two-sided platforms. In standard platform competition models (Armstrong, 2006; Rochet and Tirole, 2003), platforms compete for quantity, that is, number of sellers or buyers. For the model presented in Section 3.3, journals also have a screening role, such that they also compete for the quality of submissions. McCabe and Snyder (2005) and Jeon and Rochet (2010) modelled the quality aspect of journals. They focused on open access pricing in a model with a single monopolist journal. While the model presented in this paper does not consider pricing, it does take into account competition between *multiple* journals with respect to the quality threshold strategy. Thus, it is possible to obtain results regarding the interaction of journals with respect to their influence. McCabe and Snyder (2016) offered a model with multiple academic journals that compete for quality and quantity. In their model, journals compete in prices but do not decide on a quality threshold. In the model presented here, competition is exactly in terms of the quality threshold, which explains why journals interact with respect to journal influence, i.e. this approach links the model with the main variable of interest of the empirical analysis.

The paper proceeds as follows. Section 3.2 presents the empirical trends on journal influence since 1980. Section 3.3 constructs the theoretical model and derives the main theoretical results. Section 3.4 links the model back to the empirics, giving a potential explanation for the trends observed in Section 3.2. Section 3.5 concludes.

3.2 Trends in Journal Influence

This section presents the main empirical trends of the impact factors of economic journals. First, it describes the datasets studied, and then it displays the trends associated with the influence gap between Top 5 and other top economic journals.

3.2.1 Data Sources

The key objective of this paper is to investigate the evolution of the influence top economic journals have had in academia. In this case, influence is studied via citation-based measurements.²

²Traditionally, influence has been considered to be a function of the level of citations that a paper, author, or journal receives. Underlying this perspective is the idea that citations reflect the way in which researchers validate the quality of ideas that an article contains (Posner, 2000). However, with the emergence of new social networks where authors divulge research, several authors such as Konkiel (2016) proposed new approaches to measuring the influence of a paper including papers' outreach in social and specialised media. While these new

The citation data used for the following analysis comes from Web of Science. Web of Science is an information system containing more than 20,000 journals, books, and conference proceedings that included over 80 million records of the most relevant journals (Clarivate Analytics, 2018). Moreover, out of the 45 ranking papers reviewed and presented by Bornmann et al. (2017), 26 studies used Web of Science as the primary source of information. Hence, economists widely regard it as a reliable source of information for this type of studies.

In this paper, top journals are classified into general interest and Top Field (top field-specific) journals. Within the general interest journals, there are Top 5 and Tier 2 (second-tier general interest journals). Except for Top 5, there are no well-defined criteria separating the other groups. For comparability purposes, the journals in Tier 2 and Top Field were chosen based on Ellison (2002b). This categorisation gives a set of 16 top journals: five Top 5, three Tier 2, and eight Top Field journals. Table 3.2.1 presents the journals in each group.³

Table 3.2.1 Top journal groups

Category	Journals
Top 5	American Economic Review, Econometrica, Journal of Political Economics, Quarterly Journal of Economics, and Review of Economic Studies.
Tier 2	Economic Journal, Review of Economics and Statistics, International Economic Review.
Top Field	Journal of International Economics, Journal of Econometrics, Journal of Economic Theory, Journal of Urban Economy, Journal of Monetary Economy ⁴ , Rand Journal of Economics ⁵ , Journal of Public Economics, and Journal of Development Economics.

Given the list of top journals, and based on the data available on Web of Science, two datasets on citations were created. The first dataset, hereafter known as Top-Journals dataset, considers all the articles and proceedings papers published between 1970 and 2017 in each of

approaches mark a step forward towards understanding outreach in a world dominated by the internet, such measurements are new and are thus not available for this paper's relevant period of study.

³In recent years, new influential journals such as the American Economic Journal and the Journal of the European Economic Association have emerged. Nevertheless, they have not been operating long enough and therefore it is not possible to measure accurately the influence they have had on economic research.

⁴Also covers its predecessor Carnegie-Rochester Conference Series on Public Policy.

⁵Also covers its predecessor Bell Journal of Economics.

the aforementioned 16 journals. Top-Journals dataset includes all citations from the universe of Web of Science that each item received every year following its publication date.

The Top-journals dataset does not contain information on individual citations, however. Therefore, a second dataset was created, hereafter known as the 100-Journals dataset. This dataset includes information on individual citations by building on a universe of 100 journals (including the 16 previously mentioned). While it is not the complete universe of economic-relevant journals, the sample is large enough to cover the journals that account for the most significant share of citations that top journals receive. Moreover, there is no official list that states what are the relevant economic journals. As such, this paper selects the journals from the "Simple Rank" list of "All Years" published by IDEAS/RePEc.⁶ For each of the journals in this list, the 100-Journals dataset includes, per article and proceedings paper, how many references it has, and what articles they are mentioning. The data was retrieved from <https://ideas.repec.org/top/> in May 2018, and the complete list is in Appendix 3.A. In contrast to the Top-journals dataset, the 100-Journals dataset tracks which journals are citing each of the top journals and in what year they are citing them.

3.2.2 Journal Influence

The impact factor, first described by Garfield (1955, 1972), has been the primary indicator of a journal's influence. This measurement has been used to rank journals by several authors in economics, such as Liebowitz and Palmer (1984), Laband and Piette (1994), Kalaitzidakis et al. (2003), and Engemann and Wall (2009). Analytically, the p -years impact factor ($IF_{i,t}^p$) of journal i in year t is calculated as

$$IF_{i,t}^p = \frac{\sum_{s=t-p}^{t-1} c_{i,s,t}}{\sum_{s=t-p}^{t-1} n_{i,s}}$$

where $c_{i,s,t}$ is the number of citations that the articles of journal i published in year s received from articles published in any journal in year t ; and $n_{i,s}$ is the number of articles that journal i published in year s . This measure derives its popularity from its natural interpretation and wide availability. Intuitively, a journal that gets more citations per article is more influential in the field than a journal that receives fewer citations. Moreover, since the late 1970s, Social

⁶IDEAS is a web portal run by the Research Division of the Federal Reserve Bank of St. Louis and uses the RePEc (Research Papers in Economics) database to rank economic journals, among other objectives. Dedicated exclusively to economic research, IDEAS currently has over 2,500,000 items of research, and it is therefore considered a focal point for many economists. Due to its broad coverage and relevance in economics, IDEAS was considered the primary reference to identify the universe of journals.

Science Citation Index (SSCI) has provided a large sample of citation data that facilitates its calculation (Liebowitz and Palmer, 1984).

Following Ellison (2002b), this paper calculates the p -year impact factor ratio ($IFR_{i,t}^p$) of journal i in year t as:

$$IFR_{i,t}^p = \frac{IF_{i,t}^p}{\sum_{j \in Top5} \frac{IF_{j,t}^p}{5}}$$

This measure is used to track the evolution of the influence gap between a particular journal and Top 5 across the period of study. While many papers and bibliometric sources calculate the 2-year impact factor, Engemann and Wall (2009) claimed that two years is considered a very short period to measure influence. For this reason, as well as for comparability with Ellison (2002b), this article studies influence based on the 10-year impact factor ratio of a journal.⁷

Figure 3.2.1 uses the Top-Journals dataset to present the evolution of the average impact factor ratio of Tier 2 and Top Field journals relative to Top 5. This figure reveals two trends. First, it displays a dramatic decline in the impact factor ratio of Tier 2 and Top Field journals in the period from 1980 until 1999 relative to the impact factor of the Top 5 journals. In particular, the impact factor ratio of Tier 2 to Top 5 dropped from 0.59 to 0.22, while that of Top Field to Top 5 declined from 0.81 to 0.34. To illustrate these numbers, while the average article in a Top 5 journal received around 68% more citations than the average Tier 2 article in 1980, Top 5 articles received on average about 350% more citation than Tier 2 articles in 1999! Similarly, articles in Top 5 journals received around 23% more citations than those in Top Field journals in 1980, however, on average they received about 200% more citations than Top Field articles in 1999. These numbers show that Top 5 journals simply started to jump out in the 1980s and 1990s, explaining why the label "Top 5" appeared in around the year 2000.

However, Figure 3.2.1 depicts a reversal of this trend after 2000. By 2017, the impact factor ratio of Tier 2 was 0.49, which means that Top 5 articles received over a 100% more citations than Tier 2 articles, almost as high as in the mid-1980s. Top Field journals also closed the influence gap although their impact factor ratio did not improve as much as Tier

⁷Ellison (2002b) used a slight variation of impact factor in his calculations. His version is defined as $IF_{i,t}^p = \frac{\sum_{s=t-p-1}^t c_{i,s,t}}{n_{i,p}^*}$. There are two main differences with the standard definition. First, he considered articles that were published on the reference year. Second, due to the lack of data, he estimated the number of articles published during the t years preceding year p in journal i ($n_{i,p}^*$) based on the growth of articles of *American Economic Review*.

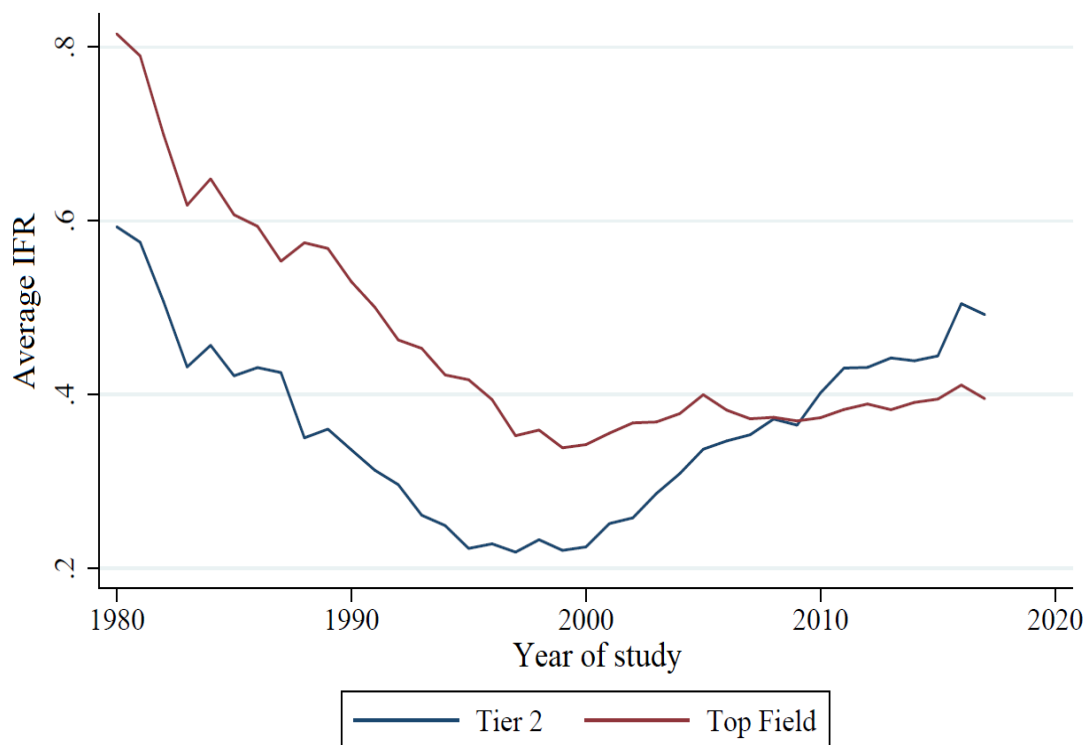
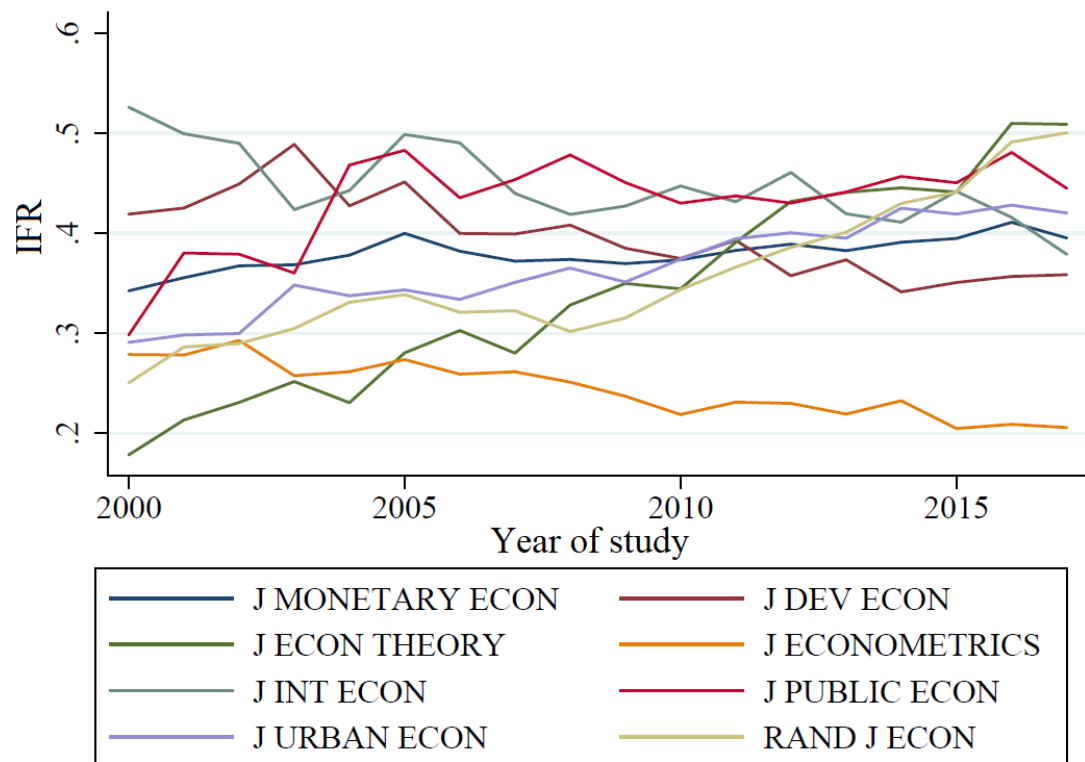


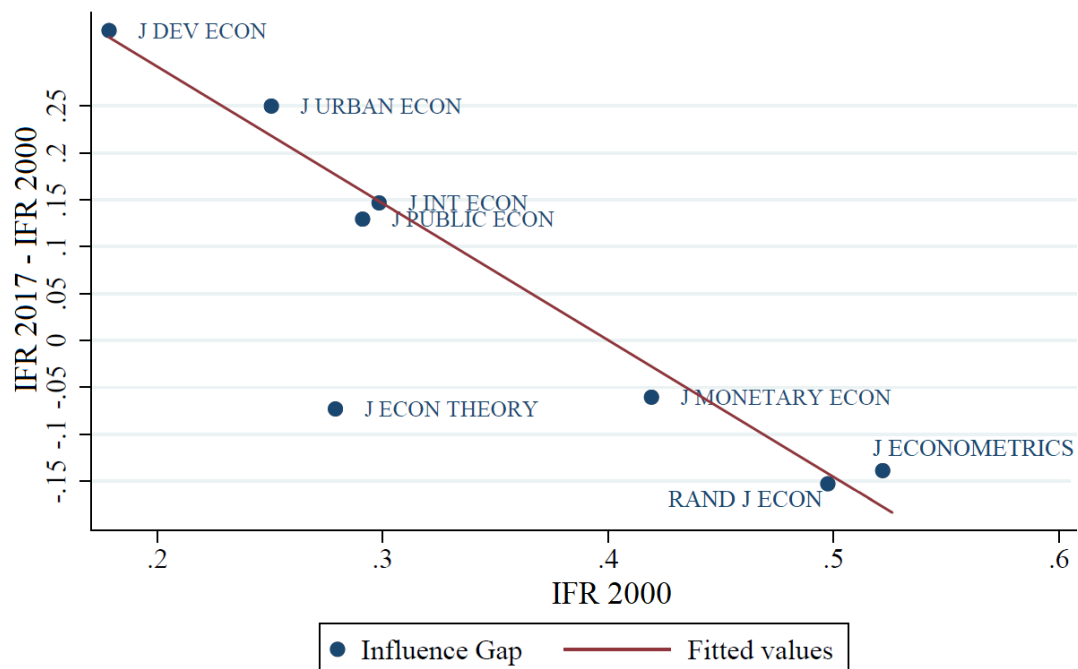
Fig. 3.2.1 Influence gap

2. By 2017, their impact factor ratio was 0.39, which means that Top 5 articles received on average over 150% more citations than Top Field articles, almost as high as it was in the mid-1990s. Moreover, 2010 was the first time since 1980 that Tier 2 journals received more citations per article than Top Field. The trends after 2000 are remarkable and against common perception, as they show that Top 5 journals are becoming considerably less central, and researchers are attaching more importance to other top journals in more recent years.

As fields in economics are not homogeneous, it is important to take a closer look at the impact factor ratio of individual field journals to analyse the diversity of Top Field journals. Prior to 2000, the impact factor ratio of all top field journals decreased and no trends in dispersion were seen. For this reason, this part focuses on the dynamics after 2000. Based on the Top Journals dataset, Figure 3.2.2a calculates the impact factor ratio of each Top Field journal after 2000. This figure reveals that the impact factor ratios of the field journals have experienced a remarkable convergence; the impact factor ratio of those fields with low impact factors is increasing, while the ratio of those fields with initially higher impact factor is decreasing. The only exception is the *Journal of Economic Theory*, which has experienced a continuous decline.



(a) Impact factor ratio per journal



(b) Change in influence

Fig. 3.2.2 Influence gap of Top Field journals

Figure 3.2.2b highlights the convergence as those journals with low influence at the beginning of the period show a strong positive change, while those with initially high influence show a strong negative change. The figure shows a scatter plot of the change in impact factor ratio from 2000 to 2017 (y-axis: $IFR_{i,2017} - IFR_{i,2000}$) against the impact factor ratio of 2000 (x-axis: $IFR_{i,2000}$). The figure also shows the best linear fit. The corresponding slope of the regression line is significantly negative, even including the outlier for *Journal of Economic Theory* ($\hat{\beta} = -1.311$, $s_{\hat{\beta}} = 0.297$, $p = 0.004$). Excluding the outlier, the significance is even stronger ($\hat{\beta} = -1.458$, $s_{\hat{\beta}} = 0.093$, $p = 0.000$).⁸

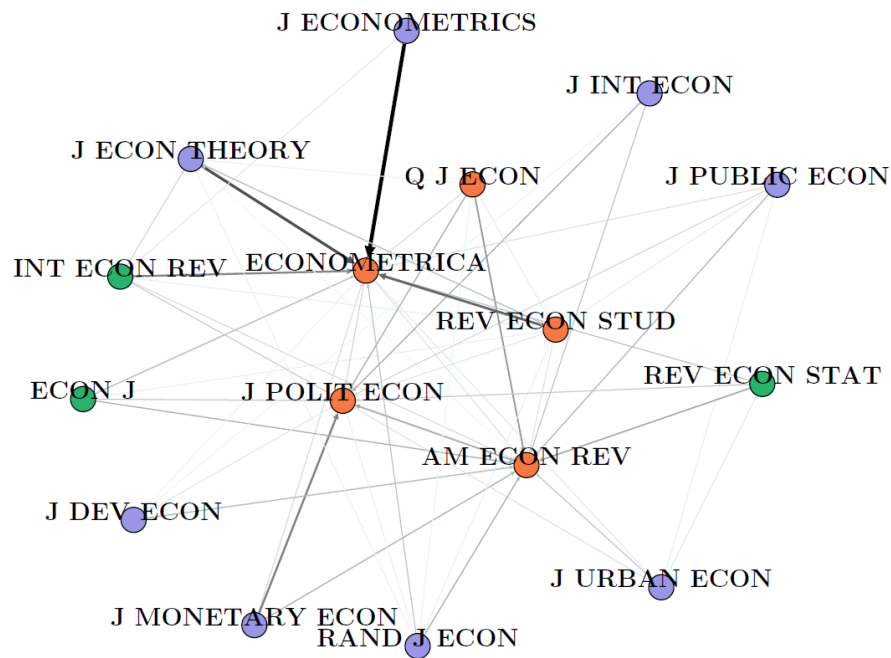
The evidence presented in this section reveals three important trends about the influence of top journals in economics. First, prior to 2000, Top 5 journals demonstrated a sharp increase in their influence and distanced themselves from other top journals. Second, after 2000, Tier 2 journals significantly reduced the influence gap with Top 5 journals. In contrast, Top Field journals kept the influence difference stable. Finally, due to the heterogeneity among economic fields, a closer look at the dynamics of Top Field journals reveals that after 2000 those journals with initially low influence are becoming more relevant, while those with initially high influence are fading out, thus creating a convergence in the influence gap among the fields.

The previous observations relied on the impact factor ratio as a measurement of the influence of a journal. Although quite popular, the impact factor ratio does have some weaknesses. In particular, it does not consider the citations' sources, the number of self-citations, and the change in the number of references per article in a journal. To prove that the results are robust to these considerations, Appendix 3.B.1 replicates the previous exercises using article influence, a measurement formalised by Thompson Reuters. Appendix 3.B.1 shows that these results are robust to the way influence is measured.

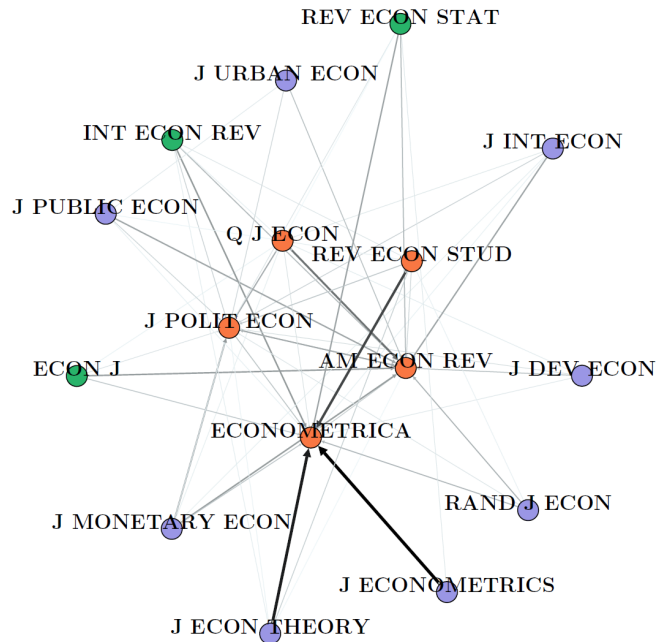
Finally, this section investigates further the citations between the top journals in the 100-Journals dataset. To be consistent with the timespan of the impact factor, the analysis only considers citations between articles that differ at most ten years from the year of publication and divides them based on their source: the same journal (journal self-citations), Top 5, Tier 2, Top Field and other journals. Table 3.2.2 shows the results.

