Essays on Commitment and Inefficiency in Political Economy

STOCKHOLM SCHOOL OF ECONOMICS
HANDELSHÖGSKOLAN I STOCKHOLM

EFI, The Economic Research Institute

EFI Mission
EFI, the Economic Research Institute at the Stockholm School of Economics, is a scientific institution that works independently of economic, political and sectional interests. It conducts theoretical and empirical research in the management and economic sciences, including selected related disciplines. The Institute encourages and assists in the publication and distribution of its research findings and is also involved in the doctoral education at the Stockholm School of Economics. At EFI, the researchers select their projects based on the need for theoretical or practical development of a research domain, on their methodological interests, and on the generality of a problem.

Research Organization
The research activities at the Institute are organized into 21 Research Centres.
Centre Directors are professors at the Stockholm School of Economics.

EFI Research Centre:	***Centre Director:***
Management and Organisation (A)	**Sven-Erik Sjöstrand**
Centre for Entrepreneurship and Business Creation (E)	Carin Holmquist
Public Management (F)	Nils Brunsson
Information Management (I)	Mats Lundeberg
Centre for People and Organization (PMO)	Andreas Werr (acting)
Centre for Innovation and Operations Management (T)	Christer Karlsson
Economic Psychology (P)	Guje Sevón
Centre for Consumer Marketing (CCM)	Magnus Söderlund
Centre for Information and Communication Research (CIC)	Per Andersson (acting)
Marketing, Distribution and Industrial Dynamics (D)	Björn Axelsson
Centre for Strategy and Competitiveness (CSC)	Örjan Sölvell
Centre for Business and Economic History (BEH)	Håkan Lindgren
Accounting and Managerial Finance (B)	Johnny Lind
Centre for Financial Analysis and Managerial Economics in Accounting (BFAC)	Kenth Skogsvik
Finance (FI)	Clas Bergström
Centre for Health Economics (CHE)	Bengt Jönsson
International Economics and Geography (IEG)	Mats Lundahl
Economics (S)	Paul Segerstrom
Economic Statistics (ES)	Anders Westlund
Law (RV)	**Johnny Herre**
Centre for Tax Law (SR)	Bertil Wiman

Chair of the Board: Professor Carin Holmquist
Director: Associate Professor Filip Wijkström

Address
EFI, Box 6501, SE-113 83 Stockholm, Sweden • Website: www.hhs.se/efi/
Telephone: +46(0)8-736 90 00 • Fax: +46(0)8-31 62 70 • E-mail efi@hhs.se

Essays on Commitment and Inefficiency in Political Economy

Elena Paltseva

STOCKHOLM SCHOOL
OF ECONOMICS
HANDELSHÖGSKOLAN I STOCKHOLM

EFI, The Economic Research Institute

Dissertation for the Degree of Doctor of Philosophy, Ph.D. in Economics
Stockholm School of Economics 2006

ISBN 91-7258-712-1

Keywords:
Institutions, Growth, Autocracy, Commitment, Resource Curse, Public Goods, Coase Theorem, Clubs, Governance, Unanimity, Enhanced Cooperation, Lobbying, Strategic Trade Policy, Substitutability, Free-riding.

Printed by:
Elanders Gotab, Stockholm 2006

Distributed by:
EFI, The Economic Research Institute
Stockholm School of Economics
P O Box 6501, SE 113 83 Stockholm, Sweden
www.hhs.se/efi

To Valentina and Alexander

Acknowledgments

In 1991, when I entered the Mechanics and Mathematics Department of the Moscow State University, I could not imagine that I would ever turn to economics. In 1996, when I started my Masters in Economics at the New Economic School, I was doubtful if I really wanted to make economics my profession. In 1998, beginning to work for the Russian European Center for Economic Policy, I was sure I would never go abroad to study.
Life has proved to be diverse beyond my expectations. Now, I am close to the completion of my PhD studies at the Stockholm School of Economics. Looking back, I realize that the day I decided to enter the PhD program, I was not fully aware of what I was going to face. In fact, during the years, it was not always easy to be enthusiastic and optimistic about my choice. But, ex post, I can say that I am happy to have taken that decision. And now I would like to thank the people who helped me during this journey.

I am greatly indebted to my supervisor, Tore Ellingsen, for his guidance and support during the process of writing this thesis. His incredible ability to shape and structure the argument and look at problems from a different perspective proved to be of great help. His enthusiastic approach to economic research was very motivating and sometimes crucial to overcome my skepticism. I was also fortunate to have Tore as the coauthor of one of the papers included in this thesis. I know I am not the first of his students to say so, but it does not make it less true: this thesis would never have been written without Tore.

I am very grateful to Jörgen Weibull for his generous advice and encouragement. Jörgen has always been ready to discuss ideas and on-going projects and suggest new ways of approaching them. In addition, I hope I learned a bit of his amazing teaching skills while being his teaching assistant.

Special thanks goes to Torsten Persson for invaluable suggestions and insightful comments. I greatly benefited from our discussions on several papers of this thesis.

I would like to express my sincere gratitude to my coauthors Erik Berglöf, Guido Friebel and Mike Burkart. I learned a lot from working with them on one of the papers in this thesis. In addition, I am grateful to Erik for encouraging me to join the PhD program and for providing advice and help both on academic and non-academic issues. I am thankful to Guido for many fruitful discussions. I am indebted to Mike for teaching me his approach to research, in particular the way to find the right balance between rigor and flexibility, for his advice, and last, but not least, for the joy of being exposed to his incredible sense of humor.

During my stay at Princeton University, I highly appreciated rewarding discussions with Gene Grossman and Giovanni Maggi. I am also grateful to Victor Polterovich for his support and advice at the early stages of this thesis.

I received help and comments from many colleagues at the Economics Department of the Stockholm School of Economics and, more recently, at the Institute for International Economic Studies. I would like to thank them for a very pleasant working and social environment. I also owe thanks to the administrative staff – Annika Andreasson, Carin Blanksvärd, Pirjo Furtenbach, Ritva Kiviharju, Christina Lönnblad, Gun Malmquist, Anneli Sandbladh and Lilian Öberg – who were always very helpful. I also want to thank Christina Lönnblad for excellent editorial assistance. Financial support from Jan Wallander's and Tom Hedelius' Research Foundation is gratefully acknowledged.

During my graduate student years, I was lucky to meet many people who made my time in the program much more interesting and fun. I am grateful to my (former and current) fellow PhD students Rudolfs Bems, Milo Bianchi, Alessandra Bonfiglioli, Marieke Bos, Emilio Calvano, Anna Dreber, Max Elger, Anders Fredriksson, Andres Gonzalez, Ola Granström, Erik Grönqvist, Peter Gustafsson, Erik Höglin, Emanuel Kohlscheen, Sandra Lerda, Therese Lindahl, Henrik Lundvall, Andreas Madestam, Agatha Murgoci, Anete Pajuste, Mauricio Prado, Carlos Razo, Witness Simbanegavi, Per Sonnerby, Björn Tyrefors, Nina Waldenström, Fredrik Wilander, and many others. It was a great pleasure to have the company of Andryi Bodnaruk, Sasha Matros, Andrei Simonov, Olga Lazareva, Oleg Zamulin, and the entire Russian-speaking crowd.

A very special thanks goes to my dear girlfriends. Daria Finocchiaro, Raquel Gaspar and Virginia Queijo von Heideken, you made my life so much more enjoyable, and you have always been there for me. Anna Breman and Chloé Le Coq, what started as simple office-sharing, turned into a great friendship. Thank you for our discussions on economics and life, for all your help and for putting up with my everyday mood swings. Caterina Mendicino and Irina Slinko, you have been with me through bad and good moments, you have shared my excitement and my stress, you have helped me with an enormous amount of things, and your care made an incredible difference! I am very grateful for your invaluable support and friendship.

Thanks to my Swedish friends outside academia for helping me improve the balance between work and leisure. I am also thankful to my friends outside Sweden who have provided a lot of encouragement despite the distance. It is always nice to have you behind me!

Last, but not least, I want to express my immense gratitude and love to my parents, Valentina and Alexander. Thank you for always believing in me and encouraging me, for all that you have given and still give to me, for your understanding and love!

Contents

Introduction and summary

Economically inefficient outcomes are widespread. Societies develop institutions that enrich a narrow group of actors at the cost of aggregate welfare. Governments pursue policies that hamper growth and development. Parties engage in transactions that result in suboptimal outcomes for all participants. Economists suggest a number of explanations for the origin of these inefficiencies, such as informational asymmetries, uncertainty, externalities etc. In particular, a wide range of studies explains the inefficiencies by the inability of actors to make a binding commitment to certain actions and policies. The discrepancy between ex-ante and ex-post incentives of agents may undermine efficient outcomes.

This dissertation is devoted to the analysis of various aspects of inefficiency in the political economy. It consists of four self-containing theoretical essays. The first two chapters deal with the interplay between inefficiency and commitment. Chapter 1 studies the problem of commitment in autocratic regimes and its implications for growth. Chapter 2 argues that the absence of commitment undermines the validity of the Coase theorem. The next two chapters address alternative sources of inefficiency, abstracting from commitment-related problems. Chapter 3 discusses inefficiencies arising in organizations whose members possess veto power and suggests a way of mitigating the problem. Finally, Chapter 4 analyzes the impact of demand linkages on the efficiency of lobbying for trade policy.

Chapter 1. "Autocracy, Devolution and Growth"

Some autocracies have sustained high economic growth for many decades; others have stagnated completely at low levels of production. Paradoxically, the stagnating autocracies appear to possess more natural resources and to be more resistant to political change than the growing autocracies.

The paper argues that the scope for capital accumulation and growth in an autocracy is largely determined by the autocrat's incentive to cling to power. In the model, an infinitely lived autocrat determines taxes and political regime. As capital accumulates, the autocrat faces an increasing temptation to expropriate the capitalists. Since expropriation eliminates growth, the autocrat may voluntarily refrain from expropriating if future growth is sufficiently large; otherwise, the temptation to expropriate can only be resisted through a credible commitment, that is, by devolving some political power. For autocrats with large benefits of political control, for example due to valuable natural resources, devolution of power may always be unattractive. As a result, capitalists realize that they will eventually be expropriated, and capital accumulation therefore never starts. On the other hand, autocrats with small benefits of control will eventually devolve power, since this commitment is necessary to sustain growth. Therefore, capitalists are willing to start accumulating despite the autocratic regime.

Natural resource availability might be the most obvious determinant of an autocrat's incentive to cling to power. If a country is sufficiently rich in natural resources, an autocratic ruler will always resist political change, because of the lost stream of revenues. In a resource poor country, on the other hand, the autocrat may prefer sustaining capital accumulation to preserving resource rents. This simple model can therefore account both for the observed negative relationship between natural resource rents and economic growth and the great heterogeneity in economic outcomes under autocracy. It also suggests why resistance to political change may be greater in autocracies that never experienced capital accumulation than in the economically more successful autocracies.

Chapter 2. "The Coase Theorem Is False" (with Tore Ellingsen)

This paper provides simple and robust counterexamples to the Coase Theorem. It shows that if the voluntariness requirement is taken seriously, the absence of a commitment mechanism may destroy the efficiency. More precisely, we show that equilibrium investments in club goods can be suboptimally small, despite the presence of well-defined and perfectly protected property rights and the absence of transaction costs and informational asymmetries. The reason is that, in equilibrium, a club of owners will typically not exercise their right to exclude outsiders, instead preferring to exercise their right to sell access. As long as the club of owners does not have all the bargaining power in such ex post access negotiations, strategic non-membership provides a valuable free-riding opportunity. Even in the case of only two persons, coercion may be needed in order to guarantee an efficient outcome.

Chapter 3. "Club-in-the-Club: Reform under Unanimity" (with Erik Berglöf, Mike Burkart and Guido Friebel)

In many organizations, decisions are taken by unanimity, giving each member veto power. Unanimity grants each member a veto right, thereby protecting her against the "tyranny of the majority". But the flipside of unanimity is slow and inflexible decision-making and underprovision of the common good: Under unanimity, the least efficient members impose their preferred effort choice on the entire organization.

This paper argues that organizations operating under the unanimity rule can provide more public goods than preferred by their least committed member. In the presence of externalities and an incomplete charter, the threat of forming an "inner organization" can undermine the veto power of the less efficient members and coerce them to exert more effort. We identify the conditions under which the threat of forming an inner organization is never executed and under which inner organizations are equilibrium outcomes. Finally, the members' welfare can both be higher and lower under the simple majority rule than under unanimity. This explains observed diversity of decision rules across organizations as well as across subject matters within a given organization. In addition, it provides a rationale for why members of an organization may voluntarily renounce their veto right and subject themselves to the will of future majorities.

Chapter 4. "Protection for Sale to Oligopolists"

The paper modifies Grossman and Helpman's canonical "Protection for Sale" model by allowing demand linkages and oligopolistic competition. It shows that increased substitutability between products weakens the interest group's incentives to lobby. If demands are interdependent, an increase in the price of a good causes demand to shift towards its substitutes. The interest group takes this shift into account when lobbying the government; hence, the lobbying strategy of an organized industry becomes less aggressive. For the case of one organized and one unorganized industry, we obtain a particularly simple result: as product substitutability increases, the protection of the organized industry's product falls, whereas the protection of the unorganized sector's product increases. Additionally, with endogenous lobby formation, fewer industries get organized and lobbying becomes less intense in the presence of substitutes.

Empirical studies of the "Protection for Sale" model have suggested that the U.S. government's trade policy decisions are overwhelmingly determined by a concern for welfare maximization; the alternative interpretation of the paper is that the original model overstates the lobby groups' desire for protection.

CHAPTER 1

Autocracy, Devolution and Growth*

Abstract

Some autocracies have sustained high economic growth for many decades; others have stagnated completely at low levels of production. Paradoxically, the stagnating autocracies appear to possess more natural resources and be more resistant to political change than the growing autocracies. The paper proposes a new explanation for these observations. In the model, an infinitely lived autocrat determines taxes and political regime. As capital accumulates, the autocrat faces an increasing temptation to expropriate the capitalists. Since expropriation eliminates growth, the autocrat may voluntarily refrain from expropriating if future growth is sufficiently large; otherwise, the temptation to expropriate can only be resisted through a credible commitment, that is, by devolving some political power. For autocrats with large rents, for example due to valuable natural resources, devolution of power may always be unattractive. As a result, capitalists realize that they will eventually be expropriated, and capital accumulation therefore never starts. On the other hand, autocrats with small resource rents will eventually devolve power, since this commitment is necessary to sustain growth. Therefore, capitalists are willing to start accumulating despite the autocratic regime. In other words, autocracies are vulnerable to the resource curse.

*I am indebted to Tore Ellingsen for his advice and encouragement. I also thank Philippe Aghion, Erik Berglöf, Mike Burkart, Nicola Gennaioli, Sergei Guriev, John Hassler, Ethan Kaplan, Assar Lindbeck, Elena Panova, Torsten Persson, Maria Petrova, Jesper Roine, Konstantin Sonin, Jörgen Weibull and Fabrizio Zilibotti for valuable comments, as well as seminar participants at CEFIR, Stockholm School of Economics and IIES. Jan Wallander's and Tom Hedelius Research Foundation is gratefully acknowledged for financial support. All remaining errors are my own.

1 Introduction

Nowadays, most economists agree that economic and political changes are intertwined.[1] For example, it is commonly argued both that protection of property rights from governmental abuse creates economic growth, and that economic growth gives rise to political freedom, constraining the discretion of the executive.[2] Still, the relationship between limited government and economic outcomes is not very well understood. Notably, autocracies are found to be both the best and worst performers in terms of growth rates.[3] Moreover, the most economically successful autocracies tend to eventually be replaced by more democratic institutions, whereas poorly performing autocracies often prevail for a very long time.

In this paper, I argue that the scope for capital accumulation and growth in an autocracy is largely determined by the autocrat's incentive to cling to power. A main result of the model is that there can be private capital accumulation only if the autocrat's benefits from political control are not too large. The reason is that eventually, as capital accumulates and growth rates decline, an unfettered autocrat's temptation to expropriate capital becomes irresistible. Therefore, capital accumulation proceeds beyond this point only after the autocrat has relinquished some power. While autocrats with small benefits of political control are willing to relinquish power once it becomes necessary in order to sustain future capital accumulation, autocrats with large benefits of control are not. Looking ahead, capitalists realizing that they will eventually be expropriated never start accumulating.

Natural resource availability might be the most obvious determinant of an autocrat's incentive to cling to power. If a country is sufficiently rich in natural resources, an autocratic ruler will always resist political change, because of the lost stream of revenues. In a resource poor country, on the other hand, the autocrat may prefer sustaining capital accumulation to preserving resource rents. This simple model can therefore account both for the observed negative relationship between natural resource rents and economic growth and the great heterogeneity in economic outcomes under autocracy. It also suggests why resistance to political change may be greater in au-

[1] E.g. "At some level, these rich dynamics of economic and political change have to be connected", (Persson and Tabellini (2006b)), or "...wealth, its distribution, and the institutions that allocate factors and distribute incomes are mutually interdependent and evolve together." (Przeworski (2004)).

[2] "The expansion of economic freedom will bring in train greater political freedoms."(Friedman (2002)).

[3] E.g. Almeida and Ferreira (2002).

tocracies that never experienced capital accumulation than in the economically more successful autocracies.

The relationship between economic growth and the autocratic rule with its predatory potential has been (and still is) a topic of an extensive debate. While it is well recognized that expropriation or other property rights violations by political rulers are detrimental for investment and growth, there is no consensus on whether the respective protection of investors has to be institutionalized. According to one view, the abuse of political power can only be effectively prevented through explicit legal limitations on rulers' authority. Institutional checks and balances work as a commitment device against expropriation, which encourages investment and gets growth started. In the modern literature, this view is associated with works by North and Thomas (1970), North (1971), North and Weingast (1989) and has inspired a considerable amount of supportive empirical work.[4]

An alternative view is that sustained growth can be in the interest of an unconstrained dictator, who therefore rationally refrains from expropriation. As emphasized by Olson (1993), this argument does not require any benevolence on the part of the autocratic ruler; a "stationary bandit" will promote growth for selfish reasons. This view is also buttressed by the evidence – indeed, the most impressive growth episodes were almost always observed in autocratic regimes.[5] Furthermore, Glaeser et al. (2004) stress that "initial levels of constraints on the executive do not predict subsequent economic growth" and "growth ... may be feasible without immediate institutional improvement".

Nonetheless, the "stationary bandit" approach does not explain why most of the well-performing autocracies eventually limit the dictator's power, e.g. through partial or full democratization. The classical argument for devolution of power – the famous Lipset (1960) hypothesis – suggests that prosperity improves political institutions through better education and the increased importance of the middle class.

In this paper, we complement Lipset's demand-driven view of democratization and instead look at the supply side of political change. More precisely, we show that a self-interested dictator may choose to relinquish power in a phase of declining growth. We thus suggest a compromise between the view of North and Weingast (1989) and the "stationary bandits" argument of Olson (1993). Olson's argument implies that the

[4]E.g. Knack and Keefer (1995), Goldsmith (1995), Hall and Jones (1999), Henisz (2000), Keefer (2004).

[5]E.g. Almeida and Ferreira (2002).

interests of the dictator and the investors are aligned and no commitment device is needed at all. On the contrary, North and Weingast argue that the incentives of the ruler are always opposed to those of private investors, so in order to initiate investment and growth, the autocrat has to commit beforehand. In our model, the conflict of interests between the autocrat and the investors intensifies over time, thereby causing a potential delay in the devolution of power.

Specifically, we assume that a country is facing an exogenous opportunity for multi-period investment that is only exploitable by the private sector. This investment is characterized by decreasing returns to scale. An autocratic ruler can tax the private sector or expropriate it, but is unable to invest and thus support growth after expropriation. The ruler can also choose to relinquish some of her power, which reduces her payoff and deprives her of the right to expropriate. While being in power, the ruler enjoys private benefits of control, (part of) which are no longer available to the ruler after devolution. That is, the ruler cannot "fine tune" the devolution decision so as to keep every possible fraction of current economic privileges. The private sector has an option to avoid taxation and expropriation by diverting resources to a less efficient alternative use. Due to decreasing returns to scale, the growth rate in this economy declines over time. As a result, the degree of alignment of government's and private agents' interests also decreases: At the early stages of growth, when the growth rate is high, an autocrat has no incentive to seize assets, because delayed expropriation significantly increases the value at stake. As growth slows down, immediate expropriation ("getting the entire pie") becomes increasingly attractive, as compared to the option of postponing it to increase the size of the pie. The private sector recognizes these incentives of the ruler, and diverts resources once the ruler is tempted to expropriate. If the ruler's private benefit of control is relatively low, she commits to non-expropriation through devolution of power and thereby keeps the private sector investing. If the private benefits of control are high, the ruler does not want to ever lose them through devolution. Realizing that capital will eventually be expropriated, the private sector thus never starts to invest. Therefore, the model generates a variety of development trajectories: The country can face an early limitation of the ruler's power and grow in a non-autocratic regime. Alternatively, the country can grow under maintained autocracy and experience delayed devolution of power. Finally, the country can stagnate under an autocratic rule with neither devolution nor growth ever taking place.

The model predicts that these stagnating economies are characterized by high private benefits of control, e.g. are abundant with appropriable natural resources. This finding parallels the "resource curse" literature, arguing that the natural resource wealth can be detrimental to countries' economic development. The early work on the "resource curse" attributes the underperformance of resource-rich countries to economic factors such as "Dutch disease" or deteriorating terms of trade for the primary commodities. Empirically, this effect is documented e.g. by Sachs and Warner (1995) and Gylfason (2000). More recent literature emphasizes the political and institutional determinants of the resource curse (see e.g. Robinson et al. (2006), Mehlum et al. (2006) and Ross (1999) for a review of both approaches), which again finds empirical support (Mehlum et al. (2006) and Boschini et al. (2003)). Finally, there are studies addressing the reverse effect – the impact of the resource curse on institutions. The effect is found to be negative and is attributed to reduced government accountability, a better ability to repressed opposition and increasing corruption and rent-seeking (see Ross (2001), Sala-i-Martin and Subramanian (2003) and Collier and Hoeffler (2005)). We propose an alternative link, suggesting that the abundant resources undermine the autocrat's incentives to relinquish power and hence, hamper capital accumulation.

The key feature of our model is that devolution of power is a commitment device against expropriation. A related strand of literature takes devolution of power as a commitment to redistribution. In Acemoglu and Robinson (2000, 2001), the rich elite relinquishes the power (by extending franchise) in case it cannot prevent social unrest through a temporary increase in taxes. Franchise extension transfers the taxation decision to the (poor) median voter, thereby working as a commitment device for more redistribution. Gradstein (2004) treats franchise extension as a commitment to private property rights. In his model, agents allocate their time between production and rent-seeking, with wealthier agents having a comparative advantage in the latter. With an income-based franchise, voting participation increases as the economy grows. The median voter becomes relatively more interested in curbing rent-seeking and protecting private property rights. Commitment to protect property rights in the future can thus be achieved by reduction in the franchise threshold. While these approaches provide an alternative explanation for devolution of power, they ignore the predatory nature of autocratic rulers, which is key to our argument.

Bates et al. (2005) do consider predatory rulers. In their model, a government can be benevolent or opportunistic, but its type is unobservable to the citizens ex ante. The opportunistic government may decide to show restrains to pretend that it is benevolent in order to stimulate private investment and increase the appropriable assets in the future. In this model, the economy reaches full development under autocracy in case of a benevolent government and collapses in case of a predatory government. Robinson (2001) associates predatory behavior with inefficiently low public investment and suggests that the development-enhancing policies may facilitate the political power contest. Therefore, the party in power may refrain from promoting development. However, these models do not address the issue of the devolution of power, which is a fundamental aspect of our framework.

Our model is consistent with a range of empirical findings. It predicts that richer countries devolve earlier, as has been found e.g. by Barro (1996). Autocratic economies are predicted to have either relatively high or relatively low growth rates, while the democratic regimes fall into an intermediate range, which is in line with Almeida and Ferreira (2002). Moreover, the model also suggests a new explanation to why many empirical studies (e.g. Barro (1999), Alvarez et al. (2000), Boix and Stokes (2003)) find that growth causes democratization. In our model, the devolution of power is preceded by a more or less extensive period of growth. Thus, if the data were to be generated according to our model, we would find that growth Granger causes democratization. Nevertheless, this conclusion is misleading: in the model, institutions of limited government and growth are determined simultaneously and endogenously. Indeed, in the absence of a possibility to relinquish power, growth would slow down if not completely stop. Similarly, in the absence of growth opportunities, the institutions of limited government would never be introduced. Therefore, by means of the model, we illustrate why the observed time pattern between growth and democratization does not reflect a causal relation. This is in line with the more recent literature (Acemoglu et al.(2004), Przeworski (2004)) that emphasizes the endogenous determination of institutions and growth.

The paper proceeds as follows. Chapter 2 describes the setup of the model and chapter 3 presents the model's solution. The predictions of the model and its comparative statics are discussed in chapter 4. Chapter 5 addresses the assumptions of the model. Chapter 6 discusses relevant case studies. Finally, Chapter 7 concludes and suggests some directions for further research.

2 The Model

An infinite horizon economy ruled by an autocrat is populated by identical atomistic citizens of mass one. As all citizens are identical, we only discuss a two-player game between the Ruler and the representative Citizen. At date $t = 0$, due to an exogenous shock, the Citizen is being exposed to a technology allowing for growth based on capital accumulation. In each period t, output is produced from capital according to the Cobb-Douglas production function

$$y_t = A_t (K_t)^{\alpha},$$

were K_t denotes the capital stock at time t. We assume that capital depreciates completely in each period; that is, the depreciations rate δ is 1.

Each period, the Citizen can stay in the market or leave the market and divert the capital to an alternative activity. If the Citizen stays, the Ruler has the power to tax the Citizen through a consumption tax τ_t. Taxation is costly for the Ruler. We assume the cost of tax collection $\Phi(\tau_t c_t)$ is proportional to the tax base:

$$\Phi(\tau_t c_t) = \phi(\tau_t) c_t,$$

where $\phi'(.) > 0, \phi"(.) > 0$, and $\phi(0) = \phi'(0) = 0$. The Ruler also has the option of expropriating the entire production process (we assume technological indivisibility of the capital stock). In the latter case, the Ruler can still employ the Cobb-Douglas production function but, unlike the Citizen, she is unable to accumulate capital. Thus, due to complete depreciation, she can only use the expropriated capital stock in one period.[6] In case of expropriation, the citizen receives zero payoff from the moment of expropriation and onwards.[7]

If the Citizen leaves, she receives a payoff of L, which is non-taxable and non-expropriable. The decision to leave is once-and-forever: if diverted, capital cannot be returned to the market sector. We assume that

$$L < \sum_{i=0}^{\infty} \beta^i \ln \frac{(1-\alpha\beta) A K_0^{\alpha}}{(1+\tau_A)}, \tag{1}$$

where τ_A solves $1 - \phi'(\tau_i)(1+\tau_i) + \phi(\tau_i) = 0$. As we shall see, (1) implies that the alternative activity is less efficient than the market activity. That is, the Citizen's

[6]The introduction of a cost of expropriation does not affect the qualitative predictions of the model as long as the assumption of indivisibility of capital holds.

[7]The taxation scheme consisting of two instruments (consumption tax and complete expropriation of capital) was chosen to keep the analysis tractable. As will be shown below, in our setting, consumption tax does not influence the capital accumulation path. This, together with the cost of taxation, allows us to avoid the time inconsistency problem, standard in infinite-horizon taxation models. We believe that the argument of the model will not be destroyed by replacing the proposed tax scheme by one distortive (capital) tax and classifying expropriation as unappropriately high level of taxes. However, we are not able to solve for the resulting equilibrium.

payoff of leaving is less than the payoff she would get by staying in the market, were the Ruler able to commit to never expropriate from her.

The Ruler can also choose to relinquish her power. We assume that this decision is equivalent to completely renouncing the right to expropriate and limiting the Ruler's power to set taxes; the Ruler cannot "fine tune" the devolution decision to keep all current economic privileges.[8] More precisely, we assume that after devolution of power, the Ruler cannot set a consumption tax above some upper bound τ_D, where

$$\tau_D < \tau_A. \tag{2}$$

We also assume that the devolution decision is once and forever, i.e., if the Ruler devolves she cannot return, and no other dictatorial ruler can take over.

While being in power, the Ruler receives private benefits of control of b units each period. For example, these benefits can represent natural resource rents. After devolution of power, these benefits are no longer available to the Ruler. Note that we do not argue that the devolution deprives the Ruler of all benefits. Instead, b reflects those benefits of control that are lost upon devolution. We also assume that if the economy reaches full development, the industrial sector is sufficiently more productive than the natural resource sector, so the steady-state post-devolution payoff of the Ruler exceeds the value of the flow of private benefits.

The Citizen's instantaneous utility is logarithmic, so she maximizes the flow of her future utilities

$$V_t = \sum_{j=1}^{\infty} \beta^j v(c_{t+j}) = \sum_{j=1}^{\infty} \beta^j \ln c_{t+j},$$

where c_t is consumption and β is the discount factor. The Ruler's utility is linear and her payoff function is denoted

$$U_t = \sum_{j=1}^{\infty} \beta^j d_{t+j},$$

where d_{t+j} is the payoff received by the Ruler in period $t + j$.

The timing of the game is as follows: The Citizen and the Ruler meet at discrete time periods $t = 0, 1, ..., \infty$. Each period has three stages. If there was no devolution of power in the past, at stage 1 of period t, the Ruler decides whether to devolve (D) or abstain from devolution (ND). At stage 2, the Citizen decides whether to stay in the market sector (S) or leave (L). At stage 3 of period t, in case the Citizen

[8] In Chapter 5, we discuss the implications of relaxing this assumption.

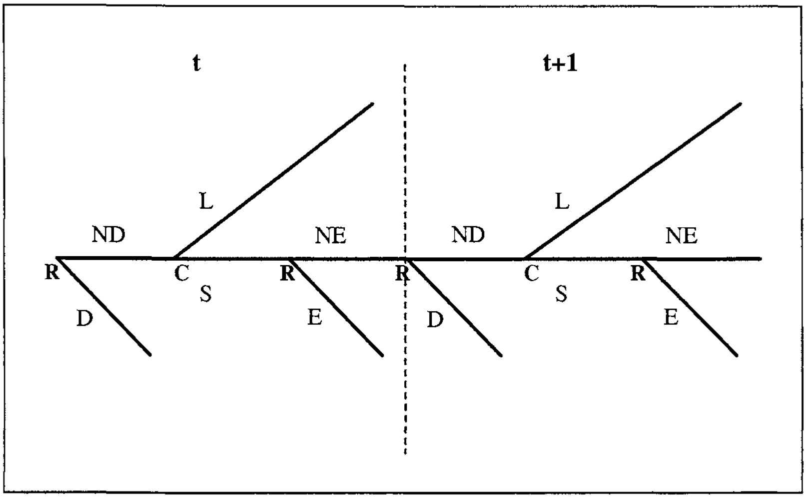


Figure 1: Timing of the game

stays in the market sector, the Ruler chooses whether or not to expropriate (E) the Citizen's capital. Expropriated capital is used for production only in period t and then it depreciates completely, as the Ruler cannot invest. If no expropriation (NE) occurs, production, consumption and investment take place and taxes are paid. Then, the game proceeds to the period $t+1$.

As in any other multi-stage game, the history of the game is a collection of all actions played up to stage t. A (behavior) pure strategy of the ruler/citizen is a function $\sigma(h^t)/\rho(h^t)$ prescribing an action in each of the ruler's/citizen's controlled game tree nodes for a given history of the game h^t. For example, conditional on the history h^t, at the first stage of each period, a pure strategy of the ruler determines a choice between D_t and ND_t. At the second stage, the strategy of citizen determines S_t or L_t taking into account all the past actions including those just played at stage 1 of period t. Finally, at the third stage, the ruler's strategy defines a choice between E_t and NE_t, again, conditional on all actions played up to stage 3 of period t and, in the latter case, τ_t.

3 Analysis of the Game

3.1 The benchmark case: no government intervention

Let us start by describing the evolution of this economy in the absence of any governmental intervention. As staying in the market sector is more productive than leaving it, the efficient outcome is for production to occur in the market sector and for the citizen to maximize her utility

$$\max \sum_{j=1}^{\infty} \beta^j \ln c_{t+j},$$

subject to the dynamic budget constraint

$$K_{t+1} = A\,(K_t)^{\alpha} - c_t.$$

This problem is identical to the standard Ramsey growth model and the solution is[9]

$$\begin{aligned} c_t &= (1-\alpha\beta)\,AK_t^{\alpha}, \\ K_{t+1} &= \alpha\beta AK_t^{\alpha}. \end{aligned} \tag{3}$$

The dynamic equilibrium of this economy is characterized by the capital and the output monotonically converging to the steady state values K^* and y^*, respectively:[10]

$$\begin{aligned} K^* &= (\alpha\beta A)^{\frac{1}{1-\alpha}}, \\ y^* &= A^{\frac{1}{1-\alpha}}\,(\alpha\beta)^{\frac{\alpha}{1-\alpha}}. \end{aligned}$$

Note that along the transition path, both capital and output are approaching the steady state at a decreasing rate. The growth rate of capital

$$\gamma_k(t) \equiv \frac{K_{t+1}-K_t}{K_t} = \alpha\beta AK_t^{\alpha-1} - 1,$$

and the growth rate of output

$$\gamma_y(t) \equiv \frac{y_{t+1}-y_t}{y_t} = [\gamma_k(t)+1]^{\alpha} - 1,$$

are both decreasing in t.

[9] It can be confirmed by the usual guess-and-verify method.

[10] We concentrate on studying the increasing branch of the saddle path.

3.2 Subgame-perfect Nash equilibrium

We turn to the analysis of the game between the Ruler and the Citizen and characterize the pure-strategy subgame-perfect Nash equilibria of this game.

We start by examining the Ruler's taxation incentives.

Lemma 1 *If the Ruler were not able to expropriate and the Citizen were not able to leave the market, the Ruler would choose the same tax rate τ_A each period, such that*

$$1 - \phi'(\tau_A)(1 + \tau_A) + \phi(\tau_A) = 0.$$

Proof. First, note that in our setting, consumption taxation does not influence the capital accumulation path (and thus the output path). Indeed, assume that the citizen is facing a flow of taxes $\{\theta_t\}, t = 1, ..., \infty$. Her problem then becomes

$$\begin{aligned} &\max \sum_{t=0}^{\infty} \beta^t \ln c_t \\ &\text{s.t.} \quad K_{t+1} = AK_t^\alpha - (1 + \theta_t) c_t, \end{aligned}$$

and the solution is

$$\begin{aligned} c_t &= \frac{(1 - \alpha\beta) AK_t^\alpha}{(1 + \theta_t)}, \\ K_{t+1} &= \alpha\beta AK_t^\alpha. \end{aligned}$$

Thus, it is only the consumption profile $\{c_t\}$ that changes as compared to the benchmark case (3), while capital accumulation is unaffected by θ_t.

As a result, at each point in time t, the Ruler solves the problem

$$\begin{aligned} &\max_{\{\tau_i\}} \sum_{i=t}^{\infty} \beta^{i-t} \left(\tau_i c_i - \Phi(\tau_i c_i)\right) = \\ &\max_{\{\tau_i\}} \sum_{i=t}^{\infty} \beta^{i-t} \left(\tau_i - \phi(\tau_i)\right) \left(\frac{(1 - \alpha\beta) AK_i^\alpha}{(1 + \tau_i)}\right). \end{aligned}$$

In the absence of the Citizen's diversion opportunity, the solution of this maximization problem does not depend on past taxes. In other words, the Ruler can decide on each period's tax separately. The respective first-order condition can be rewritten as

$$1 - \phi'(\tau_i)(1 + \tau_i) + \phi(\tau_i) = 0.$$

As $\phi' > 0$, the right-hand side of this equality is decreasing in τ_i. Moreover, it is positive at $\tau_i = 0$ and negative at $\tau_i \to \infty$. Thus, there is a unique $\tau_i \equiv \tau_A$ satisfying first-order conditions and, as second-order conditions are satisfied, it indeed determines the maximum point. ■

Lemma (1) explains why we imposed assumption (1). The assumption ensures that, if the Citizen leaves the market, she receives less than she would under Ruler's full discretion to tax (but not to expropriate). The Ruler in this case does not receive any payoff in addition to the private benefit of control, hence, overall efficiency falls. In turn, our assumption on post-devolutionary taxation (2) simply reflects the idea that devolution imposes some restrictions on the Ruler, so that she is no longer able to choose her preferred tax rate.

We proceed to characterize the equilibria.

Lemma 2 *In any SPNE in all subgames following the devolution of power, the Ruler sets the highest possible tax rate τ_D in each period and the Citizen stays in the market.*

Proof. As shown above, a consumption tax does not influence the capital accumulation path. As the citizens are atomistic, they ignore the effect of their decisions to stay or leave on the Ruler's choice of the tax rates and the timing of devolution, and take them as given. Hence, in each period after the devolution of power, the Ruler and the Citizen "share a pie" of a predetermined size, unless the Citizen diverts the capital to the non-market sector. Let us describe the unique subgame-perfect Nash equilibrium of this game. First, note that given the restrictions on the post-devolution maximum tax rate, the Citizen chooses to stay in the market for any tax schedule. Indeed, after devolution, the Ruler cannot not set the tax rate above τ_D. So, at any point in time t, the Citizen's payoff from staying in the market is at least as high as her payoff when the Ruler taxes at the rate τ_D:

$$V^S(\tau_t, \tau_{t+1}, \tau_{t+2}, ...) \geq \sum_{i=t}^{\infty} \beta^{i-t} \ln \frac{(1-\alpha\beta) AK_i^{\alpha}}{(1+\tau_D)},$$

while her payoff from leaving the market is

$$V^L = L.$$

Note that the Citizen's payoff from leaving the market is below the post-devolutionary payoff. Indeed, as $\tau_A > \tau_D$ and the capital stock increases over time,

$$\sum_{i=0}^{\infty} \beta^{i} \ln \frac{(1-\alpha\beta) AK_i^{\alpha}}{(1+\tau_A)} \leq \sum_{i=0}^{\infty} \beta^{i} \ln \frac{(1-\alpha\beta) AK_i^{\alpha}}{(1+\tau_D)} \leq \sum_{i=t}^{\infty} \beta^{i-t} \ln \frac{(1-\alpha\beta) AK_i^{\alpha}}{(1+\tau_D)}. \quad (4)$$

From our assumption (1) and inequality (4), it follows that once devolution occurs, it is never optimal for the Citizen to leave the market

$$L < \sum_{i=t}^{\infty} \beta^{i-t} \ln \frac{(1-\alpha\beta)\, AK_i^{\alpha}}{(1+\tau_D)}.$$

That is, after devolution of power, the Citizen always prefers staying in the market to leaving for the alternative sector:

$$V^S(\tau_t, \tau_{t+1}, \tau_{t+2}, \ldots) > V^L.$$

Thus, in any SPNE, the Ruler is free to choose the tax schedule that brings her the maximal payoff given the post-devolutionary tax restrictions. As $\tau_D < \tau_A$, the Ruler's payoff increases with the increase in each period's tax. So in every period t, she sets the tax rate τ_t equal to the maximum post-devolutionary tax τ_D. ■

Corollary 1 *Along the non-expropriation (sub)game path, the Citizen never leaves the market. The Ruler chooses τ_A while in power and τ_D after devolution.*

Proof. If there is no expropriation along the game path, the maximum tax the Ruler would set is τ_A. Thus, the Citizen's payoff from staying will never be below

$$\underline{V} = \sum_{i=t}^{\infty} \beta^{i-t} \ln \frac{(1-\alpha\beta)\, AK_i^{\alpha}}{(1+\tau_A)}.$$

But, as capital accumulates, the payoff from leaving is always below the payoff from staying:

$$L < \sum_{i=0}^{\infty} \beta^{i} \ln \frac{(1-\alpha\beta)\, AK_i^{\alpha}}{(1+\tau_A)} \leq \sum_{i=t}^{\infty} \beta^{i-t} \ln \frac{(1-\alpha\beta)\, AK_i^{\alpha}}{(1+\tau_A)} = \underline{V} \leq V^S(\tau_t, \tau_{t+1}, \tau_{t+2}, \ldots).$$

As a result, along the non-expropriation game path, the Ruler can always choose his preferred tax levels, that is τ_A while in power and τ_D after devolution. ■

Denote the share of output the Ruler receives while being in power by

$$\varepsilon_A = \frac{\tau_A - \phi(\tau_A)}{(1+\tau_A)}(1-\alpha\beta),$$

and the one she gets after devolution by

$$\varepsilon_D = \frac{\tau_D - \phi(\tau_D)}{(1+\tau_D)}(1-\alpha\beta).$$

Note that as the tax in autocracy is higher than after devolution, $\tau_A > \tau_D$, the same is true about the Ruler's output shares:

$$\varepsilon_A > \varepsilon_D.$$

We are now ready to study how the Ruler's incentive to expropriate evolves over time.

Lemma 3 *If* $\frac{\varepsilon_A}{(1-\beta)} < 1$, *then in any SPNE, there exists a finite time period* T *such that the Ruler prefers expropriation at stage* 3 *of period* T *over any other continuation strategy.*

Proof. We start with some introductory observations. There are three types of "exits" in this game. First, the Ruler can expropriate the capital from the Citizen. If this occurs, the Citizen is left with zero capital and cannot restart production, so that the continuation of the game is fully predetermined. Second, the Citizen can divert the capital to the shadow sector, which makes the continuation game independent of the players' actions. Finally, the Ruler can devolve power. Since this decision is irreversible, and after devolution the continuation game has a unique subgame-perfect Nash Equilibrium, the value of the remaining game is also clear at the time of devolution.

Observe that after the capital has reached a sufficiently high level, no SPNE can have exits of the first two types along the game path. Indeed, in case of expropriation, the Citizen gets zero, while if she diverts the capital at the preceding stage, she is guaranteed a positive payoff. Similarly, if the Citizen diverts the capital, the Ruler is left with the private benefits of control only. As we assume that the steady-state post-devolution payoff of the Ruler is higher than the value of the flow of private benefits, by continuity, the same holds if devolution occurs sufficiently close to the steady state. Thus, the Ruler prefers devolving power at the preceding stage rather than clinging to power and only receiving the benefits of control.

Keeping this in mind, let us analyze the subgame starting at stage 3 of some period t. At this stage, the Ruler is choosing between immediate expropriation and possible continuation options. For the reasons we have just mentioned, the continuation games to be taken into consideration are those where the game "ends" by the devolution of power at some future period $t+\widetilde{t}$ or continues forever. More precisely, all continuation strategies consistent with SPNE belong to the set:

$$S_t \equiv \left\{ \left\{ s_t^{D_{t+\widetilde{t}}} \right\}_{\widetilde{t}=1,2,..}, s_t^{ND} \right\} \qquad (5)$$

where

$$s_t^{D_{t+\widetilde{t}}} \equiv \left(\tau_t, ND_{t+i}, NE_{t+i}, \tau_{t+i}\ i = 1, ..\widetilde{t} - 1\ ,\ D_{t+\widetilde{t}}\right), \widetilde{t} = 1, 2, ...;$$

are all continuation strategies where taxation is followed by the devolution of power in period $\widetilde{t}$, and

$$s_t^{ND} \equiv (\tau_t, ND_{t+i}, NE_{t+i}, \tau_{t+i},\ i = 1, ..\infty)$$

denotes a strategy, where devolution never occurs and taxation is continued forever.

We start by finding the maximum payoff the Ruler can get if she taxes forever. As the Citizen has the option of leaving the market, the Ruler's payoff cannot exceed the payoff characterized in Lemma 1.[11] That is, the maximum utility of the Ruler in the case of eternal taxation cannot be higher than

$$U\left(s_t^{ND}\right) \leq \sum_{i=t}^{\infty} \beta^{i-t} \varepsilon_A A K_i^{\alpha} + \frac{b}{1-\beta} \equiv \overline{U}_t.$$

Consider now instead the continuation games where the Ruler eventually devolves. Once more, we are interested in the maximum payoff the Ruler can achieve in such a continuation game. Note that this problem differs from the problem faced by the Ruler under eternal taxation by an additional constraint on post-devolutionary taxes; after the devolution, the tax rate is determined by the SPNE and is equal to τ_D. Thus, we can conclude that the maximum payoff achieved by the Ruler in such a continuation game cannot be greater than that under eternal taxation

$$U\left(s_t^{D_{t+\widetilde{t}}}\right) \leq \overline{U}_t \qquad \widetilde{t} = 1, 2, ...$$

Now, we are ready to discuss the choice of the Ruler at stage 3 of some period t. Let us show that for sufficiently large t, the Ruler prefers expropriation over any other continuation strategy. We have just seen that the best continuation strategy for the Ruler brings her no more than $\overline{U}_t$.

The value of expropriating at t is

$$U(E_t) = AK_t^{\alpha} + \sum_{i=t}^{\infty} \beta^{i-t} b = y_t + \frac{b}{1-\beta}.$$

Denote by σ a positive constant

$$\sigma \equiv 1 - \frac{\varepsilon_A}{(1-\beta)} > 0.$$

[11] Including Citizen's participation constraints into the Ruler's optimization problem can only add some additional restrictions and decrease the Ruler's maximum utility.

As long as y_{t+1} is sufficiently close to y^*, so that $\Delta y_{t+1} = y^* - y_{t+1}$ is sufficiently small, we see that the maximum of the Ruler's payoff in any continuation game is always less than her payoff of expropriating at t. Indeed, the difference between the Ruler's expropriation payoff and the maximal Ruler's payoff in any continuation game is

$$\begin{aligned}
& U(E_t) - \overline{U}_t \\
= \; & y_t - \sum_{i=t}^{\infty} \beta^{i-t} \varepsilon_A A K_i^{\alpha} \\
> \; & y_t - \sum_{i=t}^{\infty} \beta^{i-t} \varepsilon_A y^* \\
= \; & y_t - \frac{\varepsilon_A}{1-\beta} y^* \\
= \; & y_t - y^* + \left(1 - \frac{\varepsilon_A}{1-\beta}\right) y^* \\
= \; & \sigma y^* - \Delta y_{t+1} > 0,
\end{aligned} \tag{6}$$

where the inequality in (6) follows from the fact that $y_i = AK_i^{\alpha}$ increases towards y^*, and thus for any i

$$y_{t+1} < y_{t+1+i} < y^*.$$

That is, there exists such a time period T that in any SPNE, the Ruler chooses to expropriate at date T. ■

This Lemma states that if the agents are not very patient, the non-market activity is not very inefficient, and/or the share of capital in production α is high, expropriation always takes place if the economy is sufficiently close to the steady state. Note that if the Ruler's value of expropriating at T is higher than the value of any alternative continuation strategy at time T, the same is true for any $t > T$, as output monotonically converges to the steady-state value y^*.

Corollary 2 *If $\frac{\varepsilon_A}{1-\beta} < 1$, and T is the time period found in Lemma 3, then any SPNE is characterized by the Ruler choosing to expropriate at stage 3 of every period $T, T+1, T+2, \ldots$ and the Citizen leaving the market to avoid expropriation.*

Throughout the paper, we assume that the conditions of Lemma 3 hold; that is, the parameters of the model are such that in the steady state, the Ruler always prefers expropriation over any other strategy. Corollary 2 thus allows us to treat period T as a "final" period and solve the game backwards from there.

Period T As shown in Corollary 2, at stage 3 of period T, the Ruler expropriates. At stage 2 of period T, the Citizen leaves the market sector to get a positive payoff, as compared to the zero payoff in case of staying and being expropriated at the next stage. As a result, at stage 1 of period T, the Ruler has to choose between devolution, in which case she receives a payoff

$$U(D_T) \equiv \tau_D \sum_{i=0}^{\infty} \beta^i c_{T+i} = \varepsilon_D \sum_{i=0}^{\infty} \beta^i y_{T+i},$$

and no devolution, which gives her a flow of future benefits of control

$$U(ND_T) \equiv \sum_{i=0}^{\infty} \beta^i b = \frac{b}{1-\beta},$$

as the Citizen leaves the market at the next stage.

Period T-1 At stage 3 of period $T-1$, the Ruler compares the option of expropriating and getting

$$U(E_{T-1}) = y_{T-1} + \frac{b}{1-\beta}$$

to the option of non-expropriating, taxing and proceeding to stage 1 of period T (where she either devolves or stays and receives a flow of benefits of control only).

If already in period T, the devolution payoff is lower than the flow of the benefits of control

$$U(D_T) < U(ND_T), \tag{7}$$

the Ruler has to decide whether to expropriate at stage 3 of period $T-1$ and receive $U(E_{T-1})$, or proceed to stage 1 of period T to get the flow of the benefits of control and receive

$$U(NE_{T-1}, \tau_A, ND_T) = \varepsilon_A y_{T-1} + b + \beta U(ND_T) = \varepsilon_A y_{T-1} + \frac{b}{1-\beta}.$$

Since

$$U(NE_{T-1}, \tau_A, ND_T) = \varepsilon_A y_{T-1} + \frac{b}{1-\beta} < y_{T-1} + \frac{b}{1-\beta} = U(E_{T-1}),$$

the Ruler prefers to expropriate at $T-1$ as well. Note that the devolution payoff $U(D_t)$ is increasing over time, while the payoff of sustaining the autocratic regime with no market production $U(ND_t)$ is constant. Thus, if inequality (7) holds, that is, the devolution payoff is lower than the accumulated private benefit in period T, the same is true for all periods $t < T$. As a result, the backward induction argument repeats itself until period $t = 0$, where at stage 3 the Ruler again chooses to expropriate, at stage 2 the Citizen leaves the market and at stage 1, the Ruler stays to enjoy the flow of private benefits of control. The Ruler's incentive to keep the private benefits of control is so strong that she never wants to devolve and lose them. Therefore, the Ruler cannot commit not to expropriate and, as a result, no investment ever takes place.

Proposition 1 *An economy with sufficiently high private benefits of control b is stuck in an "underdevelopment trap": the autocratic ruler never relinquishes power and the efficient production technologies never get implemented.*

Assume now that the private benefits of control are not too high, so that devolution is chosen at stage 1 of period T,

$$U(D_T) > U(ND_T).$$

Then, the Ruler compares the payoff of expropriation $U(E_{T-1})$ to the devolution payoff in period T

$$\begin{aligned} U(NE_{T-1}, \tau_A, D_T) &= \varepsilon_A y_{T-1} + b + \beta U(D_T) \\ &= \varepsilon_A y_{T-1} + b + \beta \left[\varepsilon_D \sum_{i=0}^{\infty} \beta^i y_{i+T} \right]. \end{aligned}$$

She prefers expropriation iff

$$U(NE_{T-1}, \tau_A, D_T) < U(E_T). \tag{8}$$

Suppose that condition (8) is met so that the ruler chooses to expropriate. Similarly to above, at stage 2 of period $T-1$, the citizen prefers to leave the market. At stage 1, the ruler once more either devolves to prevent the citizen from leaving, and gets the payoff

$$U(D_{T-1}) = \varepsilon_D \sum_{i=0}^{\infty} \beta^i y_{i+T-1},$$

or does not devolve and receives the benefit of control

$$U(ND_{T-1}) \equiv \sum_{i=0}^{\infty} \beta^i b = \frac{b}{1-\beta}.$$

As we have already discussed what happens in the case when the latter exceeds the former, assume now that

$$U(D_{T-1}) > U(ND_{T-1}).$$

Period T-2 At stage 3, the Ruler again faces the choice of expropriating the Citizen and getting

$$U(E_{T-2}) = y_{T-2} + \frac{b}{1-\beta},$$

vs. proceeding to stage 1 of period $T-1$ where she devolves, which yields

$$\begin{aligned} U(NE_{T-2}, \tau_A, D_{T-1}) &= \varepsilon_A y_{T-2} + b + \beta U(D_{T-1}) \\ &= \varepsilon_A y_{T-2} + b + \varepsilon_D \sum_{i=0}^{\infty} \beta^{i+1} y_{i+T-1}. \end{aligned}$$

As above, the Ruler expropriates iff

$$U(NE_{T-2}, \tau_A, D_{T-1}) < U(E_{T-2}). \tag{9}$$

The backward induction proceeds to stage 2 of period $T-2$, where the argument repeats.

Before we proceed, we need to establish a useful result.

Lemma 4 *For a given set of parameters, the difference $U(NE_t, \tau_A, D_{t+1}) - U(E_t)$ is single-peaked with a peak at some finite $\widetilde{t}$.*

Proof. See the Appendix. ■

As shown above, the Ruler's devolution decision is determined by the difference between her payoff from expropriation in period t and her payoff from devolution in period $t+1$. This Lemma establishes that there is a unique period of time $\widetilde{t}$ where the respective difference reaches its maximum and that it monotonously declines for $t > \widetilde{t}$.

Continuing to solving the model backwards, there are two possible scenarios: (i) either in period $\widetilde{t}$, the expropriation payoff falls below the non-expropriation payoff

$$U(E_{\widetilde{t}}) < U\left(NE_{\widetilde{t}}, \tau_A, D_{\widetilde{t}+1}\right)$$

or (ii) in each period $t \geq 0$, expropriation is preferred over non-expropriation. Denote the set of parameters supporting case (i) by Ω_1, and the set of parameters supporting case (ii) by Ω_2. It is easy to see that neither of these cases is degenerate, that is, that both sets Ω_1 and Ω_2 are non-empty.

Indeed, assume that $b = 0$, and consider the difference between the payoffs from expropriation and non-expropriation (followed by devolution) in the initial period $t = 0$.

As output is increasing over time,

$$\begin{aligned}U\left(NE_0,\tau_A,D_1\right)-U\left(E_0\right) &= \varepsilon_D\sum_{t=1}^{\infty}\beta^t y_t-(1-\varepsilon_A)y_0>\varepsilon_D\sum_{t=1}^{\infty}\beta^t y_1-(1-\varepsilon_A)y_0\\ &= y_0\left[\frac{\varepsilon_D\beta}{1-\beta}\frac{y_1}{y_0}-(1-\varepsilon_A)\right]. \qquad (10)\end{aligned}$$

For given values of $(\varepsilon_A,\ \varepsilon_D,\beta,A)$, if the initial capital is sufficiently low, the growth rate in the first period

$$\frac{y_1}{y_0}=\left(\alpha\beta AK_0^{(\alpha-1)}\right)^{\alpha}$$

is sufficiently high for expression (10) to be positive. Continuity ensures that the same holds for small positive b. By Lemma 4

$$U\left(NE_{\widetilde{t}},\tau_A,D_{\widetilde{t}+1}\right)-U\left(E_{\widetilde{t}}\right)\geq U\left(NE_0,\tau_A,D_1\right)-U\left(E_0\right)>0,$$

which proves the nonemptiness of the set Ω_1.

