

# Essays on Social Conflict and Reform

Anders Bornefalk

## AKADEMISK AVHANDLING

Som för avläggande av  
ekonomie doktorsexamen  
vid Handelshögskolan i Stockholm  
framlägges för offentlig granskning  
fredagen den 31 mars 2000, kl 15.15  
i sal Ragnar, Handelshögskolan,  
Sveavägen 65





## **Essays on Social Conflict and Reform**



# STOCKHOLM SCHOOL OF ECONOMICS

## EFI, THE ECONOMIC RESEARCH INSTITUTE

### ***EFI Mission***

EFI, the Economic Research Institute at the Stockholm School of Economics, is a scientific institution which works independently of economic, political and sectional interests. It conducts theoretical and empirical research in 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.

EFI selects its projects based on the need for theoretical or practical development of a research domain, on methodological interests, and on the generality of a problem.

### ***Research Organization***

The research activities are organized in nineteen Research Centers within eight Research Areas. Center Directors are professors at the Stockholm School of Economics.

#### ***ORGANIZATION AND MANAGEMENT***

Management and Organisation; (A)  
Center for Ethics and Economics; (CEE)  
Public Management; (F)  
Information Management; (I)  
Man and Organisation; (PMO)  
Industrial Production; (T)

Prof Sven-Erik Sjöstrand  
Adj Prof Hans de Geer  
Prof Nils Brunsson  
Prof Mats Lundeborg  
Acting Prof Jan Löwstedt  
Prof Christer Karlsson

#### ***ECONOMIC PSYCHOLOGY***

Center for Risk Research; (CFR)  
Economic Psychology; (P)

Prof Lennart Sjöberg  
Prof Lennart Sjöberg

#### ***MARKETING***

Center for Information and Communication  
Research; (CIC)  
Center for Consumer Marketing; (CCM)  
Marketing, Distribution and Industrial  
Dynamics; (D)

Adj Prof Bertil Thorngren  
Associate Prof Magnus Söderlund  
  
Prof Lars-Gunnar Mattsson

#### ***ACCOUNTING, CONTROL AND CORPORATE FINANCE***

Accounting and Managerial Finance; (B)  
Managerial Economics; (C)

Prof Lars Östman  
Prof Peter Jennergren

#### ***FINANCE***

Finance; (FI)

Prof Clas Bergström

#### ***ECONOMICS***

Center for Health Economics; (CHE)  
International Economics and Geography; (IEG)  
Economics; (S)

Prof Bengt Jönsson  
Prof Mats Lundahl  
Prof Lars Bergman

#### ***ECONOMICS STATISTICS***

Economic Statistics; (ES)

Prof Anders Westlund

#### ***LAW***

Law; (RV)

Prof Erik Nerep

*Chairman of the Board:* Prof Sven-Erik Sjöstrand. *Director:* Associate Prof Bo Sellstedt.

### ***Address***

EFI, Box 6501, S-113 83 Stockholm, Sweden • Internet: [www.hhs.se/efi/](http://www.hhs.se/efi/)  
Telephone: +46(0)8-736 90 00 • Fax: +46(0)8-31 62 70 • E-mail [efi@hhs.se](mailto:efi@hhs.se)

# Essays on Social Conflict and Reform

Anders Bornefalk



**SITE**

*Stockholm Institute of Transition Economics and East European Economies*  
STOCKHOLM SCHOOL OF ECONOMICS



**STOCKHOLM SCHOOL OF ECONOMICS**  
EFI, THE ECONOMIC RESEARCH INSTITUTE



Dissertation for the degree of Doctor of Philosophy, Ph.D.  
Stockholm School of Economics, 2000

© EFI and the author  
ISBN NR 91-7258-528-5

**Keywords:** Social conflict  
Redistribution  
Political economy  
Policy reform  
Democratic institutions  
Constitutions  
Economic transition  
Rent seeking  
Monetary unions  
Ruble Zone  
Contest success functions

**Cover:** Poland 1991: A potential social conflict.  
Photograph by Anders Bornefalk

**Distributed by:**  
EFI, Stockholm School of Economics  
Box 6501, SE-113 83 Stockholm, Sweden

Printed by Elanders Gotab, Stockholm 2000

*To Ankita and Alfred*





# Contents

<b>Acknowledgements</b>	xi
<b>Introduction and Summary</b>	1
<b>Essay 1:</b>	
<b>Social Conflict with Passive Groups</b>	
1. Introduction	1
2. The Model	4
2.1 Production and Distribution	4
2.2 Cooperation	6
3. The Cooperative Case	7
4. The Non-cooperative Case	10
5. Exit, Voice, or Loyalty?	16
6. Concluding Discussion	18
Appendix	20
References	22
<b>Essay 2:</b>	
<b>Democratization, Rent Seeking, and Economic Transition</b>	
1. Introduction	1
2. The Model	5
3. A Transition Interpretation	9
3.1 The Vulnerability of Unorganized Groups	9
3.2 The Possibility of Cooperation	13
4. Economic Performance and Reform	14
4.1 The Cooperative Case	15
4.2 The Non-cooperative Case	18
5. Political Reform	23
5.1 The Democratic Culture	23
5.2 The Cooperative Case	24
5.3 The Non-cooperative Case	25
6. Applications	28
6.1 Poland and Russia Compared	28
6.2 Rapid or Gradual Democratization?	30
7. Concluding Remarks	32
Appendix	34
References	37

### **Essay 3:**

#### **Constitutional Constraints and Redistributive Activities**

1. Introduction	2
2. The Model	5
2.1 Production and Distribution	5
2.2 Objectives of the Constitutional Assembly	7
3. Constitutions and Influence Activities	9
3.1 The Cooperative Case	10
3.2 The Non-Cooperative Case	16
4. Constitutional Reform	17
4.1 The Self-Seeking Constitutional Assembly	17
4.2 The Biased Constitutional Assembly	19
5. Concluding Remarks	28
Appendix	29
References	31

### **Essay 4:**

#### **The Break-up of the Ruble Zone: Undertaking Monetary Reform while Building Democratic Institutions**

1. Introduction	2
2. Trade and Payments in the Former Soviet Union	7
2.1 The Break-down of the Soviet Payments System	7
2.2 Options for a Market-Based Payments System	9
2.3 Russia's Choice	14
2.4 The Ruble Rally	16
2.5 Inertia in Trade Relations	19
3. The Monetary Mechanism of the Ruble Zone	22
3.1 Previous Analyses	22
3.2 Decisive Features of Membership	24
3.3 A Summary of Pros and Cons of Membership	29
4. Strategic Interaction in the Ruble Zone	31
4.1 The Unrestricted Ruble Zone	31
4.2 The Restricted Ruble Zone	38
5. Policies of Member Countries	47
5.1 Renegotiation of Credit Limits	47
5.2 Introduction of Currency Boards	47
5.3 Emergence of Hyperinflation	49
5.4 Concluding Remarks	51
6. The Political Economy of Russia's Monetary Reform	52
6.1 Strength and Interests of State Enterprise Managers	52
6.2 The Struggle over Russian Monetary Reform	55

6.3 Radical versus Gradual Dissolution	60
6.4 The Relative Success of the Break-up	63
7. Conclusions	66
Appendix	68
References	69



## Acknowledgements

Writing a dissertation tends to be a delicate matter if you do it in a field that is somewhat distant from your advisor's. It could even become a bit cumbersome if your topic also borders on other disciplines. Moreover, if your work treats issues that differ from those usually treated in this border region, you might very well be heading for trouble. An experienced and open-minded advisor could help you avoid most reefs in such murky waters. Lars Jonung has done this for me, in part by raising as many objections as possible to give me ample practice in countering them. He has always done this with a smile, which has helped me enjoy this challenging voyage.

I was fortunate to have Tore Ellingsen as co-advisor. He helped me to improve the quality of my models and to sharpen the interpretations of my results. He also encouraged Lars to accept my work when my own arguments did not suffice. Erik Berglöf offered valuable guidance. He has also made economics of transition, a central field of this dissertation, an important discipline at the Stockholm School of Economics. Karl Wärneryd taught me political economy and inspired me to do research in this field. Anders Åslund, my first advisor, introduced me to the fascinating topic of monetary reform in the former Soviet Union. He also gave me the opportunity to study the Russian reform process from within.

Many other people have contributed to this thesis, of whom I will just mention a few. Örjan Sjöberg acted as an advisor during periods when I did not have one. Together with Sten Luthman, he also contributed to an environment where I learned much about the development in central and eastern Europe. Mats Karlsson gave me an opportunity to deepen my insights into policy making, and allowed me to develop my ideas in a new environment. Vadim Ivanov helped me during a research visit in Russia. Niclas Berggren encouraged my work in constitutional economics and introduced me to fora where I could meet experts in political economy. Gordon Tullock and Axel Leijonhufvud provided my academic life with some unexpected excitement, and my fellow PhD students kept me in good shape both fysically and intellectually. I would also like to thank the staff at SITE and the department of economics, in particular Britt-Marie Eisler, Pirjo Furtenbach, and Gun Malmquist, for helping me with numerous things over the years.

I would finally like to express my sincere gratitude for generous financial support provided by SITE. Co-funding provided by the Swedish Institute and Ollie and Elof Ericssons stiftelse is also appreciated. Even more important was the patient support from Ankita, and recently Alfred.

Stockholm, February 16, 2000

Anders Bornefalk



## Introduction and Summary

The four essays of this dissertation concern the political economy of policy reform. The focus is on conflicts over the distribution of income, and on incentives to allow or to undertake reforms affecting the outcome of such conflicts. This focus is fruitful when it comes to analyzing the delay, or even reversal, of reforms that improve the protection of property rights and the general macroeconomic environment, and thereby lead to large gains for society as a whole in the long term. It also helps us to understand differences in economic performance between different countries undergoing the same type of changes, and why some countries are more plagued by corruption and other types of appropriative activities than others.

The dissertation has a unified theoretical framework where some groups engage actively in activities affecting the distribution of income, whereas other groups are passive in this respect. The explanation as to why the latter groups are passive could be that they have not overcome their collective action problems, as discussed by Olson (1965). This is a general setting, but it is particularly relevant for countries undergoing a transition from authoritarian rule and centralized control over the economy, towards democracy and market economy. The reason for this is found in the unbalanced nature of the political conditions that tends to emerge as such a system collapses. Vested interests representing major industries typically manage to grow strong as the centralized control over the economy is weakened. Independent political organizations, on the other hand, frequently remain undeveloped. Hence, as the authoritarian system collapses, vested interests are well organized and well connected, and face little constraints from political institutions.

The first essay, *Social Conflict with Passive Groups*, develops the basic model used in this dissertation. It extends the framework used in the modern theory of social conflict, as inspired by Oppenheimer (1914), with a group that does not engage in appropriative activities. The introduction of a passive group results in two non-monotonicities with respect to the effectiveness of appropriative activities.

The first type of non-monotonicity relates to the share of their resources that active groups allocate to appropriative activities, and appears when they are able to cooperate with each other. When the appropriative effectiveness is at an intermediate level, active groups need to allocate a relatively large share of their resources to appropriative activities in order to reach a situation where their marginal gains from productive and appropriative activities coincide. For high levels, a relatively small share is enough, since this secures a large enough share of society's total production to make them more interested in productive activities. For sufficiently low levels of appropriative effectiveness, finally, appropriative activities will be absent, since the marginal gains from such activities never exceed those from productive activities.

The second type of non-monotonicity concerns the payoffs of active groups, and results when they fail to cooperate with each other. This non-monotonicity is remarkable, since groups involved in a conflict typically become worse off the higher the appropriative effectiveness is, at least if they are about equally powerful. This is a strong incentive to settle for a peaceful solution, as envisioned by Hobbes (1651).

As a passive group is introduced, the payoffs of active groups are still declining in the appropriative effectiveness for relatively low levels. However, after a point where the appropriative effectiveness has become large enough to make active groups allocate most of their resources to appropriative activities, their payoffs start to increase in the appropriative effectiveness. If the share of initial resources that passive groups control is large enough, the payoffs of active groups will even exceed those received under a peaceful outcome. Hence, a Hobbesian solution to conflict is unlikely to occur if a sufficiently large share of society's resources is in the hands of passive groups.

The second essay, *Democratization, Rent Seeking, and Economic Transition*, employs the basic theoretical framework to determine the available political and economic reform space as well as the economic performance in a society where interest groups are powerful relative to the government. Assuming that people belonging to unorganized groups are more numerous than those belonging to interest groups, and holding the strength of interest groups constant, it is argued that the appropriative effectiveness is lower the more mature independent political organizations are, and the higher the degree of openness in society is. The reason is that the higher the level of development reached by democratic institutions is, the more likely it will be that a government that extends excessive favors to interest groups will be punished by the unorganized majority in upcoming elections.

The political reform space is then determined by the extent to which interest groups are willing to accept reforms that affect the appropriative effectiveness, for instance by allowing free elections to be held. This is, in turn, determined by the payoffs of interest groups under different conditions. The extent to which the appropriative effectiveness can, in fact, be changed through political reforms is also important for the political reform space. This is influenced by the level of development reached by basic democratic institutions, since this determines how political liberalizations will affect the accountability of the government to the majority.

It is found that interest groups could indeed be interested in reforms lowering the appropriative effectiveness if they are unable to cooperate with each other, but only if these reforms are sufficiently far-reaching to increase their payoffs. As made clear in Essay 1, this might be impossible to achieve when the share of the initial resources that passive groups control is large. If interest groups are able to cooperate with each other, they will have no incentives to allow reforms reducing the level of appropriative effectiveness.

The economic reform space follows from the level of appropriative



effectiveness and the political reform space. In other words, the level of protection of property rights and the degree of macroeconomic restraint are adapted to what the political conditions call for.

The economic performance is determined by the share of their resources that interest groups allocate to appropriative activities, or rent seeking. The authoritarian society is taken as a starting point. Appropriative effectiveness is high, which means that the state will only have to waste a relatively small share of its resources on appropriative activities. As authoritarian rule collapses, the deterioration in economic performance will be greatest in countries that undertake some reforms, but not comprehensive enough to support the development of a well-functioning market economy. Countries that manage to undertake a radical strengthening of the protection of property rights and to impose macroeconomic restraint will reach a situation where less resources are wasted on appropriative activities than under the authoritarian system. This follows from the first type of non-monotonicity discussed above.

The conclusions reached in the second essay provide an explanation as to why countries with relatively well-developed democratic institutions have undertaken more successful transitions than countries with less developed democratic institutions.<sup>1</sup> They also provide theoretical support for radical and far-reaching reforms. These are found to be preferable to gradual reforms, and this on three grounds. First, radical reforms lead to less resources being used for appropriative activities, and therefore more for production. Second, radical reforms reduce the extent of redistribution to politically powerful groups. Third, the prospects for undertaking a transition might deteriorate over time if the delay means that interest groups strengthen their ability to cooperate. On the other hand, the analysis has also identified conditions under which radical reforms are risky, if at all possible, to undertake.

The third essay, *Constitutional Constraints and Redistributive Activities*, shifts the focus to constitutional reforms. The basic model is extended with a parameter representing explicit restrictions against redistribution. These could, for instance, take the form of rules determining the share of votes in parliament that is necessary to pass a law affecting the distribution of income.

The analysis shows that the introduction or strengthening of constitutional restrictions against redistribution could increase the extent to which organized groups engage in non-productive activities aimed at influencing the distribution of income. This could come about when organized groups are able to secure such a large share of total production prior to the constitutional reform that they choose to limit the share of their resources that they allocate to influence activities, in order not to hurt their own interests. The reform would reduce the share received by organized groups at the

---

<sup>1</sup> See EBRD (1999), chapters 5 and 6, for an analysis of connections between political conditions and economic reforms during the first ten years of transition in central and eastern Europe, the Baltic states, and the CIS.

initial level of influence activities, which would have an adverse effect on their incentives to limit such activities. By facilitating redistribution through constitutional reform, on the other hand, countries where the initial level of redistributive effectiveness is relatively high could limit influence activities, thereby improving economic performance.

This finding is a potential drawback of constitutional constraints against redistribution. Unlike the drawback analyzed in Buchanan and Tullock (1962), that is, that decision time costs might increase as a consequence of this type of reform, the possibility that influence activities might increase is eliminated if Knut Wicksell's (1896) unanimity rule is imposed.

This drawback of constitutional constraints affect countries where the initial level of redistributive effectiveness is relatively high. Together with the analysis in Essay 2, this suggests that the mechanism analyzed in this essay is an impediment to the introduction of constitutional constraints against redistribution in countries where interest groups are strong and the democratic institutions are poorly developed. A constitutional assembly that gives high priority to the income of unorganized groups might nevertheless choose to undertake a reform of this type, as shown in the essay.

The final essay, *The Break-up of the Ruble Zone: Undertaking Monetary Reform while Building Democratic Institutions*, studies the political economy of monetary reform in the former Soviet Union. The great differences between both preconditions and outcomes for the 15 successor states make the break-up of the ruble zone appropriate as a test of the explanatory power of the theoretical framework developed in the previous essays. The break-up of the ruble zone is also a clear example of the importance of political and economic preconditions for the possibility of undertaking successful economic reforms in countries undergoing a major systemic change.

The varying degree to which the former Soviet republics delayed the delimitation of their currency areas is explained by the mechanisms and factors identified in Essays 1 through 3. It is found that the differences in preconditions determined whether powerful groups preferred a monetary system that facilitated exchange and access to new markets, or one that facilitated redistribution of wealth. Countries where interest groups preferred the latter type of monetary system preferred to remain in the ruble zone, even though this entailed exorbitant costs for the society as a whole.

Particular attention is given to the endeavors of the Russian government to break up the ruble zone. This was vital not only for Russia's change of economic system, but also to avoid hyperinflation. The Russian government was, indeed, weak relative to interest groups, and lacked support from basic democratic institutions. Despite this, it is found that it would have been possible to undertake a more rapid, and therefore less costly, dissolution of the ruble zone. This did not come about, since there was no coherent reform program during the period in which the political conditions would have allowed a radical dissolution. Once such a program had been

developed, the conditions had changed in such a way that Russian state enterprise managers preferred a monetary system that enabled them to acquire rents more effectively. The dissolution of the ruble zone therefore became a lengthy process. It is shown how the Russian reformers in the end managed to overcome the resistance against a dissolution by combining political reforms with reforms of the payments system that divided the interests of powerful groups.

## References

- Buchanan, James M. and Tullock, Gordon. *The Calculus of Consent*. University of Michigan Press, Ann Arbor, 1962.
- EBRD, *Transition Report 1999*, London, 1999.
- Gwartney, James D. and Wagner, Richard E, eds. *Public Choice and Constitutional Economics*. JAI Press Inc., London, 1988.
- Hobbes, Thomas. *Leviathan*. Viking Press, New York, 1982/1651.
- Olson, Mancur. *The Logic of Collective Action*. Harvard University Press, Cambridge, MA, 1965.
- Oppenheimer, Franz. *The State*. University of Toronto Press, Toronto, 1975/1914.
- Wicksell, Knut. "A New Principle of Just Taxation." *Finanztheoretische Untersuchungen*. Jena, 1896, reprinted in Gwartney and Wagner, 1988, pp. 117-130.



## Essay 1

### Social Conflict with Passive Groups



# Social Conflict with Passive Groups

Anders Bornefalk\*

February 8, 2000

## Abstract

This paper studies social conflicts where both groups that engage in appropriative activities and groups that are passive in this respect participate. The introduction of passive groups results in two non-monotonicities with respect to the effectiveness of appropriative activities. The first relates to the share of their resources that active groups allocate to appropriative activities when they cooperate with each other. The second concerns the payoffs of active groups when they are unable to cooperate. We show that these results hold under different methods of cooperation and different types of behavior of passive groups in response to appropriative activities. We also offer an explanation as to why passive groups remain passive.

"These Imrad are the serving class of the Asgars, who live on them, although the Imrad could put into the field ten times as many warriors." (Ratzel, cited in Oppenheimer (1975, p.29).)

## 1 Introduction

Oppenheimer (1914) views the formation of states as a result of domination by one group, or people, over another. The ruling group, typically herds-men, manages to control and extract wealth from the dominated peasants or hunters through superior organization. As this relation evolves, a primitive basis for modern society emerges. Hobbes (1651), on the other hand, views

---

\*Department of Economics and SITE, Stockholm School of Economics, Box 6501, SE-113 83, Stockholm, Sweden. Anders.Bornefalk@hhs.se. This paper has been inspired by discussions with and comments from Kai Konrad and Tore Ellingsen on a related paper. Any errors are my own. I would like to thank SITE for generous financial support.

the establishment of states as a contract between free men to avoid a conflict where all are against all. In the former case, the ruling group becomes better off by developing the effectiveness of its appropriative activities to the point where the dominated group finds it meaningless to resist exploitation. The ruling group can then use its resources to subjugate a new people or group. In the latter case, the warring individuals become better off by surrendering their power to a sovereign, thereby effectively abstaining from appropriative activities.

In the modern theory of social conflict, all participants in the conflict have a possibility of allocating some of their resources to appropriative activities. However, their comparative advantages in productive and appropriative activities may differ. For this reason, some contestants may choose not to engage in appropriative activities.<sup>1</sup> In this paper, we take Oppenheimer's observation one step further by introducing a group that does not engage in appropriative activities under any circumstances. We also extend the setting of Oppenheimer with a second group that can use some of its resources for appropriative activities.<sup>2</sup> In this way, we combine the setups of Hobbes and Oppenheimer.

We refer to the group that does not engage in appropriative activities as passive. The other two groups are active, and can either cooperate in exploiting the passive group or try to appropriate each other's income as well. In the type of society studied by Oppenheimer, the passive group represents the dominated people. The reason why some people become dominated is, according to Oppenheimer, that people engaging in different trades have different incentives and opportunities to develop an ability to undertake collective action effectively. Herdsmen benefit the most from acting together, whereas peasants and hunters perform best alone or in small groups. Over the years, herdsmen develop a superior ability to coordinate their activities. They combine this ability with greater momentum and stamina, if not speed, than people of other trades since they move about with their families and belongings. Together, this turns them into superior warriors. In a modern society, the existence of a passive group can be motivated by some members of society not belonging to any interest group or political party, or not having undertaken investments in becoming organized. This, in turn, could be a result of costs of engaging in collective action and problems with free riding,

---

<sup>1</sup>This could be their optimal strategy if, for instance, their productivity in appropriative activities is small relative to their productivity in productive activities. See Skaperdas (1992) for a treatment of social conflict with asymmetric conflict technologies.

<sup>2</sup>Oppenheimer makes occasional reference to an outside group threatening the domination of the ruling group, but does not consider the implications of its existence and activities in any depth.



as discussed by Olson (1965).

To be able to study social conflict with passive groups, we introduce a conflict technology that is a combination of the frequently applied ratio and difference types of conflict technology. With this combined contest success function (CSF), as with the CSF of difference type, a group that does not allocate any resources to appropriative activities can still receive a share of the prize. In contrast to the difference CSF, the amount of resources controlled by the group also has an impact on the distribution of the prize, or, in this paper, the distribution of income. This gives us a model that can be used to study social conflict in different types of society as well as transitions from one political system to another.

The introduction of a passive group results in two non-monotonicities, and demonstrates the importance of whether or not active groups cooperate with each other. The first non-monotonicity occurs when active groups cooperate with each other, and refers to the extent to which active groups engage in appropriative activities. When the effectiveness of appropriative activities is low, there will be relatively little rent seeking. When the effectiveness is at an intermediate level, rent seeking will be at its maximum levels. When the effectiveness is great, rent seeking will again be rather limited. This has implications for the economic performance and the extent to which corruption occurs under different political and economic conditions.

When active groups do not cooperate with each other, rent seeking becomes endemic, resulting in poor economic performance. It is now the expected payoffs of active groups that are non-monotonous with regard to the effectiveness of appropriative activities. Their payoffs will be lowest at intermediate levels of appropriative effectiveness. This finding indicates when it is possible to undertake political and economic reforms supporting productive activities, that is, protecting property rights, during a major economic crisis. These non-monotonicities are shown to be robust to changes in the way active groups cooperate and to the reaction of passive groups to appropriative activities.

We also find that the profound differences in economic performance and expected payoffs between the cooperative and non-cooperative case could explain why some groups in society choose to remain passive. By means of a numerical example, we show that passive groups might not become better off by actively resisting exploitation, even if they could engage in appropriative activities with equally great effectiveness as already active groups, and without having to invest in becoming organized. If the contenders had a possibility of escaping from the new non-cooperative situation through a Hobbesian solution, however, the expected payoffs of formerly passive groups could increase as a result of their becoming organized. Without this possi-

bility, they would be better off by reacting to appropriative activities by lowering their contribution to production, since this changes the trade-off between productive and appropriative activities of active groups in favor of the former type.

The paper proceeds with a presentation of the model. The extent of rent seeking and the expected payoffs are calculated in Sections 3 and 4 for the cooperative and non-cooperative case, respectively. The question of how passive groups could improve on their situation is addressed in Section 5. The results and possible applications are discussed in a concluding section.

## 2 The Model

### 2.1 Production and Distribution

We model social conflict as an imperfectly discriminating contest over the distribution of income. The outcome is the assignment to each contender of an expected share of total income. The contest success function (CSF) employed in this paper is of the general type

$$p_i(z_i, y_i) = \frac{f(z_i, y_i)}{\sum_{i=1}^m f(z_i, y_i)}, \quad (1)$$

where  $p_i(z_i, y_i)$  is contender  $i$ 's expected share of total income.  $f(z_i, y_i)$  is the functional form determining the impact of the group in the social conflict.  $z_i$  is the amount of resources that is controlled by group  $i$ , and  $y_i$  the amount of resources that is allocated to appropriative activities by this group. Applying the restriction  $y_i \leq z_i$ , the amount of resources that is used for productive activities by group  $i$  is given by  $x_i = z_i - y_i$ .

We use the specific functional form

$$f(z_i, y_i) = z_i e^{ky_i}, k \geq 0. \quad (2)$$

This form satisfies the requirements that  $p_i(z_i > 0, 0) > 0$ ,  $p_i(0, y_i) = 0$ , and  $\frac{\partial p_i(z_i, y_i)}{\partial z_i} > 0$ . Hence, a group that controls a positive amount of the initial resource receives a positive share of the contested prize even if it does not allocate any resources to appropriative activities. This share will be greater the more resources the group controls.

Equation (2) is a combination of a functional form of ratio type and one of difference type. The former type is given by  $f_r(x_i) = x_i^r$ ,  $r > 0$ , where  $x_i$  denotes the extent of appropriative activities. This type of CSF was introduced by Tullock (1980). We confine our treatment to the case  $r = 1$ .

This functional form falls out as a special case of equation (2) when  $k = 0$ . The latter type is given by  $f_d(y_i) = e^{ky_i}$ . It falls out as a special case of equation (2) when  $z_i = 1$ . It has been examined in, for instance, Hirshleifer (1989).<sup>3</sup>

The property that the share of the prize could be positive even if no resources are used for appropriative activities originates from the CSF of difference type. With the CSF of difference type, however, a contender who is passive in a social conflict will receive an equally large share regardless of the size of resources that the contender controls. If, for instance, the contender is a person who is inactive in a political struggle over the distribution of income, he will receive the same share of society's resources as if the contender consisted of all persons who are inactive in this contest. Another problem with the difference form, which is pointed out in Skaperdas (1996), is that  $y_1 = 1$  and  $y_2 = 2$  yield the same winning probabilities as  $y_1 = 1001$  and  $y_2 = 1002$ . The combination of the ratio and difference CSFs removes these problems.

With the interpretation that  $z_i$  represents the size of total resources that a group possesses, and  $y_i$  the amount of appropriative resources that are used in the conflict, equation (2) could be used to describe how the distribution of income in a modern welfare state is determined. The larger a particular group of identical individuals is, the larger its share of total production should be if they engage in appropriative activities to the same extent. The same reasoning applies to peasants being exploited by vassals or bandits. The larger their fields are, or the more numerous they are, the larger the share should be that they end up keeping, *ceteris paribus*.

Turning to production, we assume that each active group possesses an equally large amount of the initial resource, and normalize this amount to 1. Their decision problem is how large a share of their resources they should allocate to appropriative activities. With  $y_i$  depicting the amount used to affect the distribution of income,  $x_i = 1 - y_i$  represents the amount of resources used for production by group  $i = 1, 2$ .

Passive groups engage in no other activities than productive ones. Letting  $\lambda$  represent the amount of initial resources they possess and using  $y_3 = 0$  in equation (2), their impact in the social conflict is given by  $\lambda e^{ky_3} = \lambda$ . Their contribution to production depends on whether or not they react to appropriative activities undertaken by active groups. We assume that their

---

<sup>3</sup>With a CSF of ratio type, a contender's probability of winning is a function of the ratio of the resources that he allocates to appropriative activities to that of other contenders. When the CSF is of difference type, it is the difference between the amount of appropriative activities undertaken by different contenders that affects their probability of winning.

contribution is given by

$$x_3 = \lambda \left( 1 - \gamma \left( \frac{y_1 + y_2}{2} \right) \right), \quad (3)$$

where  $\gamma \in [0, \dots, 1]$ . When  $\gamma = 0$ , the contribution of passive groups to production is simply given by  $x_3 = \lambda$ . This is the base case of our model. We will also consider the case where  $\gamma = 1$ , which we refer to as full reaction. This means that passive groups react to rent seeking in proportion to their own amount of the initial resource and a weighted sum of the amount of resources used for appropriative activities by active groups.<sup>4</sup> One way of interpreting the reaction is that passive groups transfer resources to a sector that is beyond the reach of active groups. Another is that there are dead weight losses brought about by attempts at extracting rents, such as negative effects on the incentives of passive groups to exert effort in production when they realize that some, or much, of what they produce will go to others. The alternative with reaction to appropriative activities shows that our results do not rely on an assumption that the effort of passive groups is completely inelastic to the size of their resulting income.

We now have the inputs to production, which we assume is undertaken according to the production function

$$C(x_1, x_2, x_3) = \sum_{i=1}^3 x_i. \quad (4)$$

## 2.2 Cooperation

We consider two methods of cooperation. In the first type, active groups pool their resources and are able to fully exploit the potential of their common resources, as reflected by the convexity of the difference part of the CSF. In other words, they engage in the conflict as one actor with two units of the initial resource and a maximum of two units used for appropriative activities. We refer to this method of cooperation as additive.<sup>5</sup> The impact of the coalition under additive cooperation is

$$f^{AC} = (z_1 + z_2) e^{k(y_1 + y_2)} = 2e^{2ky}, \quad (5)$$

where we have used the fact that symmetry implies that  $y_1 = y_2 \equiv y$ .

The other method of cooperation is less far-reaching. The cooperating groups are unable to exploit the gains from convexity. They take part in

---

<sup>4</sup>We divide by 2 to avoid the possibility of negative contributions to production.

<sup>5</sup>This is the method used by Skaperdas (1996) to model coalition formation.

the conflict as separate entities, and divide the shares they have managed to secure between themselves. Using symmetry, we obtain the following impact of the coalition under what we refer to as additive functional cooperation:

$$f^{AFC}(z_i, y_i) = z_1 e^{ky_1} + z_2 e^{ky_2} = 2e^{ky}. \quad (6)$$

Additive functional cooperation could be interpreted as cooperation between entities that are unable to communicate effectively with each other during the conflict. They will therefore not be able to coordinate their activities as well as if they were able to communicate. The groups could also lack trust for each other, and therefore prefer to operate on their own since they do not want to rely on a potential rival. The ability to communicate with, and trust, people belonging to the same group is likely to be greater than with people belonging to other groups because of, for instance, more extensive previous interaction. By examining this alternative method of cooperation, we see that our results do not rely on an assumption that people belonging to different groups are equally effective when they cooperate as people belonging to the same group.

### 3 The Cooperative Case

We will now determine the extent to which active groups engage in appropriative activities when they are able to cooperate. We concentrate on the case where passive groups do not react to rent seeking and active groups are able to exploit convexities fully. The analysis of the alternative formulations shows that our results are robust against changes weakening the position of active groups.

The expected payoff of each active group when they cooperate additively and passive groups do not react to appropriative activities is given by equations (1), (4), and (5):

$$\begin{aligned} V_A^{AC}(y) &= p_a^{AC}(y) C(y, \lambda) = \frac{e^{2ky}}{2e^{2ky} + \lambda e^{ky_3}} (2(1-y) + \lambda) \\ &= \frac{e^{2ky}}{2e^{2ky} + \lambda} (2(1-y) + \lambda), \end{aligned} \quad (7)$$

where we use  $y_3 = 0$  in deriving the impact of passive groups.

To determine the extent to which active groups engage in appropriative activities, we differentiate equation (7) with respect to  $y$ , set the resulting expression equal to zero, and solve for the optimal level of  $y$ . We refer to this value as  $y^*$ . Figure 1 shows  $y^*$  as a function of  $k$  for the case  $\lambda = 2$ :

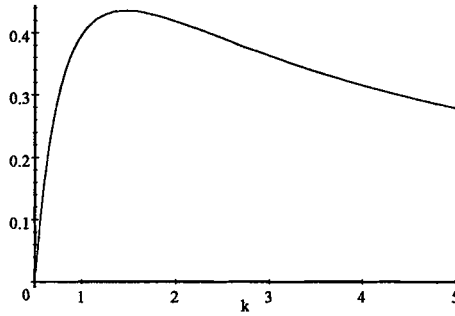


Figure 1. Appropriative activities under additive cooperation for  $\lambda = 2$ .

The non-monotonicity of  $y^*$  with respect to  $k$  is a general feature of social conflict with passive groups when active groups cooperate with each other. Starting with the lower end of the scale, there will always be a  $k$  low enough to result in a situation where active groups use all their resources for productive purposes. We refer to the maximum level of  $k$  resulting in  $y^* = 0$  as  $k_{cutoff}$ . Under additive cooperation and without reaction to rent seeking,  $k_{cutoff}^{AC}(\lambda) = \frac{1}{\lambda}$ . In the case of Figure 1,  $k_{cutoff} = \frac{1}{2}$ . For levels of  $k < k_{cutoff}$ , appropriative activities yield a lower marginal increment to the expected payoffs of active groups than productive activities, since the increase in the share of production received is too small to compensate for the loss in total production brought about by the use of resources for appropriative activities. For somewhat greater levels of appropriative effectiveness, it pays for active groups to use a limited amount of their resources for appropriative activities.

When  $k$  is high, active groups find it optimal to limit their appropriative activities to relatively low levels. The reason for this is that this will be enough for them to acquire such a large share of production that any further expenditures on appropriative activities would result in a greater loss, brought about by the drop in production, than the greater share could compensate for. In the limit, active groups will be able to secure all production by allocating an infinitesimal amount to appropriative activities. For intermediate levels of  $k$ , active groups need to allocate a relatively large share of their resources to appropriative activities in order to reach a situation where their marginal gains from productive and appropriative activities coincide.

The non-monotonicity remains regardless of the size of resources that are controlled by passive groups relative to that of active groups. Again starting with the lower end of the scale, the greater  $\lambda$  is, the closer  $k$  needs to be to zero to deter appropriative activities. The reason for this is that the loss of production brought about by appropriative activities becomes less harmful to the active groups themselves as their share of initial resources falls, since

this gives them a smaller share of total production. Therefore, engaging in appropriative activities is profitable for lower levels of  $k$ . Nevertheless, no matter how large  $\lambda$  is, the range within which  $k$  is low enough to support an equilibrium without rent seeking is non-empty. Hence, rent seeking is absent if  $k$  is low enough.

In addition to a lower cutoff level of  $k$ , a greater level of  $\lambda$  gives a higher level of appropriative activities for all levels of  $k$  unless  $k < k_{cutoff}$  also for the greater level of  $\lambda$ . Active groups do, however, reach a level where they find it optimal to lower their effort on appropriative activities at a lower level of  $k$  when  $\lambda$  is large than when  $\lambda$  is small. In other words,  $y^{\max}$  occurs at a lower level of  $k$ . These differences are brought forward by a comparison between Figure 1 and Figure 2, the latter showing the optimal level of appropriative activities for  $\lambda = 4$ .

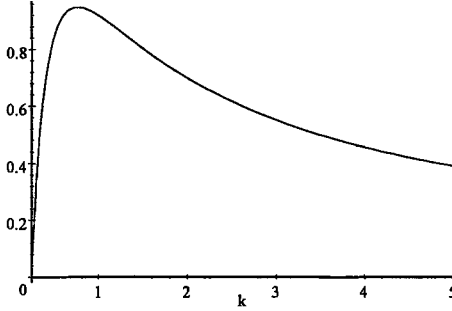


Figure 2. Appropriative activities under additive cooperation for  $\lambda = 4$ .

The findings reported so far allow us to formulate the following proposition:

**Proposition 1** *The extent to which appropriative activities are undertaken when active groups cooperate is non-monotonous with respect to the effectiveness of these activities. When  $k$  is relatively low, appropriative activities will be absent or relatively low; when  $k$  is at an intermediate level, appropriative activities are moderate or relatively extensive; when  $k$  is large, appropriative activities are moderate or relatively low.*

**Proof.** See the Appendix. ■

Proposition 1 also holds when passive groups react to appropriative activities, and when active groups are unable to cooperate fully. What these changes will accomplish is essentially that the extent to which appropriative

activities are undertaken will be reduced throughout the range of  $k$ . See the appendix for details.

Active groups that are able to cooperate are better off the greater the effectiveness of appropriative activities is, unless  $k < k_{cutoff}$  so that  $y = 0$ . This is seen by taking the derivative of the combined impact of active groups with respect to  $k$ . Under additive cooperation, we have:

$$\frac{\partial 2e^{2ky}}{\partial k} = 4ye^{2ky}.$$

This result is not surprising since a higher  $k$  unambiguously improves the opportunities for active groups as long as they cooperate with each other.

## 4 The Non-cooperative Case

In the non-cooperative case, the expected payoff of contender 1 when passive groups do not react to appropriative activities is given by equations (1), (2), and (4):

$$V_1^{NC}(y_1, y_2) = p_1^{NC}(y_1, y_2)C(y_1, y_2) = \frac{e^{ky_1}}{e^{ky_1} + e^{ky_2} + \lambda} (2 - y_1 - y_2 + \lambda), \quad (8)$$

where we again use  $y_3 = 0$ . The expected payoffs of contender 2 and passive groups are given by replacing  $e^{ky_1}$  in the numerator with  $e^{ky_2}$  and  $\lambda$ , respectively. When active groups cannot cooperate, they have to determine the level of rent seeking without knowing the allocation of their opponent. They optimize their expected payoffs given their beliefs about their opponent's decision and the behavior of unorganized producers. Hence, the equilibrium concept is Nash.

To determine the level of appropriative activities, we first derive the best reply function of group 1. Figure 3 shows this function for different values of  $y_2$ , that is,  $y_1(y_2)$ , for  $k = \frac{1}{2}$  and  $\lambda = 2$ .