Table 3.2.2 shows that the number of references has grown across all the top journals. Moreover, the number of references from Top Field to Top 5 journals changed from 2.491 in the 2000s to 3.160 in the 2010s, an increase of 27% in one decade. This evidence suggests that the average field researcher is citing more Top 5 journals.

⁸As the sample size is small and as the Breusch-Pagan test does not detect significant heteroskedasticity, normal OLS standard errors are reported, rather than robust standard errors.

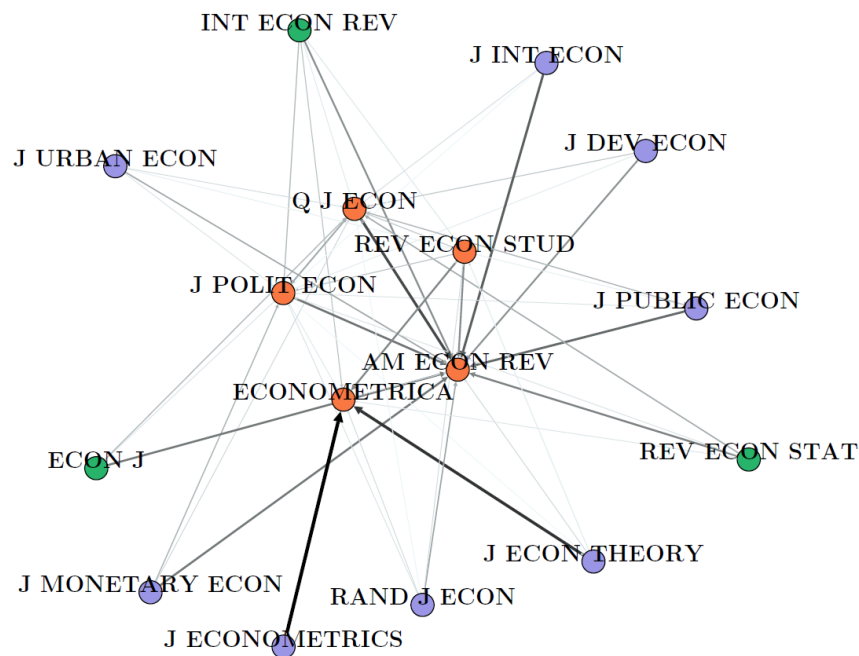


(a) Citation Network 1980

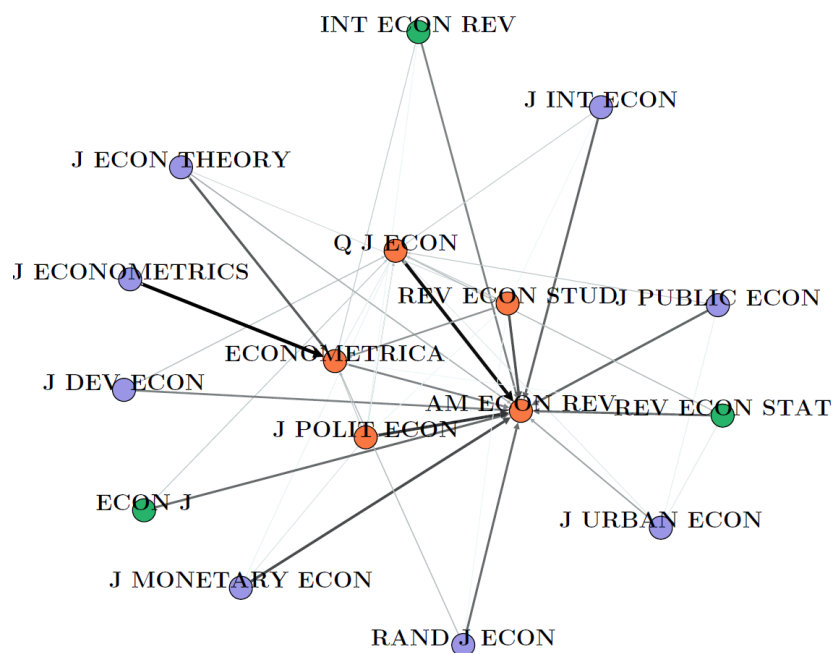


(b) Citation Network 1990

Fig. 3.2.3 Journals' citation network before 2000



(a) Citation Network 2000



(b) Citation Network 2010

Fig. 3.2.4 Journals' citation network after 2000

Table 3.2.2 Evolution of references

Citations from	Citations to	Average number of references			
		1980-1989	1990-1999	2000-2009	2010-2017
Top 5	Self-citations	1.48	1.513	1.488	1.69
	Top 5	2.246	2.436	2.329	3.061
	Tier 2	0.384	0.305	0.407	0.504
	Top Field	1.342	1.263	1.208	1.505
	Other journals	6.639	7.701	8.095	10.073
Tier 2	Self-citations	0.817	0.667	0.665	0.686
	Top 5	2.945	2.803	2.913	3.415
	Tier 2	0.243	0.224	0.309	0.395
	Top Field	1.252	1.312	1.362	1.807
	Other Journals	6.556	7.407	8.647	10.568
Top Field	Self-citations	1.62	1.611	1.575	1.946
	Top 5	2.462	2.648	2.491	3.160
	Tier 2	0.427	0.379	0.506	0.645
	Top Field	0.505	0.599	0.646	0.737
	Other journals	6.084	7.391	8.121	10.244

Note: Citations from the same group of journals exclude journal self-citations, that is, from and to the same particular journal, as those appear under self-citations.

However, the share of citations towards Top 5 decreased from 43% in the 1980s to 37.4% during 2010-2017. Thus, authors publishing in Top Field journals are citing more Top 5 journals, but they are also looking for new sources of ideas in journals that have not been traditionally at the top of their fields.

Table 3.2.2 also shows that Top Field journals have a strong tendency for self-citations. During the 1980s, more than a quarter of the citations were self-citations. This tendency has reduced to a fifth in 2010-2017; albeit still at high levels. Complementing the previous observation, the share of citations that goes to themselves or Top 5 was above 70% in the 1980s, and although it declined by half in 2010-2017, it still means that their primary sources of ideas are Top 5 and themselves. In contrast, self-citations in Top 5 have been around a fifth throughout the study period. However, the references to Top 5, including themselves, represent more than half of their sources. Thus, from the perspective of Top 5, most of their ideas come from their group. Finally, for Tier 2, the core of their references comes from Top 5 and Top Field. In the 1980s, these two groups represented two-thirds of Tier 2's references. While these numbers increased in absolute numbers, their share halved after 2010. Therefore, as in the case of Top Field, Tier 2 authors are reading more of these journals, but they are also reading more new sources. While the scope of this paper is limited to traditional top

journals due to data availability and comparability purposes, Table 3.2.2 strongly advises that it is essential to extend these results as new data emerges to incorporate the rise of new top journals.

Figures 3.2.3 and 3.2.4 aggregate the previous information and display how top journals organised themselves in a core-periphery citation network. In these figures, the colours of the edges are associated with the share of references that the source journal dedicates to the target journal. Also to keep the graphs as clean as possible, the edges that represent less than 5% of the citations of the journal were not included. These graphs highlight how Top 5 journals actively cite each other. In contrast, Top Field and Tier 2 have a few citations among themselves. This evidence suggests that, within a field, authors mostly recognise their own top journal and a group of general interest journals, i.e. the Top 5, but there is almost no recognition for the work in other fields. This observation is the core concept underlying the model presented in the next section.

3.3 Publication Platforms Model

This section presents a theoretical model of journals as platforms matching readers and authors. Section 3.4 uses this model to describe how changes in the publication capacity of journals, the number of authors in economics, and the number of people reading general interest journals explain the evolution of the three trends developed in Section 3.2.

3.3.1 General Environment

There are $\mathcal{F} = \{1, 2, \dots, F\}$ fields of research in economics. For each field $f \in \mathcal{F}$, there is a single journal f and a continuum of authors of measure $n_f > 0$, also referred to as the field size. These authors are also the readers of the papers published in journal f . In addition, there is a general interest journal g that attracts readers from all fields. In particular, a fraction $\alpha_f > 0$ of the mass of authors in field $f \in \mathcal{F}$ reads the general journal, such that the readership of the general journal is $n_g = \sum_{f \in \mathcal{F}} \alpha_f n_f$. The cases where $\alpha_f \geq 1$ represent those fields where the authors of a field read the general interest journal more than their own field journal. Finally, the next section will also consider a special case of the model where $F = 1$. This case represents a single discipline (e.g. economics) with two journals, for which the journal g has a wider readership than journal f , and it will be used to represent a simplified market with Top 5 and Tier 2 journals. In this way the model is able to study both

the dynamics between Top Field journals and Top 5, as well as the dynamics between Tier 2 and Top 5, described in Section 3.2.

In general, within each field f , author i is endowed with an original idea of value $v_i \in [0, \infty)$. The mass of authors with an idea of value v_i is $n_f h(v_i)$, where $h(v_i)$ is a probability density function.⁹ The cumulative distribution associated with $h(v)$ is denoted by $H(v)$, and $h(v)$ is assumed to be positive for each $v \in [0, \infty)$ and its hazard rate, $\frac{h(v)}{1-H(v)}$, is assumed to be decreasing.¹⁰

Authors write papers about their ideas and decide to submit them to either the field journal f , the general interest journal g , or not to submit at all.

There are many authors and, due to capacity constraints, journals are not able to accept all submissions. For that reason, they put requirements on the quality of their published articles to ensure that the number of submissions does not exceed the capacity.

These observations are modelled as follows. Regarding the capacity, journal $j \in \mathcal{F} \cup \{g\}$ can publish at most a mass of κ_j articles, where $\forall f \in \mathcal{F}, \kappa_f + \kappa_g \ll n_f$. In other words, the mass of authors in any field is greater than the publication capacity of the field and general interest journal combined. To keep the model parsimonious, the publication capacity of each journal is exogenous. Regarding the publication requirements, each journal $j \in \mathcal{F} \cup \{g\}$ chooses a threshold $t_j \in [0, \infty)$. The journal directly accepts those submissions that meet the quality threshold, $v_i \geq t_j$. For those submissions that do not meet the threshold, $v_i < t_j$, the journal requests a revision from the author and then the paper is accepted. This revision does not increase the value of the paper, but it does impose a cost $c(t_j - v_i)$ on the author where $c(\cdot)$ is a convex function such that $c(0) = 0$, $c'(x) > 0$, and $c''(x) > 0$.¹¹

This setup is in line with Ellison (2002a) q, r -theory of academic publishing, which postulates that an article's quality is determined by two dimensions, q and r , where q reflects the contribution of the main idea and r the value of robustness checks, typically involved in revisions. Journals accept papers that lack quality in dimension q only if they compensate in dimension r . The assumption that the revision does not increase the value of the paper is

⁹The words *value* and *quality* are used interchangeably in this paper.

¹⁰A decreasing hazard rate holds for many well known distributions that can represent high degrees of skewness such as the exponential, power-law, Weibull, Gamma, and half-normal distribution. As citations are known to follow a highly skewed distribution (Price, 1965), this assumption is likely to hold for academic publications.

¹¹In this model, journals do not reject submissions, but instead set the threshold for publication high enough to ensure that submissions do not exceed the journal's capacity, see Equations (3.1) and (3.2) in Section 3.3.1. A model in which journals do not request revisions, but either reject or accept papers is a special case in which the cost function $c(0) = 0$ and $c(x) = \infty$ for $x > 0$.

made for tractability of the model, and highlights the value that the innovative part of the paper (q) has over the robustness and revision part (r) for influence purposes.

The game has two stages. In the first stage, each journal $j \in \mathcal{F} \cup \{g\}$ chooses a threshold $t_j \in [0, \infty)$ simultaneously with all other journals, where $t := (t_1, \dots, t_F, t_g)$ denotes the vector of thresholds of all journals. In the second stage, all authors in all fields observe t , and simultaneously decide to which journal they wish to submit their paper: the field journal f , the general interest journal g , or not to publish at all. Hence, the strategy of an author of field f with an idea of quality v_i , is a function of t , denoted by $d_f(v_i|t) \in \{f, g, \emptyset\}$, where $d_f(v_i|t) = \emptyset$ describes the choice not to publish at all.

The remainder of this subsection defines the payoff function of the journals and the authors. For that purpose, the following definition is presented:

Definition 3.1 (Journal influence) *The journal influence of $j \in \mathcal{F} \cup \{g\}$ on field $f \in \mathcal{F}$ is defined as*

$$I_j^f(t, d_f(\cdot)) := \mathbb{E}[v | d_f(v|t) = j] = \frac{\int_0^\infty v \mathbb{1}_{d_f(v|t)=j} h(v) dv}{\int_0^\infty \mathbb{1}_{d_f(v|t)=j} h(v) dv}$$

where $\mathbb{1}$ is the indicator function ($\mathbb{1}_A = 1$ if A is true, and 0 otherwise).

Definition 3.1 states that the influence of a journal is equivalent to the expected value of the ideas it contains. For this definition, influence is a field-dependent variable. This feature aims to describe that different fields have different standards. In particular, the general interest journal can be very influential for a field and not influential for other fields. This difference depends on the quality of articles that it contains from each field. Finally, journal influence is constructed to be the theoretical equivalent to the impact factor as both indicators attempt to measure the average value of an article in a journal. However, impact factor uses citation counts to proxy influence, as the real value of an article is unobserved.

In general, higher quality ideas provide more benefits to the journal. These economic benefits may come from each time an article gets viewed. More articles provide more views than less articles, and high-quality articles generate more views than those with low quality. Therefore, this model assumes that journals want to maximise the number of articles it publishes times the journal influence subject to the capacity constraint; i.e. journals want to have as many good articles as possible.

Formally, define $A_j^f(t, d_f(\cdot)) = \int_0^\infty \mathbb{1}_{d_f(v|t)=j} n_f h(v) dv$ as the mass of articles that authors from field f submit to journal j , and let $A_j^f(t) := A_j^f(t, d_f(\cdot))$ and $I_j^f(t) := I_j^f(t, d_f(\cdot))$ be shortcut notations. Anticipating the authors' submission decisions $d(\cdot)$ in stage 2, each field journal $f \in \mathcal{F}$ solves:

$$\max_{t_f \in [0, \infty)} A_f^f(t) I_f^f(t) \quad \text{subject to} \quad A_f^f(t) \leq \kappa_f \quad (3.1)$$

Let $R(t_{-f})$ be the reaction function of field journal f , that is the choice of t_f that solves Problem (3.1) given t_{-f} , where t_{-f} is the vector of all thresholds except journal f .

Similarly, the general interest journal g solves:

$$\max_{t_g \in [0, \infty)} \sum_{f \in \mathcal{F}} A_g^f(t) I_g^f(t) \quad \text{subject to} \quad \sum_{f \in \mathcal{F}} A_g^f(t) \leq \kappa_g \quad (3.2)$$

Expression (3.2) shows that the general interest journal receives benefits from all the articles it publishes, independent of their field. Its corresponding reaction function is denoted by $R(t_{-g})$.

In the second stage, each author i of field f observes the thresholds t and decides whether to publish in the corresponding field journal, the general interest journal, or not to publish at all. The utility of an author depends on publishing in journals with:

1. Larger readership: Authors prefer journals that are read by more authors as their ideas can have greater impact.
2. Higher prestige: Authors prefer journals that, on average, publish better articles in their field because that provides a quality reference of their work.

These preferences naturally lead to the following utility function for authors. The utility of publishing in f is:

$$U_f(v_i|t) = n_f I_f^f(t) - c(\max\{t_f - v_i, 0\})$$

The utility of publishing in g is:

$$U_g(v_i|t) = n_g I_g^f(t) - c(\max\{t_g - v_i, 0\})$$

And the utility of not publishing is:

$$U_\emptyset(v_i) = 0$$

Hence, for both field and general interest journals, the utility is the product of the readership of the journal and its influence, minus the cost of obtaining the publication.

Given the previous three publication alternatives, the author's utility maximisation problem is:

$$\max_{j \in \{f, g, \emptyset\}} U_j(v_i|t) \quad (3.3)$$

and the corresponding reaction function is

$$p_f(v_i|t) = \operatorname{argmax}_{j \in \{f, g, \emptyset\}} U_j(v_i|t) \quad (3.4)$$

For analytic simplicity, authors are assumed to have a tie-breaking rule: their first preference is to publish in a field journal, their second is to publish in a general interest journal, and their third preference is not to publish.

Summarising, the model presents journals as platforms connecting authors and readers. While the reader side is fixed (i.e. a field journal is read by a single field, and the general interest journal is read by a fraction of all fields), the author side is explicitly developed. Note that the decision by an author to use a journal (i.e. submit a paper) creates an externality on other users (i.e. authors). This is a typical characteristic of models of platform competition. However, the nature of the externality in this model is different as it depends on the quality of the user's product; that is, users with high-quality ideas impose a *positive* externality on the other platform users, and users with low-quality ideas (lower than the platform average) impose a *negative* externality on the other platform users.

As a consequence, journals compete for the authors with the best ideas by setting their thresholds strategically. There is vertical competition between the field and general interest journal. Moreover, as the general interest journal is common to all fields and strategically interacts with all field journals, it indirectly creates horizontal competition between field journals.

3.3.2 Equilibrium Analysis

This subsection characterises the equilibrium of the game and provides illustrative examples of the comparative statics.

Definition 3.2 (Equilibrium) *The equilibrium is defined by a vector $t \in \mathbb{R}_+^{F+1}$ and a set of functions $d_f(v|t)$ such that $\forall j \in \mathcal{F} \cup \{g\}$:*

$$t_j \in R(t_{-j})$$

and $\forall f \in \mathcal{F}$ and $\forall t' \in \mathbb{R}_+^{F+1}$

$$d_f(v|t') = p_f(v|t')$$

Definition 3.2 is in fact the definition of a subgame perfect Nash equilibrium of the game. While easy to describe, the current game produces multiple sets of equilibria. Example 3.1 illustrates two of them.

Example 3.1 (Multiple equilibria)

Consider a setting with F fields, each with a mass of $n_f = 100$ authors. The quality of ideas in each field is distributed as $h(x) = e^{-x}$ (i.e. exponential distribution with $\lambda = 1$). The readership share of the general interest journal is $\alpha_f = 0.05$, independent of the field. All journals j (field and general interest) have a publication capacity of $\kappa_j = 20$. Finally, the cost function of revising a paper is given by $c(x) = x^2$. The game contains multiple

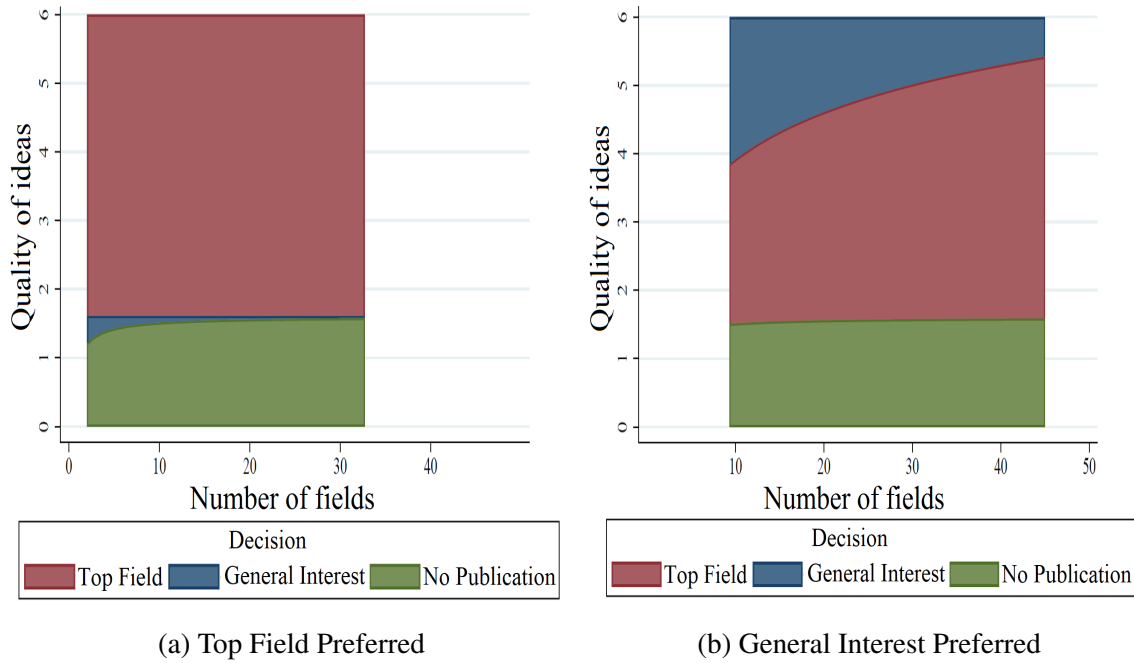
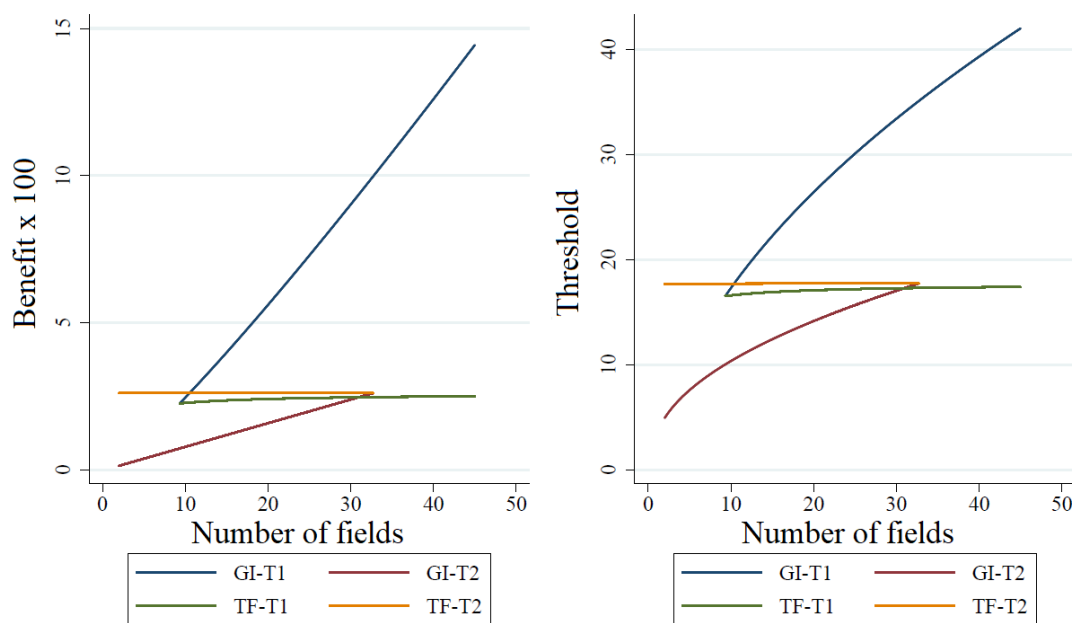


Fig. 3.3.1 Authors decisions

equilibria. This example focuses on two symmetric equilibria where the authors' decision rule is independent of their field. Figure 3.3.1a shows that in the first equilibrium, denoted as "Top Field preferred", authors with the highest quality ideas publish in field journals, authors with middle-quality ideas publish in general interest journals, and those with the lowest quality ideas do not publish at all. In this example, this equilibrium exists in games where the number of fields is less than 33. Beyond this number, the readership of the general interest journal becomes so large that authors would always prefer submitting their paper to the general interest journal rather than the field journal due to the potential outreach they

could have. Figure 3.3.1a illustrates that, as the number of fields increases, the range of quality of papers submitted to field journals does not change, while the quality of the papers submitted to general interest journals witnesses a sharp reduction in its variance.