Alternatively, consider the peak period $\widetilde{t}$. As output increases towards the steady-state value y^*,

$$\begin{aligned}U\left(NE_{\widetilde{t}},\tau_A,D_{\widetilde{t}+1}\right)-U\left(E_{\widetilde{t}}\right) &= \varepsilon_D\sum_{i=1}^{\infty}\beta^i y_{\widetilde{t}+i}-(1-\varepsilon_A)y_{\widetilde{t}}-\frac{\beta b}{1-\beta}\\ &< \varepsilon_D\sum_{i=1}^{\infty}\beta^i y^*-(1-\varepsilon_A)y_{\widetilde{t}}-\frac{\beta b}{1-\beta}. \qquad (11)\end{aligned}$$

Note that due to the separability of the payoff function, $\widetilde{t}$ is independent of b. Given $(\varepsilon_A,\ \varepsilon_D,\beta,A,K_0)$, one can choose a sufficiently high b so that

$$0<\varepsilon_D\sum_{i=1}^{\infty}\beta^i y^*-\frac{\beta b}{1-\beta}<(1-\varepsilon_A)y_{\widetilde{t}}.$$

Hence, expression (11) is negative, i.e., the considered parameters belong to the set Ω_2. We proceed by analyzing these two cases separately.

Case i

By construction, the payoff of expropriation in $\widetilde{t}$ is less than the payoff of devolution in $\widetilde{t}+1$. Moreover, we know that the Ruler prefers expropriation to devolution at the steady state, i.e., when $t\to\infty$. By Lemma 4, there exist $\widehat{t}\geq\widetilde{t}$ such that the Ruler prefers to expropriate at stage 3 of any period $t>\widehat{t}$, but not to expropriate at stage 3 of period $\widehat{t}$. It implies that the Ruler devolves at stage 1 of period $\widehat{t}+1$. The reason is simple: if the Ruler were to stay in power at stage 1 of period $\widehat{t}+1$, by non-expropriating at stage stage 3 of period $\widehat{t}$, she would receive a flow of private benefits of control as well as tax revenue in period $\widehat{t}$. Thus, her payoff from not expropriating in $\widehat{t}$ falls short of the expropriation payoff as the latter includes both the private benefits and the entire output in period $\widehat{t}$. That is,

$$U\left(NE_{\widehat{t}}, \tau_A, D_{\widehat{t}+1}\right) > U\left(E_{\widehat{t}}\right). \tag{12}$$

By backward induction, the Citizen stays in the market at stage 2 of period $\widehat{t}$. Indeed, no expropriation takes place at stage 3 and the devolution occurs at stage 1 of period $\widehat{t}+1$. Hence, there is no expropriation along this (sub)game path and, by Corollary 1, the Ruler sets $\tau_{\widehat{t}} = \tau_A$, and the Citizen chooses to stay. At stage 1 of period $\widehat{t}$, the Ruler compares the option of devolving to the option of proceeding to the next stage. Devolution yields the Ruler's payoff

$$U\left(D_{\widehat{t}}\right) = \varepsilon_D \sum_{i=0}^{\infty} \beta^i A K_{i+\widehat{t}}^{\alpha}. \tag{13}$$

Non-devolution followed by taxation entails a positive tax revenue in this period and the devolution payoff from the next period on

$$U\left(ND_{\widehat{t}}, NE_{\widehat{t}}, \tau_A, D_{\widehat{t}+1}\right) = \varepsilon_A A K_{\widehat{t}}^{\alpha} + b + \varepsilon_D \sum_{i=1}^{\infty} \beta^i A K_{i+\widehat{t}}^{\alpha}. \tag{14}$$

As taxes under autocracy are higher than after devolution, $\varepsilon_A > \varepsilon_D$ and $b \geq 0$, the Ruler does not devolve at stage 1 of period $\widehat{t}$.

Before proceeding backwards to period $\widehat{t}-1$, we need to establish an intermediate result. Consider the Ruler's net payoff

Lemma 5 *As time passes, the Ruler's payoff from expropriation net of private benefits of control becomes more attractive relative to the payoff from devolution net of private benefits of control.*

Proof. See the Appendix. ∎

Intuitively, the further away is the economy from the steady state, the longer is the growth horizon and the more appealing it is to get a share of future increasing profits (through devolution of power), as compared to grabbing the entire pie today.

Now, we are ready to discuss the choice of Ruler in the period $\widehat{t}-1$.

Lemma 6 *At stage 3 of period $\widehat{t}-1$, the Ruler prefers non-expropriation over expropriation.*

Proof. See the Appendix. ∎

By definition, at stage 3 of period $\widehat{t}$, the Ruler prefers non-expropriation, followed by devolution, to expropriation. The Ruler's choice between non-expropriation and expropriation is determined by two factors: the growth rate of output and the private benefits of control that are lost upon devolution. Now turn to period $\widehat{t}-1$. By choosing not to expropriate in period $\widehat{t}-1$, the Ruler retains the benefits of control for period $\widehat{t}$, as she does not devolve in that period. Therefore, as compared to period $\widehat{t}$, expropriation at $\widehat{t}-1$ is associated with a smaller gain in the private benefits of control. In addition, the growth rate of output in period $\widehat{t}-1$ is higher than at $\widehat{t}$, thereby providing the Ruler with additional incentive to refrain from expropriation. Hence, the Ruler chooses not to expropriate at stage 3 of period $\widehat{t}-1$,

$$U\left(NE_{\widehat{t}-1}, \tau_A, ND_{\widehat{t}}, NE_{\widehat{t}}, \tau_A, D_{\widehat{t}+1}\right) > U\left(E_{\widehat{t}-1}\right).$$

As above, at stage 2 of period $\widehat{t}-1$, the Citizen decides to stay. At stage 1 of period $\widehat{t}-1$ the Ruler again chooses whether to devolve. If she does not devolve, her payoff is

$$\begin{aligned} U\left(ND_{\widehat{t}-1}, NE_{\widehat{t}-1}, \tau_A, ND_{\widehat{t}}, NE_{\widehat{t}}, \tau_A, D_{\widehat{t}+1}\right) &= \varepsilon_A AK^{\alpha}_{\widehat{t}-1} + b + \beta\left(\varepsilon_A AK^{\alpha}_{\widehat{t}} + b\right) \\ &+ \sum_{i=1}^{\infty} \beta^i \varepsilon_D AK^{\alpha}_{i+\widehat{t}}. \end{aligned}$$

An immediate devolution yields the same flow of payments from period $\widehat{t}+1$ onwards, but lower payoffs in periods $\widehat{t}-1$ and $\widehat{t}$

$$U\left(D_{\widehat{t}}\right) = \varepsilon_D AK^{\alpha}_{\widehat{t}-1} + \beta\varepsilon_D AK^{\alpha}_{\widehat{t}} + \sum_{i=1}^{\infty} \beta^i \varepsilon_D AK^{\alpha}_{i+\widehat{t}}.$$

Therefore, the Ruler chooses not to devolve at stage 1 of period $\widehat{t}-1$. At stage 3 of period $\widehat{t}-2$, we repeat the argument of Lemma 6 to conclude that the Ruler again prefers not to expropriate.

We continue solving the model backwards along the same lines until we reach period $t=0$. In all these steps, the optimal strategy for the Ruler is to tax the citizen without devolution or expropriation. The optimal strategy for the Citizen is to stay in the market sector. The tax rate chosen by the Ruler is set to $\tau_t = \tau_A$ in each period $t = 1, 2, ...\widehat{t} - 1$.

Thus, we can conclude that there is a unique pure strategy SPNE where the Ruler taxes the Citizen while retaining autocratic power and not expropriating up to period $\widehat{t}$, and devolves in period $\widehat{t}$; that is

$$\sigma(h^t) = \left(\begin{array}{l} (ND_t, NE_t)\,,\ t = 0, ..., \widehat{t}, \quad D_{\widehat{t}+1}; \\ \tau_t = \begin{cases} \tau_A, & t = 0, ..., \widehat{t}; \\ \tau_D, & i = \widehat{t} + 1, ..., \infty \end{cases} \end{array} \right).$$

The citizen always stays in the market along the game path:

$$\rho(h^t) = (S_t,\ t = 0, ..\infty)\,.$$

We summarize our findings in the following proposition:

Proposition 2 *There exists a non-empty set of model parameters Ω_1, such that the Ruler prefers to devolve rather than expropriate in at least one period $\widetilde{t}$. If the parameters belong to the set Ω_1, then in a unique pure strategy SPNE, the Ruler taxes the Citizen until period $\widehat{t}$ and devolves in period $\widehat{t}$, where $\widehat{t}$ is the latest period in which the Ruler prefers devolution over expropriation.*

Case ii

In this case, in each period $t \geq 0$ expropriation is preferred over non-expropriation followed by devolution in the next period. The resulting equilibrium depends on the relation between the Ruler's payoff to devolution in the initial period and the value of the future flow of private benefits of control.

More precisely, suppose that the parameter values are such that in period $t = 0$, the Rulers's valuation of immediate devolution exceeds the valuation of the flow of control benefits:

$$U(D_0) > \frac{b}{1 - \beta}. \tag{15}$$

As the payoff to devolution increases over time, the same holds for every subsequent period $t > 0$:

$$U(D_t) > \frac{b}{1 - \beta}.$$

Thus, when we solve the game backwards, at stage 3 of every period the Ruler prefers to expropriate, at stage 2 the Citizen leaves the market and at stage 1 the Ruler chooses to devolve. This also holds for the very first period $t = 0$, which means that the devolution occurs immediately, before any growth in this economy takes place. Here, the Ruler values growth after devolution more than stagnation in an autocratic regime. To achieve any growth, she must relinquish her power in the very first period, otherwise the citizen immediately switches to a non-market activity.

In this case, in the unique pure strategy SPNE, the ruler devolves in the initial period $t = 0$, that is

$$\sigma(h^t) = \begin{pmatrix} D_0; \\ \tau_t = \tau_D, t = 0, ..\infty \end{pmatrix}.$$

The citizen again never leaves the market sector along the game path:

$$\rho(h^t) = (S_t, \ t = 0, ..\infty).$$

Conditions (13) and (15) suggest that such an equilibrium can take place in economies where e.g. the initial capital is relatively high, the private benefits of control are moderate, or the Ruler's power to tax after devolution is relatively large. By construction, the set of parameters supporting this equilibrium is a subset of Ω_2. Denote this set by Ω_2^d. This set is non-empty. For example, this outcome can be observed in economies that start up very close to the steady state. As shown above, the Ruler always prefers to expropriate; the efficiency assumption ensures that the flow of private benefits of control falls short of the devolution payoff in the steady state which, by continuity, also holds in a neighborhood of the steady state.

Proposition 3 *There exists a non-empty set of model parameters Ω_2^d, such that the Ruler's payoff from expropriation is always higher than the payoff from the devolution in the next period, but the payoff from devolution in $t = 0$ exceeds the value of the flow of private benefits of control. If the parameters belong to the set Ω_2^d, devolution occurs in the very first period and the economy fully realizes its growth potential.*

Finally, consider the situation where, in period $t = 0$, the Ruler's payoff of immediate devolution is lower than the valuation of the flow of control benefits:

$$U(D_0) < \frac{b}{1-\beta}. \tag{16}$$

Here, when we solve the game backward, the Ruler still expropriates at stage 3 of each period t. This holds in period T by Lemma 3. It also holds in all earlier periods irrespective of the Ruler's choice at stage 1 of period $t + 1$. Therefore, the Citizen again leaves at stage 2. But now the Ruler prefers not to devolve at stage 1 of early periods, as devolution is not sufficiently attractive. In particular, there is no devolution in the initial period $t = 0$ where inequality (16) holds. Therefore, such an economy also becomes locked in an underdevelopment trap: at the beginning of growth, the post-devolutionary future does not look sufficiently tempting to the Ruler. Thus, she prefers to retain all her political power to receive the private benefits of control, even though the Citizen immediately leaves the market to avoid expropriation. As a result, growth in this economy never occurs.

Formally, in the unique pure strategy SPNE along the game path, the ruler does not devolve in period $t = 0$:

$$\sigma(h^t) = (ND_0),$$

and the citizen leaves the market sector in the very initial period:

$$\rho(h^t) = (L_0).$$

Similarly to the above, conditions (13) and (15) suggest that this equilibrium outcome is observed in economies with low initial capital, high benefits of control and relatively limited Ruler's power to tax after devolution. Denote the respective set of parameters by $\Omega_2^u \subset \Omega_2$.

Proposition 4 *There exists a non-empty set of model parameters Ω_2^u such that the Ruler's payoff from expropriation is always higher than the payoff from the devolution in the next period, and the value of the flow of private benefits of control exceeds the payoff from devolution at $t = 0$. If the parameters belong to the set Ω_2^u, the economy is caught in an underdevelopment trap: no growth or devolution of power ever occurs.*

4 Predictions of the model

Now we are ready to discuss the model's predictions and comparative statics, as well as the importance of our modelling assumptions.

Proposition 5 *In an economy that is growing under autocratic rule, a higher level of initial capital entails earlier devolution of power. In an economy locked in an underdevelopment trap, an increase in initial capital may entail early devolution of power and growth.*

The formal proof can be found in the Appendix. Informally, consider two economies, one starting with the initial capital K_0 and another - with the capital that the first economy would reach in period $t = 1$,

$$K_0' = K_1 > K_0.$$

The backward induction procedure described above implies that the choices made by the agents in the first economy in period t are identical to the choices made by the agents in the second economy in period $t - 1$. Thus, if in the former economy the devolution of power occurs at date $\widehat{t}$, in the latter economy it occurs at date $\widehat{t} - 1$. Intuitively, an increase in the initial capital, other things equal, implies that the economy starts closer to the steady state and experiences lower growth rates throughout its development path. As a result, the future does not look that tempting for the Ruler and her incentive to grab at each point in time increases. Thus, in order to persuade the citizens to remain in the market, the Ruler needs to devolve power earlier.

Alternatively, assume that the former economy is in the "underdevelopment trap". That is, in any period $t > 0$, the Ruler prefers expropriation at stage 3 of period t to devolution at stage 1 of period $t + 1$ and her devolution payoff in period $t = 0$ is less than the flow of private benefits of control. Then, an increase in initial capital to $K_0' = K_1$ does not influence the relative attractiveness of expropriation, as compared to the devolution in the next period. Indeed, as mentioned above, the decisions made in the latter economy at time t replicate the decisions made in the former at time $t + 1$, so the Ruler of the economy starting with K_0' still prefers expropriation at any $t' > 0$. Therefore, the institutions are fully determined by the Ruler's devolution decision in period $t = 0$. Note that the Ruler's devolution payoff is increasing in the level of the initial capital, while the flow of private benefits of control is constant. As $K_0' = K_1 > K_0$, the devolution of power in period $t = 0$ in the latter economy brings the Ruler as much as the devolution in period $t = 1$ in the former economy, which is higher than the payoff to devolution in the initial period in the former economy:

$$U\left(D_0|K_0'\right) = U\left(D_1|K_0\right) > U\left(D_0|K_0\right).$$

Hence, it may be the case that the devolution payoff at K_0' exceeds the flow of private benefits of control and thus, in the latter economy, the devolution of power occurs in the very initial period $t' = 0$.

The above discussion implies the following two corollaries.

Corollary 3 *Poorer economies are more likely to be locked in an underdevelopment trap.*

An underdevelopment trap equilibrium outcome can only arise in an economy where the set parameters $(\alpha, A, \beta, \tau_D, \tau_A, b) \in \Omega_2^u$, so that the Ruler never prefers devolution over expropriation. As shown above, among these economies, higher initial capital leads to early devolution and growth, while lower capital blocks the economic and institutional development.

Corollary 4 *Among growing economies, wealthier economies experience earlier devolution of power.*

Next, let us address the impact of private benefits of control on the devolution of power. The Ruler retains the benefits of control only while being in power, whereas these benefits are no longer available to her after devolution. Thus, private benefits do not influence the Ruler's trade-off between early and late expropriation. Instead, they only have an impact on the Ruler's incentive to devolve. If the benefits are low, the Ruler can delay devolution for a long time, because the Citizen realizes that the Ruler will not cling to power the day she needs to commit not to expropriate to avoid losing investment. On the contrary, if private benefits of control are very high, the Ruler will not be willing to ever give up power, no matter how much capital accumulation is lost. Recognizing this, the Citizen never invests in the market sector and there is neither devolution nor growth. Thus, an abundance of natural resources has a detrimental effect on growth. In the intermediate range, as the private benefit of control increases, the Ruler's relative incentive to expropriate, as opposed to devolution, increases too. As a result, in order to keep capital accumulation going, the Ruler needs to devolve earlier since, at later stages, she always prefers to stay in power, and no commitment is possible. Thus, in this range, an increase in private benefits speeds up the devolution of power. Therefore, the model predicts a non-linear effect of the private benefits of control on the devolution of power. This prediction of the model parallels the arguments of the resource curse literature – that abundant natural resources may hinder growth[12] and capital accumulation.[13] Moreover, it suggests that the resource curse should also be non-linear: an increase in resource size does not have any impact on growth, until the resource rents become sufficiently large to completely kill growth.[14] We summarize

[12]E.g. see Sachs and Warner(1995) and Gylfason (2001)

[13]See Gylfason and Zoega (2001)

[14]A non-linear (negative) effect of the resource curse on growth is found by e.g. Sala-i-Martin and Subramanian (2003)

our findings in the following Proposition.

Proposition 6 *If the Ruler's private benefits of control are sufficiently small, an increase in the private benefits causes earlier devolution of power. Eventually, a further increase in the private benefits of control locks an economy in an underdevelopment trap with neither growth nor devolution.*

Proof. See the Appendix. ■

Note that the maximum post-devolutionary tax τ_D available to the Ruler and the private benefits of control are two sides of the same coin. That is, both of them reflect the Ruler's loss associated with giving up power. Thus, the effect of τ_D should be similar to that of the benefit of control. Indeed, a higher post-devolutionary tax rate makes the devolution option more attractive relative to the expropriation. Thus, the Ruler can credibly postpone devolution without jeopardizing industrialization. On the other hand, if the Ruler's payoff after devolution is very low (or very uncertain), she has no incentive to devolve and the economy is in the underdevelopment trap.[15] For such an economy, a sufficient increase in τ_D will cause an eventual devolution of power.

Proposition 7 *If the post-devolutionary tax rate τ_D is sufficiently low, the economy is locked in the underdevelopment trap. An increase in the post-devolutionary tax rate first causes devolution to occur and then delays it.*

Proof. See the Appendix. ■

Similarly, a change in the cost of taxation $\phi(.)$ inducing an increase in τ_A and Ruler's tax revenues received under autocracy has the same impact on the timing of devolution as an increase in τ_D. (Here we only consider an increase in τ_A which does not change the Ruler's incentive to expropriate in the steady state, and the incentive of the citizen to stay in the market as long as there is no expropriation, so that the assumption of Lemma 3 and condition (10) continue to hold). This result has a very simple explanation. When the Ruler decides whether or not to expropriate, she weights the expropriation payoff against the payoff from non-expropriation today and devolution next period. If she chooses to devolve, she retains the today's autocratic tax revenue and receives the devolution payoff from tomorrow onwards. Therefore, by expropriating at period t, she foregoes the tax revenue $\varepsilon_A y_t$. The higher is this revenue, the weaker are her incentives

[15] E.g. consider an extreme case when the Ruler cannot be credibly guaranteed any post-devolutionary payment from the citizen.

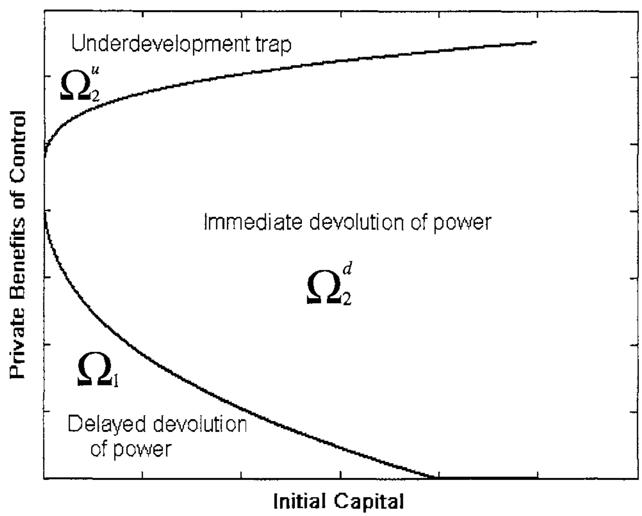


Figure 2: Development trajectories

to expropriate and the longer she can stay in power without using the commitment device. On the other hand, if the economy is locked in an underdevelopment trap, the tax revenues from capital accumulation under autocratic regime are too low to presuade the Ruler to forgo the private benefits of control. In this case an increase in τ_A may bring about an eventual devolution.

Proposition 8 *If the autocratic tax rate τ_A is sufficiently low, the economy is locked in the underdevelopment trap. In this case a change in the cost of taxation $\phi(.)$ inducing an increase in τ_A causes eventual devolution of power. In a growing economy an increase in τ_A delays devolution..*

Proof. See the Appendix. ■

To illustrate our findings, we consider numerical simulations with parameters $A = 1$, $\alpha = 0.36$, $\tau_D = 0.35$, $\tau_A = 0.4$ and $\beta = 0.7$.[16] In figure 2, we graph the set of equilibria in our economy as a function of the initial capital and the private benefits of control.

[16]The values of A and a are standard for numerical simulations of the Ramsey model. The value of the time discount β captures the fact that in our model we have complete depreciation over one period.

We see that low values of initial capital and private benefits of control result in an equilibrium with delayed devolution of power. An increase in either initial capital or the benefit of control brings the economy into the area of immediate devolution of power. A further increase in the private benefit of control leads to the "underdevelopment trap" equilibria with neither growth nor devolution of power.

Now consider a technological change – an increase in the total factor productivity parameter A. Intuitively, a country with a higher total factor productivity has a higher growth rate in each period and steady-state capital. So, at each point in time, this country's future growth potential weakens the incentives to expropriate. As a result, we expect the devolution of power to be delayed. On the other hand, the value of devolution relative to the value of expropriation increases with TFP (e.g. due to a higher growth rate in each period). Thus, higher total factor productivity may improve the chances for eventual devolution in economies in an underdevelopment trap.

Proposition 9 *If an economy is in the underdevelopment trap, higher total factor productivity* ***may create the*** *devolution of power and growth. For two growing economies, an economy with higher total factor productivity, other things equal, experiences later devolution of power.*

Proof. See the Appendix. ■

Other things equal, higher total factor productivity translates into a higher growth rate. Hence, we have an immediate Corollary:

Corollary 5 *Among growing economies, autocracies are more likely to experience higher growth rates than are less autocratic regimes.*

An increase in labor intensity α or in discount factor β has an ambiguous effect.

What are the predictions on the relationship between growth and the political regime? In our model, depending on the parameter values, there are two possible regimes: Either an economy is locked in the underdevelopment trap and neither growth nor devolution occurs, or the economy sustains an autocratic regime at higher growth rates and devolves as growth slows down. Thus, we see that the dictatorships are characterized either by no growth or by high growth rates, while less autocratic regimes fall in an intermediate range. The source of this cross-sectional variation may be different initial conditions, different stages of development (that is, the time of acquiring the technology) or a difference in technology per se. This prediction is consistent with the finding of Almeida and Ferreira (2002), who show that the cross-country variability of growth rates is higher among autocracies, and that autocracies are likely to be the best and the worst performers in terms of growth.

Second, according to the model, devolution of power is often preceded by several periods of growth. Therefore, it may look as if, in line with the results of Barro (1999) and others, the model establishes a causal relationship from growth to democratization, at least in the Granger-sense. But this conclusion is misleading: in the model, institutions of limited government and growth are determined simultaneously and endogenously. Indeed, if the institutions facilitating the devolution of power are missing in an economy, so that the power cannot be credibly relinquished, growth will not occur. On the other hand, in the absence of growth opportunities, the government would never self-impose any checks and balances. Therefore, the observed time pattern between growth and democratization does not necessarily reflect a causal relation.

5 Discussion

In this section, we discuss the key assumptions of our model and their implication for the results. In the model, we abstract from open conflict. That is, the only threat that the Citizen can make to the Ruler is that of exit. Introducing the possibility of conflict into the model would not change the nature of its predictions, but might bring some additional insights. Assume that the Citizen can struggle with the Ruler in order to force her to relinquish the power. Then, as growth declines and the Ruler is prepared to give up power in the near future, we may expect both the Ruler and the Citizen to fight less intensively. Thus, unlike in a stationary setting where the intensity of struggle would typically be constant, we may find that as growth slows down, uproars become less violent. This extension could be helpful in relating the model to the evidence, which suggests that even peaceful devolutions of power are normally preceded by some pressure on the ruler.

Another key assumption is that the ruler can commit to relinquish the power. In other words, the institutional structure in the economy allows for the devolution of power. The formation of these institutions is beyond the scope of our analysis. Clearly, if such institutions are lacking, the Ruler cannot credibly commit not to expropriate. Similarly, we assume that the citizens can, in turn, guarantee the ruler a "safe haven" after she devolves. If such an institution is missing, the autocrat has no incentive to devolve. Therefore, in the absence of either of these institutions, the model would predict any economy to be locked in an underdevelopment trap.

What happens if we allow the Ruler to "fine tune" the devolution decision? That is, suppose the Ruler may choose the post-devolutionary tax level (while keeping the assumption that devolution is associated with the absence of expropriation). Under this relaxed assumption, the Ruler does not impose any restrictions on the post-devolutionary tax return. The devolution of power occurs in any period between $t = 0$ and the period when the payoff of expropriation is just below the payoff of eternal taxation. Indeed, the Ruler can now make devolution as profitable as taxation and is thus indifferent concerning the timing of the devolution, as long as the Citizen does not leave the market. For example, if the Citizen has a weak preference for devolution, the Ruler is ready to devolve in the very initial period. That is, such "fine tuning" prevents us from predicting the precise timing of devolution. However, as long as the "fine tuning" implies post-devolutionary loss in private benefits of control, the inefficient "underdevelopment trap" equilibrium outcome continues to exist.

Some of the less realistic predictions of the model are artifacts of simplification. For example, along the equilibrium path, no expropriation occurs in the model, while we do observe examples of the government's predative behavior in real life. Lack of expropriation in the model is due to the fact that we have assumed perfect information and no uncertainty. If we instead assume that the Citizen is imperfectly informed, or there are random shocks to the production function, we extend the set of SPNE by including equilibria involving expropriation of capital along the equilibrium path. Similarly, uncertainty can yield "revolutionary" equilibria, that is, equilibria with conflict-driven devolution of power.

In our pure-strategy SPNE, no autocracy can survive in the long run, while we do observe non-collapsing autocracies in the real world (e.g. consider China). However, this does not imply that the model contradicts the evidence. The growth prospects might be sufficiently good, so that the devolution stage has not yet been reached. Also, we have deliberately confined the attention to a set of parameters under which expropriation is preferred in the steady state. Relaxing this assumption can produce equilibria with an eternal growing autocracy.

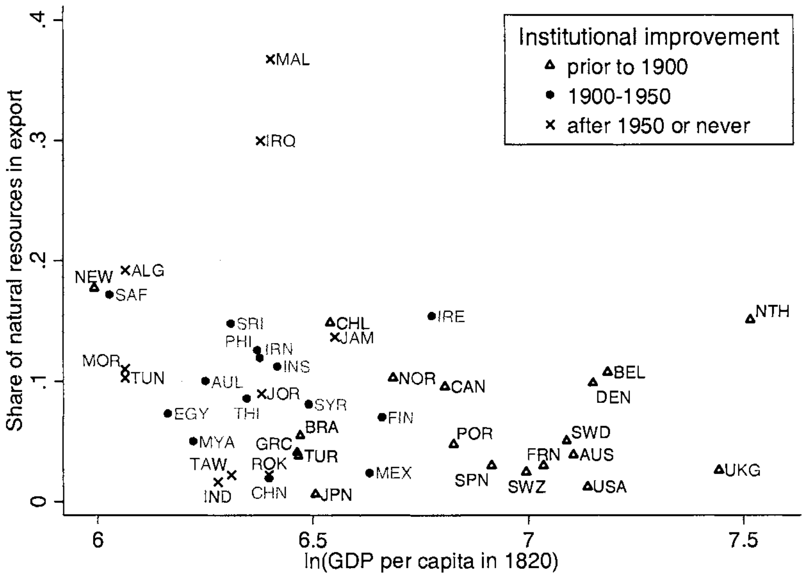
Finally, the model predicts that the growth rate declines after the devolution of power. Note that in our model, the devolution of power is, in fact, represented by an improvement in property rights protection. Most empirical studies record the opposite effect – better property right protection spurs growth. This effect can easily be incorporated in the basic framework. For example, we might extend the model to allow for more sectors. If some sectors do not have a diversion opportunity, these sectors will start to accumulate only in the absence of the expropriation threat. That is, the devolution of power will give rise to an additional wave of investment and growth, not attainable under autocracy.

6 Case studies

In this section, while not aiming at systematic empirical analysis, we relate the predictions of the model to a few observed patterns. As a preliminary step, we try to replicate the pattern of Figure 2, which links the initial capital and private benefits of control to the institutional transformation. For a set of countries, we plot the proxy of the initial capital against the proxy of the private benefits, grouping the countries by the time of their first institutional improvement after 1820. The result is presented in Figure 3. We see that, indeed, countries that have higher initial capital and moderate levels of natural resources experience earlier institutional change. In particular, countries that improved their institutions before 1900 are concentrated in the south-east corner of the figure. Countries that experienced institutional improvement between 1900 and 1950, tend to have lower capital, concentrating in the south and south-west part of the figure. Finally, countries that had their first increase in Polity after 1950 and countries that did not ever face an increase in Polity occupy the south-west and the north-west part of the picture with low initial capital and/or higher natural resources. Therefore, the data roughly corresponds to the model-generated Figure 2. However, we must mention that data on natural resources in 1820 (which is taken to be the basis year) is not available. With the use of resource data from 1970, we are likely to end up with a notably noisy diagram.

Therefore, we proceed by discussing a few cases providing support to the patterns predicted by the model. The model generates two general classes of development trajectories: either the economy stagnates under an autocratic rule, or it starts growing. In the latter case, the economy may experience early or late devolution of power.

There are numerous examples of countries stagnating under a kleptocratic autocracy. Consider, for example, the Democratic Republic of the Congo (former Zaire). This country, abundant in natural resources such as diamonds, uranium, copper and cobalt, until very recently was suffering from extreme inequality and poverty, having an average per capita GDP growth of –2.8% over the last 30 years (see Figure 4). For 32 years (1965-1997), it was under the dictatorship of Joseph Mobutu-Sese Seko. He



— Institutional improvement is measured as an increase by a minimum of 3 units in the *Change* variable the from Polity IV data set. It ranges from -20 to 20, and is equal to the net difference between the last recorded polity value and the new polity value across a continuous polity change. Positive values correspond to transitions towards less autocratic regimes. Countries that were highly democratic by the initial time of entry into the Polity IV database are classified as having their first institutional improvement at the moment of entry.

— Initial capital is proxied by the log of a country's GDP per capita in 1820 (Maddison (2006))

— Private benefits of control are measured by the country's abundance of natural resources (which is proxied by the share of natural resources in total export in 1970).

Figure 3

started his rule by nationalizing foreign-owned firms and handing their management to relatives and close associates who stole the companies' assets. He captured the control over the resource sector and heavily exploited it. By the early 1980s, his personal wealth was estimated at $5 billion (Leslie 1987), while the rest of the country was basically a subsistence economy (only five per cent of the population were estimated to work in the formal sector during the 1990's). Why was this stagnant path chosen? We propose a two-fold answer: The private benefits of being in control of such a resource-rich economy were high. In addition, there was no institutional way in Congo to guarantee Mobutu a sufficient part of the returns (including the natural sector rents)

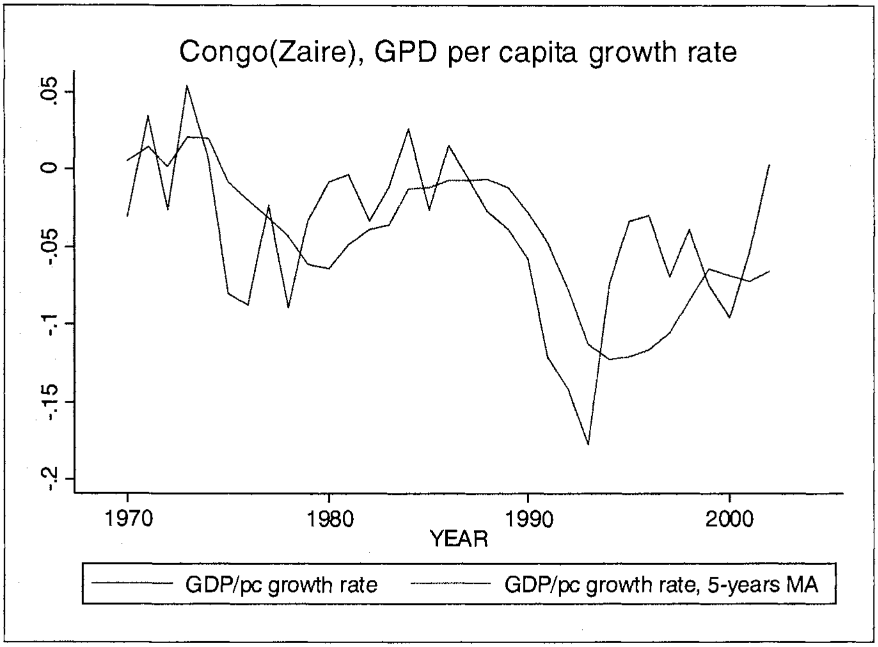


Figure 4

in case he were to limit his expropriative power.[17] Therefore, Mobutu was willing to forgo the potential future gains from capital accumulation for the immediate benefits of the resource extraction.

The People's Republic of China provides an example of the development trajectory where the country grows under autocratic rule and the devolution stage has not yet arrived. Since the country's establishment in 1949, it is ruled by the Communist Party of China (CPC) under a one-party system. While there are some limitations on the Party's power, the regime is still considered to be largely autocratic (the corresponding Polity score from the Polity IV dataset ranges between −9 in 1969 and −7 since 1976). However, the country is recognized as one of the most impressive economic performers, growing at the average annual GDP rate of 9.4% and an average annual GDP per capita rate of 8% for the past 25 years (WDI database). The pattern of GDP per capita growth rates in China is presented in Figure 5. One can observe an increase in the growth

[17] In this sense, Congo differed markedly from Botswana where, as argued by Acemoglu et al. (2003), the institutional reforms were not challenging the stability of the elite.

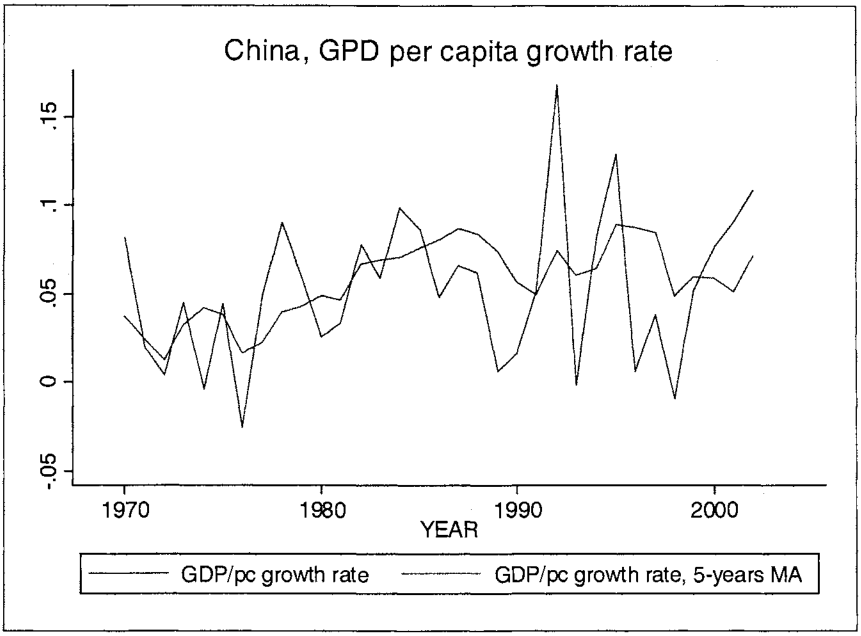
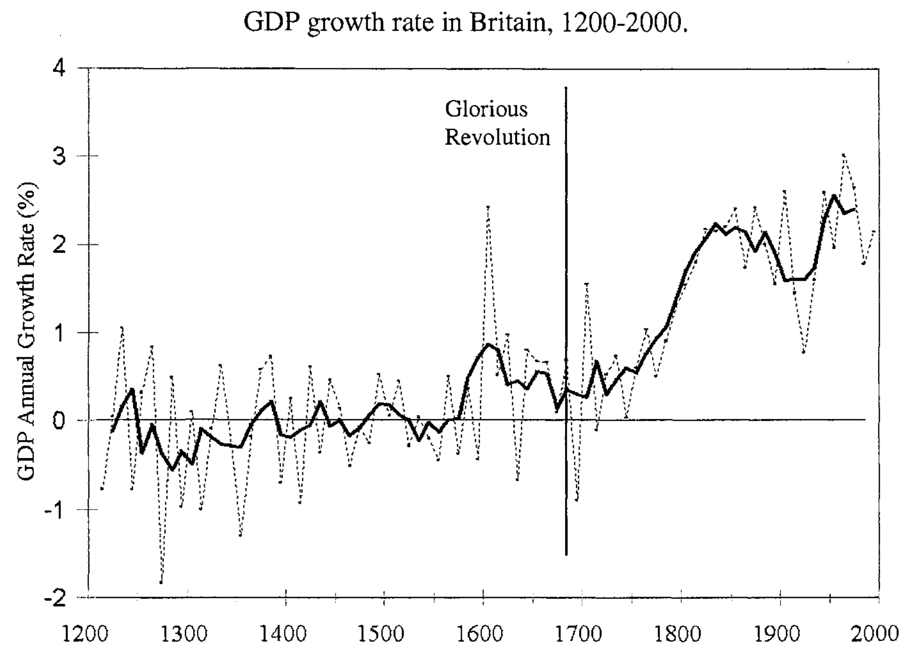


Figure 5

rate until the end of the 1980s followed by the relative stabilization, perhaps with a slightly negative trend. In terms of our model, the growth opportunities for China are sufficiently promising (i.e. the growth rate is still fairly high and its decline is moderate) for devolution to not yet be necessary to induce capital accumulation.[18]

Finally, let us turn to the examples of devolution. Given the model's somewhat Marxian spirit, notably that the change of regime is caused by capital accumulation, it is natural to look for supportive evidence in the historical examples of bourgeois state transformations. Consider the Glorious Revolution of 1688 in Britain. As argued in the seminal paper of North and Weingast (1989), the Glorious Revolution established the institutions of limited autocracy, allowing the government to credibly commit to secure property rights and eventually leading to economic growth. However, the historical data on economic growth in Britain (see Figure (6)) suggests that not only did the

[18]Alternatively, we are only discussing the set of parameters under which in the steady state expropriation is preferred over all other continuations of the game. Relaxing this assumption (i.e. assuming a very patient ruler) can produce equilibria with a sustained autocracy as the outcome.



From Clark (2005), ch. 10 fig. 4.
Dotted line corresponds to the growth rate of real output per year from the previous decade.
Solid line corresponds to a 50-year moving average

Figure 6

economy start to grow well before the end of the 17th century but, more interestingly, the growth rates were declining towards the time of the Glorious Revolution and not accelerating until another 50 years after it. That is, the constraints on autocratic power brought by the Glorious Revolution were imposed in the declining growth phase of development, which supports the proposed mechanism.[19]

[19]Obviously, the reality is much more complex than the model. We cannot claim that it was (only) the slowdown in growth that caused the institutional transformation. Many other factors and shocks were probably involved. In addition, while the confiscatory power of the Crown clearly existed (and was exercised in the first half of the 17th century through forced loans, enclosure fines etc.), it is not clear what is the empirical counterpart of the model's expropriation and accumulation of assets. One potential candidate would be investments in the quality of land (such as enclosures and the development of new agricultural techniques), being hampered by heavy and unpredictable taxation. Alternatively, it could be argued that the expansion of the East India Company, and its close association with the Crown (see, e.g., Pincus (2002), had put at risk established trade and

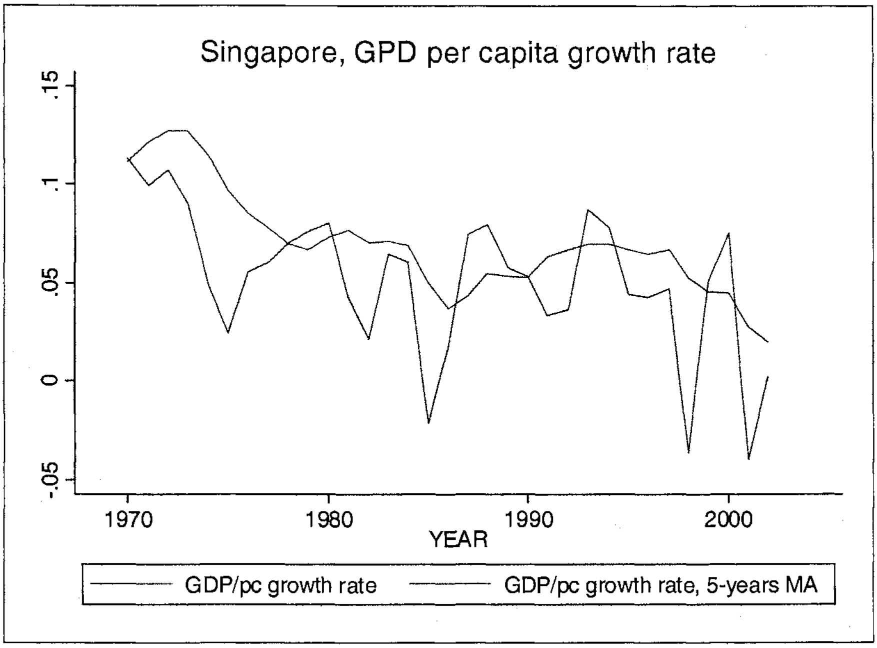


Figure 7

There are, however, more recent examples allowing us to observe the devolution mechanism in practice. Singapore, being a one-party state with Prime Minister Lee Kuan Yew holding his position from 1959 to 1990, is believed to be quite authoritarian. (The Polity IV database measures the constraints on executive power by 3 out of 7). However, as documented by Yap (2003), there are several episodes in the recent history of Singapore when the government creates commitment mechanisms to avoid private sector divestment. Consider, in particular, the period 1985-1986. As can be seen from Picture 4, the per capita GDP growth rates in Singapore were declining from the early 1970 towards the mid-1980s. In 1985, after a period of growth rate decline, the government introduced several policy changes aimed at increasing the private sector's monitoring of and participation in policy setting. To name just a few, the government replaced the Finance Minister with a former private sector leader, while returning budgetary policy-making to the Finance Ministry. It pursued a policy of divestment of

trade-related capital.

state ownership in the public-private joint-ownership enterprises. It created an Economic Review Committee, comprising six business representatives and six government representatives to reexamine the government's ten-year Economic Development Plan. While these measures did not change the actual political regime in Singapore, they clearly imposed additional constraints on the government's authority, demonstrating its commitment to non-expropriative policies. Moreover, the attempts to improve the credibility were not a regular practice of the Singaporean government (Yap (2003)), but rather a peculiar characteristic of some historical episodes. This piece of evidence, and especially its timing, illustrates the delayed devolution trajectory.

7 Conclusions and Extensions

We have built a model that addresses the interplay between devolution of political power and economic growth. The model suggests that if there are decreasing returns to capital accumulation, the ruler will be tempted to expropriate at high levels of capital and lower levels of growth. Foreseeing this, the investors cease to invest unless the ruler credibly commits not to expropriate. If being in power is not associated with high private benefits, the ruler self-imposes institutional checks and balances to protect entrepreneurs' property rights. In this case, the devolution of power occurs after a period of sustained growth, unless the initial capital is so high that the ruler is tempted to expropriate even before growth starts. In the latter case, devolution of power precedes growth. If instead the benefits of control are high, the autocrat sacrifices capital accumulation to keep these benefits. Such an economy never develops.

The model can be extended in several directions. One extension is to study how competition between rulers influences the incentives to cling to power. The threat of being overthrown tomorrow increases the relative value of current payoff and strengthens the incentive to expropriate which, in turn, is recognized by the private sector. This leads to earlier devolution. At the same time, the prospects of future devolution weaken the incentives to cling to power. Therefore, in our setting, competition for power is expected to yield earlier devolution in growing economies. The intensity of the power struggle would also depend on the development path – stagnating economies are expected to face more violent power conflicts, while growing economies would experience less violent conflicts as the time of devolution approaches. Alternatively, one can study oscillating regimes, allowing for a conflict technology so that citizens can replace the ruler and the ruler can mount a coup to return to power.

Another extension is to consider different types of growth. If growth is driven by improvement in the quality of products, each of the products is on the market only for a limited period of time. Thus, the incentive to grab for a ruler increases as the product may not be around tomorrow. This would lead to earlier devolution of power. One important prediction of this extension would be as follows: Acemoglu, Aghion and Zilibotti (2003) suggest the growth strategy of less developed countries is investment-based, while more developed economies switch to innovation-based growth. Thus, we may expect earlier devolution of power in more technologically advanced countries.

References

[1] Daron Acemoglu. Why not a political Coase theorem? Social conflict, commitment, and politics. *Journal of Comparative Economics*, 31(4):620–652, 2003.

[2] Daron Acemoglu, Phillippe Aghion, and Fabrizio Zilibotti. Distance to frontier, selection, and economic growth. NBER Working Papers 9066, National Bureau of Economic Research, 2002.

[3] Daron Acemoglu, Simon Johnson, and James Robinson. Institutions as the fundamental cause of long-run growth. NBER Working Papers 10481, National Bureau of Economic Research, 2004.

[4] Daron Acemoglu, Simon Johnson, James Robinson, and Pierre Yared. Income and democracy. NBER Working Papers 11205, National Bureau of Economic Research, 2005.

[5] Daron Acemoglu, Simon Johnson, and James A. Robinson. Reversal of fortune: Geography and institutions in the making of the modern world income distribution. *Quarterly Journal of Economics*, 118:1231–1294, 2002.

[6] Daron Acemoglu, Simon Johnson, and James A. Robinson. The rise of Europe: Atlantic trade, institutional change and economic growth. NBER Working Papers 9378, National Bureau of Economic Research, 2002.

[7] Daron Acemoglu, Simon Johnson, and James A. Robinson. *An African Success Story: Botswana.* Princeton University Press, 2003.

[8] Daron Acemoglu and James A. Robinson. Why did the west extend the franchise? democracy, inequality, and growth in historical perspective. *The Quarterly Journal of Economics*, 115(4):1167–1199, 2000.

[9] Daron Acemoglu and James A. Robinson. A theory of political transitions. *American Economic Review*, 91(4):938–963, 2001.

[10] Heitor Almeida and Daniel Ferreira. Democracy and the variability of economic performance. *Economics and Politics*, 14(3):225–257, 2002.

[11] Michael E. Alvarez, Adam Przeworski, Jose Antonio Cheibub, and Fernando Limongi. *Democracy and Development: Political Institutions and Well-Being in the World, 1950-1990.* Cambridge: Cambridge University Press, 2000.

[12] Jean-Paul Azam, Robert H Bates, and Bruno Biais. Political predation and economic development. CEPR Discussion Papers 5062, C.E.P.R., 2005.

[13] Robert J Barro. Democracy and growth. *Journal of Economic Growth*, 1(1):1–27, 1996.

[14] Robert J. Barro. Determinants of democracy. *Journal of Political Economy*, 107(6):158–29, 1999.

[15] Timothy Besley. Monopsony and time-consistency: Sustainable pricing policies for perennial crops. *Review of Development Economics*, 1(1):57–70, 1997.

[16] Charles Boix and Susan Stokes. Endogenous democratisation. *World Politics*, 55(4):517–549, 2003.

[17] Anne Boschini, Jan Pettersson, and Jesper Roine. Resource curse or not: A question of appropriability. Working Paper Series in Economics and Finance 534, Stockholm School of Economics, 2003.

[18] Christopher Clague, Philip Keefer, Stephen Knack, and Mancur Olson. Property and contract rights in autocracies and democracies. *Journal of Economic Growth*, 1(2):243–76, 1996.

[19] Gregory Clark. The conquest of nature: A brief economic history of the world, 10,000 bc-2000 ad. Unpublished manuscript, University of California, Davis, 2005.

[20] Paul Collier and Anke Hoeffler. Democracy and resource rents. GPRG Working Papers 016, Global Poverty Research Group, Oxford, UK, 2005.

[21] Simeon Djankov, Edward L. Glaeser, Rafael La Porta, Florencio Lopez de Silanes, and Andrei Shleifer. The new comparative economics. NBER Working Papers 9608, National Bureau of Economic Research, 2003.

[22] Milton Friedman. Economic freedom behind the scenes. In *Economic Freedom of the World Report: 2002 Annual Report.* Cato Institute, 2002.

[23] Edward L. Glaeser, Rafael La Porta, Florencio Lopez de Silane, and Andrei Shleifer. Do institutions cause growth? *Journal of Economic Growth*, 9(3):271–303, 2004.

[24] Arthur A. Goldsmith. Democracy, property rights and economic growth. *The Journal of Development Studies*, 32(2):157–174, 1995.

[25] Mark Gradstein. Inequality, democracy and the emergence of institutions. CEPR Discussion Paper 4187, C.E.P.R., 2004.

[26] Hershel I. Grossman. Production, appropriation, and land reform. *The American Economic Review*, 84(3):705–712, 1994.

[27] Thorvaldur Gylfason. Natural resources, education, and economic development. CEPR Discussion Papers 2594, C.E.P.R., 2000.

[28] Thorvaldur Gylfason and Gylfi Zoega. Natural resources and economic growth: The role of investment. EPRU Working Paper Series 01-02, Economic Policy Research Unit (EPRU), University of Copenhagen. Department of Economics (formerly Institute of Economics), 2001.

[29] Robert E. Hall and Charles I. Jones. Why do some countries produce so much more output per worker than others? *The Quarterly Journal of Economics*, 114(1):83–116, 1999.

[30] Witold J. Henisz. The intitutional environment for economic growth. *Economics and Politics*, 12(1):1–31, 2000.

[31] Philip Keefer. What does political economy tell us about economic development - and vice versa? Policy Research Working Paper Series 3250, The World Bank, 2004.

[32] Steven Knack and Philip Keefer. Institutions and economic performance: Cross-country tests using alternative measures. *Economics and Politics*, 7:207–227, 1995.

[33] Charles Kursman, Regina Werum, and Ross E. Burkhart. Democracy's effect on economic growth: A pooled time-series analysis, 1951-1980. *Studies in Comparative International Development*, 37(1):3–33, 2002.

[34] Martin C. McGuire and Mancur Olson. The economics of autocracy and majority rule: The invisible hand and the use of force. *Journal of Economic Literature*, 34(1):72–96, 1996.

[35] Halvor Mehlum, Karl Moene, and Ragnar Torvik. Institutions and the resource curse. *Economic Journal*, 116(508):1–20, 2006.

[36] Douglass C. North. Institutional change and economic growth. *The Journal of Economic History*, 21(1):118–125, 1971.

[37] Douglass C. North and Robert Paul Thomas. An economic theory of the growth of the western world. *The Economic History Review*, XXIII(1):1–17, 1970.

[38] Douglass C. North and Barry R. Weingast. Constitutions and commitment: Evolution of institutions governing public choice in seventeenth century England. *Journal of Economic History*, 49:803–832, 1989.

[39] Mancur Olson. Dictatorship, democracy, and development. *The American Political Science Review*, 87(3):567–576, 1993.

[40] Torsten Persson. Forms of democracy, policy and economic development. NBER Working Papers 11171, National Bureau of Economic Research, 2005.

[41] Torsten Persson and Guido Tabellini. Democracy and development: The devil in the details. CEPR Discussion Papers 5499, C.E.P.R., 2006.

[42] Torsten Persson and Guido Tabellini. Democratic capital: The nexus of political and economic change. Unpublished manuscript, Institute for International Economic Studies, 2006.

[43] Steven Pincus. Whigs, political economy, and the revolution of 1688-89. Unpublished manuscript, Chicago University, 2002.

[44] Adam Przeworski. The last instance: Are institutions the primary cause of economic development? *European Journal of Sociology*, 45(2):165–188, 2004.

[45] Adam Przeworski and Fernando Limongi. Political regimes and economic growth. *Journal of Economic Perspectives*, 7(3):51–69, 1993.

[46] James A. Robinson, Ragnar Torvik, and Thierry Verdier. Political foundations of the resource curse. *Journal of Development Economics*, 79(2):447–468, 2006.

[47] Michael M. Ross. The political economy of the resource curse. *World Politics*, 51:297–322, 1999.

[48] Michael M. Ross. Does oil hinder democracy? *World Politics*, 53:325–361, 2001.

[49] Jeffrey D. Sachs and Andrew M. Warner. Natural resource abundance and economic growth. NBER Working Papers 5398, National Bureau of Economic Research, 1995.

[50] Xavier Sala-i-Martin and Arvind Subramanian. Addressing the natural resource curse: An illustration from Nigeria. NBER Working Papers 9804, National Bureau of Economic Research, 2003.

A Appendix

A.1 Proof of Lemma 4

The difference between the payoff from non-expropriation and the payoff from expropriation in period t is given by

$$U(NE_t, \tau_A, D_{t+1}) - U(E_t) = \varepsilon_D \sum_{i=1}^{\infty} \beta^i y_{t+i} - (1-\varepsilon_A) y_t - B, \tag{17}$$

where B denotes the flow of the private benefits of control as of tomorrow, $B = b\beta/(1-\beta)$.