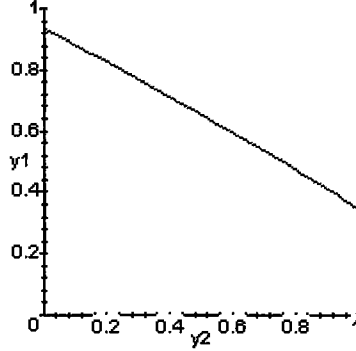


Figure 3. Best reply function of active group 1 for  $k = \frac{1}{2}$  and  $\lambda = 2$ .

We then find the value of  $y_2$  for which  $y_1(y_2) = y_2$  for these values of  $k$  and  $\lambda$ . This occurs when  $y_2 \approx .597$ , which is also indicated by Figure 3. Since the game is symmetric, this yields a Nash equilibrium, which we call  $y_{NE}^*$ . Since this is the only value where both players choose a best response to their opponent's move, it is a strict Nash equilibrium.

Nash equilibria in appropriative activities for  $k$  between 0 and 0.75 and  $\lambda = 2$  are shown in Figure 4.  $y_{NE}^*$  is quite sensitive to changes in  $k$  when passive groups do not react to rent seeking.  $y_{NE}^* = 1$  already at approximately  $k = 0.75$ , as seen in the figure. Compared with the cooperative case without reaction, the cutoff level of  $k$  shifts to the left. More precisely,  $k_{cutoff}^{NC}(\lambda) = \frac{1}{1+\lambda}$ , which in the present case yields  $k_{cutoff} = \frac{1}{3}$ .

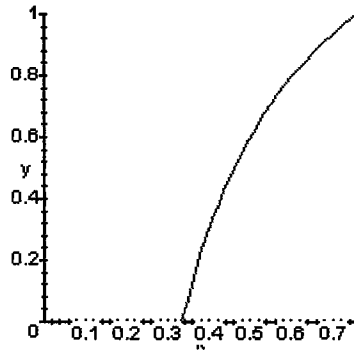


Figure 4.  $y_{NE}^*$  for  $\lambda = 2$ .

The curve tilts downwards when passive groups react to rent seeking. This, in turn, results in the cutoff level shifting to the right. With reaction,  $k_{cutoff}^{NC}(\lambda) = \frac{2+\lambda}{2(1+\lambda)}$ . A greater  $\lambda$ , on the other hand, makes the curve steeper,

and the cutoff level shifts to the left. Figure 5 shows  $y_{NE}^*$  with full reaction to rent seeking and  $\lambda = 2$ . Under these conditions,  $k_{cutoff} = \frac{2}{3}$ . We also have  $y_{NE}^* < 1$  regardless of the level of  $k$ . Active groups will, in fact, never allocate all their resources to appropriative activities under full reaction. If they did, there would be no income to distribute. This stands in sharp contrast to the case with no reaction, since we then have  $y_{NE}^* = 1$  already for low levels of  $k$ . By allowing intermediate values of  $\gamma$ , this sharp contrast is removed.<sup>6</sup>

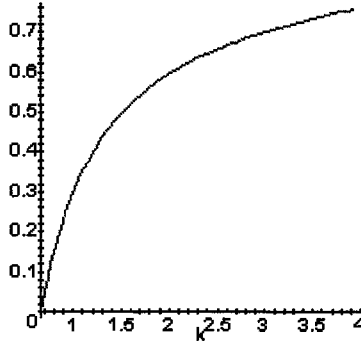


Figure 5.  $y_{NE}^*$  for  $\lambda = 2$  and reaction to rent seeking.

We have seen that rent seeking will be absent if  $k$  is low enough, in both the cooperative and non-cooperative cases. Table 1 reports cutoff levels of  $k$  for different conditions. As seen in the table,  $k_{cutoff}^{NC}(\lambda) < k_{cutoff}^{AC}(\lambda) < k_{cutoff}^{AFC}(\lambda)$ . Hence,  $k$  needs to be lower to deter rent seeking in the non-cooperative case than under both additive and additive functional cooperation, and lower under additive than additive functional cooperation. The explanation as to why  $k_{cutoff}^{NC}(\lambda)$  is lower than under cooperation lies in the inherent insecurity that active groups face when they are unable to cooperate. This means that the distortive effects of appropriative activities must now be weighed not only against the advantages in terms of redistribution from passive groups, but also against the advantages in terms of defence against attacks from the other active group. The reason why the cutoff level is lower under additive cooperation than under additive functional cooperation is that the former type is more effective than the latter type. As a result, the redistributive effects outweigh the distortive effects for lower levels of  $k$ .

We also see that the cutoff level of  $k$  is lower when there is no reaction to rent seeking than when passive groups do react. The explanation for this

---

<sup>6</sup>Bornefalk (2000) studies the possibility that the reaction depends on  $k$ .

is simply that the distortive effects are greater under reaction, which results in redistributive effects being dominated by distortive effects for lower levels of  $k$ .

	$k_{cutoff}^{NC}(\lambda)$	$k_{cutoff}^{AC}(\lambda)$	$k_{cutoff}^{AFC}(\lambda)$
no reaction	$\frac{1}{1+\lambda}$	$\frac{1}{\lambda}$	$\frac{2}{\lambda}$
full reaction	$\frac{2+\lambda}{2(1+\lambda)}$	$\frac{2+\lambda}{2\lambda}$	$\frac{2+\lambda}{\lambda}$

Table 1. The cutoff level of  $k$

**Proposition 2** *There always exists a level of  $k > 0$  below which there is no rent seeking. This level is decreasing in  $\lambda$ , lower when there is no reaction to appropriative activities, and lower in the non-cooperative case than under either method of cooperation.*

**Proof.** See the Appendix. ■

We can also show that rent seeking is more extensive in the non-cooperative case than it is under cooperation, unless  $k$  is low enough to deter rent seeking even in the non-cooperative case. Moreover, rent seeking is at least as widespread when appropriate effectiveness is great as it is when  $k$  is at an intermediate level.

**Proposition 3** *If active groups do not cooperate, they will undertake appropriative activities at lower levels of  $k$  than when they cooperate. The level of appropriative activities is higher throughout the range of  $k$ , and is strictly increasing in  $k$  except when  $y_{NE}^* = 1$  has been reached.*

**Proof.** See the Appendix. ■

We now go on to study the expected payoffs of active groups. We are particularly concerned with finding the conditions under which appropriative activities will make active groups better off than they would be if they could credibly commit to an equilibrium without such activities. Due to symmetry, we have  $V_1^{NC}(y_1, y_2) = V_2^{NC}(y_1, y_2)$ , which we call  $V_A^{NC}(y_{NE}^*)$ . Figure 6 shows  $V_A^{NC}(y_{NE}^*)$  as a function of  $k$  when there is no reaction to rent seeking and  $\lambda = 2$ .

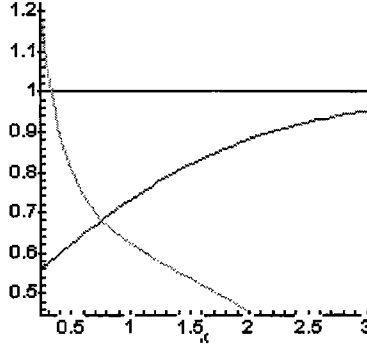


Figure 6.  $V_A^{NC}(y_{NE}^*)$  for  $\lambda = 2$ .

The expected payoffs of active groups that do not cooperate with each other are non-monotonous with respect to  $k$ . This non-monotonicity is a general feature of the non-cooperative case. There are three curves underlying Figure 6. In the interval  $k \in (0, \frac{1}{3})$ ,  $y_{NE}^* = 0$ , and expected payoffs equal 1. In the intermediate interval,  $V_A^{NC}(y_{NE}^*)$  is declining in  $k$  as  $y_{NE}^*$  is increasing in  $k$ . In the interval  $k > 0.758$ ,  $y_{NE}^* = 1$ . Hence, the distortive effects of rent seeking have already reached their maximum. Since the redistributive effects increase in  $k$ ,  $V_A^{NC}(y_{NE}^*)$  is increasing in  $k$  in this interval. No matter how large  $k$  becomes, however,  $V_A^{NC}(y_{NE}^*)$  will fall short of 1 when  $\lambda = 2$ , that is, the expected payoff in the absence of rent seeking. In fact,  $\lim_{k \rightarrow \infty} V_A^{NC}(y_{NE}^* = 1) = \frac{\lambda}{2}$ , which is seen by substituting  $y_{NE}^* = 1$  for  $y_1$  and  $y_2$  in equation (8) and calculating the limit.

For greater values of  $\lambda$  and  $k$ , active groups will become better off if they engage in appropriative activities than if they manage to commit to not doing so. Expected payoffs will not, however, start to increase in  $k$  until  $y_{NE}^* = 1$  when there is no reaction to rent seeking. The explanation for this is that active groups will choose to allocate all their resources to appropriative activities before the effectiveness of such activities has become great enough to make the redistributive effect compensate for the distortive effect.

When passive groups react to rent seeking, the non-monotonicity is manifested in a smooth, or differentiable, curve. See Figure 7 for the case of  $\lambda = 10$ .  $V_A^{NC}(y_{NE}^*)$  is now increasing in  $k$  without the level of appropriative activities having reached  $y_{NE}^* = 1$ . The redistributive effect will dominate the distortive effect when both  $\lambda$  and  $k$  are large enough. Note that active groups manage to reach a level  $V_A^{NC}(y_{NE}^*) > 1$  by engaging in appropriative activities when  $\lambda = 10$  even though there is full reaction to rent seeking. The explanation for this is simply that production is large enough when  $\lambda$  is of this size.

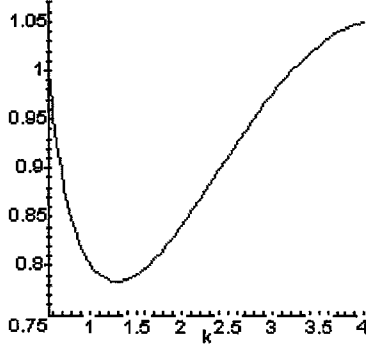


Figure 7.  $V_A^{NC}(y_{NE}^*)$  for  $\lambda = 10$  and full reaction.

The non-monotonicity of  $V_A^{NC}(y_{NE}^*)$  with respect to  $k$  will remain regardless of the size of  $\lambda$ . The expected payoffs of active groups under full reaction will never fall short of 1 when  $k < \frac{1}{2}$  since rent seeking will be absent in this range. This is seen by calculating the limit of the cutoff value of  $k$ . This gives  $\lim_{\lambda \rightarrow \infty} \frac{2+\lambda}{2(1+\lambda)} = \frac{1}{2}$ . This, in combination with  $\lim_{\lambda \rightarrow \infty} \arg \min (V_A^{NC}(y_{NE}^*)) = 1$ , establishes the non-monotonicity of expected payoffs of active groups. See also Figure 8 for an illustration of the case where  $\lambda = 100$ .

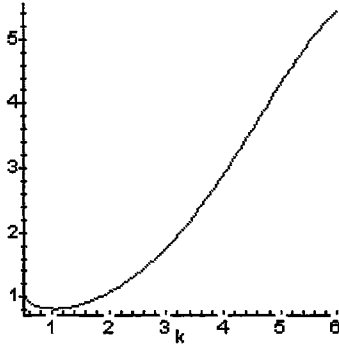


Figure 8.  $V_A^{NC}(y_{NE}^*)$  when  $\lambda = 100$  and there is full reaction to rent seeking.

The non-monotonicity of expected payoffs of active groups and the possibility that  $V_A^{NC}(y_{NE}^*) > 1$  if  $\lambda$  and  $k$  are great enough suggest that active groups that are involved in a conflict with other active groups might not be interested in a Hobbesian solution to social conflict. They might, on the contrary, prefer an increase in  $k$ . Proposition 4, which follows from the reasoning above, tells us when this will be the case:

**Proposition 4** *Active groups that are unable to cooperate with each other will prefer an increase in  $k$  to any alternative if  $\lambda$  and the increase in  $k$  are large enough.*

Hence, a Hobbesian solution, or type of state formation, is unlikely to occur if a sufficiently large share of society's resources is in the hands of passive groups. Active groups will then prefer to exploit these groups rather than to abstain from appropriative activities in return for a peaceful outcome. This is the case even if they thereby continue to be exposed to appropriative activities undertaken by the other active group. Hence, if costs of mobilization and organization are large enough to deter sufficiently many people from mobilizing and undertaking appropriative activities, a Hobbesian solution is unlikely to come about.

## 5 Exit, Voice, or Loyalty?

The existence of a passive group is crucial for the results obtained in Sections 3 and 4. In this section, passive groups will have a choice between remaining exploited and mobilizing to be able to engage in appropriative activities. If they choose not to become organized, they may choose between reacting to appropriative activities of active groups and not doing so. We abstract from costs of mobilization and reaction. Even in the absence of such costs, we will see that mobilization is not necessarily the best action that passive groups can take. Whether it will be the best option is determined by the response of already active groups and the level of appropriative effectiveness. Throughout this section, we assume that the already active groups cooperate additively with each other and that  $\lambda = 2$ .

The expected payoff of passive groups when they do not react to appropriative activities is given by substituting  $\lambda = 2$  for  $e^{2ky}$  in the numerator of equation (7). This yields

$$V_P^{AC}(y) = \frac{2}{2e^{2ky} + 2} (4 - 2y). \quad (9)$$

Calculating the optimal level of rent seeking by active groups, substituting this expression in equation (9), and plotting for  $k \in [\frac{1}{2}, \dots, 5]$  yields the curve to the left in Figure 9.

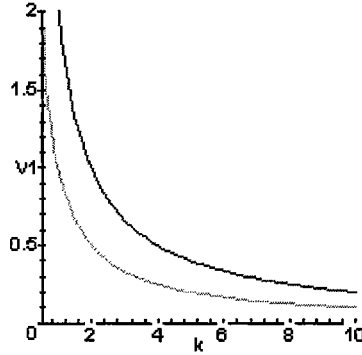


Figure 9. Payoffs of (previously) passive groups under, from the left, loyalty and voice, and exit.

We will now consider the case where the passive groups choose to mobilize and actively resist exploitation. In the terminology of Hirschman, this can be referred to as a voice strategy. If the already active groups choose continued conflict with the previously passive group but continue to cooperate with each other, there will be a conflict between two active groups, each possessing 2 units of the initial resource. The payoffs of the previously passive groups will then stay the same in the numerical example studied here. Hence, the curve to the left in Figure 9 also represents this case. The payoffs of already active groups, on the other hand, fall dramatically. This is seen by comparing the uppermost curve in Figure 10 with the lowermost curve.

Will this outcome be an equilibrium? In other words, is continued conflict a credible threat against mobilization of passive groups? The answer is no if it is possible for the contenders in the new conflict to commit to an outcome where both the previously active and passive groups abstain from appropriative activities. If they are able to do so, the payoffs of each previously active group will be 1, which is considerably more than under continued conflict. The payoffs of the previously passive group will be 2. This means that all resources are used for productive purposes, and that each group receives a share of production corresponding to its contribution.

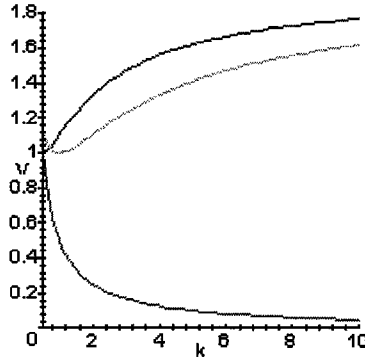


Figure 10. Payoffs of active groups under, from the top, loyalty, exit, and voice.

If the Hobbesian path cannot be treaded for some reason, the best response of passive groups to appropriative activities is to pursue an exit strategy, that is, to react to rent seeking by lowering input to production. This will lower the amount that active groups allocate to rent seeking, which in turn results in a larger expected payoff of passive groups. The curve to the right in Figure 9 shows the expected payoffs of passive groups under such conditions.<sup>7</sup> This suggests that it might be preferable for groups that dislike the redistribution of income that is typical of modern welfare states to, for instance, engage in tax evasion rather than becoming organized in order to limit redistribution by political means.

Active groups will be worse off when passive groups react to rent seeking than when they do not. The expected payoffs of active groups when passive groups react are given by the curve in the middle of Figure 10. This means that active groups have strong incentives to hinder both reaction and organization.

## 6 Concluding Discussion

This paper has shown that the existence of passive groups is important for the outcome of social conflict. If we take problems of collective action seriously, the results of this paper have implications for social conflict in contemporary societies of different types.<sup>8</sup> The results from the cooperative case provide an

<sup>7</sup>This result is robust to the size of  $\lambda$ .

<sup>8</sup>See Bornefalk (2000) for an analysis of the available economic and political reform space in countries undergoing a transition from authoritarian rule towards democracy and market economy.



explanation as to why the economic performance and the degree of corruption differ between different types of society. The results from the analysis of the non-cooperative case allow us to infer when interest groups in societies undergoing a major crisis will be interested in reforms that lower the possibility of these groups to affect the distribution of income to their advantage, or to make other groups bear the costs of stabilization. This has implications for the possibility of undertaking institutional reforms to increase the protection of property rights.

The findings in this paper also indicate some problems with early theories of social conflict. We have seen that a Hobbesian solution is unlikely to come about if a sufficiently large share of society's resources is controlled by groups that do not actively engage in appropriative activities.

Oppenheimer's conclusion, that social order is achieved—in the sense that the extent of appropriative activities falls—as the exploiting group manages to increase its effectiveness in appropriating goods from passive groups, could also be questioned. Our results indicate that Oppenheimer is right when claiming that the extent to which the exploiting group engages in appropriative activities falls as appropriative effectiveness increases. Social conflict could, nevertheless, escalate since passive groups become worse off as appropriative effectiveness increases. Their incentives to mobilize would then increase. If the already established groups would not acquiesce, or agree to, the newly mobilized groups joining the old ones in exploiting the still passive groups, a Hobbesian situation would threaten to develop. Since a smaller and smaller share of society's resources would over time be in the hands of passive groups as this development continued, active groups would in the end have incentives to accept a lowering of the effectiveness of appropriative activities. This means that a Hobbesian solution could come about if Oppenheimer were to be right about appropriative effectiveness increasing over time.

If, on the other hand, appropriative effectiveness were to fall over time, for instance because production became too complicated to facilitate simple extraction, the level of appropriative activities would fall, and the well-being of dominated groups would rise. This indicates that social order, economic development, and a fair distribution of income are best fostered by less effective oppression.

This paper has also shown that the behavior of passive groups is important. The behavior of passive groups includes complicated questions such as whether or not they should invest in becoming organized. We have not addressed these issues explicitly. To the extent that passive groups have reacted to appropriative activities undertaken by active groups, that reaction has been determined without microeconomic underpinnings. A further

step in understanding social conflict where collective action problems matter would be to derive microeconomic underpinnings of the behavior of passive groups.

## Appendix

### PROOF OF PROPOSITION 1:

Differentiating  $V_A^{AC}(y)$  in equation (7) with respect to  $y$  yields:

$$\frac{\partial V_A^{AC}(y)}{\partial y} = \frac{2e^{2ky}}{(2e^{2ky} + \lambda)^2} (2k\lambda - 2ky\lambda + k\lambda^2 - 2e^{2ky} - \lambda)$$

Equating this expression to zero and solving for  $y^*$  as a function of  $k$  and  $\lambda$  yields:<sup>9</sup>

$$y^*(k, \lambda) = -\frac{\text{LambertW}\left(\frac{2}{\lambda}e^{2k+k\lambda-1}\right) - 2k - k\lambda + 1}{2k}$$

Substituting  $\lambda = 2$  yields the curve in Figure 1, and  $\lambda = 4$  the curve in Figure 2. Proceeding in the same manner for different combinations of additive and additive functional cooperation and reaction and no reaction to rent seeking establishes the proposition. See Figure A1 for the case of additive functional cooperation and full reaction to rent seeking for  $\lambda = 2$ .

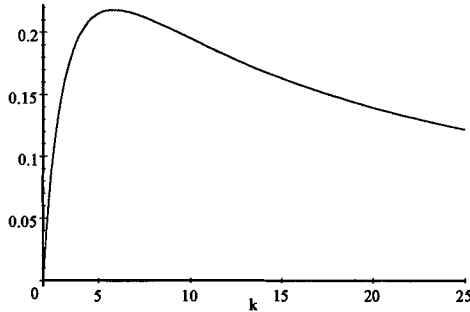


Figure A1.  $y^*$  under additive functional cooperation and full reaction for  $\lambda = 2$ .

### PROOF OF PROPOSITION 2:

To derive the cutoff level of  $k$  under additive cooperation when passive groups do not react to appropriative activities, we first derive combinations of

---

<sup>9</sup>LambertW is an implicitly defined function satisfying  $\text{LambertW}(x)e^{\text{LambertW}(x)} = x$ . The function is named after the 18th century mathematician Johann Heinrich Lambert, who was a colleague of Euler and Lagrange at the Berlin Academy of Sciences. For more on the LambertW function, see Corless et al. (1996).

$k$  and  $\lambda$  that result in a situation where active groups using all their resources for production cannot increase their expected payoffs by allocating a marginal share to appropriative activities. Differentiating  $V_A^{AC}(y)$  in equation (7) with respect to  $y$ , substituting  $y = 0$  into the resulting expression, and equating this to zero, we get:

$$\frac{\partial V_a^{AC}(0)}{\partial y} = \frac{2(k\lambda - 1)}{2 + \lambda} = 0$$

Solving for  $k$  yields:

$$k_{cutoff} = \frac{1}{\lambda}$$

Following the same procedure for other specifications completes Table 1.

### PROOF OF PROPOSITION 3:

Proposition 2 establishes that appropriative activities will be undertaken at lower levels of  $k$  when active groups are unable to cooperate with each other than when they are able to do so. To show that the level of appropriative activities is greater and strictly increasing in  $k$ , we use simulation. To derive the Nash equilibria, we first differentiate equation (8) with respect to  $y_1$ :

$$\begin{aligned} \frac{\partial V_1^{NC}(y_1, y_2)}{\partial y_1} &= \frac{e^{ky_1}}{(e^{ky_1} + e^{ky_2} + \lambda)^2} \times \\ &\times \left( \begin{aligned} &2ke^{ky_2} + 2k\lambda - ky_1e^{ky_2} - ky_1\lambda - ky_2e^{ky_2} \\ &-ky_2\lambda + k\lambda e^{ky_2} + k\lambda^2 - e^{ky_1} - e^{ky_2} - \lambda \end{aligned} \right) \end{aligned}$$

Equating to 0 and solving for  $y_1$  as a function of  $y_2$ ,  $k$ , and  $\lambda$  yields the following best reply function:

$$y_1 = \frac{-\text{LambertW}\left(\frac{1}{e^{ky_2} + \lambda} \exp(2k + k\lambda - ky_2 - 1)\right) + 2k + k\lambda - ky_2 - 1}{k}$$

We then find values of  $y_2$  for which  $y_1(y_2) = y_2$ . Plotting these values for  $\lambda = 2$  and a range of  $k$  yields the curve in Figure 4. Going through the same steps for the case when passive groups react to appropriative activities yields the curve in Figure 5.

## References

- [1] Bornefalk, Anders. "Democratization, Rent Seeking, and Economic Transition." in *Essays on Social Conflict and Reform*, Ph.D. Dissertation, Stockholm School of Economics, 2000.
- [2] Buchanan, James M., Tollison, Robert D., and Tullock, Gordon, eds. *Toward a Theory of the Rent Seeking Society*. Texas A&M Press, College Station, TX, 1980.
- [3] Corless, Robert M. et al. "On the Lambert W Function." *Advances in Computational Mathematics*, 1996, 5, 329-59.
- [4] Hirschman, Albert O. *Exit, Voice, and Loyalty*. Harvard University Press, Cambridge, MA, 1970.
- [5] Hirshleifer, Jack. "Conflict and Rent-seeking Success Functions: Ratio vs. Difference Models of Relative Success." *Public Choice*, 1989, 63, 101-112.
- [6] Hobbes, Thomas. *Leviathan*. Viking Press, New York, 1982/1651.
- [7] Olson, Mancur. *The Logic of Collective Action*. Harvard University Press, Cambridge, MA, 1965.
- [8] Oppenheimer, Franz. *The State*. University of Toronto Press, Toronto, 1975/1914.
- [9] Skaperdas, Stergios. "Contest Success Functions." *Economic Theory*, 1996, 7, 283-90.
- [10] Skaperdas, Stergios. "Cooperation, Conflict, and Power in the Absence of Property Rights." *American Economic Review*, September 1992, 82, 720-39.
- [11] Tullock, Gordon. "Efficient Rent Seeking." in Buchanan et al., eds., 1980.

## Essay 2

### Democratization, Rent Seeking, and Economic Transition



# Democratization, Rent Seeking, and Economic Transition

Anders Bornefalk\*

First version: April 30, 1998  
This version: February 8, 2000

## Abstract

This paper presents a model determining the available political and economic reform space as well as the economic performance in a society where interest groups are powerful relative to the state. Two types of determinants are derived. The first concerns the vulnerability of unorganized groups to rent seeking. The second includes the ability of interest groups to cooperate with each other and the share of the economy's resources that they control. The results are used to interpret transition experiences of countries where authoritarian rule has collapsed. (*JEL D72, E25, P26*)

## 1 Introduction

The last few decades have seen a breakdown of authoritarian rule in many countries. Democratization and economic liberalization have frequently, but not always, ensued. The collapse of the socialist system in Eastern Europe and Central Asia entailed particularly fundamental changes. The political and economic conditions also differed widely between these countries as the

---

\*SITE and Department of Economics, Stockholm School of Economics, Box 6501, SE-113 83, Stockholm, Sweden. Anders.Bornefalk@hhs.se. This paper has benefited greatly from comments from Tore Ellingsen, Kai Konrad, and Gérard Roland. Comments from Niclas Berggren, Erik Berglöf, Hans Bornefalk, Carl Andreas Claussen, Guido Friebel, Peter Hedling, Lars Jonung, Örjan Sjöberg, Karl Wärneryd, and participants at seminars at the Stockholm School of Economics and the 1998 Annual Meeting of the EPCS are also appreciated. Any errors are mine. Financial support from the Stockholm Institute of Transition Economics and East European Economies is gratefully acknowledged.

authoritarian regimes fell. This makes them ideal subjects on which to base a generalized analysis, which is what this paper attempts to do.

The initial political conditions seem to have played an important role in determining the available reform space. In countries where the transition towards democracy and market economy has been rapid, there was typically a distinct break with the communist period preceded by mounting political opposition. In countries where the transition has been slower and more costly, or even reversed, there was generally little organized opposition prior to the demise of communist rule.<sup>1</sup> Hence, as the socialist system gave way, there was no clear democratic alternative that could take over.

While this general pattern is clear, we need a better understanding of the mechanisms involved as well as of the interactions between political and other preconditions. The purpose of this paper is to provide such an understanding. In particular, it seeks to determine the available room for political and economic reform as well as the economic performance in a society where authoritarian rule has withered away or collapsed.<sup>2</sup> The question of why authoritarian rule breaks down is not specifically addressed, but can be studied within the framework that we develop. This focus has been chosen since the breakdown of authoritarian rule has often been a consequence of the end of external domination. This, in turn, explains why differences in preconditions could be so great between countries facing the same task.

To clarify our approach and some key assumptions made in this paper, we can divide the development from authoritarian rule to democracy into three main stages. In the first stage, the dictator holds a secure position, has an encompassing interest in the economy, and is able to control the producers. When repression ceases, possibly because the economy becomes more complex, the established producers can develop into powerful interest groups. If they have built close links and manage to maintain them as the authoritarian state withers away, they can secure a dominant position.<sup>3</sup> In

---

<sup>1</sup>The transition has been rapid in, for instance, Poland, the Czech Republic, and Estonia. In other countries, for instance Russia, Ukraine, and Bulgaria, the reforms have been gradual with repeated setbacks. In Belarus, Turkmenistan, and Uzbekistan, finally, little progress has been made in building the institutions of democracy and market economy.

<sup>2</sup>The focus is on the available room for reforms, not the location chosen by the reformers within the available reform space. The location, that is, the extent to which the size of the "window of opportunity" is utilized, is here considered to be determined largely by properties of the reformers, such as their knowledge and decisiveness. Hence, this paper does not attempt to provide an explanation of the outcome of reforms, but rather a scale against which outcomes can be measured.

<sup>3</sup>This is more likely to occur in socialist economies because of the great size of production units and the tight links between them that characterize such economies. Mancur Olson (1994) has argued that the Soviet Union, and then Russia, has undergone a process



this stage, interest groups can achieve great redistribution of wealth from unorganized groups through rent seeking.<sup>4</sup> This is, however, costly in terms of economic performance. Resources controlled by interest groups will be wasted, and people suffering from the redistribution are likely to respond by working less.

The third stage is a democratic society with diversified industries and an abundance of organizations. Lobbying activities of vested interests are balanced by political parties, free elections, and free media. This reflects that the ruler needs both votes from the electorate and funding and other support from organized groups to win elections. The direction that redistribution takes is then essentially determined by the logics of majority rule, with the extent of redistribution being determined by factors such as voter turnout, age structure, and the pre-tax income of the decisive voter relative to the mean pre-tax income.<sup>5</sup>

To the extent that the above description of the political transition is correct, we can regard the interim stage as a period of anarchy in the spirit of Hirshleifer (1995).<sup>6</sup> The state is weak and cannot protect property rights, while interest groups representing major industries are powerful and have ample opportunities for rent seeking.<sup>7</sup> Since democratization means that the already established interest groups lose much of their political influence to unorganized groups and to emerging interest groups, and consequently much of their possibilities to extract rents, it is unclear why they would allow a move from the second towards the third stage.

The growing literature on democratization has given little attention to

---

of this kind after the death of Josef Stalin.

<sup>4</sup>The following passage from Mau (1994, p. 436) on the policies of the Russian government in the summer of 1992 is a telling account of the weakness of governments relative to interest groups during the initial period of transition: ". . . the government was in fact weak, and had to furnish cheap credit to the enterprises not in accordance with any political criteria, but simply in proportion to the lobbying force of this or that firm (or region), without any other reasons for bias toward any of them."

<sup>5</sup>See Lindert (1996) and Meltzer and Richard (1981).

<sup>6</sup>Anarchy is not chaos but rather a spontaneous order. It is also a matter of degree. Therefore, the analysis of this paper can be extended to apply to other types of society.

<sup>7</sup>As hinted at in the description of the political transition, it is not necessarily always the case that interest groups strengthen their position during the transition: their power might wither away alongside that of the authoritarian ruler. The reasons why we still assume that interest groups are powerful are, first, that there is abundant empirical evidence of their strength in many of the countries we study. Second, we seek a firmer basis to our conclusions about the available reform space than an assumption that the government is free to undertake major reforms against the will of interest groups before democratic elections have been held, and therefore require that they approve of reforms.

this question.<sup>8</sup> Its theoretical foundation is therefore incomplete. For instance, in his seminal work, Przeworski (1991) discusses the design of a democratic institutional framework, taking the opportunity to introduce democracy as given. Likewise, research on the relationships between democratization and economic reform lacks a coherent theory. This research is therefore unable to give precise explanations of the patterns which it has identified. One such pattern, emphasized by Haggard and Webb (1994, p. 7), is that new democratic leaders facing crises managed to stabilize and undertake structural reform most successfully when they moved swiftly. But Haggard and Webb do not explain under what conditions it is possible to undertake rapid changes.<sup>9</sup>

In order to determine the political and economic reform space following the breakdown of authoritarian rule, we employ a model building on Skaperdas (1992), Hirshleifer (1989), and Tullock (1980).<sup>10</sup> The basic difference of the model we use from others studying cooperation, conflict, and power under anarchy is that we include a group that does not take part in rent seeking because of difficulties in becoming organized.<sup>11</sup> We also allow for a possibility of cooperation between interest groups. These changes allow us to study the interactions between organized and unorganized groups.

We find three major determinants of the available reform space and the economic performance. These are the ability of interest groups to cooperate with each other, the share of the economy's initial resources that they control, and the vulnerability of unorganized groups to rent seeking. These factors determine the shares of their resources that interest groups optimally devote to productive and appropriative activities, respectively. They therefore determine the available economic reform space and the economic performance

---

<sup>8</sup>Note that this problem differs from the classic problem of how cooperation between individuals, or even groups, can be achieved. The solution is, however, inspired by this field of political philosophy.

<sup>9</sup>Åslund, Boone, and Johnson (1996) find the same pattern in an analysis of economic reforms in the formerly socialist countries. Their analysis shows both theoretical weaknesses mentioned above. They acknowledge that the state was generally weak following the collapse of the socialist system, and that the opportunities for rent seeking were enormous. Nevertheless, they describe the political situation as one of a "political vacuum" in which political leaders "generally had leeway to design and carry out policies as they chose" (*ibid.*, p. 262). The only explanation as to how the political vacuum could come about is the statement that radical reforms are possible to undertake given a popular mandate in support of reform. But why would powerful interest groups allow a popular mandate to be given, that is, allow a move towards democracy, if this deprives them of much of their possibilities for rent seeking?

<sup>10</sup>See Bornefalk (2000a) for a more comprehensive analysis of the model employed here.

<sup>11</sup>For a brief review of some of these models, see Hirshleifer (1995, footnote 5).

during transition.<sup>12</sup> Since they also determine the payoffs of interest groups under different conditions, they tell us whether these groups will accept political reform.

Our major conclusion is that it is advantageous for interest groups to allow democratization if they are unable to cooperate, if the share of total resources that they control is large enough, and if the vulnerability of unorganized groups to rent seeking is low enough. This case could be described as a Hobbesian conflict in which interest groups would be willing to allow the advent of a democratic Leviathan less attentive to rent seeking in order to escape a devastating conflict.<sup>13</sup> The analysis also shows that democratization makes radical economic reform possible. If the conditions above are not fulfilled, interest groups prefer a return to an authoritarian society. The economic reform space could then also be severely restricted.

The paper proceeds as follows. Section 2 presents the model. In Section 3, the model is interpreted in terms of transition from an authoritarian system to democracy and market economy. The available economic reform space and the economic performance are derived in Section 4. In Section 5, we analyze the possibility of undertaking political reform under anarchy. In the final sections, we apply our results to interpret transition experiences of a selection of countries, discuss policy implications, and give research proposals.

## 2 The Model

We consider an economy consisting of organized and unorganized producers. The organized producers form two interest groups, labeled 1 and 2. Each interest group possesses an equally large amount of initial resources. This amount is normalized to one. Interest groups can use their resources for production or to affect the distribution of goods. Letting  $y_i$  denote the share devoted to affecting the distribution by group  $i = 1, 2$ ,  $x_i = 1 - y_i$  depicts the share used for production.

The unorganized producers are more numerous and possess a total of  $\lambda$  units of initial resources. They do not engage in rent seeking, but are potentially of great political importance because of the larger number of people belonging to this category. Their label is 3.  $\lambda$  is assumed to be exogenous, reflecting prohibitive costs of becoming organized during the limited time

---

<sup>12</sup>Other factors affecting the decision of established producers on whether or not to support reforms include their rate of time preference and the issue of timing of political and economic reform.

<sup>13</sup>See Hobbes (1651).

span covered by our model.<sup>14</sup>  $\lambda$  reflects the degree of centralization of production, the size of production units, and the degree of unionization. We will, for the most part, consider cases where  $\lambda \geq 2$ . This ensures that unorganized producers control a share of initial resources corresponding to a society where there has been little freedom of organization, and therefore little time to overcome difficulties in becoming organized, for groups other than industries.

The share of production received by a certain group is determined by the share of total resources that the group controls, the expenditures on rent seeking, and the effectiveness of appropriative activities. The share of each interest group is given by

$$p_i^{NC}(y_1, y_2) = \frac{e^{k_p y_i}}{e^{k_p y_1} + e^{k_p y_2} + \lambda e^{k_p y_3}} = \frac{e^{k_p y_i}}{e^{k_p y_1} + e^{k_p y_2} + \lambda} \quad (1)$$

for  $i = 1, 2$ , where we use that  $y_3 = 0$ .<sup>15</sup> The share received by unorganized producers is given by replacing the numerator in equation (1) with  $\lambda$ .

In the absence of rent seeking, the shares of total production received by different groups, or the equilibrium win probabilities, correspond to their respective contributions, as given by  $\lambda$ . Hence, if  $\lambda = 2$ , interest groups receive one fourth each, and unorganized producers half the production. When rent seeking takes place,  $k_p$  measures the appropriative effectiveness. In Section 3, we discuss how  $k_p$  differs between different types of society.

Production in this simple economy is given by adding inputs of different groups:

$$C(x_1, x_2, x_3) = \sum_{i=1}^3 x_i. \quad (2)$$

We will also refer to this as the economic performance.<sup>16</sup> As mentioned above, the input of each organized group is given by  $1 - y_i$ .

<sup>14</sup>Costs are high for reasons developed in Olson (1965).

<sup>15</sup>The conflict technology employed here is a combination of Gordon Tullock's (1980) contest success function of ratio type and a conflict technology belonging to the logistic family with the general form  $p_i = \frac{e^{k y_i}}{\sum_j e^{k y_j}}$ , where  $e$  is the base of the natural logarithm.

As explained by Hirshleifer (1989, p. 104), the latter type of contest success function is of difference type, and applies to combat markets with imperfections. If there are imperfections, zero resource commitments need not imply that the player loses everything, which is a requirement for the existence of an equilibrium without rent seeking. An important feature of our formulation is that it allows the amount of resources that a group controls to affect the distribution. This property comes from the ratio part.

<sup>16</sup>It might be argued that this linear and additively separable production function is overly simplistic. Under anarchy, however, the loss due to omitting investments and interrelations between inputs from different groups is likely to be relatively small, because firms

We consider three alternatives for the behavior of unorganized groups. The first is that they use all their initial resources for production. The second is that there is a decline in their contribution to production, proportional to the extent of rent seeking undertaken by interest groups. In the third alternative, the reaction is restricted by a variable  $k_m$ , which measures the extent to which interest groups are able to prevent the reaction. This will be our base case, with the alternative formulations used to check the robustness of our results. In the base case, the contribution of unorganized groups to production is given by

$$x_3 = \lambda - \frac{\lambda}{2k_m} (y_1 + y_2), \quad (3)$$

where the 2 in the denominator means that unorganized groups react to rent seeking by a weighted average of  $y_1$  and  $y_2$ .

There are two main arguments for assuming that the contribution of unorganized groups is declining in rent seeking. The first is that rent seeking lowers their share of output. It is therefore likely to have the same type of incentive effect as taxation in general. The second reason is that reaction to rent seeking is a better alternative for unorganized groups than to inelastically use all their resources for production. The reason for this is that organized groups will have to take their reaction into account when determining the shares they allocate to rent seeking. As a result, appropriative activities, and therefore also redistribution, fall.<sup>17</sup>

The main reason for assuming that the reaction is declining in  $k_m$  is that interest groups have incentives to limit the reaction of unorganized groups since reaction makes them worse off. The larger  $k_m$  is, the harder it is for unorganized groups to detect and react to rent seeking. When interpreting our results, we think of the inverse of  $k_m$  as the degree of openness in society. In order to rule out negative production, we apply the restriction  $k_m \geq 1$ .<sup>18</sup>

We can think of the remainder of the resources of unorganized groups as being diverted either to an informal sector beyond the reach of appropriative activities, or to leisure. Yet another interpretation is that there are deadweight losses brought about by appropriative activities. What matters

---

will tend to invest little and rely upon themselves since property rights are not secured.