Figure 3.3.1b presents the second equilibrium, denoted as "General Interest Preferred". The second equilibrium displays a situation where authors with the highest quality ideas publish in the general interest journal, and next group of authors publishes in the field journals, and those with lowest quality ideas do not publish. In contrast to the previous case, this equilibrium only exists when the number of fields exceeds 10. For smaller numbers, the readership of the general interest journal is very small, and due to the limited outreach, authors would always prefer the field journal. In this equilibrium, as the number of fields rises, the mass of authors that may submit their paper to the general interest journal expands as well. This means that the general interest journal is able to increase its threshold and publish higher quality ideas. This, in turn, leads authors who cannot meet the thresholds of the general interest journal to submit their paper to the field journals.



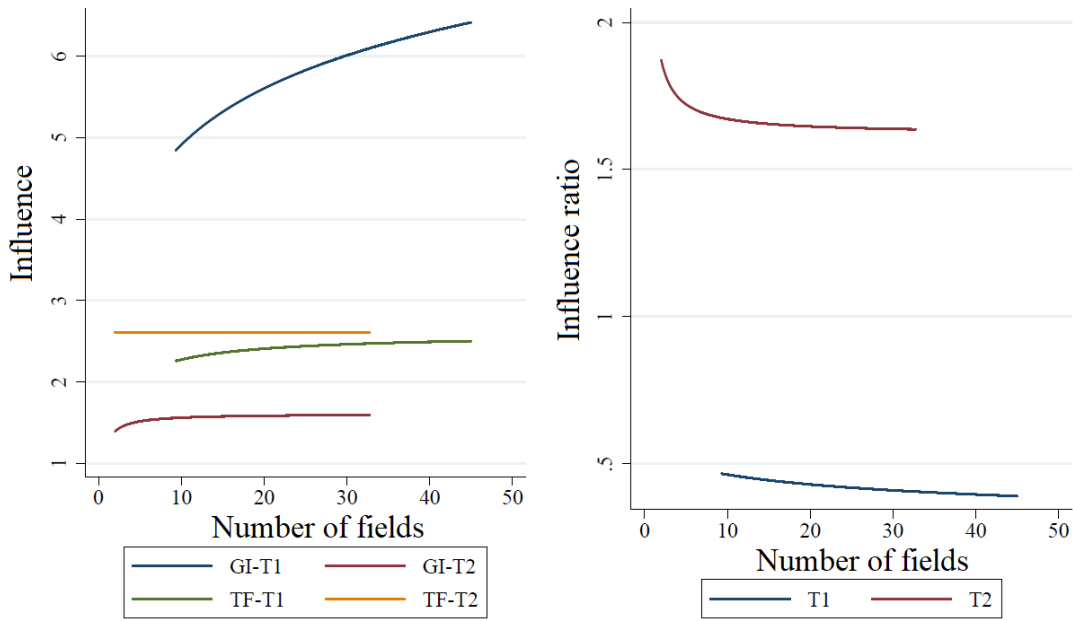
GI=General Interest, TF= Top Field, T1=GI is preferred, T2= TF is preferred

Fig. 3.3.2 Thresholds and authors benefits

Figure 3.3.2 describes the journal's equilibrium thresholds, t_f and t_g , as well as the benefits, $n_j I_j^f(t)$, that all authors receive for publishing in journal $j \in \{f, g\}$.¹² Under the

¹²By symmetry, journal influence $I_j^f(t)$ is the same for all fields $f \in \mathcal{F}$.

"Top Field Preferred" scenario, the threshold of the field journal remains roughly constant as well as the benefits that it provides, while the threshold and benefits of the general journal augments until they become similar to the field journal. Beyond that point, the equilibrium does not hold anymore. In the "General Interest Preferred" scenario, thresholds and benefits of both journal types increase, although the change in the threshold and benefits of the general interest journal outpaces that of the field journal.



GI=General Interest, TF= Top Field, T1=GI is preferred, T2= TF is preferred

Fig. 3.3.3 Influence and influence ratio

Figure 3.3.3 concludes this example by showing the influence ($I_j^f(t)$ for $j \in \{f, g\}$) and the influence ratio ($I_f^f(t)/I_g^f(t)$) of journals in both equilibria. Consistent with the previous figures, the influence of the field journal in the "Top Journal Preferred" equilibrium is independent of the number of fields, while the influence of the general interest increases with the number of fields, albeit keeping a gap with the field journal. Meanwhile, in the second equilibrium, both journals witness an increase in influence, although the influence of the general interest journal grows faster than that of the field journal. In both cases, the influence ratio decreases.

■

Example 3.1 shows that the game described in the previous section is prone to a multiplicity of equilibria that generates different, and sometimes contradictory, predictions. The

fact that there are multiple equilibria can be understood considering the strategic complementarities between journals and top authors, i.e. those with the highest v_i . These strategic complementarities create an anti-coordination game between the field journals and the general interest journal. In particular, the journal that the top authors choose has the highest influence and is therefore able to set a high-quality threshold. This again incites top authors to submit their paper to the top journal whilst discouraging low-quality authors to submit their work there, thus reinforcing the influence gap. These incentives force the alternative journal to set a lower threshold and become second tier relative to the top journal.

What is interesting, though, is that the "top journal" could very well be the field journal instead of the general interest journal. In fact, the example shows that the "Top Field Preferred" is the only symmetric equilibrium if there are less than ten fields, since the readership of the general interest journal is too small in that case to attract the top authors, even if the quality of the journal is high. On the other hand, if the number of fields exceeds 32, then the only symmetric equilibrium allocates the general interest journal as the top journal; in that case, the effect of the broad readership of the journal dominates.¹³

Naturally, whether multiple equilibria can be sustained or not depends on the readership of the different journals. However, the next theorem shows that when the game presents a large number of fields, as well as a large number of authors, the aforementioned ambiguity in equilibria disappears.

Theorem 3.1 ("General Interest Preferred" equilibrium) *Consider the game described in Section 3.3.1. There exists $\bar{n} > 0$, such that, if $n_g \left(\frac{\kappa_f}{\kappa_f + \kappa_g} \right) > n_f > \bar{n}$, then there exists a unique equilibrium, characterised by a threshold vector t and set of author submission functions $d_f(\cdot|t)$, such that $\forall f \in \mathcal{F}$:*

$$t_g > t_f$$

Moreover, for each field $f \in \mathcal{F}$ there are two quality values $c_{1,f} \geq c_{2,f} \geq 0$, such that:

$$d_f(v|t) = \begin{cases} g & \text{if } v > c_{1,f} \\ f & \text{if } c_{2,f} \leq v \leq c_{1,f} \\ \emptyset & \text{otherwise.} \end{cases}$$

Theorem 3.1 proves that, if the size of the fields is large enough and the readership of the general interest journal is substantially larger than the readership of individual field journals, then the equilibrium exists and is unique. Remember that $n_g = \sum_{f \in \mathcal{F}} \alpha_f n_f$, then the last part

¹³For some in-between parameter values there are also asymmetric equilibria in which the field journal is the top journal in some fields, but the second journal in other fields.

of the condition holds if the number of fields (F) or the fraction of readers of the general interest journal (α_f) is large enough.

Central to the proof of uniqueness is the fact that the readership of the general interest journal is larger than the readership of each field. To understand this concept, assume that the influence of the general interest journal was lower than the field journals' influence. However, the outreach of the general interest journal, as measured by its readership, outweighs its influence. Hence, the best authors prefer the general interest journal and submit their papers there, and this in turn makes the general interest journal more attractive for the other authors and more influential than field journals. Thus, at the end of the process, the general interest journal has a higher influence than the field journals.

Beyond uniqueness, this scenario exhibits positive assortativity between the quality of the original idea and the influence and thresholds of the chosen journals. As it occurred in the "General Interest Preferred" case in Example 3.1, the equilibrium mechanisms sort authors into three groups depending on the quality of their ideas. The general interest journal is the most attractive as it has higher readership. Because of this, individuals with the highest quality of ideas prefer the general interest journal to field journals, and due to their submissions, this journal becomes even more attractive than field journals. However, its capacity constraint prevents it from accepting all applications, which then drives the journal to implement a higher threshold. The field journal also places a threshold, but not as high as that of the general interest journal, which in turn attracts those authors with high-quality ideas who are not willing to pay the costs demanded by the general interest journal. This process allocates authors in each field into three groups: i) Authors above the first cut ($v_i > c_{1,f}$) have very high-quality ideas, and their expected benefits outweigh the cost of publishing in the general interest journal; ii) authors of qualities between the first and the second cuts ($c_{2,f} \leq v_i \leq c_{1,f}$) publish in the field journal as they have good ideas, but are not willing to endure the costs demanded by the general interest journal; iii) authors with ideas below the second cut ($v_i < c_{2,f}$) prefer not to publish because journals demand too much effort from them.

The model presented in this section has three groups of parameters: field size (n), journals capacity (κ), and readership shares from each field to the general interest journal (α). Example 3.2 displays how changes in these parameters affect the model's equilibrium which then helps identify the mechanisms underlying the empirical trends of Section 3.2.

Example 3.2 (Changes in parameters)

Consider a setting with 2 fields, each with a mass of $n_f = 1000$ authors. The quality of the ideas in each field distributes $h(x) = e^{-x}$ (i.e. exponential distribution with $\lambda = 1$). The readership share of the general interest journal is $\alpha_f = 0.7$ for both fields. Field journal 1 and the general interest journal have a publication capacity of $\kappa_1 = \kappa_g = 20$, while field journal 2 has a publication capacity of $\kappa_2 = 30$. Finally, the cost function of revising a paper is given by $c(x) = x^2$. The following exercises present the changes in equilibrium as certain parameters change.

The first exercise changes the general interest journal capacity (κ_g) from 20 to 40. Figure 3.3.4 shows that when the capacity of the general interest journal increases, thresholds decrease and influence ratio increases. The underlying intuition is that when a general interest journal's capacity increases then it is willing to publish more ideas, and consequently reduces its threshold. To keep some of their authors, field journals are now pressured to reduce their thresholds. However, by increasing its publications, the general interest journal accepts ideas of lower quality and reduces its influence. Meanwhile, field journals will also have to accept ideas of lower quality, albeit to a lesser extent. This difference between the journals allows the influence ratio to increase.

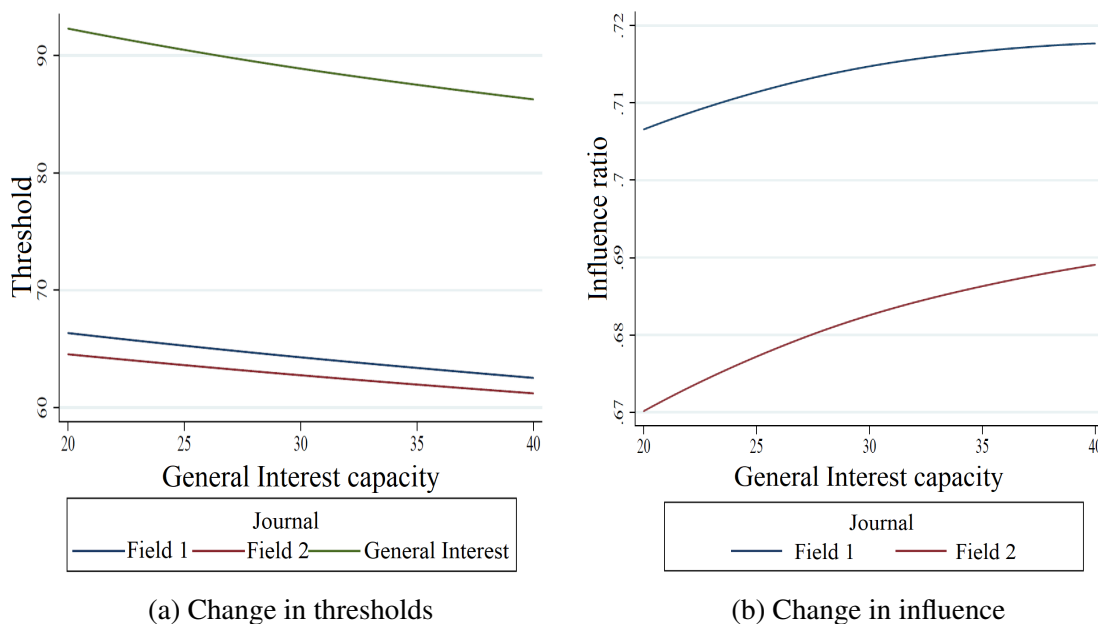


Fig. 3.3.4 Changes in general interest journal capacity (κ_g)

The second exercise changes the size of field 2 (n_2) from 750 to 1500. Figure 3.3.5 presents the influence ratios of the two field journals converge as the size of field 2 increases.

Meanwhile thresholds all rise, although the change in the threshold of field 1 is minimal compared with that of the other two thresholds. This scenario illustrates how the general interest journal creates dependencies between fields. In this case, the increase in the number of authors of field 2 boosts the competition for publishing and allows field 2 journal and the general interest journals to increase the standard quality they are willing to accept. At the same time, authors in field 1 who want to publish in the general interest journal suffer from an increase in the threshold and so decide to move towards the field journal. As a result, field 1 journal now has to increase its threshold to satisfy its capacity constraint. Regarding influence, the general interest journal faces more demand from a field. Hence it can choose better articles in both fields. As it becomes more selective, the influence ratio against field 1 decreases. In contrast, field 2 becomes more attractive as its field now has more research taking place, so it can reduce the influence gap. Therefore, the influence ratio increases.

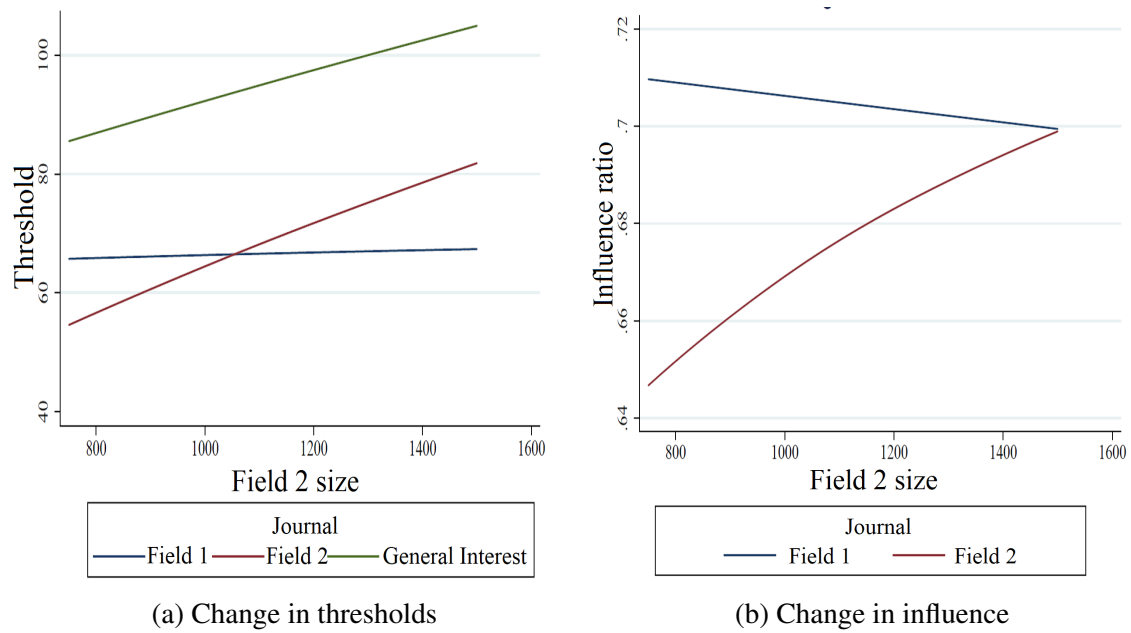


Fig. 3.3.5 Changes in field 2 size (n_2)

The last exercise changes the readership share of the general journal in each field (α_f for both $f = 1, 2$) from 0.7 to 1. Figure 3.3.6 shows that the influence ratio of the field journals converges as α_f increases. However, the mechanism behind the convergence is different from that in the previous exercise. In this case, the positive change in α_f augments the benefits from publishing in the general interest journal and thus increases the number of submissions to the general interest journal. To cope with its capacity constraint, the general interest journal increases its threshold.

As the threshold of the general interest journal increases, authors in each field face a higher cost of publishing a paper in the general interest journal. On the other hand, the benefits of publishing in the general interest journal are increasing as the readership of the general interest journal augments. As such, the difference between these magnitudes then determine what happens to the submission decisions of the authors and the threshold decisions of the field journals.

In the example, for the field where the general interest journal had the highest influence (i.e. field 2), more authors prefer to publish in the general interest journal, and therefore, the average quality of papers submitted by field 2 authors decreases and the influence of the general interest journal in field 2 is reduced. In contrast, for authors from field 1, the cost of publishing in the general interest journal increases more than the benefits, and so several field 1 authors move from the general interest journal towards the field journal. Therefore, the influence of the general interest journal increases in field 1 as it retains better quality ideas. This dynamic reduces the influence ratio ($I_1^1(t)/I_g^1(t)$) of field 1. However, noting the scale of the y-axes in Figure 3.3.6, the changes in the thresholds and influence are very small in contrast to those observed in the previous case.

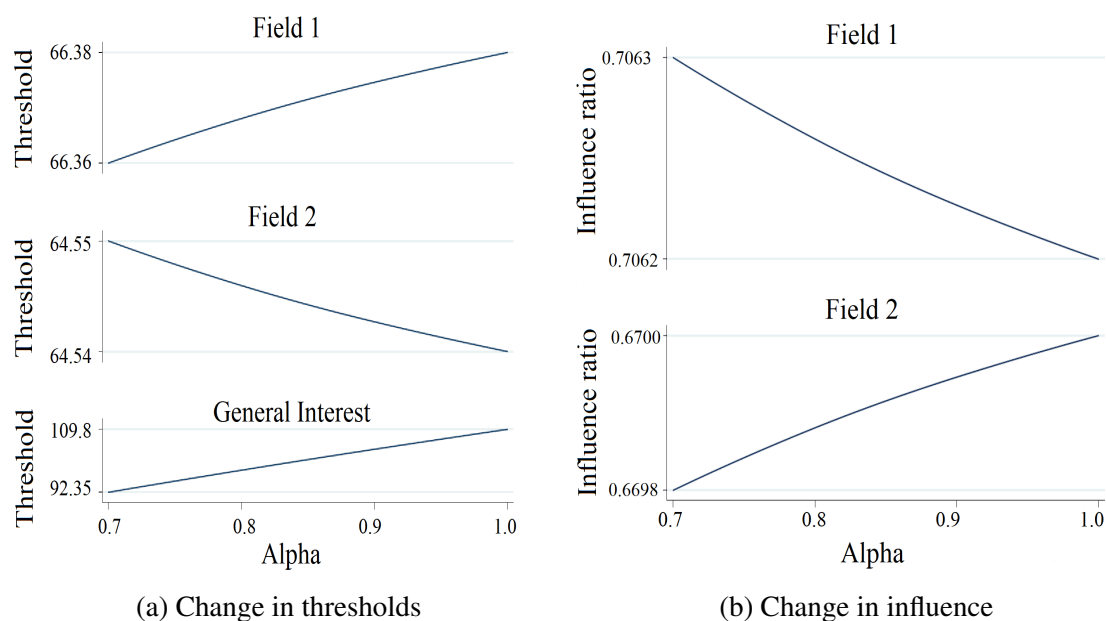


Fig. 3.3.6 Changes in readership (α)



3.4 Comparative Statics

This section applies the model developed in the previous section to explain the mechanisms driving the three empirical trends evidenced in Section 3.2. These trends are:

1. A sharp widening of the influence gap between Top 5 and Tier 2 and Top Field journals between 1980 and 1999.
2. A significant reduction in the influence gap between Tier 2 and Top 5 and the stabilization of the influence gap between Top Field and Top 5 between 2000 and 2017.
3. Convergence in the influence gaps among the Top Field journals between 2000 and 2017.

The following subsections study each of these trends from a theoretical perspective (vis-à-vis the model discussed in Section 3.3) and examine how changes in model's parameters explain the evolution of the influence gap. The comparative statics are not exhaustive as they focus only on the changes that are considered relevant to the empirical observations. However, Lemma 3.6, Lemma 3.7, Lemma 3.8, and Lemma 3.9 in Appendix 3.D provide a full characterisation of the comparative statics derived from the model. Finally, this section measures the influence gap in the following propositions as the difference in influence between journals and not as the ratio of influence. Although both measures (difference and ratio) are quite similar, the measure using the difference is preferred as it provides more general and parsimonious results.

3.4.1 Top 5 and Tier 2 Journals

To obtain a better understanding of the changes of the influence of Tier 2 journals versus Top 5 journals, consider a simplified version of the model with a single field of size n and $\alpha > 1$, and assume that both α and n are large enough to satisfy the conditions of Theorem 3.1. This scenario represents the case where there are two general interest journals, f and g , but the readership of journal g is larger, $n_g = \alpha n_f$, because it is able to attract more readers outside of the (single) field. Hence, journal g represents the set of Top 5 journals and journal f represents the set of Tier 2 journals. For clarity, this subsection refers to journal g as the "Tier 1 journal" and to journal f as the "Tier 2 journal".

For this simplified model, this subsection presents three propositions that explain the trends associated with Tier 2 journals.

Proposition 3.1 *Consider an increase in the publication capacity of the Tier 1 journal (κ_g increases). Then:*

- (a) *The threshold of the Tier 1 journal (t_g) decreases.*
- (b) *The threshold of the Tier 2 journal (t_f) decreases.*
- (c) *The difference in influence between the Tier 1 journal and the Tier 2 journal ($I_g^f(t) - I_f^f(t)$) decreases.*

Implicit in Proposition 3.1 is the fact that the Tier 1 journal attracts submissions from the authors with the best ideas. Hence, an increase in the capacity of the Tier 1 journal allows the Tier 1 journal to lower its quality threshold in order to attract more submissions. This comes at the cost of the Tier 2 journal that loses its highest quality submissions. These incentives force the Tier 2 journal to lower its quality threshold as well. Both journals lose influence, but the loss of influence of the Tier 1 journal is larger. Hence, the influence gap between the Tier 1 journal and the Tier 2 journal decreases.