As the taxation in our model does not influence the capital development path, a capital level in each time period $t+i$ is

$$K_{t+i} = (\alpha\beta A)^{\frac{1-\alpha^i}{1-\alpha}} (K_t)^{\alpha^i} \tag{18}$$

and the output is

$$y_{t+i} = AK_{t+i}^{\alpha} = A(\alpha\beta A)^{\alpha\frac{1-\alpha^i}{1-\alpha}} (K_t)^{\alpha^{i+1}}. \tag{19}$$

Thus, equation (17) is equivalent to

$$U(NE_t, \tau_A, D_{t+1}) - U(E_t) = \varepsilon_D A \sum_{i=1}^{\infty} \beta^i (\alpha\beta A)^{\alpha\frac{1-\alpha^i}{1-\alpha}} (K_t)^{\alpha^{i+1}} - (1-\varepsilon_A) AK_t^{\alpha} - B.$$

Let us introduce an auxiliary continuous function of the capital

$$F(k) = \varepsilon_D A \sum_{i=1}^{\infty} \beta^i (\alpha\beta A)^{\alpha\frac{1-\alpha^i}{1-\alpha}} k^{\alpha^{i+1}} - (1-\varepsilon_A) Ak^{\alpha} - B.$$

Note that $F(K_t)$ represents the difference between the Ruler's value of non-expropriation in period t, followed by devolution at stage 1 of period $t+1$, and expropriation at stage 3 of period t as a function of the capital K_t. We study how this function changes with k. The derivative of $F(k)$ declines over time, being positive around zero and negative at the steady state capital K^*. Indeed,

$$\begin{aligned}
\frac{\partial F(k)}{\partial k} &= \frac{\alpha}{k}\left[\varepsilon_D A \sum_{i=1}^{\infty} \beta^i \alpha^i (\alpha\beta A)^{\alpha\frac{1-\alpha^i}{1-\alpha}} (k)^{\alpha^{i+1}} - (1-\varepsilon_A) Ak^{\alpha}\right] \\
&= \frac{\alpha A}{k^{1-\alpha}}\left[\varepsilon_D \sum_{i=1}^{\infty} \beta^i \alpha^i (\alpha\beta A)^{\alpha\frac{1-\alpha^i}{1-\alpha}} (k)^{\alpha(\alpha^i-1)} - (1-\varepsilon_A)\right].
\end{aligned}$$

As $k \to 0$, the expression in square brackets becomes infinitely large. On the other hand, at $k = K^*$, the expression in brackets equals

$$\varepsilon_D \frac{\alpha\beta}{1-\alpha\beta} - (1-\varepsilon_A) < \varepsilon_A \frac{1}{1-\alpha\beta} - 1 = \frac{\tau_A - \phi(\tau_A)}{1+\tau_A} - 1 < 0.$$

As $F'(k)$ is also continuous, there exists a threshold value $\tilde{k}$, such that

$$\begin{aligned}\frac{\partial F(k)}{\partial k} &\geq 0, k \leq \tilde{k} \\ \frac{\partial F(k)}{\partial k} &\leq 0, k > \tilde{k}\end{aligned}$$

and, consequently, $F(k)$ increases for $k \leq \tilde{k}$, reaches its maximum at $\tilde{k}$ and decreases for $k > \tilde{k}$. Note that $F(0) = -B \leq 0$ and

$$F(K^*) = A(K^*)^\alpha \left(\frac{\varepsilon_D}{(1-\beta)} - 1\right) - B < 0.$$

Therefore, if in the economy in consideration $K_0 > \tilde{k}$, the peak of the expression $U(NE_t, \tau_A, D_{t+1}) - U(E_t)$ corresponds to period $\widetilde{t} = 0$. If instead $K_0 < \tilde{k}$, the maximum of $U(NE_t, \tau_A, D_{t+1}) - U(E_t)$ is achieved in period $\widetilde{t}$, such that

$$\left|K_{\widetilde{t}} - \widetilde{k}\right| = \min_t \left|K_t - \widetilde{k}\right|.$$ [20]

A.2 Proof of Lemma 5

Consider the difference between the ratio of devolution payoff to the payoff to expropriation, both net of private benefits of control, for two subsequent periods

$$\frac{U(NE_{t-1}, \tau_A, D_t) - b}{y_{t-1}} - \frac{U(NE_{t-2}, \tau_A, D_{t-1}) - b}{y_{t-2}} = \varepsilon_D \sum_{i=0}^{\infty} \beta^{i+1} \left[\frac{y_{i+t}}{y_{t-1}} - \frac{y_{i+t-1}}{y_{t-2}}\right].$$

As the growth rate of output is decreasing,

$$\frac{y_t}{y_{t-1}} > \frac{y_{i+t}}{y_{i+t-1}}$$

if and only if

$$\frac{y_{i+t}}{y_t} - \frac{y_{i+t-1}}{y_{t-1}} < 0.$$

The latter condition is equivalent to

$$\frac{U(NE_{t-1}, \tau_A, D_t) - b}{y_{t-1}} - \frac{U(NE_{t-2}, \tau_A, D_{t-1}) - b}{y_{t-2}} < 0.$$

A.3 Proof of Lemma 6

The Ruler's expropriation payoff is

$$U\left(E_{\widehat{t}-1}\right) = y_{\widehat{t}-1} + \frac{b}{1-\beta}.$$

If she does not expropriate, she sets tax τ_A and receives private benefits for periods $\widehat{t}-1$ and $\widehat{t}$, and devolves at stage 1 of period $\widehat{t}+1$

$$U\left(NE_{\widehat{t}-1}, \tau_A, ND_{\widehat{t}}, NE_{\widehat{t}}, \tau_A, D_{\widehat{t}+1}\right) = \varepsilon_A y_{\widehat{t}-1} + b + \beta\left[U\left(NE_{\widehat{t}}, \tau_A, D_{\widehat{t}+1}\right)\right]. \quad (20)$$

Using inequality (12), we see that the payoff from taxing at $\widehat{t}-1$ and $\widehat{t}$ and devolving at $\widehat{t}+1$ is higher than that from taxing at $\widehat{t}-1$ and expropriating at $\widehat{t}$

$$\begin{aligned} U\left(NE_{\widehat{t}-1}, \tau_A, ND_{\widehat{t}}, NE_{\widehat{t}}, \tau_A, D_{\widehat{t}+1}\right) &> \varepsilon_A y_{\widehat{t}-1} + b + \beta U\left(E_{\widehat{t}}\right) \\ &= \varepsilon_A y_{\widehat{t}-1} + b + \beta\left[y_{\widehat{t}} + \frac{b}{1-\beta}\right] \\ &= \varepsilon_A y_{\widehat{t}-1} + \beta y_{\widehat{t}} + \frac{b}{1-\beta}. \end{aligned}$$

If we can now show that in period $\widehat{t}-1$, the growth rate is sufficiently high, so that the Ruler gains by taxing and postponing expropriation by one period

$$\varepsilon_A y_{\widehat{t}-1} + \beta y_{\widehat{t}} > y_{\widehat{t}-1}, \quad (21)$$

we can conclude that

$$U\left(NE_{\widehat{t}-1}, \tau_A, ND_{\widehat{t}}, NE_{\widehat{t}}, \tau_A, D_{\widehat{t}+1}\right) > U\left(E_{\widehat{t}-1}\right).$$

We prove that inequality (21) holds by contradiction. Inequality (21) is equivalent to

$$\frac{y_{\widehat{t}+1}}{y_{\widehat{t}}} > \frac{1-\varepsilon_A}{\beta}. \quad (22)$$

Suppose that inequality (22) does not hold, that is

$$\frac{y_{\widehat{t}+1}}{y_{\widehat{t}}} < \frac{1-\varepsilon_A}{\beta}. \quad (23)$$

As we know, at stage 3 of period $\widehat{t}$, the Ruler prefers non-expropriation followed by devolution of power to the expropriation:

$$U\left(NE_{\widehat{t}}, \tau_A, D_{\widehat{t}+1}\right) > U\left(E_{\widehat{t}}\right) = y_{\widehat{t}} + \frac{b}{1-\beta}.$$

This implies that her devolution payoff net of the private benefits is higher than the value of the expropriated output

$$U\left(NE_{\widehat{t}}, \tau_A, D_{\widehat{t}+1}\right) - b > y_{\widehat{t}},$$

or, equivalently,

$$\varepsilon_A y_{\widehat{t}} + \varepsilon_D \sum_{i=0}^{\infty} \beta^{i+1} y_{i+\widehat{t}} > y_{\widehat{t}}.$$

From Lemma 5, it immediately follows that the same holds at stage 3 of period $\widehat{t}-1$,

$$\varepsilon_A y_{\widehat{t}-1} + \varepsilon_D \sum_{i=0}^{\infty} \beta^{i+1} y_{i+\widehat{t}} > y_{\widehat{t}-1},$$

or equivalently,

$$\sum_{i=0}^{\infty} \beta^i \varepsilon_D \frac{y_{i+\widehat{t}}}{y_{\widehat{t}-1}} > 1 - \varepsilon_A. \tag{24}$$

Remember that output in our model is growing at a decreasing rate. Using inequality (23), we have

$$\frac{y_{\widehat{t}+i}}{y_{\widehat{t}-1}} = \frac{y_{\widehat{t}+i}}{y_{\widehat{t}+i-1}} \frac{y_{\widehat{t}+i-1}}{y_{\widehat{t}+i-2}} .. \frac{y_{\widehat{t}}}{y_{\widehat{t}-1}} < \left(\frac{1-\varepsilon_A}{\beta}\right)^{i+1}.$$

As a result, at stage 3 of period $\widehat{t}-1$, the ratio of tomorrow's devolution payoff net private benefits of control to the expropriated output must be below $1-\varepsilon_A$. Indeed,

$$\begin{aligned}
\sum_{i=0}^{\infty} \beta^{i+1} \varepsilon_D \frac{y_{i+\widehat{t}}}{y_{\widehat{t}-1}} &< \sum_{i=0}^{\infty} \beta^{i+1} \varepsilon_D \left(\frac{1-\varepsilon_A}{\beta}\right)^{i+1} = \\
&= (1-\varepsilon_A) \frac{\varepsilon_D}{\varepsilon_A}.
\end{aligned}$$

As $\varepsilon_D < \varepsilon_A$, we conclude that

$$\sum_{i=0}^{\infty} \beta^i \varepsilon_D \frac{y_{i+\widehat{t}}}{y_{\widehat{t}}} < 1 - \varepsilon_A,$$

contradicting inequality (24).

A.4 Proof of Proposition 5

Consider an economy with the initial capital K_0 where the devolution of power occurs at date $\widehat{t} > 0$. The fact that devolution occurs in period $\widehat{t}$ means that

$$\begin{aligned}
U\left(NE_{\widehat{t}-1}, \tau_A, D_{\widehat{t}}\right)(K_0) - U\left(E_{\widehat{t}-1}\right)(K_0) &> 0, \\
U\left(NE_{\widehat{t}+j-1}, \tau_A, D_{\widehat{t}+j}\right)(K_0) - U\left(E_{\widehat{t}+j-1}\right)(K_0) &< 0, j = 1, ..., \infty,
\end{aligned}$$

or, equivalently,

$$\varepsilon_D \sum_{i=0}^{\infty} \beta^i y_{\hat{t}+i}(K_0) - (1-\varepsilon_A) y_{\hat{t}-1}(K_0) - B \;>\; 0, \tag{25}$$

$$\varepsilon_D \sum_{i=j}^{\infty} \beta^{i-j} y_{\hat{t}+i}(K_0) - (1-\varepsilon_A) y_{\hat{t}+j-1}(K_0) - B \;<\; 0, j = 1, ..., \infty, \tag{26}$$

where B denotes the flow of the private benefits of control as of tomorrow, $B = b\beta/(1-\beta)$

Applying expressions (18) and (19) to conditions (25) and (26) yields

$$\varepsilon_D A \sum_{i=0}^{\infty} \beta^i (\alpha\beta A)^{\alpha\frac{1-\alpha^i}{1-\alpha}} (K_{\hat{t}})^{\alpha^{i+1}} - (1-\varepsilon_A) A K_{\hat{t}-1}^{\alpha} - B \;>\; 0, \tag{27}$$

$$\varepsilon_D A \sum_{i=j}^{\infty} \beta^i (\alpha\beta A)^{\alpha\frac{1-\alpha^i}{1-\alpha}} \left(K_{\hat{t}+j}\right)^{\alpha^{i+1}} - (1-\varepsilon_A) A K_{\hat{t}+j-1}^{\alpha} - B \;<\; 0, j = 1, ..., \infty. \tag{28}$$

With the use of the auxiliary function $F(k)$ introduced in the proof of Lemma 4, that set of inequalities (27-28) is equivalent to

$$F(K_{\hat{t}-1}) \;>\; 0, \tag{29}$$

$$F(K_{\hat{t}+j}) \;<\; 0, j = 0, ..., \infty. \tag{30}$$

There are two possibilities:

Case A. If the set of parameters is such that $F(\tilde{k}) > 0$, there exist $\underline{k}$, $\overline{k}$, such that

$$\underline{k} \leq \tilde{k} \leq \overline{k},$$

and

$$F(\underline{k}) = F(\overline{k}) = 0.$$

In such an economy, if the initial capital is sufficiently low ($K_0 < \overline{k}$), the devolution of power occurs in period t such that

$$F(K_{t-1}) > 0,$$
$$K_{t-1} \leq \overline{k},$$
$$K_{t+j} > \overline{k}, j = 0, ..., \infty,$$

which is equivalent to inequalities (29-30).

If the initial capital exceeds $\overline{k}$, the ruler devolves in period $t = 0$.

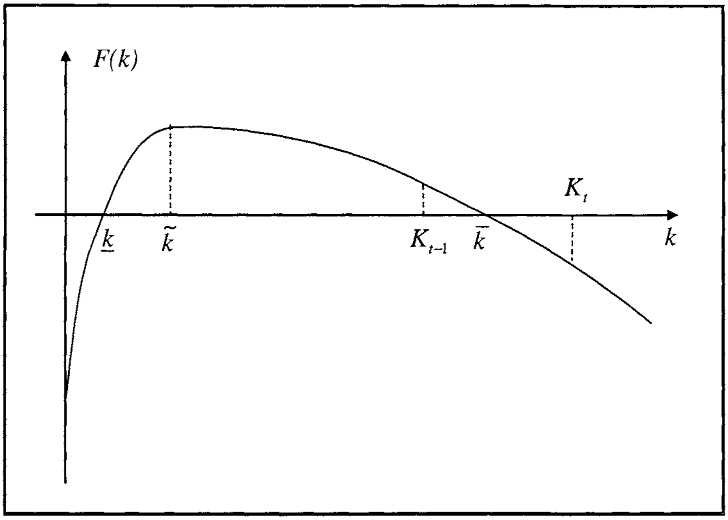


Figure 8: Case A

Case B. Alternatively, if $F(\tilde{k}) < 0$ (or, more weakly, $F(K_t) < 0$ for any $t = 0, ..., \infty$), the Ruler always prefers expropriation over devolution of power. In this case, there is a threshold level of capital k_{UD}, such that the value of devolution at k_{UD} is exactly equal to the value of the flow of private benefits:

$$U\left(D_0|k_{UD}\right) = \varepsilon_D A \sum_{i=0}^{\infty} \beta^i \left(\alpha\beta A\right)^{\alpha\frac{1-\alpha^i}{1-\alpha}} \left(k_{UD}\right)^{\alpha^{i+1}} = B.$$

As the devolution payoff increases with capital at the point of devolution, for any levels of initial capital below k_{UD}, the economy is in the "underdevelopment trap" – the capital is not accumulated and the power is never devolved. If initial capital is above k_{UD}, the devolution occurs in the initial period $t = 0$.

Now consider two economies, one starting with the initial capital K_0, and another with the initial capital $K_0' > K_0$, all other things equal. As $K_0' > K_0$, at each point in time, the capital stock (on or off market) of the second country will exceed that of the first country:

$$K_j' > K_j, j = 1, ..., \infty. \tag{31}$$

Start with the analysis of Case A. First assume that the initial capital K_0 is sufficiently

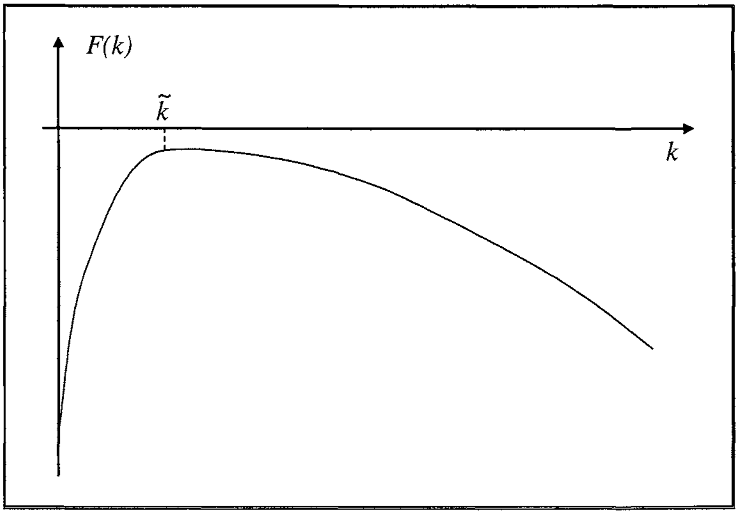


Figure 9: Case B

high

$$K_0 > \overline{k},$$

so that the devolution of power in this economy occurs in the initial period $t = 0$. It immediately follows that in the economy with the initial capital K_0', satisfying

$$K_0' > K_0 > \overline{k},$$

the devolution of power also occurs at $t = 0$.

Now assume that the initial capital K_0 is not very high so that devolution in the economy with the initial capital K_0 occurs at date $t > 0$, which is equivalent to

$$\begin{aligned} &F(K_{t-1}) > 0, \\ &K_{t-1} \leq \overline{k}, \\ &K_{t+j} > \overline{k},\ 0 = 1, ..., \infty. \end{aligned}$$

From (31) we conclude that in the economy with the initial capital K_0' devolution cannot occur after period t, as

$$K_{t+j}' > K_{t+j} > \overline{k},\ j = 0, ..., \infty.$$

Moreover, if the distance between K'_0 and K_0 is sufficiently large, it may be the case that $K'_t > \overline{k}$, or, equivalently, that $F(K'_t) < 0$, which implies that the devolution in the economy starting with K'_0 occurs strictly before period t.

There is a subtle point in this argument: if the set of parameter values is such that the segment $[\underline{k}, \overline{k}]$ where $F(k)$ is positive (i.e. the Ruler prefers devolution over expropriation), is too small, it may be the case that a particular capital accumulation path $(K'_0, K'_1, ...K'_t...)$ " misses" it. That is, there exists a t', such that the capital in period t' is below the "devolution segment" and the capital in period $t'+1$ is above it:

$$K'_0 < K'_1 < ... < K'_{t'} < \underline{k} \leq \overline{k} < K'_{t'+1} < ...$$

As a result, in such an economy, devolution of power either occurs in the very initial period (if K'_0 is sufficiently high) or never takes place. In this case, an increase in initial capital is not necessarily associated with the (weakly) earlier devolution of power. For example, it may be the case that in the economy with the initial capital K_0, the devolution of power occurs at some period $t > 0$ where

$$\underline{k} \leq K_{t-1} \leq \overline{k} < K_t,$$

while in the economy with the initial capital $K'_0 > K_0$, no devolution ever occurs. Such a situation is purely an artefact of the discrete nature of our game. One simple way of avoiding it is to assume that the adoption of technology requires a minimum initial capital/savings, this minimum being $K_{0\,\min} = \tilde{k}$ - the point of maximum of $F(k)$.[21] With this restriction for any initial capital below $K_{0\,\min}$, the country cannot accumulate capital and get growth going (and thus, no devolution of power ever occurs). If the initial capital is above $K_{0\,\min}$, $F(k)$ declines for any k and an increase in initial capital always results in a (weakly) earlier devolution of power. In our analysis, we will only consider economies with initial capital above $K_{0\,\min}$.[22]

In Case B, an increase in initial capital can affect the timing of devolution in only one situation: if K_0 is below the threshold k_{UD} and K'_0 is above it. In this case,

[21] To make this restriction only depend on the technological parameters, we assume there is a maximum tax $\overline{\tau}_d$ that the Ruler can charge after the devolution of power. As $\tilde{k}(\tau_d)$ decreases in τ_d, we set $K_{0\,\min} = \tilde{k}(\overline{\tau}_d)$.

[22] We believe that this assumption is very restrictive. For example, the numerical simulation for an economy with parameters $A = 1$, $\alpha = 0.36$, $\tau_d = 0.35$, $\tau_a = 0.4$, $\beta = 0.7$ shows that output at $\tilde{k}$ is app. 1.4% of the steady state output and that it takes app. 12 periods to (almost) reach the steady state capital $0.9999K^*$, if the economy starts from $\tilde{k}$. Note that in our model, we have complete depreciation over one period, so a period should be at least 10-15 years, which is also reflected in the value of the discount factor used in simulations.

an increase in initial capital from K_0 to $K_0' > K_0$ entails a change in the timing of devolution: in the economy with K_0, devolution of power never occurs, while the economy with K_0' faces an immediate devolution. If both K_0 and K_0' are below (or above) the "underdevelopment" threshold, a increase from K_0 to K_0' does not have any impact on devolution.

So we conclude that an increase in initial capital leads to a (weakly) earlier devolution of power.

A.5 Proof of Proposition 6

The higher is B, the lower is $F(k)$ for each level of capital k. Thus, with an increase in B the graph of $F(k)$ shifts downwards and the upper bound of the "devolution segment" $\overline{k}$ declines. Therefore, as the capital accumulation path is not affected by B, the devolution of power occurs at (weakly) lower levels of capital or, equivalently, at earlier periods in time.

If B increases even more, $F(k)$ becomes negative for any k, the "devolution segment" disappears and the economy falls into the "underdevelopment trap".

A.6 Proof of Proposition 7

If τ_D (and thus ε_D) is very low, $F(k)$ is negative for any k, and the economy falls into the "underdevelopment trap".

The higher is τ_D, the higher is $F(k)$ for each capital k. Thus, with an increase in τ_D, the graph of $F(k)$ shifts upwards and the upper bound of the "devolution segment" $\overline{k}$ increases. Therefore, as the capital accumulation path is not affected by τ_D, the devolution of power occurs at (weakly) higher levels of capital or, equivalently, at later periods in time.

A.7 Proof of Proposition 8

If τ_A (and thus ε_A) is very low, there is a parameter range so that $F(k)$ is negative for any k, and the economy falls into the "underdevelopment trap". The higher is τ_A, the higher is $F(k)$ for each capital k. Thus, with increase in τ_A the graph of $F(k)$ shifts upwards. Hence, an increase in τ_A may cause an economy to get out of an underdevelopment trap. In addition with an increase in τ_A. the upper bound of the "devolution segment" $\overline{k}$ increases. Therefore, as the capital accumulation path is not affected by τ_A, for higher τ_A devolution of power occurs at (weakly) higher levels of capital or, equivalently, at later periods in time.

A.8 Proof of Proposition 9

The Ruler devolves at the last point in time such that the devolution payoff exceeds the expropriation payoff (see conditions (25)-(26)). This can be rewritten as

$$\begin{aligned}
\varepsilon_D \sum_{i=0}^{\infty} \beta^i \frac{(\alpha\beta A)^{\frac{1-\alpha^{\hat{t}+i+1}}{1-\alpha}}}{\alpha\beta} (K_0)^{\alpha^{\hat{t}+i+1}} &> \frac{(\alpha\beta A)^{\frac{1-\alpha^{\hat{t}+1}}{1-\alpha}}}{\alpha\beta} (K_0)^{\alpha^{\hat{t}+1}} + B, \\
\varepsilon_D \sum_{i=j}^{\infty} \beta^{i-j} \frac{(\alpha\beta A)^{\frac{1-\alpha^{\hat{t}+i+1}}{1-\alpha}}}{\alpha\beta} (K_0)^{\alpha^{\hat{t}+i+1}} &< \frac{(\alpha\beta A)^{\alpha\frac{1-\alpha^{\hat{t}+j+1}}{1-\alpha}}}{\alpha\beta} (K_0)^{\alpha^{\hat{t}+j+1}} + B \\
& j = 1..\infty,
\end{aligned}$$

where K_0 is initial capital.

Consider the ratio of the Ruler's devolution payoff less the private benefit of control to the value of expropriated output in period t:

$$\frac{U(NE_{t-1}, \tau_A, D_t)(K_0) - B}{y_t(K_0)} = \frac{\varepsilon_D \sum_{i=0}^{\infty} \beta^i \frac{(\alpha\beta A)^{\frac{1-\alpha^{t+i+1}}{1-\alpha}}}{\alpha\beta} (K_0)^{\alpha^{t+i+1}} - B}{\frac{(\alpha\beta A)^{\frac{1-\alpha^{t+1}}{1-\alpha}}}{\alpha\beta} (K_0)^{\alpha^{t+1}}}. \tag{32}$$

The devolution of power occurs at the very last moment when this ratio is above 1. Note that the ratio is increasing in the productivity parameter A :

$$\frac{\partial}{\partial A}\left(\frac{U(NE_{t-1}, \tau_A, D_t)(K_0) - B}{y_t(K_0)}\right) > 0. \tag{33}$$

Consider two economies facing technologies with total factor productivity A and $A' > A$, respectively, and assume that the devolution of power in the former economy occurs at some period $\hat{t}$. Inequality (33) implies that in all time periods when the devolution of power is preferred under productivity A, it is also preferred under productivity A'. Thus, the devolution of power in the economy with A' cannot occur earlier than in $\hat{t}$.

On the other hand, assume that an economy with productivity A is in an underdevelopment trap, so that the ratio (32) is always below 1. By (33), higher productivity A' implies higher ratios (32) for all periods t and may, in fact, result in some of these ratios increasing above 1. Thus, in this case, higher productivity may cause devolution of power and growth.

CHAPTER 2

The Coase Theorem Is False*

with Tore Ellingsen

Abstract

The paper provides simple and robust counterexamples to the Coase Theorem. More precisely, we show that equilibrium investments in club goods can be suboptimally small despite the presence of well-defined and perfectly protected property rights and the absence of transaction costs and informational asymmetries. The reason is that, in equilibrium, a club of owners will typically not exercise their right to exclude outsiders, preferring instead to exercise their right to sell access. As long as the club of owners does not have all the bargaining power in such ex post access negotiations, strategic non-membership provides a valuable free-riding opportunity. Even in the case of only two persons, coercion may be needed in order to guarantee an efficient outcome.

1 Introduction

A modern statement of the Coase Theorem is that voluntary bargaining suffices to attain efficient outcomes if (i) property rights are well defined and perfectly enforced, (ii) parties have complete information, and (iii) negotiation is costless.[1] The theorem, which is based on arguments by Coase (1960), is rarely stated formally and has never

*Financial support from the Torsten and Ragnar Söderberg Foundation (Ellingsen) and the Jan Wallander's and Tom Hedelius Research Foundation (Paltseva) is gratefully acknowledged. Thanks to Avinash Dixit, Oliver Hart, and Volker Nocke for helpful discussions. We are responsible for remaining errors.

[1]The original formulations of the Coase Theorem were less restrictive. The complete information prerequisite was introduced after Myerson and Satterthwaite (1983) showed that incomplete information can lead to unavoidable efficiency losses.

been proven in full generality, yet is usually held to be true.[2] However, from the very beginning some economists have aired their doubts that such a general and strong result can hold. Recently, Dixit and Olson (2000) argued with the help of examples that if the voluntariness requirement is taken seriously, the efficient outcome is only one of many possibilities, and not a particularly likely one. In this paper we take their approach a small but crucial step further by offering a related set of examples in which *all* equilibrium outcomes are inefficient under the Coasean assumptions. That is, we conclusively prove that the Coase Theorem is false.

Our counterexamples are both simple and generic. They demonstrate that efficiency may be unattainable through voluntary negotiations when a group of people are to decide on their contributions to a club good.[3] Provision becomes inefficiently small because some people are better off by strategically abstaining from club membership, only to negotiate access to the good once it has been provided by the club's members.

Here is a sketch of our argument. Consider two families living on an island without electricity.[4] Both households would like to have an electricity cable to the island. Suppose cables come in two qualities. Each family values the high quality cable at 24 and the low quality cable at 20. The price of the low quality cable is 25 and the price of the high quality cable is 32. The jointly optimal outcome for the two families is to buy the high quality cable, because the total net benefit of $2 \cdot 24 - 32$ is greater than the total net benefit of $2 \cdot 20 - 25$. Supposing that the families has equal bargaining power, each family gets a net surplus of 8 under the efficient agreement. But the efficient agreement may not come about. Suppose one of the families, say family A, refuses to contribute anything to the purchase of the cable, relying instead on family B to purchase the cable by themselves – both families foreseeing that nothing prevents them from engaging in a new round of negotiations regarding access once the cable is in place. Sticking to the previous bargaining assumption, so that ex post negotiations also lead to an equal split of the available net surplus of 20 (the value of access to

[2] Coase's claim was baptised as theorem by George Stigler. For examples of its widespread acceptance among leading microeconomists, see for example Mas-Colell, Whinston and Greene (1995, page 357), Hart (1995, page 32), and Anderlini and Felli (2006).

[3] Club goods are defined by two properties: (i) There is no rivalry in consumption (they are public goods). (ii) It is possible to exclude potential consumers, for example on the basis of club membership criteria.

[4] In the Stockholm archipelago, there are many islands like this - with two summer houses and no electricity. One of the authors used to rent one such house. Recently, the owner of the other house wanted to have electricity, but the author's landlord did not want to chip in. The final outcome is eagerly awaited.

family A), it is straightforward to compute the outcome under this scenario. If family B buys the high quality cable, family B's surplus becomes $24 - 32 + 24/2 = 4$. If instead family B buys the low quality cable, its surplus is $20 - 25 + 20/2 = 5$. Hence, it prefers to invest a low quality cable, thus failing to maximize joint surplus. It now only remains to check that family A is wise to refuse an initial agreement. By refusing, family A's net payoff becomes $20/2 = 10$, which is greater than the 8 that it would get under the efficient initial agreement.

Compared to Dixit and Olson's examples, the present example has two distinguishing features. First, we assume that property rights contain the right to exclude outsiders, whereas Dixit and Olson focus on the case of pure public goods. Hence, we are immune to a potential critique against their model, namely that it violates Coase's property rights prerequisites. At the same time, the example demonstrates that the underlying incentive to free ride is qualitatively similar whether or not exclusion is allowed, as long as the owner does not have all the bargaining power in ex post access negotiations. Second, Dixit and Olson confine attention to the case of a binary investment choice – i.e., their investment is either zero or fully efficient. Because of the binary set–up, their example admits efficient equilibria. By allowing more than two investment levels, we ensure that all equilibria of our model are inefficient.

An objection to the above example is that we assume simultaneous moves at the club formation stage, and thereby preclude agreements in which early entrants (family B) compensate late entrants (family A) for joining the club before the investment is made.[5] We therefore also study simple sequential negotiations. Arguing that voluntary negotiation implies the right of any party to unconditionally leave the negotiation table, we find that inefficiency remains. In our example, this is most easily seen when the electricity cable, for technological reasons, has to be built at a certain date, T, whereas negotiations can proceed costlessly during an interval of time before T. Then, the first family to choose whether to initiate the project is better off by exiting the negotiation until just after date T, getting the outsider's payoff of 10, than by opting in and having to concede a net payoff of at least 10 to the other family.

The paper is organized as follows. In Section 2, we consider a simple parametric n-player simultaneous move club–formation game. We show that in equilibrium the club will typically not comprise all n players, because some players are better off

[5]Relatedly, inefficient equilibria sometimes disappear under sequential in/out decisions. For example, in the leading example of Dixit and Olson (2000) it is easy to check that only the efficient equilibrium remains if players move sequentially.

negotiating access to the club good ex post. As a result, there is underprovision of the club good. In Section 3, we consider a sequential game in which each player has exactly one opportunity to join the club. Finally, in Section 4 we assume that players negotiate sequentially, with each player having finitely many opportunities to join the club. Any player who remains at the negotiation table when it is her time to move has the choice between three options: (i) entering the club, at terms negotiated with existing club members (if any); not entering the club; (iii) leaving the negotiation table. The difference between options (ii) and (iii) is that once a player leaves the negotiating table, the player commits to stay away until the club good has been produced. Section 5 concludes.

2 The simultaneous membership model

A group of identical players are to decide on production of and access to a club good. The game has three stages. At stage 1, all players independently and simultaneously decide whether or not to join a club. At stage 2, the club members jointly decide on their contributions to the club good, and the club good is produced. The contribution of each club member is fully observable and club members can write binding and enforceable agreements among themselves.[6] Outsiders are unable to contract with club members at this stage. At stage 3, the club negotiates with outsiders the terms under which the latter are allowed access to the club good. Let N denote the set of all players and M the set of members in the club, with n and m denoting the number of elements in the respective sets. We assume that n is finite and greater than 1.

Each player is endowed with a fixed budget, b. The budget can be spent on private goods or club goods. There is only one type of each good. Private goods have a fixed quality and can be bought in any quantity at a constant price. The club good can only be produced in a fixed quantity, normalized to 1;[7] however the quality of the club good is variable and depends on the sum of the players' contributions. Let e_i denote the contribution of player i to the club good.

[6] As in Dixit and Olson (2000), one may think of the club as a jurisdiction that enables contracting among members. We refer to their paper for a careful discussion of these timing and contracting assumptions.

[7] For example, the club good may require a specific and unique site. The restriction to one club good simplifies our analysis by allowing us to disregard investment by outsiders – investment that might in principle be used to strengthen their bargaining position in ex post access negotiations.

The quality of the club good is assumed to be fixed after stage 2. Thus, it depends only on total stage 2 contributions. For simplicity we adopt the production function

$$q(e_1, ..., e_m) = \begin{cases} 2\left(\sum_{i \in M} e_i - \varepsilon\right)^{1/2} & \text{if } \sum_{i \in M} e_i - \varepsilon \geq 0; \\ 0 & \text{otherwise.} \end{cases}$$

Note that ε can be interpreted as the fixed cost associated with producing the lowest possible quality. For most of the paper we assume for convenience that $\varepsilon < 1$, which is small enough never to threaten production. The utility of any player j is assumed to be linear in the quantity of private goods and in the quality of the club good. In order to avoid corner solutions, we assume that the budget b is large enough to ensure that each player's marginal dollar is spent on private goods. Thus, player j's utility can be written

$$U(e_j, q) = k + q(e_1, ..., e_m) - e_j,$$

where k is a constant which is henceforth neglected. Note that the perfect substitutability of the players' efforts implies that it unimportant for overall efficiency which set of players actually produces the public good.

While the club good is non-rivalrous in consumption, providers have enforceable property rights to it. More precisely: each member of the provision club has access, outsiders can be denied access, and the club can sell access rights. (These properties define club goods.)

Suppose for a moment, counterfactually, that the club members were forced to deny access to any outsiders. The game would then only contain the first two stages – club formation and negotiation among club members. Suppose that at date 0 exactly m players were to enter the club. Since contributions are contractible among club members, contributions would maximize their joint utility. If positive, the contribution of member i must thus solve the problem

$$\max_{e_j} 2m\left(\sum_{i \in M} e_i - \varepsilon\right)^{1/2} - e_j,$$

which implies that

$$\sum e_i = m^2 + \varepsilon.$$

Assuming that costs are shared symmetrically, the utility of each club member j is given by

$$U(e_j, q) = m - \varepsilon/m.$$

By staying outside the club, player j would instead get the reference utility of 0. Since $\varepsilon < 1 \leq m$, we see that j is better off in the club than outside it. As the argument holds for any $m \geq 1$, in a subgame-perfect equilibrium of this game everyone wants to join the club and the production of the club good is socially efficient.

However, the premise that the club will deny access to any outsiders is untenable. *Ex post* there is always an incentive for the club to sell access to any willing outsider. Therefore, we return to the original timing, allowing the club to bargain with the outsiders at stage 3. Suppose that ex post bargaining gives the outsider a share β of the gains from trade. That is, if at stage 2 the club produces quality Q, an outsider receives a net utility of βQ and the club receives a payment of $(1-\beta)Q$. Assuming symmetric sharing of payoffs among club members, the game has the following equilibrium outcome.

Proposition 1 *The equilibrium size of the club is*

$$m^* = \min\left\{n, \ \text{integer}\left(\left(\frac{n^2(1-\beta)^2+\beta^2-\varepsilon}{\beta^2}\right)^{1/2}\right)+1\right\}.$$

Proof. First, derive the quality of club good produced at date 1, conditional on m players having decided to join the club at date 0. Taking into account the outcome of bargaining with $n-m$ outsiders at date 2, the club of m members maximizes

$$2m\left(\sum_{i\in M} e_i - \varepsilon\right)^{1/2} - \sum_{i\in M} e_i + (1-\beta)2(n-m)\left(\sum_{i\in M} e_i - \varepsilon\right)^{1/2} \tag{1}$$

subject to the non-negativity constraint. Since $\varepsilon < 1$ the solution is interior and can be written

$$\sum e_i = (\beta m + (1-\beta)n)^2 + \varepsilon. \tag{2}$$

Thus the quality produced by a club of size m is

$$q_m = 2\left(\beta m + (1-\beta)n\right). \tag{3}$$

From equation (2) it follows that the payoff of staying outside of club of size m is equal to

$$U_m^{Out} = 2\beta\left(\beta m + (1-\beta)n\right). \tag{4}$$

Given symmetric payoffs, the payoff of a member of a club of size m is

$$U_m^{In} = \frac{(\beta m + (1-\beta)n)^2 - \varepsilon}{m}. \tag{5}$$

Now turn to the participation decision at date 0. A club of size m constitutes an equilibrium if (i) no outsider wants to join, i.e., if

$$U_m^{Out} \geq U_{m+1}^{In}, \tag{6}$$

and (ii) no insider wants to leave, i.e., if

$$U_m^{In} \geq U_{m-1}^{Out}. \tag{7}$$

Let us assume for expositional simplicity that when the outsider is indifferent, she joins, while if the insider is indifferent, she stays in the club. Substituting (4) and (5) into (6) and (7) and solving the resulting system yields the optimal club size

$$m^* = \min\left[n, \text{integer}\left(\frac{\left(n^2\left(1-\beta\right)^2+\beta^2-\varepsilon\right)^{1/2}}{\beta}\right)+1\right],$$

(see Appendix for details). The quality of the club good produced in equilibrium q^* is given by formula (3) for $m = m^*$. ■

We are now ready to characterize some of the properties of the equilibrium.

Proposition 2 *The equilibrium size of the club m^* and the equilibrium quality provision q^* are weakly increasing in the size of the society n.*

Proof. See Appendix. ■

The intuition is quite simple. For a given size of the club and a given quality of the club good, a larger population means that the number of outsiders increases. This raises the payoff of insiders while keeping the outsiders' payoff unchanged. Therefore outsiders have additional incentive to join the club. At the same time, for a given club size the insiders' marginal return from increasing provision and selling access goes up, therefore amplifying the insiders' incentive to invest. If the additional player joins the club, the latter effect can only be strengthened, because the cost per member goes down.

Proposition 3 *The equilibrium size of the club m^* and the equilibrium level of club good provision q^* are weakly decreasing in the bargaining power of the outsiders β.*

Proof. See Appendix. ■

For a given size of the club and quality level of the club good, higher bargaining power of outsiders increases their payoff and lowers the payoff of the insiders. Therefore, the marginal club member may now be better off outside the club. For a given club size the marginal return to quality investment goes down due to the increased bargaining power of the outsiders. If any member leaves, the club invests even less, because the leaving player ends up contributing less toward financing the marginal quality unit.

Proposition 4 *In the simultaneous club formation game there exists a threshold $\widehat{\beta} < 1$ such that for any $\beta \in \left[\widehat{\beta}, 1\right]$ the unique subgame perfect equilibrium outcome is a club size $m^* < n$. In this equilibrium quality provision is inefficiently low.*

Proof. Suppose that at date 0 player n expects all other $n-1$ players to join the club. From formula (2) it follows that a club of the size $n-1$ produces the quality

$$\beta(n-1) + (1-\beta)n = n - \beta.$$

Thus if player n decides to stay outside, her payoff is

$$U_{n-1}^{Out} = 2\beta(n-\beta).$$

If instead she decides to join the club of size $n-1$ (so that it becomes the club of size n), she gets the payoff

$$U_n^{In} = n - \frac{\varepsilon}{n}.$$

Thus she prefers to stay outside if and only if

$$2\beta(n-\beta) > n - \frac{\varepsilon}{n}. \tag{8}$$

As $\beta \in [0,1]$, the expression $2\beta(n-\beta)$ reaches its maximum at $\beta_{\max} = 1$. For this $\beta_{\max}$ the inequality (8) becomes

$$2(n-1) > n - \frac{\varepsilon}{n},$$

or, equivalently

$$n > 2 - \frac{\varepsilon}{n}. \tag{9}$$

As long as $\varepsilon > 0$, inequality (9) holds for any $n \geq 2$. By continuity, inequality (8) is satisfied for some segment $\left[\widehat{\beta}, 1\right]$. This, in turn, implies that for any $\beta \in \left[\widehat{\beta}, 1\right]$ the n-th player chooses not to join the club. ■

Again, the intuition is simple. As long as outsiders have enough bargaining power, it is more tempting to stay outside than to join the club at date 0. If n and/or ε is large the critical bargaining power (the threshold $\widehat{\beta}$) can in fact be quite low for free riding incentives to preclude the efficient outcome.

While the results were derived under the assumption that club members divide utility equally, inefficiency can only get worse if asymmetric splitting is assumed. The reason is that players who expect to carry more than the average burden have a stronger incentive to stay outside. Asymmetric splitting can thus only be helpful in combination with sequential moves at the club formation stage.

3 The single-round sequential membership model

To demonstrate that the inefficiency is not due to simultaneous moves in the coalition formation game, we now assume instead that the players enter the club formation negotiations in a sequential order. At each stage, if no club was created before, player i has a choice of organizing the club or staying outside. If the club is already created, the organizer of the club makes an unconditional monetary offer to the new potential member i. If player i accepts the offer, she joins the club; otherwise she stays outside. Finally, once all players have had their opportunity to join, the club produces the club good whereafter the club organizer bargains with the outsiders over access.[8] As above, β denotes the bargaining power of the outsiders. We assume, until noted differently, that the club does not have all the bargaining power in ex post access negotiations ($\beta > 0$).

Lemma 1 *If player 1 organizes the club, she signs up player n in any subgame-perfect equilibrium.*

Proof. For each branch of the game, let Θ denote the "history" of that branch - that is, the collection of actions of players $1, ..., n$ along the branch. Let $P(i|\Theta)$ denote the payoff of player i conditional on Θ, and let $V(m)$ denote the aggregate payoff of a club of m members.

Consider a game tree node X where player 1 makes an offer to player n. Let Θ_{n-1} denote the "history" of this node - that is, the actions of players $i \in \{1, 2, 3, ..., n-1\}$. In other words, Θ_{n-1} summarizes the offers from player 1 to players $i \in \{2, 3, ..., n-1\}$

[8] At first sight, it may seem like a contradiction that club insiders can give take-it-or-leave-it offers in membership negotiations but not in access negotiations. However, the assumption has a game theoretic justification. The outsider's surplus under ex post access is an outside option in the membership negotiation. To the extent that outside options affect bargaining outcomes in non-cooperative alternating offers bargaining, they act as constraints. In the membership negotiation, the outsider's outside option is typically more attractive than is the insider's, and in many examples the outsider's outside option would indeed be binding if we were to consider an alternating offers bargaining game.

and whether they accepted or rejected these offers. Suppose that $m \leq n-2$ players accepted the offer along this game branch before it reached the node X, and let $M(\Theta_{n-1})$ denote the set of players who accepted the offer.

If player 1 want to sign player n in, she must offer player n at least her outside option. Among the continuations of the game beyond node X at which player n joins the club, the only candidate for SPNE is the one bringing n the minimum "joining" payoff. A history Θ_{n-1} followed by this continuation is denoted Θ_n^{In}. All the continuations of the game beyond node X at which player n does not join, are payoff-equivalent for all participants, so we denote them all Θ_n^{Out}. The minimum payoff to sign in player n is thus determined by the equation

$$P(n|\Theta_n^{In}) = P(n|\Theta_n^{Out}). \tag{10}$$

Note that the payments to the club members $i \in M(\Theta_{n-1})$ are already set when the game reaches node X and do not depend on whether player n accept or rejects the offer. That is,

$$P(i|\Theta_n^{In}) = P(i|\Theta_n^{Out}). \tag{11}$$

Consider the difference between the maximum payoff of player 1 in case she signs in player n and her maximum payoff in case she keeps player n outside:

$$\begin{aligned} P(1|\Theta_n^{In}) - P(1|\Theta_n^{Out}) &= \left(V(m+1) - \sum_{i \in M(\Theta_{n-1})} P(i|\Theta_n^{In}) - P(n|\Theta_n^{In}) \right) \\ &\quad - \left(V(m) - \sum_{i \in M(\Theta_{n-1})} P(i|\Theta_n^{Out}) \right). \end{aligned} \tag{12}$$

Applying formulas (10), (11) to equality (12) transforms it into

$$P(1|\Theta_n^{In}) - P(1|\Theta_n^{Out}) = V(m+1) - V(m) - P(n|\Theta_n^{Out}). \tag{13}$$

The payoff to the coalition of m members is

$$V(m) = (\beta m + (1-\beta)n)^2 - \varepsilon, \tag{14}$$

and the payoff of being outside such a coalition is

$$P_m^{Out} = 2\beta\left(\beta m + (1-\beta)n\right). \tag{15}$$

Using (14) and (15) to substitute for the terms in equation (13) yields

$$\begin{aligned} P(1|\Theta_n^{In}) - P(1|\Theta_n^{Out}) &= \left(\beta^2 (2m+1) + 2n\beta(1-\beta) - 2\beta (\beta m + (1-\beta)n)\right) \\ &= \beta^2 > 0. \end{aligned}$$

That is, no matter how large the current club is, player 1 is always better off by signing in player n, than she is by keeping n outside. ■

The intuition is that when the club's founder bargains with the last potential member, the founder is already committed and cannot escape. Since player n is willing to join if offered the utility emanating from staying outside, and since both parties can gain from a higher quality level than that which will result if n does stay outside, the founder will sign in player n.

However, a similar logic does not hold when player 1 negotiates with other players. The reason is that each entering club member is associated with higher quality, and thus improves outside options for subsequent potential entrants, making it more costly for the founder to sign them into the club.

Lemma 2 *If player 1 organizes the club, she does not sign up player $n-1$ in any subgame-perfect equilibrium.*

Proof. Consider a game tree node X' where player 1 makes an offer to player $n-1$. Similarly to above, denote by Θ_{n-2} the "history" of node X'. Denote the set of players who accepted the offer along this game branch before it reached node X' by $M'(\Theta_{n-2})$, with m' being its cardinality.

Again, to sign player $n-1$ in, player 1 should offer player $n-1$ at least her outside option. According to Lemma 1, in any SPNE player n is always signed in. If player $n-1$ stays outside, in a SPNE player 1 signs in player n for the minimum payoff, so that the final club size is $m'+1$. The respective game tree branch is denoted by Θ_{n-1}^{Out}. If player $n-1$ accepts the offer, the subgame-perfect continuation implies a club of $m'+2$ members. Again, among all the continuation games with both players $n-1$ and n signed player 1 prefers the one where both $n-1$ and n join for the minimum acceptable payoff. This game branch is denoted by Θ_{n-1}^{In}. So player $n-1$ joins if she is offered at least

$$P(n-1|\Theta_{n-1}^{In}) = P(n-1|\Theta_{n-1}^{Out}). \tag{16}$$

Again, like in (11) the payments to the club members joining the club prior to the node X' being reached, $i \in M'(\Theta_{n-2})$ do not depend on the continuation of the game, so

$$P(i|\Theta_{n-1}^{In}) = P(i|\Theta_{n-1}^{Out}). \tag{17}$$

Compare the payoff of player 1 along Θ_{n-1}^{In} and Θ_{n-1}^{Out} respectively:

$$\begin{aligned} P(1|\Theta_{n-1}^{In}) - P(1|\Theta_{n-1}^{Out}) &= \left(V(m'+2) - \sum_{i \in M'(\Theta_{n-2}) \cup \{n\}} P(i|\Theta_{n-1}^{In}) - P(n-1|\Theta_{n-1}^{In}) \right) \\ &\quad - \left(V(m'+1) - \sum_{i \in M'(\Theta_{n-2}) \cup \{n\}} P(i|\Theta_{n-1}^{Out}) \right), \end{aligned}$$

or equivalently,

$$\begin{aligned} P(1|\Theta_{n-1}^{In}) - P(1|\Theta_{n-1}^{Out}) &= \left(V(m'+2) - P(n|\Theta_{n-1}^{In}) - P(n-1|\Theta_{n-1}^{In}) \right) \\ &\quad - \left(V(m'+1) - P(n|\Theta_{n-1}^{Out}) \right), \end{aligned} \tag{18}$$

where the last equality follows from (17). Substituting (16) into (18) we get

$$\begin{aligned} P(1|\Theta_{n-1}^{In}) - P(1|\Theta_{n-1}^{Out}) &= V(m'+2) - P(n|\Theta_{n-1}^{In}) \\ &\quad - \left(V(m'+1) - P(n|\Theta_{n-1}^{Out}) \right) - P(n-1|\Theta_{n-1}^{Out}). \end{aligned}$$

There are $m'+2$ club members along the game tree branch Θ_{n-1}^{In}, so $P(n|\Theta_{n-1}^{In})$ can be found from (10) setting $m = m'+1$. Similarly, $m'+1$ players join the club along the path Θ_{n-1}^{Out} and (10) for $m = m'$ determines $P(n|\Theta_{n-1}^{Out})$. Finally, $P(n-1|\Theta_{n-1}^{Out})$ is the outsider's payoff when the club consists of $m'+1$ members. Using these observations as well as relations (14) and (15) we obtain

$$P(1|\Theta_{n-1}^{In}) - P(1|\Theta_{n-1}^{Out}) = -\beta^2 < 0. \tag{19}$$

■

We make two observations before proceeding to the next step of the argument. First, for a given parameters set, the payoff of player n is fully determined by the number of players who join the club before she gets to choose, and not affected by their identity. Second, Lemma 2 holds independently of the number of players who already agreed to join the club prior to the node X' being reached. These two observations together with backward induction allow us to extend the statement of Lemma 2 to all remaining players $i = n-2, n-3, ..., 2$.

Lemma 3 *If player* 1 *organizes the club, she does not sign up player* $i = 2, ..., n-2$ *in any subgame-perfect equilibrium.*

Proof. As above, consider a game tree node X'' where player 1 makes an offer to player $n-2$. Lemmas 1 and 2 describe the SPNE continuation of the subgame starting in node X''. As the player $n-1$ does not join the club in any SPNE, she does not impact on the allocation of the payoffs within the club. Thus, we can repeat the argument in the proof of Lemma 2, replacing $n-1$ by $n-2$. As a result, we get an analogy of property (19): Player 1 is better off by not signing in player $n-2$, because

$$P(1|\Theta_{n-2}^{In}) - P(1|\Theta_{n-2}^{Out}) = -\beta^2 < 0. \tag{20}$$

The result for $i = n-3, ..., 2$ is obtained in exactly the same way. ■

An immediate consequence of last three Lemmas is that whenever player 1 organizes the club, the resulting club only consists of two players: player 1 and player n.

Corollary 1 *As long as outsiders have some bargaining power, the grand coalition never forms.*

One might suspect that failure to form the grand coalition entails inefficiency. Is that correct, and if so how severe is the problem? To tackle the question, we proceed to derive the pure-strategy SPNE of the game. So far, we have merely supposed that player 1 chooses to organize the club, but would she do so in equilibrium? Clearly, in making the decision player 1 compares the payoff of being a club organizer (and signing player n only) to the payoff of staying outside and allowing subsequent players to form the club. Let us therefore first derive the subgame perfect equilibrium of the continuation game that follows if player 1 decides not to organize the club.

Again, once the player opts out of the club she does not influence the subsequent distribution of payoff in the club. Thus we can extend the logic of Lemmas 1, 2 and 3 to the case where player k gets to organize the club.

Lemma 4 *If players $i = 1, ..., k-1$ choose not to participate in the production of the public good, and player $k < n$ organizes the club, she only signs in player n in any SPNE.*

It remains to investigate whether anybody but player n will be in the club, and if so who. Consider first the node where player $n-1$ gets her chance to organize a club. If she chooses to organize the club, she becomes the residual claimant in a club of size 2, as players $i = 1, ..., n-2$ have already chosen to be outside, and player n is going to be signed in by Lemma (4). If instead player $n-1$ chooses not to organize the club, she receives the payoff of an outsider of a club of size $m = 1$, as player n obviously would organize the club in this case (the alternative is to have no good produced at all). So player $n-1$ will choose not to organize the club if and only if

$$V(2) - P_1^{Out} \leq P_1^{Out},$$

where P_1^{Out} is the minimum payment required to attract player n. Using (14) and (15) and simplifying, we find that player $n-1$ will pass up the opportunity to organize the club if and only if

$$n^2(1-\beta)^2 - \varepsilon \leq 0. \tag{21}$$

Now move one step backwards – to the node where player $n-2$ gets to organize the club. If she chooses to do so, she becomes the residual claimant of a club of size 2, as players $i = 1, ..., n-3$ already chose to be outside, and only player n will be subsequently signed (by Lemma 4). If she decides to opt out, there are two possible continuation games depending on a parameter values: In the case that (21) holds, the continuation game only has player n joining the club. Thus by opting out player $n-2$ receives a payoff of an outsider of a club of size 1. Note that in this case the choice of player $n-2$ is exactly identical to the choice of player $n-1$, and is also determined by (21). As a result, player $n-2$ chooses to stay outside too. By backward induction the same result holds for players $n-3, ..., 1$. So for these parameter values the SPNE of the game is highly inefficient; the club consist of player n only, and the quality of the club good is merely

$$\underline{q} = 2\left(\beta + (1-\beta)n\right),$$

as given by equation (3).

If instead the parameters induce player $n-1$ to organize the club (so that (21) is violated), by opting out player $n-2$ becomes an outsider of a club of size 2. Thus she opts out if and only if

$$V(2) - P_1^{Out} \leq P_2^{Out}. \tag{22}$$

Using (14) and (15) and simplifying, we conclude that player $n-2$ decides not to organize the club if and only if

$$n^2(1-\beta)^2 - 2\beta^2 - \varepsilon \leq 0. \tag{23}$$

Note that independently of the decision of player $n-2$ the club that follows her organization decision has 2 members: either player $n-2$ and player n, or player $n-1$

and player n. Thus the choice of player $n-3$ in the node where she gets to organize the club is also determined by condition (23). By backward induction, the players $n-4, ..., 1$ face the same choice as player $n-2$, and consequently, make the same decision. So if condition (21) fails but condition (23) holds, it is better to be an outsider of a club of size 2, than to be its organizer. Thus, in the SPNE the club comprises players $n-1$ and n and produces

$$\overline{q} = 2(2\beta + (1-\beta)n),$$

according to equation (3).

If both (21) and (23) fail, each player $i = n-1, n-2, ..., 1$ prefers being a residual claimant in a club of size 2 to being an outsider of a club of the same size. Thus, in this case the club consists of two members, 1 and n, and again produces the quality $\overline{q}$. This completes our characterization.

Proposition 5 *Let $\beta \in (0,1)$. (a) If the outsider's bargaining power is sufficiently high, i.e., if $n^2(1-\beta)^2 - \varepsilon \leq 0$, the club consists of player n only and provides the quality $\underline{q} = 2(\beta + (1-\beta)n)$. (b) For intermediate values of β, satisfying $n^2(1-\beta)^2 - \varepsilon > 0 \geq n^2(1-\beta)^2 - 2\beta^2 - \varepsilon$, the club comprises players $n-1$ and n and provides the quality $\overline{q} = 2(2\beta + (1-\beta)n)$.(c) For lower values of β, the club comprises players 1 and n and provides the quality $\overline{q}$.*

The intuition for why the club becomes so small under sequential negotiations is that with each additional member signed up by the founder it becomes more expensive to sign up subsequent members. A striking fact is that the grand coalition fails to form even when the club's outsiders has very little bargaining power (β is close to zero). The result is not as paradoxical as it first appears, because as β goes to zero the equilibrium quality converges monotonically to the optimal level $2n$. That is, even if the equilibrium club size is small, the efficiency level can be large.

4 Multi-round membership negotiations.

An objection to the single-round membership negotiation model is that it precludes relative losers, notably the last player, from making efficiency-enhancing offers to the relative winners at the club formation stage. What would happen if we added bargaining rounds to the above game? Under some circumstances, efficiency may then be attainable. Suppose in particular that the decision to enter the club is irreversible, whereas the decision not to enter is fully reversible. Then, it can be shown that the unique subgame perfect equilibrium of the game entails full efficiency provided that there are sufficiently many bargaining rounds.

We think that this objection is inadmissible, as it contradicts the idea of voluntariness. By adding rounds to the bargaining game, while insisting that the decision not to enter is only binding over a single negotiation round, the modeller implicitly forces reluctant players to participate in a negotiation that they may prefer to leave. In our view, a more reasonable definition of voluntary bargaining is that the players have the ability to exit the negotiation. In other words, a player ought to have the ability to not make or take offers.