<sup>17</sup>See Bornefalk (2000a), who also shows that reacting to rent seeking could yield a better outcome than becoming organized and participating in the struggle over distribution even if it is costless to become organized.

<sup>18</sup>This approach can be compared to that of Ellingsen (1991). In an analysis of the social cost of monopoly, he assumes that buyers can become organized and undertake costly lobbying against monopolies. In a transition context, however, there is insufficient time for large groups to overcome difficulties in becoming organized.

for our purposes is that interest groups receive no utility from this part of initial resources, and that their employment does not affect the distribution of goods produced according to equation (2).<sup>19</sup>

In the conflict technology represented by equation (1), interest groups also compete with one another in the struggle over distribution. They would be better off if instead they could commit to cooperate in extracting rents from unorganized producers. We assume that interest groups collude on a common level of rent seeking, but keep operating as separate entities in the conflict. Letting  $y$  denote the joint share of their resources that interest groups allocate to rent seeking, the share of production received by each group,  $p_1^C(y) = p_2^C(y) \equiv p_{IG}^C(y)$ , is given by

$$p_{IG}^C(y) = \frac{1}{2} \frac{e^{k_p y} + e^{k_p y}}{e^{k_p y} + e^{k_p y} + \lambda} = \frac{e^{k_p y}}{2e^{k_p y} + \lambda}. \quad (4)$$

This method of limited cooperation is appropriate when the cooperating groups are unable, or unwilling, to merge into a new entity.<sup>20</sup> When determining the share of their resources that they allocate to rent seeking, they take into account the effects of their own activities on the other group. The two groups then divide the share of total production that they have secured into two parts. Because of symmetry, these parts will be equally large in our case. The share of unorganized producers is again given by substituting  $\lambda$  for the expression in the numerator.

Our model covers one period and starts immediately after the demise of authoritarian rule. Interest groups make their resource allocations in the beginning of the period, and then produce. The decisions are made under perfect information about the form of the production and conflict technologies and the size of the parameters. Unorganized producers base their decision on how large a share of their resources they should contribute to production

---

<sup>19</sup>With the interpretation that unorganized producers react to rent seeking by diverting resources to the informal sector, the second term on the right hand side of equation 3 represents the share of their resources that are used for production in the informal sector. This amount is increasing in rent seeking since rent seeking lowers the relative payoff to production in the formal sector for unorganized producers. There are two reasons for this. The first is that rent seeking lowers the contribution of organized groups to the formal sector. Hence, there will be less to divide. The second is that rent seeking affects the distribution in favor of interest groups. That this is the case, and that unorganized producers realize this, follows from a rationality assumption. By assuming that there is a productivity differential between the formal and informal sectors, and that there are costs of diverting resources and of investing in becoming organized, this model could be used to examine the degree of organization and the extent of tax evasion in an economy.

<sup>20</sup>Cooperation in conflicts could be modeled in different ways. See Bornefalk (2000a), who discusses what he refers to as additive and additive functional cooperation. The method used in this paper is additive functional cooperation.

on the imperfectly observed allocations of interest groups. At the end of the period, the participants of the game consume the shares of production that they have managed to secure.<sup>21</sup>

### 3 A Transition Interpretation

The model presented in Section 2 can be used to analyze social conflicts with unorganized groups in different types of society. We will now see how certain elements of the model could be interpreted to make it capture essential aspects of a transition from a collapsed authoritarian system to democracy and market economy. We will concentrate on what we refer to as the vulnerability of unorganized groups to appropriative activities and the ability of interest groups to cooperate.

Let us first make two remarks on the applicability of this static and fairly simple model when it comes to understanding the dynamic and complicated phenomenon of transition from one political and economic system to another. The reason why we can use a static model to understand transition is that there are certain key decisions or events that determine the direction of change. Decisions to allow free elections and to abolish the severe restrictions on private economic activities that characterize soviet-type economies belong to this category. What our model tells us in this respect is essentially when interest groups will have incentives to accept such decisions.

Considering the simplicity of the model, our assumptions do not make it easier to arrive at conditions under which powerful interest groups are willing to accept the introduction of democracy and market economy. The requirement that both interest groups agree on changes, for instance, makes it more difficult to find such conditions. In reality, there have been cases where some fairly powerful groups have lost, and others gained, during the transition. Even without the possibility that support from one interest group is enough, we are able to identify conditions under which radical political and economic reforms can be undertaken.

#### 3.1 The Vulnerability of Unorganized Groups

The extent to which unorganized groups are vulnerable to activities undertaken by organized groups differs between different economic and political systems. Our model captures two aspects of this vulnerability. The first is

---

<sup>21</sup>The only assumptions we make about utility functions are that the marginal utility of consumption is positive and that the degree of diversification of production does not affect utilities.

represented by  $k_m$ , which measures the difficulty of diverting resources beyond the reach of appropriative activities. The second is  $k_p$ , the effectiveness of appropriative activities.

The determinants of  $k_m$  include the thickness of the veil behind which rent seeking takes place and the cost of lowering effort. The freedom and impact of mass media and the rate of literacy are determinants of the cost of acquiring information indicating that rent seeking is taking place. The cost of lowering effort could be great because of close surveillance or lack of mobility, for geographic or other reasons. It is, for instance, easier to tax peasants living in landlocked areas than the highly mobile shipping industry. When  $k_m = 1$ , information is freely available and there are no costs of lowering effort in the production of goods that are subject to redistribution.

Moving to appropriative effectiveness,  $k_p$  is, by definition, large in authoritarian states with a powerful ruler. The variable then reflects the ease with which the ruler can extract rents for himself from the economy using resources of various sectors of the economy. He might, however, be threatened by rivals to the throne, but this is just an extreme form of rent seeking. As the grip of the ruler weakens and he becomes dependent on interest groups growing in strength in different sectors,  $k_p$  denotes the responsiveness of the ruler to appropriative activities undertaken by these groups. The responsiveness will be rather large as long as the ruler does not have to be concerned about winning any free elections, and does not risk being criticized by mass media.

The appropriative effectiveness will generally fall as political freedoms have been, or are about to be, introduced. The more mature political parties are, and the better they span the political field, the lower  $k_p$  will be, *ceteris paribus*.<sup>22</sup> The maturity of political parties matters since they typically serve as a way of building and maintaining a reputation for pursuing certain policies.<sup>23</sup> Mass media play an important role in this process by examining and transmitting information on the behavior of political parties. Hence, in countries where political parties have had little time or freedom to develop, and where mass media have a limited reach, appropriative effectiveness will be relatively high even though free elections and freedom of speech have been introduced. The reason for this is that there is little credibility to be lost and therefore parliamentarians and ministers easily become corrupt. It is also unlikely that there are considerably better alternatives available to the electorate. In a mature democracy, on the other hand, a party that is overly

---

<sup>22</sup>This need not be the case if a large share of the voters are organized in interest groups, as they tend to be in modern welfare states.

<sup>23</sup>See Brennan and Kliemt (1994), and Wärneryd (1994).



generous against vested interests will be punished in upcoming elections if evidence comes out in the media. Consequently rent seeking of the kind discussed here results in relatively little redistribution.

The degree to which the legal system is tuned to support productive economic activity and the effectiveness of the public administration in extracting wealth from different groups are other determinants of the possibilities to achieve redistribution. Since we are considering formerly authoritarian societies, the possibility of a constitution effectively protecting property rights can safely be ignored. It can also be assumed that the formerly repressive authorities are still sufficiently strong to achieve the desired degree of redistribution. Hence, private property rights receive little protection by the more permanent aspects of the legal and administrative systems. Other aspects of these systems, such as the formation of specific laws, are more easily changed by the ruler or the parliament. They are therefore secondary to the freedom and maturity of political organizations and the freedom and impact of media.

We assume that  $k_p \geq 1$ .<sup>24</sup> Since the freedom and impact of media affect both  $k_p$  and  $k_m$ , we will assume that  $k_p = k_m \equiv k$  in the analysis. In this way,  $k$  represents the vulnerability of unorganized groups to rent seeking. This emphasizes the strong interconnections between both different aspects of political freedoms and the development of political parties and mass media. The interactions between these two foundations of a democratic society are worth stressing. For one thing, mass media are necessary for political parties to reach the electorate effectively. At the same time, political parties could very well play a role in limiting information costs. Once a political party has built a reputation for pursuing policies that favor a certain group, people belonging to this group could afford to spend less on information gathering and faithfully vote for that party.

The mechanisms we refer to are of a similar nature as those emphasized by Amartya Sen in his entitlement approach to starvation. Drèze and Sen (1990, p. 6) argue that "...it appears that no country with a free press *and* scope for oppositional politics has ever experienced a major famine." Sen (1983, p. 757) explains why China and not India has experienced a gigantic famine in the following way: "...with the present political system in India, it is almost impossible for a famine to take place. The pressure of newspapers and diverse political parties make it imperative for the government in power

---

<sup>24</sup>The lower bound on the responsiveness to rent seeking corresponds to the assumption that  $\alpha = 1$  in the following extended conflict technology:  $p_i^{NG}(y_1, y_2) = \frac{e^{\alpha k_p y_i}}{e^{\alpha k_p y_1} + e^{\alpha k_p y_2} + \lambda}$ , where  $\alpha \geq 0$  and  $k_p \geq 1$ .  $\alpha$  represents the degree of constitutional protection of property rights and the ability of authorities to collect taxes. The assumption that  $\alpha = 1$  can be made in a transition context for reasons mentioned above. See Bornefalk (2000b) for an analysis of the effects of constitutional reform based on the extended conflict technology.

to organise swift relief. It has to act to retain credibility.”

Our discussion on the vulnerability of unorganized groups to appropriative activities is summarized in Figure 1. The maturity of free political organizations and the degree of openness determine what we will refer to as the democratic culture. The more mature free political organizations are and the more open the society is, the greater the depth of the democratic culture will be. Note that the vulnerability is great although interest groups are typically weak in authoritarian societies. During transition, interest groups are generally strong.<sup>25</sup> The difference between what we refer to as Transitions 1 and 2 is found mainly in the depth of the democratic culture. It is therefore essentially the depth of the democratic culture that determines the vulnerability of unorganized groups during transition. To the extent that the strength of interest groups differs between different transition economies, the degree of vulnerability is determined by the balance between the strength of interest groups and the depth of the democratic culture. In mature democracies, finally, the vulnerability of the majority to appropriative activities of the kind modelled here is low because of the impact of the median voter. The reason for this is that, by assumption, he is unorganized.

---

<sup>25</sup>The following examples relating to formerly socialist economies show the weak position of the government and therefore dependence on interest groups during the transition. First, given the poorly developed capital markets, the government needs to keep the budget deficit down and the balance of payments deficit low. Until a tax reform has been implemented and new firms have gained in importance, vested interests control most of taxable production and exports, and can therefore threaten to cause a financial and monetary chaos. Second, interest groups can gain control of much of the media because of their financial strength, which would make them a fierce opponent. For instance, if the government is unwilling to abide by the requirements of factory managers, they can call out their workers on the streets to protest against the government, which would weaken its position further. The weakness of independent media makes it unlikely that a reasonably correct picture of the development will emerge to the workers.

<i>Political system</i>	<i>Free political organizations</i>	<i>Openness</i>	<i>Interest groups</i>	<i>Vulnerability</i>
Authoritarian	Non-existent	Low	Weak	High
Transition 1	Relatively immature	Relatively low	Strong	High / intermediate
Transition 2	Relatively mature	Relatively high	Strong	Intermediate / low
Mature democracy	Mature	High	Weak / strong	Low

Table 1. Vulnerability of unorganized groups to appropriative activities

### 3.2 The Possibility of Cooperation

Commitment to cooperation requires commonality of interests and an ability to overcome prisoner's dilemma type of problems. In an authoritarian system with a little developed economy, the ruler can force producers to cooperate, that is, to make fulfillment of the ruler's wishes their number one priority.<sup>26</sup> As the economy becomes more developed, and therefore complex, it will be increasingly difficult for the ruler to maintain control, and therefore to secure cooperation, through repression. As we will see, this could in itself lead to a deteriorating economic situation bringing the authoritarian system to an end.

During transition, the key to the ability of vested interests to cooperate is found in the structure of output and the way the economy is, or used to be, organized. We could think of the extent to which interests of different industries coincide as being determined by the degree of diversification of production, which in turn corresponds to the level of economic development.<sup>27</sup> If production is diversified, producers have greater incentives to compete with each other for rents than if it is homogenous.<sup>28</sup> Hence, industries in rela-

<sup>26</sup>Note that authoritarian control is possible under both a market and a command system.

<sup>27</sup>See Smith (1776).

<sup>28</sup>For an empirical analysis lending support to this assumption, and to the fundamental importance of interest groups in the Russian economic transition, see Åslund (1994, pp. 300-303).

tively well developed transition economies should have weaker incentives to cooperate.

Considering the *ability* to cooperate, this is determined by the cooperative structures emanating from the authoritarian system and the possibility to invest in structures supporting cooperation, such as common lobbying organizations.<sup>29</sup> As the centralized command system was abolished, each industry had become closely knit together since the economies had been divided into more or less autonomous branches headed by ministries with a high degree of centralized control.<sup>30</sup> Contacts with other groups had, on the other hand, been limited. Still, the fact that entities were internally well organized and large—and therefore relatively few—suggests that it would be easier to establish cooperation with other groups than if this had not been the case. It could nevertheless have been difficult to establish cooperation under such conditions during a limited transition period unless groups representing different industries had invested in establishing cooperation during the demise of the authoritarian system.

Interest groups will be most likely to cooperate in a limited fashion under the conditions mentioned above. Hence, the method of cooperation represented by equation (4) should be appropriate for the situation we are interested in. This method of limited cooperation also enables us to do without an assumption that people belonging to different groups are just as effective when cooperating as people belonging to the same group, regardless of the time spent in investing in this relationship.

## 4 Economic Performance and Reform

In this section, we will derive the economic performance and the available economic reform space. If the level of rent seeking chosen by interest groups and the resulting extent of redistribution are both low, the available economic reform space is wide. It would then be possible for a government to undertake radical economic reforms even though it is weak relative to interest groups. Otherwise, the economic reform space is narrow.  $k$  is taken as given, which leaves no scope for reforms to improve the protection of property rights beyond the available economic reform space. We refer to such reforms as political, and postpone our analysis of them to the following section.

---

<sup>29</sup>Under repeated interaction, a punishment strategy could, in principle, secure cooperation.

<sup>30</sup>See Kornai (1992) for an analysis of the Soviet economic system.

## 4.1 The Cooperative Case

When interest groups cooperate, their expected payoffs,  $V_1^C(y) = V_2^C(y) \equiv V_{IG}^C(y)$ , are given by equations (2), (3), and (4), with  $x_i = 1 - y_i$  and  $y_i = y$  for  $i = 1, 2$ . Each of the two groups receives the expected payoff<sup>31</sup>

$$V_{IG}^C(y) = p_{IG}^C(y)C(y) = \frac{e^{k_p y}}{2e^{k_p y} + \lambda} \left( 2 - 2y + \lambda - \frac{\lambda y}{k_m} \right). \quad (5)$$

Likewise, the expected payoff of unorganized producers is given by

$$V_3^C(y) = p_3^C(y)C(y) = \frac{\lambda}{2e^{k_p y} + \lambda} \left( 2 - 2y + \lambda - \frac{\lambda y}{k_m} \right). \quad (6)$$

To derive the level of rent seeking that interest groups find optimal under cooperation,  $y^*$ , we differentiate  $V_{IG}^C(y)$  with respect to  $y$ , set the resulting expression equal to zero, and solve. With  $\lambda = 2$  and  $k_p = k_m = k$ ,  $y^*$  is a function of  $k$  alone, as shown in Figure 1. The hump shape, or the non-monotonicity of  $y^*$  with respect to  $k$ , is a striking feature of the cooperative case. It is a consequence of rent seeking being non-profitable for sufficiently low levels of  $k$ , and low levels of rent seeking being enough to secure most of production when  $k$  is high.

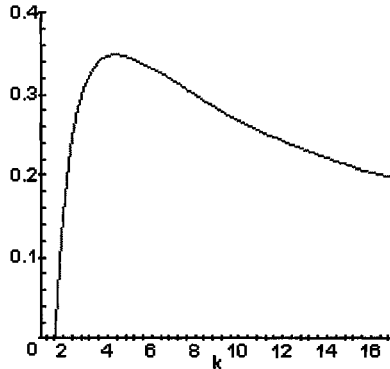


Figure 1. Optimal rent seeking in the cooperative case for  $k \geq 1$ .

The non-monotonicity is a general feature of the cooperative case. It holds also when there is no reaction and full reaction to rent seeking, and when interest groups are able to cooperate fully, that is, to merge into a

<sup>31</sup>Note that the 2 in the denominator of equation (3) disappears in the cooperative case since  $y_1 + y_2$  is replaced by  $y$ .

new entity.<sup>32</sup> The reaction's dependence on  $k$  tilts the curve upwards in comparison with the case of full reaction. Hence, there will be less rent seeking when there is full reaction except at the point where  $k = 1$  where the curves take on the same value.

In our base case where the reaction depends on  $k$  and  $k$  is restricted to the range  $k \geq 1$ ,  $y^*(k)$  will always be hump-shaped if the unorganized producers' reaction to rent seeking is sufficiently great. Since the reaction is related to  $y$  and  $\lambda$ , and since there is nothing deflating the reaction other than  $k$ , this requirement is fulfilled in our case. Another, and closely related, sufficient condition for non-monotonicity is that the share of initial resources that is controlled by interest groups is large enough. This translates to  $\lambda$  being sufficiently small. If there is no reaction to rent seeking and if  $\lambda$  is large, however, rent seeking is at its maximum level already when  $k = 1$  and then falls monotonously as  $k$  increases. But if there is no reaction to rent seeking, there is no point in imposing the restriction  $k \geq 1$ . By extending the range to  $k > 0$ ,  $y^*(k)$  will be non-monotonous in this case also.

Defining levels of  $k \in [1, 2]$  as low, levels of  $k \in (2, 15)$  as intermediate, and levels of  $k \geq 15$  as high, we can formulate the result as in Proposition 1:

**Proposition 1** *If interest groups cooperate, optimal levels of rent seeking are low to moderate when  $k$  is low, moderate to high when  $k$  is intermediate, and moderate when  $k$  is high.*

**Proof.** See the Appendix for a derivation of Figure 1.

For reasons developed in Section 3.1, we can think of relatively low levels of  $k$  as corresponding to a transition country with a considerable depth of the democratic culture, intermediate levels to a transition country where the democratic culture lacks depth, and high levels of  $k$  as corresponding to a transition country with only rudimentary democratic institutions.

Of particular interest in Proposition 1 is the finding that rent seeking is lower when  $k$  is high, that is, when redistributive effects are high and reaction to rent seeking is difficult, than when  $k$  is intermediate. To understand why this is the case, it is helpful to examine expected payoffs, shares received by different groups, and economic performance for different levels of  $k$ .<sup>33</sup> In order to do this, we simply insert  $y = y^*(k)$  and a selection of  $k$  into equations (2), (3), (5), and (6). The result when  $\lambda = 2$  is shown in Table 2.

<sup>32</sup>See Bornefalk (2000a) for details.

<sup>33</sup>Production, or economic performance, can be interpreted as total surplus net of what is required for the survival of producers. The same interpretation applies to payoffs of different groups.

$k$	$y^*$	$V_{IG}^C(y^*)$	$V_3^C(y^*)$	$p_{IG}^C(y^*)$	$C(y^*)$
$\frac{1+\sqrt{5}}{2}$	0	1	2	.25	4
2	.154	1.02	1.50	.288	3.54
3	.313	1.14	.889	.360	3.16
4	.347	1.25	.625	.400	3.13
8	.298	1.52	.281	.458	3.33
15	.212	1.70	.142	.480	3.55

Table 2. Payoffs and production in the cooperative case for  $\lambda = 2$

When  $k \leq \frac{1+\sqrt{5}}{2}$ , there is no rent seeking and  $V_{IG}^C(y^*) = 1$ .<sup>34</sup> In this interval, the distortive effects of rent seeking dominate the redistributive effects. We will refer to the highest value of  $k$  for which rent seeking is absent as the cutoff level of  $k$ . For values of  $k$  up to around 4, a greater  $k$  leads to more rent seeking. In this interval, the increase in the share received by interest groups more than offsets the effects of the fall in production on the payoffs of these groups. When  $k = 3$ , for instance, vested interests use about 31 percent of their resources for rent seeking. They thereby increase their expected payoffs to 1.14. This improvement comes from the larger size of the share they secure, 0.360 instead of 0.25. Production, however, falls from 4 to 3.16. These changes make the expected payoffs of unorganized producers plummet from 2 to 0.889. For  $k$  greater than around 4, rent seeking is declining in  $k$ . The reason why a greater  $k$  leads to less rent seeking in this interval is that the redistributive effects are great enough to enable interest groups to secure a large enough share of production so that the distortive effects of rent seeking dominate the redistributive effects on the margin. As  $k$  approaches infinity, interest groups will be able to secure almost all net production by allocating an infinitesimal share of their resources to rent seeking.

The result of these dynamics is that economic performance is relatively high for low and high values of  $k$ , and low for intermediate levels of  $k$ . Expected payoffs, on the other hand, have monotonous trends above the cutoff level of  $k$ . The higher  $k$  is, the higher are the expected payoffs of interest groups, and the lower are those of unorganized groups. Finally, the connection to the available space for economic reforms improving economic performance and limiting redistribution is immediate, since vested interests will not allow reforms beyond the level supporting  $V_{IG}^C(y^*)$ . For instance, if seigniorage were to be a preferred way of extracting rents, a credible fixed

---

<sup>34</sup>To see that  $k = \frac{1+\sqrt{5}}{2}$  yields an equilibrium without rent seeking, see the proof of Proposition 2.

exchange rate could only be introduced if  $k$  were low enough.

As noted above, rent seeking is absent for  $k \in \left[1, \frac{1+\sqrt{5}}{2}\right]$  when  $\lambda = 2$ . The upper value in this range, that is, the cutoff value of  $k$ , is a function of  $\lambda$ . The relationship is a simple one: the greater  $\lambda$  is, the closer  $k$  has to be to 1 to deter rent seeking. The reason for this is that the loss of production brought about by rent seeking becomes less harmful to the vested interests themselves as their share of initial resources falls, since this gives them a smaller share of total production. Therefore, engaging in rent seeking could be profitable even when  $k$  is low. Nevertheless, no matter how large  $\lambda$  is, the range within which  $k$  is low enough to support an equilibrium without rent seeking is non-empty.<sup>35</sup> Hence, rent seeking is absent if  $k$  is low enough. This is the second main characteristic of the cooperative case. It is also the message of our second proposition:

**Proposition 2** *When interest groups cooperate, there always exists a level of  $k > 1$  below which there is no rent seeking. This level is decreasing in  $\lambda$ .*

**Proof.** See the Appendix.

## 4.2 The Non-cooperative Case

In the non-cooperative case, the expected payoff of interest group 1 is given by equations (1), (2), and (3):

$$\begin{aligned} V_1^{NC}(y_1, y_2) &= p_1^{NC}(y_1, y_2)C(y_1, y_2) \\ &= \frac{e^{k_p y_1}}{e^{k_p y_1} + e^{k_p y_2} + \lambda} \left[ 2 - y_1 - y_2 + \lambda - \frac{\lambda}{2k_m} (y_1 + y_2) \right]. \end{aligned} \quad (7)$$

The expected payoffs of interest group 2 and unorganized producers,  $V_2^{NC}$  and  $V_3^{NC}$ , are given by replacing  $e^{k_p y_1}$  in the numerator with  $e^{k_p y_2}$  and  $\lambda$ , respectively. When interest groups cannot cooperate, they have to determine the level of rent seeking without knowing the allocation of their opponent. They optimize their expected payoffs given their beliefs about their opponent's decision and the behavior of unorganized producers. Hence, the equilibrium concept is Nash.

The first step in determining the Nash equilibria is to derive a best reply function. Again assuming that  $\lambda = 2$  and  $k_p = k_m = k$ , Figure 2 shows the best reply function of interest group 1 for different values of  $y_2$ , that is,

---

<sup>35</sup>This also holds for alternative specifications of the model if the reaction of unorganized groups is great enough, or, if there is no reaction, if we also allow  $k \in (0, 1)$ .



$y_1(y_2)$ , for  $k = 1.5$ . To solve for  $y_1^*$ , the optimal level of rent seeking for interest group 1, when  $k = 1.5$ , we first substitute 1.5 for  $k$  in the expression for  $y_1(y_2)$ , and then find the value of  $y_2$  for which  $y_1(y_2) = y_2$ . This occurs when  $y_2 \approx .674$ , which is also indicated by Figure 2. To derive  $y_2^*$ , we proceed in the same manner. Since the game is symmetric, we must have  $y_2^* = y_1^*$ , which we call  $y_{NE}^*$ . This value, which is a function of  $k$ , is the only one where both players choose a best response to their opponent's move, and is therefore a strict Nash equilibrium. We therefore also have  $V_1^{NC}(y_1, y_2) = V_2^{NC}(y_1, y_2)$ , which we call  $V_{IG}^{NC}(y_{NE}^*)$ . Proceeding in the same way for  $k$  between 1 and 2.25 yields the Nash equilibria in rent seeking shown in Figure 3.<sup>36</sup>

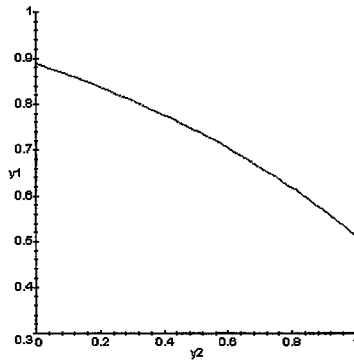


Figure 2. The best reply function of interest group 1 for  $k = 1.5$

We can now formulate Proposition 3. It tells us that rent seeking is much more extensive in the non-cooperative case than it is under cooperation, and that it is at least as widespread in an authoritarian society as it is when  $k$  is at an intermediate level.

**Proposition 3** *If interest groups do not cooperate, rent seeking is high when  $k$  is low, and is strictly increasing in  $k$  until  $y_{NE}^* = 1$ .*

**Proof.** See the Appendix.

<sup>36</sup>Without reaction of unorganized producers to rent seeking, the Nash equilibria would be extremely sensitive to changes in  $k$ . In other words, reaction tilts the curve downwards. The reaction's dependence on  $k$  makes this difference less pronounced. A greater  $\lambda$  makes the curve steeper.

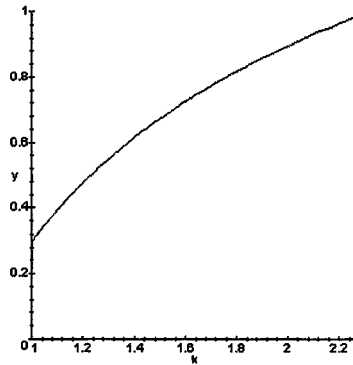


Figure 3. Nash equilibria in rent seeking,  $y = y_{NE}^*$ , for  $k \geq 1$

Let us now compare expected payoffs and economic performance for different values of  $k$ . Consider first the case of a small  $\lambda$ . Setting  $\lambda = 2$  and substituting  $y_{NE}^*$  for  $y_1$  and  $y_2$  for a selection of  $k$  in equations (2), (3), and (7), we get the results in Table 3. For values in the table with  $k \geq 2$ ,  $y_{NE}^*$  is paralyzingly high. For example, when  $k = 2.25$ , vested interests use almost all of their resources for rent seeking.<sup>37</sup> As a result, they secure 45 percent each of (net) production. Although unorganized producers contribute almost all productive effort, they receive as little as 10 percent of (net) production. On top of that, production is far below the value in an economy without rent seeking. Interestingly, the payoffs of vested interests are also falling up to a point where they allocate all their resources to rent seeking. The same is true for production. After this point has been reached, their payoffs and the production are increasing in  $k$  since the unorganized producers' reaction to rent seeking is falling in  $k$ . Despite this increase, vested interests are worse off in any equilibrium with rent seeking than in one without. As the attentive reader remembers, they receive 1 in the latter case.

<sup>37</sup>This might sound unreasonable, but in a transition context, this can be interpreted as there being hardly any effort exerted in adapting production to market demand. Instead, the old type of production is maintained and the labor force kept intact in order to exert maximum pressure on the government to give support to the firm, or even return to the old system. Those who are skeptical about the existence of such high levels of rent seeking in other than transition economies are recommended to read Klitgaard's (1991) account of attempts at reforming the economy of Equatorial Guinea.

$k$	$y_{NE}^*$	$V_{IG}^{NC}(y_{NE}^*)$	$V_3^{NC}(y_{NE}^*)$	$p_{IG}^{NC}(y_{NE}^*)$	$C(y_{NE}^*)$
1	.299	.805	1.19	.287	2.81
1.5	.674	.643	.468	.367	1.75
2	.896	.562	.188	.429	1.31
2.25	.980	.526	.116	.450	1.17
3	1	.635	.063	.476	1.33
5	1	.795	.011	.497	1.6
15	1	.933	.000	.500	1.87

Table 3. Payoffs and production in the non-cooperative case for  $\lambda = 2$

The picture in terms of expected payoffs differs in two respects when interest groups control a small share of initial resources. Table 4 shows the case of  $\lambda = 10$ . The first difference is that  $V_{IG}^{NC}(y_{NE}^*)$  starts to increase in  $k$  at a lower level of  $k$  than when  $\lambda$  is small.<sup>38</sup> The second difference is that  $V_{IG}^{NC}(y_{NE}^*) > 1$  if  $k$  is sufficiently high. For low levels of  $k$ , however, they would still be better off if they could abstain from rent seeking.

$k$	$y_{NE}^*$	$V_{IG}^{NC}(y_{NE}^*)$	$V_3^{NC}(y_{NE}^*)$	$p_{IG}^{NC}(y_{NE}^*)$	$C(y_{NE}^*)$
1	.433	.802	5.20	.118	6.80
1.1	.557	.785	4.26	.135	5.83
1.2	.669	.785	3.52	.154	5.09
1.3	.771	.798	2.93	.176	4.53
1.5	.953	.851	2.04	.228	3.74
2	1	1.49	2.02	.298	5
5	1	3.87	.261	.484	8
15	1	4.67	.000	.500	9.33

Table 4. Payoffs and production in the non-cooperative case for  $\lambda = 10$

The possibility for non-cooperating interest groups that control a small share of initial resources to improve their situation by engaging in appropriative activities is drastically reduced when the reaction of unorganized groups is independent of  $k$ . Table 5 shows that we could have  $V_{IG}^{NC}(y_{NE}^*) < 1$  even for high levels of  $k$  and  $\lambda$  under such conditions. It is still possible that interest groups manage to secure a higher payoff than they would receive without

<sup>38</sup>The payoffs of interest groups are also non-monotonous with respect to  $k$  for large levels of  $\lambda$  when there is no or full reaction to rent seeking and the range  $k > 0$  is allowed.

rent seeking, as seen by  $V_{IG}^{NC}(y_{NE}^*) = 1.03$  when  $k = 5$ . For higher levels of  $k$ , however, payoffs of interest groups are falling in  $k$ , and falling short of what they would receive without rent seeking. The explanation for this is found in the monotonic, and drastic, reduction of production. This indicates that it could be of vital importance for interest groups, or an authoritarian ruler, to reduce the possibility for unorganized groups to react to exploitation.<sup>39</sup> The following proposition captures the possibilities for non-cooperating interest groups to improve their outcome by engaging in appropriative activities under different types of reaction:

**Proposition 4** *If there is limited or no reaction to rent seeking, the non-cooperative payoffs of interest groups will be greater than what they would receive without rent seeking if  $\lambda$  and  $k$  are sufficiently large. If there is full reaction to rent seeking, the non-cooperative payoffs of interest groups could be falling in  $k$  even for high levels of  $k$  and  $\lambda$ .*

**Proof.** See Tables 3, 4, and 5.

$k$	$y_{NE}^*$	$V_{IG}^{NC}(y_{NE}^*)$	$V_3^{NC}(y_{NE}^*)$	$p_{IG}^{NC}(y_{NE}^*)$	$C(y_{NE}^*)$
1	.433	.802	5.20	.118	6.80
1.1	.480	.790	4.66	.127	6.24
1.3	.550	.783	3.83	.145	5.40
1.5	.601	.790	3.21	.165	4.79
2	.680	.841	2.16	.219	3.84
5	.815	1.03	.175	.461	2.23
10	.900	.600	.001	.500	1.20
15	.933	.400	.000	.500	.800

Table 5. Payoffs and production in the non-cooperative case for  $\lambda = 10$  and full reaction to rent seeking

An important similarity between the cases studied in the non-cooperative case is that the high levels of optimal rent seeking and the utterly skewed distribution of income make the available economic reform space severely restricted. Allowing also for  $\lambda < 2$ , we find that there will always be rent seeking if interest groups are unable to cooperate with each other as long as unorganized producers control a positive share of initial resources:

**Proposition 5** *If interest groups do not cooperate,  $y_{NE}^* > 0$  for any combination of  $\lambda > 0$  and  $k \geq 1$ .*

---

<sup>39</sup>The collectivization of agriculture is one example of how communist regimes limited the possibility for workers to reduce their effort in response to appropriation of their production.

**Proof.** See the Appendix.

A key result highlighted by our three cases is that  $V_{IG}^{NC}(y_{NE}^*) < 1$  is possible under a wide range of conditions. Still, vested interests do play a best response when they participate in the struggle for rents under these conditions. Hence, they find themselves in a classical tragedy of commons type of situation. They would be better off if they could commit to an equilibrium with less, or even without, rent seeking. However, if one interest group were to unilaterally commit to increasing the share used for production, its payoff would only worsen. The way out of this dilemma is to allow political reform taking them from anarchy to hierarchy, in the form of either democracy or a resurrected authoritarian system.

## 5 Political Reform

### 5.1 The Democratic Culture

The vulnerability of unorganized groups to rent seeking can be divided into aspects that develop continuously, and aspects that can be changed discretely through political reform affecting political freedoms. The maturity of political organizations and the level of development of mass media belong to the first category, which we refer to as the democratic culture.  $k$  is then a function of the extent of political freedoms and the depth of the democratic culture, where the least developed factor determines its size. This specification allows us to study the implications of the depth of the democratic culture for the possibility of extending political freedoms.

Discrete changes could be either radical or limited.<sup>40</sup> After the demise of authoritarian rule, a radical lowering of  $k$  can be achieved by introducing free elections and free media if one of the following two conditions is fulfilled. The first is that political organizations have already reached a certain level of maturity, and that mass media are already in a position where they can utilize their new freedom effectively. The second is that the country has previous experience of democracy, which makes a rapid reintroduction of democratic institutions possible since people are already familiar with their way of functioning. If neither of these conditions is met, the introduction of free elections and free media will result in a limited lowering of  $k$ . The speed of the continuous development of the democratic culture will increase in both cases.<sup>41</sup>

---

<sup>40</sup>A marginal change of  $k$  represents a continuous development, but this will not be studied here.

<sup>41</sup>Free elections create opportunities for political parties to strengthen their position.

To allow for discrete changes, we assume that there is an agent with agenda-setting power.<sup>42</sup> He can suggest a lowering of  $k$  to  $\underline{k}$ , where the depth of the democratic culture determines the lower bound of  $\underline{k}$ . He can also suggest a jump to  $\bar{k}$ , which corresponds to political reform in an authoritarian direction. This can be accomplished by reducing political freedoms and increasing the repression of political organizations.<sup>43</sup>

In this section, we will for the most part limit our attention to the case where unorganized groups react to rent seeking according to equation (3). The available political reform space would be greater with full reaction, and more limited with no reaction to rent seeking.

## 5.2 The Cooperative Case

When vested interests cooperate and  $k > k_{cutoff}(\lambda)$ , rent seeking increases their expected payoffs. As indicated in Table 1, their payoffs are greater the higher  $k$  is in this range. They will therefore have incentives to promote a movement towards less freedom of organization and more control over media. For levels of  $k < k_{cutoff}(\lambda)$ , their payoffs are always equal to one. These observations take us to our first proposition on political reform.

**Proposition 6** *If interest groups cooperate and  $k > k_{cutoff}(\lambda)$ , they will resist democratization. They will support any move in an authoritarian direction in this range. If  $k < k_{cutoff}(\lambda)$ , interest groups are indifferent to further democratization.*

**Proof.** See the Appendix.

The intuition behind Proposition 6 is quite simple. A higher level of  $k$  facilitates redistribution from unorganized groups and reduces the distortive effects of rent seeking without causing any negative effects, as long as interest groups are able to cooperate. Interest groups will therefore be better off the greater  $k$  is, as long as  $k$  is great enough to make rent seeking profitable. Proposition 6 allows us to conclude that the scope for democratic reforms is quite limited when interest groups cooperate, unless  $k$  is low enough to yield an equilibrium without rent seeking prior to the reforms. The political

---

The adoption of a democratic constitution stating that elections will be recurrent gives political parties incentives to improve their credibility. Likewise, the demand for media coverage increases, enabling mass media to become more developed.

<sup>42</sup>For a discussion of agenda-setting in economic theory, see Dewatripont and Roland (1997) who also employ this concept in a transition context.

<sup>43</sup>Increased repression of labor unions would also result in an increase in  $\lambda$ . This would make other organized groups better off, but will not be studied here.

reform space is more limited the lower the degree of organization is, that is, the higher  $\lambda$  is. The reason for this is that  $k_{cutoff}(\lambda)$  is falling in  $\lambda$ .

If democratic institutions have not become strong enough to withstand attempts at eroding them, democratization set off by the demise of authoritarian rule will be reversed. Interest groups could even *become* the state if they manage to appoint a ruler from their own ranks, who then maximizes the expected payoffs of himself and his cronies.<sup>44</sup> One particularly effective way of doing this is to (re)create a socialist system. But this requires harsh political repression.<sup>45</sup>

When  $k \leq k_{cutoff}(\lambda)$ , anarchy is stable in the sense that interest groups are satisfied with the status quo. Hence, there is no need for a state. Since rent seeking is absent in this case, there is no struggle over distribution either. In other words, we have derived a continuum of utopian communist, or anarcho-liberal, societies.

### 5.3 The Non-cooperative Case

If interest groups are unable to cooperate and  $\lambda$  and  $k$  are not sufficiently high, rent seeking lowers their expected payoffs, as shown in Section 4. By allowing democratization, they can credibly commit to reducing expenditures on rent seeking since the ruler will provide less rents when  $k$  is lower. Since they benefit from less rent seeking, consecutive economic reforms supporting the new equilibrium would also be allowed. Hence, a radical reformer could call for free elections and introduce complete liberalization of the media. If the democratic culture is deep enough, that is, if the resulting decline in  $k$  is sufficiently great, he can also implement a full-fledged economic reform program.<sup>46</sup> In such a case, it would appear as if the radical reformer could operate in a political vacuum, and as if there were a social consensus in favor of reform. It would also appear as if interest groups were not powerful, but this need not be the case. Pursuing a Hobbesian parallel, this could be described as an agreement between powerful interest groups on a social

---

<sup>44</sup>Our result for rent seeking should then be interpreted as the share of resources controlled by vested interests that the ruler, on behalf of his principals, uses for extracting resources from the rest of the economy.