Proposition 3.2 *Consider an increase of the publication capacity of the Tier 2 journal (κ_f increases). Then:*

- (a) *The threshold of the Tier 1 journal (t_g) increases.*
- (b) *The threshold of the Tier 2 journal (t_f) decreases.*
- (c) *The difference in influence between the Tier 1 journal and the Tier 2 journal ($I_g^f(t) - I_f^f(t)$) increases.*

Proposition 3.2 describes a similar story but from the perspective of the Tier 2 journal. In this case, the increase in capacity of the Tier 2 journal reduces its influence because it accepts articles with lower ideas. This change, in turn, incentivises some authors to move from the Tier 2 journal to the Tier 1 journal. The latter needs to increase its threshold to satisfy its capacity constraint. As a result, the influence of the Tier 1 journal and the influence gap between Tier 1 and Tier 2 increases.

Proposition 3.3 *Consider an increase in the field size (the mass of authors n). Then:*

- (a) *The threshold of the Tier 1 journal (t_g) increases.*
- (b) *The threshold of the Tier 2 journal (t_f) increases.*
- (c) *The difference in influence between the Tier 1 journal and the Tier 2 journal ($I_g^f(t) - I_f^f(t)$) decreases.*

Adding to the previous propositions, Proposition 3.3 shows that if the size of the field (i.e. the number of authors) increases, then competition among authors increases and therefore both journals can demand higher quality ideas. As such, the Tier 2 journal increases its influence more, and the influence gap between journals becomes smaller.

Now, the above propositions are used to explain the dynamics of the impact factors of the Top 5 and Tier 2 journals. Remember that, in Section 3.2, the trends associated with Tier 2 were:

1. The influence gap between Top 5 and Tier 2 journals expanded in the 1980s and 1990s.
2. The influence gap partially rebounded in the 2000s and 2010s.

Based on the previous propositions, the first trend might well be explained by the changes in the publication capacities of these journals. Figure 3.4.1 displays the average number of articles per journal that each of the top journal groups published. Before 2000, the number of articles in the average Top 5 journals declined, while the number of articles in the average Tier 2 journal increased. As a result of these changes, Propositions 3.1 and 3.2 predict an increase in the influence gap between Tier 2 and Top 5 journals.

After 2000, Figure 3.4.1 shows that the average Top 5 journal increased its publication capacity. This tendency is consistent with the information presented by Card and DellaVigna (2013) who explained that Top 5 journals have been systematically decreasing their publications. However, since the late 1990s, *American Economic Review* has steadily increased its number of issues per year and is now responsible for the majority of Top 5 publications. In the case of Tier 2, the capacity changes are ambiguous. On average, publications increased by less than those of the Top 5, however the evolution of the publications was very volatile. Therefore, it is not clear that the effect of the capacity is the only driver of the gap between Tier 2 and Top 5 journals.

There has been a remarkable surge in economic research since 2000 that can complement the explanation of the second trend. Figure 3.4.2 uses the Top-Journals dataset and displays

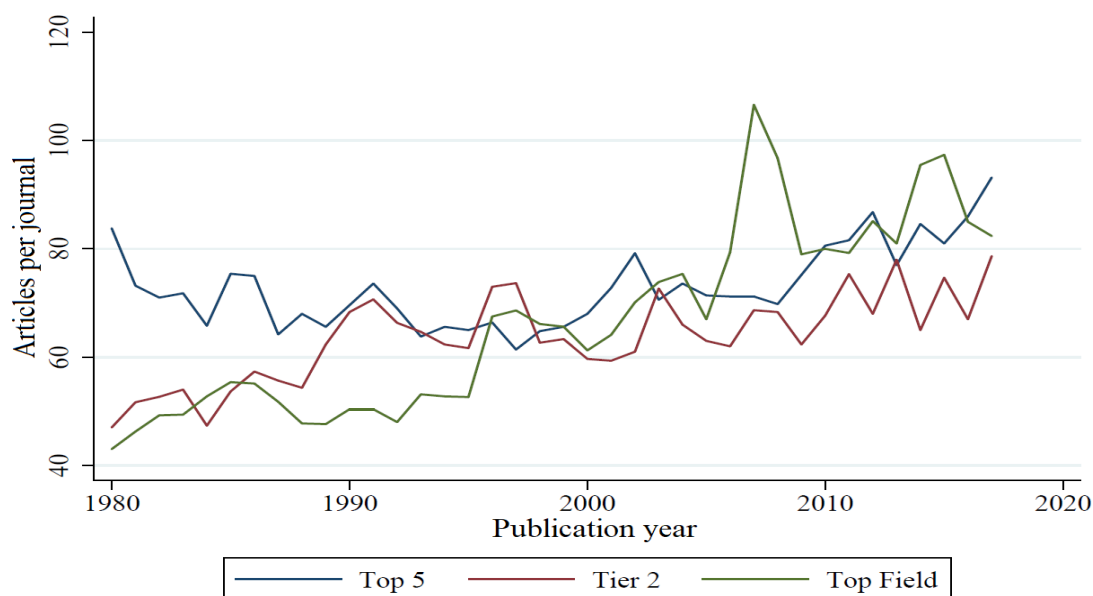


Fig. 3.4.1 Articles per year

the number of times that top journal articles were cited by any journal from the universe of Web of Science. Until the end of the 1990s, the number of citations was growing at a stable pace. However, after 2000, there has been a steep increase in them. The reason behind this change is twofold: First, people are citing more papers; this trend was reported by Ellison (2002b) and can explain the growth in citations until 2000. Second, the number of researchers interested in economics is growing; therefore papers receive more citations. If the second reason is driving the rise in citations, then Proposition 3.3 suggests that this increase may have driven the partial comeback of Tier 2 journals after 2000.

Three empirical observations suggest that the number of researchers interested in economics is growing.

The first evidence that the number of researchers is growing comes from the growth of the number of economic students. Based on the study of Siegfried (2017), between 1990 and 2015, the number of economic baccalaureate degrees awarded in the US grew by 57.7%, with the largest growth episode occurring after 2000. Moreover, the number of PhD students from public institutions grew by 67%, and the number of PhD students from private institutions grew by 69%. In a similar topic, Johnston et al. (2014) showed how the number of economic students in the UK has almost doubled since 2000. The fact that the number of students is growing by these high rates suggests that the number of researchers is growing in a similar way.

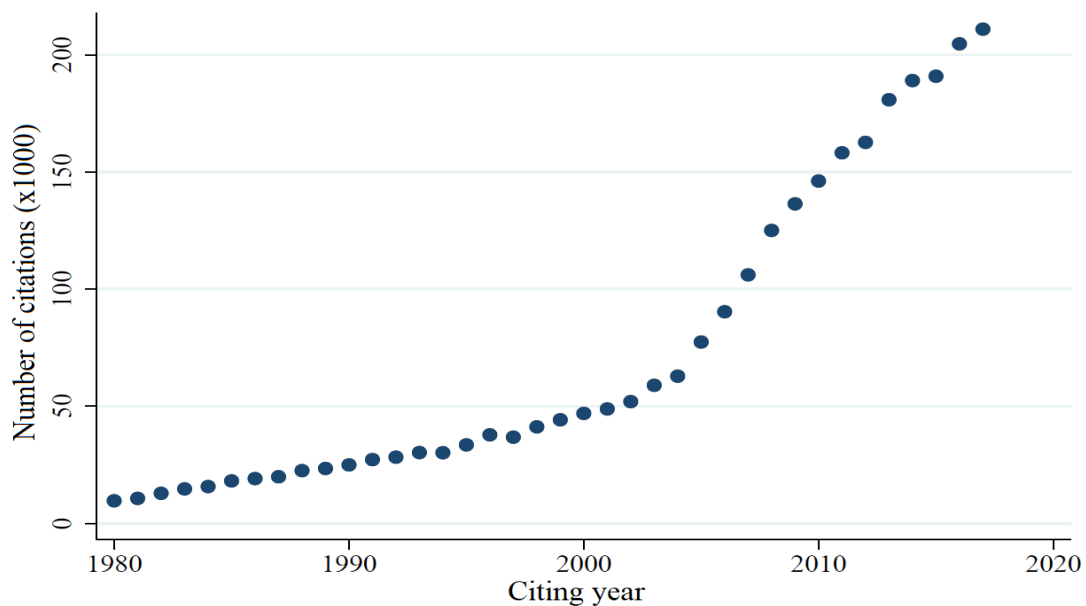


Fig. 3.4.2 Citations per year

The second evidence that the number of researchers is growing comes from the number of authors publishing in economics. Figure 3.4.3 shows that the number of authors who publish in economics is growing at a faster rate than the number of articles. The data used for this graph contains information on 646,927 articles published between 1970 and 2011 in journals listed in EconLit, a bibliography of over 1,000 journals compiled by editors of the *Journal of Economic Literature*.¹⁴

Note that the number of authors in a field is the number of researchers in a field truncated by the capacity of the journals in which they can publish. Hence, the fact that the number of authors is growing can be a consequence of an increase in the field size or an increase in the number of journals (or in their capacities). However, the fact that the growth of the number of authors is higher than the growth of the number of articles, and that the difference significantly increased after the nineties, suggests that there are more papers with multiple authors. Several authors, such as Card and DellaVigna (2013), have claimed that the intense competition between authors drives the increase in co-authorship. As the co-authorship increased more than the capacity of the journals (measured by the number of articles), it is possible to infer that even when capacity constraints are loosened, the number of researchers

¹⁴Econlit dataset is not suitable for the citation analysis performed in Section 3.2. However, its broad universe of journals makes it a very accurate source to approximate the number of authors and relevant journals in economics. Hence, it was used in this part of the analysis.

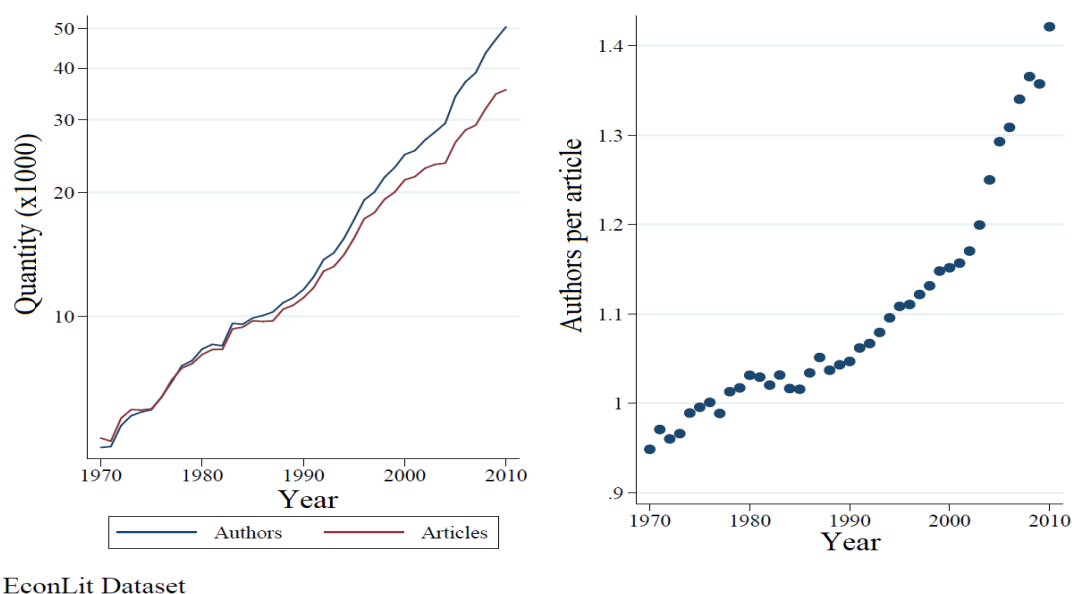


Fig. 3.4.3 Evolution of articles and authors

interested in the field has increased at a rate that makes competition for publication even tougher than before.

The third evidence that the number of researchers is growing comes from the way that old articles are being cited. Figure 3.4.4 uses the Top-Journals dataset to present the life cycle of citations of articles in 1980 and 1990.¹⁵

For each year, all the articles published in top journals were ranked according to the total citations they have in 2017. Then, these articles were divided into ten groups of equal size groups. Decile 10 encompasses the 10% of authors that received the highest number of citations, decile 9 has the next 10% and so on. Figure 3.4.4 plots the average number of citations that each of the three highest deciles received since the publication date. It also plots the evolution of the average number of citations of an article.

In theory, the average life cycle of an article resembles the life cycle of a commercial product. Similar to the logic of the Bass model (Bass, 1969), in the beginning, the article usually has few citations as it is starting to get known in academia. Once a given number of individuals knows the content of the research, it appears more often in academic discussions and starts gaining more citations. Then, the paper becomes less cited as time passes and new research emerges. Thus, when citations are plotted against years after publication, it is expected that this curve displays an inverted U-shape. However, Figure 3.4.4 shows that

¹⁵This analysis was performed for all the years of the sample, and the results did not change.

all these groups follow an inverted U-shape until years around 2000. After that, the articles started witnessing increasing citations again.

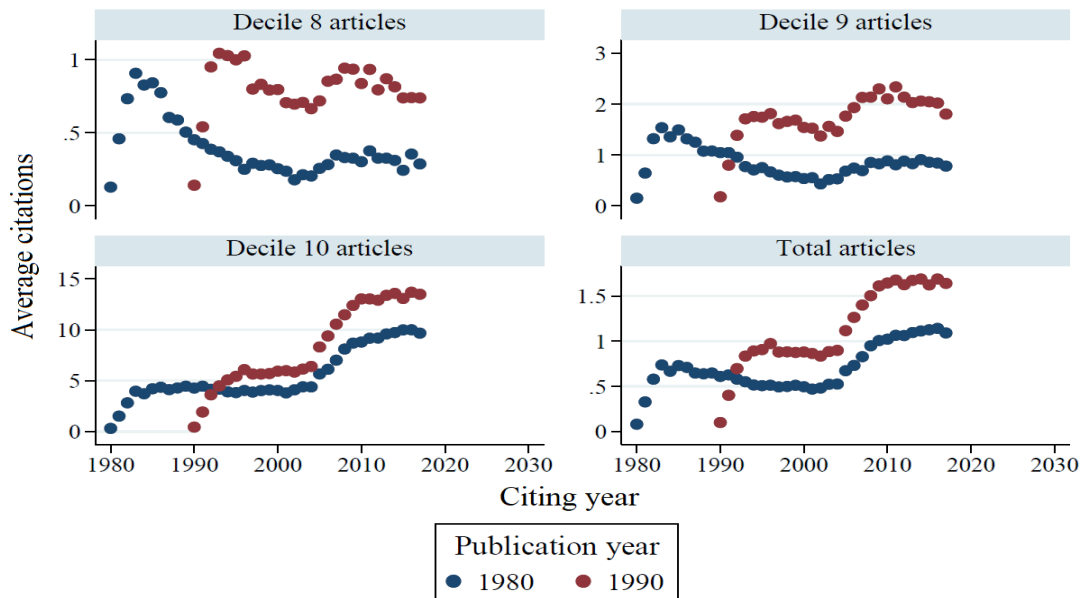


Fig. 3.4.4 Citation evolution for 1980 and 1990

This double peak in old cited papers challenges the idea that the increase in the number of references is enough to explain Figure 3.4.2. If that was the case, new references would go only to current topics and there is no reason for them to be citing old papers. Unless a paper is seminal for a field, it is very unlikely that authors keep on citing it after new research appears. However, it is common for new researchers in a field to read previous literature and find alternative ways to use old results. From this perspective, Figure 3.4.4 suggests an influx of new authors into economic research as a reasonable explanation for the double peak phenomena.

Although none of the previous three observations is conclusive, they strongly advocate for growth in the profession, especially after 2000. Based on Proposition 3.3, the main consequences of the increase in the number of authors are the reduction in the influence gap between Top 5 and Tier 2 journals as well as the increase in the quality requirements of these journals. Therefore, this mechanism provides a reasonable explanation behind the increase in Tier 2 journals' influence after 2000. Moreover, this mechanism also explains that the slowdown of the publication process (i.e. the rise in the number of revisions and time costs) can be explained by a change in the structural parameters, and in this way it complements Ellison (2002b), who claimed that the slowdown was mainly driven by a change in the way that the editors perceived the publication process.

3.4.2 Top 5 and Top Field Journals

The observed trends of journal influence for the Top Field journals relative to the Top 5 trends are:

1. A sharp widening in the influence gap between Top 5 and Top Field journals between 1980 and 1999.
2. A stabilisation of the influence gap between Top Field and Top 5 between 2000 and 2017.
3. Convergence in the influence gaps among the Top Field journals between 2000 and 2017.

The first two trends are similar to the trend of the average influence ratio of the Top 5 versus Tier 2 journals and the explanation is likely to be the same; i.e. the journal capacity of Top 5 journals changed relative to both Tier 2 and Top Field journals. First, note that the intuition of the simplified model of Section 3.4.1 still holds. Proposition 3.1 is trivially generalised to multiple fields (as Example 3.2 suggests). Hence, a decrease in the capacity of Top 5 journals and/or an increase in the capacities of the Top Field journals lead to a widening in the influence gaps. Second, Figure 3.4.1 reveals that these trends in journal capacities have taken place. In 1980, Top 5 journal published on average 80 articles per year, while the average Top Field journal published around 40 articles per year. These number had converged by 1999 when both Top 5 and Top Field journals published around 65 articles per year. Hence, it is plausible that changes in journal capacity explained the observed widening of the influence gap. Appendix 3.B.2 provides further statistical evidence to confirm that journal capacities and journal influence ratios are indeed negatively related, also at the individual journal level. Note however, that in contrast to Tier 2, Top Field journals stabilised after 2000. To explain this, Figure 3.4.1 showed that both the capacity of Top 5 and the capacity of Top Field journals increased. Ergo, these two mechanisms cancelled each other and induced the stabilization process.

The third trend focuses on the heterogeneity among Top Field journals. To understand this trend, it is necessary to go beyond the simplified model of Section 3.4.1 into a model with multiple fields. Hence, the following analysis considers a scenario with $F > 1$ fields whereby the size of each field is large enough to satisfy the conditions of Theorem 3.1. For the purpose of notation, the index of the fields is ordered based on the influence that the

general interest journal has in each field at equilibrium, i.e.:

$$I_g^f(t) \geq I_g^{f'}(t) \Leftrightarrow f \geq f'$$

Proposition 3.4 *Consider the vector $\alpha = \{\alpha_f\}_{f \in \mathcal{F}}$ of fractions of authors in each field f that read the general interest journal. Consider an increase of α to α' , such that $\forall f \in \mathcal{F} : \alpha'_f \geq \alpha_f$ and $\exists f \in \mathcal{F} : \alpha'_f > \alpha_f$. Then, the threshold of the general interest journal (t_g) rises, and there exists a field f^* such that for all $f \in \mathcal{F}$:*

- (a) *If $f < f^*$, then the threshold of field journal f (t_f) increases, and its influence gap ($I_g^f(t) - I_f^f(t)$) increases.*
- (b) *If $f \geq f^*$, then the threshold of field journal f (t_f) decreases, and its influence gap ($I_g^f(t) - I_f^f(t)$) decreases.*

Proposition 3.4 explains the dynamics observed in the third exercise of Example 3.2. If α increases, the general interest journal becomes more attractive to authors, and therefore, the general interest journal needs to raise its threshold to satisfy its capacity constraint. In each field, authors face higher benefits and higher costs to publish in the general interest journals. This trade-off leads towards the convergence of the influence gap. This happens because, in the fields where the field journal has low influence compared with the general interest journal, the increase in the benefits of the general interest journal outweighs the costs, so more authors publish there and therefore drive down its quality. In contrast, in the fields where the field journal has high influence compared with the general interest journal, the increase in the thresholds of the general interest journal outweighs the benefits, so several authors move to the field journal. This incentive structure implies that the general interest journal has only the best of those articles, then its influence will grow. As mentioned in Section 3.2, the label Top 5 appeared at the end of the nineties, so it is probable that this led to the increase in readership of these journals. The information presented in Table 3.2.2 also shows that the average top article is citing more Top 5 papers, which also suggests that the readership increased. However, the evidence is limited. Therefore, it is not possible to attribute the convergence of the field gaps solely to the increase in the readership. Moreover, Example 3.2 showed that the magnitude of the convergence due to changes in this parameter might be very small. For that reason, the following proposition suggests a complementary mechanism of convergence based on the change in the size of those fields that had the largest influence gap.

Proposition 3.5 Consider an increase in the size of field F in which the general interest journal has the highest influence (n_F). There exists an $\bar{\alpha} \in (0, 1)$, such that, if $\alpha_f > \bar{\alpha}$ for every field $f \in \mathcal{F}$, then the threshold of the general interest journal (t_g) increases with n_F , the influence gap between the general interest journal and field F ($I_g^F(t) - I_F^F(t)$) decreases, and, for the other journals, there exists a field f^* such that for all $f \in \mathcal{F} \setminus \{F\}$:

- (a) If $f < f^*$, then the threshold of field journal f (t_f) increases, and its influence gap ($I_g^f(t) - I_f^f(t)$) increases.
- (b) If $f \geq f^*$, then the threshold of field journal f (t_f) decreases, and its influence gap ($I_g^f(t) - I_f^f(t)$) decreases.

Proposition 3.5 formalises the mechanism behind the second exercise of Example 3.2. It describes how the increase in the size of the field with the highest influence gap, field F , generates differentiated benefits and costs of publishing in the field journals. As it was mentioned in Example 3.2, the mechanism differs from the convergence derived from an increase in readership α_f . However, both describe the convergence of the influence ratios of the different fields and an increase of the quality threshold of the general interest journal.

Proposition 3.5 would provide a plausible mechanism of the empirically observed convergence of field journal impact factor ratios, if the field with the largest influence gap, i.e. the smallest impact factor ratio, has expanded in recent years. Has this happened?

To address this question, recall that economics is becoming more empirical. Angrist et al. (2017) showed that during 1980-2000, economic research was predominantly theoretical, but during the 1990s, empirical methods proliferated and by 2000 they were more frequently used than theoretical approaches. Moreover, Hamermesh (2013) focused on three of the Top 5 journals and concluded that there was a shift towards empirical research that is partially due to the emergence of self-collected datasets owing to reduced costs and new technologies. As a consequence, there has been growing research tackling topics of development, finance, labour, and other empirical and applied fields (Card and DellaVigna, 2013).