To formalize our argument, we assume that a modified version of the above sequential game is repeated a finite number of times K. Each player has K options to move and within each round the moves are always in the same sequence. As above, at each stage, if no club was created before, player i has a choice between organizing the club or staying outside. Since the game is repeated, we need to make explicit our commitment assumptions. We assume that the decision to organize or to join the club is binding, whereas the decision to stay outside may or may not be binding. If player i who chooses to stay outside at time t receives another offer later in the game, or chooses to organize a club (in case it is not created yet), she is free to do so. At the same time, players cannot be forced to sit at the bargaining table. That is, at her move, in addition to the options of creating the club and staying outside, an uncommitted player i can choose to exit the bargaining process until after date K, by which time the good has to be provided.[9] We refer to the three possible actions for an uncommitted player as "in", "wait", and "exit". Observe that we do not allow players to make conditional commitments to future choices. For example, it is impossible to commit to exit at a future date unless some prerequisite has been fulfilled by that date. Any promise or threat is credible only to the extent that it is actually desirable to carry it out when the time arrives.[10] Moreover, the option to leave the negotiation table may also be available to the club's founder and not only to presently uncommitted players. However, it is easily

[9] The assumption that there is a specific provision date greatly simplifies the analysis.

[10] In our introductory example, efficiency can be attained by letting family B commit to only take part in provision if A also does so. The problem is that carrying out such a commitment is typically not credible once it is clear that family A will not be providing. (This is not to say that conditional commitment is always impossible. In some cases one party has a strong reputation or can rely on other sources of authority. For a general analysis of which outcomes can be attained under this assumption, see Segal,1999.)

seen that the founder will not ever want to exit negotiations, either before or after date K. If the founder exits the negotiation before interacting with all players, she loses some potentially valuable members; if instead she exits after date K but before having negotiated with all outsiders, she loses revenues from selling access.

Except for the distinction between waiting and exiting, which could not be made in a single-round framework, everything is as before: The organizer of the club makes unconditional monetary offers to the potential members, and once the club is created (that is, either the final period of the game is over or every player committed to join the club or to exit), the organizer bargains with any remaining outsiders over the terms of access.

As it turns out, the case of $K = 2$ rounds contains virtually all the insights of the general case. Moreover, the final outcome in the $K = 2$ case is identical to the $K = 1$ case.

Lemma 5 *If player* 1 *chooses to organize the club at her first move, then players* $2, ..., n-1$ *choose to exit in the first round of the game and player* n *joins the club. Thus, the resulting club consist of players* 1 *and* n *only.*

Proof. Consider the choice of player n at her first opportunity to move, that is, at the end of the first round. First, suppose that by this time every player $2, ..., n-1$ already either joined the club or exited. Assume that the number of club members is given by m. Then if player n chooses to exit, nobody joins the club in the next period and thus player n becomes an outsider of a club of size m with the payoff P_m^{Out}. If player n waits, she gets signed in the second round of the game for the payoff of P_m^{Out} and the resulting club size is $m + 1$. Thus, as above, player 1 signs in player n for the payoff of P_m^{Out} and the resulting club has the size $m + 1$.

Suppose now that a set of players $W \subseteq \{2, ..., n-1\}$ waited in the first round. Let $j = \max W$. Again, assume that the current club has m members. If player n exits, by Lemma 1 player j gets signed up in the second round of the game for the payoff P_m^{Out}. Thus the resulting club consists of $m + 1$ members and, being an outsider, player n receives P_{m+1}^{Out}. If player n waits instead, she will get signed up in the second round of the game, receiving $P_m^{Out} < P_{m+1}^{Out}$. (We know from the single-round game that neither player j not any other player who waited will be signed up in the last round.) That is, for player n in round 1 waiting is strictly dominated by exiting. Thus, to sign player n at the end of the first round costs P_{m+1}^{Out}. If player 1 chooses to do so, she continues by signing player j in the second round for the payoff P_{m+1}^{Out}. The resulting club consists of $m+2$ members. Therefore, the difference between the payoffs of player 1 when she signs player n in the first round and when she does not is given by

$$\left[V(m+2) - P_{m+1}^{Out} - P_{m+1}^{Out}\right] - \left[V(m+1) - P_{m}^{Out}\right] = -\beta^2 < 0. \tag{24}$$

Thus, in this case player 1 does not sign player n in the first round, and player n chooses to exit. Player j gets signed in the second round and the resulting club again consists of $m+1$ members.

We proceed to the choice of player $n-1$ in the first round of the game. Suppose m players joined the club before player $n-1$ gets her first move. If she exits, we have just shown that in the continuation of the game player 1 signs in one more player for the payoff of P_m^{Out}. The resulting club will consist of $m+1$ members and player $n-1$ receives P_{m+1}^{Out}.

If player $n-1$ waits, then player n exits in the first round and player $n-1$ gets signed in the second round for the payoff $P_m^{Out} < P_{m+1}^{Out}$, so player $n-1$ never chooses this strategy. Thus, to sign in player $n-1$ in the first round player 1 has to pay P_{m+1}^{Out}. If she signs player $n-1$, the club expands to $m+1$. As we shown above, player 1 then proceeds by signing in one more player for the payoff of the outsider of the club of size $m+1$, P_{m+1}^{Out}. Therefore, the net benefit to player 1 from signing player $n-1$, is again given by expression (24). We see that it is too costly for the player 1 to sign player $n-1$ in the first round. Thus, player 1 does not sign in player $n-1$, and player $n-1$ chooses to exit.

By backward induction the same result holds for players $n-2, ..., 2$. To summarize, along the branch where player 1 organizes the club, players $2, ..., n-1$ exit at the first round, and player n joins the club. ■

The result continues to hold when the club is organized by player i, $1 < i < n$.

Lemma 6 *If player i, $1 < i < n$, chooses to organize the club at her first move, the resulting club comprises only player i and player n. Players $1, ..., i-1, i+1, ..., n-1$ exit in the first negotiation round.*

Proof. This result follows from two observations. First, it can be proven exactly along the lines of the Lemma 5, that players $i+1, ..., n-1$ choose to exit.

Suppose that each player $1, ..., i-1$ has chosen to exit too. Then we showed above that player n joins the club, so that the resulting club consists of player i and player n.

Suppose now that a set of players $W' \subseteq \{1, ..., i-1\}$ waited in the first round. Let $k = \max W'$. Then player n exits in the first round, and the second round of the game results in the club consisting of player i and player k. Therefore, player k receives the payoff of an outsider of the club of size 1. But by choosing to exit, she would get the payoff of an outsider of the club of size 2. Thus, in equilibrium player k never prefers waiting to exiting, which contradicts the definition of player k. ■

The situation may potentially be different when player n gets her first chance to organize the club. To analyze the resulting equilibria, we consider three cases depending on the parameter values (n, β, ε). These three cases are motivated by the one-round game of the previous section. Denote by P_2^{Org} the payoff of the organizer of club of size 2,

$$P_2^{Org} = V(2) - P_1^{Out},$$

where P_i^{Out} denotes the payoff of an outside of a club of size i, recall formula (15).

Case 1.

Assume that the parameters n, β and ε satisfy the inequality

$$n^2(1-\beta)^2 - 2\beta^2 - \varepsilon > 0. \tag{25}$$

This case corresponds to the equilibrium of the one-round game when the club consists of players 1 and n. As we have seen above, in this equilibrium the payoff to an outsider of club of size 2 is below the payoff to the organizer of the club of size 2,

$$P_2^{Org} > P_2^{Out}. \tag{26}$$

There are three possibilities: First, assume that every player $1, ..., n-1$ has chosen to exit. Then in equilibrium player n chooses to be the single provider of the public good, and players $1, ..., n-1$ get the payoff of an outsider of club of size 1.

Second, suppose that only one player $h \in \{1, ..., n-1\}$ waited. If player n chooses to organize the club, she signs in player h, creates the club of size 2 and gets P_2^{Org}. If she chooses not to organize, but does not exit, she gets signed in by player h in the second round for the payoff P_1^{Out} (as the alternative would be to be an outsider of club of size 1). If she exits, she also receives P_1^{Out}, which is less than P_2^{Out}. By assumption (26) player n chooses to organize the club in the very first round, if she gets this opportunity.

Third, assume that there are more than one player $1 \leq h_1, ..., h_j < n$ that have waited (and the remaining players exited). The only difference with the previous case appears when player n chooses to exit. In this case by the argument of the previous section the remaining players organize a club of size 2 in the second round and player n receives P_2^{Out}. But due to the assumption (26) this option is dominated by her preferred choice - to organize the club.

Now let us return to the options of the player $n-1$. If she does not organize the club, it follows from the above discussion that her payoff is at most that of an outsider of a club of size 2. If she organizes the club (of 2 members), she receives $P_2^{Org} > P_2^{Out}$ by assumption (26). Thus, player $n-1$ would choose to organize the club in the first round, if ever given this opportunity. By the same logic we see that for this set of parameters the equilibrium is characterized by player 1 organizing the club of size 2 at the very first round, players $2, ..., n-1$ immediately exiting and player n joining the club. That is, for this set of parameters the outcome of the two-rounds game exactly replicates the outcome of the one-round game.

Case 2.

Assume now instead that

$$n^2(1-\beta)^2 - \varepsilon > 0 \geq n^2(1-\beta)^2 - 2\beta^2 - \varepsilon. \tag{27}$$

In this case the payoff of an outsider of a club of size 2 is higher than the payoff of the organizer of the club of size 2, which is in turn higher than the payoff of an outsider of the club of size 1 :

$$P_2^{Out} > P_2^{Org} > P_1^{Out}. \tag{28}$$

In the equilibrium of the corresponding one-stage game, the club consists of players $n-1$ and n.

Again, we consider three possibilities: If every player $1, ..., n-1$ has chosen to exit, then in equilibrium player n chooses to be the single provider of the public good. Then players $1, ..., n-1$ get the payoff P_1^{Out}.

Suppose now that only one player $h \in \{1, ..., n-1\}$ waited. If player n organizes the club, she signs in player h into the club of size 2 and gets the payoff P_2^{Org}. If n chooses to wait, she gets signed by h in the second round and receives the payoff of an outsider of the club of size 1. If n exits, she also receives P_1^{Out}. Also in this case, by assumption (28) player n chooses to organize the club in the very first round if she gets this opportunity.

The difference with the previous case arises when there are more than one players $1 \leq h_1 < ... < h_j < n$ that have waited (while the other players exited). If player n

chooses to exit, players h_{j-1} and h_j organize the club of size 2 in the second round. Therefore player n receives P_2^{Out}, which, by assumption (28) exceeds her payoff under the other two options. Thus, player n exits in this case. Now let us move backwards. In round 2, player h_j joins the club and receives no more than the maximum of the payoff of the organizer of club of size 1 (if she is the only one to join) and that of an outsider of the club of size 1 (if she is the last one to join the club of size 2). Suppose that instead of waiting, player h_j exits in the first round before player n gets to move. As there was at least one more person h_1 who has chosen not to organize the club and not to exit, player h_j receives P_2^{Out}. By assumption (28) she thus exits in round 1, which contradicts the definition of h_j. More generally, in equilibrium no player $i = 1, ..., n-1$ chooses to wait if any of her predecessors waited.

Now consider the options of the player $n-1$. If everyone exited before, she chooses to organize the club, as this brings her P_2^{Org}, while alternative options bring her P_1^{Out}. If instead there is a player $h_1 < n-1$ who already has chosen to wait, we are back to our discussion of the incentives of the player h_j for $h_j = n-1$. Thus in this case player $n-1$ exits and receives P_2^{Out}.

Let us proceed backwards. If player $n-2$ exits, she gets the payoff P_2^{Out}. If she organizes the club, she signs up player n and receives the payoff P_2^{Org}, which is below the payoff of the exit option. Finally, if she waited (which she would only do if all players $1, ..., n-3$ exited), she gets signed in the club by player n and receives the payoff P_1^{Out}. Thus, player $n-2$ chooses to exit. By backward induction, so do the players $n-3, ..., 1$. Thus, in the resulting equilibrium players $1, ..., n-2$ exit and player $n-1$ organizes the club of size 2 by signing player n already in the first round. Again, this equilibrium outcome exactly replicates that of the single-round game.

Case 3.

Finally, assume that

$$n^2(1-\beta)^2 - \varepsilon \leq 0. \tag{29}$$

In the single-round game, the club in this case consists of player n only, and the payoff of the outsider of club of size 1 exceeds the payoff of the organizer of the club of size 2,i.e.,

$$P_2^{Out} > P_1^{Out} > P_2^{Org}. \tag{30}$$

Thus, no player wants to be the club organizer, as this brings the lowest payoff.

If every player $1, ..., n-1$ has chosen to exit, then in equilibrium player n chooses to be the single provider of the public good. Then players $1, ..., n-1$ get the payoff

P_1^{Out}.

If there is any player $h \in \{1, ..., n-1\}$ that has chosen to wait, player n chooses to exit. By assumption (30) she does not want to organize the club. If she stays, the SPNE of the one-stage game will force her to do so in the second round.

Thus, staying outside without exit entails a loss of the first mover advantage – if player $h < n$ has chosen to stay, every player $h+1, ..., n$ immediately exits. On the other hand, an early exit insures the player the payoff of an outsider of the club of size 1. In the respective equilibrium players $1, ..., n-1$ exit in the first round of the game and player n organizes the club of size 1.

So we demonstrated that in the SPNE of the 2-rounds game every player makes a binding decision already by the end of the round 1 and the respective equilibrium outcomes exactly replicate the ones of the one-round game. By backward induction, the same is true for a game with any finite number of rounds K.

Proposition 6 *In the multi-round sequential game with a finite number of rounds, the grand coalition never forms if $n > 2$. If $n = 2$, the grand coalition may or may not form, depending on parameters. In either case the equilibrium size and composition of the club, as well as the club good provision level, replicate the outcomes of the one-stage game.*

In some respects, Case 3 above is perhaps the most important. It identifies parameters such that the grand coalition fails to form even when $n = 2$. Since the scope for varying the extensive form is quite limited in bilateral bargaining games, we are confident that the two-player counterexamples to the Coase Theorem are highly robust. For the sake of concreteness, suppose that $n = 2$, $\beta = 1/2$, and $\varepsilon \in (1, 9/4)$, satisfying the inequality that defines Case 3. (Moreover, even though $\varepsilon > 1$, it is easy to check that the solution remains interior.) For all our three extensive forms, the only equilibrium outcome is for one of the players to provide the club good and invest in quality $q = 3$, instead of the optimum quality $q = 4$; both numbers follow from equation (3). Under the chosen parameters, we believe that this is the unique kind of equilibrium outcome of any bargaining game with voluntary participation: Both players cannot be better off as joint providers than as lone outsiders, and each player prefers being the lone provider to having no provision.

5 Conclusion

Dixit and Olson (2000) asked whether voluntary participation undermines the Coase Theorem. We find that the answer is affirmative. Private property rights in combination with costless bargaining fail to preclude strategic non-membership in provision clubs, because ownership alone does not suffice to extract all the surplus arising from future trade with non-members.

References

[1] Anderlini, Luca and Leonardo Felli, 2006, Transaction Costs and the Robustness of the Coase Theorem, Economic Journal 116, 223-245.

[2] Coase, Ronald, 1960, The Problem of Social Cost, Journal of Law and Economics 3, 1-44.

[3] Dixit Avinash K. and Mancur Olson, 2000, Does Voluntary Participation Undermine the Coase Theorem? Journal of Public Economics 76, 309-335.

[4] Hart, Oliver, 1995, Firms, Contracts, and Financial Structure, Oxford: Clarendon Press.

[5] Mas-Colell, Andreu, Michael D. Winston, and Jerry R. Green, 1995, Microeconomic Theory, Oxford: Oxford University Press.

[6] Myerson, Roger B. and Mark A. Satterthwaite, 1983, Efficient Mechanisms for Bilateral Trading, Journal of Economic Theory 28, 265-281.

[7] Segal, Ilya, 1999, Contracting with Externalities, Quarterly Journal of Economics 114, 337-388.

A Appendix

A.1 Proof of Proposition (1).

Using (4) and (5) equation (6) takes the form

$$2\beta\left(\beta m+(1-\beta)n\right)>\frac{\left(\beta\left(m+1\right)+(1-\beta)n\right)^{2}-\varepsilon}{m+1}. \tag{31}$$

Solving for m yields the following restriction on the size of the club m:

$$m>\frac{\left(n^{2}\beta^{2}+\beta^{2}-2n^{2}\beta+n^{2}-\varepsilon\right)^{1/2}}{\beta} \tag{32}$$

Similarly, (7) can be rewritten as

$$\frac{\left(\beta m+(1-\beta)n\right)^{2}-\varepsilon}{m}\geq 2\beta\left(\beta\left(m-1\right)+(1-\beta)n\right), \tag{33}$$

which, in turn, is equivalent to

$$m\leq 1+\frac{\left(n^{2}\beta^{2}+\beta^{2}-2n^{2}\beta+n^{2}-\varepsilon\right)^{1/2}}{\beta}. \tag{34}$$

Recognizing that the equilibrium club size m^* is a natural number between 1 and n, we have

$$m^{*}=\min\left[n,\text{integer}\left(\frac{\left(n^{2}\left(1-\beta\right)^{2}+\beta^{2}-\varepsilon\right)^{1/2}}{\beta}\right)+1\right].$$

A.2 Proof of Proposition (2).

Denote

$$\widetilde{m}=\frac{\left(n^{2}\left(1-\beta\right)^{2}+\beta^{2}-\varepsilon\right)^{1/2}}{\beta^{2}},$$

so that

$$m^{*}=\min\left[n,\text{integer}\,\widetilde{m}+1\right].$$

Note that if β is sufficiently close to 0, then $\widetilde{m}\to\infty$ and $m^*=n$. Otherwise,

$$\frac{\partial\widetilde{m}}{\partial n}=\frac{n\left(\frac{1}{\beta}-1\right)^{2}}{\left(n^{2}\left(\frac{1}{\beta}-1\right)^{2}+1-\frac{\varepsilon}{\beta^{2}}\right)^{1/2}}>0,$$

and thus m^* also (weakly) increases in n, the actual increase occurring when $\widetilde{m}\in\mathbb{N}$. According to formula (3)

$$q^{*}=2\left(\beta m^{*}+(1-\beta)n\right),$$

which immediately implies that q^* also increases in n.

A.3 Proof of Proposition (3).

If β is sufficiently close to 0, then $\widetilde{m} \to \infty$ and $m^* = n$, which does not change with β. Otherwise,

$$\frac{\partial \widetilde{m}}{\partial \beta} = -\frac{\left(n^2 (1-\beta) - \varepsilon\right)}{\beta^2 \left(n^2 (1-\beta)^2 + \beta^2 - \varepsilon\right)^{1/2}}.$$

Note that whenever

$$n^2 (1-\beta) - \varepsilon > 0,$$

$\frac{\partial \widetilde{m}}{\partial \beta} < 0$ and thus m^* also (weakly) increases in β. Now consider the case when

$$n^2 (1-\beta) - \varepsilon \leq 0.$$

Then,

$$\widetilde{m} = \frac{\left(n^2 (1-\beta)^2 + \beta^2 - \varepsilon\right)^{1/2}}{\beta} < \left(\frac{\beta - \varepsilon}{\beta}\right)^{1/2} < 1.$$

Thus for all such $\widetilde{m}$, $m^* = 1$ and does not change with β[11]. We conclude that m^* always (weakly) increases in β.

Rewriting (3) we get

$$q^* = 2\left(\beta m^* + (1-\beta) n\right) = 2\left(n - \beta\left(n - m^*\right)\right). \tag{35}$$

As m^* (weakly) decreases in β, so does the term $-\beta\left(n - m^*\right)$ and thus the entire expression (35).

[11]Note that this situation, where $\frac{\partial \widetilde{n}}{\partial \beta} \geq 0$, is a result of the discrete setting. In a continious-time setting is does not appear.

CHAPTER 3

Club-in-the-Club: Reform under Unanimity*

with Erik Berglöf, Mike Burkart and Guido Friebel

Abstract

In many organizations, decisions are taken by unanimity giving each member veto power. We analyze a model of an organization in which members with heterogenous productivity privately contribute to a common good. Under unanimity, the least efficient member imposes her preferred effort choice on the entire organization. In the presence of externalities and an incomplete charter, the threat of forming an "inner organization" can undermine the veto power of the less efficient members and coerce them to exert more effort. We identify the conditions under which the threat of forming an inner organization is never executed, and under which inner organizations are equilibrium outcomes. Finally, the members' welfare can be both higher or lower under the simple majority rule than under unanimity, thereby providing a rationale for the emergence and diversity of decision rules.

1 Introduction

Coordination of individual actions is the core problem that a society must solve to assure the well-being of its members. Much of economics focuses on the markets; arguably, organizations are an equally important coordination mechanism. When studying organizations, economists typically presuppose the existence of a governance system

*We have benefited from comments by Tore Ellingsen, Chloé Le Coq, Sergei Guriev and seminar participants in Berkeley, Moscow (CEFIR), Munich, Oxford, Stockholm (SITE) and Toulouse(IDEI). Financial support from the Bank of Sweden Tercentenary Foundation (Burkart) and Jan Wallander's and Tom Hedelius Research Foundation (Paltseva) is gratefully acknowledged. All remaining errors are our own.

consisting of rules, penalties or transfers. Club theory, for instance, assumes that there is a system of transfers, taxes and entry fees that can be used to make members of a club behave in a common interest (e.g., Cornes and Sandler, 1996). Similarly, organization theory presupposes the existence of a principal who coordinates the members of an organization through the use of a host of monetary and non-monetary instruments (e.g., Milgrom and Roberts, 1992).

In many circumstances, cooperation and organizations exist in the absence of a comprehensive governance system. Rather, agents sharing a common form a "loosely knit" group. For example, sovereign states may come together to coordinate their action in specific areas, such as economic policies, protection of the environment or defence. But a priori, there is no structure in place governing how decisions are taken to pursue the common goal. The states must first sort out how to decide. Similarly, associations, political parties, NGOs all share the feature that in early stages of their existence their governance structure is highly incomplete or non-existing. In all these cases, there is initially no other decision rule but unanimity, as pointed out by Rousseau.[1]

Unanimity grants each member of an organization a veto right, thereby protecting her against coercion or what de Tocqueville (1835) called the "tyranny of the majority". But the flipside of unanimity is slow and inflexible decision-making and underprovision of the common good. Heterogeneity is key here: Members who are less committed or less productive can veto any proposal to increase contributions (effort) to the common good. The problem of holding back other more productive members becomes particularly severe when there are complementarities between the members' contributions. In the presence of such "weakest-link" effects, a member who invests little limits the amount of the common good for the entire organization.

This paper argues that organizations operating under the unanimity rule can provide more public goods than preferred by their least committed member. The mechanism that can overcome the veto power of least committed members is the threat of forming a "club-in-the-club". Suppose more committed members can form an inner club whose members cooperate more intensively. When staying outside is costly, the possible formation of an inner club constitutes a threat. In response, the least committed members are willing to contribute more to the public good in order to undo

[1] "Indeed, if there were no prior convention, where, unless the election were unanimous, would be the obligation on the minority to submit to the choice of the majority? How have a hundred men who wish for a master the right to vote on behalf of ten who do not? The law of majority voting is itself something established by convention, and presupposes unanimity, on one occasion at least." Rousseau, 1672.

the attractiveness of a club-in-the-club. Conversely, more committed members may then prefer not to execute the threat. They may be better off having single membership in the initial club at an increased effort, compared to having dual membership in both the initial and the inner club. Thus, unanimity does not necessarily constrain an organization to the pace of its least committed members.

To develop our argument we analyse the provision of a public good by an organization in which decisions are taken by unanimity. The club good is produced through a Leontief technology with each member's effort as the inputs. Though effort should be broadly interpreted as any costly contribution to a common good. The members differ in terms of their effort cost. We exclude governance or transfer mechanism that can induce club members to exert more effort than they individually prefer. Given each member has a veto power, one would expect that public good provision is determined by the "weakest-link", i.e., the member with the highest cost of effort.

In this setting, we show that the mere possibility of forming a club-in-the-club, to which only high-effort providers have access, can increase the amount of club good provided by the entire organization. Weaker members increase their effort in order to avoid that the inner organization forms. Thus, the threat of forming a club-in-the-club limits the leverage of the veto power for less committed members, and unanimity does not necessarily lead to stagnation. Despite their veto powers, weaker members exert more effort than they individually prefer.

Alternative reasons why less committed members may refrain from vetoing like reputation or log-rolling rely on repeated interaction. By contrast, our mechanism functions in a static model with heterogeneous members. The threat of forming a club-in-the-club has, however, bite only if two conditions hold. First, the statutes of the organization must not exclude the proposal to form an inner organization, because otherwise less productive members could veto it. Hence, constitutional incompleteness that makes it hard to incentivise and coordinate members in the first place, can help to reform the club through the threat of inner club. Second, in order to constitute a threat, forming a club-in-the-club must involve deadweight losses, i.e., the value from belonging to the initial club must decrease. It is important that this affects each member of the initial club.

We show that the threat of forming an inner organization may also be executed in equilibrium. The inner club is only formed when the original club is sufficiently heterogeneous, and the deadweight loss of forming an inner club is sufficiently small. Put differently, the possibility of clubs-in-clubs may lead to more integration if the off-equilibrium threat makes all members increase their effort, but also to less integration if the threat is executed in equilibrium.

Given the drawbacks of the unanimity rule, the question arises whether organizations would fare better under an alternative decision rule. We compare the welfare implications of the simple majority and unanimity rules and show that neither of them dominates. Our analysis suggests that the relative productivities of the median vs. average organization member are the crucial determinants for the welfare ranking. More specifically, when these types are relatively similar, majority rule is superior. When the median member is much more productive, unanimity dominates because the "tyranny of the majority" entails substantial costs for weak members. This explains observed diversity of decision rules across organizations as well as across subject matters within a given organization. In addition, it provides a rationale why members of an organization may voluntarily renounce their veto right and subject themselves to the will of future majorities.

Throughout the paper, we will illustrate our theory by using the example of European integration. Since the foundation of the European Union (EU), or at least since its first enlargement, the risk of the formation of an inner core has exerted considerable pressure on member states to further integrate despite veto rights. For instance, the EU appeared incapable of reforming prior to the 1992 Maastricht Treaty. Then, the economically strongest members proposed a reform, the introduction of a common currency. As in our model, the threat of inner club open only to countries that meet the Maastricht criteria led to wide-spread reform efforts, in particular in economically weaker Euro candidate countries. Recently, however, efforts to further deepen European integration have slowed down, and most observers attribute this to the increased heterogeneity following the Eastern Enlargement. In our model, higher heterogeneity goes together with an increasing risk of a club-in-the club, a risk that is reflected in the current discussions to create a two-speed Europe. Finally, this risk is also accompanied by efforts to alter the decision-making process in EU institutions. Although the attempt to install a Constitution failed, we witness a general shift from unanimity decisions to (qualified) majorities. We believe that our theory can account for important aspects of the constitutional dynamics in Europe and in other international organizations.

The threat of an inner organization and its possible execution parallel arguments put forward in the literature on secession. Our mechanism is similar to the one in Buchanan and Faith (1987) on "internal exit" as an alternative to "voting with your feet". They derive an optimal tax rate as the tax rate that maximizes tax revenue without triggering secession. In Bolton and Roland (1997) secession reduces private income as a result of lost agglomeration economies or efficiencies, but avoids the "tyranny of the majority" by creating more homogenous political entities.[2] The focus of our paper is, however, not so much on secessions – that is, the complete separation of federations – but rather on the creation of costly internal structures. Further, we emphasize the potential function of the inner organization threat as a mechanism to discipline less committed members. This contrasts with Gradstein (2004) who argues that secession rights, while protecting minority rights, involve inefficiencies in bargaining processes. In our model, internal threats can increase efficiency (because they can induce higher effort), or decrease it (as the formation of an inner club entails a deadweight loss). An additional distinction is that the above papers consider majority voting, while we concentrate on the unanimity rule, reflecting organizations with highly incomplete constitutions.

Our model predicts that organization members may benefit from foregoing veto power and subjecting themselves to the will of the majority. This result is related to a growing literature analyzing how constitutions form, in particular, what determines the voting rules of a society. Aghion and Bolton (2003) identify a trade-off between minority protection and flexibility. To adapt to changes, a society must offer transfers to some individuals to prevent them from exercising their veto right. Hence, a society may under the veil of ignorance decide to replace unanimity by some type of majority voting. Messner and Polborn (2004) take a complementary view and show why societies may opt for supermajorities rather than simple majority. In their model, young people, who vote today over tomorrow's decision rule, anticipate that they will benefit less from reforms when they are old. Hence, they want to have more power about future reforms, which gives them an incentive to agree on a supermajority rule. Erlenmaier and Gersbach (2004) argue that first best outcomes can be achieved under unanimity, provided that it is supplemented by a number of constitutional provisions, such as bundling of projects. Compared to all these papers, our model is more parsimonious, in particular excluding side payments. In addition, we focus on the effects of inner group formation on the efficiency of an organization in the absence of constitutional rules, i.e., under voluntary cooperation.

[2]Bordignon and Brusco (2001) point out that constitutionally defined secession rights involve a trade-off. While these rights reduce the cost of an actual break-up ex post, they increase the likelihood of this outcome.

Harstad (2006) compares flexible cooperation where organization members can decide on the speed of integration to rigid cooperation where all members go at the same speed. While addressing similar issues, his model does not consider the disciplining role that a threat of an inner club has on weak member. In Dixit (2003) this role is played by network externalities. Due to these externalities, agents may sequentially adopt an innovation (or join an organization) even if it would be in the collective interest that the innovation were not introduced. That is, adoption is the individually rational, unless agents can coordinate their actions. In our model, weaker members are in a similar situation – they would prefer the threat of forming an inner organization not to exist. In addition, stronger members can execute the threat and form an inner organization, a possibility not explored by Dixit.

The paper is organized as follows. Section 2 outlines and solves the basic model which abstracts from the possibility of inner clubs. Section 3 introduces this possibility and examines the impact that threat of an inner club has on the initial organization. Section 4 derives the conditions under which an inner club is formed and characterizes the equilibrium outcomes. Section 5 discusses key assumptions. Section 6 compares the organizations operating under unanimity and majority rule. Section 7 concludes. All mathematical proofs are relegated to the Appendix.

2 The Curse of Unanimity

We consider an organization with N members, who produce a club good.[3] The provision of the good increases in the size of the organization and in the effort e of the members. Inspired by Leontief partnership models (e.g., Vislie, 1994), we assume that the amount of club good is determined by the smallest effort in the organization, scaled by the size of the organization: $N \min[e_1, e_2, ...e_N]$.

The utility of each member increases in the consumption of the club good and decreases in the amount of effort provided. The benefit from consumption is the same for all members, whereas the effort cost differs across members. Member $i \in N$ has effort cost $\theta_i e^2/2$, and the type parameter θ_i follows an equidistant distribution with support $[\underline{\theta}, \bar{\theta}]$. Occasionally, we refer to $\underline{\theta}$ as the most productive or "strongest" type, and to $\bar{\theta}$ as the least productive or "weakest" type and use "stronger" and "weaker"

[3]Our interest is how an existing organization responds to new challenges for which its members have different preferences. We, hence, abstract from the question of whether any given member has an incentive to leave the organization or whether outsiders would like to join.

when comparing types. Assigning rank 1 to the strongest type $\underline{\theta}$, the cost parameter of the member with rank i is

$$\theta_i = \underline{\theta} + \frac{i-1}{N-1}\left(\bar{\theta} - \underline{\theta}\right)$$

and the productivity difference between two adjacent members is constant and equal to

$$\theta_i - \theta_{i-1} \equiv \frac{1}{N-1}\left(\bar{\theta} - \underline{\theta}\right).$$

Given that the public good is produced with a Leontief technology, member i's payoff is

$$y(\theta_i, e) = N \min\{e_1, ..e_N\} - \theta_i e_i^2/2.$$

As the members have different costs, their preferred amount of public good differs. Hence, some members could offer side payments to others in order to influence their effort choices. Since we focus on the threat of forming a "club-in-the-club" as a mechanism to overcome the opposition of individual members against reform proposals, we deliberately abstract from transfer payments.

We model the production of the club good as a two-stage game. In the first stage, members vote on a minimum effort level in the club and in the second stage each member exerts an effort.[4] Individual effort levels are verifiable and each member commits herself to exert - at least - the effort level agreed upon in the voting stage. That is, underprovision is infinitely punished. At the same time, the voting outcome is not binding from above, thereby allowing for the opportunity of joint welfare improvement. This asymmetry reflects our interest in the constraints that unanimity imposes on organizations. However, unilateral overprovision is never an equilibrium outcome due to Leontief technology.

The voting procedure parallels that of the continuous-time ascending-bid auction (Milgrom and Weber, 1982):[5] An uninterested agent ("auctioneer") proposes a sequence of continuously increasing effort levels $\{e\}$ starting with the initial level $e = 0$. After each proposal agents decide whether or not to vote in favour of a further increase in the common effort level. Once a member "leaves the auction" by voting against an increase,

[4] Given the Leontief technology, exerting more effort than other members is a pure waste. Hence, there is no loss of generality in assuming that the organization votes on a common (minimum) effort level as opposed to a menu of type-contingent efforts.

[5] In standard voting procedure agents vote over pairs of alternatives and the winner in one round is posed against a new alternative in the next round. Under the unanimity rule, this procedure may easily fail to generate a unique winner. The ascending procedure is motivated by our focus on the impact of the weakest member on the club production.

she cannot "return" by supporting any subsequent proposals. Given the majority rule m the voting continues as long as at least share m of members vote for an increase in effort. The "winning" effort level is the one at which the voting stops. Under the unanimity rule ($m = 1$), the voting stops once a single member votes against a further increase in effort. Accordingly, the option to withdraw from the voting confers a veto power on each member. After the voting stage, members simultaneously choose their effort and the public good is produced.

In this game, any outcome in which all members exert some common effort $e \in [0, N/\bar{\theta}]$ can be supported as a Nash equilibrium, where $N/\bar{\theta}$ is the effort maximizing the payoff of the weakest type. Efforts above $N/\bar{\theta}$ are never an equilibrium outcome. This is due to the Leontief technology and the punishment for underperforming, making unilateral deviations non-profitable.

More precisely, any effort $e \in [0, N/\bar{\theta}]$ can be supported in an equilibrium where at least two members withdraw from the vote at some $e^V \in [0, e]$ and where all members choose in the production stage the same effort level e. Indeed, given some e^V, member i's decision problem at the implementation stage is

$$\max_{e_i \geq e^V} \left(N \min\{e, e_i\} - \theta_i e_i^2/2\right).$$

Member i's best preferred choice $e_i^* = N/\theta_i$ exceeds e, as $e \leq N/\bar{\theta} \leq N/\theta_i$. Thus, member i always chooses $e_i = e$, since any effort $e_i - e > 0$ would be wasted. That is, unilateral overperformance ($e_i > e$) is never profitable. In this case, the voting outcome e^V needs not to be binding as all members can agree on exerting higher effort $e \geq e^V$.

At the voting stage the only deviation for member i, that influences the outcome of the game, is to withdraw prior to e^V. This deviation is profitable iff $e^V > N/\theta_i$. In fact, by withdrawing at $e_i^V \leq N/\theta_i$ and choosing $e_i^* = N/\theta_i$ in the implementation stage member i attains her first best. Since this applies to all members $i = 1, .., N$,

$$e^V \leq N/\bar{\theta} \tag{1}$$

must hold in equilibrium. Consequently, any effort $e > N/\bar{\theta}$ cannot be an equilibrium outcome. Indeed, if everyone but the member N chooses e, member N's unilateral underperformance ($e_N^* = N/\bar{\theta} < e$) is both profitable and compatible with the voting outcome as $N/\bar{\theta} \geq e^V$ by (1).

It is well known that input games for a team with a Leontief technology have a continuum of Nash equilibria and that these equilibria can be Pareto-ranked.[6] This also holds for our voting game: all members prefer the Pareto-dominant equilibrium with $e = N/\bar{\theta}$ which we use as a benchmark in the subsequent analysis.

Thus, our framework generates an outcome that is in line with the widespread belief about the drawbacks of unanimity rule:

Proposition 1 *Under unanimity, the weakest member of the organization executes her veto power, holding back the entire organization at her privately optimal choice.*

Proposition 1 captures the idea that unanimity voting may result in the weakest member blocking any attempt to increase organization-wide effort. In principle, unanimity may well favour stronger rather than weaker members of an organization. For example, more productive (and wealthier) members would exercise their veto power if the organization were to vote on redistribution. However, we have chosen to analyse the wide-spread view that unanimity protects particularly weak members and slows down reforms (e.g., Erlenmaier and Gersbach, 2004).

3 Undermining Veto Power

Under unanimity the weakest member cannot be overruled to exert more effort than her privately optimal level. This veto power is, however, undermined once some members can form or threaten to form an inner organization with the purpose of providing an additional public good. We extend the game and allow a subset of members to form an "inner organization". We will use three terms: First, "initial organization" is defined as an organization in which no inner organizations form, and no additional effort is induced by the threat; second, "reformed" organization is defined as equilibrium with no inner organization and a higher organization-wide effort due to the threat; third, "divided" organizations refers to outcomes in which the threat of forming an inner organization is executed in equilibrium.

The members of an inner organization continue to be members of the initial, henceforth "outer" organization. Instead of adopting a multi-task framework (Holmström and Milgrom, 1991) which views inner and outer club efforts as substitutes, we assume

[6] The inevitable free-riding problem in teams where the members' inputs are substitutes can be avoided when inputs are strict complements. For such teams, there exists a linear (balanced-budget) sharing rule that implements the efficient outcome as a Nash equilibrium outcome (Legros and Matthews, 1993; Vislie, 1994).

negative externality in consumption: An inner organization with n members reduces the utility of consuming the outer public good for all N agents by λn with $\lambda \geq 0$. The deadweight loss λ is meant to capture the notion that the formation of an inner club causes its members to divert attention and effort from the outer organization.

The production technology of the inner organization is similar to the one of the outer organization. Membership in the inner organization generates additional per-capita benefits of $n(e_{In} - e_{Out})$, where e_{In} (e_{Out}) denotes the minimal effort exerted by anyone who is a member of the inner (outer) organization. Notice that the efforts of members are still complements in the production functions of the inner and outer organization, but that members of the inner organization now have the possibility to exert additional effort.

To keep the analysis tractable, we abstract from the possibility of multiple inner organizations and allow for at most one inner organization. Furthermore, the inner organization must have at least two members ($n \geq 2$). This is a natural restriction because an inner organization provides a public - rather than private - good to its members.

We use the term constellation for a partitioning of members into an inner organization with $n \leq N$ members, together with the associated effort levels in the outer and in the inner organization. The payoff of type i who is a member of both the inner and the outer organization is

$$y_i = n(e_{In} - e_{Out}) + Ne_{Out} - \lambda n - \theta_i e_i^2/2.$$

The payoff of type j who is only a member of the outer organization is

$$y_j = Ne_{Out} - \lambda n - \theta_j e_j^2/2.$$

Each member can freely decide whether she wants to join the inner club. Hence, the formation of the inner club is not subject to unanimous approval, the constitution of the initial organization is incomplete.

In general, the formation of an inner organization would depend on how agents coordinate, for instance who determines the effort level e_{In} and/or the size n of the inner club. We intentionally abstract from members' coordination mechanisms and let nature choose e_{In}, instead of analyzing the game with an arbitrary agenda-setting procedure. This allows us to identify all constellations that can be supported as Nash equilibrium outcomes which we believe to be *per se* of interest. In addition, these constellations constitute the constraints of any potential decision-maker's optimization problem, who - by virtue of an imposed extensive form game - sets the agenda.

The production of the outer and possible inner club goods takes place in three stages. In the first stage, members vote on the minimum effort of the outer organization. As before, voting follows the ascending procedure under the unanimity rule. In the second stage nature draws an inner club effort level e_{In}. Following the logic of the model, we restrict the possible draws of nature to $e_{In} > e_{Out}$. Having observed e_{In}, all members have the option to simultaneously subscribe to join the inner club. By subscribing, a member commits to exert e_{In}. Otherwise, she gets infinitely punished. Non-subscribers cannot be members of the inner organization, irrespective of their subsequent effort choice. In the final stage, all members simultaneously choose their effort and the club goods are produced.

Nature's draw e_{In} is strictly binding in the sense that inner club members have to exert exactly e_{In}, neither less nor more. The draw e_{In} restricting the inner club effort levels from below parallels the assumption in the voting stage. The additional restriction of e_{In} being binding from above is a simplification, introduced to avoid multiplicity of equilibria. We discuss its implications in section 5.

Finally, we assume that in the first stage members vote non-strategically. That is, each member's decision to withdraw from the voting does not take into account its impact on the subsequent subscription decision of other members. As we show in section 5, there is scope for strategic voting but it does not change the qualitative predictions of our model.

The possibility of an inner club can increase organization-wide effort or result in a divided organization. In the remainder of this section, we analyse the first outcome and relegate the discussion of the formation of inner organizations to the next sub-section.

The assumption of sincere voting pins down a unique voting outcome, as each member i withdraws at her preferred effort level N/θ_i. So the weakest member ends the voting by exiting at $e^V = e_{Out} = N/\bar{\theta}$. That is, non-strategic voting rules out all Pareto-inferior equilibria of the basic framework.

If nature draws a moderate level of e_{In}, there exist a Nash equilibrium where all members subscribe to join the inner organization and exert exactly e_{In}. Consider the choice of the weakest member when all other members subscribe to e_{In}. If she also subscribes, she exactly matches the announced threshold e_{In} as any other effort level entails an infinite penalty. Alternatively, if she abstains from joining, an inner club of size $N-1$ forms. She then sets the outer-club effort to her most preferred level

$e_N^* = N/\bar{\theta}$. This option entails lower disutility of effort but also lower consumption and in addition the deadweight loss $\lambda(N-1)$. Comparing the respective payoffs of the weakest type

$$Ne_{In} - \bar{\theta}e_{In}^2/2 \geq N(N/\bar{\theta}) - \lambda(N-1) - \bar{\theta}\left(N/\bar{\theta}\right)^2/2 \qquad (2)$$

reveals that she prefers to join the inner organization for all efforts

$$e_{In} \in \left(N/\bar{\theta}, N/\bar{\theta} + \sqrt{2\lambda(N-1)/\bar{\theta}}\right].$$

Indeed, solving the quadratic inequality (2) for e_{In} and imposing the constraint $e_{In} > e_{Out}$ yields the above interval.

Suppose type N joins and consider type $N-1$. Similarly to the weakest type, if she does not join the inner organization, she would want to set the outer-club effort to her preferred level $e_{N-1}^* = N/\theta_{N-1}$. If $e_{In} > N/\theta_{N-1}$ this deviation cannot be profitable because type $N-1$ has to exert less "additional" effort $(e_{In} - e_{N-1}^*)$ at a lower cost than the weakest type who still prefers to join. If $e_{In} < N/\theta_{N-1}$ her preferred effort level is not even feasible as the outer-club effort cannot exceed the inner-club one. This follows from the Leontief technology and the fact that members of the inner club continue to be members of the outer club. Thus, type $N-1$ also prefers to subscribe to the inner organization. This reasoning applies to all other types $i = 1, .., N-2$, and consequently all members joining and exerting e_{In} is an equilibrium outcome for $e_{In} \in \left(N/\bar{\theta}, N/\bar{\theta} + \sqrt{2\lambda(N-1)/\bar{\theta}}\right]$. For future reference, we denote by $e^B = N/\bar{\theta} + \sqrt{2\lambda(N-1)/\bar{\theta}}$ the maximum effort level that is compatible with all N members choosing the same effort.

Proposition 2 *Reformed organizations can emerge for any $\lambda > 0$ and $e_{In} \leq e^B$. That is, the threat of forming an inner club undermines the veto power of the weakest member and increases organization-wide effort.*

Unanimity is commonly viewed as preventing majorities from coercing minorities at the cost of organizational inertia or inability of adjusting. Proposition 2 shows that this view needs to be qualified: Unanimity need not be tantamount to complete protection of the weakest members or, equivalently, to the inability to reform. The threat of forming an inner organization can undermine the veto power of each single member and may enable the organization to reform. To be an effective reform mechanism, two conditions must hold. First, the statutes of the organization must be incomplete, thereby exempting the formation of an inner organization from the unanimous approval. Otherwise, weaker members have no reason to avoid the formation of an inner organization by exerting more effort, since they can simply veto its formation. Second, the inner organization must impose some externalities on the outer organization. Otherwise, the weaker members have no incentives to increase their effort beyond their privately optimal level. The "reform potential" of an organization, measured by the difference $e^B - N/\bar{\theta}$, increases with the deadweight loss associated with an inner organization and with the size of initial organization.[7]

While the threat of an inner organization can make the initial organization more dynamic, the organization may also remain in inertia.

Corollary 1 *The initial organization is always an Nash equilibrium of the extended game, irrespective of nature's draw e_{In}.*

Even in the presence of the threat, the initial organization, that is, the outcome in which the weakest member imposes his privately optimal choice $e_N^* = N/\bar{\theta}$ on the entire organization continues to be a Nash equilibrium of the extended game. This follows from the assumption that an inner organization must have at least two members. Hence, if all other members choose not to subscribe to e_{In} and exert the weakest member's preferred effort level, no single member has an incentive to deviate from this common pattern.

European Union and Euro

The European Union offers a good illustration for reforms under unanimity. In the aftermath of the Second World War a number of initiatives were taken to bring about cooperation. In 1951 France, the Benelux countries, Germany and Italy formed the "European Coal and Steel Community" (ECSC) to coordinate in two important industries. Six years later, on March 25, 1957, the six countries signed the Treaty of Rome creating the European Economic Community (EEC) to promote trade and exchange among them.[8] Achieving integration required removing barriers to trade. In order to reassure concerned electorates it was agreed that most important decisions

[7]Let $f(N, \lambda) \equiv 2\lambda(N-1)$. Clearly, $f_N > 0$, $f_\lambda > 0$, and $e^B = e_N^*$ at $\lambda = 0$.

[8]The signatories stated their determination "to lay the foundations of an ever closer union among the peoples of Europe" and called "upon the other peoples of Europe who share their ideal to join in their efforts" thus opening up for new members to join, provided they meet certain criteria (Preamble to the Treaty of Rome).

would be taken by unanimity in the Council, the EEC's main decision-making body where the governments of all countries are represented.

Already in early stages, there were conflicts about the intensity of these efforts. For instance, a plan for further integration into a European Defence Community failed in 1954. The French Parliament did not ratify the plan, because of fears that the EDC threatened France's national sovereignty. France valued the costs of further integration efforts higher than its partners and made use of the veto right each nation had within the European Community (EC).

Despite the veto powers entrusted with individual member states, significant integration was achieved and membership doubled: Great Britain, Ireland, Denmark, Greece, Portugal and Spain joined in successive enlargements, often after substantial domestic political and economic reforms.

During the second half of 1980s, EC Commission President Jacques Delors and stronger members pushed for further integration. This process resulted in the Treaty of Maastricht, which states in article 2: "This Treaty marks a new stage in the process of creating an ever closer union among the peoples of Europe..." The core proposal was the creation of a common currency area with stringent criteria for joining the "club-in-the-club", the European Monetary Union.

Reaching the Maastricht criteria was only possible if countries made substantial and painful reform efforts. More importantly, these efforts would be more painful for countries with larger budgetary problems, such as Greece or Italy. However, the benefits of further integration and the creation of a joint currency would accrue to all participating members. Arguably, the process of reaching the criteria led to a revitalization of the European integration process and a phase of growth.

In the language of our model a group of economically stronger countries brought forward a proposal that was open to everyone. However, inclusion in the new European club Euroland was only possible after exerting substantial efforts. The threat of forming such an inner club that would have excluded the "underperformers" seems to have worked. Countries that wanted to join managed to reach the criteria.

4 Club-in-the-club Formation

We have so far only looked at equilibrium constellations in which there are no inner organizations in equilibrium. We here explore the set of constellations with divided organizations that can be supported as Nash equilibrium outcomes for any given deadweight loss λ.

Assume an inner organization exists. Because of our assumption of non-strategic voting, $e^V = N/\bar{\theta}$. All agents who do not subscribe to the inner club exert this effort level $e_{Out} = e^V = N/\bar{\theta}$. As the inner club effort level e_{In}, drawn by nature, is binding, all members who join the inner organization exert e_{In}.

For expositional simplicity, we only consider inner organization with $n \in \{3, ..., N-2\}$, where the relevant participation constraints have the same functional form. While our setting allows for $n = 2$ or $n = N - 1$, the respective participation constraints are specific to these two cases. Occasionally, we will comment on the latter cases, but if not explicitly stated our discussion refers to $n \in \{3, ..., N-2\}$.

We first establish which types are members of both the inner and outer organization.

Lemma 1 *Provided that an inner organization of size $n \in \{3, ..., N-2\}$ is an equilibrium outcome, its members are the low ranked (most productive) types $i \in \{1, ..., n\}$.*

Given that an inner organization of size n exists, two types of constraints must be satisfied. First, $N - n$ members of the outer organization must prefer staying in the outer organization rather than joining the inner organization. Second, the n members of the inner organization must prefer to stay in the inner organization.

An agent i chooses to be a member of the inner organization if the following condition holds:

$$n(e_{In} - e_{Out}) + Ne_{Out} - \lambda n - \theta_i e_{In}^2/2 \geq Ne_{Out} - \lambda(n-1) - \theta_i e_{Out}^2/2.$$

Rearranging yields

$$n(e_{In} - e_{out}) - \lambda \geq \theta_i \left[e_{In}^2 - e_{Out}^2\right]/2\ . \tag{3}$$

The LHS of this constraint is composed of the per-capita membership benefits of the inner organization net of the deadweight loss λ. Both are independent of the type, i. The RHS is type i's cost differential between exerting the inner and outer effort levels. It increases in i. Thus, if condition (3) holds for type i, it must hold for all more productive types $j = 1, ..., i-1$. That is, if type i chooses to be member of the inner organization, so do types 1 to $i-1$. Hence, given the equilibrium inner club has n members these members must be exactly the n most productive agents. The $N - n$ least productive agents are only members of the outer club.

The above result implies that an inner organization forms in equilibrium if the following two constraints are satisfied:

$$n\left(e_{In} - N/\bar{\theta}\right) - \lambda \geq \theta_n \left[e_{In}^2 - (N/\bar{\theta})^2\right]/2\ , \tag{4}$$

$$\theta_{n+1}\left[e_{In}^2-(N/\bar{\theta})^2\right]/2>(n+1)\left(e_{In}-N/\bar{\theta}\right)-\lambda. \tag{5}$$

The first condition ensures that the marginal, i.e., least productive, member of the inner organization prefers to be member in both the inner and the outer organization. The second condition ensures that the most productive member in the outer organization prefers to be member in the outer organization only. We assume that type $n+1$ does not join in case she is indifferent which accounts for the the strict inequality in the non-participation constraint (5).

The above discussion makes clear why the restriction to inner clubs of size $n \in 3,...,N-2$ simplifies matters. For these cases, the participation constraints (4) and (5) only differ with respect to the the size of the inner club and the marginal members' productivity. For the inner club with $n = N-1$ members, the non-participation constraint of type $n+1$ is different. If type N were to join the inner organization, all members would exert the same effort and an inner organization ceases to exist. Thus, type N would eliminate the entire deadweight loss, rather than increase it by one unit of λ (as in the case of all other inner organizations).[9] This modification of constraint 5 does not substantially change the analysis, but it leads to a different functional form of the set of equilibrium effort level e_{In}. Similar effect appears in the participation constraint of type 2 in the inner club of size $n = 2$. As we do not allow for inner clubs consisting of one member, if type 2 does not join the club, the club fails to form.[10]

In what follows, we first establish the conditions for the existence of divided organization and then characterize constellations supporting inner organizations of different size.

An increase in the size of the inner organization benefits all its members as it rises the amount of the inner public good. However, the less productive agents may find it too costly to join the club and exert the requested effort level e_{In}. Hence, any inner organization has to strike a balance between size and productivity of its marginal member. This trade-off has, however, no interior solution if the productivity differences among (two adjacent) members is relatively small. That is, when the N members are

[9]The respective non-participation constraint of type N becomes

$$\bar{\theta}\left[e_{In}^2-(N/\bar{\theta})^2\right]/2-(N-1)\lambda>N(e_{In}-N/\bar{\theta}).$$

[10]Type 2's participation constraint becomes

$$2\left(e_{In}-N/\bar{\theta}\right)-2\lambda\geq\theta_2\left[e_{In}^2-(N/\bar{\theta})^2\right]2.$$

relatively homogeneous, an inner organization never forms in equilibrium.

Lemma 2 *The inner club can only emerge if the agents are sufficiently heterogeneous, $N < \overline{\theta}/\underline{\theta}$.*

When members are relatively homogenous, the benefit of increasing the size of the inner organization exceeds the (possibly) adverse effect on the effort levels that are compatible with the resulting inner organization. As shown in Lemma A1 in the Appendix, for $N > \overline{\theta}/\underline{\theta}$ the best preferred effort level of the marginal member of the inner club, $e_n^* = n/\theta_n$, increases in n, the size of the inner club.[11] As the nature drawn e_{In} must exceed $N/\bar{\theta}$, no marginal member would ever subscribe to e_{In}. Indeed, her alternative is to exert the $e_{Out} = N/\bar{\theta}$, which by Lemma A1 differs less from her best preferred inner club effort level, and does not entail an externality.

For the remainder of this section, we assume that the heterogeneity condition $N < \overline{\theta}/\underline{\theta}$ holds. This is, however, only a necessary condition for the existence of inner organizations. Rather intuitively, the size of the deadweight loss and the level of the inner club effort, drawn by nature, also matter. Indeed, even when members are heterogeneous, divided organizations only exist for certain pairs of (e_{In}, λ).

Notation 1 *Denote by Ω the set of all pairs (e_{In}, λ) such that*
i) for $\lambda \in (\lambda_{\min}, \lambda_{\max}]$

$$e_{In} \in \left[\frac{3}{\theta_3} - \sqrt{\left(\frac{3}{\theta_3} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_3}}, \frac{3}{\theta_3} + \sqrt{\left(\frac{3}{\theta_3} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_3}}\right],$$

ii) for $\lambda \in (0, \lambda_{\min}]$

$$e_{In} \in \left\{ \begin{array}{l} \left[\frac{3}{\theta_3} - \sqrt{\left(\frac{3}{\theta_3} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_3}}, \frac{N-1}{\theta_{N-1}} - \sqrt{\left(\frac{N-1}{\theta_{N-1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{N-1}}}\right) \\ \cup \left(\frac{N-1}{\theta_{N-1}} + \sqrt{\left(\frac{N-1}{\theta_{N-1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{N-1}}}, \frac{3}{\theta_3} + \sqrt{\left(\frac{3}{\theta_3} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_3}}\right] \end{array} \right.,$$

[11] Given that $y_i = n(e_{In} - N/\bar{\theta}) + N(N/\bar{\theta}) - \lambda n - \theta_i e_i^2/2$ is the payoff of type i who is a member of the inner organization of size n and the outer organization, type i would choose $e_{In} = e_i^*(n) = n/\theta_i$.

iii) for $\lambda = 0$,

$$e_{In} \in \left[2\frac{N-1}{\theta_{N-1}} - \frac{N}{\overline{\theta}}, 2\frac{3}{\theta_3} - \frac{N}{\overline{\theta}}\right]$$

where

$$\lambda_{\min} = \frac{\theta_{N-1}}{2}\left(\frac{N-1}{\theta_{N-1}} - \frac{N}{\overline{\theta}}\right)^2,$$

$$\lambda_{\max} = \frac{\theta_3}{2}\left(\frac{3}{\theta_3} - \frac{N}{\overline{\theta}}\right)^2.$$

The set Ω is determined by type 3's participation constraint and type $N-1$'s non-participation constraint. These constraints define the largest set of pairs (e_{In}, λ) that support a divided organization outcome. Indeed, the non-participation constraint of member n is

$$n(e_{In} - e_{out}) - \lambda < \theta_n \left[e_{In}^2 - e_{Out}^2\right]/2.$$

Sufficient heterogeneity ($N < \overline{\theta}/\underline{\theta}$) implies $n/\theta_n > (n+1)/\theta_{n+1}$, i.e., θ_n is increasing faster than n (Lemma A1). Therefore, if type 3 does not want to be the marginal member of the inner organization of size 3, no type $k > 3$ prefers to be the marginal member of the inner club of size k. Similarly, if type n wants to be the marginal member of the inner club of size n, any more productive type $k < n$ has the same preference. Finally, member $N-1$'s non-participation constraint has to be met, as we only consider inner clubs of size $n \in \{3, .., N-2\}$.