<sup>45</sup>Interestingly, certain countries in the former Soviet Union have followed this extreme track, notably Belarus, Turkmenistan, and Uzbekistan.

<sup>46</sup>For instance, a fixed exchange rate could be introduced to support an overall tightening of monetary and fiscal policies, thereby limiting seigniorage. Transaction costs would also fall as inflation falls, which is a further incentive to welcome reform.

contract allowing free elections to be held.<sup>47,48</sup>

But there are also other possibilities for vested interests to improve their situation, as is made clear in the following propositions. Together with the size of  $k$  and  $\lambda$ , the depth of the democratic culture is a crucial determinant of the possibility of undertaking democratization since it determines the extent to which  $k$  can be lowered. Considering first the case when only limited reform can be accomplished, we can establish that:

**Proposition 7** *If interest groups do not cooperate and if  $\lambda$  is sufficiently small, an agenda-setter can find support for a limited reform in a democratic, but not in an authoritarian, direction if  $k$  is low enough. If  $k$  is sufficiently high, the opposite applies. Within a narrow interval, interest groups will support limited reform in either direction.*

**Proof.** A comparison of  $V_{IG}^{NC}(y_{NE}^*)$  in Table 3 for different values of  $k$  proves the proposition for the case of  $\lambda = 2$ . Limited reform can be undertaken in either direction in a neighborhood around  $k = 2.25$ .<sup>49</sup>

When radical political reform is possible, the available political reform space widens considerably in both directions as long as the size of the economy that is controlled by unorganized groups is small enough. This is the content of Proposition 8:

---

<sup>47</sup>It is not necessary that a contract is really signed. The issue is whether we can think of the development *as if* vested interests had the opportunity of signing a binding contract.

Hobbes's idea of a social contract saving societies from eternal conflict has been criticized by David Hume and others on two grounds. First, there is no evidence that there ever was a war of each against all in any society. Second, Hobbes's critics argue that no societies known to us have been founded on agreements of the kind Hobbes suggests.

In our setting, Hobbes's perspective might nevertheless be useful. First of all, interest groups were left in anarchy with potentially vast opportunities for rent seeking as the former ruler collapsed. Hence, the evolutionary process that terminated authoritarian rule could very well have resulted in an unstable situation that could be described as a war for rents. It is of course possible that similar episodes have suddenly arisen in early societies, but that they have left little or no historic or archaeological imprint simply because of their instability and therefore short duration. Second, difficulties in signing a binding contract between a relatively small group of people, that is, managers of the most influential industries, should be easier to overcome than those of signing a contract between many free men.

<sup>48</sup>Such a contract could, in principle, be delayed by a "war of attrition" between groups trying to escape from a part of their share of the burden of stabilization, as suggested by Alesina and Drazen (1991). In their case, hyperinflation results because the groups cannot agree, or sign a contract, on the share of the burden to be carried by each group.

<sup>49</sup>For the sake of simplicity, we refer to changes in  $k$  amounting to one step in Tables 2-5 as limited reform, and changes exceeding one step as radical reform.



**Proposition 8** *If interest groups do not cooperate and if  $\lambda$  is sufficiently small, an agenda-setter can find support for reform in a democratic direction for a higher initial value of  $k$  when radical political reform is possible than when he is constrained to limited reform. Likewise, a change in an authoritarian direction can be undertaken for a lower initial value of  $k$ . The range within which the agenda-setter can find support for radical reform in either direction contains the range within which he can find support for limited political reform in either direction.*

**Proof.** See Table 3 for the case of  $\lambda = 2$ .

Proposition 8 gives us an important argument as to why radical democratization might be preferable to limited, or gradual, democratization. This can be stated as:

**Corollary 1** *There are cases in which an agenda-setter can find support from interest groups for a far-reaching democratization but not for a less radical one.*

As an example, consider Table 3 at  $k = 3$ . A move to  $k = 1.5$ , but not to  $k = 2.25$ , would be supported. A consequence of this is that when  $k$  is too sticky to undertake radical reform, and when a small step towards democracy is not enough to make vested interests better off, democratization will not be achieved.

If  $\lambda$  and  $k$  are large enough, the expected payoffs of interest groups are increasing in  $k$  in the non-cooperative case also, as long as unorganized groups do not react fully to rent seeking. The implications for the possibilities of undertaking political reform are as follows:

**Proposition 9** *If interest groups do not cooperate and if  $\lambda$  and  $k$  are large enough, they will resist democratization and support any move in an authoritarian direction unless unorganized groups react fully to rent seeking.*

**Proof.** See Table 4 for the case of  $\lambda = 10$ .

Another way in which interest groups can improve their situation is to resurrect cooperation. This could either be accomplished by surrendering their power to an authoritarian ruler who could force them to cooperate, or by investing in cooperative structures. We define the outcome of the latter choice as corporative rule. Resurrecting cooperation is the first best option for vested interests, as is clear from a comparison between  $V_{IG}^C(y^*)$  and  $V_{IG}^{NC}(y_{NE}^*)$  for different levels of  $k$  and  $\lambda$ .<sup>50</sup> Let us now introduce a fixed

---

<sup>50</sup>See Tables 2 and 3 for the case of  $\lambda = 2$ . The result carries through also for the alternative types of behavior of unorganized groups.

cost,  $\Theta$ , for doing this. Whether interest groups will consider this option or not is then dependent on the size and flexibility of  $k$ , as shown in our final proposition.

**Proposition 10** *If interest groups do not cooperate and if it is costly to resurrect cooperation, it is more likely that authoritarian or corporative rule will be (re)established the greater  $k$  is and the less  $k$  can be lowered.*

**Proof.** Consider first the case when  $k$  cannot be lowered. If  $k$  is high, interest groups stand to gain more from resurrecting cooperation than if  $k$  is low. This is clear from a comparison between  $V_{IG}^C(y^*)$  and  $V_{IG}^{NC}(y_{NE}^*)$  for different levels of  $k$ . It is therefore more likely that  $V_{IG}^C(y^*) - \Theta > V_{IG}^{NC}(y_{NE}^*)$  when  $k$  is high, which is a necessary condition for cooperation to be resurrected. When  $k$  can be lowered to  $\underline{k}$ , the net increase in payoffs from resurrecting cooperation has to exceed the possible increase in payoffs as a result of democratization. This will be the case when  $V_{IG}^C(y^*) - \Theta > V_R^{NC}(y_{NE}^*(\underline{k}))$ .

Hence, vested interests will have less to lose on admitting a democratic Leviathan instead of overcoming their disagreements the further the democratic development has gone, and the deeper the democratic culture is. It might therefore not be worthwhile undertaking investments leading to authoritarian or corporative rule if  $k$  is too low or too flexible. Since it takes time to overcome disagreements, or for a potential authoritarian ruler to establish himself as an alternative, it is also clear that the decisiveness of the reformer matters. The faster he acts, the more likely he will be to succeed.

## 6 Applications

Our analysis has both positive and normative implications. Starting with the former category, we will now compare and interpret major aspects of the transitions in Poland and Russia in order to highlight qualitative implications of our model.

### 6.1 Poland and Russia Compared

Poland was the first country to attempt a radical transition from a command to a market economy as it launched an ambitious reform package in January 1990. It was also the first to reach economic growth. Leszek Balcerowicz (1994), the leading reformer, has described the preconditions as favorable. He argues that the second half of 1989 and early 1990 was characterized by "extraordinary politics," which he defines as a period of very clear discontinuity

in a country's history (*ibid.*, p. 176).<sup>51</sup> The former ruler is discredited, and the new political structures, including political parties and interest groups, have still not become established. The tendency of politicians and interest groups to act in favor of the common good is stronger than normal, as is the likelihood that the population at large will accept such measures. The reason given by Balcerowicz to explain why Poland succeeded in the early stages of reform is then, simply but crucially, that there were politicians in place who could utilize the beneficial conditions to implement a radical reform program.

The concept of extraordinary politics captures some of the factors that our analysis has identified as important. In particular, interest groups have to be unable to coordinate their actions. In Poland, the major interest groups included agrarians consisting of private small-scale farmers, and labor unions. The position of the workers was further strengthened by powerful workers' councils, which played a major role in most industries. The employers' organizations, on the other hand, were weak. It is no surprise that these groups were unable to cooperate with each other.

But the depth of the democratic culture and the share of resources that is controlled by interest groups are not included among the factors producing conditions for extraordinary politics. The democratic culture was indeed quite well developed. Most importantly, a political mass movement in the guise of a trade union, Solidarity, had grown to include almost half of Poland's adult population as its members (*ibid.*, p. 154). Also, mass media were well developed, and the major national newspapers came to play an important role in their support of reforms. The share of resources controlled by interest groups was also high because of the high degree of unionization and centralization, and the large size of production units. That these factors should be included in the preconditions for reform is therefore supported by the Polish reality.

Jeffrey Sachs's (1994) analysis of the preconditions for reform stands in sharp contrast to that of Balcerowicz, and even more so to ours. According to Sachs (*ibid.*, p. 505), "Poland was as likely as any country to succumb to rampant populism and wild, misdirected policies. That it found its way out of the crisis was heavily dependent upon the leadership of Leszek Balcerowicz and the critical involvement of the international community." Hence, Sachs argues that political preconditions do not matter, at least not when a severe economic crisis has erupted. Any country with an economic crisis similar to Poland's could undertake equally radical economic reforms.

---

<sup>51</sup>This could come about during a period of very deep economic crisis, because of a breakdown of the institutional system, or after liberation from external domination or the end of a war. According to Balcerowicz (*ibid.*, p. 177), these phenomena coexisted in Poland in 1989.

If we think of a severe economic crisis as one in which interest groups allocate a substantial share of their resources to appropriative activities, our analysis shows that interest groups might find it preferable to resist democratization under a wide range of conditions even in a major economic crisis. Such conditions will, for instance, come about in the non-cooperative case if the share of resources that interest groups control is low enough and the vulnerability of unorganized groups to appropriative activities is great enough.

Inspired by the Polish reforms, but cautious because of their weak political platform and the power of major industries, the Russian reform government under Yegor Gaidar launched an attempt at radical reform in January 1992. The Russian attempt was far less successful than the Polish one, and a number of concessions were made to vested interests which were even admitted into the government in the summer of 1992. The initial macroeconomic conditions were similar to those of Poland, but the political preconditions differed widely. The Russian agrarians were directors of state and collective farms, and teamed up with industrial leaders in their opposition against reform. There were no important independent labor unions, and the position of workers in individual enterprises was weak. Moreover, compared to the situation in Poland, there was little organized political activity outside of the communist party and affiliated organizations. Mass media had, however, started to develop fast in the late 1980s, but the coverage of independent media outside the major cities was small.

The outcome of the Russian reforms is well in line with the analysis in this paper. Interest groups resisted reforms and engaged in massive rent seeking, thereby increasing their share of a diminishing pie. Nevertheless, the reformers in the government eventually succeeded in putting the foundations of a market economy in place. The key explanation for this is, according to our analysis, that the vulnerability of unorganized groups to appropriative activities was lowered by a sequence of political reforms. These included, among other things, the introduction of free elections and the adoption of a new constitution by means of a referendum. The reformers also secured support for economic reforms through a referendum. They utilized the steps taken in a democratic direction to implement economic reforms which disunited interest groups. Their discord, in turn, made further economic reforms possible.<sup>52</sup>

## 6.2 Rapid or Gradual Democratization?

Turning to normative implications, our analysis supports the view that political and economic reforms should be as radical as possible during a transi-

---

<sup>52</sup>See Bornefalk (1995) for the case of the break-up of the ruble zone.

tion from an authoritarian to a democratic system. The reason is that this minimizes the costs in terms of a deteriorating economic performance and redistribution of wealth. However, our analysis does also point out the importance of initial conditions for the possibility of undertaking radical reforms. This enables us to evaluate statements that a particular reform should have been undertaken faster in a particular country. For instance, Dewatripont and Roland (1995, p. 1219) suggest that Boris Yeltsin might have made a major mistake when he did not take advantage of the abortive coup of August 1991 to push for new congressional elections and a new constitution prior to initiating further economic reform. Although our analysis shows that the possibility of successful economic reforms would indeed have been greater if the democratization process had gone further, it also suggests that such a strategy could have been risky. This would have been the case if interest groups had been able to cooperate at an early stage, or if the democratic culture had not already become deep enough. The reason for this is simply that interest groups would then have had strong incentives to resist democratization.

Instead of this risky strategy, Yeltsin followed the gradual strategy outlined above. The length of the political transition period led to a costly and sometimes chaotic economic transition. Russia did, however, become a market economy, and it is by no means certain that the costs would have been lower if the reformers had chosen to try to implement a more radical strategy.

In contrast to the countries our study focuses on, China has not commenced its political transition. The democratic culture is shallow, and much of the economy is still underdeveloped and under centralized control. According to our analysis, these factors could explain why China has been able to undergo a transition towards market economy without experiencing a fall in GDP. The only other country that has managed to do so is Vietnam, which has a similar record regarding political reform as China.

But political and economic changes have also taken place in China. As the gradual economic reforms launched in 1978 resulted in somewhat greater diversity of industries during the 1980s, and the repression was slightly diminished, macroeconomic imbalances appeared, which in turn sparked political protests. The Chinese leaders responded by tightening repression, and thereby succeeded in fending off rising inflation, just as our model would have predicted. A further prediction is that it will be more difficult to fight macroeconomic imbalances and corruption as the economy becomes more developed. Should future Chinese leaders try to undertake a gradual democratization, we would expect appropriative activities to increase further. What could save China from a chaotic transition period is either increasing

repression or a radical democratization. Unfortunately, the latter alternative seems to be quite unlikely to occur given the absence of any previous experience of democracy.

## 7 Concluding Remarks

This paper has identified determinants of the political and economic reform space and the economic performance following the demise of authoritarian rule. Our focus has been on societies where the economy has been characterized by strict centralized control and a rigid division into different sectors. There was frequently a lack of balance between the power of interest groups and unorganized groups. Interest groups were generally powerful because of the way the economy had been organized, whereas unorganized groups lacked influence because of poorly developed democratic institutions due to a long period of political oppression. This made unorganized groups rather vulnerable to appropriative activities in many transition countries. The vulnerability of unorganized groups contributed to creating conditions under which interest groups to a great extent chose to engage in appropriative rather than productive activities. This, in turn, is found to be a major factor behind the drastic output falls and the slow pace of political and economic reform that have characterized many transition countries.

The attempt at rapid democratization from above differs significantly from the gradual democratization process that modern western societies have, for the most part, undergone. In countries of the latter type, the strength of interest groups and the democratic culture have tended to develop in parallel. Appropriative effectiveness has therefore been fairly low throughout the democratization process. These countries have therefore not been struck by costs of the type and magnitude faced by the societies we study.

The costs brought about by the unbalanced nature of political conditions during the democratic transition in formerly authoritarian societies have a parallel in the rapid increase in population during the second half of the 20th century in many third world countries. In the latter case, advances in medicine and hygiene often came prior to a general increase in wealth levels rather than being introduced in parallel, as they were in the western world. As a consequence, the fall in mortality rates relative to the fall in birth rates has been greater in third world countries during the demographic transition, which has led to a rapid and sometimes problematic expansion of the population.<sup>53</sup>

---

<sup>53</sup>There is a link between the unbalanced nature of the demographic transition and the political conditions that have characterized, and still characterize, many third world

There are also similarities between the pattern we have recognized for transition countries and the development in more stable societies. For instance, the pattern we have identified between economic performance and the depth of the democratic culture following the collapse of authoritarian rule is similar to that found in research which examines interconnections between economic growth and political systems. One such finding is that economic growth is likely to be highest on average in mature democracies, although it varies greatly among different systems and over time.<sup>54</sup> McGuire and Olson (1996) argue that high growth rates are also possible in an authoritarian society, since the dictator has an interest in economic growth as long as he controls a dominant share of the economy and can use much of the production to attain his own goals. Harsh repression will in general be required to secure control, however, which might impede economic development.

These findings correspond well with our results if we are willing to draw parallels between different political systems and different conditions following the collapse of authoritarian rule. For instance, our result for the difference between a transition country with a considerable depth of the democratic culture and one with rudimentary democratic institutions, and McGuire and Olson's result for the difference between a mature democracy and an authoritarian society where the ruler has firm control over the economy coincide, since economic performance in the cooperative case is relatively good for a high  $k$ , but not as good as for a low  $k$ . As the economy of McGuire and Olson becomes more complex, it will be increasingly difficult for the dictator to maintain control, especially over emerging sectors. His interest in economic growth in these will therefore be small. Hence, the growth differential vis-à-vis a democracy will increase with greater complexity of the economy. This result is also consistent with our model as long as there is a connection between the complexity of the economy and the ability of interest groups to cooperate, or of the ruler to force them to do so.

Our results can easily be reinterpreted to shed light on a number of related issues. Questions that could be analyzed include under what conditions authoritarian rule will terminate as a result of internal pressure, and why monetary unions comprising different language areas have turned out to be unstable. The clue to the latter question is that the democratic culture is

---

countries. People living in authoritarian societies often lack an opportunity to invest in themselves and their children, and to save for the future. Moreover, they typically cannot rely on the state for support during old age. As a consequence, they choose to raise many children. This suggests that it is advantageous to undergo a democratic transition prior to a demographic transition.

<sup>54</sup>Grossman and Noh (1990) discuss factors that determine the rate of growth in a democracy.

similar within a language area, but not necessarily between different areas. Optimal levels of rent seeking therefore tend to differ between language areas.<sup>55</sup> Our results could also be used to discuss under what conditions there will be a middle class, and when starvation threatens.

The model could be extended to determine the degree of organization in a society and the degree and method of cooperation between interest groups. A more fundamental step in reaching a better understanding of the preconditions for reform could be taken by endogenizing the depth of the democratic culture. Our results also suggest where empirical research on preconditions for reform should focus. Clearly, we need to learn more about the foundations of democratic institutions and the behavior of interest groups in various environments, in order to understand when particular reform programs are likely to become successful. To allow for more radical reforms, we need to learn more about how the democratic culture could be deepened, and how interest groups blocking reforms could be divided.

## Appendix

### PROOF OF PROPOSITION 1:

Differentiating  $V_{IG}^C(y)$  in equation (5) with respect to  $y$  for  $\lambda = 2$  yields

$$\begin{aligned} \frac{\partial V_{IG}^C(y)}{\partial y} &= \frac{ke^{ky}}{2e^{ky} + 2} \left( 4 - 2y - 2\frac{y}{k} \right) - \\ &\quad - \frac{2ke^{2ky}}{(2e^{ky} + 2)^2} \left( 4 - 2y - 2\frac{y}{k} \right) - \frac{e^{ky}}{2e^{ky} + 2} \left( 2 + \frac{2}{k} \right) \\ &= 0 \end{aligned} \tag{8}$$

at the optimum. Solving for the optimal  $y$  as an expression of  $k$  gives

$$y^*(k) = \frac{2k^2 - (k+1)LambertW\left(e^{\frac{2k^2-k-1}{k+1}}\right) - k - 1}{k(k+1)}, \tag{A2}$$

which is the expression underlying the curve in Figure 1.<sup>56</sup>

### PROOF OF PROPOSITION 2:

<sup>55</sup>See Bornefalk (2000c) for an analysis of monetary reform in the former Soviet Union along these lines.

<sup>56</sup>LambertW is an implicitly defined function satisfying  $LambertW(x)e^{LambertW(x)} = x$ . The function is named after the 18th century mathematician Johann Heinrich Lambert, who was a colleague of Euler and Lagrange at the Berlin Academy of Sciences. For more on the LambertW function, see Briggs (1998) and Corless et al. (1996).



We first derive combinations of  $k$  and  $\lambda$  that result in a situation where interest groups using all their resources for production cannot increase their expected payoffs by allocating a marginal share to rent seeking. Differentiating  $V_{IG}^C(y)$  in equation (5) with respect to  $y$ , substituting  $y = 0$  into the resulting expression, and equating this to zero, we get

$$\frac{\partial V_{IG}^C(0)}{\partial y} = \frac{\lambda k^2 - 2k - \lambda}{(2 + \lambda)k} = 0. \quad (\text{A3})$$

This expression immediately translates into

$$k^2 - \frac{2k}{\lambda} - 1 = 0, \quad (\text{A4})$$

which is the relationship we sought.

We then show that no matter how large  $\lambda$  is, there will always be a level of  $k$  below which rent seeking is absent. Solving for  $k$  yields the positive root

$$k_{cutoff}^C(\lambda) = \frac{1 + \sqrt{1 + \lambda^2}}{\lambda}, \quad (\text{A5})$$

which is the cutoff level of  $k$  in the cooperative case. We see that  $k_{cutoff}^C > 1$  for all  $\lambda > 0$ . Substituting  $\lambda = 2$  yields  $k_{cutoff}^C = \frac{1+\sqrt{5}}{2}$ , which is the cutoff value in Table 2.

To see that the cutoff level is decreasing in  $\lambda$  in the relevant range, we take the derivative of  $k_{cutoff}^C$  with respect to  $\lambda$ :

$$\frac{dk_{cutoff}^C(\lambda)}{d\lambda} = \frac{1}{\sqrt{1 + \lambda^2}} - \frac{1 + \sqrt{1 + \lambda^2}}{\lambda^2}. \quad (\text{A6})$$

This expression is negative for  $\lambda > 0$ . Hence, the greater  $\lambda$  is, the smaller  $k_{cutoff}^C$  is. As  $\lambda$  approaches infinity,  $k_{cutoff}^C$  approaches 1 asymptotically:

$$\lim_{\lambda \rightarrow \infty} k_{cutoff}^C = 1. \quad (\text{A7})$$

Likewise, as  $\lambda$  comes asymptotically close to 0,  $k_{cutoff}^C$  approaches infinity:

$$\lim_{\lambda \rightarrow 0} k_{cutoff}^C = \infty. \quad (\text{A8})$$

Proposition 2 holds for alternative specifications of the reaction to rent seeking. If the reaction is independent of  $k$  but otherwise as in equation (3), we get  $k_{cutoff}^C(\lambda) = \frac{2+\lambda}{\lambda}$ . If there is no reaction to rent seeking,  $k_{cutoff}^C(\lambda) = \frac{2}{\lambda}$ . Proposition 2 then holds if we allow  $k > 0$ .

### PROOF OF PROPOSITION 3:

To derive the Nash equilibria, we first differentiate equation (7) with respect to  $y_1$  for the case of  $\lambda = 2$ .

$$\begin{aligned} \frac{\partial V_1^{NC}}{\partial y_1} &= \frac{e^{ky_1}}{e^{ky_1} + e^{ky_2} + 2} \left[ 4 - y_1 - y_2 - \frac{y_1 + y_2}{k} \right] k - \\ &\quad - \frac{e^{2ky_1}}{(e^{ky_1} + e^{ky_2} + 2)^2} \left[ 4 - y_1 - y_2 - \frac{y_1 + y_2}{k} \right] k - \\ &\quad - \frac{e^{ky_1}}{e^{ky_1} + e^{ky_2} + 2} \left( 1 + \frac{1}{k} \right). \end{aligned} \quad (9)$$

Equating this derivative to zero, we then solve for  $y_1$  as a function of  $y_2$  and  $k$ . We get

$$\begin{aligned} y_1^*(k, y_2) &= \frac{4k^2 - k^2y_2 - ky_2 - k - 1}{k(k+1)} - \\ &\quad - \frac{1}{k} \text{LambertW} \left( \frac{e^{\frac{4k^2 - k^2y_2 - ky_2 - k - 1}{k+1}}}{e^{ky_2} + 2} \right). \end{aligned} \quad (10)$$

This is the expression underlying the reaction function in Figure 2. We then find values of  $y_2$  for which  $y_1^*(k, y_2) = y_2$ . Plotting these values for different values of  $k$  yields the curve in Figure 3.

### PROOF OF PROPOSITION 5:

Differentiating  $V_1^{NC}(y_1, y_2)$  in equation (7) with respect to  $y_1$ , substituting  $y_1 = y_2 = 0$  into the resulting expression, and equating this to zero, we arrive at the following expression for combinations of  $k$  and  $\lambda$  that result in a situation where non-cooperating interest groups using all their resources for production cannot increase their expected payoffs by allocating a marginal share to rent seeking:

$$k^2 - \frac{k}{1 + \lambda} - \frac{\lambda}{2(1 + \lambda)} = 0. \quad (A11)$$

Solving for  $k$  yields the positive root:

$$k_{cutoff}^{NC}(\lambda) = \frac{1 + \sqrt{1 + 2\lambda + 2\lambda^2}}{2 + 2\lambda}. \quad (A12)$$

At  $\lambda = 0$ ,  $k_{cutoff}^{NC}(\lambda) = 1$ . This is a maximum of  $k_{cutoff}^{NC}(\lambda)$  for  $\lambda \geq 0$  since  $k_{cutoff}^{NC}(\lambda)$  is declining in  $\lambda$  in this range. It is therefore the only value of  $k_{cutoff}^{NC}(\lambda)$  for  $\lambda \geq 0$  for which the restriction  $k \geq 1$  is satisfied. Hence, rent

seeking is never absent in the non-cooperative case when unorganized groups control a positive share of initial resources. This also holds for alternative specifications of the reaction to rent seeking. When the reaction is independent of  $k$ ,  $k_{cutoff}^{NC}(\lambda) = \frac{2+\lambda}{2(1+\lambda)}$ . When there is no reaction to rent seeking,  $k_{cutoff}^{NC}(\lambda) = \frac{1}{1+\lambda}$ .

#### PROOF OF PROPOSITION 6:

It can be shown numerically that  $\frac{\partial V_{IG}^C(y^*)}{\partial k} > 0$  when  $k > k_{cutoff}(\lambda)$ . This is done by first substituting  $y^*(k)$  (see equation (A2) for the case of  $\lambda = 2$ ) for  $y$  in equation (5) and then differentiating the resulting expression with respect to  $k$ . Plotting  $\frac{\partial V_{IG}^C(y^*)}{\partial k}$  for different combinations of  $k > k_{cutoff}(\lambda)$  and  $\lambda$  reveals that the derivative is positive in the relevant range.

## References

- [1] Alesina, Alberto, and Drazen, Allan. "Why Are Stabilizations Delayed?" *American Economic Review*, December 1991, 81, 1170-88.
- [2] Åslund, Anders. *How Russia Became a Market Economy*. The Brookings Institution, Washington, D.C., 1995.
- [3] Åslund, Anders, Boone, Peter, and Johnson, Simon. "How to Stabilize: Lessons from Post-communist Countries." *Brookings Papers on Economic Activity*, 1996, 217-313.
- [4] Balcerowicz, Leszek. "Poland." in Williamson, ed., 1994.
- [5] Bornefalk, Anders. "Social Conflict with Passive Groups." in *Essays on Social Conflict and Reform*, Ph.D. Dissertation, Stockholm School of Economics and SITE, 2000a.
- [6] Bornefalk, Anders. "Constitutional Constraints and Redistributive Activities." in *Essays on Social Conflict and Reform*, Ph.D. Dissertation, Stockholm School of Economics and SITE, 2000b.
- [7] Bornefalk, Anders. "The Break-up of the Ruble Zone: Undertaking Monetary Reform while Building Democratic Institutions." in *Essays on Social Conflict and Reform*, Ph.D. Dissertation, Stockholm School of Economics and SITE, 2000c.
- [8] Bornefalk, Anders. "The Ruble Zone: A Case of Irrationality?" *Ukrainian Economic Review*, 1995, 1, 96-118.

- [9] Brennan, Geoffrey and Kliemt, Hartmut. "Finite Lives and Social Institutions." *Kyklos*. 1994, 47, 551-571.
- [10] Briggs, Keith. "W-ology or Some Exactly Solvable Growth Models." available online at <http://epidem13.plantsci.cam.ac.uk/~kbriggs/W-ology.html>, 1998 [1999-]
- [11] Buchanan, James M., Tollison, Robert D., and Tullock, Gordon, eds. *Toward a Theory of the Rent Seeking Society*. Texas A&M Press, College Station, TX, 1980.
- [12] Corless, Robert M. et al. "On the Lambert W Function." *Advances in Computational Mathematics*, 1996, 5, 329-59.
- [13] Dewatripont, Mathias and Roland, Gérard. "Transition as a Process of Large-scale Institutional Change." in Kreps and Wallis, eds., 1997.
- [14] Dewatripont, Mathias and Roland, Gérard. "The Design of Reform Packages under Uncertainty." *American Economic Review*, December 1995, 85, 1207-23.
- [15] Drèze, Jean and Sen, Amartya K. *The Political Economy of Hunger. Volume I: Entitlement and Well-Being*. Clarendon Press, Oxford, 1990.
- [16] Ellingsen, Tore. "Strategic Buyers and the Social Cost of Monopoly." *American Economic Review*, June 1991, 81, 648-57.
- [17] Grossman, Herschel I. and Noh, Suk Jae. "A Theory of Cleptocracy with Probabilistic Survival and Reputation." *Economics and Politics*, July 1990, 2, 157-171.
- [18] Haggard, Stephan and Webb, Steven B. *Voting for Reform. Democracy, Political Liberalization, and Economic Adjustment*. Oxford University Press, New York, N.Y., 1994.
- [19] Hirshleifer, Jack. "Conflict and Rent-seeking Success Functions: Ratio vs. Difference Models of Relative Success." *Public Choice*, 1989, 63, 101-112.
- [20] Hirshleifer, Jack. "Anarchy and its Breakdown." *Journal of Political Economy*, 1995, 103, 26-52.
- [21] Hobbes, Thomas. *Leviathan*. Viking Press, 1982/1651.

- [22] Klitgaard, Robert. *Tropical Gangsters*. I. B. Tauris & Co Ltd, London, 1991.
- [23] Kornai, Janos. *The Socialist System. The Political Economy of Communism*. Clarendon Press, Oxford, 1992.
- [24] Kreps, D. and Wallis, K., eds. *Advances in Economic Theory*. Cambridge University Press, Cambridge, 1997, II, 240-278.
- [25] Lindert, Peter H. "What Limits Social Spending?" *Explorations in Economic History*, 1996, 33, 1-34.
- [26] Mau, Vladimir. "Russia." in Williamson, ed., 1994.
- [27] McGuire, Martin and Olson, Mancur. "The Economics of Autocracy and Majority Rule: The Invisible Hand and the Use of Force." *Journal of Economic Literature*, March 1996, XXXIV, 72-96.
- [28] Meltzer, Allan H. and Richard, Scott F. "A Rational Theory of the Size of Government." *Journal of Political Economy*, 1981, 89, 914-27.
- [29] Olson, Mancur. "Why Is Economic Performance Even Worse When Communism Is Abolished?" Paper presented at the Institute for the Economy in Transition, Moscow, March 1994.
- [30] Olson, Mancur. *The Logic of Collective Action*. Harvard University Press, Cambridge, MA, 1965.
- [31] Przeworski, Adam. *Democracy and the Market. Political and Economic Reforms in Eastern Europe and Latin America*. Cambridge University Press, Cambridge, 1991.
- [32] Sachs, Jeffrey D. "Life in the Economic Emergency Room." in Williamson, ed., 1994.
- [33] Sen, Amartya K. "Development: Which Way Now?" *The Economic Journal*, December 1983, 93, 745-62.
- [34] Skaperdas, Stergios. "Cooperation, Conflict, and Power in the Absence of Property Rights." *American Economic Review*, September 1992, 82, 720-39.
- [35] Smith, Adam. *The Wealth of Nations*. Edited by Edwin Cannan. Modern Library, New York, 1937/1776.

- [36] Tullock, Gordon. "Efficient Rent Seeking." in Buchanan et al., eds., 1980.
- [37] Williamson, John, ed. *The Political Economy of Policy Reform*. Institute for International Economics, Washington, D.C., 1994.
- [38] Wärneryd, Karl. "Partisanship as Information." *Public Choice*, 1994, 80, 371-380.

## Essay 3

### Constitutional Constraints and Redistributive Activities





# Constitutional Constraints and Redistributive Activities

Anders Bornefalk\*

February 8, 2000

## Abstract

This paper develops a theoretical argument suggesting that non-productive activities aimed at influencing the distribution of income might increase when constitutional restrictions against redistribution are imposed or strengthened. This could come about when organized groups are able to secure such a large share of total production prior to the constitutional reform that they choose to limit the share of their resources that they allocate to influence activities, in order not to hurt their own interests. The reform would reduce the share received by organized groups at the initial level of influence activities, which in turn would lower their incentive to limit such activities. By facilitating redistribution through constitutional reform, on the other hand, countries where the initial redistributive effectiveness is relatively large could limit influence activities, thereby improving economic performance. These findings are used together with other factors, such as the share of the economy's initial resources that is controlled by organized groups and the priorities of the constitutional assembly, to analyze constitutional choice under different conditions. We find that the mechanism analyzed in this paper is an impediment to the introduction of constitutional restrictions against redistribution, particularly in countries with poorly developed democratic institutions.  
(JEL: D72, P50)

---

\*Department of Economics and SITE, Stockholm School of Economics, Box 6501, SE-113 83, Stockholm, Sweden. Anders.Bornefalk@hhs.se. I would like to thank Lars Jonung, Niclas Berggren, Tore Ellingsen, Torsten Persson, Gordon Tullock, Jörgen Weibull, and participants at seminars at the Stockholm School of Economics and the 1999 PCS Annual Meeting for comments on earlier versions of this paper. The responsibility for any errors is mine.

# 1 Introduction

Knut Wicksell's (1896) idea of a unanimity rule in decisions on government expenditures and their financing addresses a perceived need to protect minorities from exploitation, and to avoid a struggle between different groups over the distribution of benefits and costs of government activities. This idea has inspired a number of scholars to propose constitutional constraints against government activities with redistributive effects. A prominent candidate is the requirement of some qualified majority, say two-thirds of the votes in parliament, for decisions causing redistribution to be enacted.<sup>1</sup> In this literature, it is typically taken for granted that interest groups will engage less in unproductive activities affecting the distribution of benefits and costs of government activities when constitutional constraints make such changes more difficult to accomplish.<sup>2</sup> But do we always try less when we meet more resistance? If the answer is no, this simple argument does not allow us to conclude that constitutional limitations on redistribution unambiguously reduce such activities.

This paper seeks to determine whether the introduction or sharpening of binding constitutional constraints against redistribution always reduces unproductive influence activities. The extent to which interest groups can affect the distribution of income prior to the constitutional reform turns out to be a key factor for the effects of constitutional reform of this kind. This is determined by the strength of interest groups and the implicit restrictions that the government faces when engaging in redistribution. The implicit restrictions against redistribution are determined, as argued in Bornefalk (2000b), by aspects of the political system such as the extent of political freedoms and the depth of the democratic culture. The maturity of political parties, the impact of mass media, and the degree of literacy are important parts of the latter factor.

The development of democratic institutions that had been taking place during the 19th century was, indeed, emphasized by Wicksell in conjunction with his proposal of a unanimity rule:

The movement which has nearly everywhere shaped the political history of this century has been steady progress toward parliamentary and democratic forms of public life. One of the prime movers is modern general education in the widest sense of the term, since this enables even the lowest classes of the people to participate

---

<sup>1</sup>Berggren (1996) presents a selection of constitutions with varying degrees of restrictions against redistribution. See also Mueller (1996) and Mueller (1989) for examples of such proposals.

<sup>2</sup>Mueller (1989, p. 245-46) is one example.

more and more in political life; another is the tremendous development of the press, whereby everyone has gained unprecedented access to knowledge of public affairs. (Wicksell, 1988, p. 118)

Hence, Wicksell proposed his new principle for just taxation after a deepening of the democratic culture had taken place. The depth of the democratic culture is a neglected dimension in later research in the field, but this paper presents a theoretical argument suggesting that the balance between the strength of interest groups and the depth of the democratic culture is important for the effects of constitutional constraints against redistribution.<sup>3</sup>

We focus on the effects of constitutional reform on the share of their resources that interest groups devote to production as opposed to influence activities. We also consider the extent to which redistribution takes place.<sup>4</sup> Influence activities and redistribution are both parts of what we refer to as redistributive activities. The analysis is undertaken in two stages. In the first stage, the explicit, that is, constitutional, restrictions against redistribution are used as one of the determinants to derive the share of their resources that interest groups devote to influence activities. In the second stage, the explicit restrictions against redistribution are a decision variable.

A major result is that interest groups might very well increase the share of their resources that they devote to influence activities rather than production, if constitutional constraints against redistribution are introduced or sharpened. This could be the outcome when interest groups cooperate with each other in extracting rents from unorganized groups, but not if they are

---

<sup>3</sup>This neglect could be due to the fact that many researchers have learned about Wicksell's unanimity rule through the work of Buchanan and Tullock (1962). In *The Calculus of Consent*, where they develop Wicksell's insight that the strict unanimity rule would make collective decision making ineffective, and suggest a number of decision rules that only approach unanimity, they assume for the most part that decisions are made by direct popular vote. They consequently say little about political parties and elected decision makers. When they do discuss elected legislatures, they assume that the representative votes according to the majority preference in his district, thereby neglecting special interest groups. Hence, they treat the behavior of politicians as a constant and investigate the role of constitutions in maximizing the extent to which the voter achieves his goals given this behavior. See Tullock (1962, p. 338).

<sup>4</sup>Wicksell (1896) considers the degree of fairness in taxation and the efficient level of government activities. He shows that the unanimity rule, together with the requirements that public expenditures and their financing be decided on simultaneously and in an open collective process, would secure fairness and an efficient level of government activities. No one would have to pay for anything that is not in his interest, and all activities for which there is a sufficient demand to cover the costs would be undertaken. Wicksell also argued that corrections of what was perceived as unfair distributions of income and wealth should be made subject to a requirement of some qualified majority, if only to secure political stability. See Söderström (1986) for more on this.

unable to do so. If their ability to affect the distribution of income is relatively extensive prior to the constitutional reform, sharpened constitutional constraints will increase influence activities and therefore lower economic performance. If their ability is relatively weak, on the other hand, sharpened constitutional constraints will increase economic performance.

The former case corresponds to societies with relatively powerful interest groups. This case applies, for instance, to countries where democracy has been introduced after a collapse of authoritarian rule, particularly if the economy was centrally organized and controlled. Under such conditions, the democratic culture is typically relatively shallow, whereas interest groups have been able to grow strong.<sup>5</sup> The latter case is, arguably, typical of countries where democratic institutions have developed in parallel with interest groups. The analysis also shows that the smaller the share of the economy's initial resources that is controlled by interest groups, the weaker the initial ability of interest groups to affect the distribution of income needs to be for strengthened constitutional constraints to actually have the effect of limiting influence activities.