Complementing these observations from the literature, Figure 3.2.2a reveals that the two field journals with the lowest impact factors were the *Journal of Development Economics* and the *Journal of Urban Economics*. These are precisely the two fields that are empirical in nature, and that have expanded significantly. Thus, it is likely that the mechanism underlying Proposition 3.5 applies and contributes to the explanation of the influence gap convergence between fields.

On a final note, both mechanisms that explain convergence suggest that Top 5 journals are continuously raising their thresholds, and thus it is becoming more difficult to publish

in these journals. This observation supports Card and DellaVigna (2013) as these authors observed that the acceptance rate of papers in Top 5 has been on a trend decline.

3.5 Conclusions

Top 5 journals have been, without doubt, central to economic researchers. However, the evidence presented in this paper shows that tendencies are changing. Prior to 2000, Top 5 were distinguishing themselves from other top journals by being more selective and filtering the best articles. In the late 1990s, these tendencies changed and the influence gap between them and Tier 2 journals started declining. As to Top Field journals, the influence gap stopped its increasing trend but did not recover. Furthermore, a more detailed study to each of the members of the Top Field journals shows that after 2000 the influence gap among the Top Field journals is converging.

The model presented in this paper provides important insights into the mechanisms underlying these trends. With respect to the influence gap prior to 2000, the model suggests that it was driven by changes in the journals' publication capacity. After 2000, the model explains the gap reduction between Tier 2 and Top 5 and the gap stabilization between Top Field and Top 5 via capacity changes and an increase in the number of economists. Finally, the model describes the convergence among Top Field journals in terms of changes in readership and increase in the size of specific research fields.

Economics is a dynamic research field and, as this paper shows, it is continuously changing. The current paper provides mechanisms that explain the changes between 1980-2017. However, the paper does so by focusing on the traditional top journals only. While new journals have been emerging recently, this paper could not explore them due to their novelty and thus lack of data relevant to them. However, as new data becomes available, it will be possible to extend the current research and deepen the understanding of the dynamics underlying ranking positions.

This research paves the way for two open questions that will be central for the discipline moving forward. First, while the "Top 5" label emerged by end-1990s due to the enormous influence this group of journals had, it is not clear how this label will be affected with the rise of second-tier interest journals. Second, with the emergence of the *American Economic Journals*, it is unclear whether these journals should be considered independent entities or parts of conglomerates. In particular, the creation of many journals linked to the same brand can be a new strategy of current academic associations, and it is important to understand the effects that this phenomenon will have on the structure of economic

publications hereafter. Solving these questions will mark the next step toward understanding the mechanisms underlying journals' and academic publications' decisions, which in the end, are the generators of knowledge in any research field.

Appendix to Chapter 3

Appendix 3.A Journal List

- | | |
|--|--|
| 1. AEJ-Applied Economics | 16. Economic Theory |
| 2. AEJ-Macroeconomics | 17. Econometric Theory |
| 3. American Economic Review | 18. Econometrica |
| 4. American J. Agricultural Economics | 19. Economica |
| 5. American Political Science Review | 20. Empirical Economics |
| 6. Brookings Papers On Economic Activity | 21. Energy Economics |
| 7. Cambridge J. Economics | 22. Energy Policy |
| 8. Canadian J. Economics | 23. Energy Journal |
| 9. J. Monetary Economics | 24. Environmental & Resource Economics |
| 10. Ecological Economics | 25. European Economic Review |
| 11. Economic Development And Cultural Change | 26. European J. Political Economy |
| 12. Economic Inquiry | 27. Experimental Economics |
| 13. Economic Journal | 28. Games And Economic Behavior |
| 14. Economics Letters | 29. ILR Review |
| 15. Economic Policy | 30. IMF Economic Review |
| | 31. Industrial And Corporate Change |
| | 32. International Economic Review |

-
- | | |
|--|---|
| 33. International J. Industrial Organization | 56. J. Environmental Economics And Management |
| 34. J. Accounting & Economics | 57. J. The European Economic Association |
| 35. J. Accounting Research | 58. J. Finance |
| 36. J. Applied Econometrics | 59. J. Financial Economics |
| 37. J. Banking & Finance | 60. J. Financial Intermediation |
| 38. J. Business | 61. J. Financial And Quantitative Analysis |
| 39. J. Business & Economic Statistics | 62. J. Health Economics |
| 40. J. Business Venturing | 63. J. Human Resources |
| 41. J. Comparative Economics | 64. J. Industrial Economics |
| 42. J. Consumer Research | 65. J. International Business Studies |
| 43. J. Development Economics | 66. J. International Economics |
| 44. J. Development Studies | 67. J. International Money And Finance |
| 45. J. Economic Behavior & Organization | 68. J. Labor Economics |
| 46. J. Economic Dynamics & Control | 69. J. Law & Economics |
| 47. J. Economic Geography | 70. J. Law Economics & Organization |
| 48. J. Economic Growth | 71. J. Money Credit And Banking |
| 49. J. Economic Literature | 72. J. Political Economy |
| 50. J. Economics & Management Strategy | 73. J. Population Economics |
| 51. J. Economic Perspectives | 74. J. Public Economics |
| 52. J. Economic Surveys | 75. J. Risk And Uncertainty |
| 53. J. Economic Theory | 76. J. Urban Economics |
| 54. J. Econometrics | 77. Labour Economics |
| 55. J. Empirical Finance | 78. Land Economics |
| | 79. Management Science |

- | | |
|---|--|
| 80. Marketing Science | 90. Research Policy |
| 81. Mathematical Finance | 91. Resource And Energy Economics |
| 82. Oxford Bulletin Of Economics And Statistics | 92. Review Of Economic Dynamics |
| 83. Oxford Economic Papers-New Series | 93. Review Of Economics And Statistics |
| 84. Oxford Review Of Economic Policy | 94. Review Of Economic Studies |
| 85. Public Choice | 95. Review Of Financial Studies |
| 86. Quarterly J. Economics | 96. Review Of International Economics |
| 87. RAND J. Economics | 97. Review Of World Economics |
| 88. Regional Science And Urban Economics | 98. Scandinavian J. Economics |
| 89. Regional Studies | 99. World Bank Economic Review |
| | 100. World Development |

Appendix 3.B Robustness Analysis

3.B.1 Influence Using Article Influence Score

The impact factor does not take into account the source of the citations. As an alternative, Pinski and Narin (1976) developed an indicator, called Influence weight, which gives more weight to citations from journals that themselves have a high Influence weight. Formally, for journals $i \in \mathcal{J}$ and $j \in \mathcal{J}$, let c_{ij} be the number of references in journal i that cite journal j , and let $s_i = \sum_j c_{ij}$ be the total number of references in journal i . Then, the Influence weight IW_i of journal i is the solution to

$$IW_i = \sum_{j \in \mathcal{J}} \frac{c_{ji}}{s_j} IW_j$$

that is, the principal eigenvector of the row-normalized citation matrix $C := (\frac{c_{ij}}{s_i})_{i,j \in \mathcal{J}}$. Normalising this measure by the number of articles a_i of a journal, Pinski and Narin (1976) obtained a measure called Influence Per Publication (IPP), that is, $IPP_i = IW_i/a_i$. Palacios-Huerta and Volij (2004) showed that the IPP indicator is the unique indicator satisfying

invariance axioms of reference intensity and splitting journals, as well as weak homogeneity and consistency.

The IPP indicator has a practical implementation in the Article Influence Score (AIS) of the Eigenfactor™ project (Bergstrom et al., 2008). AIS defines the citation matrix in a similar way as the impact factor, but corrects for journal self-citations and also ensures that the citation matrix is ergodic, see Bergstrom and West (2008) for details.¹⁶ This is the implementation followed for this study.

As with the Impact Factor, this paper defines elements in the citation matrix C as the fraction of references in articles published in journal i in year t , that cite articles published in journal j in year $t - 10$ to $t - 1$. This study uses the 100-Journals database, as it is the only dataset to our consideration that contains individual citations from article to article, see Section 3.2.1 for a description.

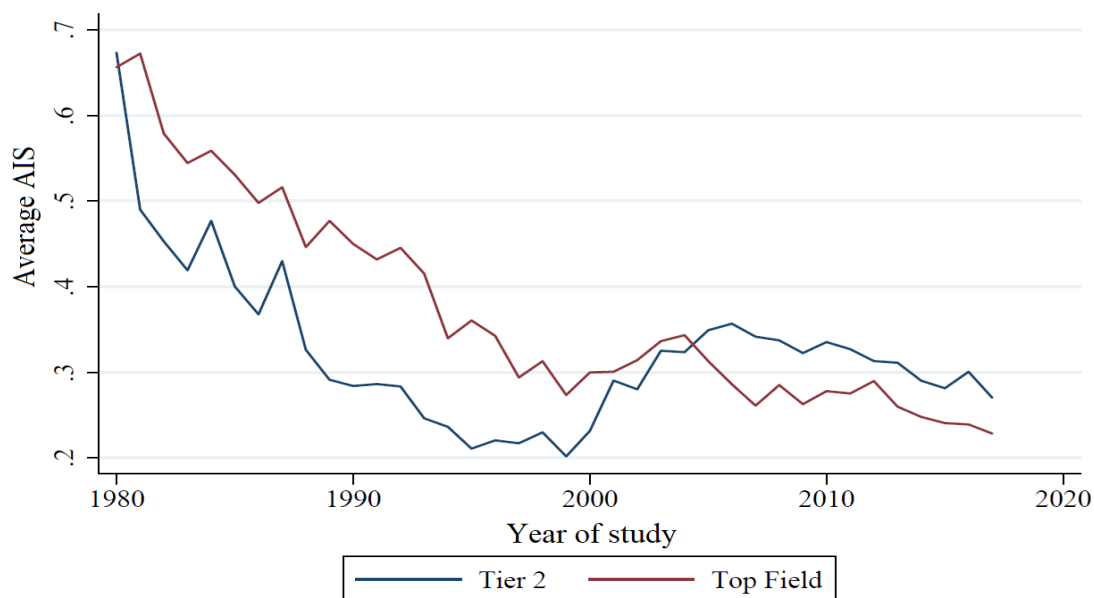
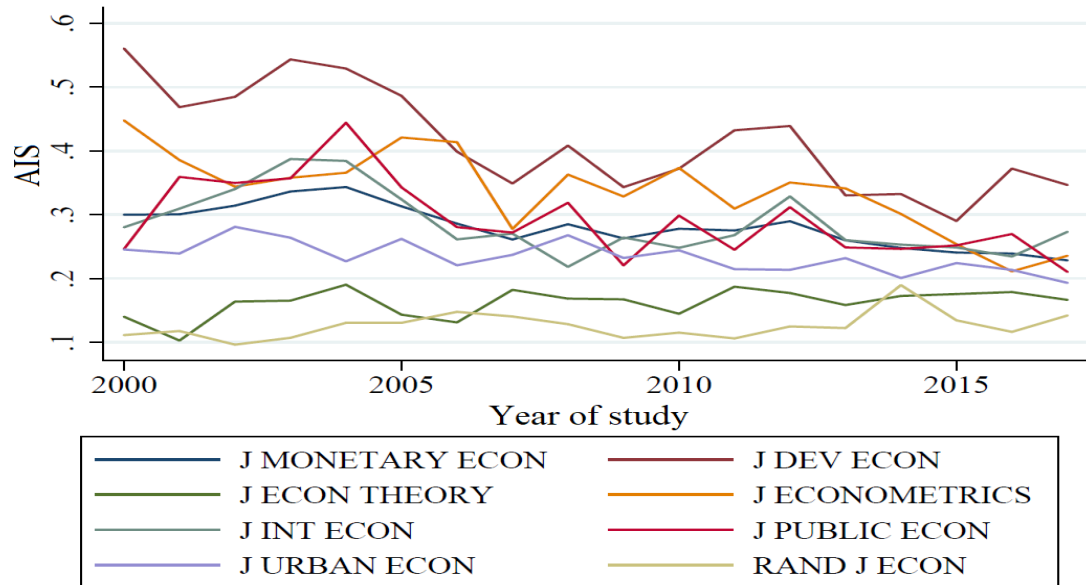


Fig. 3.B.1 Influence gap using article influence

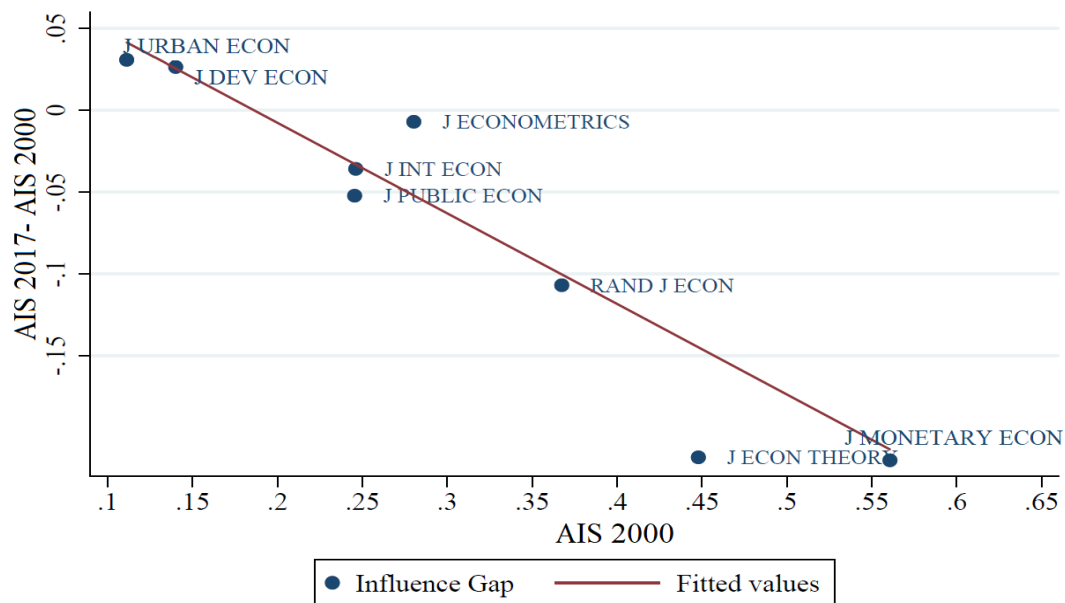
Figure 3.B.1 shows the trends of the ratios of the average AIS of Tier 2 and Top Field journals relative to the average AIS of the Top 5 journals, corresponding to the plot for Impact factors in Figure 3.2.1. Figure 3.B.1 shows that the trends observed in Figure 3.2.1 are robust to changes in the influence measure. In both cases, the influence gap between Top 5 and other top journals increased before 2000. After that, the ratio between Top 5 and Tier 2

¹⁶Bergstrom and West (2008) define Eigenfactor EF_i as the solution to $EF_i = \alpha \sum_{j \in \mathcal{J}} \frac{c_{ji}}{s_j} EF_j + (1 - \alpha) \frac{a_i}{\sum_j a_j}$, similarly to PageRank (Brin and Page, 1998). Whereas West and Bergstrom choose the PageRank value of $\alpha = 0.85$, this study sets $\alpha = 1$, making the measure similar to the Influence weight of Pinski and Narin (1976).

declined. Contrary to the impact factor measurements, the gap between Top 5 and Top Field kept growing after 2000.



(a) Article influence per journal



(b) Change in influence

Fig. 3.B.2 Influence gap of Top Field journals using article influence

Regarding the convergence of fields influence after 2000, Figure 3.B.2 shows the the AIS ratios of the individual field journals relative to the AIS of the Top 5 journals. This plot is the AIS-equivalent of Figure 3.2.2. This figure, again, displays a strong convergence of

the AIS ratios after 2000. A scatter plot of change of AIS ratios versus AIS ratios in 2000 reveals a strong and significant negative relationship ($\hat{\beta} = -0.544$, $s_{\hat{\beta}} = 0.087$, $p = 0.001$). Remarkably, the Journal of Economic Theory is not anymore a clear outlier when considering the AIS measure.

Based on these results it is possible to conclude that the trends presented in the main text are robust to the measure of influence.

3.B.2 Capacity Effects on Influence

From the theoretical model presented in Section 3.3, the publication capacity (i.e. the number of articles per year) that Top 5 publish should have a significant effect on the influence ratio. This section confronts that hypothesis with data from the Top-Journals dataset. The interest variables are the influence ratio (measured as the impact factor ratio ($IFR_{j,t}$)) as well as the number of articles that Top 5 published in that year ($\kappa_g(t)$). Due to the panel characteristics of this exercise, two steps were performed to check whether there is a relationship between these two variables.

The first step, displayed in Table 3.B.1, confirmed that these variables have unit roots. This paper uses the Im-Pesaran-Shin test for panel data to verify this hypothesis because it admits unbalanced panels.

Table 3.B.1 Stationarity tests

Im-Pesaran-Shin unit-root test for	P-Value
$IFR_{j,t}$	0.1338
$\kappa_g(t)$	0.2508
$\Delta IFR_{j,t}$	0.0000
$\Delta \kappa_g(t)$	0.0000
Ho: All panels contain unit roots	
Ha: Some panels are stationary	

Based on the previous results, the second step replicates a co-integration test to identify the relationship between a given journal's capacity and influence. Co-integration was tested using Kao procedure. This method presumes that the co-integration factor does not depend on the journal and therefore identifies the overall relationship between capacity and impact. Table 3.B.2 shows four different versions of the Dickey-Fuller statistical tests.

All the tests support the hypothesis that the variables are co-integrated by having p-values lower than 10%, albeit to varying degrees of significance. The co-integration coefficient is

Table 3.B.2 Cointegration tests

Estimator	P-Value
Modified Dickey-Fuller t	0.0984
Dickey-Fuller t	0.0000
Unadjusted modified Dickey-Fuller	0.0659
Unadjusted Dickey-Fuller t	0.0000

Ho: No cointegration

Ha: All panels are cointegrated

estimated to be 0.036. Its positive sign points that negative changes in the capacity of the general interest journal are associated with negative changes in its influence ratio, i.e. the statistical evidence supports the claim that the negative change in the publication capacity of Top 5 journals drove the increase in the influence gap.

Appendix 3.C Math Notation

This section defines the technical notation that the paper uses in the main text and in the proofs. The notation used in the proofs of Section 3.D, that is not mentioned in this section, was explicitly defined in the main text.

- $h(x)$ is a probability density function with support $[0, \infty)$ such that $h'(x) < 0$.
- $H(x)$ is the corresponding cumulative function.
- $S(x)$ is the corresponding survival function.
- $s(x)$ is the corresponding hazard rate.
- $q(c) = \int_c^\infty \frac{xh(x)}{S(c)} dx$.
- $q_f(c_1, c_2) = \int_{c_2}^{c_1} \frac{xh(x)}{S(c_2) - S(c_1)} dx$.

Appendix 3.D Proofs

Lemma 3.1 *If $s'(c) \leq 0$ for all c , then:*

$$(a) \quad \frac{dq(c)}{dc} \geq 1.$$

$$(b) \quad \lim_{c \rightarrow \infty} \frac{dq(c)}{dc} = 1.$$

$$(c) \lim_{c \rightarrow \infty} s(c)(q(c) - c) = 1.$$

Proof:

Using integration by parts,

$$q(c) = c + \int_c^\infty \frac{S(x)}{S(c)} dx$$

Also, note that

$$\lim_{c \rightarrow \infty} \left(s(c) \int_c^\infty \frac{S(x)}{S(c)} dx \right) = \lim_{c \rightarrow \infty} \left(s(c) \frac{S(c)}{h(c)} \right) = 1$$

On the other hand,

$$\frac{dq(c)}{dc} = s(c)(q(c) - c)$$

Then, by the previous results,

$$\lim_{c \rightarrow \infty} \frac{dq(c)}{dc} = 1$$

Finally,

$$\frac{d^2q(c)}{(dc)^2} = s'(c)(q(c) - c) + s(c)(s(c)(q(c) - c) - 1)$$

Given that $s(x)$ is decreasing, if there exists x such that $\frac{dq(c)}{dc}|_{c=x} \leq 1$, then $\frac{d^2q(c)}{(dc)^2}|_{c=x} < 1$ at that point. Hence $\forall x^* \geq x, \exists \varepsilon \geq 0 : \frac{dq(c)}{dc}|_{c=x^*} \leq 1 - \varepsilon$. But this is a contradiction because $\lim_{c \rightarrow \infty} \frac{dq(c)}{dc} = 1$. Thus, $\frac{dq(c)}{dc} \geq 1$. ■

Lemma 3.2 Let $\kappa_1, \kappa_2 \geq 0$. Then,

$$\lim_{n \rightarrow \infty} \frac{q(H^{-1}(1 - \frac{\kappa_2}{n}))}{q_f(H^{-1}(1 - \frac{\kappa_2}{n}), H^{-1}(1 - \frac{\kappa_2 + \kappa_1}{n}))} \leq 1 + \frac{\kappa_1}{\kappa_2}$$

Proof:

For this proof, and to simplify notation, the arguments of q and q_f are omitted.

Define $c_1 = H^{-1}(1 - \frac{\kappa_2}{n})$ and $c_2 = H^{-1}(1 - \frac{\kappa_2 + \kappa_1}{n})$. From L'Hopital Rule

$$\lim_{n \rightarrow \infty} \frac{q}{q_f} = \lim_{n \rightarrow \infty} \frac{q - c_1}{q_f - c_2 + \frac{c_1 S(c_1) - c_2 S(c_1)}{H(c_1) - H(c_2)}} = \lim_{n \rightarrow \infty} \frac{s(c_1)(q - c_1)}{s(c_1)(q_f - c_2) + \frac{h(c_1)(c_1 - c_2)}{H(c_1) - H(c_2)}}$$

Given that $H(x)$ is increasing and concave, then $\frac{1}{h(c_2)} \leq \frac{h(c_1)(c_1 - c_2)}{H(c_1) - H(c_2)} \leq \frac{1}{h(c_1)}$.

Also, $\frac{\kappa_2}{\kappa_1 + \kappa_2} = \lim_{n \rightarrow \infty} \frac{S(c_1)}{S(c_2)} = \lim_{n \rightarrow \infty} \frac{h(c_1)}{h(c_2)}$. Therefore, using the previous equations and Lemma 3.1

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{q}{q_f} &\leq \lim_{n \rightarrow \infty} \frac{s(c_1)(q - c_1)}{s(c_1)(q_f - c_2) + \frac{h(c_1)}{h(c_2)}} \leq \lim_{n \rightarrow \infty} \frac{1}{s(c_1)(q_f - c_2) + \frac{h(c_1)}{h(c_2)}} \\ &\leq \lim_{n \rightarrow \infty} \frac{h(c_2)}{h(c_1)} = 1 + \frac{\kappa_1}{\kappa_2} \end{aligned}$$

■

The following lemmas assume the uniqueness conditions of Theorem 3.1 (i.e. large number of fields and large size of the fields).