Proposition 3 *Provided that types are heterogenous, an inner club of size* $n \in \{3, .., N-2\}$ *can form in equilibrium iff the pair* (e_{In}, λ) *belongs to the set* Ω*. Moreover, for each pair* (e_{In}, λ) *the size and composition of the inner club is unique.*

If the deadweight loss λ is very high, type 3 refrains from subscribing to the inner club, choosing to be held back in the outer organization. Similarly, signing up for very high effort level e_{In} becomes too costly for type 3, so, again, she stays in the initial organization. Finally, if the inner club effort level is not very different from the outer club one, there is little value for the type 3 in joining the inner club and suffering the dead-weight loss λ. As explained above, type 3's willingness to join is crucial for the existence of divided organization of any size.

Figure 1 depicts the set of pairs (λ, e_{In}) that are consistent with the divided organizations. Drawn in the λ-e_{In} space, the boundaries of this set are two parabolas

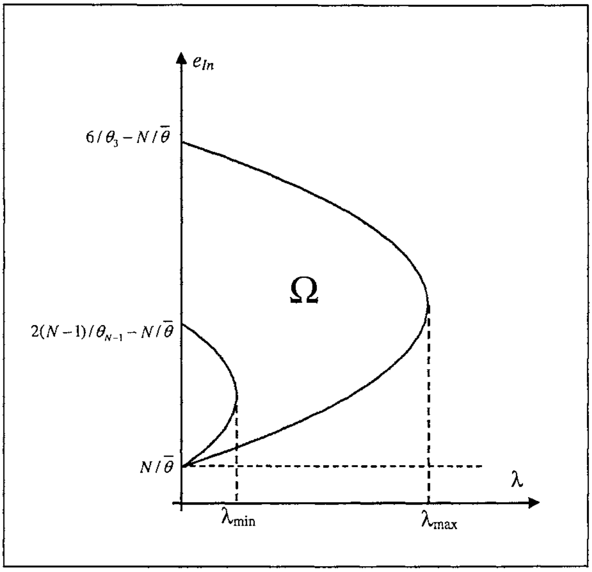


Figure 1:

that are tipped to the vertical axis. The outer parabola is determined by the participation constraint (condition 4) of type 3 and the inner parabola by the non-participation constraint (condition 5) of type $N-1$.

The uniqueness is easily understood by comparing the conditions for the existence of two inner organizations that differ in size by one member. On the one hand, type k prefers to be a member of the inner organization of size k if

$$k\left(e_{In} - N/\bar{\theta}\right) - \lambda \geq \theta_k \left[e_{In}^2 - (N/\bar{\theta})^2\right]/2$$

(condition 4) holds. On the other hand, type k prefers not to join the inner organization of size $k-1$ if

$$k\left(e_{In} - N/\bar{\theta}\right) - \lambda < \theta_k \left[e_{In}^2 - (N/\bar{\theta})^2\right]/2$$

(condition 5) holds. The inequality sign apart, these two constraints are identical. Accordingly, there exists no effort level e_{In} that can be supported for a given deadweight loss by more than one inner organization.

More generally, heterogeneity implies that if type k does not want to be the marginal member of an inner organization of size k, neither of types $k+1, k+2, ...$ wants to be the marginal member of the inner club of respective size. Therefore, inner clubs with non-adjacent size cannot coexist.

By Lemma 1 an inner organization of size n comprises the n most productive members, which establishes the unique composition of the club.

Now we are ready to study the comparative statics of the divided organization's equilibria with respect to the deadweight loss λ and inner club effort level e_{In}. From Proposition 3 we know that each pair (λ, e_{In}) corresponds to a unique inner organization of size $n(\lambda, e_{In})$. Given that the size of the inner club is a discrete variable, we do not study marginal effects. Instead, we describe the set of all the pairs (λ, e_{In}) that lead to formation of the inner club of size n in equilibrium.

We partition the set Ω into subsets Ω_n, each corresponding to the pairs (λ, e_{In}) consistent with an inner organization of size n. Define $\lambda_n \equiv \theta_n/2\left(n/\theta_n - N/\overline{\theta}\right)^2$. Lemma A2 in Appendix A.4 establishes that λ_n decreases in n for $N < \overline{\theta}/\underline{\theta}$.

Notation 2 *For $n \in \{3, ..., N-2\}$ denote by Ω_n the set of all pairs (λ, e_{In}) such that i) for $\lambda \in (\lambda_{n+1}, \lambda_n]$*

$$e_{In} \in \left[\frac{n}{\theta_n} - \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}}, \frac{n}{\theta_n} + \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}}\right],$$

ii) for $\lambda \in (0, \lambda_{n+1}]$

$$e_{In} \in \begin{cases} \left[\frac{n}{\theta_n} - \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}}, \frac{n+1}{\theta_{n+1}} - \sqrt{\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{n+1}}}\right) \\ \cup \left(\frac{n+1}{\theta_{n+1}} + \sqrt{\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{n+1}}}, \frac{n}{\theta_n} + \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}}\right] \end{cases},$$

iii) for $\lambda = 0$

$$e_{In} \in \left[2\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}, 2\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right].$$

Each set Ω_n is determined by type n's participation constraint and type $n+1$'s non-participation constraint.

Proposition 4 *Provided that types are heterogenous, an inner club of size $n \in \{3, ..., N-2\}$ can form in equilibrium iff the pair (e_{In}, λ) belongs to the set Ω_n.*

To understand Proposition 4, suppose first, that the nature draws $e_{In} = n/\theta_n$, and an inner organization of size n forms. By Lemma A1 in the Appendix, the best preferred effort choice of the marginal member $e_n^*(n) = n/\theta_n$ decreases in the size of the inner organization. This in turn implies that $N/\overline{\theta} < n/\theta_n < N/\theta_n$. Thus, when leaving the inner organization, type n would reduce the deadweight loss by one unit of λ, but would also move further away from his optimal effort level. Hence, whether or not type n's participation constraint (condition 4) is satisfied depends on the size of the deadweight loss.

When the deadweight loss is relatively large ($\lambda > \lambda_n$), type n strictly prefers to reduce the deadweight loss rather than being a member of the inner organization, even if nature drew type n's preferred effort level, $e_n^*(n) = n/\theta_n$. As the deadweight loss decreases ($\lambda \in (\lambda_{n+1}, \lambda_n]$), there exist values of e_{In}, deviating from n/θ_n, such that the benefits of being in the inner organization exceed the additional deadweight loss λ. Since type n's payoff is symmetric and single-peaked in effort, the mid-point of feasible e_{In} values is n/θ_n, type n's preferred effort level. As the deadweight loss decreases further ($\lambda \in (0, \lambda_{n+1}]$), the range of e_{In} values for which type n prefers to be part of the inner organization becomes larger and includes values for which type $n+1$ would also want to join the inner organization. To restore the non-participation constraint of type $n+1$ (condition 5), the interval

$$\left(\frac{n+1}{\theta_{n+1}} - \sqrt{\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{n+1}}}, \frac{n+1}{\theta_{n+1}} + \sqrt{\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{n+1}}}\right)$$

has to be removed from the solution set. Consequently, the solutions to conditions (4) and (5) are two disjoint intervals when $\lambda_{n+1} \geq \lambda > 0$. Finally, for $\lambda = 0$ solutions are also two disjoint intervals,

$$e_{In} \in \left\{N/\overline{\theta}\right\} \cup \left[2\,(n+1)\,/\theta_{n+1} - N/\overline{\theta}, 2n/\theta_n - N/\overline{\theta}\right],$$

but $e_{In} = N/\overline{\theta}$ violates the constraint that the effort in the inner organization must be strictly larger than effort in the outer organization.

A noteworthy feature of Proposition 4 is that the maximum sustainable size of an inner organization decreases with the deadweight loss λ. This follows directly from the

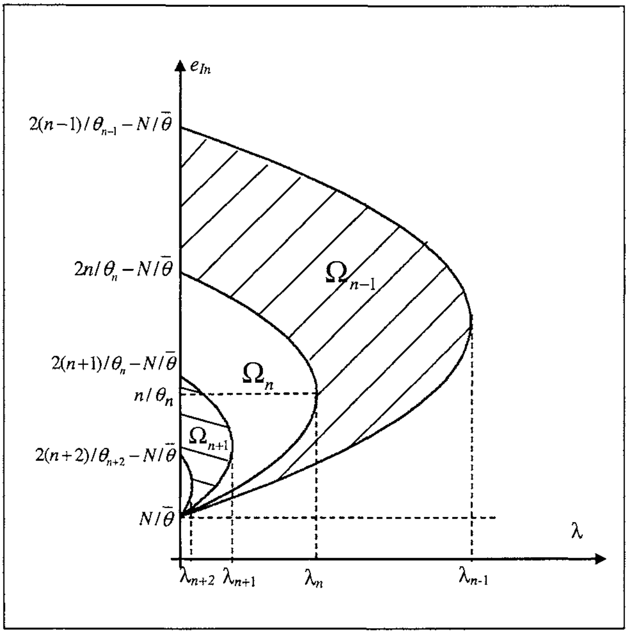


Figure 2:

fact that the critical value of the deadweight loss λ_n below which effort levels e_{In} exist that support an inner organization of size n decreases in n.

Figure 2 shows the equilibrium outcomes for inner organization of different sizes. Plotted in the λ-e_{In} space, the sets of pairs (e_{In}, λ) that support inner organizations of size $n-1$, n, $n+1$, and $n+2$ have an "onion-like" shape with the outer layers enclosing the pairs compatible with smaller inner organizations. The layers do not intersect due to the uniqueness established in Proposition 3.

Figure 2 helps to evaluate the impact of the deadweight loss λ on the equilibrium constellations with divided organizations. For a given λ inner clubs of different size can form depending on the generated inner club effort level e_{In}.

Corollary 2 *Given a deadweight loss λ, if there is an effort level $e_{In}(n)$ compatible with an inner organization of size n, then there exist effort levels $e_{In}(n')$ supporting each of the smaller inner organizations of size $n' < n$.*

The intuition for this result is as follows: for a given λ a club of size n forms in equilibrium if and only if the marginal type n prefers to join the club under her best preferred effort $e_{In} = e_n^*(n) = n/\theta_n$. Indeed, as long as this condition holds, an inner club effort level e_{In} supporting the club of size n is among those that are sufficiently close to $e_n^*(n) = n/\theta_n$, so that n joins the club, and sufficiently far from the $e_{n+1}^*(n+1) = (n+1)/\theta_{n+1}$, so that $n+1$ stays outside. Assume that a club of size n forms in equilibrium. As we have seen in the proof of Proposition 3, if, given e_{In}, type n prefers to join an inner club of size n, type $n-1$ prefers to join a club of size $n-1$. It immediately follows that type $n-1$ prefers to join a club of size $n-1$ under her best preferred effort $e_{n-1}^*(n-1) = (n-1)/\theta_{n-1}$. Thus, there exists a e_{In} such that a club of size $n-1$ forms in equilibrium.

Corollary 3 *The maximum size of an inner organization that can be sustained decreases (weakly) in the deadweight loss λ.*

Indeed, a marginal member of an inner club prefers to join the club and face additional dead-weight loss λ to being hold back in an outer club. Thus, for high λ the difference in the effort costs between the marginal inner club member θ_n and the least productive type $\overline{\theta}$ should be be sufficiently high, otherwise θ_n would stay outside. For a given type distribution it means that the size of the maximum supportable inner club shrinks as λ increases.

To complete the analysis we address the coexistence of initial, reformed and divided organizations. For any parameter values, the initial organization is an equilibrium of the extended game. If in addition, the effort e_{In} is not too high, the reformed organization can co-exist in equilibrium. The divided organization requires sufficiently heterogeneous members in combination with moderate levels of e_{In} and deadweight loss λ. Therefore, all three organization outcomes only coexist if the heterogeneity condition holds and the pair (e_{In}, λ) belongs to a subset of Ω.

More specifically, a pair $\left\{(e_{In}, \lambda) \in \Omega : e_{In} \leq e_B \equiv N/\overline{\theta} + \sqrt{2\lambda(N-1)/\overline{\theta}}\right\}$ can result both in an equilibrium with a unique inner club and a reformed organization with all N members exerting e_{In}. (Clearly, the initial organization is also supported.) The coexistence is due to the fact that an inner club gives rise to a deadweight loss. The least efficient member may prefer to stay outside if at least one more member stays outside (that is, $n \leq N-2$). In this case joining the inner club has double cost: it involves exerting a higher effort e_{In} and bearing a higher deadweight loss. Instead, if the other $N-1$ types exert e_{In}, by joining them the least efficient type faces the cost of exerting too high effort but at the same time eliminates the inner club externality. Thus, the weakest member may choose to join and exert e_{In} if all the other types join. Figure 3 depicts the coexistence region of reformed and divided organization.[12]

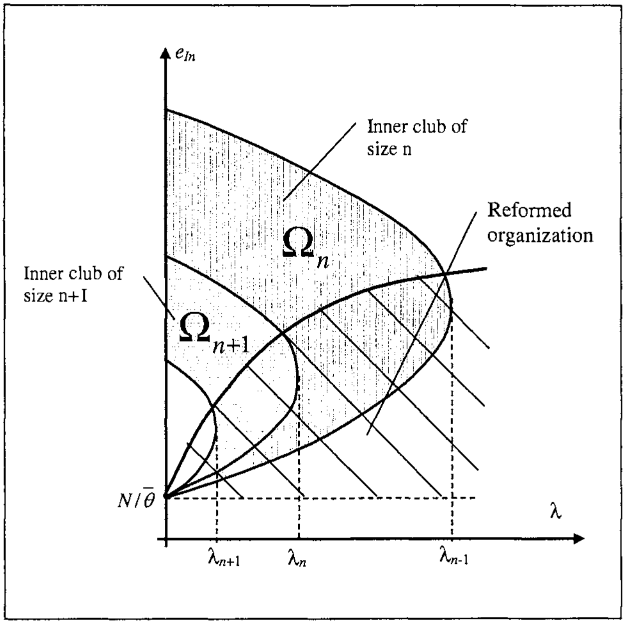


Figure 3:

Heterogeneity and The Threat of a Two-speed Europe

The founding EU members had at the time quite similar aims and economic structures, such as the importance of heavy industry sectors and the ambition to coordinate the steal and coal industries. Through a number of enlargement waves, the heterogeneity of EU members increased dramatically, thereby altering matters considerably. With the southern periphery joining, a new challenged has to be confronted: rather backward economies had to be developed with the help of structural fonds. Eastern enlargement increased these necessities, and more funds had to be moved from other budgetary

[12]This does not hold for the inner organization of size $N-1$, as member N's decision to exert either e_{In} or $N/\bar{\theta}$ fully determines the equilibrium constellation. That is, divided and reformed organization cannot co-exist. This constitutes the only qualitative difference of the inner club of size $N-1$ relative to all smaller inner clubs.

items, increasing the potential for disagreement about each member's financial contributions to the Union. In addition, the Eastern enlargement created or intensified disputes over other issues, such as social standards and labor market liberalisation.

These developments have resulted in a growing concern about paralysis in the EU, and it is on this background that discussions about a two-speed Europe have come up. Interestingly, it was representatives of the founding members, France and Germany (President Chirac and Former Foreign Minister Joschka Fischer) who proposed to allow a subset of EU members to cooperate and integrate more. As in our model, more heterogeneity increases the likelihood of the divided organization outcome. While such a "club-in-the-club" would facilitate decision-making and stimulate cooperation among its members, those countries left out would suffer from trade diversion and exclusion from important decisions.

The intergovernmental conference in Nice in 2001 was supposed to resolve these problems by facilitating decision-making in the new larger Union and by regulating the formation of inner clubs. The Nice Treaty succeeded in regulating the formation of inner clubs in specific areas of cooperation, so-called enhanced cooperation, giving the right to initiate to the European Commission. However, the attempt to make decision-making more efficient was widely seen as a failure (Berglöf et al., 2003). As a result, the European Council in Laeken in 2002 decided to call a convention to draft a new constitution for Europe. The delegates met under the threat that, if they failed to agree, a group of countries would form a "club-in-the-club" and proceed with integration in many areas at the expense of countries left out. The Draft Treaty regulates the instrument of enhanced cooperation among sub-groups of countries, and reinforces the sole right of the Commission to formally propose such initiatives (Part III, Title VI, Chapter III, Article III-322-329).

In the light of our model, there was hence an attempt to mitigate the threat of inner clubs by reducing the constitutional incompleteness. The Draft Treaty establishing a Constitution for Europe was admittedly a break-through in the sense that all members acknowledged the need for a written constitution, but it failed in the referenda held subsequently.

5 Discussion

Throughout the analysis, we rely on several core assumptions to keep the model tractable. We now discuss their implications for the results. The assumption of equidistantly distributed types delivers a generic functional form for the participation constraints of the inner club members. The essential feature ensuring the formation of divided organization is the heterogeneity of types, that is, the best preferred effort of the marginal inner club member $e_n^*(n) = n/\theta_n$ decreasing in the club size. We are confident that any distribution satisfying this property can generate divided organization equilibria, though the "thickness" of the Ω_n-layers would differ as compared to the equidistant distribution. The distribution has no impact on the formation of the initial and reformed organizations in equilibrium, as they are solely determined by the decisions of the least productive member.

Nature's draw e_{In} is assumed to be binding not only from below but - unlike in the voting stage - also from above. This simplifying assumption ensures that the members of the inner club exert precisely e_{In} which for a given deadweight loss λ yields a unique divided organization. If inner club members were free to exert a higher effort than nature's draw, inner clubs of different size could emerge in equilibrium. Suppose that a pair (e_{In}, λ) supports an inner club of size n. Then, by coordinating to work harder than e_{In}, the most productive $m < n$ members can form an inner organization in equilibrium. To see this, consider a point $\left(\tilde{e}_{In}, \tilde{\lambda}\right) \in \Omega_n$ in Figure 2. The ray along the vertical line $\lambda = \tilde{\lambda}$ starting at $\tilde{e}_{In}$ and corresponding to an increase in e_{In}, crosses all the sets Ω_{n-1},Ω_{n-2},...,Ω_3.

While we restrict our analysis to a single inner club, the logic of our model seems compatible with multiple inner organizations. For example, if nature draws two e_{In}, a plausible equilibrium candidate is a constellation with two inner clubs, the most productive types being members of both inner clubs, intermediate types joining the "outer-inner" club and the least productive types being only in the outer organization. However, the outcomes in an extended framework will depend on the modelling details such as the assumed interaction between the deadweight loss of different inner clubs, and single vs. multiple inner club membership.

Finally, we turn to the assumption of non-strategic voting. While it is a standard assumption in many political economy models, it may be limiting in our framework. Indeed, the prospect of an inner club provides the members of the initial organization with the incentive to behave strategically in the voting stage. On the one hand, more productive types may choose to withdraw from the voting before the least productive member would pull out when voting sincerely. While this reduces the provision of the outer club good, it may induce more members to subscribe to the emerging inner club. On the other hand, less productive members may remain in the voting beyond their best preferred level. Though costly, the extra effort reduces the attractiveness of an inner club, thereby lowering the number of its potential members and the consequent deadweight loss, or even preventing its formation altogether.

Therefore, allowing for strategic voting would likely alter the effort level of the outer club and the size of the inner organization. Nonetheless, we would expect to observe the same types of organizational outcomes: divided as well as reformed and initial organizations. In addition, the game may feature an equilibrium in which all members exert an effort below the best preferred level of the least productive member.

It is worth noting that strategic voting entails certain costs but uncertain benefits. When a highly productive member withdraws early, the outer club good is provided at the lower level. At the same time, a larger inner club may or may not materialize depending on the draw of e_{In}. More generally, the benefits of strategic voting depend on the extent to which the agent can influence or correctly anticipate the subsequent decision (i.e., the level e_{In}). Our setting abstracts from any specific agenda setting procedure and lets e_{In} be randomly chosen by "nature". In this hard-to-predict environment, the benefits of strategic voting seem particularly limited, making the sincere voting assumption less restrictive than it may seem at first glance.

6 Unanimity vs Majority Rule

The preceding sections show that unanimity gives rise to two potential costs. If it is not possible to form an inner organization, it results in holding back the more efficient members. If an inner organization can be (and is) formed, the unanimity rule brings about a deadweight loss associated with divided organizations. This suggests that the unanimity rule may be inferior to other decision rules, such as the simple majority. In what follows we compare organizations operating under the simple majority and unanimity rule.

We consider two N-member organizations with different voting rules, each playing the extended game of section 3. As before, voting follows the ascending procedure, but under majority rule it ends once the median type has chosen to “leave the auction”. Again, we assume that in the second stage nature draws a (potential) inner club effort that exceeds the one voted upon in the first stage, i.e., $e_{In} > e_{Out}$. The purpose of our analysis is to compare the ex-post social welfare in these two organizations. To measure welfare we use the sum of all members' utilities (a Benthamite welfare function).[13]

[13]While this choice may be questioned as we assume non-transferable utility, there seems no obvious

To keep the comparison tractable, we consider an organization consisting of three types: $\theta_1 = \underline{\theta}$, $\theta_3 = \overline{\theta}$ and $\theta_2 = \alpha\underline{\theta}+(1-\alpha)\overline{\theta}$. Subsequently, we refer to this organization as the organization $\mathbb{A}(\alpha)$. In addition, we impose two restrictions. First, we assume that the members are sufficiently heterogenous, i.e., $3 < \overline{\theta}/\underline{\theta}$. This assumption replicates the condition in Lemma 2 ensuring that divided organizations are feasible under unanimity. Second, we only consider $\alpha \in [1/2, 1]$. For $\alpha = 1/2$, organization $\mathbb{A}(\alpha)$ is simply the three-type version of the previous equidistant distribution. For $\alpha > 1/2$, the distribution becomes skewed to the left. That is, the median type is closer to the most productive type. We abstract from cases $\alpha < 1/2$ where the majority of members is less productive than the (hypothetical) average type. This restriction is in the spirit of our paper that focuses on ways of overcoming the power of weak members to hold back the organization.

If an inner club emerges in equilibrium, the welfare of organization $\mathbb{A}(\alpha)$ is

$$W^{n=2} = \sum_{i=1}^{3} \left(2(e_{In} - e_{Out}) + 3e_{Out} - 2\lambda - \theta_i e_{In}^2/2\right) .$$

If instead all members exert the same effort in equilibrium, the resulting welfare is

$$W^{n=\emptyset} = \sum_{i=1}^{3} \left(3e - \theta_i e^2/2\right) .$$

This case comprises two equilibriums outcomes: the initial organization with $e = e_{Out}$, and the reformed organization with $e > e_{Out}$.

To evaluate the welfare under the two voting rules, we need to characterize the set of equilibrium outcomes in either case. Consider first the unanimity rule.

Lemma 3 *In organization $\mathbb{A}(\alpha)$ under the unanimity rule, initial, reformed and divided organizations can all be supported as equilibrium outcomes.*

Similarly to the N-type case, voting in the first stage results in $e_{Out} = 3/\overline{\theta}$, which is the effort exerted in the initial organization. The reformed organization can be supported for any $e_{In} \in \left(3/\overline{\theta}, 3/\overline{\theta} + \sqrt{4\lambda/\overline{\theta}}\right]$. As we do not allow for inner clubs of size 1, the only possible inner club of organization $\mathbb{A}(\alpha)$ has exactly 2 members.

alternative welfare function. For instance, one typically cannot Pareto-rank the outcomes under these voting rules.

Adapting the participation constraints (4) and (5) to organization $\mathbb{A}(\alpha)$, we obtain the conditions for the existence of an inner club $n = 2$

$$2\left(e_{In} - 3/\bar{\theta}\right) - 2\lambda \geq \theta_2\left[e_{In}^2 - (3/\bar{\theta})^2\right]/2 \tag{6}$$

$$\bar{\theta}\left[e_{In}^2 - (3/\bar{\theta})^2\right]/2 \geq 3\left(e_{In} - 3/\bar{\theta}\right) + 2\lambda \tag{7}$$

Solving the system (6) and (7) we get that an inner club can emerge for $\lambda \leq \frac{\theta_2}{4}\left(\frac{2}{\theta_2} - \frac{3}{\overline{\overline{\theta}}}\right)^2$ and

$$e_{In} \in \left[\max\left\{\frac{3}{\overline{\overline{\theta}}} + \sqrt{\frac{4\lambda}{\overline{\overline{\theta}}}}, \frac{2}{\theta_2} - \sqrt{\left(\frac{2}{\theta_2} - \frac{3}{\overline{\overline{\theta}}}\right)^2 - \frac{4\lambda}{\theta_2}}\right\}, \frac{2}{\theta_2} + \sqrt{\left(\frac{2}{\theta_2} - \frac{3}{\overline{\overline{\theta}}}\right)^2 - \frac{4\lambda}{\theta_2}}\right].$$

While the productivity of type 2, i.e., value of α, affects the range of e_{In} in divided organizations, it has no impact on the initial and reformed organization. In these latter cases, the outcome is solely determined by the weakest type.

As type 2 is the median voter in organization $\mathbb{A}(\alpha)$, the outcome of the voting at stage 1 under majority rule is equal to her preferred effort level $3/\theta_2$. Therefore, the outer-club effort exerted under the majority rule is higher than the one under the unanimity rule. Not surprisingly, this affects the feasibility of an inner club.

Lemma 4 *In organization $\mathbb{A}(\alpha)$ under the majority rule, both initial and reformed organizations are equilibrium outcomes, whereas a divided organization never emerges.*

Indeed, the only feasible inner club in this case is the one of two members, with type 2 being the marginal member. Since type 2 is also the median member, she can impose her best preferred effort choice on the initial organization. Given that the provision of the good depends on the size of the club, type 2 would never opt for a divided organization.

Initial and reformed organizations can form for the same reason as under the unanimity rule. Given that the other 2 members exert effort e_{In}, the weakest member's decision is determined by the participation constraint (2) adapted to $e_{Out} = 3/\theta_2$:

$$3e_{In} - \bar{\theta}e_{In}^2/2 \geq 3\left(3/\theta_2\right) - 2\lambda - \bar{\theta}\left(3/\theta_2\right)^2.$$

The resulting set of effort draws $e_{In} \in \left(3/\theta_2, \frac{3}{\bar{\theta}} + \sqrt{9\left(\frac{\bar{\theta}-\theta_2}{\bar{\theta}\theta_2}\right)^2 + \frac{4\lambda}{\bar{\theta}}}\right]$ supports reformed organization outcomes.

Lemmas 3 and 4 enable us to compare the welfare under the two voting rules. Clearly, the social welfare under either rule depends on the parameters, notably the deadweight loss and the productivity of type 2. This also applies to the ranking of the two rules. We consider two particular cases: An organization with the median type being close to the average, and another organization with the median type being close to the most productive type. In either case, we assume zero deadweight loss associated with an inner club.

Lemma 5 *For $\lambda = 0$, there exist sufficiently small non-negative ε and δ, such that in the organization $\mathbb{A}(\alpha)$*
a) if $\alpha = 1/2 + \varepsilon$ the majority rule results in higher welfare than the unanimity rule.
b) if instead $\alpha = 1 - \delta$, the unanimity rule dominates.

Note that $\lambda = 0$ implies that the reformed organization ceases to exist under either rule. Therefore, the only feasible constellation under the majority rule is the initial organization with welfare

$$W_{Maj} = 3 * 3\frac{3}{\theta_2} - \frac{(\underline{\theta} + \theta_2 + \overline{\theta})}{2}\left(\frac{3}{\theta_2}\right)^2.$$

Under the unanimity rule, there are two possible outcomes: the initial organization with welfare

$$W_{Un}^{n=\emptyset} = 3 * 3\frac{3}{\overline{\overline{\theta}}} - \frac{(\underline{\theta} + \theta_2 + \overline{\theta})}{2}\left(\frac{3}{\overline{\overline{\theta}}}\right)^2$$

and the divided organization with welfare

$$W_{Un}^{n=2}(e_{In}) = \sum_{i=1}^{2}\left(2\left(e_{In} - \frac{3}{\overline{\overline{\theta}}}\right) + 3\frac{3}{\overline{\overline{\theta}}} - \frac{\theta_i}{2}e_{In}^2\right) + \left(3\frac{3}{\overline{\overline{\theta}}} - \frac{\overline{\theta}}{2}\left(\frac{3}{\overline{\overline{\theta}}}\right)^2\right).$$

In all three cases welfare depends on the distribution of types (captured by the parameter α), which we omit for notational simplicity.

The welfare ranking between the outcomes under the unanimity follows from a revealed preference argument. Since types 2 and 1 self-select into the inner club, they must be better off than in the original organization. As the formation of inner club does not impose any externality ($\lambda = 0$), type 3 is equally well off in either organization. Hence, for any effort level e_{In} that supports a divided organization, $W_{Un}^{n=2}(e_{In}) > W_{Un}^{n=\emptyset}$. Consequently, the majority rule dominates if W_{Maj} exceeds the largest possible value of $W_{Un}^{n=2}(e_{In})$. Similarly, unanimity is superior if $W_{Un}^{n=\emptyset}$ is larger than W_{Maj}.

Lemma 5 is best understood by considering first why $W_{Un}^{n=\emptyset} > W_{Maj}$, that is, why unanimity dominates when the median type is relatively productive (result b). The social cost of the unanimity rule is that types 2 and 1 are held back in the initial organization. The social cost under the majority rule is the underprovision by type 1, and the overprovision by type 3. For α close to 1 the best preferred efforts of types 1 and 2 are very similar. Under majority rule, underprovision by type 1 becomes negligible, while type 3 overprovides almost the entire difference between type 1's best preferred choice and his own. Under the unanimity rule, the two efficient types deviate from their preferred level by nearly the same absolute amount. Since a given deviation is less costly for more productive types, the two deviations of types 1 and 2 under the unanimity have lower social costs than the one deviation of the least efficient type under majority. This is an clear-cut example of the tyranny of the majority: the outcome under majority not only hurts the weakest member, but also yields a lower welfare for the entire organization. Nonetheless, it is preferred by the majority of its members.

To establish the dominance of the majority rule (result a), we need to compare the welfare of the divided organization under unanimity $W_{Un}^{n=2}(e_{In})$ with the welfare of the initial organization under majority W_{Maj}. When the types are equidistantly distributed ($\alpha = 1/2$), the cost of type 3's deviation under majority is not too large. At the same time, the consumption benefits are considerable as the amount of the outer public good is set by the median type for the entire organization. Under unanimity less outer public good is produced (level effect) and the benefits of the inner club only accrue to type 1 and 2 (size effect). For $\alpha = 1/2$, the size of these two effects ensures that welfare under unanimity is lower than under majority.[14] By continuity, introduction of small positive deadweight loss λ preserves the ranking in both results a) and b).

While the above examples are specific, the respective ranking of the decision rules can be attributed to effects that we believe to be sufficiently general. To see this, consider organizations with many members whose distribution has the following properties: It is unimodal or equidistant, the median type is at least as productive as the average type and the distribution tails are sufficiently heavy, that is, sufficiently many members are either highly productive or unproductive.

The equilibrium outcomes for these organizations parallel those in Lemmas 3 and 4: All three types of organization can form under unanimity, while only initial and reformed organizations emerge under majority rule. In fact, the inexistence of a divided

[14]The deviation cost under unanimity may be higher or lower depending on the drawn e_{In} but it never outweighs the consumption benefit effects.

organization under majority can be proven for the equidistant distributions (Appendix A.8). Intuitively, skewed distributions can only reinforce this result: Higher productivity of the median type reduces the incentive to form an inner organization. Thus, the model implies that enhanced cooperation among a subset of members (divided organization) emerges only in organizations operating under the unanimity rule.

As for the welfare ranking, we expect the unanimity rule to dominate when the median member is considerably more productive than the hypothetical average member. In this case, the best preferred effort choice of the less productive members and of the median type differ substantially. Hence, majority entails high social cost of overprovision, while unanimity avoids it.[15] In addition, the formation of an inner organization partially solves the problem of more productive members being held back under unanimity. This advantage fades as the externality associated with the formation of the inner organization increases, making a divided organization more costly.[16]

By contrast, the majority rule is superior when the median type is close to the average type. Here, the potential cost advantage of the unanimity rule is outweighed by the consumption benefits of the majority rule. Both the level effect and the size effect are present also in many-members organization. Moreover, the presence of the deadweight loss in the divided organization under unanimity further favours the majority rule.

This reasoning and Lemma 5 lead us to conclude that there is no universally optimal rule for decision making in different organizations.

Proposition 5 *The majority rule may be superior or inferior to the unanimity rule.*

This result is compatible with the existence of organizations operating under different decision rules, including unanimity and simple majority. Moreover, organizations which happened to adopt an inferior decision rule perform worse. Hence, one would expect those to be eventually crowded out. Based on such a selection argument, our model predicts that organizations in which median and average member differ substantially operate under the unanimity rule. If instead median and average member are relatively similar, we expect the organization to be governed by a majority rule. This line of reasoning also applies to organizations with multidimensional agendas. Within

[15] A similar logic also applies to the reformed organization outcomes under the two voting rules for moderate values of λ.

[16] As shown above, the initial organization under unanimity rule dominates the majority outcome if the distribution of types is highly skewed. Hence, large values of λ do not necessarily imply that unanimity is inferior.

a given organization, the decision rule may vary across issues, reflecting the differences in distribution of the members' preferences over each agenda issue.

As organizations evolve over time so may their decision rules. For instance, members may become more heterogeneous, say because their productivity develops differently or because new members join the organization (for reasons outside of our model). If the organization operates under unanimity, these changes increase the likelihood of a divided organization outcome with the associated deadweight loss. To avoid this loss, members could respond by "coordinating" on the reformed organization outcome. Alternatively, they may agree to adopt a majority rule, as it precludes the formation of an inner club. This constitutional change results in a higher organization-wide effort, and may be interpreted as a way of institutionalizing the reformed organization outcome, feasible under unanimity. However, the abolition of the unanimity rule may well reflect the power of some members rather than overall welfare considerations.

Finally, Proposition 5 provides a rationale why members of an organization may voluntarily renounce their veto right and subject themselves to the will of future majorities. Suppose members of the organization do not know their type at the time of writing the constitution. Assume further that the agenda setting procedure is sufficiently complex and hard-to-predict, eliminating all incentives to vote strategically. Being ex ante identical, all members have the same preferences over the decision rules, determined by the expected payoff of the average type. Scalar $1/N$ apart, this objective function coincides with the chosen Benthamite welfare function. Consequently, the members' choice follows the above welfare ranking, and for some distributions of types they all unanimously choose the majority rule. Hence, our model can albeit in a restrictive setting explain the emergence of decision rules.

7 Concluding Remarks

The paper presents a theory of loosely-knit organizations. While members have a common interest, there is no governance mechanism in place that enforces contributions to the common good. Hence, organization-wide decisions must be taken unanimously, granting each member veto power. We show that there are nonetheless ways for such organizations to avoid being held back by its weakest members. By threatening to form a club-in-the-club, members that are less interested or less productive can be induced to contribute more to the common good than their best-preferred level. Key for this mechanism is that the formation of a club-in-the-club imposes a deadweight loss on all members, but benefits only for those who join the inner club. Thus, unanimity does not preclude reform, in the sense of all members exerting more effort than is preferred by its weakest members.

We also show that identical organizations can end up quite differently: some may stagnate at the level preferred by its weakest members, others may reform, and yet others may be divided by the formation of an inner club. Furthermore, the divided organization outcome is more likely, when members are more heterogenous.

By contrast, organizations operating under majority rule never divide and thus avoid the deadweight loss. Nonetheless, majority rule does not necessarily dominate the unanimity. The welfare ranking of the two decision rules depends on the difference between the median and the average organization member. This provides a rationale for the emergence and diversity of decision rules in different organizations.

We have illustrated our theory by the process of European integration: the EMU and the introduction of the Euro seem to have worked very much like a threat of an inner organization, which was avoided, because candidate countries made massive reform efforts. However, with increasing heterogeneity, EU decision-makers seems to have realized the risks of creating inner organizations - the two-speed Europe. They do counter this risk by two instruments: trying to channel the desires of members who want to go faster by the so-called enhanced cooperation, and a general tendency towards replacing unanimity votes by qualified majorities.

We believe that there are many more applications of our theory: the dynamics of research centres often exhibits similar tensions between individuals who are more or less committed to research, which often leads to infighting and the creation of sub-research centres. Many sport leagues suffer from a similar tension between high performance teams and those that lag behind, and there has been a threat of top teams to create their own superleagues, be it in basketball or football.

Our paper is only a first step to a more systematic analysis of loosely knit organizations and the club-in-the-club phenomenon. Strategic voting and the possibility of multiple competing inner clubs are extensions that we believe to be particularly interesting.

References

[1] Aghion, Philippe and Patrick Bolton, 2003, Incomplete Social Contracts, *Journal of the European Economic Association*, 1(1): 38-67.

[2] Berglöf, Erik Gérard Roland, Guido Tabellini and Charles Wyplosz, 2003, *Built to Last: A Political Architecture for Europe*, London Centre for Economic Policy Research.

[3] Bolton, Patrick and Gerard Roland, 1997, The Breakup of Nations: A Political Economy Analysis, *Quarterly Journal of Economics*, 112(4): 1057–90.

[4] Bordignon, Massimo amd Sandro Brusco, 2001, Optimal Secession Rules, *European Economic Review*, 45(10): 1811-1834.

[5] Buchanan, James and Roger L. Faith, 1987, Secession and the Limits of Taxation: towrd a Theory of Internal Exit, *American Economic Review*, 77(5): 1023-1031.

[6] Cornbes, Richard and Todd Sandler, 1996, *The Theory of Externalities, Public Goods, and Club Goods*, Cambridge: Cambridge University Press.

[7] Dixit, Avinash, 2003, Clubs with Entrapment, *American Economic Review*, 93(5): 1824-1829.

[8] Erlenmaier, Ulrich and Hans Gersbach, 2004, "Bundling and the Unanimity Rule," CEPR Discussion Paper 4808.

[9] European Convention, 2003, Draft Treaty establishing a Constitution for Europe (18.07.2003), http://european-convention.eu.int

[10] Gradstein, Mark, 2004, Political Bargaining in a Federation: Buchanan meets Coase, *European Economic Review*, 48(5): 983-999.

[11] Harstad, Bård, 2006, Flexible Cooperation: Mandatory and Minimum Participation Rules, Working Paper.

[12] Holmström, Bengt and Paul Milgrom, 1991, Multi-Task Principal-Agent Problems: Incentive Contracts, Asset Ownership, and Job Design, *Journal of Law, Economics and Organization*, 7(1): 24-52

[13] Legros, Patrick and Steven A. Matthews, 1993, Efficient and Nearly Efficient Partnerships, *Review of Economic Studies*, 60(3): 599-611.

[14] Messner, Matthias and Mattias Polborn, 2004, Voting on Majority Rules, *Review of Economic Studies*, 71(1): 115-132.

[15] Milgrom, Paul R. and John Roberts, 1992, *Economics, Organization and Management*, Prentice-Hall International, Englewood Cliffs, New Jersey.

[16] Milgrom, Paul R. and Robert J. Weber, 1982, A Theory of Auctions and Competitive Bidding, *Econometrica*, 50(5): 1089-1122.

[17] Rousseau Jean-Jacques, 1762, *Contract Social*, translated by Maurice Cranston, Penguin Books, London 1986.

[18] Vislie, Jon, 1994, Efficiency and Equilibria in Complementary Teams, *Journal of Economic Behavior and Organization*, 23(1): 83-91.

A APPENDIX

A.1 Lemma A1

Here we characterize how the preferred effort choice of the marginal member of the inner organization $e_n^*(n)$ varies with the size of the inner organization and the productivity difference between two adjacent members. Given an inner organization of size n exists,

$$e_n^*(n) = \arg\max \left\{ n(e_n - N/\bar{\theta}) + N^2/\bar{\theta} - \lambda n - \theta_n e_n^2/2 = n/\theta_n \right\}$$

Lemma A1: *For $N < \bar{\theta}/\underline{\theta}$, n/θ_n increases with n. Otherwise n/θ_n decreases with n.*

Proof. Substracting $(n+1)/\theta_{n+1}$ from n/θ_n yields

$$\frac{n\theta_{n+1} - (n+1)\theta_n}{\theta_n\theta_{n+1}} = \frac{[(n+1)(\theta_{n+1} - \theta_n) - \theta_{n+1}]}{\theta_n\theta_{n+1}}.$$

Using the definition $\theta_{n+1} = \underline{\theta} + n(\bar{\theta} - \underline{\theta})/(N-1)$ and $\theta_n - \theta_{n+1} = (\bar{\theta} - \underline{\theta})/(N-1)$ we obtain

$$\frac{1}{(N-1)\theta_n\theta_{n+1}} \left[(n+1)(\bar{\theta} - \underline{\theta}) - (N-1)\underline{\theta} - n(\bar{\theta} - \underline{\theta})\right]$$
$$= \frac{1}{(N-1)\theta_n\theta_{n+1}} \left[\bar{\theta} - N\underline{\theta}\right] > 0, \text{ if and only if } N\underline{\theta} - \bar{\theta} < 0.$$

■

A.2 Proof of Lemma 2:

Proof by contradiction. If an equilibrium with an inner organization of size $n < N$ exists,

$$n(e_{In} - N/\bar{\theta}) - \lambda > \theta_n \left[e_{In}^2 - \left[N/\bar{\theta}\right]^2\right]/2$$

and

$$e_{In} > e_{Out}$$

must hold. Setting $e_{In} = N/\bar{\theta} + \delta$ and inserting it in the first condition yields

$$\delta(n - N\frac{\theta_n}{\bar{\theta}}) > \frac{\theta_n\delta^2}{2} + \lambda.$$

This can only hold if $(n - N\theta_n/\bar{\theta}) > 0$ or, equivalently, $(n/\theta_n - N/\bar{\theta}) > 0$. Using the definition

$$\theta_n = \frac{1}{N-1}\left[(N-n)\underline{\theta} + (n-1)\bar{\theta}\right],$$

the difference $(n/\theta_n - N/\bar{\theta})$ can be written as

$$\frac{(N-1)n}{\left[(N-n)\underline{\theta} + (n-1)\bar{\theta}\right]} - \frac{N}{\bar{\theta}} > 0.$$

Rearranging yields

$$\frac{(N-n)}{\bar{\theta}\left[(N-n)\underline{\theta} + (n-1)\bar{\theta}\right]}\left(\bar{\theta} - N\underline{\theta}\right) > 0$$

which contradicts $N \geq \bar{\theta}/\underline{\theta}$.

A.3 Proof of Proposition 3:

The participation constraint (4) of type θ_n, the marginal member in an inner club of size n, can be rewritten as

$$n\left(e_{In} - N/\bar{\theta}\right) - \lambda \geq \frac{\bar{\theta}(n-1) + \underline{\theta}(N-n)}{N-1}\left[e_{In}^2 - (N/\bar{\theta})^2\right]/2$$

or, equivalently

$$n\left(e_{In} - N/\bar{\theta}\right)\left(1 - \frac{\bar{\theta} - \underline{\theta}}{N-1}\frac{\left(e_{In} + N/\bar{\theta}\right)}{2}\right) \geq \frac{N\underline{\theta} - \bar{\theta}}{N-1}\left[e_{In}^2 - (N/\bar{\theta})^2\right]/2 + \lambda. \tag{8}$$

Similarly the non-participation constraint of the type θ_{n+1} can be written as

$$[n+1]\left(e_{In} - N/\bar{\theta}\right)\left(1 - \frac{\bar{\theta} - \underline{\theta}}{N-1}\frac{\left(e_{In} + N/\bar{\theta}\right)}{2}\right) < \frac{N\underline{\theta} - \bar{\theta}}{N-1}\left[e_{In}^2 - (N/\bar{\theta})^2\right]/2 + \lambda. \tag{9}$$

(As before, if type $n+1$ is indifferent, she does not join). Define a function of x

$$F(x) = x\left(e_{In} - N/\bar{\theta}\right)\left(1 - \frac{\bar{\theta} - \underline{\theta}}{N-1}\frac{\left(e_{In} + N/\bar{\theta}\right)}{2}\right).$$

As the LHS of inequalities (8) and (9) coincide with $F(n)$ and $F(n+1)$ respectively, an inner club of size n can form in equilibrium if

$$F(x) \geq \frac{N\underline{\theta} - \bar{\theta}}{N-1}\left[e_{In}^2 - (N/\bar{\theta})^2\right]/2 + \lambda \tag{10}$$

holds for $x = n$, but fails for $x = n+1$.

We begin by proving uniqueness. As by construction $e_{In} > N/\bar{\theta}$,

$$1 - \frac{\bar{\theta} - \underline{\theta}}{N-1}\frac{\left(e_{In} + N/\bar{\theta}\right)}{2} < 1 - \frac{\bar{\theta} - \underline{\theta}}{N-1}\left(\frac{N/\bar{\theta} + N/\bar{\theta}}{2}\right).$$

Given that the types are heterogeneous ($N\underline{\theta} \leq \bar{\theta}$),

$$1 - \frac{\bar{\theta} - \underline{\theta}}{N-1}\left(\frac{N/\bar{\theta} + N/\bar{\theta}}{2}\right) = \frac{(N-1)\,\bar{\theta} - N\bar{\theta} + N\underline{\theta}}{N-1} = \frac{N\underline{\theta} - \bar{\theta}}{N-1} < 0.$$

Thus, the coefficient of x in $F(x)$ is negative, that is, $F(x)$ is decreasing in x. As the RHS of (10) is a constant for given model parameters and e_{In}, there will be at most one n such that

$$\begin{aligned} F(n) \ &\geq \ \frac{N\underline{\theta} - \bar{\theta}}{N-1}\left[e_{In}^2 - (N/\bar{\theta})^2\right]/2 + \lambda, \\ F(n+1) \ &< \ \frac{N\underline{\theta} - \bar{\theta}}{N-1}\left[e_{In}^2 - (N/\bar{\theta})^2\right]/2 + \lambda, \end{aligned} \tag{11}$$

which proves the uniqueness part.

To prove existence, we need to show that for any pair $(e_{In}, \lambda) \in \Omega$ the following two conditions hold:

$$F(3) \ \geq \ \frac{N\underline{\theta} - \bar{\theta}}{N-1}\left[e_{In}^2 - (N/\bar{\theta})^2\right]/2 + \lambda, \tag{12}$$

$$F(N-1) \ < \ \frac{N\underline{\theta} - \bar{\theta}}{N-1}\left[e_{In}^2 - (N/\bar{\theta})^2\right]/2 + \lambda, \tag{13}$$

From the continuity of $F(.)$ it follows that there exists a $n \in [3, N-2]$ such that the system (11) holds, which, in turn, implies that this n is the equilibrium size of the inner club.

We start by showing that inequality (12) holds for any $(e_{In}, \lambda) \in \Omega$. By definition of $F(.)$ inequality (12) can be rewritten as

$$3\left(e_{In} - N/\bar{\theta}\right) - \lambda \geq \theta_3\left[e_{In}^2 - (N/\bar{\theta})^2\right]/2.$$

Solving for e_{In} shows that inequality (12) is satisfied for any (λ, e_{In}) such that

$$\lambda \leq \frac{\theta_3}{2}\left(\frac{3}{\theta_3} - \frac{N}{\bar{\theta}}\right)^2 \tag{14}$$

and

$$e_{In} \in \left[\frac{3}{\theta_3} - \sqrt{\left(\frac{3}{\theta_3} - \frac{N}{\bar{\theta}}\right)^2 - \frac{2\lambda}{\theta_3}}, \frac{3}{\theta_3} + \sqrt{\left(\frac{3}{\theta_3} - \frac{N}{\bar{\theta}}\right)^2 - \frac{2\lambda}{\theta_3}}\right]. \tag{15}$$

The set Ω is clearly a subset of the set defined by (14) and (15) as the set Ω is determined by further restrictions in addition to the restrictions (14) and (15). Thus, as the inequality (12) holds for any pair (e_{In}, λ) satisfying (14) and (15), it also holds for any pair $(e_{In}, \lambda) \in \Omega$.

Now we show that condition (13) holds for any pair $(e_{In}, \lambda) \in \Omega$. Similarly to above, (13) is equivalent to

$$(N-1)\left(e_{In} - N/\bar{\theta}\right) - \lambda < \theta_{N-1}\left[e_{In}^2 - (N/\bar{\theta})^2\right]/2.$$

Solving for e_{In} yields that the inequality (13) is satisfies for any (λ, e_{In}) such that

$$\lambda > \frac{\theta_{N-1}}{2}\left(\frac{N-1}{\theta_{N-1}} - \frac{N}{\overline{\theta}}\right)^2 \tag{16}$$

and

$$e \in (-\infty, -\infty), \tag{17}$$

or for any pair (λ, e_{In}) such that

$$\lambda \leq \frac{\theta_{N-1}}{2}\left(\frac{N-1}{\theta_{N-1}} - \frac{N}{\overline{\theta}}\right)^2 \tag{18}$$

and

$$e_{In} \in \left\{ \begin{array}{l} \left(-\infty, \frac{N-1}{\theta_{N-1}} - \sqrt{\left(\frac{N-1}{\theta_{N-1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{N-1}}}\right) \\ \cup \left(\frac{N-1}{\theta_{N-1}} + \sqrt{\left(\frac{N-1}{\theta_{N-1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{N-1}}}, \infty\right). \end{array} \right. \tag{19}$$

As above, the set Ω is a subset of the set determined by inequalities (16),(17), (18) and (19). Indeed, for both $\lambda > \frac{\theta_{N-1}}{2}\left(\frac{N-1}{\theta_{N-1}} - \frac{N}{\overline{\theta}}\right)^2$ and $\lambda \leq \frac{\theta_{N-1}}{2}\left(\frac{N-1}{\theta_{N-1}} - \frac{N}{\overline{\theta}}\right)^2$ there are additional restrictions imposed on e_{In} in the set Ω. Thus, inequality (13) is satisfies for any pair $(e_{In}, \lambda) \in \Omega$.

A.4 Proof of Proposition 4

An inner organization of size $n \in [3, N-2]$ is a Nash equilibrium if the constraints (4) and (5) are satisfied and $e_{In} > N/\overline{\theta}$ holds. Solving inequality (4) we obtain

$$e_{In} \in \left\{ \begin{array}{ll} \emptyset & \text{for } \lambda > \lambda_n, \\ \left[\frac{n}{\theta_n} - \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}}, \frac{n}{\theta_n} + \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}}\right] & \text{for } \lambda_n \geq \lambda. \end{array} \right. \tag{20}$$

Similarly, inequality (5) yields

$$e_{In} \in \begin{cases} (-\infty, \infty) & \text{for } \lambda > \lambda_{n+1}, \\ \left(-\infty, \frac{n+1}{\theta_{n+1}} - \sqrt{\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{n+1}}}\right] & \\ \cup \left[\frac{n+1}{\theta_{n+1}} + \sqrt{\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{n+1}}}, \infty\right) & \text{for } \lambda_{n+1} \geq \lambda. \end{cases} \tag{21}$$

Before solving the system (20) and (21) and checking whether $e_{In} > N/\overline{\theta}$, we establish two useful results.

Lemma A2: $\lambda_n > \lambda_{n+1}$

Proof. Consider the function

$$\Lambda(x) = \frac{1}{2}\left(\frac{N-x}{N-1}\underline{\theta} + \frac{x-1}{N-1}\overline{\theta}\right)\left(\frac{x}{\left(\frac{N-x}{N-1}\underline{\theta} + \frac{x-1}{N-1}\overline{\theta}\right)} - \frac{N}{\overline{\theta}}\right)^2$$

and note that $\Lambda(n) = \lambda_n$ and $\Lambda(n+1) = \lambda_{n+1}$.

$$\frac{\partial \Lambda(x)}{\partial x} = -\frac{\left(\overline{\theta} - N\underline{\theta}\right)^2}{2\overline{\theta}^2}\frac{2(N-x)\left((N-x)\underline{\theta} + (N+x-2)\overline{\theta}\right)}{(N-1)\left(\frac{N-x}{N-1}\underline{\theta} + \frac{x-1}{N-1}\overline{\theta}\right)^2} < 0 \quad \text{if} \quad 2 \leq x < N$$

Thus $\lambda_n > \lambda_{n+1}$. ∎

Lemma A3: *For every $\lambda \leq \lambda_{n+1}$, it holds that*

$$\frac{n+1}{\theta_{n+1}} - \sqrt{\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{n+1}}} > \frac{n}{\theta_n} - \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}} \tag{22}$$

and

$$\frac{n}{\theta_n} + \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}} > \frac{n+1}{\theta_{n+1}} + \sqrt{\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{n+1}}} \tag{23}$$

Proof. Consider first inequality (22). We define a function $F_1(\lambda)$ on $[0, \lambda_{n+1}]$ where

$$F_1(\lambda) = \frac{n}{\theta_n} - \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}} - \left(\frac{n+1}{\theta_{n+1}} - \sqrt{\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{n+1}}}\right)$$

At $\lambda = 0$

$$F_1(0) = \frac{n}{\theta_n} - \left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right) - \left(\frac{n+1}{\theta_{n+1}} - \left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}\right)\right) = 0$$

Moreover, whenever defined

$$\frac{\partial F_1(\lambda)}{\partial \lambda} = \frac{1}{\theta_n\sqrt{\left(\frac{n}{\theta_n}-\frac{N}{\overline{\theta}}\right)^2-\frac{2\lambda}{\theta_n}}} - \frac{1}{\theta_{n+1}\sqrt{\left(\frac{n+1}{\theta_{n+1}}-\frac{N}{\overline{\theta}}\right)^2-\frac{2\lambda}{\theta_{n+1}}}} < 0$$

Indeed

$$\theta_n\sqrt{\left(\frac{n}{\theta_n}-\frac{N}{\overline{\theta}}\right)^2-\frac{2\lambda}{\theta_n}} > \theta_{n+1}\sqrt{\left(\frac{n+1}{\theta_{n+1}}-\frac{N}{\overline{\theta}}\right)^2-\frac{2\lambda}{\theta_{n+1}}}$$

as

$$\theta_n^2\left(\left(\frac{n}{\theta_n}-\frac{N}{\overline{\theta}}\right)^2-\frac{2\lambda}{\theta_n}\right) - \theta_{n+1}^2\left(\left(\frac{n+1}{\theta_{n+1}}-\frac{N}{\overline{\theta}}\right)^2-\frac{2\lambda}{\theta_{n+1}}\right)$$

$$= \left(\frac{(N-n)\left(\overline{\theta}-N\underline{\theta}\right)}{\overline{\theta}(N-1)}\right)^2 - \left(\frac{(N-n-1)\left(\overline{\theta}-N\underline{\theta}\right)}{\overline{\theta}(N-1)}\right)^2 - 2\lambda\left[\theta_n-\theta_{n+1}\right]$$

$$= \left(\frac{\left(\overline{\theta}-N\underline{\theta}\right)}{\overline{\theta}(N-1)}\right)^2 (2N-2n-1) + 2\lambda\frac{\left(\overline{\theta}-\underline{\theta}\right)}{(N-1)} > 0$$

Thus, $F_1(\lambda)$ is a decreasing function of λ and $F_1(\lambda)|_{\lambda>0} < 0$. This proves inequality (22).