A general finding, then, is that similar constitutional features could improve economic performance under certain political and economic conditions and worsen it in others. This is, perhaps, a less intuitive insight than its frequently mentioned counterpart, that is, that constitutions affect the development of the political and economic system. Moreover, the possibility that unproductive influence activities might increase when explicit restrictions against redistribution are strengthened is added to increased decision time costs as a potential drawback of constitutional reform in this direction.

The results concerning the effects of constitutional reform are used to analyze constitutional choice by a hypothetical constitutional assembly. The model is appropriate for the study of constitutional choice, since it assumes little about the political structure. The possibility that the assembly could be dominated by factions that have other aims than an improved economic performance and a distribution of income corresponding to the contribution to production is one obstacle to the introduction of constitutional constraints against redistribution. The finding that constitutional reform of this kind might lower economic performance is a further obstacle. Hence, not even an assembly with the sole ambition of improving economic performance will always prefer to strengthen the constitutional protection of property rights. At the same time, assemblies dominated by groups benefiting from redistribution or influence activities could indeed find it optimal to introduce

---

<sup>5</sup>See Bornefalk (2000b) for an analysis of problems typical for a democratic transition of this kind.

constitutional limitations against redistribution under certain conditions.

The model is presented in the next section. In Section 3, optimal levels of influence activities with and without constitutional constraints against redistribution are derived for different preconditions. Outcomes for different types of constitutional assemblies are then evaluated in Section 4. The outcomes are used to draw conclusions regarding the choice of constitution under different political and economic conditions. Section 5 concludes.

## 2 The Model

### 2.1 Production and Distribution

Our model extends the simple economy considered in Bornefalk (2000a,b) with a constitutional dimension.<sup>6</sup> A part of the economy's initial resources is controlled by organized interests which form two groups, labeled 1 and 2. The remainder is controlled by groups which have not solved their collective action problems.<sup>7</sup> Producers belonging to the latter category are, however, more numerous, which gives them a potentially great political influence. The interest groups possess one unit each of an initial resource. They can use their resources for production or to affect the distribution of income through influence activities aimed at what we refer to, for convenience, as the government. The only function of the government is to determine the distribution of goods or, equivalently, income.<sup>8</sup>

With  $y_i$  representing the share used for influence activities by group  $i = 1, 2$ ,  $x_i = 1 - y_i$  represents the share used for production. Unorganized groups, labeled 3, do not engage in influence activities. They use all their initial resources, given by  $\lambda$ , for production. In other words, their work

---

<sup>6</sup>We choose the model of Bornefalk (2000a,b) as a basis for our analysis since it contains few assumptions concerning the political structure. This makes it suitable for studying constitutional choice as a function of underlying factors, where the choice of constitution, in turn, shapes the political structure. This approach could then serve as a complement to a related field of research where detailed structural assumptions are made about, for instance, the allocation of proposal and veto powers, but where issues of positive constitutional choice are not addressed. However, the advantage of the structural approach followed by, for instance, Persson, Roland, and Tabellini (1997a,b), is that it enables a more careful study of how constitutions shape policy choices.

<sup>7</sup>See Olson (1965).

<sup>8</sup>Note that we only consider direct redistribution of income since there is no government sector in the model. Hence, we abstract from redistributive effects that result when government activities are financed according to the ability to pay rather than the benefit principle.

effort is assumed to be perfectly inelastic to the extent of redistribution.<sup>9</sup> Production, or economic performance, is then simply given by

$$C(x_1, x_2) = x_1 + x_2 + \lambda. \quad (1)$$

The government bases its decision about the distribution of income on the share of initial resources that different groups control and their expenditures on influence activities. The share of production received by each interest group in the non-cooperative case, that is, when each group seeks rents on its own, is given by

$$p_i^{NC}(y_1, y_2) = \frac{e^{\alpha k y_i}}{e^{\alpha k y_1} + e^{\alpha k y_2} + \lambda e^{\alpha k y_3}} = \frac{e^{\alpha k y_i}}{e^{\alpha k y_1} + e^{\alpha k y_2} + \lambda}. \quad (2)$$

for  $i = 1, 2$ , where we use  $y_3 = 0$ .<sup>10</sup> The share received by unorganized producers is given by replacing the numerator in equation (2) with  $\lambda$ .

The degree to which the government is able to respond to influence activities is determined by the strength of interest groups and the implicit and explicit restrictions against redistribution that the government faces. The joint effect of the first two factors is represented by  $k$  in equation (2).  $\alpha$  represents the explicit restrictions. We refer to the product  $\alpha k$  as the redistributive effectiveness in the economy. The redistributive effectiveness is assumed to be the same for both interest groups.

In a democracy, implicit restrictions against redistribution correspond to the need to be reelected so as to enjoy benefits channeled through the political process. They could measure the extent to which a given amount of campaign contributions to (presumptive) legislators affects their voting on policies of interest to the contributing group. Determinants of  $k$  could then, together with the strength of interest groups, be the extent of political freedoms and the depth of the democratic culture, as defined in the introduction. Explicit restrictions could take the form of constitutional provisions impairing or facilitating redistribution. They could, for instance, reflect the share of votes in parliament that is needed to adopt a law affecting distribution.

---

<sup>9</sup>Bornefalk (2000a,b) considers different cases of elastic contribution of unorganized groups to production. The effect of the reaction to appropriative activities is essentially equivalent to that of the size of  $\lambda$  for our purposes. The explanation for this is that  $\lambda$  reflects the share of the economy that is affected by appropriative activities.

<sup>10</sup>The conflict technology employed here belongs to the logistic family with the general form  $p_i = \frac{e^{k y_i}}{\sum_j e^{k y_j}}$ , where  $e$  is the base of the natural logarithm. As explained by Hirshleifer (1989, p. 104), this conflict technology applies to combat markets with imperfections. If there are imperfections, zero resource commitments need not imply that the player loses everything.

Hence, if  $k$  represents the ease with which an individual legislator can be bought by interest groups,  $\alpha$  reflects the number of legislators that must be bought to affect a decision.

We impose the restrictions  $k \geq 1$  and  $\alpha \geq 0$ . In terms of voting rules,  $\alpha = 1$  corresponds to the simple majority rule and  $\alpha < 1$  to a requirement of a qualified majority of some percentage. An example of constitutional reform resulting in  $\alpha > 1$  would be if the government is given proposal power on issues of redistribution that the parliament can only veto if, for instance, two thirds of the votes are against the proposal.

The expected payoff of interest group 1 in the non-cooperative case is given by equations (1) and (2) with  $x_i = 1 - y_i$ :

$$\begin{aligned} V_1^{NC}(y_1, y_2) &= p_1^{NC}(y_1, y_2)C(y_1, y_2) \\ &= \frac{e^{\alpha k y_1}}{e^{\alpha k y_1} + e^{\alpha k y_2} + \lambda} (2 - y_1 - y_2 + \lambda). \end{aligned} \quad (3)$$

Expected payoffs of interest group 2 and unorganized producers,  $V_2^{NC}(y_1, y_2)$  and  $V_3^{NC}(y_1, y_2)$ , are given by replacing  $e^{\alpha k y_1}$  in the numerator with  $e^{\alpha k y_2}$  and  $\lambda$ , respectively.

When interest groups cooperate with each other in seeking rents, they control two units of the initial resource and allocate a share  $y$  to influence activities.<sup>11</sup> The share of each interest group is then given by

$$p_i^C(y) = \frac{1}{2} \frac{2e^{\alpha k y}}{2e^{\alpha k y} + \lambda} = \frac{e^{\alpha k y}}{2e^{\alpha k y} + \lambda} \quad (4)$$

for  $i = 1, 2$ . The expected payoffs of interest groups are given by equations (1) and (4), with  $x_i = 1 - y$  for  $i = 1, 2$ . Each interest group receives

$$V_i^C(y) = p_i^C(y)C(y) = \frac{e^{\alpha k y}}{2e^{\alpha k y} + \lambda} (2 - 2y + \lambda). \quad (5)$$

The expected payoffs of unorganized groups in the cooperative case,  $V_3^C(y)$ , are given by replacing  $e^{\alpha k y}$  in the numerator in equation (5) with  $\lambda$ .

## 2.2 Objectives of the Constitutional Assembly

Having defined the behavior of the government and different groups in society, we now turn to the constitutional assembly. If every person were to

---

<sup>11</sup>The method of cooperation employed here is referred to as additive functional cooperation in Bornefalk (2000a). Our results are robust to the method of cooperation. See Bornefalk (2000b) for a discussion of the conditions under which interest groups will be able to cooperate.

participate with equal rights and had access to full information on the production and conflict technologies, and if votes could not be bought, the assembly would choose to set  $\alpha = 0$  to prevent redistribution from unorganized groups, since these form a majority. In addition to a fair distribution of income—in the sense that payoffs of different groups correspond to what these groups contribute to production—this would result in economic performance being at its maximum level since there would be no influence activities. If, however, the constitutional assembly consists of a subset of the population, it could very well choose  $\alpha > 0$ .<sup>12</sup> This would certainly be the case if organized groups were to dominate the assembly. But it could also happen if the assembly were to be drawn from the population at large in a fair way. In this case, it suffices that being a member of government opens up a possibility to benefit from influence activities, and that the representatives expect to become members of the government with a higher probability than those they represent.

We refer to the constitutional assembly that is dominated by a certain group as biased. It cares about the distribution and receives no direct benefits from resources spent on influence activities. It maximizes a weighted average of the economic performance and the expected payoffs of the group it is biased towards. In the non-cooperative case, this index is given by

$$A_{B_i}^{NC}(\alpha) = \eta C(y_1(\alpha), y_2(\alpha)) + (1 - \eta) V_i^{NC}(y_1(\alpha), y_2(\alpha)) + E \quad (6)$$

for  $i = 1, 2, 3$ , where the weight is given by  $\eta \in [0, 1]$ .

The last term in the objective function,  $E$ , represents exogenous rents from holding office, such as esteem and any material advantages tied to the office. The members of the constitutional assembly that form the government receive these rents if the government behaves according to the conflict technology represented by equations (2) and (4). If it transgresses, that is, behaves in any other way, it is replaced by an identical government drawn from the population rather than the constitutional assembly.  $E$  is assumed to be high enough to make the government prefer to remain in the system rather than seek to exploit its position by transgressing, but is normalized to 0. Hence, the government will follow the same policy on distribution regardless of which group it represents. The only way the constitutional assembly can benefit its own group is therefore through the choice of a constitution, as represented by assigning a value to  $\alpha$ .

---

<sup>12</sup>Constitutions or constitutional reforms are often subject to approval in a referendum. Still, it is the constitutional assembly that formulates the proposal and indeed determines whether or not there will be a proposal of reform.



We refer to the second type of constitutional assembly as self-seeking; it values resources spent by organized groups on affecting the distribution of income and does not identify with any of the groups. The motivation for this is that in this case there is insufficient party or organizational discipline to make the representatives of the constitutional assembly that form the government favor anything but themselves. The self-seeking constitutional assembly has the objective function

$$A_S^{NC}(\alpha) = \eta C(y_1(\alpha), y_2(\alpha)) + \gamma(1 - \eta)(y_1(\alpha) + y_2(\alpha)) + E \quad (7)$$

in the non-cooperative case. The economic performance is again given a weight  $\eta \in [0, 1]$ .<sup>13</sup>  $\gamma \in [0, 1]$  represents the share of the expenditures on influence activities that the government is able to transform to goods benefiting only itself. The more transparent the political system is, the more difficult diversion becomes. This is reflected in a smaller  $\gamma$ . When  $\gamma$  is small, the government might have to accept goods or services that cost more to supply than they are worth to the government, instead of accepting pecuniary bribes. An equivalent interpretation is that a larger part of the bribes might have to be sacrificed to conceal the corruptive activities. The government gives the weight  $1 - \eta$  to the endogenous rents, that is, to expenditures on influence activities net of transaction costs in corruption.<sup>14</sup> We again assume that the exogenous rents are high enough to make the government behave according to the model.<sup>15</sup>

### 3 Constitutions and Influence Activities

We now proceed to derive the extent to which influence activities are carried out under different conditions. The aim is to establish when constitutional reform that increases the explicit restrictions against redistribution lowers the share of resources spent on affecting the distribution of income. Since

---

<sup>13</sup>Note that the two types of constitutional convention are identical when  $\eta = 1$ .

<sup>14</sup>It is conceivable that the type of government and the size of  $\eta$  and  $\gamma$  depend on the level of development reached by different democratic institutions. For instance, a government would be more likely to be self-seeking during a transition from an authoritarian political system to a democratic one, since political organizations will then typically not have had time to grow strong enough to control individual representatives to any greater extent.  $\gamma$  is also likely to be relatively small when  $k$  is small, while  $\eta$  is likely to be fairly high when  $k$  is small. However, these relationships will not be explored in detail in this paper.

<sup>15</sup>In a repeated model, exogenous rents could incorporate the expected value of future endogenous rents.

the analysis and the results differ sharply between the cooperative and non-cooperative cases, we treat the two cases separately.

### 3.1 The Cooperative Case

We consider two levels of  $\lambda$  in the cooperative case. The first is  $\lambda = 1$ , which means that organized groups control a relatively large share of initial resources. Keeping in mind that unorganized producers are more numerous than organized producers, this could signify that the degree of organization is fairly high, or that initial resources are highly concentrated to individuals who are organized.<sup>16</sup>

Differentiating  $V_i^C(y)$  with respect to  $y$  and solving for  $y$  yields the level of influence activities that interest groups find optimal,  $y^*$ , as a function of  $\alpha$  and  $k$ . We first consider the case of  $\alpha = 1$ . This represents a neutral constitution in the sense that the redistributive effectiveness is determined by the strength of interest groups and the implicit restrictions against redistribution, that is, by  $k$  alone. Figure 1 shows  $y^*(k)$  for  $\alpha = 1$ .<sup>17</sup> Maximum values of  $y^*$  occur for intermediate values of  $k$ . When  $k$  is high, less resources are wasted on activities aimed at affecting the distribution than when  $k$  is on an intermediate level. The lowest values of  $y^*$  occur when  $k$  is relatively small. For  $k$  below a certain level, no influence activities are undertaken. This level, referred to as the cutoff level of  $k$ , is given by  $k = 2$  in the case represented by Figure 1.

Before considering the effects of changes in  $\alpha$ , we will just note that the non-monotonous pattern of  $y^*$  with respect to the redistributive effectiveness is found under general conditions in the cooperative case.<sup>18</sup> The explanation for the non-monotonicity is a simple one.  $y^* = 0$  when the redistributive effectiveness is too low to compensate for the loss of production brought about by influence activities. For high levels of redistributive effectiveness, interest groups will only have to allocate a small share of their resources to influence activities in order to secure most of the production. For values of  $\alpha k$  large enough to generate influence activities and up to the value yielding the maximum level of influence activities,  $y^{\max}$ , a higher redistributive effectiveness leads to more influence activities since the increase in the share received by

---

<sup>16</sup>If unorganized individuals outnumber organized individuals by ten to one, the latter have to control twenty times more initial resources on average than the former to arrive at  $\lambda = 1$ . If organized producers are almost as numerous as unorganized producers, they only control about twice as much initial resources per person as do unorganized producers when  $\lambda = 1$ .

<sup>17</sup>See the proof of Proposition 2 for a derivation of Figures 1 through 3.

<sup>18</sup>This is made clear in Bornefalk (2000 a).

interest groups more than offsets the effects of the fall in production on the payoffs of these groups. As  $\alpha k$  increases from the level corresponding to  $y^{\max}$  influence activities fall, since the share of production received by interest groups is large enough to make the distortive effect of influence activities be more perceptible on the margin than the redistributive effect.

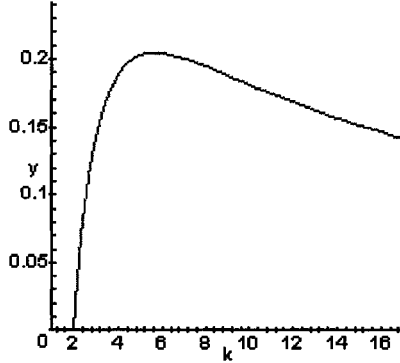


Figure 1:  $y^*(k)$  when  $\alpha = 1$  and  $\lambda = 1$ .

When the explicit restrictions against redistribution are sharpened,  $y^*(k)$  becomes more protracted. Hence, when  $\alpha$  falls,  $k$  needs to increase in order to arrive at the same level of redistributive effectiveness. This is illustrated by the rightmost curve in Figure 2, which represents the case of  $\alpha = 0.5$ . The implication for the economic performance depends on the level of  $k$ . If  $k$  is relatively low, sharpened explicit restrictions lead to less resources being wasted on influence activities. If  $k$  is relatively large, on the other hand,  $y^*$  will increase following constitutional reform in this direction. The increase will be larger the higher  $k$  is in this range, due to the protraction. The cutoff level of  $k$  also shifts to the right, from  $k_{cutoff} = 2$  when  $\alpha = 1$  to  $k_{cutoff} = 4$  when  $\alpha = 0.5$ . Hence, influence activities will be absent within a wider range of  $k$  than prior to the reform.

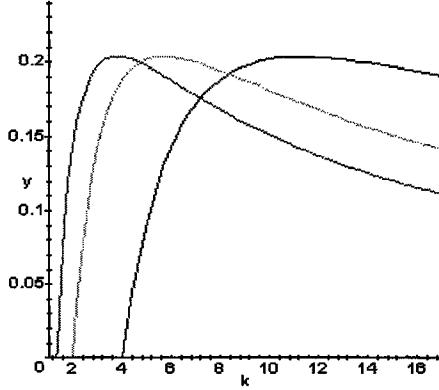


Figure 2:  $y^*(k)$  when  $\lambda = 1$  and  $\alpha = 1.5, 1$ , and  $0.5$  from left to right.

Constitutional reform can also increase the ability of the government to achieve redistribution. When the explicit restrictions against redistribution are lowered,  $y^*(k)$  becomes more contracted. This is illustrated by the left-most curve in Figure 2, which represents the case of  $\alpha = 1.5$ . Hence, influence activities increase for relatively low levels of  $k$ , and fall for relatively high levels. Constitutional reform in this direction makes influence activities relatively more profitable for low values of  $k$  than previously. As a consequence,  $k_{cutoff}$  shifts to the left. The same is true for  $y^{\max}$ , since interest groups reach a point where they receive a sufficiently large share of production so as to suffer from the distortive effects of influence activities for a lower level of  $k$ .

The pattern brought forward by Figure 2 is a general one. To express this, we define  $\tilde{k}$  as the level of  $k$  where two curves corresponding to different  $\alpha$  intersect each other.  $\tilde{k}(\underline{\alpha}, \hat{\alpha})$  represents the intersection between curves corresponding to a lower and an intermediate value of  $\alpha$ , and  $\tilde{k}(\hat{\alpha}, \bar{\alpha})$  that of an intermediate and a higher value of  $\alpha$ . We also define  $\underline{k}_{cutoff}$  as the cutoff level of  $k$  corresponding to the highest  $\alpha$  in a comparison.  $\underline{k}_{cutoff}$  will then be the lowest cutoff level. We can now establish the following lemmas, which hold as long as  $y^*(k)$  has a non-monotonous shape of the kind shown in Figures 1—3.

**Lemma 1** *In the interval  $k \in (\underline{k}_{cutoff}, \tilde{k}(\underline{\alpha}, \hat{\alpha}))$ , increasing the explicit restrictions against redistribution from  $\hat{\alpha}$  to  $\underline{\alpha}$  reduces influence activities. For  $k \geq \tilde{k}(\underline{\alpha}, \hat{\alpha})$ , influence activities will increase. The increase will be larger the higher  $k$  is in the upper interval.*

**Proof.** Immediate.

In the case represented by Figure 2,  $\tilde{k}(\underline{\alpha}, \hat{\alpha}) = 4 + \frac{4}{3} \ln \left(1 + \frac{1}{2}\sqrt{6}\right) + \frac{4}{3}\sqrt{6} \approx 8.33$ .

**Lemma 2** *In the interval  $k \in (\underline{k}_{cutoff}, \tilde{k}(\hat{\alpha}, \bar{\alpha})]$ , reducing the explicit restrictions against redistribution from  $\hat{\alpha}$  to  $\bar{\alpha}$  increases influence activities. For  $k \geq \tilde{k}(\hat{\alpha}, \bar{\alpha})$ , influence activities will fall. The reduction will be larger the higher  $k$  is in the upper interval.*

**Proof.** Immediate.

In the case represented by Figure 2,  $\tilde{k}(\hat{\alpha}, \bar{\alpha}) \approx 4.70$ .

While  $\alpha = 1$  yields a better economic performance than  $\alpha = 0.5$  when  $k > 8.33$  and a better economic performance than  $\alpha = 1.5$  when  $k < 4.70$ , it will never yield an economic performance that is better than both alternatives. Moreover, in the interval  $k \in (4.70, 8.33)$ ,  $\alpha = 1$  yields more influence activities than both the alternatives. The same pattern holds for any intermediate  $\alpha$ . This is made clear in the following proposition, where we refer to the alternative to  $\hat{\alpha}$  that yields least influence activities as the best alternative:

**Proposition 1** *In the interval  $k \in [\tilde{k}(\hat{\alpha}, \bar{\alpha}), \tilde{k}(\underline{\alpha}, \hat{\alpha})]$ , the intermediate  $\alpha$  yields more influence activities than both  $\underline{\alpha}$  and  $\bar{\alpha}$ .  $\hat{\alpha}$  never yields less influence activities than the best alternative.  $y^*(\hat{\alpha})$  is only equal to  $y^*$  for the best alternative if  $k \leq k_{cutoff}(\hat{\alpha})$ .*

**Proof.** Lemma 1 and 2 establish the first part of Proposition 1. The second part is established by the fact that minimum levels of influence activities for the cases  $\underline{\alpha}$ ,  $\hat{\alpha}$ , and  $\bar{\alpha}$  are given by  $\underline{\alpha}$  in the range  $k \in [1, \tilde{k}(\underline{\alpha}, \bar{\alpha})]$ , and by  $\bar{\alpha}$  in the range  $k \geq \tilde{k}(\underline{\alpha}, \bar{\alpha})$ . The third part is given by the observation that the interval  $k \in [1, k_{cutoff}(\hat{\alpha})]$  is contained in the interval  $k \in [1, k_{cutoff}(\underline{\alpha})]$  since  $k_{cutoff} = \frac{2}{\alpha\lambda}$ .<sup>19</sup>

$\tilde{k}(\underline{\alpha}, \bar{\alpha}) \approx 7.19$  when  $\underline{\alpha} = 0.5$  and  $\bar{\alpha} = 1.5$ . Hence, a constitution corresponding to  $\underline{\alpha} = 0.5$  yields the least influence activities of the ones considered here when  $k < 7.19$ . When  $k > 7.19$ ,  $\bar{\alpha} = 1.5$  yields the least influence activities. Proposition 1 has a Corollary which is good news for any constitutional economist:

**Corollary 1** *Constitutional reform can always reduce influence activities if  $y^* > 0$ .*

---

<sup>19</sup>See the proof of Proposition 4.

We now go on to analyze the case where interest groups control a small share of initial resources. Figure 3 shows  $y^*(k)$  when  $\lambda = 4$  for  $\alpha = 1$ ,  $\alpha = 0.5$ , and  $\alpha = 1.5$ , respectively.  $y^*(k)$  changes in three respects when  $\lambda$  is increased. The first is that  $k_{cutoff}$  falls, the second that  $y^{\max}$  increases, and the third that the curve becomes more contracted, which is seen in pairwise comparisons between curves corresponding to the same  $\alpha$  in Figure 2 and Figure 3. The explanation for these changes is that the redistributive effect becomes relatively more important than the distortive effect when unorganized groups control a larger share of the resources, as long as interest groups have not managed to secure a high enough share of final production to make them suffer from the distortive effects. Hence, influence activities are profitable for lower levels of  $k$  and with greater intensity than when  $\lambda$  is smaller, since the share of the inputs to production that is negatively affected by influence activities has become smaller. This, in turn, allows interest groups to reach a point where they secure a sufficiently large share of final production to suffer from the distortive effects of influence activities at a lower level of  $k$  than before. The implications for the effect of constitutional reform in terms of the extent to which influence activities take place can be expressed as follows:

**Proposition 2** *The higher  $\lambda$  is, the lower  $\tilde{k}(\hat{\alpha}, \bar{\alpha})$ ,  $\tilde{k}(\underline{\alpha}, \bar{\alpha})$  and  $\tilde{k}(\underline{\alpha}, \hat{\alpha})$  are.*

**Proof.** See the Appendix.

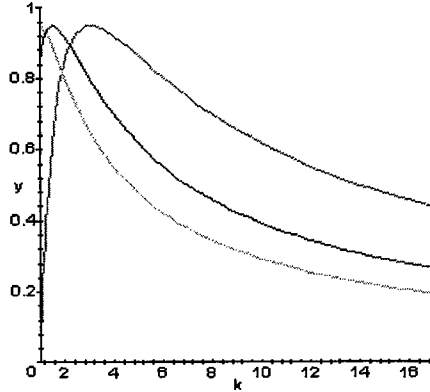


Figure 3:  $y^*(k)$  when  $\lambda = 4$  and  $\alpha = 1.5$ , 1, and 0.5 from left to right.

Hence, the smaller the share of initial resources controlled by interest groups is, the lower their ability to affect the distribution of income needs to be for sharpened explicit restrictions against redistribution to lower influence

activities. The interaction between  $\lambda$  and  $\alpha$  is worth emphasizing. When  $\alpha$  is increased, the redistributive effects of influence activities become relatively more important than the distortive effects, thereby reinforcing the effect of a larger  $\lambda$ . Hence, facilitation of redistribution when  $\lambda$  is large yields a situation where interest groups secure a sufficiently large share of final production to make the distortive effects dominate the redistributive effects even for relatively low levels of  $k$ . Only for low levels of  $k$  will strengthened restrictions against redistribution result in distortive effects being large enough to dominate over the redistributive effects when interest groups receive a relatively small share of the final production.

Proposition 2 implies what could be called the Catch 22 of constitutional reform. When  $\lambda$  is small,  $k$  needs to be quite low to make a lowering of  $\alpha$  lead to a reduction of influence activities. The problem is that the implicit restrictions against redistribution are likely to be weak when few groups have had time to overcome their collective action problems. The reason is that it takes time for political parties and a free press to develop, just to mention two important pillars of a democratic culture. Hence, if interest groups are relatively effective but small, and democratic institutions are little developed, an introduction, or strengthening, of constitutional constraints against redistribution could very well impel these groups to waste more resources on unproductive influence activities. Moreover, once the share of the economy that is controlled by organized groups has become sufficiently large, and the implicit restrictions against redistribution sufficiently strong, to make a strengthening of the explicit restrictions against redistribution lead to less influence activities, interest groups could very well be influential enough to resist such a reform successfully since they now represent a large share of the society. Bornefalk (2000a) shows that they have incentives to do so as long as they are able to cooperate with each other.

The chances to strengthen the explicit restrictions against redistribution and thereby improve the economic performance, when interest groups are able to cooperate with each other, would then be highest when the democratic culture is fairly well developed and the influence of interest groups is relatively small. This situation is not a common one, but could come about as a consequence of a major crisis, such as a war. Germany in the late 1940s could be an example of such a country.<sup>20</sup> This mechanism could then, together with the fact that Germany experienced hyperinflation in the early 1920s, be part of an explanation as to why Germany managed to secure a

---

<sup>20</sup>Olson (1982, pp. 75-76) points out the importance of the weakened position of German interest groups, which is explained by Hitler's totalitarian rule and the consequent foreign occupation, for that country's rapid postwar economic growth.

protection of the value of its currency in its constitution of 1949. Note that an explanation based only on Germany's experiencing hyperinflation is insufficient since most countries that have experienced hyperinflation lack such constitutional provisions. Before turning to the topic of constitutional choice in Section 4, we will consider the case when interest groups are unable to cooperate with each other.

### 3.2 The Non-Cooperative Case

Interest groups that do not cooperate have to optimize their expected payoffs given their beliefs about their opponent's decision. They therefore look for Nash equilibria,  $y_i^*$ ,  $i = 1, 2$ , for different  $\lambda$ ,  $k$ , and  $\alpha$ . Since the game is symmetric,  $y_1^* = y_2^*$ . We refer to this value as  $y_{NE}^*$ . The curve to the left in Figure 4 shows  $y_{NE}^*(k)$  for  $\lambda = 1$  and  $\alpha = 1$ .<sup>21</sup>

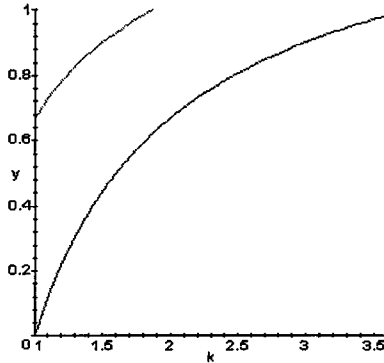


Figure 4:  $y_{NE}^*(k)$  when  $\lambda = 1$  and  $\alpha = 1$  and  $0.5$  from left to right.

The monotonicity of  $y_{NE}^*(k)$  makes the effects of strengthened explicit restrictions against redistribution unambiguous. When the explicit restrictions increase, that is, when  $\alpha$  is lowered,  $y_{NE}^*(k)$  becomes more protracted. When the explicit restrictions are reduced, this effect is reversed. The implications are as follows:

**Proposition 3** *When interest groups do not cooperate, a move from  $\hat{\alpha}$  to  $\underline{\alpha}$  will lower influence activities if  $\underline{\alpha}$  is low enough to bring down  $y_{NE}^*$  below 1. Influence activities will never increase when  $\alpha$  is lowered. A move from  $\hat{\alpha}$  to  $\bar{\alpha}$  will increase influence activities unless  $y_{NE}^*(k) = 1$  at the initial level of explicit restrictions.*

<sup>21</sup>See the Appendix for a derivation of Figure 4.



The curve to the right in Figure 4 shows  $y_{NE}^*(k)$  for  $\lambda = 1$  when  $\alpha$  has been lowered to 0.5. This reduces influence activities in the range  $k \in (1, 3.6)$ . Note that the level of  $k$  determines how extensive the constitutional reform needs to be to accomplish a fall in influence activities.

## 4 Constitutional Reform

Having studied the effects on the economic performance of constitutional reform under different conditions, we can now ask what types of constitutional assembly are willing to undertake what type of constitutional reform. We will consider both the case where any  $\alpha$  can be chosen and the case where the constitutional assembly is constrained in its choice to constitutional reform corresponding to certain values of  $\alpha$ .

### 4.1 The Self-Seeking Constitutional Assembly

To analyze the preferences of the self-seeking constitutional assembly, the following definition is helpful:

**Definition 1** *A performance motivated self-seeking constitutional assembly prefers productive activities to influence activities. A corrupt constitutional assembly prefers influence activities to productive activities.*

With a linear production function of our type, it is easily established under what conditions the assembly will be performance motivated:

**Lemma 3** *A self-seeking constitutional assembly is performance motivated when  $\gamma < \frac{\eta}{1-\eta}$ . It is corrupt when  $\gamma > \frac{\eta}{1-\eta}$ .*

**Proof.** See the Appendix.

Hence, constitutional assemblies that attach equal weight to production could be either performance motivated or corrupt, depending on the size of the transaction costs in corruption that they face. Lemma 3 reveals that a self-seeking constitutional assembly that attaches a larger weight to production than  $\eta = \frac{1}{2}$  is always performance motivated with our linear production function since  $\gamma \in [0, 1]$ . When  $\eta < \frac{1}{2}$ , the higher  $\gamma$  is, the higher  $\eta$  must be for the constitutional assembly to be performance motivated. Lemma 3 is valid for both the cooperative and the non-cooperative case.

We can now establish the following propositions for assemblies that can assign any value to  $\alpha$ :

**Proposition 4** *An unconstrained self-seeking constitutional assembly will choose  $\alpha$  such that no influence activities take place if it is performance motivated. In the cooperative case, any  $\alpha \in [0, \alpha_C^*]$ , where  $\alpha_C^* = \frac{2}{k\lambda}$ , yields  $y^*(\alpha) = 0$ . In the non-cooperative case, any  $\alpha \in [0, \alpha_{NC}^*]$ , where  $\alpha_{NC}^* = \frac{1}{k(1+\lambda)}$ , yields  $y_{NE}^*(\alpha) = 0$ .*

**Proof.** See the Appendix.

Proposition 4 tells us that an unconstrained performance motivated assembly can always find an  $\alpha$  that allows it to abolish influence activities. Since  $\frac{2}{k\lambda} > \frac{1}{k(1+\lambda)}$ , we can also conclude that the highest level of  $\alpha$  for which influence activities are absent is higher in the cooperative than in the non-cooperative case. A corrupt assembly will act as follows:

**Proposition 5** *A corrupt constitutional assembly will choose  $\alpha$  such that  $y^{\max}$  is realized in the cooperative case if it is unconstrained. In the non-cooperative case, it will choose  $\alpha$  such that  $y_{NE}^*(\alpha) = 1$ . The optimal constitution is in each case a function of  $k$  and  $\lambda$ .*

We are unable to solve for  $y^{\max}$  explicitly, but we have seen that changes in  $\lambda$  affect its location, which establishes the proposition.

The assumption that the constitutional assembly can engage in constitutional fine tuning is likely to be unrealistic in many cases. Values of  $\alpha$  that are below or above certain points could also be unfeasible. In the former case, this could be due to excessive decision time costs. We therefore go on to analyze constitutional reform when the assembly is constrained to certain values of  $\alpha$ .

Recapitulating Proposition 1, we realize that a performance motivated constrained constitutional assembly dealing with interest groups that cooperate prefers to strengthen the constitutional constraints against redistribution when the ability of interest groups to affect the distribution of income is relatively poor, and to reduce the explicit restrictions when their ability is relatively strong. When  $k$  is in the intermediate range, that is, when  $k \in [\tilde{k}(\hat{\alpha}, \bar{\alpha}), \tilde{k}(\underline{\alpha}, \hat{\alpha})]$ , both alternatives are preferred to  $\hat{\alpha}$ , the initial level of explicit restrictions, since both alternatives yield less influence activities. If the assembly is dealing with interest groups that do not cooperate, it prefers to strengthen the explicit restrictions against redistribution provided that the constitutional reform under consideration is of sufficient magnitude to make  $y_{NE}^* < 1$ . Assuming that  $y^*(\hat{\alpha}) > 0$  in the cooperative case and that  $y_{NE}^*(\hat{\alpha}) > 0$  in the non-cooperative case, we have:

**Proposition 6** *In the cooperative case, a performance motivated self-seeking constitutional assembly that is constrained in its constitutional choice to  $\underline{\alpha}$ ,*

$\hat{\alpha}$ , and  $\bar{\alpha}$  prefers both  $\underline{\alpha}$  and  $\bar{\alpha}$  to  $\hat{\alpha}$  when  $k \in [\tilde{k}(\hat{\alpha}, \bar{\alpha}), \tilde{k}(\underline{\alpha}, \hat{\alpha})]$ . It will choose  $\underline{\alpha}$  when  $k \in [1, \tilde{k}(\underline{\alpha}, \bar{\alpha})]$ . When  $k \geq \tilde{k}(\underline{\alpha}, \bar{\alpha})$ , it will choose  $\bar{\alpha}$ . In the non-cooperative case, it will choose  $\underline{\alpha}$ .

A corrupt constrained constitutional assembly dealing with cooperating interest groups prefers status quo when  $k$  is in the intermediate range since this gives maximum levels of influence activities. When the implicit restrictions are relatively weak, it prefers to strengthen the constitutional constraints against redistribution. When  $k$  is relatively low, finally, it prefers a constitution that facilitates redistribution. If it is dealing with interest groups that do not cooperate, it always prefers to facilitate redistribution unless we already have  $y_{NE}^* = 1$ . Hence:

**Proposition 7** *In the cooperative case, a corrupt constitutional assembly that is constrained to  $\underline{\alpha}$ ,  $\hat{\alpha}$ , and  $\bar{\alpha}$  will remain at  $\hat{\alpha}$  when  $k \in [\tilde{k}(\hat{\alpha}, \bar{\alpha}), \tilde{k}(\underline{\alpha}, \hat{\alpha})]$ . When  $k > \tilde{k}(\underline{\alpha}, \hat{\alpha})$ , it will choose  $\underline{\alpha}$ . When  $k < \tilde{k}(\hat{\alpha}, \bar{\alpha})$ , it will choose  $\bar{\alpha}$ . In the non-cooperative case, it will choose  $\bar{\alpha}$  unless  $y_{NE}^*(\hat{\alpha}) = 1$ .*

The finding that the ideal constitution for a corrupt constitutional assembly is the one that makes interest groups choose  $y^{\max}$  in the cooperative case and  $y_{NE}^* = 1$  in the non-cooperative case provides an argument in favor of the view that constitutional reform should be an affair for a constitutional assembly representing broad layers of the population. We now go on to consider the constitutional assembly that cares about the distribution of income.

## 4.2 The Biased Constitutional Assembly

The evaluation of the preferences for the biased constitutional assembly is more complicated than for the self-seeking one. We will therefore make use of simulated outcomes for different types of assembly and different values of  $\lambda$  and  $\eta$  in our analysis. We also limit our analysis to assemblies that are constrained to three levels of  $\alpha$  in their constitutional choice.

Starting with the cooperative case, Figures 5–8 show outcomes for constitutional assemblies favoring interest groups, that is,  $A_{B_{IG}}^C$ , as a function of  $k$  for  $\alpha = 2, 1$ , and  $\frac{1}{2}$ , and different combinations of  $\lambda$  and  $\eta$ . Figures 9–12 show outcomes for assemblies favoring unorganized groups, that is,  $A_{B_{UG}}^C$ . To derive these curves, we first substitute the expression for the level of influence activities chosen by interest groups, that is,  $y^*(\alpha, k, \lambda)$ <sup>22</sup> in the cooperative

<sup>22</sup>See the proof of Proposition 2 in the Appendix.

version of equation (6). We then use different combinations of  $\lambda$ ,  $\alpha$ , and  $\eta$  in the resulting evaluation function, and end up with curves showing how the outcomes for the constitutional assembly depend on the interplay between  $k$  and  $\alpha$ .

The location of the point of intersection for curves corresponding to different  $\alpha$  again plays an important role in the analysis. We refer to this point as  $\tilde{k}(A_{B_{IG}}^C)$  when we compare outcomes for an assembly biased towards interest groups, and  $\tilde{k}(A_{B_{UG}}^C)$  when we compare outcomes for an assembly biased towards unorganized groups. For  $k < \tilde{k}$ , both types of assembly prefer the lower level of  $\alpha$ .

A common characteristic of all curves is the more or less pronounced non-monotonicity, which is explained by the non-monotonicity in the allocative choice of interest groups, as discussed in Section 3. It is this non-monotonicity that makes the constitution preferred by the constitutional assembly a function of  $k$ . Since the sizes of  $\eta$  and  $\lambda$  affect the degree of non-monotonicity, they also affect the location of  $\tilde{k}$ , and thereby the constitutional choice.