Lemma 3.3 *Consider three quality levels $v_1 < v_2 < v_3$. Assume there is an equilibrium of the game defined by a vector t and a set of functions $d_f(v|t)$. Then,*

$$d_f(v_1|t) = d_f(v_3|t) \Rightarrow d_f(v_1|t) = d_f(v_2|t)$$

Proof:

There are three cases for this proof.

Case 1: $d_f(v_1|t) = d_f(v_3|t) = g$

Due to the monotonicity of $c(x)$,

$$\begin{aligned} U_g(v_i|t) - U_f(v_i|t) &= n_g \mathbb{E}[v|d_f(v|t) = g] \\ &\quad - c(\max\{t_g - v_i, 0\}) - n_f \mathbb{E}[v|d_f(v|t) = f] + c(\max\{t_f - v_i, 0\}) \end{aligned}$$

is monotonous in v_i . Therefore, if $d_f(v_1|t) = d_f(v_3|t) = g$, then $U_g(v_1|t) - U_f(v_1|t) > 0$ and $U_g(v_3|t) - U_f(v_3|t) > 0$. This implies that $U_g(v_2|t) - U_f(v_2|t) > 0 \Rightarrow d_f(v_2|t) = g$.

Case 2: $d_f(v_1|t) = d_f(v_3|t) = f$

Analogous to the previous proof.

Case 3: $d_f(v_1|t) = d_f(v_3|t) = 0$

For an individual of quality v_i , choosing not to publish means:

$$U_g(v_i|t) = n_g \mathbb{E}[v|d_f(v|t) = g] - c(\max\{t_g - v_i, 0\}) < 0$$

and

$$U_f(v_i|t) = n_f \mathbb{E}[v|d_f(v|t) = f] - c(\max\{t_f - v_i, 0\}) < 0$$

As in the previous case, both functions are monotonous in v_i . Hence, if $U_g(v_1|t)$, $U_g(v_3|t)$, $U_f(v_1|t)$, and $U_f(v_3|t)$ are less than 0, then $U_g(v_2|t), U_f(v_2|t) < 0$. Thus, $d_f(v_2|t) = 0$. ■

Lemma 3.4 *Assume there is an equilibrium of the game defined by a vector t and a set of functions $d_f(v|t)$. Then, $\exists c_{1,f} \geq c_{2,f} \geq 0$ such that:*

- (a) $v_i > c_{1,f} \Rightarrow d_f(v_i|t) = g$.
- (b) $c_{1,f} \geq v_i \geq c_{2,f} \Rightarrow d_f(v_i|t) = f$.
- (c) $v_i < c_{2,f} \Rightarrow d_f(v_i|t) = \emptyset$.
- (d) $n_f(H(c_{1,f}) - H(c_{2,f})) = \kappa_f$.

Proof:

Lemma 3.3 implies that there exist three convex sets of values, one for each choice $\{f, g, \emptyset\}$. Hence, per field, there exist two values $c_{1,f} \geq c_{2,f}$ that divide the interval $[0, \infty)$ in three groups, each associated with a decision.

The remainder part is to define which decision corresponds to each set. Regarding the decision not to publish, based on the proof of Lemma 3.3, $U_g(v_i|t)$ and $U_f(v_i|t)$ are increasing functions in v_i , and by construction, if $v_i \geq \max\{t_f, t_g\}$, then $U_g(v_i|t)$ and $U_f(v_i|t)$ are greater than 0. Thus, for high quality values, an author will always prefer to publish than not to publish. Ergo, the set with the smallest qualities (i.e. $[0, c_{2,f})$) contains all the people not willing to publish.

Due to the capacity constraint of the field journal, $n_f(H(c_{1,f}) - H(c_{2,f}))$ can not be greater than κ_f . Moreover,

$$\begin{aligned} U_f(v_i|t) - U_g(v_i|t) &= -(n_g \mathbb{E}[v|d_f(v|t) = g] \\ &\quad - c(\max\{t_g - v_i, 0\}) - n_f \mathbb{E}[v|d_f(v|t) = f] + c(\max\{t_f - v_i, 0\})) \end{aligned}$$

and

$$U_g(v_i|t) = n_g \mathbb{E}[v|d_f(v|t) = g] - c(\max\{t_f - v_i, 0\}) < 0$$

are both monotonously decreasing in t_f . Thence, if the field journal is not satisfying its capacity, then it has incentives to reduce the threshold as this will induce more authors to publish there. Ergo, its benefit increases. As t is already an equilibrium vector, there are no incentives to deviate. Therefore, the only option is that $n_f(H(c_{1,f}) - H(c_{2,f})) = \kappa_f$.

Due to Lemma 3.2, $n_g \mathbb{E}[v|d_f(v|t) = g] - c(\max\{t_j - v_i, 0\}) - n_f \mathbb{E}[v|d_f(v|t) = f] > 0$. Then, for an individual with quality $v_i \geq \max\{t_f, t_g\}$, the general interest journal is strictly preferred to the field journal. Hence, the only possible order is the one stated by the current lemma. ■

Lemma 3.5 *Assume there is an equilibrium of the game defined by a vector t and a set of functions $d_f(v|t)$. Then, for all $f \in \mathcal{F}$ it holds that $t_g \geq t_f$. Furthermore,*

$$(a) \quad t_g > c_{1,f}.$$

$$(b) \quad t_f > c_{2,f}.$$

Proof:

Based on the proof of Lemma 3.4, in equilibrium $n_g \mathbb{E}[v|d_f(v|t) = g] - n_f \mathbb{E}[v|d_f(v|t) = f] > 0$. Hence, if $t_f \geq t_g$ then, for all v_i , $U_g(v_i|t) \geq U_f(v_i|t)$. This implies that no one will choose the field journal, but this contradicts Lemma 3.4 part (d). Therefore, $t_g \geq t_f$.

Using the notation of Lemma 3.4, assume $t_f \leq c_{2,f}$. Then, an individual with quality $v_i < t_f$ will prefer not to publish.

However, $U_f(t_f) = n_f \int_{c_{2,f}}^{c_{1,f}} \frac{xh(x)}{H(c_{1,f}) - H(c_{2,f})} dx > 0$. Moreover, there exists $\varepsilon > 0$ such that $U_f(v'_i) > 0$ for $v'_i = t_f - \varepsilon$. Thus, this individual prefers to publish in the field journal that not to publish, which is a contradiction with Lemma 3.4. Hence, $t_f > c_{2,f}$. The proof of $t_g > c_{1,f}$ mimics the same steps, and therefore is not repeated. ■

Based on the previous results, if there exist an equilibrium, there exists quantities $c_{1,f}$ and $c_{2,f}$ such that:

$$\mathbb{E}[v|d_f(v|t) = g] = \int_{c_{1,f}}^{\infty} \frac{xh(x)}{1 - H(c_{1,f})} dx$$

and

$$\mathbb{E}[v|d_f(v|t) = f] = \int_{c_{2,f}}^{c_{1,f}} \frac{xh(x)}{H(c_{1,f}) - H(c_{2,f})} dx$$

Thus, using the notation defined at the beginning of the section, the influence of a journal can be written as

$$q(c_{1,f}) = \mathbb{E}[v|d_f(v|t) = g] = I_g^f(t)$$

and

$$q_f(c_{1,f}, c_{2,f}) = \mathbb{E}[v|d_f(v|t) = f] = I_f^f(t)$$

The proofs of the theorem and the propositions will use these functions to avoid an excess of notation.

Proof: Theorem 3.1

The proof of this theorem is divided into three parts. The first part defines $\bar{n}$. The second part focuses on identifying an equilibrium of the game. Finally, the third part proves that it is unique. Then, knowing that there exist a unique equilibrium, Lemma 3.5 explains how the thresholds are organised, and Lemma 3.4 provides the information about $c_{1,f}$ and $c_{2,f}$.

Part 1: Characterisation of $\bar{n}$.

Central to the proof of uniqueness is the fact that the audience at each field is large and that there are enough fields such that the audience of the general interest journal is large compared to the audience of each field journal.

To understand the meaning of "large" in this context, consider equation (3.5).

$$\begin{aligned} & \frac{q(H^{-1}(1 - \frac{\kappa_g}{n_f})) - H^{-1}(1 - \frac{\kappa_g}{n_f})}{1 - H(H^{-1}(1 - \frac{\kappa_g}{n_f}))} - \frac{H^{-1}(1 - \frac{\kappa_g}{n}) - H^{-1}(1 - \frac{\kappa_g + \kappa_f}{n_f})}{H(H^{-1}(1 - \frac{\kappa_g}{n_f})) - H(H^{-1}(1 - \frac{\kappa_g + \kappa_f}{n_f}))} \\ &= \frac{q(H^{-1}(1 - \frac{\kappa_g}{n_f})) - H^{-1}(1 - \frac{\kappa_g}{n_f})}{\frac{\kappa_g}{n_f}} - \frac{H^{-1}(1 - \frac{\kappa_g}{n}) - H^{-1}(1 - \frac{\kappa_g + \kappa_f}{n_f})}{\frac{\kappa_f}{n_f}} \end{aligned} \quad (3.5)$$

Due to the concavity of H , and based on Lemma 3.1, when n_f goes to infinity equation (3.5) is greater than 0. Hence, there exists a minimum $\bar{n}_1$ such that, for all fields, if $n_f > \bar{n}_1$, equation (3.5) is positive.

Now, consider equation (3.6).

$$n_f q\left(H^{-1}\left(1 - \frac{\kappa_f}{n_f}\right)\right) - n_g q_f\left(H^{-1}\left(1 - \frac{\kappa_f}{n_f}\right), H^{-1}\left(1 - \frac{\kappa_f + \kappa_g}{n_f}\right)\right) \quad (3.6)$$

As long as $\frac{n_g}{n_f} > 1 + \frac{\kappa_g}{\kappa_f}$, Lemma 3.1, implies that, there exists a minimum $\bar{n}_2$ such that, for all fields, if $n_f > \bar{n}_2$, equation (3.6) is negative.

For the rest of the proof it is necessary to have equation (3.5) positive and equation (3.6) negative. Therefore, the technical requirement behind the theorem is that for each field $n_f > \max\{\bar{n}_1, \bar{n}_2\} = \bar{n}$ and that there are enough fields or providing enough readers to the general interest journal such that $\frac{n_g}{n_f} > 1 + \frac{\kappa_g}{\kappa_f}$. This is summarised by $n_g \left(\frac{\kappa_f}{\kappa_f + \kappa_g} \right) > n_f > \bar{n}$.

Part 2: Existence.

Consider the following four equations per field:

$$n_f(1 - H(c_{1,f})) = \kappa_g^f \quad (3.7)$$

$$n_f(1 - H(c_{2,f})) = \kappa_g^f + \kappa_f \quad (3.8)$$

$$n_f q_f(c_{1,f}, c_{2,f}) - c(t_f - c_{2,f}) = 0 \quad (3.9)$$

$$n_g q(c_{1,f}) - c(t_g - c_{1,f}) = n_f q_f(c_{1,f}, c_{2,f}) - c(\max\{t_f - c_{1,f}, 0\}) \quad (3.10)$$

The parameters of the game are n_f , n_g , κ_g , and κ_f . Whereas κ_g^f has not been defined yet, notice that, if this value is known, equation (3.7) uniquely defines $c_{1,f}$. Then, equation (3.8) uniquely defines $c_{2,f}$. For equation (3.9), the left hand side is a monotonous function in t_f after it is greater than $c_{2,f}$, therefore it is invertible for these values, i.e. equation (3.9) uniquely defines $t_f \geq 0$. Finally, using the previous logic, equation (3.10), uniquely defines $t_g \geq 0$.

Let $K_f(x) : [0, \min_{f \in \mathcal{F}} \{n_f - \kappa_f\}] \rightarrow [0, \infty)$, such that $t_g = K_f(\kappa_g^f)$ satisfy equations (3.7) to (3.10) when κ_g^f is given. From the previous results about the way each equation uniquely determines a parameter, $K_f(x)$ is well defined.

Let $p \in \Delta$, where Δ is an n -dimensional simplex, and define $\kappa_g^f = p_f \kappa_g$. Let $\Gamma(p) : \Delta \rightarrow \Delta$, where

$$\Gamma(p)_f = \frac{K_f(\kappa_g p_f) + \max\{K_f(\kappa_g p_f) - \frac{\sum_{j \in \mathcal{F}} K_j(\kappa_g p_j)}{F}, 0\}}{\sum_{i \in \mathcal{F}} (K_i(\kappa_g p_i) + \max\{K_i(\kappa_g p_i) - \frac{\sum_{j \in \mathcal{F}} K_j(\kappa_g p_j)}{F}, 0\})}$$

Notice that, if there is a vector p such that m of its entries ($\mathcal{S} = \{\pi(1), \pi(2), \dots, \pi(m)\}$) are 0 (i.e. $p_f = 0 \Leftrightarrow f \in \mathcal{S}$). Then, for all $f \in \mathcal{S}$, it holds that $\Gamma(p)_f = \frac{1}{m}$.

Given that Γ is continuous and that it goes from a compact into itself, Brouwer's Theorem states that it has at least a fix point. In this context a fix point means that $\forall f, j \in \mathcal{F}, K_f(\kappa_g^f) = K_j(\kappa_g^j)$. Let p^* be a fix point. Then, from equations (3.7) to (3.10), it is possible to define the values of $c_{1,f}$, $c_{2,f}$, t_f , κ_g^f for each f , and also it is possible to define t_g , which my construction is independent of the field. Furthermore, p^* guarantees that $\sum_{f \in \mathcal{F}} \kappa_g^f = \kappa_g$.

Finally, define t as the vector that include t_g and all the different t_f , and

$$d_f(v_i|t) = \begin{cases} g, & \text{for } v_i \geq c_{1,f} \\ f, & \text{for } c_{2,f} \geq v_i \geq c_{1,f} \\ \emptyset, & \text{for } c_{2,f} \geq v_i \end{cases}$$

As defined above, t and d_f satisfy all the incentives and participation constraints. In particular, from equations (3.9) and (3.10) no individual wants to change its choice. On the other side, from equation (3.8) and the way κ_g^f was defined, all the journals are satisfying its capacity.

Thus vector t and the set of functions $d_f(v_i|t)$ define an equilibrium of the game.

Part 3: Uniqueness.

Given that equations (3.7) to (3.10) define an equilibrium, Lemma 3.5 guarantees that for every field f , $t_g > c_{1,f}$ and $t_f > c_{2,f}$. Moreover, from equation (3.7),

$$1 = -n_f h(c_{1,f}) \frac{dc_{1,f}}{d\kappa_g^f}$$

From equation (3.8),

$$1 = -n_f h(c_{2,f}) \frac{dc_{2,f}}{d\kappa_g^f}$$

From equation (3.9),

$$0 = n_f \frac{dq_f(c_{1,f}, c_{2,f})}{d\kappa_g^f} - c'(t_f - c_{2,f}) \left(\frac{dt_f}{d\kappa_g^f} - \frac{dc_{2,f}}{d\kappa_g^f} \right)$$

From equation (3.10),

$$\begin{aligned} & n_g \frac{dq(c_{1,f})}{d\kappa_g^f} - c'(t_g - c_{1,f}) \left(\frac{dt_g}{d\kappa_g^f} - \frac{dc_{1,f}}{d\kappa_g^f} \right) \\ &= n_f \frac{dq_f(c_{1,f}, c_{2,f})}{d\kappa_g^f} - c'(\max\{t_f - c_{1,f}, 0\}) \left(\frac{dt_f}{d\kappa_g^f} - \frac{dc_{1,f}}{d\kappa_g^f} \right) \end{aligned}$$

Therefore, by replacing the values,

$$\begin{aligned}
& -c'(t_g - c_{1,f}) \frac{dK_f}{d\kappa_g^f} = \frac{n_g(q(c_{1,f}) - c_{1,f})}{n_f(1 - H(c_{1,f}))} + \frac{c'(t_g - c_{1,f})}{n_f h(c_{1,f})} \\
& - \frac{c_{1,f} - c_{2,f}}{H(c_{1,f}) - H(c_{2,f})} \left(1 - \frac{c'(\max\{t_f - c_{1,f}\})}{c'(t_f - c_{2,f})} \right) + \frac{c'(\max\{t_f - c_{1,f}\})}{n_f h(c_{2,f})} \\
& - \frac{c'(\max\{t_f - c_{1,f}\})}{n_f h(c_{1,f})}
\end{aligned}$$

Due to part 1 and Lemma 3.1, $\frac{(q_g - c_{1,f})}{(1 - H(c_{1,f}))} - \frac{c_{1,f} - c_{2,f}}{H(c_{1,f}) - H(c_{2,f})} > 0$, and from Lemma 3.5, $\frac{c'(t_g - c_{1,f})}{n_f h(c_{1,f})} - \frac{c'(\max\{t_f - c_{1,f}\})}{n_f h(c_{1,f})} > 0$. Ergo, $\frac{dK_f}{d\kappa_g^f} < 0$. This implies that $K_f(x)$ is invertible.

Define $T(t_g) := \sum_{j \in \mathcal{F}} K_f^{-1}(t_g)$. From the definition of $K_f(t_g)$, it holds that $T(0) = \sum_{j \in \mathcal{F}} n_f$, $\lim_{t_g \rightarrow \infty} T(t_g) = 0$ and $T'(t_g) < 0$. Therefore, for a given κ_g , there exists a unique t_g that allows the system of equations to have a solution, and as κ_g is a parameter of the model, then there exists a unique t_g in the game that satisfy all the equations. So any alternative equilibrium will not satisfy one of these equations:

If equation (3.10) is not satisfied, then $\exists \varepsilon > 0$ such that either an individual with an idea of quality $c_{1,f} + \varepsilon$ will deviate from g to f , or an individual with idea of quality $c_{1,f} - \varepsilon$ will deviate from f to g .

If equation (3.9) is not satisfied, then $\exists \varepsilon > 0$ such that either an individual with an idea of quality $c_{2,f} + \varepsilon$ will deviate from f to $\emptyset$, or an individual with idea of quality $c_{2,f} - \varepsilon$ will deviate from $\emptyset$ to f .

If equation (3.8) is not satisfied, then either the field journal is publishing more than its capacity, so the equilibrium is not possible, or it produces fewer articles than desired, in which case it would like to deviate to a lower threshold.

If equation (3.7) is not satisfied, then either the general interest journal is publishing more than its capacity, so the equilibrium is not possible, or it produces fewer articles than desired, in which case it would like to deviate to a lower threshold.

For these reasons, all equilibria have to bind equations (3.7) to (3.10). However, there is a unique solution for the system of equations. Hence, the equilibrium is unique.

With the proof of existence and uniqueness, Lemma 3.5 and Lemma 3.4 finish the proof of the theorem. ■

From Theorem 3.1, the unique equilibrium of the game is characterized by t_f , t_g , κ_g^f , $c_{1,f}$, and $c_{2,f}$. The following lemmas describe how these variables change in response to three

model parameters: α_s , n_s , and κ_g . The outcomes of these proofs are the main inputs to prove Proposition 3.1 to Proposition 3.5.

For these proofs, the following auxiliary variables are defined:

$$A_f = \frac{n_g(q(c_{1,f}) - c_{1,f})}{n_f(1 - H(c_{1,f}))} + \frac{c'(\max\{t_f - c_{1,f}, 0\})}{n_f h(c_{2,f})} - \frac{c'(\max\{t_f - c_{1,f}, 0\})}{n_f h(c_{1,f})} \quad (3.11)$$

$$- \frac{c_{1,f} - c_{2,f}}{H(c_{1,f}) - H(c_{2,f})} \left(1 - \frac{c'(\max\{t_f - c_{1,f}, 0\})}{c'(t_f - c_{2,f})}\right) + \frac{c'(t_g - c_{1,f})}{n_f h(c_{1,f})}$$

$$X_f = \frac{q(c_{1,f})}{A_f} \quad (3.12)$$

$$W_f = \frac{c'(t_g - c_{1,f})}{A_f} \quad (3.13)$$

$$V_f = \frac{1}{A_f} \left(c'(\max\{t_f - c_{1,f}, 0\}) \right. \quad (3.14)$$

$$\left. + \left(1 - \frac{c'(\max\{t_f - c_{1,f}, 0\})}{c'(t_f - c_{2,f})}\right) \left(\frac{q_f(c_{1,f}, c_{2,f}) - c_{2,f}}{H(c_{1,f}) - H(c_{2,f})} \right) \right)$$

$$B_f = \frac{1}{A_f} \left(\left(1 - \frac{c'(\max\{t_f - c_{1,f}, 0\})}{c'(t_f - c_{2,s})}\right) \frac{c_{1,f}(1 - H(c_{1,f})) - c_{2,f}(1 - H(c_{2,f}))}{H(c_{1,f}) - H(c_{2,f})} \right. \quad (3.15)$$

$$\left. - \left(1 - \frac{c'(\max\{t_f - c_{1,f}, 0\})}{c'(t_f - c_{2,s})}\right) 2q_f(c_{1,f}, c_{2,f}) \right.$$

$$+ \frac{n_g}{n_f} (q(c_{1,f}) - c_{1,f}) + c'(t_g - c_{1,f}) \frac{1 - H(c_{1,s})}{h(c_{1,f})n_f}$$

$$\left. + c'(\max\{t_f - c_{1,f}, 0\}) \frac{1 - H(c_{2,f})}{h(c_{2,s})n_f} - \frac{(1 - H(c_{1,f}))}{h(c_{1,f})n_f} \right)$$

The conditions on the size of each field stated by Theorem 3.1 guarantee that $A_f > 0$. Therefore, $X_f, W_f \geq 0$.