Similarly, to prove inequality (23) we define a function $F_2(\lambda)$ on $[0, \lambda_{n+1}]$ where

$$F_2(\lambda) = \frac{n}{\theta_n} + \sqrt{\left(\frac{n}{\theta_n}-\frac{N}{\overline{\theta}}\right)^2-\frac{2\lambda}{\theta_n}} - \left(\frac{n+1}{\theta_{n+1}} + \sqrt{\left(\frac{n+1}{\theta_{n+1}}-\frac{N}{\overline{\theta}}\right)^2-\frac{2\lambda}{\theta_{n+1}}}\right)$$

From Lemma A1 it follows that at $\lambda = 0$

$$F_2(0) = 2\left(\frac{n}{\theta_n} - \frac{n+1}{\theta_{n+1}}\right) > 0$$

Moreover,

$$\frac{\partial F_2(\lambda)}{\partial \lambda} = -\frac{1}{\theta_n\sqrt{\left(\frac{n}{\theta_n}-\frac{N}{\overline{\theta}}\right)^2-\frac{2\lambda}{\theta_n}}} + \frac{1}{\theta_{n+1}\sqrt{\left(\frac{n+1}{\theta_{n+1}}-\frac{N}{\overline{\theta}}\right)^2-\frac{2\lambda}{\theta_{n+1}}}} = -\frac{\partial F_1(\lambda)}{\partial \lambda} > 0$$

Thus, $F_2(\lambda)$ is a increasing function of λ and $F_2(\lambda)|_{\lambda>0} > 0$. This is equivalent to

$$\frac{n}{\theta_n} + \sqrt{\left(\frac{n}{\theta_n}-\frac{N}{\overline{\theta}}\right)^2-\frac{2\lambda}{\theta_n}} > \frac{n+1}{\theta_{n+1}} + \sqrt{\left(\frac{n+1}{\theta_{n+1}}-\frac{N}{\overline{\theta}}\right)^2-\frac{2\lambda}{\theta_{n+1}}},$$

thereby proving inequality (23). ■

Using Lemma A2 and A3, we can now describe the entire set of joint solutions for (20) and (21). For $\lambda > \lambda_n$, the inequality (20) and hence the system has no solution.

For $\lambda_n \geq \lambda > \lambda_{n+1}$, the intersection of (20) and (21) results in

$$e_{In} \in \left[\frac{n}{\theta_n} - \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}}, \frac{n}{\theta_n} + \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}}\right] \tag{24}$$

According to Lemma A3,

$$\frac{n+1}{\theta_{n+1}} - \sqrt{\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{n+1}}} > \frac{n}{\theta_n} - \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}}$$

and

$$\frac{n}{\theta_n} + \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}} > \frac{n+1}{\theta_{n+1}} + \sqrt{\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{n+1}}}.$$

Hence, the intersection of (20) and (21) for $\lambda_{n+1} \geq \lambda > 0$ is

$$e_{In} \in \left\{ \begin{array}{l} \left[\frac{n}{\theta_n} - \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}}, \frac{n+1}{\theta_{n+1}} - \sqrt{\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{n+1}}}\right) \\ \cup \left(\frac{n+1}{\theta_{n+1}} + \sqrt{\left(\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_{n+1}}}, \frac{n}{\theta_n} + \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}}\right] \end{array} \right. \tag{25}$$

For each $\lambda > 0$

$$\frac{n}{\theta_n} - \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2 - \frac{2\lambda}{\theta_n}} > \frac{n}{\theta_n} - \sqrt{\left(\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right)^2} = \frac{N}{\overline{\theta}}.$$

Thus, $e_{In} > N/\overline{\theta}$ is satisfied for any e_{In} belonging to the sets (24) and (25).

Imposing the restriction $e_{In} > N/\overline{\theta}$ on the intersection of (20) and (21) for $\lambda = 0$ gives

$$e_{In} \in \left[2\frac{n+1}{\theta_{n+1}} - \frac{N}{\overline{\theta}}, 2\frac{n}{\theta_n} - \frac{N}{\overline{\theta}}\right]. \tag{26}$$

Thus, (24), (25) and (26) describe the set Ω_n. This concludes the proof of Proposition 4.

A.5 Proof of Lemma 3

Existence of the inner organization for any draw e_{In} follows from the assumption that the inner club must have two members. Solving the participation constraint (2) for

$N = 3$ yields the maximum e_{In} compatible with all three members exerting the same effort. As for the divided organization, solving inequality (6) yields

$$e_{In} \in \begin{cases} \emptyset & \text{for } \lambda > \lambda_2 \\ \left[\frac{2}{\theta_2} - \sqrt{\left(\frac{2}{\theta_2} - \frac{3}{\bar{\bar{\theta}}} \right)^2 - \frac{4\lambda}{\theta_2}}, \frac{2}{\theta_2} + \sqrt{\left(\frac{2}{\theta_2} - \frac{3}{\bar{\bar{\theta}}} \right)^2 - \frac{4\lambda}{\theta_2}} \right] & \text{for } \lambda \leq \lambda_2 \end{cases}$$

where $\lambda_2 = \frac{\theta_2}{4} \left(\frac{2}{\theta_2} - \frac{3}{\bar{\bar{\theta}}} \right)^2$. Solving inequality (7) we obtain

$$e_{In} \in \left(-\infty, \frac{3}{\bar{\bar{\theta}}} - \sqrt{\frac{4\lambda}{\bar{\theta}}} \right] \cup \left[\frac{3}{\bar{\bar{\theta}}} + \sqrt{\frac{4\lambda}{\bar{\theta}}}, \infty \right).$$

For $\lambda \leq \lambda_2$, the solution to the system (6), (7) and the constraint $e_{In} > e_{Out}$ is

$$e_{In} \in \left[\max \left\{ \frac{3}{\bar{\bar{\theta}}} + \sqrt{\frac{4\lambda}{\bar{\theta}}}, \frac{2}{\theta_2} - \sqrt{\left(\frac{2}{\theta_2} - \frac{3}{\bar{\bar{\theta}}} \right)^2 - \frac{4\lambda}{\theta_2}} \right\}, \frac{2}{\theta_2} + \sqrt{\left(\frac{2}{\theta_2} - \frac{3}{\bar{\bar{\theta}}} \right)^2 - \frac{4\lambda}{\theta_2}} \right].$$

A.6 Proof of Lemma 4:

Existence of the initial organization follows from the same argument as in proof of Lemma 3. is trivial. The effort in the reformed organization must satisfy the inequality

$$3\left(e_{In} - 3/\theta_2\right) - \lambda \geq \theta_3 \left[e_{In}^2 - (3/\theta_2)^2 \right] /2.$$

as under the majority rule $e_{Out} = 3/\theta_2$. Solving it for e_{In}, we obtain

$$e_{In} \in \left[\frac{3}{\bar{\bar{\theta}}} - \sqrt{9 \left(\frac{\bar{\theta} - \theta_2}{\bar{\theta}\theta_2} \right)^2 + \frac{4\lambda}{\bar{\theta}}}, \frac{3}{\bar{\bar{\theta}}} + \sqrt{9 \left(\frac{\bar{\theta} - \theta_2}{\bar{\theta}\theta_2} \right)^2 + \frac{4\lambda}{\bar{\theta}}} \right].$$

Moreover, the constraint $e_{In} > e_{Out} = 3/\theta_2$ must hold. As for non-negative λ

$$\frac{3}{\bar{\bar{\theta}}} + \sqrt{9 \left(\frac{\bar{\theta} - \theta_2}{\bar{\theta}\theta_2} \right)^2 + \frac{4\lambda}{\bar{\theta}}} > \frac{3}{\bar{\bar{\theta}}} + \sqrt{9 \left(\frac{\bar{\theta} - \theta_2}{\bar{\theta}\theta_2} \right)^2} = \frac{3}{\theta_2}$$

and

$$\frac{3}{\bar{\bar{\theta}}} - \sqrt{9 \left(\frac{\bar{\theta} - \theta_2}{\bar{\theta}\theta_2} \right)^2 + \frac{4\lambda}{\bar{\theta}}} < \frac{3}{\bar{\bar{\theta}}} < \frac{3}{\theta_2},$$

the reformed organization can form for the effort levels

$$e_{In} \in \left(3/\theta_2, \frac{3}{\bar{\bar{\theta}}} + \sqrt{9 \left(\frac{\bar{\theta} - \theta_2}{\bar{\theta}\theta_2} \right)^2 + \frac{4\lambda}{\bar{\theta}}} \right].$$

We prove the inexistence of a divided organization by contradiction. If there are only 3 types, the only inner organization possible is the one of size $n = 2$. If it forms, the marginal member is type 2, and the following conditions must hold:

$$2(e_{In} - e_{Out}) - 2\lambda > \theta_2 \left[e_{In}^2 - [e_{Out}]^2\right] /2 , \tag{27}$$

and

$$e_{In} > e_{Out},$$

where $e_{Out} = 3/\theta_2$, the best preferred outer club effort level chosen by the median type - type 2. Setting $e_{In} = 3/\theta_2 + \delta, \delta > 0$ and inserting it in equation (27) yields

$$2\delta - 2\lambda \geq \frac{\theta_2}{2}\left(2\delta\frac{3}{\theta_2} + \delta^2\right),$$

or equivalently,

$$0 \geq \delta + 2\lambda + \frac{\theta_2}{2}\delta^2,$$

which never holds, and consequently the constraint (27) is also violated.

A.7 Proof of Lemma 5:

For $\lambda = 0$, the solution of (6), (7) and the constraint $e_{In} > 3/\bar{\theta}$ yields

$$e_{In} \in \left(\frac{3}{\bar{\bar{\theta}}}, \frac{4}{\theta_2} - \frac{3}{\bar{\bar{\theta}}}\right]. \tag{28}$$

Denote the upper bound of the solution to (28) by

$$\overline{e}_{In}^{\theta_2} = \frac{4}{\theta_2} - \frac{3}{\bar{\bar{\theta}}}$$

The social welfare under the unanimity rule in case an inner club forms is equal to

$$\begin{aligned} W_{Un}^{n=2}(e_{In}) &= \sum_{i=1}^{2}\left(2\left(e_{In} - \frac{3}{\bar{\bar{\theta}}}\right) + 3\frac{3}{\bar{\bar{\theta}}} - \frac{\theta_i}{2}e_{In}^2\right) + \left(3\frac{3}{\bar{\bar{\theta}}} - \frac{\overline{\theta}}{2}\left(\frac{3}{\bar{\bar{\theta}}}\right)^2\right) \\ &= 4e_{In} + \frac{21}{2\bar{\bar{\theta}}} - \frac{e_{In}^2}{2}\left(\underline{\theta} + \theta_2\right) \end{aligned}$$

The social welfare under the unanimity rule if an inner club does not form is

$$\begin{aligned} W_{Un}^{n=\emptyset} &= 3 * 3\frac{3}{\bar{\bar{\theta}}} - \frac{\left(\underline{\theta} + \theta_2 + \overline{\theta}\right)}{2}\left(\frac{3}{\bar{\bar{\theta}}}\right)^2 \\ &= \frac{1}{2}\left(\frac{3}{\bar{\bar{\theta}}}\right)^2\left(5\bar{\bar{\theta}} - \left(\underline{\theta} + \theta_2\right)\right). \end{aligned}$$

Note that for any e_{In} satisfying (28)

$$W_{Un}^{n=2}(e_{In}) > W_{Un}^{n=\emptyset} \tag{29}$$

as in the absence of deadweight loss ($\lambda = 0$) the utility of types $\underline{\theta}$ and θ_2 is higher when they join the inner club, and the utility of the least productive type does not differe in these two cases.

Since an inner club never emerges under the majority rule (4), the social welfare is equal to

$$\begin{aligned} W_{Maj} &= 3 * 3\frac{3}{\theta_2} - \frac{\left(\underline{\theta} + \theta_2 + \overline{\theta}\right)}{2}\left(\frac{3}{\theta_2}\right)^2 \\ &= \frac{1}{2}\left(\frac{3}{\theta_2}\right)^2\left(5\theta_2 - \left(\underline{\theta} + \overline{\theta}\right)\right). \end{aligned}$$

The following proof is divided into two parts. Part i) shows that if the distribution of types is equidistant ($\alpha = 1/2$) or sufficiently close to it, the social welfare under majority rule exceeds the maximum social welfare under the unanimity rule. Part ii) establishes that if the median type is very productive (α close to 1), the social welfare under the unanimity rule is never less than the one under the majority rule.

i) We start by showing that for $\alpha = 1/2$ the majority rule dominates:

$$W_{Maj} > \max_{e_{In}} W_{Un}^{n=2}(e_{In}).$$

Lemma A4: *For $\alpha = 1/2$ the effort level that maximizes $W_{Un}^{n=2}(e_{In})$ does not support an inner club $n = 2$.*

Proof. The function $W_{Un}^{n=2}(e_{In})$ is a quadratic function of e_{In} and reaches its maximum at $e_{In}^{*} = 4/\left(\underline{\theta} + \theta_2\right)$. However, e_{In}^{*} exceeds $\overline{e}_{In}^{\theta_2}$:

$$\begin{aligned} e_{In}^{*} - \overline{e}_{In}^{\theta_2} &= \frac{8}{3\underline{\theta} + \overline{\theta}} - \frac{8}{\underline{\theta} + \overline{\theta}} + \frac{3}{\overline{\theta}} \\ &= \frac{3\overline{\theta}^2 + 9\underline{\theta}^2 - 4\underline{\theta}\overline{\theta}}{\left(3\underline{\theta} + \overline{\theta}\right)\left(\underline{\theta} + \overline{\theta}\right)\overline{\theta}} \\ &= \frac{2\overline{\theta}^2 + 5\underline{\theta}^2 + \left(2\underline{\theta} - \overline{\theta}\right)^2}{\left(3\underline{\theta} + \overline{\theta}\right)\left(\underline{\theta} + \overline{\theta}\right)\overline{\theta}} > 0 \end{aligned}$$

■

Thus, the maximum social welfare in case inner club forms is achieved under $e_{In} = \overline{e}^{\theta_2}_{In}$, and is equal to

$$W^{n=2}_{Un}\left(\overline{e}^{\theta_2}_{In}\right) = 4\left(\frac{8}{\underline{\theta}+\overline{\theta}} - \frac{3}{\overline{\overline{\theta}}}\right) + \frac{21}{2\overline{\overline{\theta}}} - \frac{1}{4}\left(\frac{8}{\underline{\theta}+\overline{\theta}} - \frac{3}{\overline{\overline{\theta}}}\right)^2 \left(3\underline{\theta}+\overline{\theta}\right)$$

Consider the difference between the social welfare under the majority rule and the unanimity rule:

$$\begin{aligned}
W_{Maj} - W_{Un} &\geq W_{Maj} - W^{n=2}_{Un}\left(\overline{e}^{\theta_2}_{In}\right) \\
&= \frac{27}{\underline{\theta}+\overline{\theta}} - \left[4\left(\frac{8}{\underline{\theta}+\overline{\theta}} - \frac{3}{\overline{\overline{\theta}}}\right) + \frac{21}{2\overline{\overline{\theta}}} - \frac{1}{4}\left(\frac{8}{\underline{\theta}+\overline{\theta}} - \frac{3}{\overline{\overline{\theta}}}\right)^2 \left(3\underline{\theta}+\overline{\theta}\right)\right] \\
&= \frac{1}{4}\left(\overline{\theta}-\underline{\theta}\right)\frac{48\overline{\theta}\underline{\theta} - 27\underline{\theta}^2 + 11\overline{\theta}^2}{\overline{\theta}^2\left(\overline{\theta}+\underline{\theta}\right)^2} \\
&> \frac{1}{4}\left(\overline{\theta}-\underline{\theta}\right)\frac{48\underline{\theta}^2 - 27\underline{\theta}^2 + 11\overline{\theta}^2}{\overline{\theta}^2\left(\overline{\theta}+\underline{\theta}\right)^2} \qquad (30) \\
&= \frac{1}{4}\left(\overline{\theta}-\underline{\theta}\right)\frac{21\underline{\theta}^2 + 11\overline{\theta}^2}{\overline{\theta}^2\left(\overline{\theta}+\underline{\theta}\right)^2} > 0
\end{aligned}$$

where the inequality in (30) follows from $\overline{\theta} > \underline{\theta}$. Hence, $W_{Maj} > W_{Un}$ for $\alpha = 1/2$, and by continuity we can conclude that the result also holds for sufficiently small $\lambda > 0$ and $\delta > 0$, such that $\alpha = 1/2 + \delta$.

ii) Now we want to show that for α sufficiently close to 1 and sufficiently heterogeneous types unanimity is more efficient than majority. Let's compare the social welfare under majority rule with the social welfare of the original club (that is, the welfare under unanimity rule when the inner club does not form):

Lemma A5. *For $\lambda = 0$ and $\overline{\theta} > 2\underline{\theta}$, there exists a non-empty set $F(\underline{\theta}, \overline{\theta}) = \{\alpha : \alpha \in (\alpha_F, 1]\}$ with*

$$\alpha_F \doteqdot \frac{-(2\overline{\theta}-\underline{\theta}) + \sqrt{12\overline{\theta}^2 - 12\overline{\theta}\underline{\theta} + \underline{\theta}^2}}{2(\overline{\theta}-\underline{\theta})} < 1$$

such that

$$W^{n=\emptyset}_{Un} > W_{Maj}. \qquad (31)$$

Proof. As $\theta_2 = \alpha\underline{\theta} + (1-\alpha)\overline{\theta}$, $W_{Un}^{n=\emptyset} - W_{Maj}$ can be rewritten as

$$\begin{aligned} W_{Un}^{n=\emptyset} - W_{Maj} &= \frac{1}{2}\left(\frac{3}{\overline{\theta}}\right)^2 \left(5\overline{\theta} - (\underline{\theta} + \theta_2)\right) - \frac{1}{2}\left(\frac{3}{\theta_2}\right)^2 \left(5\theta_2 - (\underline{\theta} + \overline{\theta})\right) \\ &= \frac{9\alpha}{2\overline{\theta}^2\theta_2^2}\left(\overline{\theta} - \underline{\theta}\right)^2 \left[\left(\overline{\theta} - \underline{\theta}\right)\alpha^2 + \left(2\overline{\theta} - \underline{\theta}\right)\alpha - 2\overline{\theta}\right] \end{aligned}$$

So $W_{Un}^{n=\emptyset} - W_{Maj} > 0$ for the values of α satisfying

$$\left(\overline{\theta} - \underline{\theta}\right)\alpha^2 + \left(2\overline{\theta} - \underline{\theta}\right)\alpha - 2\overline{\theta} > 0$$

Since $\alpha \in [0,1]$, the set $F(\underline{\theta}, \overline{\theta})$ of α such that inequality (31) holds is non-empty iff

$$\frac{-(2\overline{\theta} - \underline{\theta}) + \sqrt{12\overline{\theta}^2 - 12\overline{\theta}\underline{\theta} + \underline{\theta}^2}}{2(\overline{\theta} - \underline{\theta})} < 1$$

which is equivalent to

$$\left(\overline{\theta} - 2\underline{\theta}\right)\left(\overline{\theta} - \underline{\theta}\right) > 0.$$

■

As shown above (see (29)), $W_{Un}^{n=\emptyset} < W_{Un}^{n=2}(e_{In})$. Thus, it follows from Lemma A.7 and inequality (29) that if $\alpha \in F(\underline{\theta}, \overline{\theta})$ the social welfare under unanimity (no matter if the inner club forms or does not) exceeds the one under majority.

$$W_{Un}^{n=2}(e_{In}) > W_{Un}^{n=\emptyset} > W_{Maj}.$$

By continuity, this result also holds for sufficiently small positive λ.

A.8 Inexistence of Divided Organizations under Majority

Proof by contradiction. If an equilibrium with an inner organization of size $n < N$ exists,

$$n(e_{In} - e_{Out}) - \lambda > \theta_n \left[e_{In}^2 - e_{Out}^2\right]/2 \tag{32}$$

and

$$e_{In} > e_{Out}$$

must hold where $e_{Out} = 2N/(\underline{\theta} + \overline{\theta})$. Setting $e_{In} = 2N/(\underline{\theta} + \overline{\theta}) + \delta$ and inserting it in equation (32) yields

$$n\delta - \lambda \geq \frac{\theta_n}{2}\left(2\delta\frac{2N}{\underline{\theta} + \overline{\theta}} + \delta^2\right),$$

or equivalently,

$$\delta\left(\frac{n}{\theta_n} - \frac{2N}{\underline{\theta}+\overline{\theta}}\right) \geq \frac{\delta^2}{2} + \frac{\lambda}{\theta_n} > 0.$$

This condition can only be satisfied if

$$\frac{n}{\theta_n} > \frac{2N}{\underline{\theta}+\overline{\theta}}. \tag{33}$$

Given $N < \overline{\theta}/\underline{\theta}$, n/θ_n decreases with n (Lemma A1) and

$$\max_n \frac{n}{\theta_n} = \frac{2}{\theta_2} = \frac{2(N-1)}{(N-2)\underline{\theta}+\overline{\theta}}.$$

Hence, if the condition (33) holds for some n, it must be true that

$$\frac{2}{\theta_2} > \frac{2N}{\underline{\theta}+\overline{\theta}}.$$

Inserting the explicit expression for $2/\theta_2$ and rearranging yields

$$(N-1)(\underline{\theta}+\overline{\theta}) > N\left[(N-2)\underline{\theta}+\overline{\theta}\right] \Leftrightarrow 0 > (N^2-3N+1)\underline{\theta}+\overline{\theta}.$$

Since $N^2 - 3N + 1 > 0$ for $N > 2$, condition (33) never holds, and consequently the constraint (32) is also violated.

CHAPTER 4

Protection for Sale to Oligopolists*

Abstract

This paper modifies Grossman and Helpman's canonical "Protection for Sale" model by allowing demand linkages and oligopolistic competition. It shows that increased substitutability between products weakens interest groups' incentives to lobby. For the case of one organized and one unorganized industry, we obtain a particularly simple result: as product substitutability increases, the protection of the organized industry's product falls, whereas the protection of the unorganized sector's product increases.

Empirical studies of the "Protection for Sale" model have suggested that the U.S. government's trade policy decisions are overwhelmingly determined by a concern for welfare maximization; the alternative interpretation of the paper is that the original model overstates the lobby groups' desire for protection.

1 Introduction

Recent theories of endogenous trade policy suggest that trade policy is determined through interaction between organized lobby groups and privately interested policy-makers, rather than by a benevolent social-welfare maximizer. While there are many variations on the theme (see Rodrik (1995) for a review), Grossman and Helpman's (1994) "Protection for Sale" model has been the most influential by far. Grossman

*I am grateful to Tore Ellingsen and Victor Polterovich for advice and encouragement. I also thank Gene Grossman, Giovanni Maggi and Torsten Persson for valuable comments, as well as seminar participants at Stockholm School of Economics and Princeton University. Jan Wallander's and Tom Hedelius' Research Foundation is gratefully acknowledged for financial support. All remaining errors are my own.

and Helpman (henceforth GH) explicitly describe a mechanism through which interest groups' contributions influence the policy-maker's decision for trade protection, providing micro-foundations for the previous approaches. The model's predictions for the equilibrium protection pattern relate the industry's trade tariff to import demand elasticity, state of organization, import penetration and other variables, thereby providing a coherent framework for empirical testing.

As is well known, GH neglect some important issues. In particular, they abstract from production linkages and strategic market interactions. Indeed, factor-specific production in GH implies that different lobbies do not compete in the factor market. Hence, their interests are only opposed to the extent that each industry wants to increase its profits by raising the price of its own good, while all other organized industries aim at reducing the same price in order to increase their members' consumer surplus. That is, political competition among the lobbies arises purely from the desire of the members of different industry groups to protect their interests as ordinary consumers. This feature of GH has caused concerns about the interpretation of the model.[1] To add a more realistic justification for the political rivalry among the organized groups, Gawande and Bandyopadhyay (2000) introduce supply-side interactions through a single importable intermediate input. Cadot, de Melo and Ollarreaga (2001) strengthen the degree of inter-industry interactions even further, assuming that the industries compete for a common scare resource and each good is produced using other goods as intermediate inputs. These extensions help obtain predictions consistent with the observed stylized facts, e.g., escalation of protection rates with the degree of processing, or higher average protection for poor countries.

This paper instead focuses on demand-side interactions, addressing the rivalry between the organized groups that arises due to the substitutability between goods. It studies the impact of demand linkages on the determination of trade policy, and the intensity of inter-industry lobbying competition.

The consumers' utility function in the original GH model is assumed to be quasilinear, implying that the demand for each product is independent of the prices of the other products. Besides, the GH model considers a small and open economy, so that the producers face perfectly elastic demand. Thus, the prices in other industries do not affect the producers' incentives to lobby. To address the inter-industry competition in lobbying and its impact on the trade policy, resulting from the demand-side interac-

[1] See e.g. Baldwin and Robert-Nicoud (2006).

tions, I relax these assumptions and allow for both demand linkages and imperfectly competitive industries. More precisely, I consider a utility function that admits cross-price effects on demand and assume that each good is produced by an international oligopoly and sold in internationally segmented markets.

The first part of the paper shows that the presence of substitutes may reduce interest group incentives to lobby due to competition in the goods market. If demands are interdependent, an increase in the price of a good causes demand to shift towards its substitutes. A decrease in the price of the substitute has a similar effect. The interest group takes this shift into account when lobbying the government to increase the price of its own good (as a producer receiving profit from selling this good) and reduce the price of all other goods (as a consumer who wants to maximize her utility). Therefore, the lobbying strategy of an organized industry becomes less aggressive. To put it differently, introducing substitutability produces more aligned interests between the interest groups. As a result, other things equal, economies producing closer substitutes (or having a more competitive industry structure) should experience more moderate protection rates. That is, with an increase in the degree of substitutability, the protection of the organized industries falls, and the protection of the non-organized industries increases, relative to the first-best benchmark.

The result suggests a new explanation for an observed empirical puzzle. Studies devoted to empirical testing of the GH theory find that the government puts very low weight on the campaign contributions relative to the welfare loss. That is, the government behaves almost as a social welfare maximizer and trade policy deviates surprisingly little from the first-best,[2] thus causing a concern about the empirical significance of the GH model. E.g., Gawande and Krishna (2002) write that "..it is enough to cast doubt on the value of viewing trade policy determination through this political economy lens". There have been several attempts at modifying the model to explain low protection rates. For example, Gawande and Bandyopadhyay (2000) introduce political competition between the upstream and downstream producers, and Gawande, Krishna and Robbins (2004) introduce counter-lobbying by the foreign organized groups. This paper provides an alternative explanation, suggesting that smaller deviations from the first-best protection rates may result from the weaker incentives to lobby caused by the substitutability effects.

The second part of the paper addresses the impact of product substitutability on

[2] E.g. Goldberg and Maggi (1999); Gawande and Bandyopadhyay (2000).

lobby formation. The original GH model assumes an exogenous lobby group structure. A natural question to ask is thus why are some industries organized and others not. In the GH model, the interests of different industry groups are opposed to each other, so if an additional lobby formation stage is introduced in their model, all industries would get organized (in the absence of the lobby formation costs).[3]

However, in the presence of demand linkages, at a sufficiently high degree of substitution a non-organized industry becomes protected without paying for it. The reason is that the organized industry cannot lobby to increase its own price without losing a substantial part of its consumers switching to the cheaper (but similar) good. For the same reason, it cannot lobby for dropping the price of the substitute. Thus, the non-lobbying industry gets a free ride on the lobbying industry efforts. If instead both these substitute-producing industries are organized, they both contribute to the government for (potentially higher) protection. Comparing these two outcomes, it turns out that the industry may better off not being organized. As a result, with endogenous lobby formation, fewer industries get organized and lobbying becomes less intense. We find this free-riding effect to be present for industries with relatively dispersed ownership, which is in line with Olson's (1965) interest group size hypothesis.

This paper is not the first to endogenize lobby formation in the GH model or discuss the possibility of free-riding. Mitra (1999) adds an initial stage to the GH model, letting the owners of each specific factor decide in Nash equilibrium whether it is profitable to incur a fixed cost of forming a lobby. Magee (2002) employs a two-stage game, where in the first stage, industry representatives and the policy maker determine the tariff schedule and, in the second stage, every firm in the industry decides whether to contribute to the lobbying effort (detecting is infinitely punished). Both these papers address the collective action problem at the intra-industry level. We, instead, relate lobby formation to the inter-industry demand links, thereby avoiding the issue of exogenous fixed costs.

The remainder of the paper is organized as follows: Section 2 describes the model setup, Section 3 discusses the equilibrium structure of protection, Section 4 analyzes the effect of the degree of substitution on the extent of protection, Section 5 studies the question of the lobby formation in the presence of substitute goods and Section 6 concludes.

[3]Like GH, we neglect any intra-industry organizational conflict.

2 The model

There are $m+1$ goods being produced in an open economy. The domestic price of good i is denoted p_i. Good 0 is taken to be a numeraire, so $p_0 = 1$. The individuals populating the economy have identical preferences represented by the quasi-linear utility function

$$U(x_0, x_1, ..., x_m) = x_0 + \hat{U}(x_1, ..., x_m), \tag{1}$$

where x_i denotes consumption of good i. The original GH model uses the separable utility function

$$U(x_0, x_1, ..., x_n) = x_0 + \sum_{i=1}^{n} \hat{u}_i(x_1, ..., x_n),$$

which implies that demand functions are independent of the prices of the other goods. I relax this assumption allowing for the cross-price effects. More precisely, I adopt the quadratic sub-utility function:

$$\hat{U}(x_1, ..., x_m) = \sum_{k=1}^{m} x_k - 1/2 \sum_{k=1}^{m} x_k^2 - \sigma \sum_{k=1}^{m} \sum_{j=k+1}^{m} x_k x_j,$$

where $\sigma \in [0,1]$ reflects the substitutability between goods $x_1, ...x_m$. As is well known, the associated inverse demand function for each non-numeraire good i is

$$p_i = 1 - x_i - \sigma \sum_{j \neq i}^{m} x_j.$$

For $\sigma \in (0,1)$, considering only interior solutions, the resulting domestic demand function for good i is

$$d_i(p_1, ..., p_m) = \frac{(1-\sigma) - ((m-2)\sigma + 1)\, p_k + \sigma \sum_{j \neq k}^{m} p_j}{((m-1)\sigma + 1)\,(1-\sigma)}.$$

In words, domestic demand is linear in prices of non-numeraire goods, decreasing in the price of the own good and increasing in the prices of the other goods. That is, good i and goods $1, 2..i-1, i+1, ..m$ are imperfect substitutes.

Due to the quasi-linearity of the utility function, any income effect is totally captured by the consumption of good 0. That is, when optimally spending an amount E, each individual obtains demand functions

$$x_i = d_i(\mathbf{p}), \;\; i \in \{1, ..., m\}$$
$$x_0 = E - \sum_{k=1}^{m} p_k d_k(\mathbf{p}),$$

where $\mathbf{p}$ is the vector of domestic prices $(p_1, ..., p_m)$. The associated indirect utility function is

$$V(\mathbf{p}) = E - \sum_{k=1}^{m} p_k d_k(\mathbf{p}) + \hat{U}(d_1(\mathbf{p}), ..., d_m(\mathbf{p})).$$

The numeraire good is produced using labor only, with an input-output coefficient of 1. As the non-numeraire good is freely traded in a perfectly competitive international market, the wage rate in this economy is equal to 1. The other m goods are produced by a CRS technology using labor alone, but are sold at internationally segmented oligopolistic markets. More precisely, good k is supplied by $n_k > 0$ identical domestic firms and $n_k^* > 0$ identical international firms, competing in quantities. Thus, the total number of firms operating in sector k is given by

$$N_k = n_k + n_k^*.$$

It takes c_k units of labor to produce one unit of good k, both at home and abroad.[4] We assume that each firm is a pure profit maximizer and that the number of firms is fixed, so that no entry or exit decisions are taken.

Each individual is endowed with some labor and may also own some claims to the profit of a firm in at most one industry. These claims are indivisible and non-tradable.

The government introduces trade taxes and/or subsidies and redistributes the resulting net revenue from all these taxes equally among the voting population. For each non-numeraire good, the domestic and the international market are segmented and production is characterized by the constant marginal costs, c_k. As a result, firms' production decisions are made separately for the domestic and the international market. The trade tax revenue is separable in the import and export tax and, as a result, so is consumer welfare. Thus, the government can set import and export trade policies separately. In this paper, we concentrate on the government decision about import tariffs, which determines the production and consumption in the domestic market. We also assume that the foreign government does not impose any export tariffs on its firm, so that in the absence of trade intervention of the domestic government, firms at home and abroad are equally efficient. As we do not study the strategic interaction between the home government and foreign governments in setting trade policy, this assumption does not affect the model's qualitative results.

[4]The assumption that the domestic and **the** foreign firm have the same unit cost c_k is made for computational convenience. It does not change the predictions of the model.

We denote the trade tariff in sector k by τ_k. Then, in Cournot-Nash equilibrium, each of the n_k domestic firms in sector k solves the profit maximization problem

$$\max_{q_k^i} \pi_k^i = q_k^i \left[1 - \sum_{j=1}^{n_k} q_k^j - \sum_{j=1}^{n_k^*} q_k^{j*} - \sigma \sum_{s \neq k} \left(\sum_{j=1}^{n_s} q_s^j + \sum_{j=1}^{n_s^*} q_s^{j*} \right) \right] - c_k q_k^i, \tag{2}$$

where q_k^i denotes the production of domestic firm i in sector k, and q_k^{j*} denotes the production of foreign firm i in sector k. Similarly, each of the n_k^* foreign firms in sector k solves the problem

$$\max_{q_k^{i*}} \pi_k^{i*} = q_k^{i*} \left[1 - \sum_{j=1}^{n_k} q_k^j - \sum_{j=1}^{n_k^*} q_k^{j*} - \sigma \sum_{s \neq k} \left(\sum_{j=1}^{n_s} q_s^j + \sum_{j=1}^{n_s^*} q_s^{j*} \right) \right] - (c_k + \tau_k)\, q_k^{j^*}. \tag{3}$$

Solving the system of these equations for all sectors yields the equilibrium quantities produced by each firm. The symmetry assumption implies that firms within each sector choose the same equilibrium output at home,

$$q_k^i(\boldsymbol{\tau}) = q_k^j(\boldsymbol{\tau}) \equiv q_k(\boldsymbol{\tau}),$$

and abroad

$$q_k^{i*}(\boldsymbol{\tau}) = q_k^{j*}(\boldsymbol{\tau}) \equiv q_k^*(\boldsymbol{\tau}),$$

where $\boldsymbol{\tau}$ denotes the vector of the trade tariffs $(\tau_1, ...\tau_m)$. Similarly, in industry k, each firm in the same country earns the same profits, $\pi_k(\boldsymbol{\tau})$ for a domestic firm and $\pi_k^*(\boldsymbol{\tau})$ for a foreign firm, respectively. Moreover, from the first-order conditions of the profit maximization problems for the domestic firm (2) and the foreign firm (3) in sector k, it follows that the profit is given by

$$\begin{aligned} \pi_k(\boldsymbol{\tau}) &= q_k^2, \\ \pi_k^*(\boldsymbol{\tau}) &= (q_k^*)^2 . \end{aligned} \tag{4}$$

Note also that the market clearing condition implies that

$$d_k(\mathbf{p}(\boldsymbol{\tau})) = n_k q_k(\boldsymbol{\tau}) + n_k^* q_k^*(\boldsymbol{\tau}).$$

The amount of the tax revenue collected by the government (and redistributed to the citizens) is equal to

$$r(\boldsymbol{\tau}) = \sum_{k=1}^{m} \tau_k n_k^* q_k^*(\boldsymbol{\tau}). \tag{5}$$

Thus, individual income is the sum of wages, government transfers and possibly claims to a domestic firm's profit.

The owners of firms in the same industry may choose to organize and form a lobby group trying to influence the government in its decision about trade policies. The joint welfare of the members of such a group i comprising share α_i of the total population is

$$W_i(\boldsymbol{\tau}) = l_i + n_i\pi_i(\boldsymbol{\tau}) + \alpha_i \left[r(\boldsymbol{\tau}) - \sum_{k=1}^{m} p_k d_k(\mathbf{p}) + \hat{U}(d_1(\mathbf{p}(\boldsymbol{\tau})), ..., d_m(\mathbf{p}(\boldsymbol{\tau}))) \right], \quad (6)$$

where l_i denotes the total labor endowment of group i members. Note that the vector of equilibrium prices is, in turn, a function of trade tariffs.

Each lobby i may contribute to the government an amount $C_i(\boldsymbol{\tau})$ conditional on the trade policy vector, or, equivalently, the domestic price vector.

The objective function of the government is

$$G(\boldsymbol{\tau}) = \sum_{i \in L} C_i(\boldsymbol{\tau}) + aW(\boldsymbol{\tau}),$$

where L is the exogenously given set of organized sectors, $a > 0$ is the weight the government attaches to aggregate welfare in the economy, and

$$W(\boldsymbol{\tau}) = l + \sum_{k=1}^{m} n_k \pi_k(\boldsymbol{\tau}) + \left[r(\boldsymbol{\tau}) - \sum_{k=1}^{m} p_k d_k(\mathbf{p}(\boldsymbol{\tau})) + \hat{U}(d_1(\mathbf{p}(\boldsymbol{\tau})), ..., d_m(\mathbf{p}(\boldsymbol{\tau}))) \right] \quad (7)$$

is aggregate welfare in the economy.

Trade policy is determined in a two-stage game. In the first stage, lobbies simultaneously announce their contribution schedules, i.e., the amount contributed as a function of the import tariffs vector (note that *each* strategy of a lobby is a *function* on the price simplex). In the second stage, the government chooses policy by maximizing its objective function over the suggested contribution schedules.

The resulting equilibrium is a set of contribution functions, each maximizing the welfare of the respective lobby members given the schedules of all other lobbies and the anticipated tariff decision of the government and the price vector maximizing the government's objective function under these contribution schedules.

3 The structure of protection

The problem has the structure of a menu-auction, as characterized by Bernheim and Whinston (1986). Hence, if contribution schedules are differentiable, they are *locally* *truthful* around equilibrium, i.e., a marginal change in each lobby's contribution to the government resulting from a small policy change is exactly equal to the respective marginal change of this lobby's welfare.

Local truthfulness combined with the optimality of the equilibrium price vector for the government results in the condition (equation (12) in GH)

$$\sum_{i\in L} \nabla W_i(\boldsymbol{\tau}) + a\nabla W(\boldsymbol{\tau}) = \mathbf{0}. \tag{8}$$

Substituting the tariff revenue (5) and the expression for lobby i's welfare (6) into (8) yields

$$\begin{aligned}(a+\alpha_L)\nabla\left[\sum_{k=1}^{m}\tau_k n_k^* q_k^*(\boldsymbol{\tau}) - \sum_{k=1}^{m} p_k d_k(\mathbf{p}(\boldsymbol{\tau})) + \hat{U}(d_1(\mathbf{p}(\boldsymbol{\tau})),..,d_m(\mathbf{p}(\boldsymbol{\tau})))\right] \\ + \sum_{k=1}^{m}(I_k+a)n_k\nabla\pi_k(\boldsymbol{\tau}) \;=\; \mathbf{0},\end{aligned} \tag{9}$$

where I_k is an indicator function taking the value of 1 if industry k is organized, and 0 otherwise, and $\alpha_L = \sum_{i\in L}\alpha_i$ is the total share of population in the organized industries. Simplifying the system (9) (the detailed derivation can be found in the Appendix) entails the following form for each equation j in (9):

$$-\sum_{k=1}^{m}\tau_k\frac{\partial m_k(\boldsymbol{\tau})}{\partial\tau_j} = \sum_{k=1}^{m}\frac{(I_k+a)}{(a+\alpha_L)}n_k\frac{\partial\pi_k(\boldsymbol{\tau})}{\partial\tau_j} - \sum_{k=1}^{m}\frac{\partial p_k}{\partial\tau_j}d_k(\mathbf{p}(\boldsymbol{\tau})) + n_j^*q_j^*(\boldsymbol{\tau}), \tag{10}$$

where

$$m_k(\boldsymbol{\tau}) = n_k^* q_k^*(\boldsymbol{\tau})$$

denotes total imports in sector k. In what follows, we shall study the properties of the system of equations (10).

Before proceeding, we need to establish several intermediate results concerning the responsiveness of domestic and foreign supply to the trade tariffs.

Lemma 1 *The equilibrium quantity produced by the foreign firm in sector k decreases in the trade tariff on its own good, τ_k, and weakly increases in the trade tariff on the substitute goods, τ_{-k}. That is, $\partial q_k^*(\boldsymbol{\tau})/\partial\tau_j < 0$ if $i = k$ and $\partial q_k^*(\boldsymbol{\tau})/\partial\tau_j \geq 0$ if $i \neq k$. The equilibrium quantity produced by the domestic firm in sector k increases in the trade tariff on each good $i = 1, ..., m$. That is, $\partial q_k(\boldsymbol{\tau})/\partial\tau_j > 0$ if $i = k$ and $\partial q_k(\boldsymbol{\tau})/\partial\tau_j \geq 0$ if $i \neq k$.*

Proof. See the Appendix. ■

Lemma 1 follows from two observations. First, consider the own-tariff effect. Due to Cournot competition and linear demand functions, the home and foreign outputs in sector k are strategic substitutes. A domestic import tariff in sector k shifts the foreign firm's response curve in the south-west direction. Therefore, an increase in τ_k reduces the output of the foreign firm and increases the output of the domestic firm in sector k. Now, turn to the cross-tariff effect. The output produced by either a domestic or a foreign firm in sector k and the output produced by a foreign firm in sector j are strategic substitutes. As a result, an increase in τ_j induces higher output of both domestic and foreign firms in sector k.

As the domestic profits are given by relation (4), we immediately obtain the following corollary.

Corollary 2 *The equilibrium profit of the domestic firm in sector k increases in the trade tariff on each good $i = 1, ..., m$. That is, $\partial \pi_k(\boldsymbol{\tau})/\partial \tau_j > 0$ for all k, j.*

Let us now turn to the analysis of equation (10). First, consider the case of a separable utility function, so that $\sigma = 0$.[5] This immediately implies that $\partial m_k(\boldsymbol{\tau})/\partial \tau_j = \partial \pi_k(\boldsymbol{\tau})/\partial \tau_j = \partial p_k/\partial \tau_j = 0$ for $k \neq j$. In this case, equation (10) becomes

$$\tau_j = \frac{1}{(-\partial m_j(\boldsymbol{\tau})/\partial \tau_j)} \left[\frac{(I_j + a)}{(a + \alpha_L)} n_j \frac{\partial \pi_j(\boldsymbol{\tau})}{\partial \tau_j} - \left(\frac{\partial p_j}{\partial \tau_j} d_j(\mathbf{p}(\boldsymbol{\tau})) - n_j^* q_j^*(\boldsymbol{\tau}) \right) \right], ^{6} \qquad (11)$$

and in the absence of lobbying, the first-best trade tariffs are determined by

$$\tau_j^0 = \frac{1}{(-\partial m_j(\boldsymbol{\tau})/\partial \tau_j)} \left[n_j \frac{\partial \pi_j(\boldsymbol{\tau})}{\partial \tau_j} - \left(\frac{\partial p_j}{\partial \tau_j} d_j(\mathbf{p}(\boldsymbol{\tau})) - n_j^* q_j^*(\boldsymbol{\tau}) \right) \right]. \qquad (12)$$

Due to imperfect competition in the goods markets, free trade is no longer socially optimal. But, as in the original GH model, the organized industries experience higher protection than in the first-best equilibrium. In turn, the non-organized industries are

[5] It parallels the original GH setting in an imperfectly competitive environment.

[6] It is equivalent to

$$\frac{\tau_k}{p_k} = \frac{(2I_k + 2a)}{(a + \alpha_L)} \frac{X_k(\boldsymbol{\tau})}{m_k(\boldsymbol{\tau})} \frac{1}{\left| \frac{\partial m_k(\boldsymbol{\tau})}{\partial \tau_k} / \frac{m_k(\boldsymbol{\tau})}{p_k} \right|} + \frac{1}{p_k \frac{\partial m_k(\boldsymbol{\tau})}{\partial \tau_k}} \left(\frac{\partial p_k}{\partial \tau_k} d_k(\mathbf{p}(\boldsymbol{\tau})) - n_k^* q_k^*(\boldsymbol{\tau}) \right),$$

where $X_k(\boldsymbol{\tau})$ denotes total home production of good k, $X_k(\boldsymbol{\tau}) = n_k q_k(\boldsymbol{\tau})$. That is, we replicate the result (12) in Gawande, Krishna and Robbins (2004), up to the foreign lobbies component and the approximation they use ($\frac{\partial p_k}{\partial \tau_k} d_k(\mathbf{p}(\boldsymbol{\tau})) \approx n_k^* q_k^*(\boldsymbol{\tau})$).

underprotected. Indeed, other things equal, equations (11) and (12) differ by the term

$$\frac{1}{(-\partial m_j(\boldsymbol{\tau})/\partial \tau_j)} n_j \frac{\partial \pi_j(\boldsymbol{\tau})}{\partial \tau_j}, \tag{13}$$

having the factor $(1+a)/(a+\alpha_L)$ in the latter equation. From Lemma 1, it follows that $-\partial m_j(\tau)/\partial \tau_j > 0$, $\partial \pi_j(\tau)/\partial \tau_j > 0$ and thus, the entire term (13) is positive. As long as the organized industries do not comprise the entire population ($\alpha_L < 1$), the factor $(I_j + a) / (a + \alpha_L)$ exceeds 1 for an organized industry ($I_j = 1$) but falls below 1 for an unorganized industry ($I_j = 0$). Therefore, other things equal, for an organized industry the tariff is higher than the first-best tariff. Similarly, for a non-organized industry the tariff set in the presence of lobby groups is lower than the first-best tariff. Note that if all industries are organized and every voter belongs to some lobby ($\alpha_L = I_j = 1\ \forall j$), the protection rates are equal to the first-best ones. That is, like in GH, the lobbying efforts of different industries exactly offset each other.

Now, let us see how the relaxation of utility separability influences the equilibrium tariffs. Once more, we start by establishing an auxiliary result – we want to characterize the matrix of the derivatives of the import with respect to the trade tariffs. Denote it by M

$$M = (M_{kj}) = \left(\frac{\partial m_k}{\partial \tau_j}\right), \quad 1 \le k, j \le N.$$

Lemma 3 *The matrix M is invertible and its inverse M^{-1} is non-positive, i.e., all entries of M^{-1} are non-positive.*

Proof. See the Appendix. ■

We use Lemma 3 to obtain the expression for the trade tariffs from the system (10). System (10) in matrix form becomes

$$\begin{aligned} -M\boldsymbol{\tau} &= -\left(\frac{\partial m_k}{\partial \tau_j}\right)(\tau_j) \\ &= \left(\sum_{k=1}^{m} \frac{(I_k + a)}{(a+\alpha_L)} n_k \frac{\partial \pi_k(\boldsymbol{\tau})}{\partial \tau_j} - \sum_{k=1}^{m} \frac{\partial p_k}{\partial \tau_j} d_k(\mathbf{p}(\boldsymbol{\tau})) + n_j^* q_j^*(\boldsymbol{\tau})\right). \end{aligned} \tag{14}$$

Denote the negative of the matrix M^{-1} by

$$B = -(M)^{-1} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \dots & \dots & \dots & \dots \\ b_{m1} & b_{m2} & \dots & b_{mm} \end{pmatrix}.$$

Lemma 3 implies that each element of B is nonnegative, $b_{ij} \geq 0$. Multiplying equation (14) by B yields

$$\tau_i = \sum_{j=1}^{m} b_{ij} \left(\sum_{k=1}^{m} \frac{(I_k + a)}{(a + \alpha_L)} n_k \frac{\partial \pi_k(\boldsymbol{\tau})}{\partial \tau_j} - \sum_{k=1}^{m} \frac{\partial p_k}{\partial \tau_j} d_k(\mathbf{p}(\boldsymbol{\tau})) + n_j^* q_j^*(\boldsymbol{\tau}) \right), \qquad (15)$$

which is equivalent to the system (10).

Let us analyze the system (15). First, we note that the first-best trade tariffs are given by equation

$$\tau_i^0 = \sum_{j=1}^{m} b_{ij} \left(\sum_{k=1}^{m} n_k \frac{\partial \pi_k(\boldsymbol{\tau})}{\partial \tau_j} - \sum_{k=1}^{m} \frac{\partial p_k}{\partial \tau_j} d_k(\mathbf{p}(\boldsymbol{\tau})) + n_j^* q_j^*(\boldsymbol{\tau}) \right).$$

As above, we find that if all industries are organized and every voter belongs to some lobby ($\alpha_L = I_j = 1 \; \forall j$), the equilibrium protection rates are first-best.

However, unlike the case with a separable utility (system (11)), it does not imply that the non-organized industries are always underprotected and the organized industries are overprotected. Indeed, we see that in the presence of substitutability between products, the trade tariffs are directly affected by the sensitivity of supply and demand in the other industries and the organizational status of these industries. That is, the terms corresponding to the reaction of the other industries to the increase in τ_i, directly appear in the tariff equation for industry i. In particular, the negative effect of industry k being organized on industry i's protection is weaker if industries i and k produce substitutes. To see this, compare equations (11) and (15) for industry i. First, assume that the goods are independent ($\sigma = 0$), so that the tariff for industry i is given by equation (11) for $j = i$. Other things equal, industry k getting organized reduces the protection for industry i only through an increase in α_L (as the corresponding factor $(I_i + a)/(a + \alpha_L)$ decreases and $n_i \left(\partial \pi_i(\boldsymbol{\tau})/\partial \tau_i\right) > 0$). Now, instead, consider the case when the goods are substitutes and the tariff for industry i is given by equation (15)). As above, industry k getting organized reduces the industry i tariff through higher α_L in factors $(I_s + a)/(a + \alpha_L)$ for all $s \in \{1, ..., m\}$, as $\partial \pi_s(\tau)/\partial \tau_j \geq 0$ and $b_{ij} \geq 0$ for all i, j. However, it also has a positive effect on industry i's protection through an increase in I_k from 0 to 1 in factor $(I_k + a)/(a + \alpha_L)$. It is unclear which effect dominates, but this argument suggests that the overall negative impact on protection of good i resulting from industry k being organized is weaker in case of substitutability between the goods.

The intuition behind this effect is straightforward: in the absence of substitution, that is, if the utility function is quasi-linear and separable in goods $i = 1, ..., m$, the consumption of each non-numeraire good is fully determined by the price of this good only. As a result, each organized group has two goals. First, it tries to raise the trade tariff in its own sector, which increases its market share, the own good price and, thus, the profit of the lobby. At the same time, it attempts to reduce the tariffs on all other goods it consumes, as lower tariffs entail lower consumption prices. However, in the presence of substitution between the goods, such a lobbying strategy may cause consumers to switch consumption from the most highly protected goods to less protected ones. To limit substitution, organized industries tend to apply more "moderate" lobbying strategies. That is, they try to maintain a balance between decreasing the price of the other goods and increasing its own price.

This effect is best understood in case the ownership of industry k is very highly concentrated, that is, the share α_k of the population entitled to its profit is zero. Then, the total share of the population in the organized industries α_L does not change when industry k becomes organized. In this case, the effect of other industries' price change on lobby k's welfare is negligible, so that they have no consumer welfare gain from the trade interventions in other sectors. Therefore, in the absence of substitutability, if they get organized, it does not influence the protection of any other industry $j \neq k$ (indeed, as α_L does not change, equations (11) for tariffs in industries $j \neq k$ are unaffected). However, with substitutability between the goods, industry k still lobbies for additional protection of all substitute-producing industries because it is concerned with maintaining demand for its own good. That is, it increases industry i protection, as the term

$$\sum_{j=1}^{m} b_{ij} \frac{I_k}{(a+\alpha_L)} n_k \frac{\partial \pi_k(\boldsymbol{\tau})}{\partial \tau_j} = \sum_{j=1}^{m} b_{ij} \frac{1}{(a+\alpha_L)} n_k \frac{\partial \pi_k(\boldsymbol{\tau})}{\partial \tau_j}$$

entering the equation for the equilibrium tariff τ_i is positive.

In GH, the interests of different industry groups are opposed to each other. Here, we see that non-organized industries may benefit from the contributions of organized ones by "exploiting" the demand properties. We return to this discussion in the subsequent sections and show that this effect caused by substitutability of products can lead to free-riding in the lobbying behavior.

The above results are obtained for Cournot competition. However, it is easily seen that they also hold if we keep the assumption of linear demand and constant marginal costs, but allow firms to compete in prices in a Bertrand-fashion. Also in this case does a higher tariff in industry i convert into higher prices and demand shifting away from sector i, while higher tariffs in other sectors attract more demand into sector i.

Indeed, prices are strategic compliments and therefore, they increase in each sector's tariffs. From the profit maximization problem it follows that, in equilibrium, higher domestic prices entail higher domestic outputs

$$q_i = (p_i - c_i)\left(-\frac{\partial q_i}{\partial p_i}\right),$$

and the same is true for foreign output in sector i for a given tariff τ_i

$$q_i^* = (p_i - c_i - \tau_i)\left(-\frac{\partial q_i}{\partial p_i}\right). \tag{16}$$

Therefore, domestic output increases in each sector's tariff and foreign output increases in all other sectors' tariffs. As the sensitivity of the foreign price with respect to the own tariff is never above one (an increase in industry i's tariff is fully captured by an increase in price only in case domestic and foreign goods in industry i are perfect substitutes), formula (16) implies that foreign output is decreasing in its own tariff. It can be shown that under reasonable assumptions, the corresponding matrix M is invertible and thus the qualitative results do not change.

4 The level of protection

We are not ready to investigate whether, all else equal, more substitutability in consumer preferences causes less protection. To make the analysis tractable, we assume that there is only one domestic and one foreign firm operating in each imperfectly competitive sector, so that the respective non-numeraire good k is supplied by a duopoly, i.e.,

$$\begin{aligned} n_k &= n_k^* = 1, \\ N_k &= 2. \end{aligned}$$

Furthermore, we limit ourselves to the case of two non-numeraire goods, $m = 2$, and assume the marginal costs of production to be equal across sectors, $c_1 = c_2$. We concentrate on the interior solutions.[7]

In this case, the profit maximization problem of each domestic firm takes the form

$$\max_{q_k} \pi_k = q_k\left[1 - (q_k + q_k^*) - \sigma\left(q_{-k} + q_{-k}^*\right)\right] - cq_k, \quad k = 1, 2.$$

[7]It will be shown in the proof of Proposition 4 that a necessary condition for an interior solution is $a + 3\alpha \geq 2$.