We first consider the choice of constitution by an assembly favoring interest groups when these control a share of initial resources corresponding to  $\lambda = 1$ . In Figure 5, the assembly gives equal weight to total production and the expected payoffs of the interest group it favors. Hence,  $\eta = \frac{1}{2}$ . In Figure 6,  $\eta = \frac{9}{10}$ , which means that the assembly is only slightly biased.

To understand the effect of  $\eta$  on the choice of constitution, we first sort out a couple of less interesting properties of the curves. First, the levels of  $k$  below which influence activities do not take place for different  $\alpha$  are the same for  $\eta = \frac{1}{2}$  and  $\eta = \frac{9}{10}$ , since interest groups do not consider the priorities of the government when making their allocative decisions.<sup>23</sup> Second, the value of  $A_{B_{IG}}^C$  for values of  $k \leq k_{cutoff}$  is bigger when  $\eta = \frac{9}{10}$  than when  $\eta = \frac{1}{2}$ . The explanation for this is simply that total production,  $C(y(\alpha))$ , which is then given a larger weight, is larger than the part of production that is received by the group that the assembly favors, that is,  $V_i^C(y(\alpha))$  for  $i = 1, 2$ .

---

<sup>23</sup>The relationship  $k_{cutoff} = \frac{2}{\alpha\lambda}$  determines these levels, as is made clear in the proof of Proposition 4.

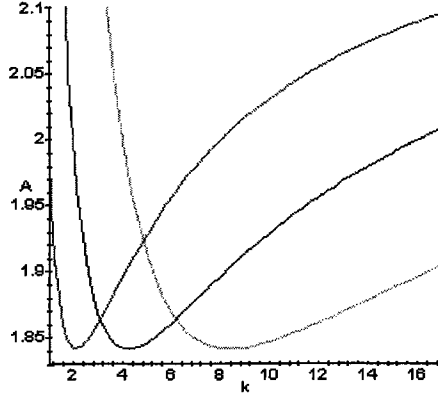


Figure 5:  $A_{BIG}^C$  for  $\lambda = 1$ ,  $\eta = \frac{1}{2}$ , and  $\alpha = 2, 1$ , and  $\frac{1}{2}$  from left to right.

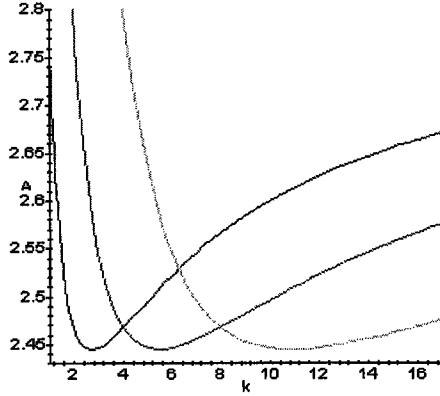


Figure 6:  $A_{BIG}^C$  for  $\lambda = 1$ ,  $\eta = \frac{9}{10}$ , and  $\alpha = 2, 1$ , and  $\frac{1}{2}$  from left to right.

When values of  $A_{BIG}^C$  are compared for high and low levels of  $k$ , we see that the outcomes for the assembly with  $\eta = \frac{1}{2}$  are relatively more preferable for high than for low levels of  $k$  compared with the case of  $\eta = \frac{9}{10}$ . The same tendency, albeit less pronounced, is seen in a comparison between Figures 7 and 8, which show  $A_{BIG}^C$  when  $\lambda = 4$  for  $\eta = \frac{1}{2}$  and  $\eta = \frac{9}{10}$ , respectively. In general, the lower  $\eta$  is, the larger  $A_{BIG}^C$  for a relatively high level of  $k$  will be relative to  $A_{BIG}^C$  for a relatively low level of  $k$ . This effect of  $\eta$  on the relative size of  $A_{BIG}^C$  for different levels of  $k$  will, in turn, affect the outcome of constitutional reform. To show this, we will make use of the following lemma:

**Lemma 4** *The larger  $k$  is, the larger  $V_i^C(y(\alpha))$ ,  $i = 1, 2$ , will be relative to  $C(y(\alpha))$ .*

**Proof.** See the Appendix.

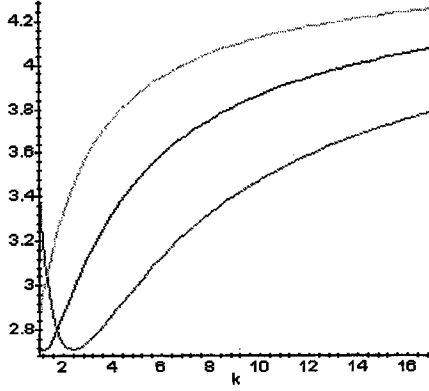


Figure 7:  $A_{B_{IG}}^C$  for  $\lambda = 4$ ,  $\eta = \frac{1}{2}$ , and  $\alpha = 2, 1$ , and  $\frac{1}{2}$  from left to right.

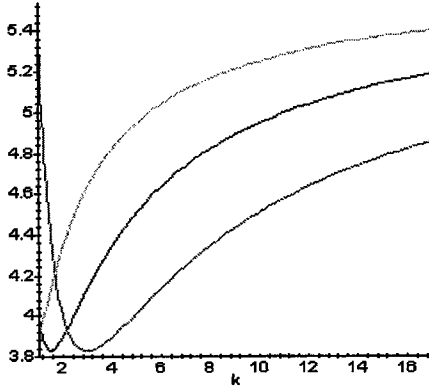


Figure 8:  $A_{B_{IG}}^C$  for  $\lambda = 4$ ,  $\eta = \frac{9}{10}$ , and  $\alpha = 2, 1$ , and  $\frac{1}{2}$  from left to right.

The intuition behind the proof of the lemma is quite simple. Since we evaluate a quotient where the size of total production appears in both the numerator and the denominator, the quotient reduces to  $p_i^C(y(\alpha))$ . Hence, we only need to determine how the share received by interest groups depends on  $k$ . Since a higher level of  $k$  enhances the possibilities for interest groups to influence the distribution of income, and since these always prefer a larger share to a smaller one,  $p_i^C(y(\alpha))$  is increasing in  $k$ . It then follows that the size of the share received by interest groups relative to the size of total production is also increasing in  $k$ .

How, then, does  $\eta$  enter in the constitutional choice? We first note that since  $V_i^C(y(\alpha))$  is counted twice for  $i = 1, 2$  by the constitutional assembly favoring interest groups, and since this part of total production is increasing in  $k$ , the share of total production that is counted twice is also increasing in  $k$ . We then make use of the fact that the lower  $\eta$  is, the higher the assembly values the increase in the share of production that is counted twice. Since total production also increases after  $y^{\max}$  has been reached, this explains why the curves corresponding to  $\eta = \frac{1}{2}$  are quite steep to the right of their troughs. When  $\eta = \frac{9}{10}$ ,  $A_{BIG}^C$  still bends upwards for relatively large values of  $k$ , but now mostly because total production is larger for values of  $k$  higher than the one yielding  $y^{\max}$ . To the left of the troughs, curves corresponding to  $\eta = \frac{9}{10}$  are steeper than those corresponding to  $\eta = \frac{1}{2}$ . The explanation is as follows. The increase in the part of production that is received by interest groups as  $k$  is increased is valued less by the assembly that is only slightly biased towards interest groups. Hence, the increase does less to neutralize the accompanying fall in total production. The fall in total production is, accordingly, felt more by the slightly biased assembly.

The changes in the shape of the curves indicate that the smaller  $\eta$  is, the lower  $k$  needs to be in order to make a constitutional assembly biased towards an organized group prefer to lower  $\alpha$ . This is seen by comparing the intersections between curves corresponding to different  $\alpha$ . Hence, the smaller  $\eta$  is, the more  $A_{BIG}^C$  bends upwards for large values of  $k$ , and the lower  $\tilde{k}(A_{BIG}^C)$  therefore becomes. This result also holds when  $\lambda = 4$ , as can be seen by making a comparison between Figures 7 and 8.

Comparing Figures 5 and 7, and Figures 6 and 8, we see that  $\tilde{k}(A_{BIG}^C)$  shifts to the left as  $\lambda$  is increased. The explanation for this is found in the way  $y^*(k)$  changes as  $\lambda$  is increased, as discussed in connection with Proposition 2. The increase in  $\lambda$  shifts  $k_{cutoff}$  to the left, increases  $y^{\max}$ , and makes  $y^*(k)$  more contracted. As a result,  $A_{BIG}^C$  also shifts to the left and becomes more contracted. Moreover, the size of  $A_{BIG}^C$  for high levels of  $k$  relative to that for low levels increases as  $\lambda$  increases. This is explained by the fact that the size of the quotient  $\frac{V_i^C(y(\alpha))}{C(y(\alpha))}$  for high relative to low levels of  $k$  is larger the larger  $\lambda$  is. When  $\lambda$  is big, interest groups only receive a small share of total production when  $k$  is low. There is therefore only a small fraction of total production that is counted twice.  $C(y(\alpha))$  is also comparatively low for small levels of  $k$  since influence activities are already profitable for low levels of  $k$  when  $\lambda$  is big. We can now formulate the following proposition:

**Proposition 8** *When the constitutional assembly is biased towards interest groups and these cooperate,  $k$  needs to be lower to make the assembly prefer to lower  $\alpha$  the lower  $\eta$  is and the larger  $\lambda$  is.*

Mancur Olson's (1982) concept of the encompassing interest plays a role for the result conveyed in Proposition 8.<sup>24</sup> The role is, however, more subtle than it might first appear. In our model, the degree of encompassing interest is determined not only by  $\lambda$ , but also by the effectiveness of redistributive activities. For a given level of  $\alpha k$ , a larger  $\lambda$  corresponds to interest groups bearing a lower proportion of the losses brought about by influence activities. In other words, they have less of an encompassing interest in the economy the larger  $\lambda$  is, *ceteris paribus*. However, the interest groups have already taken the effect of  $\lambda$  into account in their choice of  $y^*$ , which in turn affects total production. Hence, it is not the interest groups themselves that are less inclined to lower  $\alpha$  as  $\lambda$  increases. Cooperating interest groups will, in fact, never prefer a lowering of the effectiveness with which they can redistribute income from unorganized groups to themselves, as shown in Bornefalk (2000a,b). Instead, their encompassing interest exerts its influence on the preferred  $\alpha$  via  $\eta$ , which represents the direct interest in total production that the constitutional assembly has. The larger  $\lambda$  is for a given level of  $k$ , that is, the lower the encompassing interest is, the greater the interest of the constitutional assembly in total production needs to be to make it prefer to lower  $\alpha$ . With  $\eta = 0$ , the assembly would never prefer to lower  $\alpha$  as long as interest groups were able to cooperate, since the assembly would then not value total production per se.

When the constitutional assembly favors unorganized groups, the effect of the size of  $\eta$  on the intersection between curves corresponding to different  $\alpha$  is reversed. The larger  $\eta$  is, that is, the less biased the assembly is towards unorganized groups, the lower  $k$  needs to be to make the assembly prefer to lower  $\alpha$ . This can be seen by comparing Figure 9 with Figure 10, and Figure 11 with Figure 12. The explanation is as follows. We first note that, by Lemma 4, the higher  $k$  is, the smaller is the size of the part of production received by unorganized groups. When this size is given a relatively large weight in the evaluation function of the constitutional assembly, that is, when  $\eta$  is relatively low, the increase in production for values of  $k$  higher than the one yielding  $y^{\max}$  is not valued high enough to make the curve bend up to any considerable extent. When  $\eta$  is relatively high, on the other hand, the increase in total production is valued highly enough to compensate for the deteriorating size of the part of production received by unorganized groups, which makes the curve bend up more. This, in turn, shifts  $\tilde{k}(A_{BUG}^C)$  to the left. Note also that the curve again bends upwards more sharply when  $\lambda$

---

<sup>24</sup>McGuire and Olson (1996) analyze redistribution from a minority to a majority by applying the concept of encompassing interest. Calmfors and Driffill (1988) use this concept to study the impact of the bargaining structure on macroeconomic performance.



is larger, as in Figures 11 and 12. The reason for this is essentially that the increase in total production is worth more than when  $\lambda$  is small since total production is larger. We can now formulate our second proposition on constitutional reform in the cooperative case:

**Proposition 9** *When the constitutional assembly is biased towards unorganized groups and interest groups cooperate,  $k$  needs to be lower to make the assembly prefer to lower  $\alpha$  the larger  $\eta$  and  $\lambda$  are.*

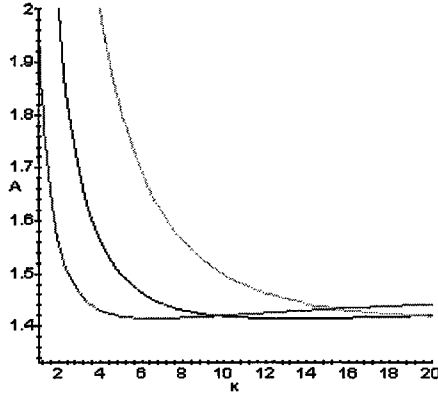


Figure 9:  $A_{BUG}^C$  for  $\lambda = 1$ ,  $\eta = \frac{1}{2}$ , and  $\alpha = 2, 1$ , and  $\frac{1}{2}$  from left to right.

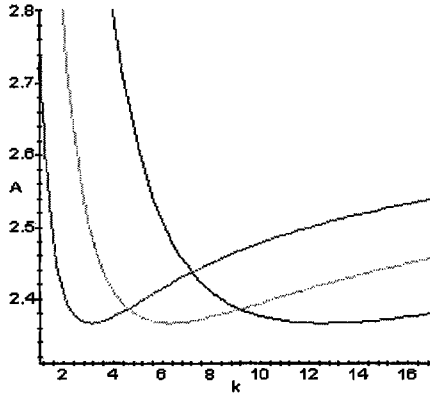


Figure 10:  $A_{BUG}^C$  for  $\lambda = 1$ ,  $\eta = \frac{9}{10}$ , and  $\alpha = 2, 1$ , and  $\frac{1}{2}$  from left to right.

The results conveyed in Propositions 8 and 9 are, as such, not surprising. Nevertheless, they do show that a constitutional assembly favoring interest groups might indeed choose to limit the possibilities for these to extract rents

from unorganized groups if it attaches a positive weight to total production. Likewise, a constitutional assembly favoring unorganized groups might under certain conditions choose to facilitate redistribution from these to interest groups. This is so despite the fact that unorganized groups become worse off when redistribution is facilitated through constitutional reform.

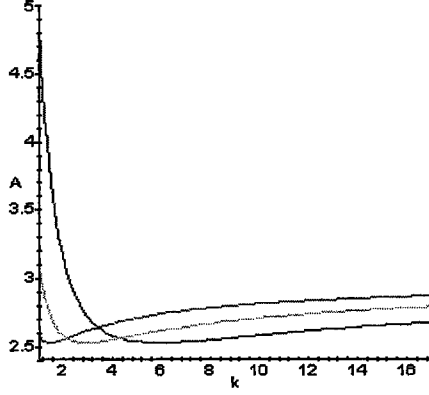


Figure 11:  $A_{BUG}^C$  for  $\lambda = 4$ ,  $\eta = \frac{1}{2}$ , and  $\alpha = 2, 1$ , and  $\frac{1}{2}$  from left to right.

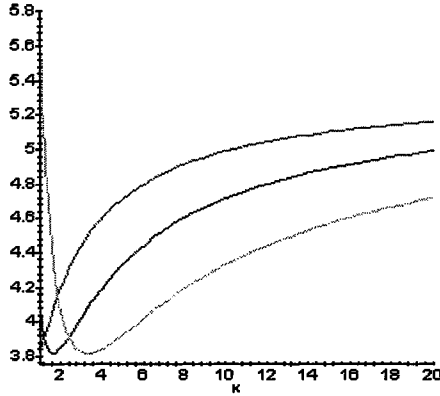


Figure 12:  $A_{BUG}^C$  for  $\lambda = 4$ ,  $\eta = \frac{9}{10}$ , and  $\alpha = 2, 1$ , and  $\frac{1}{2}$  from left to right.

When interest groups are unable to cooperate with each other, a strengthened constitutional protection of property rights will be a preferred action for a wide range of biased constitutional assemblies. Figure 13 shows a typical case where a constitutional assembly biased towards interest groups prefers to strengthen the constitutional constraints against redistribution. The lowering of  $\alpha$  leads to an increase in  $A_{BIG}^C$  throughout the range of  $k$ . The

explanation for this is that the resulting reduction in redistributive effectiveness leads to less resources being used for influence activities. This, in turn, raises production. But it also leads to an increase in the payoffs of interest groups since their payoffs are higher the lower the redistributive effectiveness is when they are unable to cooperate. See Bornefalk (2000a,b) for details.

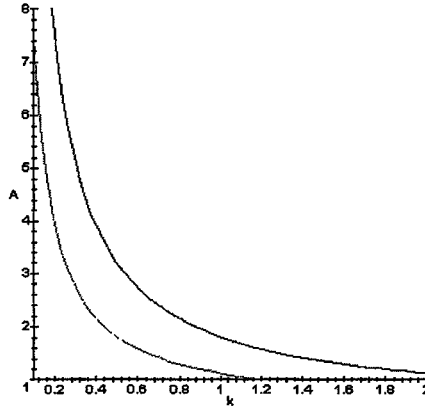


Figure 13:  $A_{B_{IG}}^C$  for  $\lambda = 1$ ,  $\eta = \frac{1}{2}$ , and  $\alpha = 1$  and  $\frac{1}{2}$  from left to right.

There is, nevertheless, a possibility that strengthened constitutional constraints against redistribution could lead to lower payoffs for a constitutional assembly that is biased towards organized groups in the non-cooperative case. This will occur when  $\lambda$  is high enough and  $\eta$  is low enough. This is a result of the payoffs of interest groups being non-monotonous with respect to the redistributive effectiveness when  $\lambda$  is large enough, as shown in Bornefalk (2000a). Figure 14 shows such a case. At the same time, the multiple intersections in Figure 14 suggest that constitutional reform is an option with little value for a society of this type.

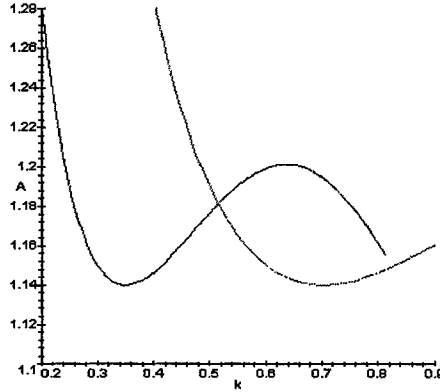


Figure 14:  $A_{BIG}^C$  for  $\lambda = 10$ ,  $\eta = \frac{1}{10}$ , and  $\alpha = 1$  and  $\frac{1}{2}$  from left to right.

## 5 Concluding Remarks

Our analysis of the impact of the initial ability of interest groups to affect the distribution of income on the effects of constitutional reform could be applied to discuss constitutional arrangements in countries with differing political and economic systems. The choice of Finland and Germany to strengthen the protection of property rights by constitutional provisions could be rationalized, as could the systematic neglect of similar constitutional provisions in less democratic countries.<sup>25</sup> Our analysis could also be used to explain why countries with a relatively well-developed democratic culture, such as Estonia and Lithuania, could benefit from introducing currency boards at an early stage of the transition to democracy and market economy, while Russia had to wait until its democratic pillars had strengthened before this option could be seriously considered although it faced similar economic problems from the outset.

The adoption of constitutional clauses facilitating redistribution could also be explained by the mechanisms studied in this paper. The hegemony of the communist party in centrally planned economies made channeling of resources to sectors given priority by the leaders more effective. Likewise, Indonesia's choice under president Suharto to rule out economic reforms threatening the dominant position of the leading clan made it easier for the government to benefit its own supporters. Our analysis suggests that economic

<sup>25</sup>See Voigt (1998) for a discussion of the poor economic performance and the lack of rule of law in many Latin American countries despite the formal similarity of their constitutions with that of the United States.

performance was improved in both cases because of the strong redistributive impetus in these societies.

An issue not touched upon in this essay is the effect of constitutional constraints on redistribution on the degree of organization. To analyze this issue, we could allow unorganized groups to have the opportunity of investing in overcoming their collective action problems. Since constitutional reform affects the relative payoffs of different groups, it should also affect the incentives to invest in organization.

## Appendix

### PROOF OF PROPOSITION 2:

We prove the proposition for the case of initial explicit restrictions represented by  $\hat{\alpha} = 1$  and the proposed new level corresponding to  $\underline{\alpha} = .5$ . To compute  $\tilde{k}(\underline{\alpha}, \hat{\alpha})$ , we first differentiate equation (5) with respect to  $y$ :

$$\begin{aligned} \frac{\partial V_i^C(y)}{\partial y} &= \alpha k \frac{e^{\alpha k y}}{2e^{\alpha k y} + \lambda} (2 - 2y + \lambda) \\ &\quad - 2 \frac{e^{2(\alpha k y)}}{(2e^{\alpha k y} + \lambda)^2} (2 - 2y + \lambda) \alpha k - 2 \frac{e^{\alpha k y}}{2e^{\alpha k y} + \lambda}. \end{aligned}$$

Equating the derivative to zero and solving for  $y^*$  as a function of  $\alpha$ ,  $k$ , and  $\lambda$  yields<sup>26</sup>

$$y^*(\alpha, k, \lambda) = - \frac{\text{LambertW}\left(\frac{2e^{\frac{1}{2}\alpha k \lambda + \alpha k - 1}}{\lambda}\right) - \frac{1}{2}\alpha k \lambda - \alpha k + 1}{\alpha k}.$$

This is the function underlying Figures 1—3. Substituting  $\alpha = 1$  and  $\alpha = .5$ , equating the two resulting expressions, and solving for  $k$  yields

$$\tilde{k}(\lambda) = \frac{4 + 4 \ln\left(1 + \frac{1}{2}\sqrt{4 + 2\lambda}\right) + \frac{8 + 4\sqrt{4 + 2\lambda}}{\lambda}}{2 + \lambda}.$$

We then differentiate  $\tilde{k}(\lambda)$  with respect to  $\lambda$ :

---

<sup>26</sup>LambertW is an implicitly defined function satisfying  $\text{LambertW}(x) e^{\text{LambertW}(x)} = x$ . The function is named after the 18th century mathematician Johann Heinrich Lambert, who was a colleague of Euler and Lagrange at the Berlin Academy of Sciences. For more on the LambertW function, see Briggs (1999) and Corless et al. (1996).

$$\frac{\partial \tilde{k}(\lambda)}{\partial \lambda} = \frac{\frac{2}{\sqrt{4+2\lambda}(1+\frac{1}{2}\sqrt{4+2\lambda})} + \frac{4}{\lambda\sqrt{4+2\lambda}} - \frac{8+4\sqrt{4+2\lambda}}{\lambda^2}}{2+\lambda} - \frac{4+4\ln\left(1+\frac{1}{2}\sqrt{4+2\lambda}\right) + \frac{8+4\sqrt{4+2\lambda}}{\lambda}}{(2+\lambda)^2}.$$

Evaluating this expression numerically for  $\lambda > 0$ , we find that  $\frac{\partial \tilde{k}(\lambda)}{\partial \lambda} < 0$  in this range.

#### DERIVATION OF FIGURE 4:

To derive the Nash equilibria, we first differentiate equation (3) with respect to  $y_1$ :

$$\begin{aligned} \frac{\partial V_1^{NC}}{\partial y_1} &= \alpha k \frac{e^{\alpha k y_1}}{e^{\alpha k y_1} + e^{\alpha k y_2} + \lambda} (2 - y_1 - y_2 + \lambda) - \\ &\quad \frac{e^{2\alpha k y_1}}{(e^{\alpha k y_1} + e^{\alpha k y_2} + \lambda)^2} (2 - y_1 - y_2 + \lambda) \alpha k - \frac{e^{\alpha k y_1}}{e^{\alpha k y_1} + e^{\alpha k y_2} + \lambda}. \end{aligned}$$

We then equate this derivative to zero and solve for  $y_1$  as a function of  $y_2$ ,  $\alpha$ , and  $k$ . This yields the best reply function

$$y_1^* = - \frac{\text{LambertW}\left(\frac{-e^{(2\alpha k - \alpha k y_2 + \alpha k \lambda - 1)}}{-e^{\alpha k y_2} - \lambda}\right) - 2\alpha k + \alpha k y_2 - \alpha k \lambda + 1}{\alpha k}.$$

Substituting  $\lambda = 1$  and  $\alpha = 1$ , we then find values of  $y_2$  for which  $y_1^*(k, y_2) = y_2$ . Plotting these values yields the curve to the left in Figure 4. Proceeding in the same way for  $\alpha = .5$  yields the curve to the right.

#### PROOF OF LEMMA 3:

Substituting  $x_i = 1 - y_i$  in equation (7) for  $i = 1, 2$  and then differentiating with respect to  $(y_1 + y_2)$  yields

$$\frac{\partial A_S^{NC}}{\partial (y_1 + y_2)} = -\eta + \gamma - \gamma\eta.$$

Equating this to zero and solving for  $\gamma$  yields  $\gamma = \frac{\eta}{1-\eta}$ . This expression shows when  $\gamma$  and  $\eta$  are such that the self-seeking constitutional assembly is indifferent to whether interest groups use their resources for production or influence activities. For higher levels of  $\gamma$  the constitutional assembly prefers

influence activities, and for lower levels it prefers productive activities. The proof is essentially identical for the cooperative case.

PROOF OF PROPOSITION 4:

Setting  $y^*(\alpha, k, \lambda) = 0$  and solving for  $\alpha$  yields  $\alpha_C^* = \frac{2}{k\lambda}$ . To derive the highest level of  $\alpha$  that yields no influence activities in the non-cooperative case, we first find an expression for  $y_{NE}^*$  by finding values of  $y_2$  for which  $y_1^*(k, y_2) = y_2$ . Setting  $y_{NE}^* = 0$  and solving for  $\alpha$  then yields  $\alpha_{NC}^* = \frac{1}{k(1+\lambda)}$ .

PROOF OF LEMMA 4:

By equation (5) we have

$$\frac{V_i^C(y(\alpha), \lambda)}{C(y(\alpha), \lambda)} = p_i^C(y(\alpha), \lambda) = \frac{e^{\alpha k y}}{2e^{\alpha k y} + \lambda}$$

for  $i = 1, 2$ . Substituting  $y^*(\alpha, k, \lambda)$  for  $y$ , and differentiating the resulting expression with respect to  $k$ , yields

$$\frac{\partial p_i^C(y^*(\alpha), \lambda)}{\partial k} = \frac{\alpha(2 + \lambda) \text{LambertW}\left(\frac{2e^{\frac{1}{2}\alpha k \lambda + \alpha k - 1}}{\lambda}\right)}{4\left(1 + \text{LambertW}\left(\frac{2e^{\frac{1}{2}\alpha k \lambda + \alpha k - 1}}{\lambda}\right)\right)^3}.$$

This expression, which is defined for  $\lambda \neq 0$ , is positive unless  $\alpha = 0$ . To see this, consider the definition of the *LambertW* function:

$$\text{LambertW}(x) e^{\text{LambertW}(x)} = x.$$

It is clear that  $\text{LambertW}(x)$  is positive if  $x$  is positive since  $e^{\text{LambertW}(x)}$  is positive by the definition of the base for the natural logarithm. Hence, since  $\left(\frac{2e^{\frac{1}{2}\alpha k \lambda + \alpha k - 1}}{\lambda}\right)$  is positive for  $\lambda > 0$ ,  $\text{LambertW}\left(\frac{2e^{\frac{1}{2}\alpha k \lambda + \alpha k - 1}}{\lambda}\right)$  is positive for  $\lambda > 0$ .

## References

- [1] Berggren, Niclas. "Social Order through Constitutional Choice: A Contractarian Proposal." *Public Choice*, 1996, 89, 339-361.
- [2] Bornefalk, Anders. "Social Conflict with Passive Groups." in *Essays on Social Conflict and Reform*, Ph.D. Dissertation, Stockholm School of Economics, 2000a.

- [3] Bornefalk, Anders. "Democratization, Rent Seeking, and Economic Transition." in *Essays on Social Conflict and Reform*, Ph.D. Dissertation, Stockholm School of Economics, 2000b.
- [4] Briggs, Keith. "W-ology or Some Exactly Solvable Growth Models." available [online:] at <http://epidem13.plantsci.cam.ac.uk/~kbriggs/W-ology.html>, 1998/11/12.
- [5] Buchanan, James M. and Tullock, Gordon. *The Calculus of Consent*. University of Michigan Press, Ann Arbor, 1962.
- [6] Calmfors, Lars and Driffill, John. "Bargaining Structure, Corporatism and Macroeconomic Performance." *Economic Policy*, 1988, 3, 14-61.
- [7] Corless, Robert M. et al. "On the Lambert W Function." *Advances in Computational Mathematics*, 1996, 5, 329-59.
- [8] Gwartney, James D. and Wagner, Richard E, eds. *Public Choice and Constitutional Economics*. JAI Press Inc., London, 1988.
- [9] Hirshleifer, Jack. "Conflict and Rent-seeking Success Functions: Ratio vs. Difference Models of Relative Success." *Public Choice*, 1989, 63, 101-112.
- [10] Mueller, Dennis C. *Constitutional Democracy*. Oxford University Press, Oxford, 1996.
- [11] Mueller, Dennis C. *Public Choice II*. Cambridge University Press, Cambridge, 1989.
- [12] McGuire, Martin and Olson, Mancur. "The Economics of Autocracy and Majority Rule: The Invisible Hand and the Use of Force." *Journal of Economic Literature*, March 1996, XXXIV, 72-96.
- [13] Olson, Mancur. *The Rise and Decline of Nations*. Yale University Press, New Haven and London, 1982.
- [14] Olson, Mancur. *The Logic of Collective Action*. Harvard University Press, Cambridge, MA, 1965.
- [15] Persson, Torsten, Roland, Gerard, and Tabellini, Guido. "Separation of Powers and Political Accountability." *Quarterly Journal of Economics*, 1997a, 112, 1163-1202.



- [16] Persson, Torsten, Roland, Gerard, and Tabellini, Guido. "Comparative Politics and Public Finance." mimeo, 1997b.
- [17] Söderström, Lars. "The Wicksellian Alternative." *Skandinaviska Enskilda Banken Quarterly Review*, 1986, 88-96.
- [18] Tullock, Gordon. "Theoretical Forerunners." Appendix 2 in Buchanan and Tullock, 1962.
- [19] Voigt, Stefan. "Making Constitutions Work: Conditions for Maintaining the Rule of Law." *Cato Journal*, 1998, 18, 191-208.
- [20] Wicksell, Knut. "A New Principle of Just Taxation." *Finanztheoretische Untersuchungen*. Jena, 1896, reprinted in Gwartney and Wagner, 1988, pp. 117-130.



## Essay 4

# **The Break-up of the Ruble Zone**

## **Undertaking Monetary Reform while Building Democratic Institutions**



# The Break-up of the Ruble Zone

## Undertaking Monetary Reform while Building Democratic Institutions

Anders Bornefalk\*

February 10, 2000

### Abstract

This essay studies the development of a market-based payments system in the former Soviet Union. Particular attention is given to the costly delay in establishing well-delimited currency areas. Differences in policies between the newly independent states are explained by differences in political and economic preconditions. These differences determined whether powerful groups preferred a monetary system that facilitated exchange and access to new markets, or one that facilitated redistribution of wealth. Considering Russia's monetary reform, we find two main reasons for the slow pace. The first is the lack of a coherent reform program during a period in which conditions would have allowed a radical dissolution of the ruble zone. The second is that once such a program had been developed, conditions had changed in such a way that Russian state enterprise managers preferred a monetary system that enabled them to acquire rents more effectively. The great differences between both outcomes and preconditions for the 15 successor states make the break-up of the ruble zone a clear example of the importance of political and economic preconditions for the possibility of undertaking successful economic reforms in countries undergoing a major systemic change.

---

\* SITE and Department of Economics, Stockholm School of Economics, Box 6501, SE-113 83 Stockholm, Sweden. [Anders.Bornefalk@hhs.se](mailto:Anders.Bornefalk@hhs.se) Research was carried out at the Institute for the Economy in Transition, Moscow, Russian Federation. Discussions with Anders Åslund, Brigitte Granville, Grigory Gritsenko, Ardo Hansson, Vadim Ivanov, Andrei Lushin, Sten Luthman, Örjan Sjöberg, Sergei Tagor and Pavel Teplukhin were helpful. Comments from Lars Jonung on different versions of this paper were particularly valuable. Comments from Erik Berglöf and participants at seminars at the Stockholm School of Economics are also appreciated. The responsibility for any errors is mine. Financial support from SITE and the Swedish Institute is gratefully acknowledged.

## 1. INTRODUCTION

As the Soviet Union was broken up on December 8, 1991, all 15 successor states kept the Soviet ruble as their currency. There was no effective coordination of monetary and fiscal policies despite the well-known fact that a monetary union requires credible commitment among member states to sound fiscal and monetary policies to function well. On the contrary, each member state had established a national central bank which could issue ruble credits without restriction. The Central Bank of Russia (CBR) was, however, in charge of the emission of ruble bank notes for the whole ruble area since it controlled the printing presses.<sup>1</sup> With several central banks in a currency area, each with an independent authority to issue credits, incentives to free ride on other member states will necessarily be overwhelming.<sup>2</sup> High inflation rates will result, which in turn will bring the monetary union to an abrupt end.<sup>3</sup> Since these faults could not be overcome, the ruble zone had to fall apart.

The dissolution became a lengthy process despite the obvious weaknesses of the currency regime. Russia had to pay a high price for the delay in delimiting its currency area. In 1992, Russia provided the other former Soviet republics (FSRs) with trade credits amounting to 8.5 percent of its GDP at highly negative real interest rates and with low expectations of repayment. The resulting drain of resources from Russia to the other member states contributed to inflation in Russia and made stabilization impossible. Despite the immense subsidies, trade between Russia and the other FSRs fell dramatically. Furthermore, the ruble zone sparked a behavior that led to extreme levels of inflation in those non-Russian member states that did not leave it at an early stage, and postponed the introduction of economic reforms.

The delimitation of currency areas was the main task in creating a monetary system that could handle market-based payments, but there were several other complicated monetary problems facing the successor states of the Soviet Union. Despite this complexity, leading members of the Russian reform government, notably acting prime minister Yegor Gaidar, realized the inherent inconsistency between the currency area and the monetary authority within the ruble zone as early as in the fall of 1991. The radical Russian reformers also received general advice on the two possibilities to escape from the highly inflationary ruble zone and at the same time solve the underlying problem, that is, how to support market-based trade among the FSRs. The fast option was to nationalize the ruble, that is, to turn the Soviet ruble into a Russian ruble, and push for early currency reforms in the other FSRs. Multilateral settlements could then be secured by an early

---

<sup>1</sup> Several former Soviet republics announced that they would subsequently issue their own currencies, but only the Baltic countries took decisive actions during the first half of 1992.

<sup>2</sup> Milton Friedman (1992, p. 242) has stated that "The key feature of a unified currency area is that it has at most one central bank with the power to create money—'at most' because no central bank is needed with a pure commodity currency."

<sup>3</sup> There are several cases of collapsing monetary unions in the 20<sup>th</sup> century. See Dornbusch (1992), Friedman (1992), and Bordo and Jonung (1997).

move to current account convertibility, or by setting up a clearing or payments union. The other possibility was to merge those independent central banks that remained in the ruble zone into one monetary authority in which Russia would dominate because of the size of its economy. In practice, however, the latter option was politically unrealistic from the outset because of mistrust among the FSRs. The same argument applied against forming a clearing or payments union.

It therefore appears as if an early nationalization of the ruble, followed by the introduction of current account convertibility, was an evident choice. But the key decision makers in the Russian government initially held a pessimistic view about the possibility of successfully undertaking a rapid nationalization, and therefore opted for maintaining the currency union for the time being. A partial explanation for this pessimism is that they did not receive concrete advice on how the break-up could be accomplished without threatening trade links until May 1, 1992.<sup>4</sup> In fact, the ruble zone was supported by a number of independent economists and international organizations, notably the IMF.<sup>5</sup>

The delay in reaching insight on the necessity and possibility of currency reform among Russian politicians, their advisors, and international financial institutions is, in hindsight, remarkable. The delay in delimiting the Russian currency area, once the Russian reformers had come to understand that this was the only way to escape monetary chaos, is even more remarkable, particularly considering the huge costs that were incurred because of the delay.

This essay seeks to explain the delay in dissolving the ruble zone and creating viable conditions for market-based trade between the FSRs. Several authors have concluded that resistance from interest groups, in particular state enterprise managers, was essential in explaining the delay.<sup>6</sup> We will investigate why interest groups were in favor of maintaining the ruble zone, and how they could have such a great impact on policies.

The strategy used by the reformers to overcome resistance against a break-up is another topic of this essay. In previous research, policies concerning the ruble zone are described as a failure. According to Åslund (1995, p. 135), "At first glance, the dissolution of the ruble zone might appear to have been fast. However, the statistics shatter any illusion of success." With this logic, only a radical dissolution in late 1991 or early 1992 could have been considered a success. But this did not come about, which Åslund attributes to Yegor Gaidar's overestimating the difficulties involved in an early dissolution (*ibid.*, p. 114).

In this essay, decisions on the ruble zone will be evaluated against the situation that was actually at hand, not against a blueprint of how reforms should ideally be undertaken. This approach is in contrast to that of Åslund (1995). He seems to base his conclusion that the slow pace of the dissolution was a failure on a presumption that even very radical

---

<sup>4</sup> The advice was later published in Sachs and Lipton (1993).

<sup>5</sup> A number of proposals were made on how to make the currency union function properly since the former Soviet Union, or most of it, was perceived as an optimal currency area. See Goldberg et al. (1994).

<sup>6</sup> These include Goldberg et al. (1994), Åslund (1995), Bornefalk (1995), Conway (1995), Dabrowski (1995), and Granville (1995).

reforms can be undertaken as part of a major reform program if this is launched at the right moment of time, that is, after a democratic break-through. The condition would be either that vested interests are taken by surprise so that they cannot offer effective resistance, or that reforms are designed in such a way that interests of powerful groups are divided rather than united. The explanation for the delay of the dissolution would then be the mounting influence of vested interests once the government had failed to take advantage of the "window of opportunity" that supposedly opened in late 1991.

This author does not dispute that a rapid nationalization and consecutive move towards convertibility could have been undertaken as part of a comprehensive, well-prepared reform program launched immediately after a major democratic break-through. The point is, however, that it cannot be taken for granted that the available reform space was large enough to carry out radical and encompassing economic reforms. In fact, Åslund (1995) argues that the potential support for economic reforms suffered from a failure of president Jeltsin to dissolve the Russian parliament, which was inherited from Soviet times, and to call for democratic elections, thereby securing a real democratic break-through. In this way, the failure of Russian currency policies can be derived from the failure to democratize Russia from above. But this argument simply assumes that a democratic break-through could have been achieved from above. This cannot be taken for granted either. A third ambition of this essay is to determine under what conditions it would have been possible to undertake a radical and early dissolution.