Lemma 3.6 Consider an equilibrium characterised by t_f , t_g , κ_g^f , κ_f , $c_{1,f}$, and $c_{2,f}$, and a marginal change in κ_g . Then,

$$\begin{aligned} (a) \quad \frac{dt_g}{d\kappa_g} &= \frac{-1}{\sum_{j \in \mathcal{F}} W_j}. \\ (b) \quad \frac{dt_f}{d\kappa_g} &= \frac{-1}{c'(t_f - c_{2,f})} \left(\frac{c_{1,f} - c_{2,f}}{H(c_{1,f}) - H(c_{2,f})} + \frac{c'(t_f - c_{2,f})}{n_f h(c_{2,f})} \right) \frac{W_f}{\sum_{i \in \mathcal{F}} W_i}. \\ (c) \quad \frac{d(q(c_{1,f}) - q_f(c_{1,f}, c_{2,f}))}{d\kappa_g} &= - \left(\frac{(q_g - c_{1,f})}{\kappa_g^f} - \frac{(c_{1,f} - c_{2,f})}{\kappa_f} \right) \frac{W_f}{\sum_{i \in \mathcal{F}} W_i}. \end{aligned}$$

Proof:

From the proof of Theorem 3.1, the equilibrium of the game has to satisfy equations (3.7) to (3.10). When the system is derived in κ_g , the following derivatives are obtained:

From equation (3.7),

$$\frac{d\kappa_g^f}{d\kappa_g} = -n_f h(c_{1,f}) \frac{dc_{1,f}}{d\kappa_g}$$

From equation (3.8),

$$\frac{d\kappa_g^f}{d\kappa_g} = -n_f h(c_{2,f}) \frac{dc_{2,f}}{d\kappa_g}$$

From equation (3.9),

$$-c'(t_f - c_{2,f}) \frac{dt_f}{d\kappa_g} = \left(\frac{c_{1,f} - c_{2,f}}{H(c_{1,f}) - H(c_{2,f})} + \frac{c'(t_f - c_{2,f})}{n_f h(c_{2,f})} \right) \frac{d\kappa_g^f}{d\kappa_g}$$

From equation (3.10),

$$-c'(t_g - c_{1,f}) \frac{dt_g}{d\kappa_g} = A_f \frac{d\kappa_g^f}{d\kappa_g}$$

Finally, $\sum_{f \in \mathcal{F}} \kappa_g^f = \kappa_g$. Then, $\sum_{f \in \mathcal{F}} \frac{d\kappa_g^f}{d\kappa_g} = 1$. By doing algebraic substitutions using the previous equations,

$$\frac{dt_g}{d\kappa_g} = \frac{-1}{\sum_{f \in \mathcal{F}} W_f}$$

Then,

$$\frac{d\kappa_g^f}{d\kappa_g} = \frac{W_f}{\sum_{i \in \mathcal{F}} W_i}$$

$$\frac{dt_f}{d\kappa_g} = \frac{-1}{c'(t_f - c_{2,f})} \left(\frac{c_{1,f} - c_{2,f}}{H(c_{1,f}) - H(c_{2,f})} + \frac{c'(t_f - c_{2,f})}{n_f h(c_{2,f})} \right) \frac{W_f}{\sum_{i \in \mathcal{F}} W_i}$$

and

$$\frac{d(q(c_{1,f}) - q_f(c_{1,f}, c_{2,f}))}{d\kappa_g} = - \left(\frac{(q_g - c_{1,f})}{\kappa_g^f} - \frac{(c_{1,f} - c_{2,f})}{\kappa_f} \right) \frac{W_f}{\sum_{i \in \mathcal{F}} W_i}$$

■

Lemma 3.7 Consider an equilibrium characterised by t_f , t_g , κ_g^f , κ_f , $c_{1,f}$, and $c_{2,f}$, and a marginal change in α_s . Then,

$$\begin{aligned} (a) \quad \frac{dt_g}{d\alpha_s} &= \frac{\sum_{j \in \mathcal{F}} X_j}{\sum_{j \in \mathcal{F}} W_j}. \\ (b) \quad \frac{dt_f}{d\alpha_s} &= \frac{-1}{c'(t_f - c_{2,f})} \left(\frac{c_{1,f} - c_{2,f}}{H(c_{1,f}) - H(c_{2,f})} + \frac{c'(t_f - c_{2,f})}{n_f h(c_{2,f})} \right) \left(\frac{\sum_{j \in \mathcal{F}} n_s X_j W_j - \sum_{j \in \mathcal{F}} n_s X_j W_f}{\sum_{j \in \mathcal{F}} W_j} \right). \\ (c) \quad \frac{d(q(c_{1,f}) - q_f(c_{1,f}, c_{2,f}))}{d\alpha_s} &= - \left(\frac{(q_g - c_{1,f})}{\kappa_g^f} - \frac{(c_{1,f} - c_{2,f})}{\kappa_f} \right) \left(\frac{\sum_{j \in \mathcal{F}} n_s X_j W_j - \sum_{j \in \mathcal{F}} n_s X_j W_f}{\sum_{j \in \mathcal{F}} W_j} \right). \end{aligned}$$

Proof:

From Theorem 3.1, the equilibrium of the game has to satisfy equations (3.7) to (3.10). When the system is derived in α_s , the following derivatives are obtained:

From equation (3.7),

$$\frac{d\kappa_g^f}{d\alpha_s} = -n_f h(c_{1,f}) \frac{dc_{1,f}}{d\alpha_s}$$

From equation (3.8),

$$\frac{d\kappa_g^f}{d\alpha_s} = -n_f h(c_{2,f}) \frac{dc_{2,f}}{d\alpha_s}$$

From equation (3.9),

$$-c'(t_f - c_{2,f}) \frac{dt_f}{d\alpha_s} = \left(\frac{c_{1,f} - c_{2,f}}{H(c_{1,f}) - H(c_{2,f})} + \frac{c'(t_f - c_{2,f})}{n_f h(c_{2,f})} \right) \frac{d\kappa_g^f}{d\alpha_s}$$

From equation (3.10),

$$n_f q(c_{1,f}) - c'(t_g - c_{1,f}) \frac{dt_g}{d\alpha_s} = A_f \frac{d\kappa_g^f}{d\alpha_s}$$

Finally, $\sum_{f \in \mathcal{F}} \kappa_g^f = \kappa_g$. Therefore, $\sum_{f \in \mathcal{F}} \frac{d\kappa_g^f}{d\alpha_s} = 0$. By doing algebraic substitutions using the previous equations,

$$\frac{dt_g}{d\alpha_s} = \frac{\sum_{f \in \mathcal{F}} n_s X_f}{\sum_{f \in \mathcal{F}} W_f}$$

Then,

$$\frac{d\kappa_g^f}{d\alpha_s} = \frac{\sum_{j \in \mathcal{F}} n_s X_f W_j - \sum_{j \in \mathcal{F}} n_s X_j W_f}{\sum_{j \in \mathcal{F}} W_j}$$

$$\begin{aligned} \frac{dt_f}{d\alpha_s} = & \frac{-1}{c'(t_f - c_{2,f})} \left(\frac{c_{1,f} - c_{2,f}}{H(c_{1,f}) - H(c_{2,f})} \right) \left(\frac{\sum_{j \in \mathcal{F}} n_s X_f W_j - \sum_{j \in \mathcal{F}} n_s X_j W_f}{\sum_{j \in \mathcal{F}} W_j} \right) \\ & + \frac{-1}{c'(t_f - c_{2,f})} \left(\frac{c'(t_f - c_{2,f})}{n_f h(c_{2,f})} \right) \left(\frac{\sum_{j \in \mathcal{F}} n_s X_f W_j - \sum_{j \in \mathcal{F}} n_s X_j W_f}{\sum_{j \in \mathcal{F}} W_j} \right) \end{aligned}$$

and

$$\begin{aligned} \frac{d(q(c_{1,f}) - q_f(c_{1,f}, c_{2,f}))}{d\alpha_s} = & - \left(\frac{(q_g - c_{1,f})}{\kappa_g^f} - \frac{(c_{1,f} - c_{2,f})}{\kappa_f} \right) \left(\frac{\sum_{j \in \mathcal{F}} n_s X_f W_j - \sum_{j \in \mathcal{F}} n_s X_j W_f}{\sum_{j \in \mathcal{F}} W_j} \right) \end{aligned}$$

■

Lemma 3.8 Consider an equilibrium characterised by t_f , t_g , κ_g^f , κ_f , $c_{1,f}$, and $c_{2,f}$, and a parametrization of the size of all the journals, $n_f(\beta)$. If there is a marginal change in β , then:

$$(a) \quad \frac{dt_g}{d\beta} = \frac{\sum_{f \in \mathcal{F}} (\sum_{j \in \mathcal{F}} \alpha_j \frac{dn_f}{d\beta}) X_f + B_f \frac{dn_f}{d\beta}}{\sum_{f \in \mathcal{F}} W_f}.$$

$$(b) \quad \frac{d\kappa_g^f}{d\beta} = \frac{\sum_{j \in \mathcal{F}} \left((\sum_{i \in \mathcal{F}} \alpha_i \frac{dn_f}{d\beta}) X_j + B_j \frac{dn_j}{d\beta} \right) W_j - \sum_{j \in \mathcal{F}} \left((\sum_{i \in \mathcal{F}} \alpha_i \frac{dn_f}{d\beta}) X_j + B_j \frac{dn_j}{d\beta} \right) W_f}{\sum_{j \in \mathcal{F}} W_j}.$$

(c)

$$\begin{aligned} \frac{dt_f}{d\beta} = & \frac{-1}{c'(t_f - c_{2,f})} \left(\frac{c_{1,f} - c_{2,f}}{H(c_{1,f}) - H(c_{2,f})} + \frac{c'(t_f - c_{2,f})}{n_f h(c_{2,f})} \right) \frac{d\kappa_g^f}{d\beta} \\ & + \frac{1}{c'(t_f - c_{2,f})} \left(\frac{c_{1,f}(1 - H(c_{1,f})) - c_{2,f}(1 - H(c_{2,f}))}{H(c_{1,f}) - H(c_{2,f})} \right. \\ & \left. - c'(t_f - c_{2,f}) \frac{1 - H(c_{2,f})}{n_f h(c_{2,f})} + 2q_f(c_{1,f}, c_{2,f}) \right) \frac{dn_f}{d\beta}. \end{aligned}$$

(d)

$$\begin{aligned} \frac{d(q(c_{1,f}) - q_f(c_{1,f}, c_{2,f}))}{d\beta} = & - \left(\frac{q_g - c_{1,f}}{\kappa_g^f} - \frac{c_{1,f} - c_{2,f}}{\kappa_f} \right) \frac{d\kappa_g^f}{d\beta} \\ & + \frac{1}{n_f} \left(q(c_{1,f}) - q_f(c_{1,f}, c_{2,f}) - \frac{(1 - H(c_{2,f}))(c_{1,f} - c_{2,f})}{H(c_{1,f}) - H(c_{2,f})} \right) \frac{dn_f}{d\beta}. \end{aligned}$$

Proof:

From Theorem 3.1, the equilibrium of the game has to satisfy equations (3.7) to (3.10). When the system is derived in β , the following derivatives are obtained:

From equation (3.7),

$$\frac{d\kappa_g^f}{d\beta} = \left((1 - H(c_{1,f})) - n_f h(c_{1,f}) \frac{dc_{1,f}}{dn_f} \right) \frac{dn_f}{d\beta}$$

From equation (3.8),

$$\frac{d\kappa_g^f}{d\beta} = \left((1 - H(c_{2,f})) - n_f h(c_{2,f}) \frac{dc_{2,f}}{dn_f} \right) \frac{dn_f}{d\beta}$$

From equation (3.9),

$$\begin{aligned} -c'(t_f - c_{2,f}) \frac{dt_f}{d\beta} = \\ \left(\frac{c_{1,f} - c_{2,f}}{H(c_{1,f}) - H(c_{2,f})} + \frac{c'(t_f - c_{2,f})}{n_f h(c_{2,f})} \right) \frac{d\kappa_g^f}{d\beta} - \left(-2q_f(c_{1,f}, c_{2,f}) \right. \\ \left. + c'(t_f - c_{2,f}) \frac{1 - H(c_{2,f})}{n_f h(c_{2,f})} - \frac{c_{1,f}(1 - H(c_{1,f}) - c_{2,f}(1 - H(c_{2,f})))}{H(c_{1,f}) - H(c_{2,f})} \right) \frac{dn_f}{d\beta} \end{aligned}$$

From equation (3.10),

$$\left(\sum_{f \in \mathcal{F}} \alpha_f \frac{dn_f}{d\beta} \right) q(c_{1,f}) - c'(t_g - c_{1,f}) \frac{dt_g}{d\beta} = A_f B_f \frac{dn_f}{d\beta} + A_f \frac{d\kappa_g^f}{d\beta}$$

Finally, $\sum_{f \in \mathcal{F}} \kappa_g^f = \kappa_g$. Therefore, $\sum_{f \in \mathcal{F}} \frac{d\kappa_g^f}{d\beta} = 0$. By doing algebraic substitutions using the previous equations,

$$\frac{dt_g}{d\beta} = \frac{\sum_{f \in \mathcal{F}} (\sum_{j \in \mathcal{F}} \alpha_j \frac{dn_j}{d\beta}) X_f + B_f \frac{dn_f}{d\beta}}{\sum_{f \in \mathcal{F}} W_f}$$

Then,

$$\begin{aligned} \frac{d\kappa_g^f}{d\beta} = \\ \frac{\sum_{j \in \mathcal{F}} ((\sum_{i \in \mathcal{F}} \alpha_i \frac{dn_i}{d\beta}) X_f + B_f \frac{dn_f}{d\beta}) W_j - \sum_{j \in \mathcal{F}} ((\sum_{i \in \mathcal{F}} \alpha_i \frac{dn_i}{d\beta}) X_j + B_j \frac{dn_j}{d\beta}) W_f}{\sum_{j \in \mathcal{F}} W_j} \end{aligned}$$

$$\begin{aligned} \frac{dt_f}{d\alpha_s} = \frac{-1}{c'(t_f - c_{2,f})} \left(-\frac{c_{1,f}(1 - H(c_{1,f})) - c_{2,f}(1 - H(c_{2,f}))}{H(c_{1,f}) - H(c_{2,f})} \right. \\ \left. + c'(t_f - c_{2,f}) \frac{1 - H(c_{2,f})}{n_f h(c_{2,f})} - 2q_f(c_{1,f}, c_{2,f}) \right) \frac{dn_f}{d\beta} \\ - \frac{1}{c'(t_f - c_{2,f})} \left(\frac{c_{1,f} - c_{2,f}}{H(c_{1,f}) - H(c_{2,f})} + \frac{c'(t_f - c_{2,f})}{n_f h(c_{2,f})} \right) \frac{d\kappa_g^f}{d\beta} \end{aligned}$$

and

$$\begin{aligned} \frac{d(q(c_{1,f}) - q_f(c_{1,f}, c_{2,f}))}{d\alpha_s} &= - \left(\frac{(q_g - c_{1,f})}{\kappa_g^f} - \frac{(c_{1,f} - c_{2,f})}{\kappa_f} \right) \frac{d\kappa_g^f}{d\beta} \\ &+ \frac{1}{n_f} \left(q(c_{1,f}) - q_f(c_{1,f}, c_{2,f}) - \frac{(1 - H(c_{2,f}))(c_{1,f} - c_{2,f})}{H(c_{1,f}) - H(c_{2,f})} \right) \frac{dn_f}{d\beta} \end{aligned}$$

■

Lemma 3.9 Consider an equilibrium characterised by t_f , t_g , κ_g^f , κ_f , $c_{1,f}$, and $c_{2,f}$, and a marginal change in κ_s . Then,

$$(a) \quad \frac{dt_g}{d\kappa_s} = \frac{V_s}{\sum_{j \in \mathcal{F}} W_j}.$$

$$(b) \quad \forall f \neq s, \frac{dt_f}{d\kappa_s} = \frac{1}{c'(t_f - c_{2,f})} \left(\frac{c_{1,f} - c_{2,f}}{H(c_{1,f}) - H(c_{2,f})} + \frac{c'(t_f - c_{2,f})}{n_f h(c_{2,f})} \right) \frac{V_s W_f}{\sum_{i \in \mathcal{F}} W_i}.$$

$$(c) \quad \forall f \neq s, \frac{d(q(c_{1,f}) - q_f(c_{1,f}, c_{2,f}))}{d\kappa_s} = \left(\frac{(q_g - c_{1,f})}{\kappa_g^f} - \frac{(c_{1,f} - c_{2,f})}{\kappa_f} \right) \frac{V_s W_f}{\sum_{i \in \mathcal{F}} W_i}.$$

(d)

$$\begin{aligned} \frac{dt_s}{d\kappa_s} &= \frac{-1}{c'(t_s - c_{2,s})} \left(\frac{c_{1,s} - c_{2,s}}{H(c_{1,s}) - H(c_{2,s})} + \frac{c'(t_s - c_{2,s})}{n_s h(c_{2,s})} \right) V_s \left(1 - \frac{W_s}{\sum_{i \in \mathcal{F}} W_i} \right) \\ &- 1 - \left(\frac{1}{c'(t_s - c_{2,s})} \right) \left(\frac{q_f(c_{1,s}, c_{2,s}) - c_{2,s}}{H(c_{1,s}) - H(c_{2,s})} \right). \end{aligned}$$

(e)

$$\begin{aligned} \frac{d(q(c_{1,s}) - q_f(c_{1,s}, c_{2,s}))}{d\kappa_s} &= - \left(\frac{(q_g - c_{1,s})}{\kappa_g^s} - \frac{(c_{1,s} - c_{2,s})}{\kappa_s} \right) V_s \left(1 - \frac{W_s}{\sum_{i \in \mathcal{F}} W_i} \right) \\ &+ \frac{1}{n_s} \left(\frac{q_f(c_{1,s}, c_{2,s}) - c_{2,s}}{H(c_{1,s}) - H(c_{2,s})} \right). \end{aligned}$$

Proof:

From Theorem 3.1, the equilibrium of the game has to satisfy equations (3.7) to (3.10). When the system is derived in κ_s the following derivatives are obtained for any field where $f \neq s$: From equation (3.7),

$$\frac{d\kappa_g^f}{d\kappa_s} = -n_f h(c_{1,f}) \frac{dc_{1,f}}{d\kappa_s}$$

From equation (3.8),

$$\frac{d\kappa_g^f}{d\kappa_s} = -n_f h(c_{2,f}) \frac{dc_{2,f}}{d\kappa_s}$$

From equation (3.9),

$$-c'(t_f - c_{2,f}) \frac{dt_f}{d\kappa_s} = \left(\frac{c_{1,f} - c_{2,f}}{H(c_{1,f}) - H(c_{2,f})} + \frac{c'(t_f - c_{2,f})}{n_f h(c_{2,f})} \right) \frac{d\kappa_g^f}{d\kappa_s}$$

From equation (3.10),

$$-c'(t_g - c_{1,f}) \frac{dt_g}{d\kappa_s} = A_f \frac{d\kappa_g^f}{d\kappa_s}$$

Moreover, for field s , the corresponding equations are:

From equation (3.7),

$$\frac{d\kappa_g^s}{d\kappa_s} = -n_s h(c_{1,s}) \frac{dc_{1,s}}{d\kappa_s}$$

From equation (3.8),

$$\frac{d\kappa_g^s}{d\kappa_s} = -n_s h(c_{2,s}) \frac{dc_{2,s}}{d\kappa_s} - 1$$

From equation (3.9),

$$\begin{aligned} & -c'(t_s - c_{2,s}) - \frac{q_f(c_{1,s}, c_{2,s}) - c_{2,s}}{H(c_{1,s}) - H(c_{2,s})} - c'(t_s - c_{2,s}) \frac{dt_s}{d\kappa_s} = \\ & \left(\frac{c_{1,s} - c_{2,s}}{H(c_{1,s}) - H(c_{2,s})} + \frac{c'(t_s - c_{2,s})}{n_s h(c_{2,s})} \right) \frac{d\kappa_g^s}{d\kappa_s} \end{aligned}$$

From equation (3.10),

$$-c'(t_g - c_{1,s}) \frac{dt_g}{d\kappa_s} + V_s A_s = A_s \frac{d\kappa_g^s}{d\kappa_s}$$

Finally, $\sum_{f \in \mathcal{F}} \kappa_g^f = \kappa_g$. Then, $\sum_{f \in \mathcal{F}} \frac{d\kappa_g^f}{d\kappa_g} = 0$. By doing algebraic substitutions using the previous equations,

$$\frac{dt_g}{d\kappa_g} = \frac{V_s}{\sum_{f \in \mathcal{F}} W_f}$$

Then, for $f \neq s$,

$$\frac{dt_f}{d\kappa_s} = \frac{1}{c'(t_f - c_{2,f})} \left(\frac{c_{1,f} - c_{2,f}}{H(c_{1,f}) - H(c_{2,f})} + \frac{c'(t_f - c_{2,f})}{n_f h(c_{2,f})} \right) \frac{V_s W_f}{\sum_{i \in \mathcal{F}} W_i}$$

and

$$\frac{d(q(c_{1,f}) - q_f(c_{1,f}, c_{2,f}))}{d\kappa_s} = \left(\frac{(q_g - c_{1,f})}{\kappa_g^f} - \frac{(c_{1,f} - c_{2,f})}{\kappa_f} \right) \frac{V_s W_f}{\sum_{i \in \mathcal{F}} W_i}$$

while for the field s ,

$$\begin{aligned} \frac{dt_s}{d\kappa_s} &= \frac{-1}{c'(t_s - c_{2,s})} \left(\frac{c_{1,s} - c_{2,s}}{H(c_{1,s}) - H(c_{2,s})} + \frac{c'(t_s - c_{2,s})}{n_s h(c_{2,s})} \right) V_s \left(1 - \frac{W_s}{\sum_{i \in \mathcal{F}} W_i} \right) \\ &\quad - 1 - \left(\frac{1}{c'(t_s - c_{2,s})} \right) \left(\frac{q_f(c_{1,s}, c_{2,s}) - c_{2,s}}{H(c_{1,s}) - H(c_{2,s})} \right) \end{aligned}$$

and

$$\begin{aligned} \frac{d(q(c_{1,s}) - q_f(c_{1,s}, c_{2,s}))}{d\kappa_s} &= \\ &\quad - \left(\frac{(q_g - c_{1,s})}{\kappa_g^s} - \frac{(c_{1,s} - c_{2,s})}{\kappa_s} \right) V_s \left(1 - \frac{W_s}{\sum_{i \in \mathcal{F}} W_i} \right) + \frac{1}{n_s} \left(\frac{q_f(c_{1,s}, c_{2,s}) - c_{2,s}}{H(c_{1,s}) - H(c_{2,s})} \right) \end{aligned}$$

■

The remainder of this section uses the previous lemmas to prove the comparative static propositions.

Proof: Proposition 3.1

From Lemma 3.6 part (a), a marginal increase in κ_g reduces the threshold of Tier 1 journal. From Lemma 3.6 part (b), the thresholds of Tier 2 journal decreases, so authors to put less effort to publish an idea of a given quality. Finally, from Lemma 3.6 part (c), the influence difference, measured as $q(c_{1,f}) - q_f(c_{1,f}, c_{2,f})$ marginally decreases because $\left(\frac{(q_g - c_{1,f})}{\kappa_g^f} - \frac{(c_{1,f} - c_{2,f})}{\kappa_f} \right) > 0$.

■

Proof: Proposition 3.2

From Lemma 3.9 part (a), a marginal increase in κ_f increases the threshold of Tier 1 journal. From Lemma 3.9 part (d), the thresholds of Tier 2 journal decreases, so authors have to put less effort to publish an idea of a given quality. Finally, from Lemma 3.9 part (e), the influence difference, measured as $q(c_{1,f}) - q_f(c_{1,f}, c_{2,f})$ marginally increases because $\frac{1}{n_s} \left(\frac{q_f(c_{1,s}, c_{2,s}) - c_{2,s}}{H(c_{1,s}) - H(c_{2,s})} \right) \geq 0$ and $1 - \frac{W_s}{\sum_{i \in \mathcal{F}} W_i} = 0$.