Similarly, each foreign firm solves

$$\max_{q_k} \pi_k^* = q_k^* \left[1 - (q_k + q_k^*) - \sigma \left(q_{-k} + q_{-k}^*\right)\right] - (c + \tau_k)\, q_k, \quad k = 1, 2.$$

From the first-order conditions, it follows that for a given vector of trade tariffs $\boldsymbol{\tau} = (\tau_1, \tau_2)$, in an interior Nash equilibrium each domestic firm produces

$$q_k = \frac{(1-c)(3-2\sigma) + (3-2\sigma^2)\,\tau_k + \sigma\tau_{-k}}{9-4\sigma^2}, \tag{17}$$

while each foreign firm produces

$$q_k^* = \frac{(1-c)(3-2\sigma) - 2\,(3-\sigma^2)\,\tau_k + \sigma\tau_{-k}}{9-4\sigma^2}. \tag{18}$$

The profit in each case is determined by equality (4).

Substituting expressions (17), (18) and (4) into the system (10), we can solve for the equilibrium trade tariffs. If none of the industries is organized, so that $\alpha_L = I_1 = I_2 = 0$, the government implements the first-best policy and imposes import taxes

$$\tau_1^0(\sigma) = \tau_2^0(\sigma) = \frac{(1-c)}{\sigma+3}. \tag{19}$$

Now, consider the case when only industry 1 is organized and it represents a proportion α of the total population. The resulting trade tariff for industry 1 is given by

$$\begin{aligned}\tau_1(\sigma,\alpha,a) \;=\;& (1-c)\left[4\,(a+1)\,(a+2\alpha)\,\sigma^3 + 4\left(\alpha^2 - 3a\alpha - 2a - 3a^2 - 3\alpha\right)\sigma^2\right. \\ & - \left(9a^2 + 26a\alpha + 10a + 13\alpha^2 + 14\alpha\right)\sigma \\ & \left. + (3a+\alpha+2)\,(9a+11\alpha)\right] * \left[4\,(a+2\alpha-1)\,(a+2\alpha)\,\sigma^4\right. \\ & + \left(14a - 118a\alpha - 45a^2 - 81\alpha^2 + 22\alpha\right)\sigma^2 \\ & \left. + (9a+11\alpha-2)\,(9a+11\alpha)\right]^{-1},\end{aligned} \tag{20}$$

and the tariff for industry 2 is determined by

$$\begin{aligned}\tau_2(\sigma,\alpha,a) =\;& (1-c)\left[4a\,(a+2\alpha-1)\,\sigma^3 - 4\left(a + 2\alpha + 3a\alpha - \alpha^2 + 3a^2\right)\sigma^2\right. \\ & + (18a - 26a\alpha - 9a^2 - 13\alpha^2 + 14\alpha)\sigma \\ & \left. + (3a+\alpha)\,(9a+11\alpha-2)\right] * \left[4\,(a+2\alpha-1)\,(a+2\alpha)\,\sigma^4\right. \\ & + \left(14a - 118a\alpha - 45a^2 - 81\alpha^2 + 22\alpha\right)\sigma^2 \\ & \left. + (9a+11\alpha-2)\,(9a+11\alpha)\right]^{-1},\end{aligned} \tag{21}$$

(see the Appendix for the derivation of equations (19), (20) and (21)).

We would like to show that the increase in the degree of competition in the product market (represented by the increase in the degree of substitution between the products) reduces the incentives for the lobby to raise its own price and/or lower the prices of the other goods. Hence, it results in a more moderate equilibrium trade protection for the organized interest group.

As shown by equation (19), an increase in the degree of substitution decreases the protection level even in the absence of lobbying – the first-best trade tariffs decline in σ. The intuition is as follows: A domestic import tax in sector i is aimed at increasing the market share of the domestic firm.[8] As the degree of substitutability increases, the effect of the trade tax in sector i also extends to the firms in sector $-i$, thereby increasing their output and improving their market position. Hence, due to the competition arising from substitutability, the overall effect of a tax in sector i on the firm in sector i becomes weaker, leaving less room for strategic trade policy. When examining the influence of the degree of substitutability on lobbying, we want to abstract from this effect. We do it by analyzing how the trade tariffs in a lobbying equilibrium differ from the first-best trade policy. More precisely, we study the ratios

$$T_i(\sigma, \alpha, a) = \tau_i(\sigma, \alpha, a) / \tau_i^0(\sigma), \quad i = 1, 2,$$

which we henceforth refer to as relative protection, and their response to the change in the degree of substitution σ.

Proposition 4 *If industry* 1 *is organized, while industry* 2 *is not, the equilibrium relative protection of the organized industry decreases with the degree of substitution:*

$$\frac{dT_1(\sigma, \alpha, a)}{d\sigma} < 0.$$

For the non-organized industry, the effect is the opposite – the equilibrium relative protection increases with the degree of substitution:

$$\frac{dT_2(\sigma, \alpha, a)}{d\sigma} > 0.$$

Proof. See the Appendix. ∎

It is worth noting that the result of Proposition 4 does not necessarily imply that the influence-driven protection rates will converge to the socially-optimal level as the

[8] Subject to the respective loss in consumer welfare and the change in the tax revenues.

degree of substitutability increases. In fact, for a larger σ, the organized industry is interested in reducing the difference between the price of its own good and the substitute, not in achieving the first-best outcome. For $\sigma = 1$, the goods become indistinguishable from the consumer's point of view, i.e., industries 1 and 2 face a joint demand for these goods. Hence, the foreign firms and sector 1 and sector 2 become identical competitors from the point of view of the domestic organized industry 1, so it lobbies for the same trade tariff in both industries. However, depending on the concentration of industry 1 (that is, of the share of population $\alpha_1 = \alpha$ that has claims to its profit) this tariff can differ from the first-best level in either direction. If $\alpha = 1/2$, industry 1's members are exactly representative of the entire population – they own as much of the firm's profit claims [9] as does the average person. Therefore, the tariff for which the industry is lobbying perfectly matches the first-best tariff. If instead industry 1 is more concentrated, so that $\alpha < 1/2$, it cares about the domestic profits, as opposed to consumer welfare, more than does the average person. Therefore, in this case, the equilibrium protection rates exceed the socially optimal ones. Similarly, if $\alpha > 1/2$, the resulting equilibrium protection falls below the socially optimal level.

Still, if the degree of substitution is not too high, the non-organized industry is always underprotected as compared to the first-best trade tariff, because

$$T_2(0, \alpha, a) = 3\frac{(3a+\alpha)}{(9a+11\alpha)} = 1 - 8\frac{\alpha}{9a+11\alpha} < 1,$$

and the relative protection rate changes continuously. The organized industry achieves more than the first-best protection, as

$$T_1(0, \alpha, a) = 3\frac{(3a+\alpha+2)}{(9a+11\alpha-2)} = 1 + \frac{8(1-\alpha)}{9a+11\alpha-2} > 1.^{10}$$

Therefore, we have the following corollary:

Corollary 5 *If the degree of substitutability is not too high, an increase in σ shifts the influence-driven protection rates towards the socially optimal levels.*

We have shown above that competition in the lobbying market can reduce the deviation of the influence-based protection from the first-best level. For example, when

[9] Note that in case of perfect substitutability, domestic firms in sectors 1 and 2 are exactly identical as are their profit functions.

[10] As shown in the proof of Proposition 4, the necessary condition for the interior solution is $a + 3\alpha > 2$, which implies that $9a + 11\alpha - 2 > 0$.

every person owns shares of one or the other industry and both industries actively participate in lobbying, the economy ends up with the socially-optimal trade tariffs. Corollary 5 demonstrates that competitive pressure in the product market (resulting from the rise in the degree of substitution) can work in the same direction.

This result suggests a new explanation for a puzzle commonly observed in the empirical studies of GH theory. Most studies report extremely low estimates of the weight government puts on campaign contributions relative to social welfare.[11] That is, the government behaves almost as a social welfare maximizer and trade policy deviates surprisingly little from the first-best, which has caused a concern about the empirical significance of the GH model. For example, Gawande and Krishna (2001) write that "..it is enough to cast doubt on the value of viewing trade policy determination through this political economy lens". Recently, several papers have modified the GH model in order to resolve the puzzle. This research primarily emphasizes the idea of a lobbying competition. Gawande and Bandyopadhyay (2000) introduce political competition between the producers of intermediate and final goods, Gawande and Krishna (2005) go even further by bringing in the cross-sectoral connections through the input-output matrix, and Gawande, Krishna and Robbins (2004) introduce counter-lobbying by foreign organized groups. This paper provides an alternative explanation, suggesting that smaller deviations from the first-best trade policy may result from the weaker incentives to lobby, due to product substitutability. Indeed, all studies cited above are based on 3-4 digit SIC industry data, which suggests at least some degree of substitutability between the products of the different industries. Therefore, an organized industry's lobbying decisions may well be affected by the potential demand shift effects.

5 Substitutability and lobbying activity

So far, as in the original GH model, we assumed an exogenous lobby group structure. But why are only some industries organized? In the previous section, we demonstrated that if industry i is organized, the import tax for the substitute product $-i$ increases with the degree of substitutability. Hence, a sufficiently high degree of substitutability may entail a situation where a non-organized industry becomes protected without paying for it. That is, there may be free-riding in lobbying.

Let us illustrate this reasoning with an example. We continue to consider the 2-sector-2-firm model of the previous section but impose several additional assumptions.

[11] See e.g. Goldberg and Maggi (1999) and Gawande and Bandyopadhyay (2000).

First, assume that the non-numeraire products are perfect substitutes, $\sigma = 1$. Second, assume that the industries have the same size $\alpha_1 = \alpha_2 = \alpha$. We allow some of the voters to own no claims to any of the industries' profit but the labor only, so $\alpha \leq 1/2$. To study the free-riding behavior, we extend the existing game by allowing for an initial stage 0 of costless lobby formation. In this stage, industries simultaneously decide whether they are getting organized. The industries deciding to form a lobby participate in the standard lobbying game in stage 1.

Using backward induction, we start by discussing stage 1. Compare the equilibrium outcomes under the two regimes: (i) when only industry 1 is organized and (ii) when both industries 1 and 2 are organized. Once more, we concentrate on interior solutions.

If only industry 1 is organized, the trade tariffs in this economy are determined by equations (20) and (21) for $\sigma = 1$. This entails the equilibrium tariffs

$$\widetilde{\tau}_1 \left(1, \alpha, a\right) = \widetilde{\tau}_2 \left(1, \alpha, a\right) \equiv \widetilde{\tau} = \frac{(1-c)}{4} \frac{5a + \alpha + 2}{(5a + 7\alpha - 1)}. \tag{22}$$

As mentioned above, the import taxes on both goods are the same, even though only one of the industries is organized. Under perfect substitutability, industry 1 treats the foreign producers of both good 1 and good 2 as similar Cournot competitors. The symmetry of the setting implies that the preferred trade policy of the organized lobby does not differ with respect to the foreign firms in sectors 1 and 2. Thus, the government also sets identical trade tariffs for these two goods.

In this equilibrium, the non-organized industry 2 gains more from trade protection than the organized (and contributing) industry 1. Indeed, as the industries are identical and so are their trade tariffs, the *gross* welfare levels of industries 1 and 2, that is, the welfare before the lobbying contributions are paid, coincide. Formally,

$$W_1(\widetilde{\boldsymbol{\tau}}) = W_2(\widetilde{\boldsymbol{\tau}}),$$

where $\widetilde{\boldsymbol{\tau}}$ denotes the tariff vector in this equilibrium, $\widetilde{\boldsymbol{\tau}} = (\widetilde{\tau}_1, \widetilde{\tau}_2)$. However, the trade tariffs in this equilibrium differ from the first best level, which implies that the organized industry makes a positive contribution to the government. Therefore, the *net* welfare level of industry 1 is below the net welfare level of industry 2:

$$V_1(\widetilde{\boldsymbol{\tau}}) = W_1(\widetilde{\boldsymbol{\tau}}) - C_1(\widetilde{\boldsymbol{\tau}}) < W_2(\widetilde{\boldsymbol{\tau}}) = V_2(\widetilde{\boldsymbol{\tau}}).$$

In this regime, industry 2 benefits from the trade policy while not paying for it.

Now, consider the case when the domestic producers of both good 1 and good 2 are organized. If they both participate in lobbying, solving the equation (10) for $\sigma = 1$ and $\alpha_L = 2\alpha \leq 1$ yields the equilibrium import tariffs

$$\tilde{\tilde{\tau}}_1(1,\alpha,a) = \tilde{\tilde{\tau}}_2(1,\alpha,a) \equiv \tilde{\tilde{\tau}} = \frac{(1-c)}{4}\frac{5a+2\alpha+4}{5a+14\alpha-2}, \tag{23}$$

(see the Appendix). Note that with two industries actively lobbying, the import tax is further from the first-best level. Formally,

$$\tilde{\tilde{\tau}} \geq \tilde{\tau} \geq \tau_0(\sigma=1) = \frac{1-c}{4}. \tag{24}$$

[12]

Indeed, due to the perfect substitutability ($\sigma = 1$) and equal size ($\alpha_1 = \alpha_2 = \alpha$), industries 1 and 2 have exactly the same preferences. Hence, when they both actively participate in lobbying, the resulting policy is more biased towards the interest groups' preferred import tax.

We are interested in the trade-off between the costs and benefits of this bias for both industry 1, which is lobbying in either equilibrium, and for industry 2, which only contributes in the second equilibrium. In other words, we want to compare the lobbies' payoffs between these two equilibria. We consider the case when organized industries play globally truthful strategies, that is when the contributions of lobby i are (globally) equal to the excess of the lobby's welfare over a certain threshold B_j,

$$C_j(\tilde{\tau}, B_j) = \max\left[0, W_j(\tilde{\tau}) - B_j\right].$$

All truthful equilibria are also locally truthful, so in our argument we can rely on the results obtained in equations (22) and (23).

Note that the tariff given by equation (23) corresponds to the equilibrium when both organized industries actively participate in lobbying. In the original GH setting, the interests of different industries are opposed to each other so, if organized, each industry indeed prefers to buy the protection. However, in the presence of substitutes, this outcome is not necessarily unique. Indeed, if both industries are allowed to lobby, but each of them can get protected without paying for it, the game may admit equilibria where only one of two organized industries is active. Alternatively, there can be equilibria where both industries lobby but make different contributions. In what follows, we concentrate on symmetric truthful equilibria, which seems to be natural given the symmetry of the setting.

Denote the truthful equilibrium when only industry 1 is organized by E_1 and a symmetric equilibrium when both industries are organized by $E_{1\&2}$. We start by evaluating the amount of contributions and the government surplus in these two equilibria.

Lemma 6 *a) The government is equally well off in E_1 and $E_{1\&2}$. b) The total contributions to the government are greater in $E_{1\&2}$ than in E_1.*

Proof. To understand these results, we need to calculate truthful equilibrium contributions and net welfare levels. In a truthful equilibrium, each lobby j chooses a scalar anchor B_j so that the government would be just indifferent between choosing the equilibrium policy τ and the policy τ^{-j}, which is defined as the policy chosen by the government if the contributions of lobby j were zero:

$$\boldsymbol{\tau}^{-j} = \arg\max_{\tau \in \mathbf{T}} \sum_{i \in L, i \neq j} C_i(\boldsymbol{\tau}, B_i) + aW(\boldsymbol{\tau}). \tag{25}$$

Then, the contribution of lobby j solves

$$\sum_{i \in L, i \neq j} C_i(\boldsymbol{\tau}^{-j}, B_i) + aW(\boldsymbol{\tau}^{-j}) = \sum_{i \in L} C_i(\boldsymbol{\tau}, B_i) + aW(\boldsymbol{\tau}). \tag{26}$$

As a consistency check, we should have each lobby j making no contribution at the tariff vector τ^{-j}

$$W_j(\boldsymbol{\tau}^{-j}) \leq B_j, \tag{27}$$

as otherwise it can increase its reservation utility B_j at no cost.

We are now ready to apply this procedure. First, we establish that in E_1, the government receives exactly the payoff it would get in the first-best equilibrium. Indeed, if industry 1 were to contribute zero, the government would get no contributions at all, and would thus maximize the gross social welfare

$$\widetilde{\boldsymbol{\tau}}^{-j} = \arg\max_{\tau \in \mathbf{T}} aW(\boldsymbol{\tau}) = \boldsymbol{\tau}_0. \tag{28}$$

This observation and the government indifference condition (26) immediately imply that in such an equilibrium, the government gets exactly $G(\boldsymbol{\tau}_0)$.

Now, we show that the government receives exactly as much in $E_{1\&2}$. Indeed, by the construction of the equilibrium, condition (27) holds and industry 1 does not make any contribution at the tariff vector $\widetilde{\widetilde{\boldsymbol{\tau}}}^{-1}$:

$$C_1(\widetilde{\widetilde{\boldsymbol{\tau}}}^{-1}, B_1) = \max\left[0, W_1(\widetilde{\widetilde{\boldsymbol{\tau}}}^{-1}) - B_1\right] = 0. \tag{29}$$

As the goods are perfect substitutes and the two industries are exactly alike, their gross welfare is the same under policy $\widetilde{\widetilde{\boldsymbol{\tau}}}^{-1}$. Moreover, as the equilibrium $E_{1\&2}$ is symmetric, the anchors of two industries are the same ($B_1 = B_2$). Therefore, industry 2 does not contribute anything at the tariff vector $\widetilde{\widetilde{\boldsymbol{\tau}}}^{-1}$ either. Hence, if the contributions of lobby 1 were zero, the government would choose the free-trade policy

$$\widetilde{\widetilde{\boldsymbol{\tau}}}^{-1} = \arg\max_{\boldsymbol{\tau}\in\mathbf{T}} aW(\boldsymbol{\tau}) = \boldsymbol{\tau}^0. \tag{30}$$

The same result holds for trade policy $\widetilde{\widetilde{\boldsymbol{\tau}}}^{-2}$. Condition (26) immediately implies that in such an equilibrium, the government's payoff is the same as in the first-best equilibrium without any lobbying, which proves part a) of the Lemma.

Part b) follows from part a) and our observation (24). Indeed, when both industries are organized, the equilibrium tariffs are higher than in the case when only one industry is lobbying, and thus further from the first-best outcome. Thus, the government requires higher contributions to compensate for the larger social welfare loss. ∎

Thus, we see that the equilibrium with two organized industries provides the lobbies with higher gross welfare, but requires additional lobbying contributions. However, these contributions are now paid by two lobbies, as compared to the equilibrium with a single organized industry. So lobby 1, which was active in both equilibria, benefits more (or loses less) than lobby 2, as it can now share the costs. Still, it is not clear whether either of them actually wins or loses in the equilibrium $E_{1\&2}$. This question is answered in the following proposition.

Proposition 7 *a) Lobby* 1 *has a higher net welfare in $E_{1\&2}$ than in E_1 for any admissible parameter values. b) Lobby* 2 *has a lower net welfare in $E_{1\&2}$ than in E_1, if and only if $\alpha > 1/7$.*

Proof. We start by calculating lobbying contributions to the government in equilibrium E_1. From (26) and (28), it follows that

$$C_1\left(\widetilde{\boldsymbol{\tau}},\widetilde{B}_1\right) = aW\left(\boldsymbol{\tau}^0\right) - aW(\widetilde{\boldsymbol{\tau}}).$$

Hence, the net welfare of lobby 1 is

$$V_1(\widetilde{\boldsymbol{\tau}}) \equiv \widetilde{B_1} = W_1(\widetilde{\boldsymbol{\tau}}) - C_1(\widetilde{\boldsymbol{\tau}},\widetilde{B_1}) = W_1(\widetilde{\boldsymbol{\tau}}) - aW(\boldsymbol{\tau}^0) + aW(\widetilde{\boldsymbol{\tau}}). \tag{31}$$

As industry 2 is not organized, its net welfare is equal to its gross welfare,

$$V_2(\widetilde{\boldsymbol{\tau}}) = W_2(\widetilde{\boldsymbol{\tau}}). \tag{32}$$

In equilibrium $E_{1\&2}$, conditions (26), (29) and (30) determine the aggregate contributions

$$C_1(\tilde{\tilde{\tau}},\widetilde{\widetilde{B_1}}) + C_2(\tilde{\tilde{\tau}},\widetilde{\widetilde{B_2}}) = aW(\tau^0) - aW(\tilde{\tilde{\tau}}).$$

As this equilibrium is symmetric, the anchors of both lobbies are the same and so is their gross welfare. Hence, their contributions are also equal and comprise half of the aggregate amount

$$C_1(\tilde{\tilde{\tau}},\widetilde{\widetilde{B_1}}) = C_2(\tilde{\tilde{\tau}},\widetilde{\widetilde{B_2}}) = \frac{1}{2}\left(aW(\tau^0) - aW(\tilde{\tilde{\tau}})\right).$$

The payoff of either lobby group is thus

$$V_i(\tilde{\tilde{\tau}}) \equiv \widetilde{\widetilde{B_i}} = W_i(\tilde{\tilde{\tau}}) - C_i(\tilde{\tilde{\tau}},\widetilde{\widetilde{B_i}}) = W_i(\tilde{\tilde{\tau}}) - \frac{1}{2}\left(aW(\tau^0) - aW(\tilde{\tilde{\tau}})\right), \quad i = 1,2. \quad (33)$$

So in order to determine whether industry 1 gains from the equilibrium with two lobbies, payoffs (31) and (33) must be compared. It can be shown (see the Appendix) that

$$V_1(\tilde{\tilde{\tau}}) - V_1(\tilde{\tau}) = \frac{9a}{20}\frac{(1-2\alpha)^2(1-c)^2}{(5a+7\alpha-1)(5a+14\alpha-2)} \geq 0. \quad (34)$$

Similarly, the welfare difference of industry 2 is given by payoffs (32) and (33). Simplifying the expression (see the Appendix), we get

$$V_2(\tilde{\tilde{\tau}}) - V_2(\tilde{\tau}) = \frac{9a}{20}\frac{(1-2\alpha)^2(1-c)^2}{(5a+14\alpha-2)(5a+7\alpha-1)^2}(1-7\alpha) \begin{cases} < 0, \alpha < 1/7; \\ \geq 0, \alpha \geq 1/7. \end{cases} \quad (35)$$

■

So as long as the size of the industry is sufficiently large, it loses from participating in lobbying. The intuition behind this result is as follows: if the industries are very concentrated ($\alpha = 0$), they highly benefit from an increase in protection as the loss in their members' consumer welfare resulting from higher prices is negligible. With decreasing ownership concentration (higher α), more and more consumers in the industry lose from the price increase which results in a decrease in protection rates and industry welfare. And it may be the case that the gain of the industry from increased protection (due to this industry participation in lobbying) is not high enough to cover the necessary contribution for this increase.

We characterized the outcome of the symmetric truthful equilibrium when both industries are organized, *given* its existence. The following lemma establishes that it indeed exists.

Lemma 8 *In the lobbying game with perfect substitutability ($\sigma = 1$) and industries of identical size, there exists a symmetric truthful equilibrium.*

Proof. See the Appendix. ■

Now, we turn to the lobby formation stage. In this stage, each industry decides whether it gets organized (O) and buys influence in the next stage of the game, or stays non-organized (N) and remains passive in the lobbying game. The original GH setup entails a single equilibrium of a type (O,O): getting organized is a dominant strategy in that game as different lobbies' interests are strictly opposed to each other. The same happens in our setting if the industry's ownership structure is very concentrated (low α). Then, again, getting organized is a dominant strategy and a single (O,O) equilibrium emerges.

However, if an industry has a dispersed ownership structure (higher α), it prefers to commit not to lobby as long as the substitute industry will be lobbying. That is, the lobby formation stage is a "chicken game", where each industry prefers to be organized when the other is not and vise versa. The payoffs of the game are

Ind.1 / Ind.2	organized	non-organized
organized	$(A, A,)$	(B, C)
non-organized	(C, B)	(D, D)

where

$$\begin{aligned}
A &= V_i(\widetilde{\widetilde{\boldsymbol{\tau}}}) = W_i(\widetilde{\widetilde{\boldsymbol{\tau}}}) - \frac{1}{2}\left(aW(\boldsymbol{\tau}^0) - aW(\widetilde{\widetilde{\boldsymbol{\tau}}})\right), \\
B &= V_1(\widetilde{\boldsymbol{\tau}}) = W_1(\widetilde{\boldsymbol{\tau}}) - aW(\boldsymbol{\tau}^0) + aW(\widetilde{\boldsymbol{\tau}}), \\
C &= V_2(\widetilde{\boldsymbol{\tau}}) = W_i(\widetilde{\boldsymbol{\tau}}), \\
D &= W_i(\boldsymbol{\tau}^0).
\end{aligned}$$

As we have shown, $C > A > B > D$. This game has two pure strategy Nash equilibria, (O,N) and (N,O), and one mixed strategy equilibrium. The outcome (O,O) is no longer an equilibrium of the game. In other words, in the presence of substitutability, fewer industries get organized, and lobbying becomes less intense. This discussion is summarized in the following corollary.

Corollary 9 *In the game extended by the participation decision stage, a dispersed ownership structure weakens the incentives of the industries to get organized which, in turn, leads to a less intensive lobbying process.*

Note that aggregate net welfare of industries 1 and 2 is larger in $E_{1\&2}$ than in E_1, as

$$V_1(\widetilde{\widetilde{\boldsymbol{\tau}}}) + V_2(\widetilde{\widetilde{\boldsymbol{\tau}}}) - (V_1(\widetilde{\boldsymbol{\tau}}) + V_2(\widetilde{\boldsymbol{\tau}})) = \frac{9}{4}\frac{a^2(1-2\alpha)^2(1-c)^2}{(5a+14\alpha-2)(5a+7\alpha-1)^2} > 0.$$

The free-riding problem in lobbying can thus be summarized as follows: product substitutability produces a positive inter-industry externality from protection which, in turn, leads to dilution of the incentives to organize smaller protection than the industries desire. By continuity, this result also extends to markets with high, but not perfect substitutability.

6 Conclusion

This paper studies the impact of demand linkages on the determination of trade policy, and the intensity of inter-industry lobbying competition. We find that product substitutability both weakens industries' incentives to get organized and their lobbying incentives when they are organized. The finding may explain why empirical investigations based on the "Protection for Sale" model have suggested that the government is almost exclusively concerned with welfare rather than contributions by lobbies.

References

[1] Baldwin Robert and Frederic Robert-Nicoud, 2006,"Protection for Sale Made Easy", *CEPR discussion paper #5452,* CEPR.

[2] Bernheim, B. Douglas and Michael D. Whinston, 1986, "Menu Auctions, Resource Allocation and Economic Influence", *Quarterly Journal of Economics* 101(1), 1-31.

[3] Brander, James and Barbara Spencer, 1984, "Tariff Protection and Imperfect Competition," In *Monopolistic Competition and International Trade*, ed. H. Kierzkowski, Oxford, Oxford University Press.

[4] Brander, James and Barbara Spencer, 1985, "Export subsidies and international market share rivalry", *Journal of International Economics* 18, 83-100.

[5] Cadot, Olivier, Jaime de Melo and Marcelo Olarreaga, 2001, "Lobbying and the Structure of Tariff Protection in Poor and Rich Countries", mimeo.

[6] Chang, Pao-Li, 2005, "Protection for Sale under Monopolistic Competition", *Journal of International Economics* 66, 509-526.

[7] Dixit, Avinash K., 1984, "International trade policy for oligopolistic industries", *Economic Journal* 94 (supplement), 1–16.

[8] Dixit, Avinash K., Gene M. Grossman and Elhanan Helpman, 1997, "Common Agency and Coordination: General Theory and Applications to Government Policy Making", *Journal of Political Economy* 105(4), 752-769

[9] Eaton, Jonathan and Gene Grossman, 1986, "Optimal trade and industrial policy under oligopoly", *Quarterly Journal of Economics* 101, 383–406.

[10] Eicher Theo and Thomas Osang, 2002, "Protection for Sale: An Empirical Investigation: Comment", *American Economic Review* vol. 92, no. 5, 1702-1710.

[11] Findlay, Ronald and Stanislaw Wellisz, 1982, "Endogenous Tariffs, The Political Economy of Trade Restrictions and Welfare", in J.N. Bhagwati and T.N. Srinivasan, eds., *Import Competition and Response,* Chicago, University of Chicago Press, 223-234.

[12] Gawande, Kishore and Usree Bandyopadhyay, 2000, "Is Protection for Sale? Evidence on the Grossman-Helpman Theory of Endogenous Protection", *Review of Economics and Statistics* 82, 139-152.

[13] Gawande, Kishore and Pravin Krishna, 2002, "The Political Economy of Trade Policy: Empirical Approaches", in James Harrigan ed., *Handbook of International Trade.*

[14] Gawande, Kishore, Pravin Krishna and Michael J. Robbins, "Foreign Lobbies and US Trade Policy", 2004, *NBER working paper* #10205, NBER.

[15] Gawande, Kishore and Pravin Krishna, 2005, "Lobbying Competition over US Trade Policy", *NBER working paper* #11371, NBER.

[16] Goldberg, Penelopi and Giovanni Maggi, 1999, "Protection for Sale: An Empirical Investigation", *American Economics Review* 89, 1135-1155.

[17] Grossman, Gene M. and Elhanan Helpman, 1995, "Trade Wars and Trade Talks", *Journal of Political Economy* 103, 675-708.

[18] Grossman, Gene M. and Elhanan Helpman, 1994, "Protection for Sale", *American Economic Review* 84(4), 833-850.

[19] Magee, Christopher, 2002, "Endogenous Trade Policy and Lobby Formation: an Application to the Free-Rider Problem", *Journal of International Economics*, 57, 449-471.

[20] Magee, Stephen P., William A. Brock and Leslie Young, 1989, "Black Hole Tariffs and Endogenous Policy Theory: Political Economy in General Equilibrium", Cambridge: Cambridge University Press.

[21] Mayer, Wolfgang, 1984, "Endogenous Tariff Formation", *American Economics Review*, 74(5), December, 970-985.

[22] Mitra, Devashish, 1999, "Endogenous Lobby Formation and Endogenous Protection: A Long-Run Model of Trade Policy Determination", *American Economics Review*, 89, 1116-1134.

[23] Mitra, Devashish, Dimitrios D. Thomakos and Mehmet A. Ulubasoglu, 2002, "Protection for Sale in a Developing Country: Democracy Versus Dictatorship", *Review of Economics and Statistics*, 84(3), 497-508.

[24] Olson, Mancur, 1965, The Logic of Collective Action, Cambridge, MA: Harvard University Press.

[25] Pecorino, Paul, 1998, "Is there a Free-Rider Problem in Lobbying? Endogenous Lobbying, Trigger Strategies and the Number of Firms", *American Economics Review*, 88, 652-660.

[26] Persson, Torsten and Guido Tabellini, 2000, "Political Economics: Explaining Economic Policy", MIT Press: Cambridge.

[27] Rodrik, Dani, 1995, "Political Economy of Trade Policy, in G.M. Grossman and K. Rogoff (eds.), *Handbook of International Economics, Vol. III*, Amsterdam: North-Holland.

[28] Rodrik, Dani, 1986, "Tariffs, Subsidies and Welfare with Endogenous Policy", *Journal of International Economics* 21, 285-99.

A Appendix

A.1 Derivation of equation (10)

We begin by simplifying equation (9):

$$(a+\alpha_L)\nabla\left[\sum_{k=1}^{m}\tau_k n_k^* q_k^*(\boldsymbol{\tau})-\sum_{k=1}^{m}p_k d_k(\mathbf{p}(\boldsymbol{\tau}))+\hat{U}(d_1(\mathbf{p}(\boldsymbol{\tau})),..,d_m(\mathbf{p}(\boldsymbol{\tau})))\right]$$
$$+\sum_{k=1}^{m}(I_k+a)\nabla\pi_k(\boldsymbol{\tau})=\mathbf{0}.$$

Start by evaluating

$$\frac{\partial}{\partial\tau_j}\left[\sum_{k=1}^{m}\tau_k n_k^* q_k^*(\boldsymbol{\tau})-\sum_{k=1}^{m}p_k d_k(\mathbf{p}(\boldsymbol{\tau}))+\hat{U}(d_1(\mathbf{p}(\boldsymbol{\tau})),...,d_m(\mathbf{p}(\boldsymbol{\tau})))\right]. \tag{36}$$

As

$$\frac{\partial}{\partial\tau_j}\left(\sum_{k=1}^{m}\tau_k n_k^* q_k^*(\boldsymbol{\tau})\right)=n_j^* q_j^*(\boldsymbol{\tau})+\sum_{k=1}^{m}\tau_k n_k^*\frac{\partial q_k^*(\boldsymbol{\tau})}{\partial\tau_j},$$

$$\frac{\partial}{\partial\tau_j}\left(-\sum_{k=1}^{m}p_k d_k(\mathbf{p}(\boldsymbol{\tau}))\right)=\left(-\sum_{k=1}^{m}\frac{\partial p_k}{\partial\tau_j}d_k(\mathbf{p}(\boldsymbol{\tau}))-\sum_{k=1}^{m}p_k\frac{\partial d_k(\mathbf{p}(\boldsymbol{\tau}))}{\partial\tau_j}\right),$$

and

$$\begin{aligned}\frac{\partial}{\partial\tau_j}\left(\hat{U}(d_1(\mathbf{p}(\boldsymbol{\tau})),..,d_m(\mathbf{p}(\boldsymbol{\tau})))\right) &= \sum_{i=1}^{m}\frac{\partial\hat{U}(d_1(\mathbf{p}(\boldsymbol{\tau})),..,d_m(\mathbf{p}(\boldsymbol{\tau})))}{\partial\left(d_i(\mathbf{p}(\boldsymbol{\tau}))\right)}*\frac{\partial d_i(\mathbf{p}(\boldsymbol{\tau}))}{\partial\tau_j}\\ &= \sum_{i=1}^{m}p_i\frac{\partial d_i(\mathbf{p}(\boldsymbol{\tau}))}{\partial\tau_j},\end{aligned}$$

expression (36) is equivalent to

$$\frac{\partial}{\partial\tau_j}\left[\sum_{k=1}^{m}\tau_k n_k^* q_k^*(\boldsymbol{\tau})-\sum_{k=1}^{m}p_k d_k(\mathbf{p}(\boldsymbol{\tau}))+\hat{U}(d_1(\mathbf{p}(\boldsymbol{\tau})),..,d_m(\mathbf{p}(\boldsymbol{\tau})))\right]=$$
$$n_j^* q_j^*(\boldsymbol{\tau})+\sum_{k=1}^{m}\tau_k n_k^*\frac{\partial q_k^*(\boldsymbol{\tau})}{\partial\tau_j}-\sum_{k=1}^{m}\frac{\partial p_k}{\partial\tau_j}d_k(\mathbf{p}(\boldsymbol{\tau}))-\sum_{k=1}^{m}p_k\frac{\partial d_k(\mathbf{p}(\boldsymbol{\tau}))}{\partial\tau_j}+$$
$$\sum_{i=1}^{m}p_i\frac{\partial d_i(\mathbf{p}(\boldsymbol{\tau}))}{\partial\tau_j}=n_j^* q_j^*(\boldsymbol{\tau})+\sum_{k=1}^{m}\tau_k n_k^*\frac{\partial q_k^*(\boldsymbol{\tau})}{\partial\tau_j}-\sum_{k=1}^{m}\frac{\partial p_k}{\partial\tau_j}d_k(\mathbf{p}(\boldsymbol{\tau})).$$

Thus, each line of condition (9) takes the form

$$(a+\alpha_L)\left(n_j^* q_j^*(\boldsymbol{\tau}) + \sum_{k=1}^{m} \tau_k n_k^* \frac{\partial q_k^*(\boldsymbol{\tau})}{\partial \tau_j} - \sum_{k=1}^{m} \frac{\partial p_k}{\partial \tau_j} d_k(\mathbf{p}(\boldsymbol{\tau}))\right) + \sum_{k=1}^{m} (I_k + a) n_k \frac{\partial \pi_k(\boldsymbol{\tau})}{\partial \tau_j} = \mathbf{0},$$

which can be rewritten as

$$-\sum_{k=1}^{m} \tau_k n_k^* \frac{\partial q_k^*(\boldsymbol{\tau})}{\partial \tau_j} = \sum_{k=1}^{m} \frac{(I_k + a)}{(a+\alpha_L)} n_k \frac{\partial \pi_k(\boldsymbol{\tau})}{\partial \tau_j} - \sum_{k=1}^{m} \frac{\partial p_k}{\partial \tau_j} d_k(\mathbf{p}(\boldsymbol{\tau})) + n_j^* q_j^*(\boldsymbol{\tau}). \tag{37}$$

Substituting definition $m_k = n_k^* q_k^*$ into expression (37) yields

$$-\sum_{k=1}^{m} \tau_k \frac{\partial m_k(\boldsymbol{\tau})}{\partial \tau_j} = \sum_{k=1}^{m} \frac{(I_k + a)}{(a+\alpha_L)} n_k \frac{\partial \pi_k(\boldsymbol{\tau})}{\partial \tau_j} - \sum_{k=1}^{m} \frac{\partial p_k}{\partial \tau_j} d_k(\mathbf{p}(\boldsymbol{\tau})) + n_j^* q_j^*(\boldsymbol{\tau}).$$

A.2 Proof of Lemma 1.

The first-order condition associated with the profit maximization problem of any of the n_k symmetric domestic firms in sector k (2) is given by

$$1 - n_k q_k - n_k^* q_k^* - \sigma \sum_{s \neq k} (n_s q_s + n_s^* q_s^*) = c_k + q_k. \tag{38}$$

Similarly, the first-order condition for any of the n_k^* symmetric foreign firms is given by

$$1 - n_k q_k - n_k^* q_k^* - \sigma \sum_{s \neq k} (n_s q_s + n_s^* q_s^*) = c_k + q_k^* + \tau_k. \tag{39}$$

It follows that

$$q_k = q_k^* + \tau_k \quad \forall k. \tag{40}$$

Let aggregate output in industry k be denoted

$$Q_k = n_k q_k + n_k^* q_k^* = N_k q_k - n_k^* \tau_k. \tag{41}$$

Then (38) for each k takes the form:

$$1 - \left(1 + \frac{1}{N_k}\right) Q_k - \sigma \sum_{l \neq k} Q_L = c_k - \frac{n_k^*}{N_k} \tau_k, \tag{42}$$

and the entire system of equations for all k transforms into

$$\begin{pmatrix} 1+1/N_1 & \sigma & \dots & \sigma \\ \sigma & 1+1/N_2 & \dots & \sigma \\ \dots & \dots & \dots & \dots \\ \sigma & \sigma & \dots & 1+1/N_m \end{pmatrix} \begin{pmatrix} Q_1 \\ Q_2 \\ \dots \\ Q_m \end{pmatrix} = \begin{pmatrix} 1 - c_1 - \frac{n_1^*}{N_1}\tau_1 \\ 1 - c_2 - \frac{n_2^*}{N_2}\tau_2 \\ \dots \\ 1 - c_m - \frac{n_m^*}{N_m}\tau_m \end{pmatrix}. \tag{43}$$

Let us study the matrix

$$\boldsymbol{S}=\left(\begin{array}{cccc} 1+1/N_1 & \sigma & \dots & \sigma \\ \sigma & 1+1/N_2 & \dots & \sigma \\ \dots & \dots & \dots & \dots \\ \sigma & \sigma & \dots & 1+1/N_m \end{array}\right).$$

First, the determinant of $\boldsymbol{S}$ is positive:

$$\det \boldsymbol{S}=\frac{N_1+1}{N_1}\prod_{i=2}^{m}\left(\frac{N_i+1}{N_i}-\sigma\right)+\sigma\sum_{i=2}^{m}\left[\prod_{j=1,j\neq i}^{m}\left(\frac{N_j+1}{N_j}-\sigma\right)\right]>0.$$

Therefore, matrix $\boldsymbol{S}$ is invertible. Denote its inverse by $\boldsymbol{R}=\boldsymbol{S}^{-1}$. Then, the elements of matrix $\boldsymbol{R}$ can be written

$$R_{11}=\frac{1}{\det \boldsymbol{S}}\left(\frac{N_2+1}{N_2}\prod_{i=3}^{m}\left(\frac{N_i+1}{N_i}-\sigma\right)+\sigma\sum_{i=3}^{m}\left[\prod_{j=2,j\neq i}^{m}\left(\frac{N_j+1}{N_j}-\sigma\right)\right]\right)>0,$$

$$\begin{aligned} R_{kk} &= \frac{1}{\det \boldsymbol{S}}\left(\frac{N_1+1}{N_1}\prod_{i=2,i\neq k}^{m}\left(\frac{N_i+1}{N_i}-\sigma\right)\right. \\ &\left. +\sigma\sum_{i=2,i\neq k}^{m}\left[\prod_{j=1,j\neq i,j\neq k}^{m}\left(\frac{N_j+1}{N_j}-\sigma\right)\right]\right)>0, \end{aligned} \tag{44}$$

for $1<k\leq m$, and

$$R_{ki}=\frac{1}{\det \boldsymbol{S}}\left(-\sigma\prod_{j=1,j\neq i,j\neq k}^{m}\left(\frac{N_j+1}{N_j}-\sigma\right)\right)\leq 0.$$

That is, matrix $\boldsymbol{R}$ is symmetric, its diagonal elements are positive, and the off-diagonal elements are non-positive. Therefore, system (43) can be rewritten as

$$\left(\begin{array}{c} Q_1 \\ Q_2 \\ \dots \\ Q_m \end{array}\right)=\boldsymbol{R}\left(\begin{array}{c} 1-c_1-\frac{n_1^*}{N_1}\tau_1 \\ 1-c_2-\frac{n_1^*}{N_2}\tau_1 \\ \dots \\ 1-c_m-\frac{n_m^*}{N_m}\tau_m \end{array}\right), \tag{45}$$

and we see that

$$\frac{\partial Q_i}{\partial \tau_j}\begin{cases} <0, i=j \\ >0, i\neq j \end{cases}. \tag{46}$$

Moreover, from (41), it follows that

$$\left(\begin{array}{c} q_1 \\ q_2 \\ \dots \\ q_m \end{array}\right)=\left(\begin{array}{c} \left(Q_1+n_1^*\tau_1\right)/N_1 \\ \left(Q_2+n_2^*\tau_2\right)/N_2 \\ \dots \\ \left(Q_m+n_m^*\tau_m\right)/N_m \end{array}\right). \tag{47}$$

So

$$\frac{\partial q_k}{\partial \tau_k} = \frac{1}{N_k}\left(\frac{\partial Q_k}{\partial \tau_k} + n_k^*\right) = \frac{1}{N_k}\left(-\frac{n_k^*}{N_k}R_{kk} + n_k^*\right). \tag{48}$$

Next we show that for any k

$$R_{kk} < N_k. \tag{49}$$

From formula (44), inequality (49) is equivalent to

$$\left(\frac{N_1+1}{N_1}\prod_{i=2,i\neq k}^{m}\left(\frac{N_i+1}{N_i}-\sigma\right) + \sigma\sum_{i=2,i\neq k}^{m}\left[\prod_{j=1,j\neq i,j\neq k}^{m}\left(\frac{N_j+1}{N_j}-\sigma\right)\right]\right) - N_k \det \boldsymbol{S} < 0.$$

Denote this difference by D. It is equal to

$$\begin{aligned}
D &= \left(\frac{N_1+1}{N_1}\prod_{i=2,i\neq k}^{m}\left(\frac{N_i+1}{N_i}-\sigma\right) + \sigma\sum_{i=2,i\neq k}^{m}\left[\prod_{\substack{j=1,\\ j\neq i,j\neq k}}^{m}\left(\frac{N_j+1}{N_j}-\sigma\right)\right]\right) - N_k \det \boldsymbol{S} \\
&= \frac{N_1+1}{N_1}\prod_{i=2,i\neq k}^{m}\left(\frac{N_i+1}{N_i}-\sigma\right) + \sigma\sum_{i=2,i\neq k}^{m}\left[\prod_{\substack{j=1,\\ j\neq i,j\neq k}}^{m}\left(\frac{N_j+1}{N_j}-\sigma\right)\right] \\
&\quad -N_k\left(\frac{N_1+1}{N_1}\prod_{i=2}^{m}\left(\frac{N_i+1}{N_i}-\sigma\right) + \sigma\sum_{i=2}^{m}\left[\prod_{j=1,j\neq i}^{m}\left(\frac{N_j+1}{N_j}-\sigma\right)\right]\right).
\end{aligned}$$

By rearranging terms and simplifying, we rewrite it as

$$\begin{aligned}
D &= N_k(\sigma-1)\left(\frac{N_1+1}{N_1}\prod_{i=2,i\neq k}^{m}\left(\frac{N_i+1}{N_i}-\sigma\right) + \sigma\sum_{i=2,i\neq k}^{m}\left[\prod_{\substack{j=1,\\ j\neq i,j\neq k}}^{m}\left(\frac{N_j+1}{N_j}-\sigma\right)\right]\right) \\
&\quad -\sigma\prod_{j=1,j\neq k}^{m}\left(\frac{N_j+1}{N_j}-\sigma\right).
\end{aligned}$$

As $\sigma \in [0,1]$, $D < 0$, proving that $R_{kk} < N_k$ as desired.

Applying this result to equality (48), we conclude that

$$\frac{\partial q_k}{\partial \tau_k} = \frac{1}{N_k}\left(-\frac{n_k^*}{N_k}R_{kk} + n_k^*\right) > \frac{n_k^*}{N_k}\left(-\frac{N_k}{N_k}+1\right) = 0.$$

Equalities (46) and (47) imply that

$$\frac{\partial q_k}{\partial \tau_j} = \frac{1}{N_k}\frac{\partial Q_k}{\partial \tau_j} > 0,\ k \neq j.$$

Finally, relations (40) and (41) suggest that

$$\begin{pmatrix} q_1^* \\ q_2^* \\ \dots \\ q_m^* \end{pmatrix} = \begin{pmatrix} (Q_1 - n_1\tau_1)/N_1 \\ (Q_2 - n_2\tau_2)/N_2 \\ \dots \\ (Q_m - n_m\tau_m)/N_m \end{pmatrix}.$$

As a result, by using inequalities (46), we conclude that

$$\frac{\partial q_k^*}{\partial \tau_k} = \frac{1}{N_k}\left(\frac{\partial Q^k}{\partial \tau_k} - n_k\right) < 0$$

and

$$\frac{\partial q_k^*}{\partial \tau_j} = \frac{1}{N_k}\frac{\partial Q_k}{\partial \tau_j} > 0, \ k \neq j.$$

A.3 Proof of Lemma 3.

Equality (40) allows us to rewrite the first-order conditions(39) as

$$1 - (N_k + 1)\, q_k^* - \sigma \sum_{s \neq k} N_s q_s^* = c_k + (n_k + 1)\,\tau_k + \sigma \sum_{s \neq k} n_s \tau_s.$$

This can be rewritten in a matrix form as

$$\begin{pmatrix} 1+N_1 & \sigma N_2 & \dots & \sigma N_m \\ \sigma N_1 & 1+N_2 & \dots & \sigma N_m \\ \dots & \dots & \dots & \dots \\ \sigma N_1 & \sigma N_2 & \dots & 1+N_m \end{pmatrix} \begin{pmatrix} q_1^* \\ q_2^* \\ \dots \\ q_m^* \end{pmatrix} \tag{50}$$

$$= \begin{pmatrix} 1-c_1 \\ 1-c_2 \\ \dots \\ 1-c_m \end{pmatrix} - \begin{pmatrix} n_1+1 & n_2\sigma & \dots & n_m\sigma \\ n_1\sigma & n_2+1 & \dots & n_m\sigma \\ \dots & \dots & \dots & \dots \\ n_1\sigma & n_2\sigma & \dots & n_m+1 \end{pmatrix} \begin{pmatrix} \tau_1 \\ \tau_2 \\ \dots \\ \tau_m \end{pmatrix}.$$

Note that

$$\begin{pmatrix} 1+N_1 & \sigma N_2 & \dots & \sigma N_m \\ \sigma N_1 & 1+N_2 & \dots & \sigma N_m \\ \dots & \dots & \dots & \dots \\ \sigma N_1 & \sigma N_2 & \dots & 1+N_m \end{pmatrix} = \begin{pmatrix} \frac{1+N_1}{N_1} & \sigma & \dots & \sigma \\ \sigma & \frac{1+N_2}{N_2} & \dots & \sigma \\ \dots & \dots & \dots & \dots \\ \sigma & \sigma & \dots & \frac{1+N_m}{N_m} \end{pmatrix} \begin{pmatrix} N_1 & 0 & \dots & 0 \\ 0 & N_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & N_m \end{pmatrix}.$$

The first matrix in this product is matrix $\boldsymbol{S}$ defined above, and we have shown that it is invertible. Clearly, the diagonal matrix

$$\boldsymbol{N} = \begin{pmatrix} N_1 & 0 & \dots & 0 \\ 0 & N_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & N_m \end{pmatrix}$$

is invertible as $N_i > 0$ for all $i = 1, ...m$. Denote its inverse by $\boldsymbol{N}^{-1}$. Therefore, equation (50) can be rewritten as

$$\begin{pmatrix} q_1^* \\ q_2^* \\ ... \\ q_m^* \end{pmatrix} = \boldsymbol{N}^{-1}\boldsymbol{R}\left[\begin{pmatrix} 1-c_1 \\ 1-c_2 \\ ... \\ 1-c_m \end{pmatrix} - \begin{pmatrix} n_1+1 & n_2\sigma & ... & n_m\sigma \\ n_1\sigma & n_2+1 & ... & n_m\sigma \\ ... & ... & ... & ... \\ n_1\sigma & n_2\sigma & ... & n_m+1 \end{pmatrix}\begin{pmatrix} \tau_1 \\ \tau_2 \\ ... \\ \tau_m \end{pmatrix}\right].$$

It follows that the matrix of the derivatives of the output of a single foreign firm in sector k with respect to the trade tariffs in sector i is given by

$$\left(\frac{\partial q_k^*}{\partial \tau_j}\right) = -\boldsymbol{N}^{-1}\boldsymbol{R}\begin{pmatrix} n_1+1 & n_2\sigma & ... & n_m\sigma \\ n_1\sigma & n_2+1 & ... & n_m\sigma \\ ... & ... & ... & ... \\ n_1\sigma & n_2\sigma & ... & n_m+1 \end{pmatrix}.$$

As a result, the matrix of the derivatives of the aggregate import in sector k with respect to the trade tariffs is

$$\begin{aligned} M &= \begin{pmatrix} n_1^* & 0 & ... & 0 \\ 0 & n_2^* & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & n_m^* \end{pmatrix}\left(\frac{\partial q_k^*}{\partial \tau_j}\right) \\ &= -\begin{pmatrix} n_1^* & 0 & ... & 0 \\ 0 & n_2^* & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & n_m^* \end{pmatrix}\boldsymbol{N}^{-1}\boldsymbol{R}\begin{pmatrix} n_1+1 & n_2\sigma & ... & n_m\sigma \\ n_1\sigma & n_2+1 & ... & n_m\sigma \\ ... & ... & ... & ... \\ n_1\sigma & n_2\sigma & ... & n_m+1 \end{pmatrix}. \end{aligned}$$

Note that

$$\begin{pmatrix} n_1+1 & n_2\sigma & ... & n_m\sigma \\ n_1\sigma & n_2+1 & ... & n_m\sigma \\ ... & ... & ... & ... \\ n_1\sigma & n_2\sigma & ... & n_m+1 \end{pmatrix} = \begin{pmatrix} \frac{n_1+1}{n_1} & \sigma & ... & \sigma \\ \sigma & \frac{n_2+1}{n_2} & ... & \sigma \\ ... & ... & ... & ... \\ \sigma & \sigma & ... & \frac{n_m+1}{n_m} \end{pmatrix}\begin{pmatrix} n_1 & 0 & ... & 0 \\ 0 & n_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & n_m \end{pmatrix}.$$

Clearly, as $n_i > 0$, the inverse of

$$\begin{pmatrix} n_1 & 0 & ... & 0 \\ 0 & n_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & n_m \end{pmatrix}$$

is a diagonal matrix with positive elements. Similarly to our derivation in the proof of Lemma 1, the matrix

$$\boldsymbol{s} = \begin{pmatrix} \frac{n_1+1}{n_1} & \sigma & ... & \sigma \\ \sigma & \frac{n_2+1}{n_2} & ... & \sigma \\ ... & ... & ... & ... \\ \sigma & \sigma & ... & \frac{n_m+1}{n_m} \end{pmatrix}$$

is also invertible. Its determinant is positive and equal to

$$\det \boldsymbol{s} = \frac{n_1+1}{n_1}\prod_{i=2}^{m}\left(\frac{n_i+1}{n_i}-\sigma\right)+\sigma\sum_{i=2}^{m}\left[\prod_{j=1,j\neq i}^{m}\left(\frac{n_j+1}{n_j}-\sigma\right)\right]>0.$$

Denote its inverse by $\boldsymbol{r}=\boldsymbol{s}^{-1}$. Then, the elements of matrix $\boldsymbol{r}$ are given by

$$r_{11}=\frac{1}{\det \boldsymbol{s}}\left(\frac{n_2+1}{n_2}\prod_{i=3}^{m}\left(\frac{n_i+1}{n_i}-\sigma\right)+\sigma\sum_{i=3}^{m}\left[\prod_{j=2,j\neq i}^{m}\left(\frac{n_j+1}{n_j}-\sigma\right)\right]\right)>0,$$

$$r_{kk}=\frac{1}{\det \boldsymbol{s}}\left(\frac{n_1+1}{n_1}\prod_{i=2,i\neq k}^{m}\left(\frac{n_i+1}{n_i}-\sigma\right)+\sigma\sum_{i=2,i\neq k}^{m}\left[\prod_{j=1,j\neq i,j\neq k}^{m}\left(\frac{n_j+1}{n_j}-\sigma\right)\right]\right)>0,$$

and

$$r_{ki}=\frac{1}{\det \boldsymbol{s}}\left(-\sigma\prod_{j=1,j\neq i,j\neq k}^{m}\left(\frac{n_j+1}{n_j}-\sigma\right)\right)\leq 0.$$

Therefore, matrix M is invertible as a product of invertible matrices. It remains to show that all elements of its inverse are non-positive, i.e., that

$$M^{-1}=\begin{pmatrix} n_1 & 0 & \dots & 0\\ 0 & n_2 & \dots & 0\\ \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & n_m\end{pmatrix}^{-1}\boldsymbol{rSN}\begin{pmatrix} n_1^* & 0 & \dots & 0\\ 0 & n_2^* & \dots & 0\\ \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & n_m^*\end{pmatrix}^{-1}\geqq \mathbf{0}\,. \tag{51}$$

Lemma 10 *The matrix* $\boldsymbol{rSN}\geqq\mathbf{0}$.