Before we can go ahead with the analysis of decisions concerning the ruble zone, we need to understand what membership meant to different countries, and what the alternative monetary arrangements were. Theories on payments systems and currency unions will be used to identify main ways of creating an effective payments system. In doing this, political restrictions will be kept in mind. The most important one was the insufficient political will to secure proper monetary coordination within a unified monetary area. As pointed out by Goldberg et al. (1994), this made the traditional optimal currency area approach towards monetary unions inapplicable to decisions concerning the ruble zone. Instead, each potential member country had to take into account a number of unorthodox aspects when making decisions on membership and behavior in the ruble zone. Of these, previous analyses have focused on financial aspects, in particular access to trade credits on favorable conditions. We extend the analysis to include differences in political institutions and in the relative size of the economies of member states.

In our approach, we adopt a theory on democratization, rent seeking, and economic transition developed in Bornefalk (2000a,b). This approach determines the available political and economic reform space in a situation where an authoritarian state has collapsed or withered away, and vested interests have secured a dominant position. This is achieved by deriving the extent to which rent seeking is undertaken as well as the resulting economic performance and the distribution of wealth under different conditions. The determinants include the extent to which political freedoms exist, the maturity of political organizations, the level of development reached by mass media, the share of



resources that is controlled by vested interests, and the ability of different interest groups to cooperate with each other.

The various factors found to be of importance for decisions concerning the ruble zone are studied in a game theoretic analysis to understand why most non-Russian FSRs were keen on remaining in the ruble zone. The game theoretic analysis also facilitates an inquiry into how the currency regime could generate such exorbitant costs in such a short time.

The starting point of the game theoretic analysis is a simple game illustrating basic aspects of the unrestricted ruble zone, that is, the mechanisms that were in place during the first half of 1992. The solution of the game indicates that we cannot explain why the ruble zone was set up and maintained during this period without resorting to explanations based on insufficient competence and/or rent-seeking. We set up two hypotheses. The first is that the decision makers neither realized the extent of the threat imposed by the ruble zone upon Russia, nor understood that a nationalization of the Russian ruble reducing available rents and inflationary pressures could have been undertaken much faster than it was without worsening the conditions for trade. The second hypothesis is that the decision makers had enough knowledge and information, but faced political restrictions which prevented them from a rapid dissolution of the ruble zone.<sup>7</sup> The role of vested interests is central to the second hypothesis.

The second, more elaborate, game analyzes the mechanisms of the restricted ruble zone in place from mid 1992 until mid 1993.<sup>8</sup> The elaborate game incorporates the political uncertainty that a weakening of the position of the Russian reformers led to. The elaborate game helps us to interpret the key decision in the gradual dissolution process and to understand why hyperinflation emerged in those non-Russian member countries that did not leave the ruble zone voluntarily at an early stage.

The games turn out to be quite simple to solve. Both solutions involve great losses to the country setting up and, in the case of the more elaborate game, dissolving the currency union only gradually. The simplicity allows us to conclude that rational, competent, and well-informed decision makers facing situations of the types illustrated by the games would be able to foresee the exorbitant costs that such mechanisms would generate. If they aimed at maximizing social welfare, they would also realize that such a currency union should not be set up, or, in the case analyzed in the elaborate game, should immediately be dissolved, if possible.

We conclude that vested interests consisting of different industrial groups resisted a dissolution of the ruble zone since political and economic conditions were such that rent seeking was relatively more profitable than alternative activities. They therefore preferred

---

<sup>7</sup> An alternative point of view is that decision makers were maximizing something other than the well-being of society as a whole. We will, however, assume that the radical Russian reformers were indeed attempting to maximize social welfare.

<sup>8</sup> The game on the unrestricted ruble zone is a development of the one-period game in Bornefalk (1995, pp. 108-112), while the game on the restricted ruble zone is essentially new. Nevertheless, some of its aspects were discussed in the repeated game of Bornefalk (1995, p. 112).

a currency regime that offered them ample opportunities to acquire rents to one that allowed them to make payments effectively. It is also shown that the Russian reformers utilized the available reform space quite well. They eventually succeeded in dissolving the ruble zone by implementing a strategy based on furthering the democratic development and dividing the interests of different industrial groups, thereby making production relatively more profitable than rent seeking. This, in turn, increased the demand for an effective payments system and lowered tolerance for high inflation. Again referring to basic political and economic conditions, it is finally concluded that a radical dissolution might very well have been possible to undertake if it had been part of a major reform program including both a swift democratization and radical economic reforms. But during the short time period that such a program could have been implemented, it simply did not exist.

The essay proceeds as follows. Chapter 2 contains a discussion of trade and payments in the FSU. The main alternatives for a new payments system are evaluated, and an account of macroeconomic consequences of maintaining the ruble zone is given. The mechanisms of the ruble zone are examined in Chapter 3. In Chapter 4, we study the strategic interaction between member states. The results serve as a basis for the discussion of policies of different FSRs in Chapter 5. Chapter 6 analyzes and evaluates Russia's monetary reform, and Chapter 7 concludes.

## **2. TRADE AND PAYMENTS IN THE FORMER SOVIET UNION**

### **2.1 The Breakdown of the Soviet Payments System**

At the time of the dissolution of the Soviet Union, the ruble could not fulfill the most elementary functions of a currency because of a huge monetary overhang that had developed mainly in 1990 and 1991.<sup>9</sup> In the same period, the hard currency reserves had almost been depleted, which added to the difficulties in trade. Moreover, the payments system for trade within the Soviet Union was highly inadequate under market conditions. In order to secure viable conditions for trade between the FSRs, it was necessary to end the monetary and financial chaos and to create a payments system with a mechanism that could handle decentralized payments effectively and that had sufficient liquidity.

In order to reform an economic system, it is important to understand its main features as well as its major flaws. A crucial difference between a Soviet-type economy and a market economy lies in the role of money and the price mechanism.<sup>10</sup> To start with the banking system, Gosbank managed the financial system of the entire Soviet Union, but lacked even nominal autonomy since it was subordinate to the government. Gosbank served as the central bank and also controlled the various specialized banks, which included an investment bank, a foreign trade bank, and a savings bank for the public. In addition to being responsible for the emission of money, Gosbank supplied the entire state sector with credits.

Each firm kept an account at Gosbank. All money that the firm “owned” had to be deposited at that account with the exception of a controlled amount of cash, whose use was determined by the planners. The firm could not dispose freely of money on its account. Instead, the money was drawn upon to cover expenditures that had been decided by the central planners. On the other hand, when an enterprise was entitled to make a transaction, it was guaranteed necessary funds. Furthermore, money was “earmarked.” There were different subaccounts for different purposes, and money could not move freely between them. In other words, the ruble was not convertible within the country, which added to the rigidity of economic activity. The same was true to an even higher degree for conversion between the ruble and foreign currencies. For these reasons, Kornai (1992, p. 133) concludes that “money fail(ed) to perform the integration of all transactions; it (was) not actually a universal means of exchange.” With this in mind, it can be said that the role of money in the enterprise sphere was passive: it supported the planning process.

The role of prices was also more limited than in a market economy. Prices were

---

<sup>9</sup> The huge monetary overhang developed since the government was too weak to increase prices in line with the rapidly increasing money supply, let alone to liberalize them. This led to massive shortages of goods. As noted by Lipton and Sachs (1992, p. 220), the money supply increased for several reasons. Two of them were that the immense budget deficits were financed almost exclusively by money issue and that real wages climbed to artificially high levels since nominal wage increases were much larger than price increases during the last three years of communism.

<sup>10</sup> The following is based on Kornai (1992, p. 131 ff.).

determined centrally and fixed for long periods to provide stability and to ease planning by making aggregation possible. Hence, they could not carry information about relative scarcity and accordingly they did not give proper incentives. The way payments were made and credits issued did not hinge on the specific Soviet republic in which a particular enterprise was located. Since the role of money was passive, trade deficits between the Soviet republics—made arbitrarily small by artificial pricing—were more or less automatically covered by credits, which were regularly written off.

The payments system had a further peculiarity that is worth mentioning since it was inherited by the payments system used in the ruble zone: there were two types of ruble. Account rubles were used for payments between enterprises and between enterprises and the state, while cash rubles were used for retail purchases and wages. These two types were not internally convertible, just as account rubles on different accounts were not convertible. As noted by Sachs and Lipton (1993, p. 136), this meant that cash was not legal tender for all transactions, and that the banking system did not fulfill the fundamental task of allowing enterprises to withdraw deposits in the form of cash. For households, it was not always possible to convert deposits in the state savings bank into cash because of repeated shortages of cash.<sup>11</sup>

With the peculiarities of the Soviet monetary system in mind, the conclusion is obvious: the heavily centralized payments system based on quasi-money could not serve a market economy. The old payments system was especially inadequate considering the massive changes in production and trade that became necessary with the change of economic system. The central planning system of the Soviet Union created an industrial sector characterized by excessive concentration and extreme vertical integration. This had little to do with economies of scale, but facilitated central planning and commands. The location of plants was determined with little or no regard for transportation costs. Thus, goods were hauled large distances throughout the Soviet Union.

Goods and payments had to pass the borders between the newly independent states after the Soviet Union was dissolved. Furthermore, most non-Russian FSRs had to adjust to drastically falling terms of trade. The reason for this was that in the Soviet era, Russia subsidized the other Soviet republics heavily by exporting energy and raw materials at prices far below world market levels and by importing manufactured goods which could often not have been sold at all on the world market. This subsidization would eventually have to cease as the countries moved to trading under market conditions. Thus, since the systemic change led to changes in the structure of demand, massive structural changes were needed throughout the FSU. For these to take place, it was necessary to create a payments system that could handle payments between, and within, the successor states of the Soviet Union effectively.

At the time of the break-up of the Soviet Union, there was even a fear of total collapse of trade between its successor states because of the lack of an effective payments

---

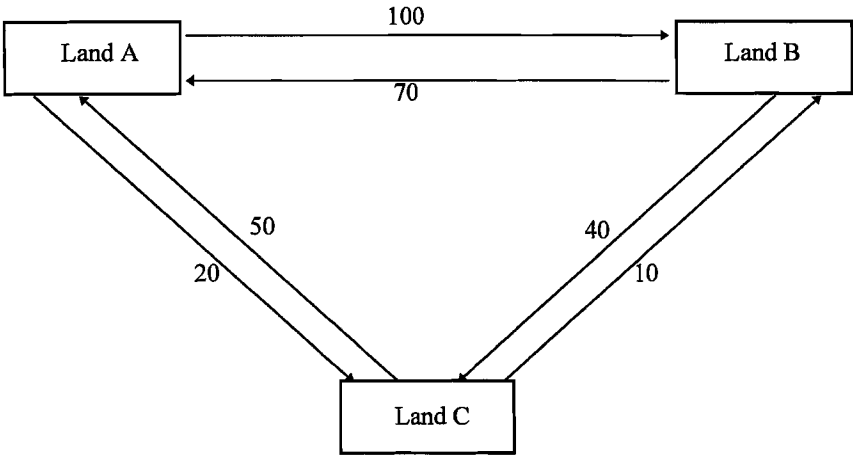
<sup>11</sup> While money was “passive” in the enterprise sphere, it was “active” in the household sector. In other words, the amount of money that the household possessed determined which transactions could potentially be

mechanism and sufficient liquidity. A collapse of trade would in turn severely hurt production, and this for two reasons. The first was that the entire production system was technologically inflexible. Second, the planners' *gigantomania* made it vulnerable—many over-sized plants depended on a single supplier for some of their raw materials or intermediate inputs and also had one dominant customer. Failure in one part of the system could therefore lead to serious supply problems, as noted by Leijonhufvud (1993).

Immediately after the change of economic system, supply shortfalls were bound to lead to substantial decline in production because of structural changes. There was a clear risk that enterprises with prospects of surviving in a market environment could be severely hurt if trade were interrupted. But rather than save the inherited trade, the task of the reform government was to create conditions in which trade on a market basis could develop between enterprises in different FSRs. Therefore, a payments system that could handle market transactions was called for in early 1992.

### 2.2 Options for a Market-Based Payments System

Two fundamental requirements for the development of market-based trade between the FSRs were that prices be freed and that agreements allowing for either free trade or trade on, for instance, a most favored nation basis be signed. A basic requirement for a payments system to handle payments effectively, that is, to provide viable conditions for market-based trade, is that it should allow multilateral settlements. The importance of multilateral settlements is illustrated in Figure 1.<sup>12</sup>



**Figure 1: The Importance of Multilateral Settlement**

Each arrow depicts an export flow. Land A is in surplus with Land B and in deficit

---

made.

<sup>12</sup> This example is due to Eichengreen (1993).

with Land C. If convertible currencies exist or multilateral clearing is allowed, each country uses its surplus to finance its deficit, and trade balances multilaterally. The total value of exports is 290. If, on the other hand, trade has to balance bilaterally, Land C, for instance, earning only the equivalent of 10 of the currency of Land B, can import only the worth of 10 of goods from Land B, and the total value of exports will have fallen to 200. The larger the initial bilateral surpluses and deficits, the greater the fall in trade when multilateral settlements are absent. Since bilateral balances between the FSRs were extremely large, the absence of multilateralism would curtail trade severely.

There are several alternative ways to make multilateral settlements possible. There are, first, the “primitive” methods of barter or payments in hard currencies. Another possibility, which is related to payments in hard currencies, would be to use one of the national currencies for trade. Then, there are the more sophisticated methods of payments through a clearing or payments union or with convertible currencies. Finally, a way to get around the problem is to form a currency union. In the case of the FSU, this initially meant the continued use of the Soviet ruble. This option is treated in the following section.

There are advantages and disadvantages with all options, and they can be judged on several criteria. The first is how effective they are in settling payments. The second is how difficult and costly the set-up and maintenance are. The third is how well they work in an inflationary environment. Further, they should not threaten monetary and fiscal discipline in the countries using the payments system, and they should facilitate correction of shocks that hit individual countries.

#### A. “Primitive” Technologies

For a payments system to be effective, it must have a functioning payments mechanism and sufficient liquidity. Barter performs poorly by both these criteria and is therefore plagued by great transaction costs. On the other hand, barter can always be employed since it requires no particular state institutions, and provides a kind of indexation in an inflationary environment. In general, it is a common choice if there is little trust in national currencies.<sup>13</sup>

Payments in hard currencies are superior to barter if there is sufficient liquidity. But there is often insufficient liquidity if the national currencies are not convertible. If they are convertible, there is no need for settlements in hard currencies, especially since it is costly for the states: seigniorage is handed over to the country printing the hard currency in use. Like barter, payments in hard currencies provide a safeguard against uncertainty brought about by high inflation. Payments in one of the national currencies could be superior to using hard currencies if this provides better liquidity and if inflation is not too high. This method might also make it easier to utilize the banking system.

---

<sup>13</sup> See Ellingsen (1998) and Åslund (1995, p. 113) for assessments of the pros and cons of barter.

### *B. A Clearing or Payments Union*

In a reasonably stable macroeconomic environment, it is possible to use more sophisticated payments systems. In a clearing union, the clearing of inconvertible currencies makes multilateral settlements possible. Transactions in a certain accounting period are expressed in an agreed unit of account, preferably a hard currency. After multilateral clearing has taken place, countries in deficit pay the amount they owe to the clearing union in hard currency. Countries in surplus are then compensated. A payments union has the additional advantage, compared to a clearing union, of letting its member countries finance temporary deficits within the union by drawing on a credit line. Clearing and payments unions save on the amount of hard currency needed to secure multilateral settlements. Payments unions have at times been used prior to the introduction of convertible currencies. One example is the European Payments Union, which was in place in the 1950s.<sup>14</sup>

A disadvantage with clearing and payments unions is that they are costly to set up and maintain. Multilateral clearing requires multilateral agreements and legal instruments that ensure the payment of credits. A payments union is also vulnerable to high inflation. Furthermore, if after multilateral clearing of bilateral deficits some members are in substantial surplus with the rest—and could be expected to continue to be in surplus—a payments union does not improve the situation substantially: there will still be a need for extensive use of hard currency.

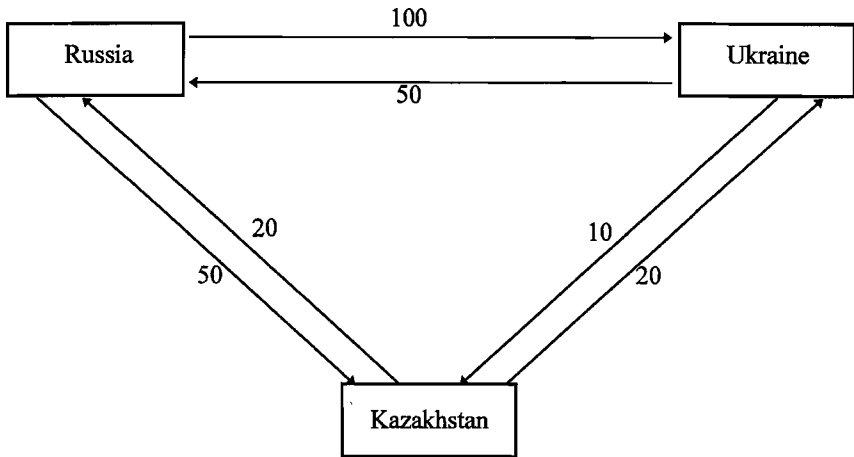
The prospects for credible multilateral agreements between the FSRs were poor, and the legal framework and enforcement was all but absent at the beginning of the systemic change. Already in early 1991, tensions among the Soviet republics began to disrupt trade. Threats to withhold deliveries were common and highly effective due to the industrial structure's vulnerability: absent deliveries from one republic could seriously disrupt production in other republics (Leijonhufvud, 1993). The lack of trust among the participating countries meant that not even a clearing union without credits would have been feasible. In the example of Figure 1, Land A would not allow Land B to use its surplus with Land C to finance its deficit with Land A, since Land A would not be certain that Land C would allow Land A to do likewise.

But even if the FSRs had been able to trust each other, a payments union would not have been adequate for the FSRs, as illustrated in Figure 2. In the example of the figure, Russia runs a large surplus with both Ukraine and Kazakhstan. Kazakhstan's surplus with Ukraine is too small to offset its deficit against Russia even if it could be collected. The chances for it to be collected are small since Ukraine runs a deficit with both countries. Thus, only bilateral clearing is really feasible. This situation resembles the one in the FSU in 1992, when Russia ran substantial surpluses with almost all other FSRs. The lack of trust and the unbalanced nature of trade meant that there was little concrete interest in a

---

<sup>14</sup> For more on the mechanisms of clearing and payments unions, the European Payments Union, and an assessment of the plausibility of a payments union for the FSU, see Eichengreen (1993).

payments union.<sup>15</sup>



**Figure 2: The Limits of Multilateral Clearing**

### *C. Convertible Currencies*

The poor prospects for clearing arrangements leaves us with payments in convertible currencies of the FSRs as the only viable sophisticated payments system. This was the main alternative to continued use of the Soviet ruble except for “primitive” methods. Hansson (1993, p. 164 ff) gives three broad advantages of introducing national currencies as compared with staying in the ruble zone.<sup>16</sup> The first is that a national currency is a way to enhance national sovereignty. It is not only an important symbol for political independence. It is also a way to strengthen economic sovereignty in that it secures national monetary authority since the respective governments acquire the basic macroeconomic policy instruments. In particular, monetary independence would help each state to choose its speed of reforms.

A second reason to introduce a national currency is that it facilitates economic stabilization since the importance of fiscal and monetary policies is made clear to each country. Countries that are able to limit their budget deficits and credits to state enterprises sufficiently will benefit from low inflation. Having a national currency also makes it easier to adjust to terms-of-trade shocks since devaluations are possible. This, in turn, makes it easier to acquire a sustainable balance of payments, which is another important factor in economic stabilization. Since terms-of-trade shocks varied greatly across the FSRs, particularly during 1992, an early move towards independent currencies

<sup>15</sup> Numerous suggestions were, however, presented by western economists arguing for various sophisticated arrangements. One example is Dornbusch (1993). These specialists in monetary economics failed to acknowledge, or draw the right conclusions from, the political and economic situation in the FSU.

<sup>16</sup> For a general account of how to introduce a national currency, see Hansson (1993, p. 172 ff). For a description of the introduction of the Estonian kroon, see Hansson (1993, p. 175 ff).



would have been most useful.

The third reason mentioned by Hansson is that a national currency can promote structural adjustment, particularly for small countries. The most important reason for this has to do with currency convertibility. As already argued, a national currency facilitates stabilization, and stabilization enhances the possibility of making currencies convertible. Convertibility, in turn, results in a close integration with the world economy since prices reflecting relative scarcity and comparative advantages are literally imported. This increases competition on the domestic market and gives incentives for restructuring in accordance with the comparative advantages of the country.

Convertible currencies were clearly the best basis for a functioning payments system between the FSRs and with the rest of the world. With convertible currencies, liquidity had been secured, and the problem with lack of trust between countries could have been overcome by direct dealing between enterprises in different countries. But few believed that convertibility could be achieved at an early stage. Indeed, many even doubted that a rapid introduction of national currencies was feasible. In particular, while the radical Russian reformers were in favor of nationalizing the ruble as soon as possible, their leader Yegor Gaidar believed that the technical preparations for a currency reform would take at least nine months.<sup>17</sup>

According to some western policy advisers, however, a nationalization of currencies in the FSU could have been accomplished in a matter of weeks. This was done in Czechoslovakia, one of the successor states of the Austro-Hungarian Empire after World War I.<sup>18</sup> By separating its currency from the Austro-Hungarian crown at an early stage, Czechoslovakia succeeded in avoiding the monetary disarray that slowed down the economic recovery of the other successor states. The currency reform was undertaken by simply stamping the notes while the borders were closed for two weeks in early 1919. Because of the nationalization of the currency, Czechoslovakia could insulate itself from the raging hyperinflation in other successor states through conservative fiscal and monetary policies. The national currency was made convertible from the moment of its introduction, which helped integrate Czechoslovakia in the world economy.

After a separation of currency areas had been achieved, the FSRs could have accomplished convertibility simply by discontinuing the state's role in the respective foreign exchange markets, as pointed out by Sachs (1993, p. 52). That would have resulted in a unified market price for foreign exchange. Reasonably stable exchange rates could then have been achieved by implementing comprehensive stabilization programs. To sum up, we can conclude that in theory, there were several possible alternatives to secure multilateral settlements between the FSRs, but in reality, the choice stood between a rapid introduction of independent national convertible currencies and a continued use of the Soviet ruble. The former alternative was clearly preferable, but that was not the path

---

<sup>17</sup> Gaidar (1993, p. 84), cited in Åslund (1995, p. 114).

<sup>18</sup> Åslund (1995, p. 115) attributes this idea to Rudiger Dornbusch, who launched it at a seminar in Moscow in September 1991. See Dornbusch (1992). Jeffrey Sachs also promoted this idea, referring to Sargent (1986, pp. 95-97) and Pasvol'sky (1928, pp. 204-31).

Russia chose for the FSU.

### 2.3 Russia's Choice

Russia tried to keep up trade by creating a currency union with the other FSRs instead of nationalizing the ruble and urging other FSRs to introduce convertible currencies.<sup>19</sup> A currency union requires much more trust and cooperation than a multilateral payments arrangement, but the ruble zone was created without any serious attempts at coordination of monetary and fiscal policies. Enterprises continued the frequent use of barter and payments in hard currencies as a necessary complement to the poorly functioning currency union. These methods had become widespread during 1991 due to the monetary chaos.

The birth of the currency union can be traced back to the coup attempt against Mikhail Gorbachev in August 1991. The weakening of the Soviet government that followed made it possible for the CBR, then headed by Georgy Matyukhin, to take over the responsibilities of Gosbank. These included management of the payments clearing mechanism. Russia welcomed the continued use by the other FSRs of the ruble as well as the payments clearing mechanism when it initiated the change of economic system on January 2, 1992. They chose to do so. At that stage, all other FSRs had established their own central banks with unlimited rights to issue ruble credits by transforming the republican Gosbank branches, and Gosbank had been dissolved.

To adapt the Soviet payments clearing mechanism to suit the ruble zone, the CBR opened correspondent accounts for each non-Russian FSR's central bank in January 1992. Purchases from Russia by any other FSR drew down that republic's account, and its balance increased with each sale to Russia. In the beginning of 1992, there was no matrix of intra-FSU accounts, only bilateral correspondent accounts between Russia and each other FSR (Eichengreen, 1993, p. 316). Later, since virtually all non-Russian FSRs continually exhausted their credit limits for reasons developed below, there were no surpluses in the accounts with the CBR to finance purchases from non-Russian FSRs. When any republic occasionally ran a surplus with Russia, Russia did not allow it to finance purchases from another FSR because of lack of trust. In this way, the payments system's inflationary bias led to insufficient liquidity, which made it ineffective.

The ineffectiveness of the payments mechanism of the ruble zone contrasts with the quite large microeconomic efficiency gains from establishing a currency union which Conway (1993) asserts that the FSRs predicted. These would have included the absence of currency conversion costs as well as greater price stability and thereby less uncertainty in trade. A currency union is, however, inappropriate if member countries face shocks of different magnitudes, unless the shocks can be offset either by sufficiently flexible markets, or by financial transfers between member states. The terms-of-trade shocks of

---

<sup>19</sup> Note that the Soviet Union was not a currency union for the simple reasons that the Soviet ruble was not a currency, and the Soviet republics were not independent states.

moving towards world market prices would hit many FSRs hard, while Russia stood to gain. Thus, a currency union would not have been appropriate or even sustainable unless Russia intended to keep supplying the other FSRs with energy priced at a few percentage points of world market prices. It is hard to find any reason why Russia should provide such massive aid to the other FSRs, especially considering its own dire financial situation, but in fact it did. Furthermore, a monetary union necessitates coordination of monetary and fiscal policies. Otherwise, member countries would try to export inflation and unemployment to each other, spreading instability throughout the currency union. But the institutions and political will required for such coordination were absent.

Thus the prospects for the currency union to hold were very poor, as was foreseen by a number of Russian radical reformers. But then, it cannot have been a careful calculation of potential microeconomic efficiency gains and possible macroeconomic costs because of the flawed institutional framework, or for that matter a political will to stay united, that lay behind the ruble zone. Instead, we can find two main reasons for the decision to continue using the Soviet ruble throughout the FSU, at least during the initial phase of reforms. The first reason is that the ruble and the foundations for the payments clearing mechanism for transactions between the FSRs, and thus a payments system, were already in place. As regards an early separation of currency areas, on the other hand, there was no concrete proposal on how this could be achieved without threatening trade links. The second reason has to do with the political economy of currency reform and the rather weak position of the Russian government.

The introduction of national currencies must have appeared too complicated to be accomplished quickly considering that only a few of the FSRs were planning serious reforms in the fall of 1991. Since Yegor Gaidar thought that the preparations for a currency reform in Russia would take at least nine months, he should have perceived that this would be too short a time for most other FSRs, notably those that lacked or were just starting to develop state institutions and had vague or non-existent reform programs. There was a clear risk that a currency reform of the type which Czechoslovakia undertook following the break-up of the Austro-Hungarian Empire would have exacerbated the already chaotic situation in trade between the FSRs, at least during a limited period.

Opponents to the reformers could then have used the drastic fall in trade and the massive supply shocks as an argument to reverse the dissolution of the ruble zone and the change of economic system as a whole. The extent to which the fall in trade resulted from the currency reform as such would, arguably, have been of little importance. The reformers would have had to take the blame also for the part of the reduction in trade that had to do with the move towards market-based relations and trade based on comparative advantages. By maintaining the key ingredients of the Soviet payments system, they hoped to secure continued payments and thereby avoid strong critique from the managers of large state owned enterprises, the group they had most to fear from. We return to these issues in Chapter 6.

Thus the Russian government postponed decisions on monetary reform as it pursued

a far-reaching liberalization, attempted to balance the budget, and prepared a radical privatization program. As we will see in the next section, the attempt to keep transaction costs in trade down by maintaining the ruble zone led to huge costs in terms of inflation and trade credits. Nevertheless, the advantages in the payments area failed to materialize.

## 2.4 The Ruble Rally

In the first half of 1992, the central banks of the FSRs were able to determine their own rates of domestic credit expansion although they remained in the ruble zone. Most non-Russian member states issued account rubles with little restraint and used them to subsidize state-owned enterprises and to finance government budget deficits and trade deficits with Russia. According to Conway (1993), the issuing of credits was less coordinated than had been agreed at meetings of ruble-zone central bankers. In fact, the agreements were more or less neglected. It was instead the different levels of subsidies to state-owned enterprises and budget and trade deficits that determined the amount of credit creation in different countries. Figures on credit creation and budget deficits during 1992 are not available for all FSRs, and trade statistics are of notoriously poor quality. Fortunately, we can get a fairly good picture of the “ruble rally” between the FSRs by looking at bilateral balances with Russia on the correspondent accounts of the CBR. But let us first take a look at its consequences in terms of inflation.

Average monthly inflation rates during 1991-1993 are shown in Table 1 for those FSRs that were not plagued by war during this period. As the Soviet republics turned republican branches of Gosbank into national banks and gained increasing control over taxation and expenditures during 1991, monetary and fiscal policies started to diverge within the common monetary area. However, since all countries shared the same currency, less reform-minded governments were virtually forced to liberalize prices after Russia had done so on January 2, 1992, although the extent of liberalization differed among countries. Countries that chose to liberalize less had to impose more restrictions on trade with other FSRs, and incurred larger budget deficits, *ceteris paribus*. The price liberalizations eliminated the monetary overhang that had accumulated in previous years and caused a jump in inflation levels between 1991 and 1992.

Inflation rates were more or less the same across countries until mid-1992 despite the divergence of monetary and fiscal policies, which increased sharply in 1992. The explanation for this is that inflation was shared among the Soviet republics in 1991 and continued being shared through the ruble zone mechanism after the Soviet Union had been dissolved. In the latter part of 1992, and even more so in 1993, inflation rates came to differ increasingly as the number of member countries in the ruble zone fell due to voluntary currency reforms or exclusions. As the sharing mechanism gradually disappeared, countries had to take on more of the inflationary burden of their policies. Countries that executed successful currency reforms, notably Estonia, Latvia and Lithuania, managed to insulate themselves from the inflationary ruble zone and benefited

from lower inflation rates.<sup>20</sup> Countries that were excluded after having issued excessive credits were thrown into hyperinflation, with Ukraine being the prime case.

**Table 1. Average Monthly Inflation Rates<sup>1</sup> in Selected FSRs in 1991-1993**

	1991	1992	1993
Russia	8	30	21
Ukraine	8	32	45
Belarus	8	28	32
Moldova	8	30	26
Kazakhstan	8	31	29
Kyrgyzstan	9	24	25
Turkmenistan	8	18	37
Uzbekistan	9	20	21
Estonia	12	22	5
Latvia	11	23	3
Lithuania	12	23	9

Source: IMF (1994:2, p. 34).

<sup>1</sup> Average monthly percentage change in CPI or retail prices

The excessive credit creation also led to a shortage of cash, which was especially severe in the non-Russian member states since Russia gave priority to its own economy once the shortages appeared. The shortages caused Ukraine to introduce a coupon as a ruble substitute on January 10, 1992. This was Ukraine's first step towards the introduction of a national currency, following the declaration in the fall of 1991 that this would be done for primarily nationalistic reasons. Several other member states followed Ukraine's initiative in the spring of 1992. They had thereby created another tool for pursuing soft monetary policies within the currency union.

Since inflation was shared among the member countries while monetary and fiscal policies differed, some countries gained at the expense of others. Those who succeeded in accumulating the largest deficits on the correspondent accounts of the CBR can be declared winners in the ruble rally. They managed to grasp a larger share of the shrinking common cake by running huge trade deficits with Russia, paying with credits extended by the national central bank. These non-Russian credits were covered by automatic Russian trade credits in the first half of 1992. The automatic credits amounted to Rbs 316 bn, or about 5 percent of Russia's GDP for the first half of 1992 (Granville and Lushin, 1993).

An attempt by Russia to limit its trade surplus and crediting of other member states

<sup>20</sup> The Estonian kroon was introduced on June 20, 1992, while Latvia and Lithuania introduced coupon currencies on July 20 and October 1, 1992, respectively.

from July 1992 by restricting the amount of available credits on the correspondent accounts made trade credits no longer automatic. Still, it did not lower the amount of credits as a percentage of GDP, as can be seen in Table 2. Credits to other FSRs in 1992 exceeded 1,5 billion rubles. This corresponded to 8.5 percent of Russia's GDP. These automatic (for the first half of 1992) and non-automatic trade credits amounted to as much as 25 percent of total CBR credits in 1992 (Sachs, 1994).

**Table 2. Financing of Selected FSRs by the Central Bank of Russia, 1992**

	End-1992 CBR correspondent account position		End-1992 CBR position adjusted to exclude currency	
	In billions of rubles	In percent of GDP	In billions of rubles	In percent of GDP
Russia	-2,109	-11,7	-1,545	-8,5
Ukraine	862	21,7	862	21,7
Belarus	102	10,7	69	7,2
Moldova	27	11,3	18	7,5
Kazakhstan	407	25,5	235	14,8
Kyrgyzstan	42	22,9	20	10,9
Turkmenistan	172	53,3	111	34,4
Uzbekistan	292	69,9	117	28,1
Estonia	4	4,0	4	4,0
Latvia	2	1,0	2	1,0
Lithuania	9	3,2	9	3,2

Source: IMF (1994:1, p. 26).

If deliveries of ruble notes are included in Russia's financing of other FSRs, the deficits reached as much as 11,7 percent of Russia's GDP in 1992. Cash took on a growing importance in the financing of trade after the limits on technical credits had been imposed in mid-1992. The reason for this was, first, as mentioned by IMF (1994:1, p. 6), that the credit limits generated discounts on account rubles in non-Russian FSRs. The discounts were larger in countries with greater excess demand for Russian financing of trade. In other words, the credit limits gave rise to a *de facto* separation of account rubles. Ruble bank notes were not, however, subject to discounts since the same notes continued to be used in the entire ruble area. Hence, cash became relatively more effective for transactions. Another, and probably more important, explanation is that, beginning in the third quarter of 1992, the CBR under the new head Viktor Gerashchenko generally provided other member countries with as many bank notes as they demanded, provided

that the printing presses could print enough. These bank notes were used to finance trade deficits. Countries that had issued currency substitutes could also relax the constraints on the amount of ruble bank notes that could be used to finance trade by printing more coupons for the internal market.

## **2.5 Inertia in Trade Relations**

The cause of the rampant inflation in the ruble zone and Russia's heavy subsidization of other members was the nature of the ruble zone, as will be explained in the following chapters. In this section, it will be shown how the deficits materialized. Structural reasons as to how deficits could become so large in such a short time period will also be discussed.

The deficits in ruble-area trade followed on from the traditionally great imbalances in trade among the Soviet republics. As we have seen, outstanding features of Soviet-type trade were that Russia supplied the members of its empire with cheap energy and raw materials, and provided them with a market for low-quality manufactured goods. The imbalances were hidden by artificial prices and automatic financing of deficits as long as the central planning system remained in place. But as soon as prices were allowed to approach world market levels, the imbalances increased sharply as energy exporters, notably Russia, drastically improved their terms of trade.

The possibility for member states to finance trade deficits by extending credits to domestic producers, so that these in turn could pay their suppliers in Russia through the correspondent accounts, aggravated the already unbalanced nature of trade. The reason was that enterprises in non-Russian member states other than the Baltic countries maintained their soft budget constraints and therefore did not have to change their behavior. Also, enterprises in Russia and other parts of the ruble zone continued to deliver their products without reasonable guarantees of payment, that is, to treat money merely as a unit of account.<sup>21</sup> Therefore, the delay in settling payments, referred to as payments arrears, became a major problem in the spring of 1992.

The behavior of Russian enterprises helped the soft budget constraints of non-Russian enterprises to cement trade flows. Thus, imbalances in the correspondent accounts as well as arrears kept growing. However, the arrears problem was less pronounced in trade with other FSRs than inside Russia, since enterprises in other non-Baltic FSRs had softer budget constraints because their governments pursued softer monetary and fiscal policies. This made Russian enterprises keen on exporting to other FSRs. Those who faced difficulties in finding buyers were especially keen. One clear example is the producers of agricultural machinery. On the other hand, Russian enterprises frequently did not pay for deliveries from other FSRs, partly because of their comparatively hard budget constraints.

The unbalanced trade meant that Russia continued to provide the other member

---

<sup>21</sup> The motives behind doing so have been studied by Boeva and Dolgopiatova (1994).

states with energy and natural resources, generally at prices well below world market levels, and frequently received payment only in inflationary credits. Thus Russia had a large positive trade balance with most other member states, despite the low prices charged for exports. But, according to Lushin and Sarafanov (1993), Russia was not discriminated against in trade with the other CIS countries through unfair price setting. Using trade data including 73 percent of Russian exports to CIS countries according to intergovernmental agreements, and most of the imports, the authors have estimated that prices of Russian exports were around 12 percent of the world level and about 11 percent for imports during 1992.<sup>22</sup> This assessment is, however, uncertain, since it is difficult to take into account that Russia's exports were more marketable than its imports.

The explanation for Russia's positive trade balance is instead, as argued by Lushin and Sarafanov (*ibid.*), found in the non-fulfillment by the other CIS countries of their obligations according to trade agreements. Therefore, Russia only received parts of the implicit trade subsidies—the difference between the world market prices and the contract prices—intended by the agreements, while it extended the contracted implicit subsidies to its partners. Russia's subsidies for the goods covered by Lushin's and Sarafanov's calculations amounted to USD 24.5 bln in 1992, while it received only USD 12.5 bln in return. According to IMF estimates, the resulting net difference of implicit subsidies, USD 12 bln, corresponds to 13 percent of Russian GDP at the average exchange rate prevailing in 1992.<sup>23</sup> Note that the size of the unintended implicit subsidies was related to the size of trade deficits, which were financed through the correspondent accounts. However, there was no perfect correlation between implicit and explicit subsidies since not all trade was state controlled.

The non-fulfillment of obligations on the part of the non-Russian CIS countries was attributed by Andrei Lushin<sup>24</sup> to the way state trading was undertaken. Roskontrakt, the Russian agency responsible for finding suppliers and customers in Russia for state controlled trade, was more effective in forcing Russian suppliers to deliver than its counterparts in the other FSRs were with their enterprises. A complementary explanation could be that the non-Russian enterprises were often not paid, due to the relatively hard budget constraints of the Russian enterprises, and therefore cut deliveries. This fits well with the huge payments arrears from Russian to non-Russian enterprises. At any rate, enterprises in other FSRs that were able to find other customers on the world market often traded with them instead. One example is the cotton producers in the Central Asian countries. But, as is argued by Åslund (1993), the state trading system as a whole supported the old structures in the other republics and was designed to maintain as much as possible of the previous trade. As Russian enterprises turned more independent, they became less willing to participate in state trading if they had other alternatives. But then they faced difficulties in finding means of payments, and often turned to barter.

---

<sup>22</sup> Although these prices were low, they still meant a big increase compared to inter-republican trade in the Soviet Union.