■

Proof: Proposition 3.3

This proposition is a direct implication of Proposition 3.5

■

Proof: Proposition 3.4

This proof shows that the increase of the readership of a single field has the effects described in the current proposition. After this proof is presented, it is trivial to generalise it to an increase of the readership vector of several fields. Hence, that part is omitted.

From Lemma 3.7 part (a), a marginal increase in a particular α_s increases the threshold of the general interest journal.

For the second part of the proposition,

$$\begin{aligned} \sum_{j \in \mathcal{F}} n_s X_f W_j - n_s X_j W_f = \\ A_f^{-1} n_s \sum_{j \in \mathcal{F}} A_j^{-1} \left(q(c_{1,f}) c'(t_g - c_{1,j}) - \sum_{j \in \mathcal{F}} q(c_{1,j}) c'(t_g - c_{1,f}) \right) \end{aligned}$$

where $q(c_{1,f})$ is increasing in $c_{1,f}$ and $c'(t_g - c_{1,f})$ is weakly decreasing in $c_{1,f}$. From that observation, $\sum_{j \in \mathcal{F}} n_s X_f W_j - n_s X_j W_f \geq 0$ for the field with the highest $c_{1,f}$ (i.e. the highest $q(c_{1,f})$), and $\sum_{j \in \mathcal{F}} n_s X_f W_j - n_s X_j W_f \leq 0$ for the field with the lowest $c_{1,f}$, with equality if all the fields have the same value of $c_{1,f}$. Moreover, due to the ordering of f in terms of the value of $q(c_{1,f})$, there exists f^* such that for every $f \geq f^*$, $\sum_{j \in \mathcal{F}} n_s X_f W_j - n_s X_j W_f \geq 0$, and for every $f \leq f^*$, $\sum_{j \in \mathcal{F}} n_s X_f W_j - n_s X_j W_f \leq 0$.

Based Lemma 3.7 part (b), the previous result implies that $f \geq f^* \Leftrightarrow \frac{dt_f}{d\alpha_s} < 0$ and based on Lemma 3.7 part (c) and Theorem 3.1, $f \geq f^* \Leftrightarrow \frac{d(q(c_{1,f}) - q_f(c_{1,f}, c_{2,f}))}{d\alpha_s} < 0$.

■

Proof: Proposition 3.5

As in the previous cases, this proof relies on Lemma 3.8. But before its application, it is important to prove that, when α is large enough, then $\alpha_f X_f + B_f \geq 0$. For this purpose,

$$\begin{aligned} A_f(\alpha_f X_f + B_f) = & \alpha_f q(c_{1,f}) + \frac{n_g}{n_f} (q(c_{1,f}) - c_{1,f}) + c'(t_g - c_{1,f}) \frac{1 - H(c_{1,s})}{h(c_{1,f})n_f} \\ & - (1 - \frac{c'(\max\{t_f - c_{1,f}, 0\})}{c'(t_f - c_{2,s})}) (2q_f(c_{1,f}, c_{2,f}) \\ & + \frac{c_{1,f}(1 - H(c_{1,f})) - c_{2,f}(1 - H(c_{2,f}))}{H(c_{1,f}) - H(c_{2,f})}) \\ & + c'(\max\{t_f - c_{1,f}, 0\}) \frac{1 - H(c_{2,f})}{h(c_{2,s})n_f} - \frac{(1 - H(c_{1,f}))}{h(c_{1,f})n_f} \end{aligned}$$

Therefore,

$$\begin{aligned} A_f(\alpha_f X_f + B_f) \geq & \alpha_f q(c_{1,f}) + \frac{n_g}{n_f} (q(c_{1,f}) - c_{1,f}) \\ & - 2q_f(c_{1,f}, c_{2,f}) - \frac{c_{1,f}(1 - H(c_{1,f})) - c_{2,f}(1 - H(c_{2,f}))}{H(c_{1,f}) - H(c_{2,f})} \end{aligned}$$

Moreover, due to the concavity of H ,

$$\mathbb{E}[v | c_{1,f} \geq v \geq c_{2,f}] \leq \frac{c_{1,f} + c_{2,f}}{2}$$

When this inequality is substituted in the previous equation,

$$\begin{aligned} A_f(\alpha_f X_f + B_f) \geq & \alpha_f q(c_{1,f}) + \frac{n_g}{n_f} (q(c_{1,f}) - c_{1,f}) - c_{1,f} + \frac{c_{1,f}(1 - H(c_{1,f})) - c_{2,f}(1 - H(c_{1,f}))}{H(c_{1,f}) - H(c_{2,f})} \end{aligned}$$

where Theorem 3.1 guarantees that

$$\frac{n_g}{n_f} (q(c_{1,f}) - c_{1,f}) - \frac{c_{1,f}(1 - H(c_{1,f})) - c_{2,f}(1 - H(c_{1,f}))}{H(c_{1,f}) - H(c_{2,f})} \geq 0$$

Hence, $\alpha_f q(c_{1,f}) - c_{1,f} \geq 0$ when $\alpha_f \rightarrow 1$. Therefore, there exists $\bar{\alpha}$, such that for $\alpha > \bar{\alpha}$, $\alpha_f X_f + B_f \geq 0$.

Based on the previous results, the current proposition is a particular case of Lemma 3.8, in which $\frac{dn_F}{\beta} = 1$ and $\forall f < F$ it holds that $\frac{dn_F}{\beta} = 0$. Then, part (a) implies that

$$\frac{dt_g}{dn_F} = \frac{B_F + \sum_{f \in \mathcal{F}} \alpha_F X_f}{\sum_{f \in \mathcal{F}} W_f} \geq 0$$

due to the previous results.

Repeating the proof of Proposition 3.4 for every $f \neq F$, it is true that $f \geq f^* \Leftrightarrow \frac{dt_f}{d\alpha_s} < 0$ and $f \geq f^* \Leftrightarrow \frac{d(q(c_{1,f}) - q_f(c_{1,f}, c_{2,f}))}{d\alpha_s} < 0$.

For the case of F , Lemma 3.8 proves that $\frac{d\kappa_g^F}{dn_F} > 0$. Furthermore,

$$\begin{aligned} q(c_{1,f}) - q_f(c_{1,f}, c_{2,f}) - \frac{c_{1,f}(1 - H(c_{2,f})) - c_{2,f}(1 - H(c_{2,f}))}{H(c_{1,f}) - H(c_{2,f})} \\ \leq q(c_{2,f}) - c_{2,f} - \frac{(1 - H(c_{2,f}))}{h(c_{2,f})} \leq 0 \end{aligned}$$

where the first inequality uses the concavity of H and the second uses Lemma 3.1. Therefore, from Lemma 3.8 part (d),

$$\frac{d(q(c_{1,f}) - q_f(c_{1,f}, c_{2,f}))}{dn_F} < 0$$

■

References

- Abramitzky, R. (2008). The limits of equality: Insights from the israeli kibbutz. *The Quarterly Journal of Economics*, 123(3):1111–1159.
- Acemoglu, D., Carvalho, V. M., Ozdaglar, A., and Tahbaz-Salehi, A. (2012). The network origins of aggregate fluctuations. *Econometrica*, 80(5):1977–2016.
- Acemoglu, D. and Zilibotti, F. (2001). Productivity differences. *The Quarterly Journal of Economics*, 116(2):563–606.
- Alesina, A. and La Ferrara, E. (2000). Participation in heterogeneous communities. *The Quarterly Journal of Economics*, 115(3):847–904.
- Ambrus, A., Mobius, M., and Szeidl, A. (2014). Consumption risk-sharing in social networks. *American Economic Review*, 104(1):149–82.
- Angrist, J., Azoulay, P., Ellison, G., Hill, R., and Lu, S. F. (2017). Economic research evolves: Fields and styles. *American Economic Review*, 107(5):293–97.
- Antras, P. and Helpman, E. (2004). Global sourcing. *Journal of Political Economy*, 112(3):552–580.
- Armstrong, M. (2006). Competition in two-sided markets. *The RAND Journal of Economics*, 37(3):668–691.
- Arrow, K. (1951). Proofs of substitution theorem: General case. In Koopmans, T., editor, *Activity Analysis of Production and Allocation*, pages 155–164. Yale University Press.
- Baland, J.-M. and Platteau, J.-P. (1999). The ambiguous impact of inequality on local resource management. *World development*, 27(5):773–788.
- Baland, J.-M. and Platteau, J.-P. (2007). Collective action on the commons: The role of inequality. In Baland, J.-M., Bardhan, P., and Bowles, S., editors, *Inequality, Cooperation, and Environmental Sustainability*, pages 10–35. Princeton University Press.
- Baldwin, R. and Robert-Nicoud, F. (2014). Trade-in-goods and trade-in-tasks: An integrating framework. *Journal of International Economics*, 92(1):51–62.
- Baldwin, R. and Venables, A. J. (2013). Spiders and snakes: offshoring and agglomeration in the global economy. *Journal of International Economics*, 90(2):245–254.
- Banerjee, A. V. and Newman, A. F. (1998). Information, the dual economy, and development. *The Review of Economic Studies*, 65(4):631–653.

- Bardhan, P. (2000). Irrigation and cooperation: An empirical analysis of 48 irrigation communities in south india. *Economic Development and Cultural Change*, 48(4):847–865.
- Bardhan, P. and Dayton-Johnson, J. (2002). Heterogeneity and commons management in large-scale multivariate research. In Ostrom, E. E., Dietz, T. E., Dolšak, N. E., Stern, P. C., Stonich, S. E., and Weber, E. U., editors, *The Drama of the Commons*, pages 87–112. National Academy Press.
- Bass, F. M. (1969). A new product growth for model consumer durables. *Management Science*, 15(5):215–227.
- Bates, R. H. (1990). Capital, kinship, and conflict: The structuring influence of capital in kinship societies. *Canadian Journal of African Studies/La Revue canadienne des études africaines*, 24(2):151–164.
- Bergin, P. R., Feenstra, R. C., and Hanson, G. H. (2009). Offshoring and volatility: evidence from mexico’s maquiladora industry. *American Economic Review*, 99(4):1664–71.
- Bergstrom, C. T. and West, J. D. (2008). Assessing citations with the eigenfactor™ metrics. *Neurology*, 71(23):1850–1851.
- Bergstrom, C. T., West, J. D., and Wiseman, M. A. (2008). The eigenfactor™ metrics. *Journal of Neuroscience*, 28(45):11433–11434.
- Bloch, F., Genicot, G., and Ray, D. (2008). Informal insurance in social networks. *Journal of Economic Theory*, 143(1):36–58.
- Borch, K. (1962). Equilibrium in a reinsurance market. *Econometrica*, pages 424–444.
- Bornmann, L., Butz, A., and Wohlrabe, K. (2017). What are the Top Five Journals in Economics? A New Meta-ranking. MPRA Paper 79176, University Library of Munich, Germany.
- Bowles, S. and Gintis, H. (2002). Social capital and community governance. *The Economic Journal*, 112(483):F419–F436.
- Bradley, S., Hax, A., and Magnanti, T. (1977). *Applied Mathematical Programming*. Addison Wesley.
- Bramoullé, Y. and Kranton, R. (2007). Risk-sharing networks. *Journal of Economic Behavior & Organization*, 64(3-4):275–294.
- Brin, S. and Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, 30(1-7):107–117.
- Card, D. and DellaVigna, S. (2013). Nine facts about top journals in economics. *Journal of Economic Literature*, 51(1):144–61.
- Cardenas, J.-C. (2005). Overcoming asymmetries in the commons with sub-optimal strategies. In Basili, M., Franzini, M., and Vercelli, A., editors, *Environment, Inequality and Collective Action*, pages 155–175. Routledge.

- Clarivate Analytics (2018). Web of science platform: Summary of coverage. Online: <https://clarivate.libguides.com/webofscienceplatform/coverage>. Accessed: 2018-03-01.
- Coate, S. and Ravallion, M. (1993). Reciprocity without commitment: Characterization and performance of informal insurance arrangements. *Journal of Development Economics*, 40(1):1–24.
- Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94:S95–S120.
- Costinot, A., Vogel, J., and Wang, S. (2012). An elementary theory of global supply chains. *The Review of Economic Studies*, 80(1):109–144.
- Dasgupta, P. (1995). *An inquiry into well-being and destitution*. Oxford University Press on Demand.
- Dasgupta, P. (2009). Trust and cooperation among economic agents. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 364(1533):3301–3309.
- Dayton-Johnson, J. and Bardhan, P. (2002). Inequality and conservation on the local commons: a theoretical exercise. *The Economic Journal*, 112(481):577–602.
- Deardorff, A. V. (2005). Ricardian comparative advantage with intermediate inputs. *The North American Journal of Economics and Finance*, 16(1):11–34.
- Di Falco, S. and Bulte, E. (2011). A dark side of social capital? kinship, consumption, and savings. *Journal of Development Studies*, 47(8):1128–1151.
- Elbers, C., Lanjouw, P., Mistiaen, J., Simler, K., and Özler, B. (2004). *On the unequal inequality of poor communities*. The World Bank.
- Ellison, G. (2002a). Evolving standards for academic publishing: A q-r theory. *Journal of Political Economy*, 110(5):994–1034.
- Ellison, G. (2002b). The slowdown of the economics publishing process. *Journal of Political Economy*, 110(5):947–993.
- Engemann, K. M. and Wall, H. J. (2009). A journal ranking for the ambitious economist. *Review of the Federal Reserve Bank of St. Louis*, pages 127–140.
- Ethier, W. (1984). Higher dimensional issues in trade theory. In Jones, R. and Kenen, P., editors, *Handbook of International Economics*, pages 191–184. Elsevier.
- Fafchamps, M. (1992). Solidarity networks in preindustrial societies: Rational peasants with a moral economy. *Economic development and cultural change*, 41(1):147–174.
- Fafchamps, M. and Gubert, F. (2007). The formation of risk sharing networks. *Journal of Development Economics*, 83(2):326–350.
- Fafchamps, M. and Lund, S. (2003). Risk-sharing networks in rural philippines. *Journal of Development Economics*, 71(2):261–287.

- Felipe, J., Kumar, U., Abdon, A., and Bacate, M. (2012). Product complexity and economic development. *Structural Change and Economic Dynamics*, 23(1):36–68.
- Fourcade, M., Ollion, E., and Algan, Y. (2015). The superiority of economists. *Journal of Economic Perspectives*, 29(1):89–114.
- Gagnon, J. and Goyal, S. (2017). Networks, markets, and inequality. *American Economic Review*, 107(1):1–30.
- Garfield, E. (1955). Citation indexes for science: A new dimension in documentation through association of ideas. *Science*, 122(3159):108–111.
- Garfield, E. (1972). Citation analysis as a tool in journal evaluation. *Science*, 178(4060):471–479.
- Genicot, G. and Ray, D. (2003). Group formation in risk-sharing arrangements. *The Review of Economic Studies*, 70(1):87–113.
- Grossman, G. M. and Helpman, E. (2002). Integration versus outsourcing in industry equilibrium. *The Quarterly Journal of Economics*, 117(1):85–120.
- Grossman, G. M. and Rossi-Hansberg, E. (2008). Trading tasks: A simple theory of offshoring. *American Economic Review*, 98(5):1978–97.
- Grossman, G. M. and Rossi-Hansberg, E. (2012). Task trade between similar countries. *Econometrica*, 80(2):593–629.
- Hamermesh, D. S. (2013). Six decades of top economics publishing: Who and how? *Journal of Economic Literature*, 51(1):162–72.
- Jackson, M. O., Rodriguez-Barraquer, T., and Tan, X. (2012). Social capital and social quilts: Network patterns of favor exchange. *American Economic Review*, 102(5):1857–97.
- Jeon, D.-S. and Rochet, J.-C. (2010). The pricing of academic journals: A two-sided market perspective. *American Economic Journal: Microeconomics*, 2(2):222–55.
- Johnston, J., Reeves, A., and Talbot, S. (2014). Has economics become an elite subject for elite uk universities? *Oxford Review of Education*, 40(5):590–609.
- Kalai, E. and Stanford, W. (1988). Finite rationality and interpersonal complexity in repeated games. *Econometrica*, pages 397–410.
- Kalaitzidakis, P., Mamuneas, T. P., and Stengos, T. (2003). Rankings of academic journals and institutions in economics. *Journal of the European Economic Association*, 1(6):1346–1366.
- Kocherlakota, N. R. (1996). Implications of efficient risk sharing without commitment. *The Review of Economic Studies*, 63(4):595–609.
- Kohler, W. (2004). Aspects of international fragmentation. *Review of International Economics*, 12(5):793–816.

- Konkiel, S. (2016). Altimetrics: diversifying the understanding of influential scholarship. *Palgrave Communications*, 2(1):1–7.
- Kremer, M. (1993). The o-ring theory of economic development. *The Quarterly Journal of Economics*, 108(3):551–575.
- Krueger, D. and Perri, F. (2006). Does income inequality lead to consumption inequality? evidence and theory. *The Review of Economic Studies*, 73(1):163–193.
- La Ferrara, E. (2002). Inequality and group participation: theory and evidence from rural tanzania. *Journal of Public Economics*, 85(2):235–273.
- Laband, D. and Piette, M. J. (1994). The relative impacts of economics journals: 1970-1990. *Journal of Economic Literature*, 32(2):640–66.
- Leontief, W. (1953). Domestic production and foreign trade; the american capital position re-examined. *Proceedings of the American Philosophical Society*, 97(4):332–349.
- Liebowitz, S. J. and Palmer, J. P. (1984). Assessing the relative impacts of economics journals. *Journal of Economic Literature*, 22(1):77–88.
- Ligon, E., Thomas, J. P., and Worrall, T. (2002). Informal insurance arrangements with limited commitment: Theory and evidence from village economies. *The Review of Economic Studies*, 69(1):209–244.
- Luke, N. and Munshi, K. (2006). New roles for marriage in urban africa: Kinship networks and the labor market in kenya. *The Review of Economics and Statistics*, 88(2):264–282.
- Marchiori, C. (2014). Inequality and rules in the governance of water resources. *Ecological Economics*, 105:124–129.
- Markusen, J. (2006). Modeling the offshoring of white-collar services: From comparative advantage to the new theories of trade and fdi. In Collins, S. and Brainard, L., editors, *Brookings Trade Forum 2005*, page a. Brookings Institution Press.
- McCabe, M. J. and Snyder, C. M. (2005). Open access and academic journal quality. *American Economic Review*, 95(2):453–459.
- McCabe, M. J. and Snyder, C. M. (2016). Open access as a crude solution to a hold-up problem in the two-sided market for academic journals. Technical report, National Bureau of Economic Research.
- Mukhopadhyay, P. (2008). Heterogeneity, commons, and privatization: Agrarian institutional change in goa. In Ghate, R., Jodha, N., and Mukhopadhyay, P., editors, *Promise, trust and evolution: managing the commons of South Asia*, pages 213–237. Oxford University Press.
- Mulder, M. B., Bowles, S., Hertz, T., Bell, A., Beise, J., Clark, G., Fazzio, I., Gurven, M., Hill, K., Hooper, P. L., et al. (2009). Intergenerational wealth transmission and the dynamics of inequality in small-scale societies. *science*, 326(5953):682–688.

- Munshi, K. and Rosenzweig, M. (2007). Network decay in traditional economies. In Rauch, J. E., editor, *The missing links: Formation and decay of economic networks*. Russell Sage Foundation.
- Munshi, K. and Rosenzweig, M. (2016). Networks and misallocation: Insurance, migration, and the rural-urban wage gap. *American Economic Review*, 106(1):46–98.
- Naidu, S. C. (2009). Heterogeneity and collective management: Evidence from common forests in himachal pradesh, india. *World Development*, 37(3):676–686.
- Olson, M. (1971). *The logic of collective action: public goods and the theory of groups* (revised edition).
- Ostrom, E. (2000). Collective action and the evolution of social norms. *Journal of Economic Perspectives*, 14(3):137–158.
- Ostrovsky, M. (2008). Stability in supply chain networks. *American Economic Review*, 98(3):897–923.
- Palacios-Huerta, I. and Volij, O. (2004). The measurement of intellectual influence. *Econometrica*, 72(3):963–977.
- Pinski, G. and Narin, F. (1976). Citation influence for journal aggregates of scientific publications: Theory, with application to the literature of physics. *Information Processing & Management*, 12(5):297–312.
- Posner, R. (2000). An economic analysis of the use of citations in the law. *American Law and Economics Review*, 2(2):381–406.
- Price, D. J. D. S. (1965). Networks of scientific papers. *Science*, pages 510–515.
- Rochet, J.-C. and Tirole, J. (2003). Platform competition in two-sided markets. *Journal of the European Economic Association*, 1(4):990–1029.
- Ruttan, L. M. (2008). Economic heterogeneity and the commons: Effects on collective action and collective goods provisioning. *World Development*, 36(5):969–985.
- Samuelson, P. A. (1948). International trade and the equalisation of factor prices. *The Economic Journal*, 58(230):163–184.
- Samuelson, P. A. (2001). A ricardo-sraffa paradigm comparing gains from trade in inputs and finished goods. *Journal of Economic Literature*, 39(4):1204–1214.
- Scott, J. C. (1977). *The moral economy of the peasant: Rebellion and subsistence in Southeast Asia*. Yale University Press.
- Sethi, R. and Somanathan, E. (1996). The evolution of social norms in common property resource use. *American Economic Review*, pages 766–788.
- Shiozawa, Y. (2007). A new construction of ricardian trade theory—a many-country, many-commodity case with intermediate goods and choice of production techniques—. *Evolutionary and Institutional Economics Review*, 3(2):141–187.

- Siegfried, J. J. (2017). Trends in undergraduate economics degrees, 1991–2016. *The Journal of Economic Education*, 48(3):238–242.
- Thomas, J. and Worrall, T. (1988). Self-enforcing wage contracts. *The Review of Economic Studies*, 55(4):541–554.
- Thomas, J. and Worrall, T. (1990). Income fluctuations and asymmetric information. *Journal of Economic Theory*, 51:367–390.
- Timmer, M. P., Erumban, A. A., Los, B., Stehrer, R., and de Vries, G. J. (2014). Slicing up global value chains. *Journal of Economic Perspectives*, 28(2):99–118.
- Townsend, R. M. (1994). Risk and insurance in village india. *Econometrica*, pages 539–591.
- Trefler, D. and Zhu, S. C. (2010). The structure of factor content predictions. *Journal of International Economics*, 82(2):195–207.
- Udry, C. (1994). Risk and insurance in a rural credit market: An empirical investigation in northern nigeria. *The Review of Economic Studies*, 61(3):495–526.
- Wade, R. (1987). The management of common property resources: Finding a cooperative solution. *The World Bank Research Observer*, 2(2):219–234.