Proof. As matrix $\boldsymbol{SN}$ is

$$\boldsymbol{SN}=\begin{pmatrix} 1+N_1 & \sigma N_2 & \dots & \sigma N_m\\ \sigma N_1 & 1+N_2 & \dots & \sigma N_m\\ \dots & \dots & \dots & \dots\\ \sigma N_1 & \sigma N_2 & \dots & 1+N_m\end{pmatrix},$$

the diagonal element (k,k) of the matrix $\boldsymbol{rSN}$ is given by

$$\left((1+N_k)r_{kk}+\sigma N_k\sum_{i=1,i\neq k}^{m}r_{ki}\right).$$

As the determinant of the matrix $\boldsymbol{s}=\boldsymbol{r}^{-1}$ is positive, the sign of this element does not change from multiplication by $\det \boldsymbol{s}$. Therefore, the sign of the diagonal element (k,k)

of the matrix $\boldsymbol{rSN}$ is equal to the sign of the following expression:

$$
\begin{aligned}
& \det \boldsymbol{s}* \left((1+N_k) r_{kk} + \sigma N_k \sum_{i=1, i\neq k}^{m} r_{ki} \right) \\
= \ & (1+N_k) \left(\frac{n_1+1}{n_1} \prod_{i=2, i\neq k}^{m} \left(\frac{n_i+1}{n_i} - \sigma \right) + \sigma \sum_{i=2, i\neq k}^{m} \left[\prod_{j=1, j\neq i, j\neq k}^{m} \left(\frac{n_j+1}{n_j} - \sigma \right) \right] \right) \\
& - \sigma^2 N_k \sum_{i=1, i\neq k}^{m} \prod_{j=1, j\neq i, j\neq k}^{m} \left(\frac{n_j+1}{n_j} - \sigma \right) \\
> \ & (1+N_k) \left(\frac{n_1+1}{n_1} \prod_{i=2, i\neq k}^{m} \left(\frac{n_i+1}{n_i} - \sigma \right) \right) - \sigma^2 N_k \prod_{j=2,, j\neq k}^{m} \left(\frac{n_j+1}{n_j} - \sigma \right) \qquad (52) \\
= \ & \left((1+N_k) \frac{n_1+1}{n_1} - \sigma^2 N_k \right) \left(\frac{n_1+1}{n_1} \prod_{i=2, i\neq k}^{m} \left(\frac{n_i+1}{n_i} - \sigma \right) \right) > 0,
\end{aligned}
$$

where the inequality in (52) follows from $\sigma > \sigma^2$, $1 + N_k > N_k$ and

$$
\prod_{j=1, j\neq i, j\neq k}^{m} \left(\frac{n_j+1}{n_j} - \sigma \right) > 0.
$$

Similarly, the sign of the off-diagonal element (k, j) is equal to the sign of the expression

$$
\begin{aligned}
& \det \boldsymbol{s}* \left((1+N_j) r_{kj} + \sigma N_j \sum_{i=1, i\neq j, i\neq k}^{m} r_{ki} + \sigma N_j r_{kk} \right) \\
= \ & -(1+N_j)\,\sigma \prod_{i=1, i\neq j, i\neq k}^{m} \left(\frac{n_i+1}{n_i} - \sigma \right) - \sigma^2 N_j \sum_{i=1, i\neq k, i\neq j}^{m} \prod_{l=1, l\neq i, l\neq k}^{m} \left(\frac{n_l+1}{n_l} - \sigma \right) \\
& + \sigma N_j \left(\frac{n_1+1}{n_1} \prod_{i=2, i\neq k}^{m} \left(\frac{n_i+1}{n_i} - \sigma \right) + \sigma \sum_{i=2, i\neq k}^{m} \left[\prod_{l=1, l\neq i, l\neq k}^{m} \left(\frac{n_l+1}{n_l} - \sigma \right) \right] \right) \\
= \ & \left(-\sigma (1+N_j) \left(\frac{n_1+1}{n_1} - \sigma \right) - \sigma^2 N_j \left(\frac{n_j+1}{n_j} - \sigma \right) \right. \\
& \left. + \sigma N_j \left(\frac{n_1+1}{n_1} - \sigma \right) + N_j \left(\frac{n_1+1}{n_1} \right) \left(\frac{n_j+1}{n_j} - \sigma \right) \right) \left(\prod_{i=2, i\neq k, i\neq j}^{m} \left(\frac{n_i+1}{n_i} - \sigma \right) \right) \\
= \ & \sigma \left((N_j - n_j) \frac{n_1(1-\sigma)+1}{n_1 n_j} \right) \left(\prod_{i=2, i\neq k, i\neq j}^{m} \left(\frac{n_i+1}{n_i} - \sigma \right) \right) \geq 0
\end{aligned}
$$

Therefore, all elements of matrix $\boldsymbol{rSN}$ are nonnegative. ■

Clearly, both

$$\begin{pmatrix} n_1 & 0 & \dots & 0 \\ 0 & n_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & n_m \end{pmatrix}^{-1}$$

and

$$\begin{pmatrix} n_1^* & 0 & \dots & 0 \\ 0 & n_2^* & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & n_m^* \end{pmatrix}^{-1}$$

are positive diagonal matrices, so the product of these and a nonnegative matrix is once more non-negative. From Lemma 10, it immediately follows that matrix M^{-1} determined by equality (51) is non-positive.

A.4 Derivation of equation (19).

The output levels of the foreign and domestic firms in sectors 1 and 2 are given by equations (17) and (18). Using them, we obtain the following formulas for the tariff sensitivity of the imports

$$\frac{\partial m_k(\boldsymbol{\tau})}{\partial \tau_k} = \frac{-2\left(3-\sigma^2\right)}{9-4\sigma^2},$$

$$\frac{\partial m_k(\boldsymbol{\tau})}{\partial \tau_{-k}} = \frac{\sigma}{9-4\sigma^2};$$

domestic profits

$$\frac{\partial \pi_k(\boldsymbol{\tau})}{\partial \tau_k} = 2\left(3-2\sigma^2\right)\frac{(1-c)\left(3-2\sigma\right)+\left(3-2\sigma^2\right)\tau_k+\sigma\tau_{-k}}{\left(9-4\sigma^2\right)^2},$$

$$\frac{\partial \pi_k(\boldsymbol{\tau})}{\partial \tau_{-k}} = 2\sigma\frac{(1-c)\left(3-2\sigma\right)+\left(3-2\sigma^2\right)\tau_k+\sigma\tau_{-k}}{\left(9-4\sigma^2\right)^2};$$

price

$$\frac{\partial p_k(\boldsymbol{\tau})}{\partial \tau_k} = \frac{\left(3-2\sigma^2\right)}{9-4\sigma^2},$$

$$\frac{\partial p_k(\boldsymbol{\tau})}{\partial \tau_k} = \frac{\sigma}{9-4\sigma^2};$$

and aggregate domestic consumption in sectors 1 and 2

$$d_k(\mathbf{p}(\boldsymbol{\tau})) = \frac{2(3-2\sigma)(1-c) - 3\tau_k + 2\sigma\tau_{-k}}{9-4\sigma^2}.$$

Substituting these relations into the system (10) yields a linear system

$$\begin{aligned}
& 2\left(3-\sigma^2\right)\tau_k - \sigma\tau_{-k} \\
= \; & \frac{(I_k+a)}{(a+\alpha_L)} 2\left(3-2\sigma^2\right)\frac{(1-c)(3-2\sigma)+\left(3-2\sigma^2\right)\tau_k + \sigma\tau_{-k}}{(9-4\sigma^2)} \\
& + \frac{(I_{-k}+a)}{(a+\alpha_L)} 2\sigma \frac{(1-c)(3-2\sigma)+\left(3-2\sigma^2\right)\tau_k + \sigma\tau_{-k}}{(9-4\sigma^2)} \\
& - \left(3-2\sigma^2\right)\frac{2(3-2\sigma)(1-c)-3\tau_k+2\sigma\tau_{-k}}{9-4\sigma^2} \\
& - \sigma\frac{2(3-2\sigma)(1-c)-3\tau_{-k}+2\sigma\tau_k}{9-4\sigma^2} + (1-c)(3-2\sigma) - 2\left(3-\sigma^2\right)\tau_k + \sigma\tau_{-k},
\end{aligned}$$

$k = 1, 2$. Collecting terms and multiplying by a common factor, we get an equivalent system

$$\begin{aligned}
& \left[-2\left(3-2\sigma^2\right)^2\frac{(I_k+a)}{(a+\alpha_L)} - 2\sigma\left(3-2\sigma^2\right)\frac{(I_{-k}+a)}{(a+\alpha_L)} - 76\sigma^2 + 16\sigma^4 + 99\right]\tau_k \\
& -\sigma\left[2\left(3-2\sigma^2\right)\frac{(a+I_k)}{(a+\alpha_L)} + 2\sigma\frac{(a+I_{-k})}{(a+\alpha_L)} + 15 - 4\sigma^2\right]\tau_{-k} \qquad (53) \\
= \; & (1-c)(3-2\sigma)\left[3 - 2\sigma + 2\sigma\frac{(a+I_{-k})}{(a+\alpha_L)} + 2\left(3-2\sigma^2\right)\frac{(a+I_k)}{(a+\alpha_L)}\right],
\end{aligned}$$

which determines trade tariffs τ_1 and τ_2.

In the absence of organized lobbies, $I_1 = I_2 = \alpha_L = 0$. Therefore, $(I_k+a)/(a+\alpha_L) = 1$ for $k = 1, 2$, and the system can be simplified into two symmetric equations

$$\left(-20\sigma - 4\sigma^2 + 4\sigma^3 + 27\right)\tau_k - \sigma(7-4\sigma)\tau_{-k} = (1-c)(3-2\sigma)^2$$

for $k = 1, 2$. Solving this yields

$$\tau_1 = \tau_2 = \frac{(1-c)}{3+\sigma}.$$

A.5 Derivation of equations (20) and (21).

Formulas (20) and (21) follow from solving system (53) for $I_1 = 1$, $\alpha_L = \alpha$, and $I_2 = 0$.

A.6 Proof of Proposition 4.

To show that the ratio $\frac{\tau_1(\sigma,\alpha,a)}{\tau_1^0(\sigma)}$ decreases with σ for any (a,α), we proceed in six steps.

1. Describe the necessary condition for the interior solution for the tariff.

2. Compose the ratio $T(\sigma,\alpha,a) = \frac{\tau_1(\sigma,\alpha,a)}{\tau_1^0(\sigma)}$, represent it as a ratio of two polynomials $T(\sigma,\alpha,a) = \frac{N(\sigma,\alpha,a)}{D(\sigma,\alpha,a)}$ and show that both the numerator $N(\sigma,\alpha,a)$ and the denominator $D(\sigma,\alpha,a)$ decline in σ.

3. Show that $T(\sigma,\alpha,a)$ declines in σ on $[0,1]$ iff the ratio of the derivatives of $N(\sigma,\alpha,a)$ and $D(\sigma,\alpha,a)$, $R(\sigma,\alpha,a) \equiv \frac{\partial N(\sigma,\alpha,a)}{\partial\sigma} \big/ \frac{\partial D(\sigma,\alpha,a)}{\partial\sigma}$ is higher than $T(\sigma,\alpha,a)$ for all $\sigma \in [0,1]$ and (a,α) delivering an interior solution.

4. Introduce an auxiliary linear function $A(\sigma,\alpha,a)$ and show that $R(\sigma,\alpha,a) \geq A(\sigma,\alpha,a)$ for any admissible (σ,α,a).

5. Show that $A(\sigma,\alpha,a) \geq T(\sigma,\alpha,a)$ for any admissible (σ,α,a).

6. From steps 3, 4 and 5, conclude that $R(\sigma,\alpha,a) \geq T(\sigma,\alpha,a)$.

A.6.1 Necessary condition for the interior solution for the tariff

The industry 1 tariff $\tau_1(\sigma,\alpha,a)$ is given by equation (20) and the first-best tariff (in the absence of any lobbying) $\tau_1^0(\sigma)$ is determined by equation (19). So the ratio of the lobbying tariff to the first-best is

$$
\begin{aligned}
T(\sigma,\alpha,a) &= \frac{\tau_1(\sigma,\alpha,a)}{\tau_1^0(\sigma)} = \\
&(3+\sigma)\left[4(a+1)(a+2\alpha)\sigma^3 + \left(4\alpha^2 - 12a\alpha - 8a - 12a^2 - 12\alpha\right)\sigma^2\right. \\
&\left. - \left(10a + 14\alpha + 26a\alpha + 13\alpha^2 + 9a^2\right)\sigma + (3a+\alpha+2)(9a+11\alpha)\right] / \\
&\left[4\sigma^4(a+2\alpha-1)(a+2\alpha) + \sigma^2\left(14a + 22\alpha - 118a\alpha - 81\alpha^2 - 45a^2\right)\right. \\
&\left. + (9a+11\alpha-2)(9a+11\alpha)\right].
\end{aligned}
$$

We are interested in an interior solution. A necessary condition is that the foreign firm produces a non-negative amount of the good:

$$q^*(0,\tau_1,\tau_2) = \frac{(1-c)-2\tau_1}{3} \geq 0.$$

This condition is equivalent to

$$
\begin{gathered}
\tau_1(0,\alpha,a) \leq \frac{(1-c)}{2} \Leftrightarrow \\
\frac{3a+\alpha+2}{9a+11\alpha-2} \leq \frac{1}{2} \Leftrightarrow
\end{gathered}
$$

$$a + 3\alpha \geq 2. \tag{54}$$

We assume that the necessary condition for the interior solution (54) always holds.

A.6.2 Ratio $T(\sigma, \alpha, a) = N(\sigma, \alpha, a) / D(\sigma, \alpha, a)$.

Denote the numerator of $T(\sigma, \alpha, a)$ by

$$\begin{aligned} N(\sigma, \alpha, a) &= (3+\sigma)\left[4(a+1)(a+2\alpha)\sigma^3 + \left(4\alpha^2 - 12a\alpha - 8a - 12a^2 - 12\alpha\right)\sigma^2 \right. \\ &\quad \left. - \left(10a + 14\alpha + 26a\alpha + 13\alpha^2 + 9a^2\right)\sigma + (3a+\alpha+2)(9a+11\alpha)\right], \end{aligned}$$

and the denominator by

$$\begin{aligned} D(\sigma, \alpha, a) &= \left[4\sigma^4(a+2\alpha-1)(a+2\alpha) + \sigma^2\left(14a + 22\alpha - 118a\alpha - 81\alpha^2 - 45a^2\right) \right. \\ &\quad \left. + (9a+11\alpha-2)(9a+11\alpha)\right], \end{aligned}$$

so that

$$T(\sigma, \alpha, a) = \frac{N(\sigma, \alpha, a)}{D(\sigma, \alpha, a)}.$$

Note that both $N(\sigma, \alpha, a)$ and $D(\sigma, \alpha, a)$ are decreasing in σ for any a, α. Indeed

$$\begin{aligned} N(\sigma, \alpha, a) &= 3(3a+\alpha+2)(9a+11\alpha) - 4\sigma\left(3a + 5\alpha + 9a\alpha + 7\alpha^2\right) \\ &\quad -\sigma^2\left(34a + 50\alpha + 62a\alpha + \alpha^2 + 45a^2\right) + 4\sigma^3\left(a + 3\alpha + 3a\alpha + \alpha^2\right) \\ &\quad +4\sigma^4(a+1)(a+2\alpha). \end{aligned}$$

The first derivative of $N(\sigma, \alpha, a)$ with respect to σ is

$$\begin{aligned} \frac{\partial N(\sigma, \alpha, a)}{\partial \sigma} &= -4\left(3a + 5\alpha + 9a\alpha + 7\alpha^2\right) - 2\sigma\left(34a + 50\alpha + 62a\alpha + \alpha^2 + 45a^2\right) \\ &\quad +12\sigma^2\left(a + 3\alpha + 3a\alpha + \alpha^2\right) + 16\sigma^3(a+1)(a+2\alpha). \end{aligned}$$

The second derivative of $N(\sigma, \alpha, a)$ is

$$\begin{aligned} \frac{\partial^2 N(\sigma, \alpha, a)}{\partial \sigma^2} &= -2\left(34a + 50\alpha + 62a\alpha + \alpha^2 + 45a^2\right) + 24\sigma\left(a + 3\alpha + 3a\alpha + \alpha^2\right) \\ &\quad +48\sigma^2(a+1)(a+2\alpha). \end{aligned}$$

As $(a+1)(a+2\alpha) > 0$, it is a convex quadratic parabola with the global minimum at

$$\sigma = -\frac{a + 3\alpha + 3a\alpha + \alpha^2}{(a+1)(a+2\alpha)} < 0,$$

so that the second derivative of $N(\sigma,\alpha,a)$ increases on $\sigma \in [0,1]$. This derivative is negative at $\sigma = 0$

$$\frac{\partial^2 N(0,\alpha,a)}{\partial\sigma^2} = -2\left(34a + 50\alpha + 62a\alpha + \alpha^2 + 45a^2\right) < 0,$$

and can be of either sign at $\sigma = 1$

$$\frac{\partial^2 N(1,\alpha,a)}{\partial\sigma^2} = 2\left(2a + 34\alpha + 22a\alpha + 11\alpha^2 - 21a^2\right).$$

If it is negative at $\sigma = 1$,

$$\frac{\partial^2 N(1,\alpha,a)}{\partial\sigma^2} < 0,$$

it is also negative at the entire $\sigma \in [0.1]$, meaning that the first derivative of $N(\sigma,\alpha,a)$ decreases in $\sigma \in [0,1]$. If

$$\frac{\partial^2 N(1,\alpha,a)}{\partial\sigma^2} > 0,$$

it changes sign only once, so the first derivative of $N(1,\alpha,a)$ with respect to σ, $\frac{\partial N(1,\alpha,a)}{\partial\sigma}$, first declines and then increases on $\sigma \in [0,1]$. As

$$\frac{\partial N(0,\alpha,a)}{\partial\sigma} = -4\left(3a + 5\alpha + 9a\alpha + 7\alpha^2\right) < 0$$

and

$$\frac{\partial N(1,\alpha,a)}{\partial\sigma} = -2(37a + 9\alpha + 26)(a + \alpha) < 0,$$

we can conclude that $\frac{\partial N(\sigma,\alpha,a)}{\partial\sigma} < 0$ for any $\sigma \in [0,1]$, and, thus, that $N(\sigma,\alpha,a)$ is decreasing in σ.

Similarly,

$$\begin{aligned} D(\sigma,\alpha,a) &= (9a + 11\alpha - 2)(9a + 11\alpha) + \\ &\quad \sigma^2\left(14a + 22\alpha - 118a\alpha - 81\alpha^2 - 45a^2\right) + 4\sigma^4(a + 2\alpha - 1)(a + 2\alpha). \end{aligned}$$

The first derivative of $D(\sigma,\alpha,a)$ is

$$\frac{\partial D(\sigma,\alpha,a)}{\partial\sigma} = 2\sigma\left(14a + 22\alpha - 118a\alpha - 81\alpha^2 - 45a^2\right) + 16\sigma^3(a + 2\alpha - 1)(a + 2\alpha).$$

The second derivative of $D(\sigma,\alpha,a)$ is

$$\frac{\partial^2 D(\sigma,\alpha,a)}{\partial\sigma^2} = 2\left(14a + 22\alpha - 118a\alpha - 81\alpha^2 - 45a^2\right) + 48\sigma^2(a + 2\alpha - 1)(a + 2\alpha).$$

It is once more a convex quadratic parabola with the global minimum at $\sigma = 0$, so it increases on $\sigma \in [0,1]$. As it is negative at $\sigma = 1$

$$\frac{\partial^2 D(1,\alpha,a)}{\partial\sigma^2} = -2\left(10a + 26\alpha + 22a\alpha - 15\alpha^2 + 21a^2\right) < 0,$$

it is thus negative at the entire segment $\sigma \in [0,1]$. As a result, the first derivative $\frac{\partial D(\sigma,\alpha,a)}{\partial\sigma}$ decreases over σ. As

$$\frac{\partial D(0,\alpha,a)}{\partial\sigma} = 0,$$

we conclude that for any $\sigma \in [0,1]$,

$$\frac{\partial D(\sigma,\alpha,a)}{\partial\sigma} \leq 0,$$

and thus, $D(\sigma,\alpha,a)$ is decreasing in σ.

A.6.3 Necessary and sufficient condition for $T(\sigma,\alpha,a)$ to decline in σ

The derivative of $T(\sigma,\alpha,a)$ is

$$\frac{\partial T(\sigma,\alpha,a)}{\partial\sigma} = \frac{\frac{\partial N(\sigma,\alpha,a)}{\partial\sigma} D(\sigma,\alpha,a) - \frac{\partial D(\sigma,\alpha,a)}{\partial\sigma} N(\sigma,\alpha,a)}{D^2(\sigma,\alpha,a)}.$$

It is negative if and only iff

$$\frac{\partial N(\sigma,\alpha,a)}{\partial\sigma} D(\sigma,\alpha,a) < \frac{\partial D(\sigma,\alpha,a)}{\partial\sigma} N(\sigma,\alpha,a). \tag{55}$$

As both $\frac{\partial N(\sigma,\alpha,a)}{\partial\sigma}$ and $\frac{\partial D(\sigma,\alpha,a)}{\partial\sigma}$ are negative, inequality (55) is equivalent to

$$\frac{\partial N(\sigma,\alpha,a)}{\partial\sigma} \Big/ \frac{\partial D(\sigma,\alpha,a)}{\partial\sigma} > \frac{N(\sigma,\alpha,a)}{D(\sigma,\alpha,a)} \equiv T(\sigma,\alpha,a). \tag{56}$$

Denote

$$\begin{aligned} R(\sigma,\alpha,a) \equiv{}& \frac{\partial N(\sigma,\alpha,a)}{\partial\sigma} \Big/ \frac{\partial D(\sigma,\alpha,a)}{\partial\sigma} = \\ & \left(-4\left(3a + 5\alpha + 9a\alpha + 7\alpha^2\right) - 2\sigma\left(34a + 50\alpha + 62a\alpha + \alpha^2 + 45a^2\right) + \right. \\ & \left. 12\sigma^2\left(a + 3\alpha + 3a\alpha + \alpha^2\right) + 16\sigma^3(a+1)(a+2\alpha)\right) / \\ & \left(2\sigma\left(14a + 22\alpha - 118a\alpha - 81\alpha^2 - 45a^2\right) + 16\sigma^3(a + 2\alpha - 1)(a + 2\alpha)\right). \end{aligned}$$

Substituting this notation into inequality (56), we see that $T(\sigma,\alpha,a)$ declines if and only if

$$R(\sigma,\alpha,a) > T(\sigma,\alpha,a).$$

A.6.4 Auxiliary function $A(\sigma,\alpha,a)$, such that $R(\sigma,\alpha,a) \geq A(\sigma,\alpha,a)$.

Define a linear function of σ $A(a,\alpha,\sigma)$,

$$A(a,\alpha,\sigma) = 3\frac{3a+\alpha+2}{9a+11\alpha-2} - 8\frac{a+\alpha+8a\alpha+6\alpha^2+2}{(9a+11\alpha-2)(37a+49\alpha-6)}\sigma.$$

Let us show that for any $\sigma \in [0,1]$

$$R(\sigma,\alpha,a) > A(a,\alpha,\sigma).$$

Consider

$$\begin{aligned}
\Delta_{RA} &= R(\sigma,\alpha,a) - A(a,\alpha,\sigma) \\
&= \frac{2(\sigma-1)}{\sigma(9a+11\alpha-2)(37a+49\alpha-6)} * \\
&\quad *\left[\left(14a+22\alpha-118a\alpha-81\alpha^2-45a^2\right)\right. \\
&\quad \left.+8\sigma^2(a+2\alpha-1)(a+2\alpha)\right]^{-1} \\
&\quad *\left\{\left(3a+5\alpha+9a\alpha+7\alpha^2\right)(37a+49\alpha-6)(9a+11\alpha-2)\right. \\
&\quad +\sigma(37a+49\alpha-6)(2a+6\alpha+176a\alpha+105\alpha^2 \\
&\quad -39\alpha^3-34a\alpha^2+9a^2\alpha+63a^2) \\
&\quad +8\sigma^2(a+2\alpha)\left(21a+3\alpha-17a\alpha-9\alpha^2+10\right)(9a+11\alpha-2) \\
&\quad \left.+32\sigma^3(a+2\alpha)(a+2\alpha-1)\left(a+\alpha+8a\alpha+6\alpha^2+2\right)\right\}
\end{aligned} \tag{57}$$

Let us look at the signs of components of the product in (57). Note that

$$\frac{2(\sigma-1)}{\sigma\left((14a+22\alpha-118a\alpha-81\alpha^2-45a^2)+8\sigma^2(a+2\alpha-1)(a+2\alpha)\right)} = \frac{4(\sigma-1)}{D'_\sigma(\sigma,\alpha,a)} \geq 0.$$

The product $(9a+11\alpha-2)(37a+49\alpha-6)$ is positive as each of the factors is positive (a and α are nonnegative and we employ condition (54)):

$$9a+11\alpha-2 = 8(a+\alpha)+(a+3\alpha-2) > 0,$$

$$37a+49\alpha-6 = 34a+40\alpha+3(a+3\alpha-2) > 0.$$

Similarly, the coefficients of the cubic polynomial with respect to σ in the numerator are all positive: the constant term

$$\left(3a+5\alpha+9a\alpha+7\alpha^2\right)(37a+49\alpha-6)(9a+11\alpha-2) > 0.$$

The coefficient by σ is positive,

$$(37a+49\alpha-6)\left(2a+6\alpha+176a\alpha+105\alpha^2-39\alpha^3-34a\alpha^2+9a^2\alpha+63a^2\right)>0$$

as

$$\begin{aligned}
2a+6\alpha+176a\alpha+105\alpha^2-39\alpha^3-34a\alpha^2+9a^2\alpha+63a^2 &\geq \\
2a+6\alpha+176a\alpha+105\alpha^2-39\alpha^2-34a\alpha+9a^2\alpha+63a^2 &= \\
a+6\alpha+142a\alpha+66\alpha^2+9a^2\alpha+63a^2 &> 0.
\end{aligned}$$

The coefficient by σ^2 is also positive,

$$8\,(a+2\alpha)\left(21a+3\alpha-17a\alpha-9\alpha^2+10\right)(9a+11\alpha-2)>0,$$

as

$$21a+3\alpha-17a\alpha-9\alpha^2+10>21a+3\alpha-17a-9+10=4a+3\alpha+1>0.$$

And, finally, the coefficient by σ^3 is positive,

$$32\sigma^3\,(a+2\alpha)\,(a+2\alpha-1)\left(a+\alpha+8a\alpha+6\alpha^2+2\right)>0,$$

as

$$a+2\alpha-1>a+3\alpha-2\geq 0.$$

Thus, we can conclude that for any $\sigma\in[0,1]$

$$\Delta_{RA}=R\,(\sigma,\alpha,a)-A(a,\alpha,\sigma)\geq 0. \tag{58}$$

A.6.5 Proof of $A\,(\sigma,\alpha,a)\geq T\,(\sigma,\alpha,a)$

Now we show that

$$A(a,\alpha,\sigma)\geq T\,(\sigma,\alpha,a)\,.$$

Consider the difference Δ_{TA} between $T(\sigma,\alpha,a)$ and $A(a,\alpha,\sigma)$:

$$\begin{aligned}
\Delta_{TA} &= T(\sigma,\alpha,a) - A(a,\alpha,\sigma) \\
&= \frac{4\sigma}{(9a+11\alpha-2)(37a+49\alpha-6)} / [(9a+11\alpha-2)(9a+11\alpha) \\
&\quad +\sigma^2\left(14a+22\alpha-118a\alpha-81\alpha^2-45a^2\right)+4\sigma^4(a+2\alpha-1)(a+2\alpha)] \\
&\quad *((9a+11\alpha-2)(54a+74\alpha-238a\alpha-181\alpha^2 \\
&\quad -211\alpha^3-416a\alpha^2-189a^2\alpha-93a^2)+2\sigma(37a+49\alpha-6) \qquad (59)\\
&\quad *\left(-2a-4\alpha-29a\alpha-16\alpha^2+29\alpha^3+49a\alpha^2+18a^2\alpha-9a^2\right) \\
&\quad +\sigma^2(68a+124\alpha+449a^2\alpha^2-912a\alpha-760\alpha^2+1555\alpha^3-433\alpha^4 \\
&\quad +2585a\alpha^2+1361a^2\alpha-247a\alpha^3+279a^3\alpha-280a^2+243a^3) \\
&\quad +2\sigma^3(1-\alpha)(37a+49\alpha-6)(5a+3\alpha+2)(a+2\alpha) \\
&\quad +8\sigma^4(a+2\alpha-1)\left(a+\alpha+8a\alpha+6\alpha^2+2\right)(a+2\alpha)).
\end{aligned}$$

Once more, we determine the signs of the components of the product in (59). We already know that

$$(9a+11\alpha-2)(37a+49\alpha-6)>0.$$

Moreover,

$$\begin{aligned}
&4\sigma/[(9a+11\alpha-2)(9a+11\alpha)+\sigma^2\left(14a+22\alpha-118a\alpha-81\alpha^2-45a^2\right) \\
&+4\sigma^4(a+2\alpha-1)(a+2\alpha)] = \frac{4\sigma}{D(\sigma,\alpha,a)} \geq 0.
\end{aligned}$$

Consider the remaining component of the product. Denote it by

$$\begin{aligned}
M(\sigma,\alpha,a) = {} & (9a+11\alpha-2)(54a+74\alpha-238a\alpha-181\alpha^2 \\
& -211\alpha^3-416a\alpha^2-189a^2\alpha-93a^2)+2\sigma(37a+49\alpha-6) \\
& *\left(-2a-4\alpha-29a\alpha-16\alpha^2+29\alpha^3+49a\alpha^2+18a^2\alpha-9a^2\right) \\
& +\sigma^2(68a+124\alpha+449a^2\alpha^2-912a\alpha-760\alpha^2+1555\alpha^3-433\alpha^4 \\
& +2585a\alpha^2+1361a^2\alpha-247a\alpha^3+279a^3\alpha-280a^2+243a^3) \\
& +2\sigma^3(1-\alpha)(37a+49\alpha-6)(5a+3\alpha+2)(a+2\alpha) \\
& +8\sigma^4(a+2\alpha-1)\left(a+\alpha+8a\alpha+6\alpha^2+2\right)(a+2\alpha).
\end{aligned}$$

Its second derivative with respect to σ is a quadratic parabola,

$$\begin{aligned}\frac{\partial^2 M(\sigma,\alpha,a)}{\partial\sigma^2} &= 2(68a+124\alpha+449a^2\alpha^2-912a\alpha-760\alpha^2+1555\alpha^3-433\alpha^4+\\&\quad 2585a\alpha^2+1361a^2\alpha-247a\alpha^3+279a^3\alpha-280a^2+243a^3)+\\&\quad 12\sigma(1-\alpha)(37a+49\alpha-6)(5a+3\alpha+2)(a+2\alpha)+\\&\quad 96\sigma^2(a+2\alpha-1)\left(a+\alpha+8a\alpha+6\alpha^2+2\right)(a+2\alpha).\end{aligned}$$

As $(a+2\alpha-1)(a+\alpha+8a\alpha+6\alpha^2+2)(a+2\alpha)>0$, this parabola is convex and has its minimum at $\sigma_{\min}=-\frac{(1-\alpha)(37a+49\alpha-6)(5a+3\alpha+2)}{16(a+2\alpha-1)(a+\alpha+8a\alpha+6\alpha^2+2)}<0$. Thus, it is increasing at $\sigma\in[0,1]$. It is positive at $\sigma=0$,

$$\begin{aligned}\frac{\partial^2 M(0,\alpha,a)}{\partial\sigma^2} &= 2(68a+124\alpha+449a^2\alpha^2-912a\alpha-760\alpha^2+1555\alpha^3-433\alpha^4\\&\quad +2585a\alpha^2+1361a^2\alpha-247a\alpha^3+279a^3\alpha-280a^2+243a^3)\\&= 2[(140a^2+456a\alpha+380\alpha^2)(a+3\alpha-2)+68a+124\alpha+449a^2\alpha^2\\&\quad +415\alpha^3-433\alpha^4+837a\alpha^2+485a^2\alpha-247a\alpha^3+279a^3\alpha+103a^3]\\&> 2\left(124\alpha^4+415\alpha^4-433\alpha^4+837a\alpha^2-247a\alpha^2\right)\\&= 4\alpha^2\left(295a+53\alpha^2\right)>0,\end{aligned}$$

and increasing in σ. So we conclude that for any $\sigma\in[0,1]$,

$$\frac{\partial^2 M(\sigma,\alpha,a)}{\partial\sigma^2}>0.$$

As a result, $M(\sigma,\alpha,a)$ is a convex function of σ at $[0,1]$ for any (α,a) and reaches its maximum at (either of) the corner points of the segment.

But $M(\sigma,\alpha,a)$ is negative both at $\sigma=0$ and $\sigma=1$. Indeed,

$$\begin{aligned}M(0,\alpha,a) &= (9a+11\alpha-2)[54a+74\alpha-238a\alpha\\&\quad -181\alpha^2-211\alpha^3-416a\alpha^2-189a^2\alpha-93a^2]<0,\end{aligned}$$

as

$$\begin{aligned}&54a+74\alpha-238a\alpha-181\alpha^2-211\alpha^3-416a\alpha^2-189a^2\alpha-93a^2=\\&(27a+37\alpha)(2-3\alpha-a)-\left(120a\alpha+70\alpha^2+211\alpha^3+416a\alpha^2+189a^2\alpha+66a^2\right)<0.\end{aligned}$$

Moreover,

$$M(1,\alpha,a)=-2(a+\alpha)(9a+11\alpha-2)\left(49a+81\alpha+22a\alpha+14\alpha^2-14\right)<0,$$

as

$$49a + 81\alpha + 22a\alpha + 14\alpha^2 - 14 = 7(a + 3\alpha - 2) + 42a + 60\alpha + 22a\alpha + 14\alpha^2 > 0.$$

So we conclude that

$$M(\sigma, \alpha, a) \leq 0,$$

which is equivalent to

$$\Delta_{TA} = T(\sigma, \alpha, a) - A(a, \alpha, \sigma) \leq 0. \tag{61}$$

A.6.6 Conclusion: $R(\sigma, \alpha, a) \geq A(\sigma, \alpha, a) \geq T(\sigma, \alpha, a)$

From (58) and (61), it follows that

$$R(\sigma, \alpha, a) \geq T(\sigma, \alpha, a).$$

which implies that $\frac{\tau_1(\sigma,\alpha,a)}{\tau_1^0(\sigma)}$ decreases with σ for any admissible (a, α).

The result that the ratio $\frac{\tau_2(\sigma,\alpha,a)}{\tau_2^0(\sigma)}$ increases with σ for any (a, α) is proven in a similar way.

A.7 Derivation of equation (23)

Equation (23) results from solving system (53) for $I_1 = I_2 = 1$, $\sigma = 1$and $\alpha_L = 2\alpha$. For these parameters, the system consists of two symmetric equations

$$\left[-4\frac{(1+a)}{(a+2\alpha)} + 39\right]\tau_k - \left[4\frac{(1+a)}{(a+2\alpha)} + 11\right]\tau_{-k} = (1-c)\left[1 + 4\frac{(1+a)}{(a+2\alpha)}\right],$$

which yields

$$\tilde{\tilde{\tau}}_1(1, \alpha, a) = \tilde{\tilde{\tau}}_2(1, \alpha, a) = \frac{(1-c)}{4}\frac{5a + 2\alpha + 4}{5a + 14\alpha - 2}.$$

A.8 Derivation of equations (34) and (35)

From equations (31) and (33), it follows that

$$\begin{aligned} V_1(\tilde{\tilde{\tau}}) - V_1(\tilde{\tau}) &= W_1(\tilde{\tilde{\tau}}) - \frac{1}{2}\left(aW(\tau^0) - aW(\tilde{\tilde{\tau}})\right) - \left(W_1(\tilde{\tau}) - aW(\tau^0) + aW(\tilde{\tau})\right) \\ &= W_1(\tilde{\tilde{\tau}}) - W_1(\tilde{\tau}) + \frac{1}{2}a\left(W(\tau^0) + W(\tilde{\tilde{\tau}}) - 2W(\tilde{\tau})\right). \end{aligned} \tag{62}$$

Substituting the expressions for the outputs, profits, volume of imports and domestic consumption into the formulas for the lobby $i = 1, 2$, welfare (6) and aggregate social welfare (7), and simplifying the resulting expressions yields

$$W_i(\widetilde{\widetilde{\boldsymbol{\tau}}}) = \frac{9}{4}(1-c)^2(a+2\alpha)\frac{a-2\alpha+14\alpha^2+3a\alpha}{(5a+14\alpha-2)^2},$$

$$W_i(\widetilde{\boldsymbol{\tau}}) = \frac{9}{4}(1-c)^2(a+\alpha)\frac{a-\alpha+7\alpha^2+3a\alpha}{(5a+7\alpha-1)^2},$$

$$W(\boldsymbol{\tau}^0) = \frac{9}{20}(1-c)^2,$$

$$W(\widetilde{\boldsymbol{\tau}}) = \frac{9}{4}(1-c)^2(a+\alpha)\frac{5a+9\alpha-2}{(5a+7\alpha-1)^2},$$

$$W(\widetilde{\widetilde{\boldsymbol{\tau}}}) = \frac{9}{4}(1-c)^2(a+2\alpha)\frac{5a+18\alpha-4}{(5a+14\alpha-2)^2}.$$

Inserting into (62), we obtain

$$V_1(\widetilde{\widetilde{\boldsymbol{\tau}}}) - V_1(\widetilde{\boldsymbol{\tau}}) = \frac{9}{20}\frac{(1-c)^2 a(1-2\alpha)^2}{(5a+7\alpha-1)(5a+14\alpha-2)}.$$

Similarly, equations (32), (33) and the formulas above imply that

$$\begin{aligned} V_2(\widetilde{\widetilde{\boldsymbol{\tau}}}) - V_2(\widetilde{\boldsymbol{\tau}}) &= W_1(\widetilde{\widetilde{\boldsymbol{\tau}}}) - \frac{1}{2}\left(aW(\boldsymbol{\tau}^0) - aW(\widetilde{\widetilde{\boldsymbol{\tau}}})\right) - W_2(\widetilde{\boldsymbol{\tau}}) \\ &= \frac{9}{20}\frac{(1-c)^2 a(1-7\alpha)(1-2\alpha)^2}{(5a+14\alpha-2)(5a+7\alpha-1)^2}. \end{aligned}$$

A.9 Proof of Lemma (8)

If industry 2 chooses $\widetilde{\widetilde{B_2}}$ defined by expression (33), industry 1's best response is precisely setting $\widetilde{\widetilde{B_1}} = \widetilde{\widetilde{B_2}}$. Clearly, it is not in industry 1's interest to decrease $\widetilde{\widetilde{B_1}}$ (and hence its payoff, in case $\widetilde{\widetilde{B_1}}$ does not bind). We need to consider two cases. Assume first that $\alpha > 1/7$ so that

$$\widetilde{\widetilde{B_i}} = W_i(\widetilde{\widetilde{\boldsymbol{\tau}}}) - \frac{1}{2}\left(aW(\boldsymbol{\tau}^0) - aW(\widetilde{\widetilde{\boldsymbol{\tau}}})\right) < W_1(\widetilde{\boldsymbol{\tau}}),$$

which implies that the government gets positive contributions both at policies $\widetilde{\widetilde{\boldsymbol{\tau}}}$ and $\widetilde{\boldsymbol{\tau}}$. If industry 1 chooses $\widetilde{\widetilde{B_1}} > \widetilde{\widetilde{B_2}}$, the government prefers tariff $\boldsymbol{\tau}^0$ to $\widetilde{\widetilde{\boldsymbol{\tau}}}$ and $\widetilde{\boldsymbol{\tau}}$. Indeed, by equation (8) $\tilde{\tilde{\boldsymbol{\tau}}}$ maximizes the sum of social welfare and the lobbies' welfare, so that the government prefers $\tilde{\tilde{\boldsymbol{\tau}}}$ to $\tilde{\boldsymbol{\tau}}$:

$$
\begin{aligned}
G(\tilde{\boldsymbol{\tau}}) &= aW(\tilde{\boldsymbol{\tau}}) + \left(W_1(\tilde{\boldsymbol{\tau}}) - \tilde{\tilde{\tilde{B}}}_1\right) + \left(W_2(\tilde{\boldsymbol{\tau}}) - \tilde{\tilde{\tilde{B}}}_2\right) \\
&< aW(\tilde{\tilde{\boldsymbol{\tau}}}) + \left(W_1(\tilde{\tilde{\boldsymbol{\tau}}}) - \tilde{\tilde{\tilde{B}}}_1\right) + \left(W_2(\tilde{\tilde{\boldsymbol{\tau}}}) - \tilde{\tilde{\tilde{B}}}_2\right) = G(\tilde{\tilde{\boldsymbol{\tau}}}).
\end{aligned}
$$

In turn, due to our assumption $\tilde{\tilde{\tilde{B}}}_1 > \tilde{\tilde{\tilde{B}}}_2$, policy $\tilde{\tilde{\boldsymbol{\tau}}}$ does not pay the government enough to deviate from the first-best policy:

$$
\begin{aligned}
G(\tilde{\tilde{\boldsymbol{\tau}}}) &= aW(\tilde{\tilde{\boldsymbol{\tau}}}) + \left(W_1(\tilde{\tilde{\boldsymbol{\tau}}}) - \tilde{\tilde{\tilde{B}}}_1\right) + \frac{1}{2}\left(aW(\boldsymbol{\tau}^0) - aW(\tilde{\tilde{\boldsymbol{\tau}}})\right) \\
&= \left(W_1(\tilde{\tilde{\boldsymbol{\tau}}}) - \tilde{\tilde{\tilde{B}}}_1\right) + \frac{1}{2}\left(aW(\boldsymbol{\tau}^0) + aW(\tilde{\tilde{\boldsymbol{\tau}}})\right) < aW(\boldsymbol{\tau}^0).
\end{aligned}
$$

If instead $\alpha \leq 1/7$, implying that

$$
\tilde{\tilde{\tilde{B}}}_i = W_i(\tilde{\tilde{\boldsymbol{\tau}}}) - \frac{1}{2}\left(aW(\boldsymbol{\tau}^0) - aW(\tilde{\tilde{\boldsymbol{\tau}}})\right) > W_1(\tilde{\boldsymbol{\tau}}),
$$

the government never sets $\tilde{\boldsymbol{\tau}}$ as it does not get any contributions for this policy. Similarly to above, the government prefers tariff $\boldsymbol{\tau}^0$ to $\tilde{\tilde{\boldsymbol{\tau}}}$. Therefore, by increasing $\tilde{\tilde{\tilde{B}}}_1$ industry 1 receives a payoff $W_1(\boldsymbol{\tau}_0) < \tilde{\tilde{\tilde{B}}}_1$, so this cannot be a best response.

EFI, The Economic Research Institute

Published in the language indicated by the title. A complete publication list can be found at *www.hhs.se/efi* Books and dissertations can be ordered from EFI via e-mail: EFI.Publications@hhs.se

Reports since 2002

2006

Books

Thodenius, Björn. *Organisering av kunskap. En studie av Wallenberg Consortium North.*

Wijkström, Filip och Torbjörn Einarsson. *Från nationalstat till näringsliv? Det civila samhällets organisationsliv i förändring.*

Östman, Lars. *Lysande ögonblick och finansiella kriser. Dramaten under ett sekel.*

Dissertations

Argenton, Cedric. *Quality provision in duopoly.*

Beckerman, Carina. *The Clinical Eye. Constructiong and Computerizing an Anesthesia Patient Record.*

Borglund, Tommy. *Aktievärden i fokus – internationell påverkan på intressentrelationer vid förvärv och fusion.*

Breman, Anna. *The Economics of altruism, paternalism and self-control.*

Edquist, Harald. *Technological breakthroughs and productivity growth.*

Frostenson, Magnus. *Legitimitetskontrollen – en studie av etiska värderingars roll i gränsöverskridande förvärv och fusioner.*

Gaspar, Raquel M. *Credit Risk and Forward Price Models.*

Hopkins, Elisabeth. *Is a higher degree of local currency pricing associated with lower exchange rate pass-through? A study of import pricing in 51 Swedish industries.*

Kling, Ragnar. *Developing Product Development in Times of Brutal Change.*

Langenskiöld, Sophie. *Peer Influence on Smoking: Causation or Correlation?*

Lychnell, Lars-Olof. *IT-relaterat förändringsarbete I informationsintensiva småföretag – en klinisk fallstudie från resebranschen.*

Meitz, Mika. *Five Contributions to Econometric Theory and the Econometrics of Ultra-High-Frequency Data.*

Mendicino, Caterina. *Financial Market Imperfections, Business Cycle Fluctuations and Economic Growth.*

Ovanfors, Anna. *Essays on Nonlinear Time Series Analysis and Health Economics.*

Rogberg, Martin. *Den modeföljande organisationen. Om acceptansen av TQM och andra populära managementmodeller.*

Slinko, Irina. *Essays in Option Pricing and Interest Rate Models.*

Wilander, Fredrik. *Essays on Exchange Rates and Prices.*

2005

Books

Andersson, Per, Susanne Hertz and Susanne Sweet (eds). *Perspectives on market networks – boundaries and new connections.*

Charpentier, Claes. *IT inom omsorgen. Förväntade effekter av införande av IT-system för utförarna inom äldre- och handikappomsorgen.*

Dembrower, Maria. *Entreprenörskap i industriella nätverk.*
Lind, Johnny och Göran Nilsson (redaktörer). *Ekonomistyrningens metoder, sammanhang och utveckling – En vänbok till Lars A Samuelson.*
Samuelson, Lars A. *Organizational governance and control – a summary of research in the Swedish society.*
Svedberg Nilsson, Karin, Roger Henning och Karin Fernler (eds). *En illusion av frihet? Företag och organisationer i regelsamhället.* EFIs Årsbok 2005. EFI/Studentlitteratur.

Dissertations

Andersson, Martin. *Making a Difference – Project Result Improvement in Organizations.*
Arvidsson, Per. *Styrning med belöningssystem – Två fallstudier om effekter av belöningssystem som styrmedel.*
Bems, Rudolfs. *Essays in International Macroeconomics.*
Berg-Suurwee, Ulrika. *Nya trender, nya nämnder – effekter av en stadsdelsnämndsreform inom kultur och fritid.*
Björkman, Hans. *Learning from members – Tools for strategic positioning and service innovation in trade unions.*
Bodnaruk, Andriy. *Essays in Empirical Corporate Finance and Portfolio Choice.*
Clapham, Eric. *Essays in Real Estate Finance.*
Dareblom, Jeanette. *Prat, politik och praktik – Om individers möten med strukturer i en kommunal satsning på kvinnors företagande.*
Fromm, Jana. *Risk Denial and Neglect: Studies in Risk Perception.*
Hjelström, Anja. *Understanding International Accounting Standard Setting – A Case Study of IAS 12, Accounting for Deferred Tax.*
Hortlund, Per. *Studies on Swedish Banking 1870-2001.*
Lindahl, Therese. *Strategic and Environmental Uncertainty in Social Dilemmas.*
Linnarsson, Håkan. *Alliance for Innovation. A structural perspective on new business development in cooperative ventures.*
Madestam, Andreas. *Developing Credit Markets.*
Nilsson, Roland. *The Market Impact of Short-Sale Constraints.*
Nordfält, Jens. *Is consumer decision-making out of control? Non-conscious influences on the consumer decision-making process for fast moving consumer goods.*
Nordin, Fredrik. *Externalising Services – Walking a Tightrope between Industrial and Service Logics.*
Parmler, Johan. *Essays in Empirical Asset Pricing.*
Simbanegavi, Witness. *Price Discrimination, Advertising and Competition.*
Thodenius, Björn. *Användning av ledningsinformationssystem: en longitudinell studie av svenska storföretag.*
Tolis, Christofer. *Framing the Business – Business Modelling for Business Development.*
Östberg, Per. *Corporate Disclosure and Investor Recognition.*

2004

Books

Ahrne, Göran och Nils Brunsson (red). *Regelexplosionen.*
Lind, Johnny. *Strategi och ekonomistyrning. En studie av sambanden mellan koncernstrategi, affärsstrategi och ekonomistyrning.*
Lind, Johnny och Walter Schuster (red). *Redovisningens teori, praktik och pedagogik. En vänbok till Lars Östman.*
Sevón, Guje och Lennart Sjöberg (red). *Emotioner och värderingar i näringslivet.* EFIs Årsbok 2004.
Wijkström, Filip and Stefan Einarsson. *Foundations in Sweden – Their scope, roles and visions.*

Dissertations

Anderson, Anders. *Essays in Behavioral Finance.*

Balsvik, Gudrun. *Information Technology Users: Studies of Self-Efficacy and Creativity among Swedish Newspaper Journalists.*

Blix, Magnus. *Essays in Mathematical Finance – Modelling the Futures Price.*

González Gómez, Andrés. *Nonlinear dynamics and smooth transition models.*

Grönqvist, Erik. *Selection and Moral Hazard in Health Insurance: Taking Contract Theory to the Data.*

Ivarsson Westerberg, Anders. *Papperspolisen – varför ökar administrationen i moderna organisationer.*

Jutterström, Mats. *Att påverka beslut – företag i EUs regelsättande.*

Jönsson, Kristian. *Macroeconomic Aspects of Capital Flows to Small Open Economies in Transition.*

Larsson, Pär. *Förändringens villkor. En studie av organisatoriskt lärande och förändring inom skolan.*

Lagerwall, Björn. *Empirical Studies of Portfolio Choice and Asset Prices.*

Malmsten, Hans. *Properties and Evaluation of Volatility Models.*

Marshall, Cassandra. *Dating for Innovation. Creating and Recognizing Opportunities through Collaborative Interorganizational Relationships in Fluid Environments.*

Mattsson, Susanna. *På gränsen mellan ordning och oordning – tingens betydelse vid marknadsombildningar. En studie av svenska postväsendets ombildning under 1990-talet.*

Nilsson, Charlotte. *Studies in Environmental Economics: Numerical Analysis of Greenhouse Gas Policies.*

Nilsson, Hans. *Medborgaren i styrsystemet – beskrivning av VAD och HUR i styrning av kommunal verksamhet.*

Nystedt, Jens. *Competition, Regulation and Integration in International Financial Markets.*

Pajuste, Anete. *Corporate Governance and Controlling Shareholders.*

Richtnér, Anders. *Balancing Knowledge Creation. Organizational Slack and Knowledge Creation in Product Development.*

Salabasis, Mickael. *Bayesian Time Series and Panel Models – Unit Roots, Dynamics and Random Effects.*

Sandberg, Rickard. *Testing the Unit Root Hypothesis in Nonlinear Time Series and Panel Models.*

Skallsjö, Sven. *Essays on Term Structure and Monetary Policy.*

Strikholm, Birgit. *Essays on Nonlinear Time Series Modelling and Hypothesis Testing.*

Söderström, John. *Från Produkt till Tjänst. Utveckling av affärs- och miljöstrategier i produktorienterade företag.*

Talia, Krim. *The Scandinavian Currency Union, 1873–1924 – Studies in Monetary Integration and Disintegration.*

2003

Books

Lundahl, Mats (editor). *Globalization and Its Enemies.* EFIs Årsbok 2003.

Sundgren, Bo, Pär Mårtensson, Magnus Mähring and Kristina Nilsson (editors). *Exploring Patterns in Information Management. Concepts and Perspectives for Understanding IT-Related Change.*

Dissertations

Andersson, Henrik. *Valuation and Hedging of Long-Term Asset-Linked Contracts.*

Bergman, Malin. *Essays on Human Capital and Wage Formation.*

Damsgaard, Niclas. *Deregulation and Regulation of Electricity Markets.*

Eklund, Bruno. *Four Contributions to Statistical Inference in Econometrics.*

Hakkala, Katariina. *Essays on Restructuring and Production Decisions in Multi-Plant Firms.*

Holgersson, Charlotte. *Rekrytering av företagsledare. En studie i homosocialitet.*

Ivaschenko, Iryna. *Essays on Corporate Risk, U.S. Business Cycles, International Spillovers of Stock Returns, and Dual Listing.*

Lange, Fredrik. *Brand Choice in Goal-derived Categories – What are the Determinants?*

Le Coq, Chloé. *Quantity Choices and Market Power in Electricity Market.*

Magnusson, Peter R. *Customer-Oriented Product Development – Experiments Involving Users in Service Innovation.*

Meisiek, Stefan. *Beyond the Emotional Work Event Social Sharing of Emotion in Organizations.*

Mårtensson, Anders. *Managing Mission-Critical IT in the Financial Industry.*

Nilsson, Göran. *Processorientering och styrning – Regler, mål eller värderingar?*

Sandberg, Robert. *Corporate Consulting for Customer Solutions Bridging Diverging Business Logics.*

Sturluson, Jon Thor. *Topics in the Industrial Organization of Electricity Markets.*

Tillberg, Ulrika. *Ledarskap och samarbete – En jämförande fallstudie i tre skolor.*

Waldenström, Daniel. *Essays in Historical Finance.*

Wallén, Ulrika. *Effektivitet i grundskolan i anslutning till en stadsdelsnämndsreform.*

Ögren, Anders. *Empirical Studies in Money, Credit and Banking – The Swedish Credit Market in Transition under the Silver and the Gold Standards, 1834–1913.*

2002

Books

Schuster, Walter. *Företagets Valutarisk – En studie av horisontella och vertikala styrprocesser.*

Sjöstrand, Sven-Erik och Pernilla Petrelius. *Rekrytering av koncernstyrelsen – Nomineringsförfaranden och styrelsesammansättning med fokus på kvinnors ställning och möjligheter.* EFI/SNS Förlag.

Löwstedt, Jan och Bengt Stymne (red). *Scener ur ett företag – Organiseringsteori för kunskapssamhället.* EFIs Årsbok 2002. EFI/Studentlitteratur.

Dissertations

Barinaga, Ester. *Levelling Vagueness – A Study of Cultural Diversity in an International Project Group.*

Berglund, Johan. *De otillräckliga – En studie av personalspecialisternas kamp för erkännande och status.*

Bolander, Pernilla. *Anställningsbilder och rekryteringsbeslut.*

Damjanovic, Tatiana. *Essays in Public Finance.*

Ekman, Mattias. *Studies in Health Economics – Modelling and Data Analysis of Costs and Survival.*

Heyman, Fredrik. *Empirical Studies on Wages, Firm Performance and Job Turnover.*

Kallifatides, Markus. *Modern företagsledning och omoderna företagsledare.*

Kaplan, Michael. *Acquisition of Electronic Commerce Capability – The Cases of Compaq and Dell in Sweden.*

Mähring, Magnus. *IT Project Governance.*

Nilsson, Mattias. *Essays in Empirical Corporate Finance and Governance.*

Schenkel, Andrew. *Communities of Practice or Communities of Discipline – Managing Deviations at the Øresund Bridge.*

Skogsvik, Stina. *Redovisningsmått, värderelevans och informationseffektivitet.*

Sundén, David. *The Dynamics of Pension Reform.*

Ternström, Ingela. *The Management of Common-Pool Resources – Theoretical Essays and Empirical Evidence.*

Tullberg, Jan. *Reciprocitet – Etiska normer och praktiskt samarbete.*

Westling, Gunnar. *Balancing Innovation and Control – The Role of Face-to-face Meetings in Complex Product Development Projects.*

Viklund, Mattias. *Risk Policy – Trust, Risk Perception, and Attitudes.*

Vlachos, Jonas. *Risk Matters – Studies in Finance, Trade and Politics.*