<sup>23</sup> The estimate is, however, extremely sensitive to the exchange rate used in the calculations.

<sup>24</sup> In an interview on March 24, 1994.



The huge trade subsidies show that it was exorbitantly costly for Russia to maintain the ruble zone and the state trading system. But what did Russia receive in return for its support to the other FSRs? The long delays in settling transactions and the mounting payments arrears show that Russia did not create a functioning payments system. Partly because of that, trade between the FSRs fell dramatically despite Russia's massive trade credits at low nominal interest rates during rapid inflation and uncertain repayment. It is also hard to think of any other benefits, political or other, that could have compensated Russia for the huge costs.

Some politicians did, however, aim for a resurrected Soviet Union and saw the ruble zone as a means of reaching that goal. Others looked upon the ruble zone as a way to secure continued payments of pensions and other support to the approximately 25 million Russians living in non-Russian FSRs at the time of the break-up of the Soviet Union. That could, however, have been solved at a lower cost to society than the one caused by the ruble zone. It can thus be concluded that the ruble zone entailed vast costs but few benefits for Russia. The next step in our analysis of policies concerning the ruble zone is to examine its monetary mechanism. In doing this, we will extend an earlier analysis stressing its "free rider" nature. This will increase our understanding of what membership meant to different countries.

### 3. THE MONETARY MECHANISM OF THE RUBLE ZONE

#### 3.1 Previous Analyses

##### *A. Independent Economists*

Several authors have studied monetary aspects of the ruble zone with emphasis on its free rider problem. Lipton and Sachs (1992, p. 237) provide a starting point by relating to a model in which several independent central banks in a currency union have independent authority to issue credit. They argue that there is no realistic possibility of controlling credit in such a system since each state would like to free ride by issuing credits at the expense of the rest of the system. In contrast to some other researchers, they do not, however, argue that this model describes the situation in the ruble zone adequately.<sup>25</sup> Indeed, Sachs (1994, p.43) leaves the possibility open that Russia could have had incentives to restrain the issue of credits even though it was a member of the ruble zone: "No single central bank outside of Russia had the incentive to maintain tight monetary policies in the face of this collective non-system."

Bornefalk (1995, p. 106) points out factors that made it possible for non-Russian member states to gain on the monetary mechanism of the ruble zone. He also applies the concept of the prisoner's dilemma to underline the inherent weakness of a monetary regime with several independent central banks. Starting with the latter aspect, he argues that the states in the hypothetical currency union in the model referred to by Lipton and Sachs (1992) face a prisoner's dilemma in determining the rate of domestic credit expansion if they are all equally large and have the same preferences. All will lose in comparison with the outcome in which every country follows agreements to limit credit expansion, since they all have the same incentive to try to free ride. And if any country should not issue excessive credits, it would lose even more. The states will therefore be "forced losers" once they are in the system, voluntarily or not.

The situation in the ruble zone differed from the model outlined above in that Russia's size and ambition to undertake a systemic change made it less inclined to issue credits than most of the other member states. The large size of Russia's share in the ruble zone economy meant that it had to carry a substantial part of the inflationary costs that it caused. The other member states, on the other hand, only had to face a small share of the inflationary costs of their credit creation.<sup>26</sup> Together with Russia's ambition to undertake a systemic change, Bornefalk (ibid.) argues, this made Russia's optimal rate of credit expansion substantially lower than that of most other FSRs. This, in turn, made it possible

---

<sup>25</sup> Conway (1993), referring to the same model as Lipton and Sachs (1992), argues that little shifting of inflation taxes would occur in the ruble zone since all member countries had the same incentive to issue excessive credits. See also Cheasty and Spencer (1993) below.

<sup>26</sup> Ukraine, being the second largest country, would of course have to carry a larger but still fairly small part of the costs of its actions.

for non-Russian member states to gain on the monetary aspects of membership.

These aspects of membership will be analyzed in greater depth in Section 3.2. The results will be used as inputs in the game theoretic analysis of the unrestricted ruble zone in Section 4.1, and the restricted ruble zone in Section 4.2. In this way, we reach a better understanding of the strategic interaction between member countries, which enables us to understand key decisions behind the ruble zone.<sup>27</sup> Before doing this, we will take a closer look at the position of the IMF regarding the ruble zone.

## B. The IMF

During most of 1992, the IMF was strongly in favor of a common currency area for all successor states of the Soviet Union. The following statement in IMF (1994:1, p. 6) reflects the positive view held by the IMF:

Under the right circumstances, a common currency area would have simplified payments among states of the former Soviet Union and provided a solid basis for macroeconomic stabilization. Despite various attempts at interstate coordination, however, a common monetary policy was never effectively applied in the ruble area. Indeed, during 1992-93, divergent credit policies and the emerging capacity to control credit extended in connection with payments across the ruble area, combined with competitive inflation that exacerbated financing imbalances, meant that most benefits of a common currency area were not being realized.

In other words, the IMF tried to reach an unattainable goal, that is, to secure coordination of monetary policies within a unified currency area where several central banks had the power to create money.

This aim has a distinct political flavor. Indeed, it contradicts basic insights into the mechanisms of a currency union. The position of the IMF indicates that it was constrained in some important respect, since it is inconceivable that its qualified economists would not realize the inherent weaknesses of a currency union of the type of the ruble zone. One possibility is that the IMF was guided by political concerns. It has been suggested that the G7 demanded that the IMF should try to force Russia to continue supporting the other FSRs in order to avoid political disruptions and possibly civil wars in those countries. It is, however, difficult to understand why Russia should provide such massive support in the midst of a major macroeconomic crisis while the G7 did little or nothing during 1992. Another explanation could simply be that the IMF failed to realize the realities of the situation in the FSU since it lacked competence on the political and economic situation in the FSU. After all, Russia became a member of the IMF only in April 1992, and the IMF has not been known for paying close attention to political factors.

The poor macroeconomic outcome throughout the ruble zone and the successful Estonian monetary reform in June 1992 made the IMF change its opinion. An article by

---

<sup>27</sup> Conway (1995) and Goldberg et al. (1994) have also analyzed certain aspects of the strategic interaction between the FSRs. We return to this in the following chapter.

the IMF economists Cheasty and Spencer (1993, p. 3) reflects the new attitude of the IMF. But it also indicates that the IMF either overlooked, or drew different conclusions from, the differences in relative size and political conditions that our analysis points out as being of major importance.

Cheasty and Spencer (*ibid.*, p. 3) acknowledge the inherent weakness of the ruble zone mechanism and argue that the absence of a single monetary authority to coordinate national policies resulted in no individual state in the ruble zone having an incentive to limit credit expansion. Instead, they claim, each state had an incentive to maximize its claim on ruble area resources by expanding credits at a faster rate than the area average. They also argue that the overall monetary policy in the ruble area remained more or less under control in the early months of 1992, while it deteriorated later on. The reason they give for this is that Russia pursued a relatively restrained monetary policy during the first part of 1992, which, together with the size of its economy, was enough to impose relative monetary prudence on the other member countries.

Cheasty and Spencer (*ibid.*) do not explain why the free-riding nature of the ruble zone would have been cured by Russia's relative monetary restraint during the first months of 1992. Moreover, they give no explanation as to how Russia's large size could induce other member states to pursue strict monetary policies. It is, however, evident that if a large country expands credits, the others will be more likely to follow suit than if a small country had expanded credits at the same rate. The reason is that the loss they will suffer from receiving a smaller share of the common cake is greater in the former case. The Russian price liberalization on January 2, 1992, is a similar case. As already noted, it was followed shortly afterwards by price liberalizations in other FSRs. If prices had not been freed in the other countries, much of their goods would have disappeared to Russia to fetch a higher price.

A final question brought about by the analysis of Cheasty and Spencer (*ibid.*) is why Russia would be interested in limiting its own credits in the first place if it were a member of a monetary union in which no single country had an incentive to limit credit expansion. Hence, while their statement that monetary policies in the ruble area deteriorated as the year progressed is true, their explanation, as outlined above, could be questioned. We now move on to a more detailed analysis of important aspects of membership.

### **3.2 Decisive Features of Membership**

#### *A. The Relative Size of Member States*

The major shortcoming of analyses focusing only on the free-rider aspects of the ruble zone is that they are unable to explain why any FSR would have wanted to be a member unless the non-monetary advantages were sufficiently large. We will now see that differences in size and political conditions between member states made it possible for some FSRs to gain by issuing excessive credits.

The importance of the relative size of different member states for their optimal rates of credit expansion can be captured in the following common pool model.<sup>28</sup> Assume that the currency area is divided into distinct districts with independent ability to expand the monetary base. Assume also that all people living in the currency area are identical, and that the districts  $i = 1, \dots, I$  only differ in their number of inhabitants, as given by  $N_i$ . The total number of people living in the currency area is then given by  $\sum_i N_i = N$ .

The change in utility of each person living in district  $i$  as a consequence of this district's change in the monetary base is given by:

$$u_i(\pi_i) = S(\pi_i) - C(\pi_i) \quad (1)$$

where  $\pi_i$  is the expansion in the monetary base.  $S(\pi_i)$  is an increasing (concave) function representing the increase in utility for each individual given by the monetary expansion. This increase stems from the amount of seigniorage per inhabitant of district  $i$  that this district levies on other districts.  $C(\pi_i)$  is an increasing (convex) function representing the costs for an individual in district  $i$  incurred in the process.

If each district had to face the costs that its monetary expansion caused,  $\pi_i$  would be set to zero since extracting seigniorage from oneself yields no advantage but certainly a cost. If, however, the costs were to be shared jointly by all districts in the currency area, individual districts could find it in their interest to engage in unilateral monetary expansion. To see this, assume that each inhabitant of the currency area suffers a cost  $Z$  from monetary emissions in all districts. The major part of this cost has to do with the loss of purchasing power of the initial amount of money that the inhabitant has. This cost is determined as follows:

$$NZ = \sum_i N_i C(\pi_i) \quad (2)$$

The change in utility of individuals in district  $i$  is now given by:

$$u_i(\pi_i) = S(\pi_i) - Z \quad (3)$$

---

<sup>28</sup> Readers who prefer a verbal treatment can consider the following example: Think of two countries, one consuming half of the total cake of the currency union, the other just one percent. In order to double consumption per capita in the latter country by ordering the national central bank to extend more credits, an increase amounting to a bit more than 100 percent would suffice if we assume a fixed rate of credit expansion in other member states. This would affect the common price level only marginally. To double consumption per capita in the large country would be impossible. But by letting credits increase towards infinity, the country could come arbitrarily close to capturing 100 percent of the total cake, that is, twice its original share, given the heroic assumptions that the other member states do not expand credits and that production is unaffected by the resulting hyperinflation taking the common price level towards infinity.

Substituting for  $Z$  from equation 2 yields the following optimization function for the decision maker in district  $i$ , who is identical to his subjects:

$$u_i(\pi_i) = S(\pi_i) - \sum_j \frac{N_j}{N} C(\pi_j) \quad (4)$$

Differentiating with respect to  $\pi_i$  and solving yields the optimum condition:

$$\frac{N_i}{N} = \frac{\frac{\partial S(\pi_i)}{\partial \pi_i}}{\frac{\partial C(\pi_i)}{\partial \pi_i}} \quad (5)$$

Hence, the smaller  $N_i$  is, the smaller the marginal gain from monetary expansion will be relative to the marginal cost when unilateral credit expansion is determined optimally from the perspective of individual districts. This means that smaller districts will expand the monetary base at a higher rate than larger districts. If a district is sufficiently large and if the curvatures of  $S(\pi_i)$  and  $C(\pi_i)$  are such that the right-hand side of equation (5) never exceeds the left-hand side, that district will not engage in monetary expansion at all. Even if conditions are such that it does expand the monetary base, it will lose since the smaller districts will do this at a more rapid pace. If there are any winners, they consist of the smallest districts.

### B. Political Conditions

The determination of Russia's reformers to undertake a change of economic system was another factor that made them reluctant to create inflation. In addition to being socially and politically destabilizing, high inflation makes privatization and the growth of the private sector more difficult. The reformist faction of the government also had little interest in redistributing wealth from unorganized to organized groups, for which inflation is an effective tool.

The Baltic countries were reluctant to create inflation for essentially the same reasons. Their determination to redirect trade towards western countries made them even more unwilling to give support to their enterprises, since this would ease the pressure on them to restructure. The fact that their shares of the total economy were so small that unilateral credit expansion would affect the common price level only marginally was not enough to make them interested in remaining in the ruble zone. The reason for this is that they realized that inflation in the ruble zone would nonetheless be too high for their

preferences.<sup>29</sup>

But why did the ambitions of the Russian reform government and the Baltic countries, on the one hand, and of most other FSRs, on the other hand, differ? An approach developed in Bornefalk (2000a,b) can be used to shed light on this issue. It seeks to determine the extent and direction of redistribution in a society where authoritarian rule has collapsed and the principles of majority rule and the median voter have not become established. In doing this, it also determines whether powerful groups will support political and economic reforms or not.

It is found that the extent of appropriative activities and redistribution is determined by the level of development reached by democratic institutions and certain characteristics of interest groups under such conditions. The democratic institutions and the strength of interest groups determine the vulnerability of unorganized groups to appropriative activities. The ability of interest groups to cooperate with each other in seeking rents and the share of the economy's resources that is controlled by interest groups also affect the outcome.<sup>30</sup>

Bornefalk (2000b) argues that the ability of politicians to respond to rent seeking by offering favors to organized groups is lower the more transparent the mechanisms for redistribution are and the greater the expected costs faced by politicians when engaging in redistribution are. Determinants of the degree of transparency are the freedom and impact of mass media, but also the degree of complexity of the economic system. These factors also affect the expected costs that politicians have to face when they reallocate resources from the majority to the organized minority.

The maturity of political organizations is another determinant of the expected costs faced by politicians. The reason for this is twofold. The first is that the more mature the organization is, the more credibility it will have built up. By favoring interest groups, the organization risks damaging its credibility, thereby lowering its chances to be elected in upcoming elections. Second, the longer freedom of organization has existed, the more likely it is that voters will have a better alternative to cast their votes on if the government injures the interests of the majority. Hence, the vulnerability of unorganized groups to appropriative activities will typically be lower the more political organizations have invested in building their reputation. Prior to the introduction of political freedoms, political organizations will in general have had little opportunity, or incentives, to build a reputation for defending the interests of the unorganized majority. The vulnerability of unorganized groups to appropriative activities undertaken by organized groups, or the government itself, is therefore typically great under such conditions.

When interest groups are able to cooperate with each other, they will resist economic and political reforms that make redistribution more difficult. If they are unable

---

<sup>29</sup> The analysis in Hansson (1993) shows that the Baltic governments were well aware of the trouble with the ruble.

<sup>30</sup> The model is valid for countries where a majority of the population does not belong to interest groups. This condition was certainly fulfilled in the FSU during the period we consider since the restrictions against political organization were quite effective and far-reaching in the Soviet Union.

to cooperate, it will typically be in their interest to support reforms limiting the effectiveness of appropriative activities since this also means that the threat against their own resources from other organized groups will be lower. This, in turn, allows interest groups to spend less on appropriative activities. They will, however, resist reform also in the non-cooperative case if the vulnerability of unorganized groups to appropriative activities is great enough and the share of resources that is controlled by interest groups is small enough.

The relevance of this for the development in the ruble zone is found in the differences between member countries in the level of development reached by democratic institutions and different characteristics of interest groups. In the Baltic countries, mass movements started to gain strength at the end of the 1980s as *glasnost* was launched. The leading role of the Communist Party in Soviet Estonia was terminated already on February 23, 1990. According to Lagerspetz and Vogt (1998, p. 67), the number of political parties mushroomed immediately afterwards. In early 1992, mass media and political parties had already reached a considerable level of development. The previous experience with basic democratic institutions that these countries had gained prior to world war two probably contributed to making this process more rapid, although this period was rather limited. A desire to reestablish close contacts with countries in the west, particularly neighbors in northern Europe, also contributed to a rapid development of democratic institutions. Moreover, many Soviet institutions were abolished and ties between enterprises in the Baltic countries and their ministries in Moscow were cut off. This weakened the influence of interest groups and their ability to cooperate with each other. They therefore became less interested in stalling political and economic reforms.

In Russia, the vulnerability of unorganized groups was rather high although political freedoms had been introduced. The reason for this was that popular mass movements were weaker and political parties less mature, and that mass media had a limited reach beyond major cities. Still, a government devoted to radical economic reform had come to power in the fall of 1991. Following the reasoning in Bornefalk (2000b), this can be explained by the inability of interest groups in Russia to cooperate during this period of time. A contributing factor for this was that the Soviet Union was breaking apart during the spring and summer of 1991. Powerful interest groups resisted this development. In August 1991, a coup aimed at halting the disintegration process failed. As a result, major parts of the establishment were weakened. Most importantly, the communist party with all its connections to various sectors in the economy was declared illegal. Interest groups therefore had to find new ways of influencing politics. This opened a window of opportunity for political and economic reform, which brought a team of radical reformers to power. We will return to the strength and activities of Russian interest groups in Chapter 6.

In most FSRs other than Russia and the Baltic countries, democratic institutions were less well developed, and Soviet era political and economic structures remained more intact. This meant that the vulnerability of unorganized groups to rent seeking was greater



in these countries. It also meant that interest groups were able to cooperate with each other to a greater extent, and therefore opposed reforms lowering this vulnerability. This, in turn, posed quite severe restrictions on the economic reform space in these countries. As regards monetary reform, these governments were accordingly not interested in implementing monetary reform limiting their ability to transfer wealth to organized groups, even though this would lead to less rent seeking and a less dramatic fall in production.

Table 3 reports assessments by the Freedom House of political rights and civil liberties for Estonia, Russia, Ukraine, and Kazakhstan for 1991-94. Estonia has higher ratings than the other countries, but Russia and Ukraine are not far behind during 1991-93.<sup>31</sup> These figures do not, however, reflect the accumulated effect of previous experience of political freedom. Since Estonia had more experience than the other countries, the figures in the table underestimate the difference in terms of the level of development of democratic institutions. Such an index should reflect the combined effect of the extent of political freedoms and the time that has elapsed in the presence of political freedoms. Other factors that should be incorporated include the purchasing power of consumers and the relative payoff of production and investments in organization. The reason why the purchasing power matters is that this affects the possibilities for independent media to develop.

**Table 3. Freedom Ratings of Selected FSRs for 1991-1994**

	Political rights			Civil liberties		
	1991-92	1992-93	1993-94	1991-92	1992-93	1993-94
Estonia	2	3	3	3	3	2
Russia	3	3	3	3	4	4
Ukraine	3	3	4	3	3	4
Kazakhstan	5	5	6	4	5	4

Source: Nations in Transit 1998, Freedom House (1999).

### 3.3 A Summary of Pros and Cons of Membership

We have seen that Russia's size and interest in systemic change made it possible for other member states to gain by free riding. Therefore the ruble zone freed them to a large extent from their balance of payments restrictions and thereby from the need to adapt to their sharply falling terms of trade. Although member states other than Russia and the Baltic countries had greater incentives to expand credits, they were not totally unrestrained. Since Russia set up the system, it had the opportunity to dissolve it or change its rules or membership in the event that its original institutional design did not serve its interests, for

<sup>31</sup> The differences increased during the following years, which is an implication of the model of Bornefalk (2000a,b).

example by creating excessive inflation. However, Russia's role as a gatekeeper provided only weak incentives to limit credits.<sup>32</sup> The reasons for this will be made clear in the following game theoretic analysis.

Apart from this "macro advantage," there was the "micro advantage" of retaining the payments system that was already in place. There were also real advantages in belonging to the ruble zone for non-Russian FSRs. These were access to cheap inputs as well as export markets, that is, the possibility to continue Soviet-type trade. The real advantages were of a temporary nature since it was unclear at what pace prices within the ruble zone would climb towards world market levels. Despite their temporary nature, these advantages played a major role in most countries' decisions on whether or not to stay in the ruble zone, but they were not tied directly to the payments mechanism. For instance, Ukraine was not cut off from access to cheap inputs, mainly oil and gas, when it was expelled from the ruble zone in the fall of 1992. At any rate, countries hoping to gain by continued Soviet-type trade with Russia were more willing to accept drawbacks of the ruble zone, such as high inflation.

Thus for non-Russian FSRs, membership in the ruble zone had much to offer, although macroeconomic stabilization would have to be sacrificed. It would otherwise have been irrational for them to participate, at least if we look upon them as decision making entities. If they had suffered from the monetary aspects, they would have tried to renegotiate the way the ruble zone was constructed. Clearly, there was no need for that since the non-Russian member states gained from both the monetary aspects—as shown by the great explicit trade subsidies—and the continued state trading with its huge implicit subsidies. But this does not mean that people in individual countries benefited from the membership, since Russia's subsidies were not used for the benefit of society as a whole. It therefore appears as if interest groups played an important role in decisions on the ruble zone in most FSRs. Indeed, Åslund (1993, p. 21) argues that:

Against the Russian national interest in minimizing costly subsidized credits to other republics stand not so much the interest of rent-seeking republics as rent-seeking state enterprises in all republics. To them, all kinds of subsidized state credits are beneficial. Russian state enterprise managers are naturally in favor of additional cheap credits to other republics, because it means that they can continue "selling" with a minimum of efforts, since the bill is eventually paid by the Russian state. Part of the explanation why the various former Soviet republics benefited so little from the massive Russian subsidies in 1992 is that they went to rent-seeking strata in recalcitrant state enterprises that utilized the subsidies to keep going without adjusting.

Our next step is to combine different aspects of membership in the ruble zone for different FSRs. This will allow us to understand their decisions on membership and behavior in the ruble zone.

---

<sup>32</sup> Note that this mechanism implies an upper limit for monetary restraint for non-Russian member states corresponding to Russia's choice of monetary policy. Countries wanting to pursue a stricter monetary policy than Russia should therefore leave the ruble zone. This mechanism therefore suffers from an adverse selection problem.

## 4. STRATEGIC INTERACTION IN THE RUBLE ZONE

### 4.1 The Unrestricted Ruble Zone

During the first half of 1992, there were no explicit restrictions on the deficits of non-Russian FSRs on the correspondent accounts with Russia. The only outside restraint on the issue of credits of non-Russian member countries was the risk of exclusion from the ruble zone. The unrestricted currency regime is analyzed in this section. The currency regime with limits on the amounts of so-called technical credits that Russia offered other member states is analyzed in Section 4.2.

In the following analysis, the ruble zone is viewed as a game between one large player, whose share of the total economy of the currency union is one half, and three small players, each representing one sixth of the economy.<sup>33</sup> The types have different valuations of various aspects of membership. These are discussed in subsection A. In subsection B, the order of moves in the game is specified along with the information and choices that are available to each player who has to make a move. After a specification of a payoff function associating valuations with different strategy combinations, the game is played. Finally, the relevance of this analysis for the development in the real ruble zone is discussed.

#### *A. Valuations of Different Types of Member States*

The risk of exclusion from the ruble zone provided the non-Russian member states with an incentive to take their reputations into account. Countries more interested in continued cooperation with Russia were more interested in maintaining a good reputation, and therefore less inclined to issue credits. Countries abusing the payments mechanism by issuing excessive credits had to risk exclusion from access to cheap inputs and export markets by being excluded from the ruble zone. Thus individual member states making decisions on the amount of credits to issue had to weigh the benefit from increased credits against the risk of suffering losses from exclusion from the payments mechanism and the Soviet-type trade. At the same time, countries that did not issue excessive credits could not be sure that the ruble zone would remain in place. Furthermore, it was not known for how long Soviet-type trading would be continued, and at what pace prices would increase to world market levels. This lowered the incentives to build a reputation as a country that issued credits with restraint.

This reasoning implies that the non-Russian FSRs had to choose among three alternative strategies for their relations with Russia. The first was to leave the ruble zone and introduce a national currency. This was necessary for countries attempting to stabilize

---

<sup>33</sup> The sizes of the different types have been chosen to ease computations. What matters is that there is one player who is considerably bigger than the others.

faster than Russia did. Faster stabilization, in turn, would make redirection of trade towards world markets easier. The timing would be determined by when the advantages of leaving outweighed those of staying. Since the advantages of leaving also depended on the way the country left the ruble zone, it could be advantageous to remain until preparations for a successful monetary reform had been completed, and in the meantime benefit from implicit and explicit trade subsidies. The Baltic countries pursued this strategy. In the game, they are represented by *E*. The second alternative was to try to stick to Russia as long as possible, hoping for continued Russian subsidies in addition to access to a fairly effective payments system. Kazakhstan and Belarus belong to this group, which is represented by *K*. The third alternative was to cheat as much as possible so as to achieve short-term gains.<sup>34</sup> Ukraine pursued this strategy most consistently. Hence, *U* is used as a label for this type.<sup>35</sup>

To capture the different aspects of membership in a coherent framework, we estimate relative valuations of various aspects of membership for different types of countries at the time of the break-up of the Soviet Union. The estimates are presented in Table 4. The assigned valuations have no absolute meaning; it is their relative sizes that matter. The purpose is to show that plausible answers to questions asked in this paper can be given, questions which other analyses have left unanswered.

The different types of non-Russian member states are represented by *E*, *K*, and *U*, respectively, while *R* is a stylized version of Russia.<sup>36</sup> The pros and cons of membership are measured in relation to the valuation of defect, which is to leave the common currency area. The valuation of defect is normalized to 0. *E* was less dependent on having a payments system set up by *R* than other types. The reason for this was that it had come further towards preparing a currency reform and the introduction of current account convertibility. Its valuation of the ruble zone payments system is therefore set to 1, while the valuation of the other types is set to 2. *K* was most dependent on continued Soviet-type trade, while *R* would gain less because of the composition of trade. Nevertheless, many industries in *R* were dependent on semi-finished goods produced in other FSRs, which explains the 1. The 2 for *E* signifies that countries belonging to this category wanted to redirect trade towards world markets, but were heavily dependent on imports of energy and raw materials from Russia. There was thus a conflict between interests. In early phases of the transition, it is safe to conclude that these countries would gain by keeping the old ties since that would give them some time to adapt to the new situation.

---

<sup>34</sup> For a more detailed classification of different types of FSRs, see Åslund (1994).

<sup>35</sup> The time horizon of decision makers was most likely of relevance for decisions on what strategy to pursue. The attempts to achieve short-term gains—which generally resulted in medium-term disaster—could be explained by the policy makers of these countries having short time horizons. For instance, if they were maximizing rents, it is possible that they concluded that it was preferable to grab as much as they could before they were ousted after elections.

<sup>36</sup> One FSR, Kirgizstan, does not fit into any of these types since it was reform-minded but dependent on keeping the old ties with Russia, to a great extent because of its geographical location.

**Table 4. Valuations and Payoffs in the Multilateral Unrestricted Game**

	Type of country:			
	<i>R</i>	<i>U</i>	<i>K</i>	<i>E</i>
<b>Valuation of:</b>				
Payments system	2	2	2	1
Keeping old ties	1	2	3	2
<b>Inflation taxation:</b>				
a) gain/loss from unilateral inflation taxation increasing share of common cake by 10%	-7	5	4	3
b) costs per type transgressing (other than <i>R</i> )	-3	-1	-2	-3
<b>Total payoff if: (<i>R</i> is assumed not to transgress)</b>				
All cooperate	3	4	5	3
The country transgresses, others cooperate	0 <sup>1</sup>	9	9	6
The country cooperates, others transgress	-3-ρ	2-υ	1-κ	-3-ε
All transgress	-6-ρ	7-υ	5-κ	0-ε
The country defects	0	0	0	0

<sup>1</sup>A type other than *R* transgresses.

When it comes to the possibility of levying inflation taxes on other member countries, *R* is the great loser. Because of *R*'s size and interest in economic reforms, it would actually lose 7 in the example of the table by unilateral credit expansion at a rate sufficiently high to increase *R*'s share of the common cake by 10 percent. The value is derived from the valuation of the increased share, which is set to 5, and *R*'s valuation of the resulting inflation, which is set to -12. The great magnitude of the cost is explained by the interaction of Russia's size, as shown in the common pool analysis, and Russia's interest in economic reforms, as shown in the analysis of the available economic reform space.

The value of the other types depends on how they perceive the inflation they cause. If we attach the same value as for *R*, which is 5, to an increase in the country's share of the common cake by 10 percent, and then subtract the negative value of inflation caused by the unilateral credit expansion, we end up with 5, 4, and 3 for *U*, *K* and *E*, respectively. It is here assumed that *U* does not care about inflation. One reason for this is that the leaders are uninterested in a change of economic system, another that the leaders and their constituencies can increase their rents if they allow inflation to rise. This would outweigh the negative effects of inflation.

The reform-minded *E* would lose more for every percentage point that its credits raised the common price level, since stability is needed for successful reforms. Furthermore, enterprises in *E* would not be forced to restructure as rapidly as if they had not received any subsidies, as mentioned above. We therefore subtract 2 from the gain in

the share of the cake that the unilateral credit expansion would cause. Note that this is much less than was subtracted for  $R$ . The reasons for this are found in the smaller size of  $E$  and the non-linearities in the seigniorage and cost functions. The former factor makes the resulting inflation from an attempt to increase the country's share of the cake substantially lower. The latter factor makes the gain from issuing excessive credits relative to the costs of doing so smaller, the higher inflation becomes.

A similar reasoning applies to the costs per other type transgressing, that is, issuing excessive credits.  $R$ , which is assumed not to transgress since that would clearly hurt itself, would lose 1 because of loss of goods to the country transgressing, and 2 because of the increase in the common price level. Since  $U$  does not care about inflation, it would only suffer a loss of 1.  $K$ , and even more so  $E$ , would also suffer from an increase in inflation. Indeed,  $E$ 's valuation of inflation is identical to that of  $R$ . Payoffs for different strategy combinations in the game will be explained in the following subsection.

If the game is played only once with this set-up of valuations, decisions on behavior in the currency union have an entirely static character. To incorporate dynamic aspects, the one-period game can be turned into a repeated game. That the game is repeated means that decision makers consider what will happen in future periods given the history, that is, how the game has been played in previous periods. In the case of the ruble zone,  $K$ , for instance, would have to take into account the risk that the ruble zone is dissolved or changed in such a way that future payoffs are affected when the rate of credit expansion is decided. If it is dissolved, it would no longer be possible to levy inflation taxes on  $R$ . Access to the payments system as well as to underpriced imports would also be denied.

These losses, caused by the break-up of the currency union after the first period, are denoted by  $\rho$  for  $R$ ,  $\upsilon$  for  $U$ ,  $\kappa$  for  $K$  and  $\varepsilon$  for  $E$ . Note that the currency union is dissolved, or its mechanisms changed, if more than one type transgresses since  $R$  then makes a loss. As has already been indicated, the valuation of continued Soviet-type trading in periods following the first is not necessarily identical to that of the first period, since prices within the ruble zone could be expected to continue rising. Likewise, the valuation of the payments mechanism in the ruble zone is also likely to decline since new alternatives could be developed.

### B. *A One-Shot Game*

The structure of the game is as follows.  $R$  sets the rules of the game. In particular,  $R$  decides which other players will be allowed to participate. The admitted players then choose whether or not to participate, and, if they participate, how to play. The options for  $U$ ,  $K$  and  $E$  are to cooperate, transgress, or defect. As regards the information structure, we assume common knowledge of which type individual members belong to, including the sizes of  $\upsilon$ ,  $\kappa$  and  $\varepsilon$ , respectively. This can be defended by the fact that the prospective members were involved in lengthy discussions covering their future cooperation by the

time of the break-up of the Soviet Union.<sup>37</sup> In making this assumption, we also assume that the non-Russian member states realized the specific situation and ambition of their dominant neighbor, that is, Russia's size and willingness to reform.<sup>38</sup>

However, there is no need to make an assumption of perfect information as regards credit emissions in member countries, just to mention the most important decision variable. Such an assumption would also clearly be wrong. It was only possible to get a partial picture of other member countries' monetary policies by the deficits on the correspondent accounts with Russia, and to some extent by the inflation levels. The information obtained in this way was not only imprecise, it was also delayed since there is a lag between the moment credits are expanded and when the effects on trade and inflation occur. As regards inflation, the lag has been measured at between three and four months for Russia. This lack of information was a central aspect of the ruble zone and contributed to its demise.

A game tree is shown in Figure 3. *R* makes the first move and thereby determines the possible strategies for the other countries.<sup>39</sup> The analysis of *R*'s choice of strategy concentrates on the one of admitting all other FSRs, since that was what Russia chose to do in reality. The size and goals of *R* are assumed to be enough to take its behavior as a member of the ruble zone as given: *R* will not issue excessive credits since that would hurt the country, as seen in Table 4.<sup>40</sup>

When *U*, *K* and *E* make their choices, they do not know each other's strategies, which is indicated by the ovals in Figure 3. If they choose to stay, they could either issue credits in line with the loose agreements that were made,<sup>41</sup> which would be to cooperate, or issue excessive credits, which would be to transgress. If they leave, they choose the strategy defect. This could be done in a cooperative fashion, which would be to notify the other member countries prior to leaving and to refrain from policies that would hurt them, such as flooding their territories with ruble notes that have been exchanged for the new currency. The option to defect will not be considered in the analysis since it is weakly dominated by joining the ruble zone and then transgressing, at least if  $\upsilon$ ,  $\kappa$  and  $\varepsilon$  are sufficiently small.<sup>42</sup>

---

<sup>37</sup> It could be argued that in reality there was incomplete information about what type a particular country belonged to. In order to handle this incomplete information about valuation of payoffs, it could be assumed that countries considering membership shared a prior distribution over probabilities that certain countries would belong to certain groups. For instance, the probability that Estonia would be of *U* type was 10 percent, of *K* type also 10 percent, and of *E* type 80 percent. The types would then be determined by a move prior to those of other players by a fifth player, nature. This transforms the problem of incomplete information into one of imperfect information, which can be handled by techniques used in this essay.

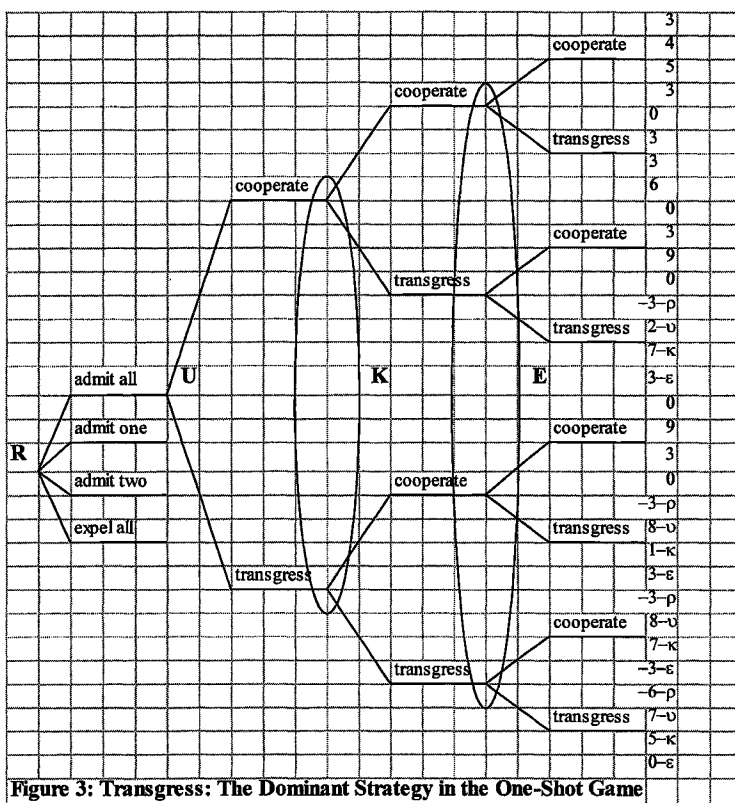
<sup>38</sup> Without this knowledge, they would not believe that they could gain by the ruble zone mechanism, and would therefore have been more reluctant to retain membership.

<sup>39</sup> The game is, however, solved backwards.

<sup>40</sup> In *R*'s case, excessive credits would mean substantially more than agreed to, so as to make the other countries want to leave the currency union.

<sup>41</sup> See Conway (1993).

<sup>42</sup> See Table 3.



If the countries cooperate, it will be possible to keep the payments system. If they choose to transgress, the result will be high inflation and unintended trade subsidies running from *R* to the other members. If *R* loses more than it gains from allowing the other countries to stay in the currency union, it will change the rules of the game, for example by excluding the others. Hence, if the small players join, they have the choice of either cooperating or transgressing, that is, of issuing credits either below or above the limit where Russia would change the rules of the game.<sup>43</sup> Note that there were no other institutional arrangements for enforcement of monetary agreements between member states than this implicit game.

The digits in the game tree of Figure 3 represent payoffs connected to different strategy sets. These payoffs are shown in Table 4, and are sums of valuations of various aspects of membership.<sup>44</sup> The digits in each quadruple give the payoffs of *R*, *U*, *K*, and *E*

<sup>43</sup> Of course, in reality no such limits were set or could be foreseen with certainty, but this is a simplifying abstraction. What the abstraction essentially does is to transform the decision problems of the countries from continuous actions and times to choices of one out of two distinct alternatives.

<sup>44</sup> If *U*, *K* and *E* cooperate, the countries will gain 3, 4, 5 and 3 since they can then continue to use the payments system and benefit from favorable aspects of the continued Soviet-type trade, such as low energy prices and access to intermediate goods. Also, inflation will not be excessive. By excessive inflation is meant high inflation created by the ruble zone mechanism. If *U* transgresses while *K* and *E* cooperate, *U* will gain 9



in descending order. The loss of  $R$  will be large if the others transgress. It will lose seigniorage to the other member states, and suffer from excessive inflation. On top of that, it might lose the payments system as it will have to change the institutional framework of the currency union. That could either entail trying to support the efficient, or cooperative, equilibrium through a punishment or trigger strategy, or excluding  $U$ ,  $K$  and  $E$ . The analysis of the ruble zone with a trigger mechanism is postponed to the following section.

Each player chooses a strategy so as to maximize expected utility given its beliefs regarding the other players' actions.<sup>45</sup> Examination of the game shows that it is rational to transgress for both  $U$  and  $K$ , since this is a dominant strategy. For  $E$ , joining and transgressing weakly dominates defecting if  $\varepsilon$  is less than or equal to zero. Hence, rational behavior of  $U$ ,  $K$  and  $E$  necessarily leads to a situation where  $R$  loses and decides to dissolve the currency union. Hence,  $R$  should not have admitted the others. With payoffs as in Table 4, both  $U$  and  $K$  benefit from membership.<sup>46</sup> For other reasonable choices of payoffs, the equilibrium would still have  $U$ ,  $K$  and  $E$  transgressing, and  $R$  changing the rules of the game.

### *C. Policies of Member Countries*

The results of our game theoretic analysis fit well with the behavior of non-Russian member states. Estonia realized the inherently inflationary nature of the ruble zone and was determined to change the economic system and move closer to the West. Hence, it decided to leave the currency union at an early stage. It did, however, remain during almost all of the time the ruble zone was unrestricted, which is in line with the result that

---

while  $K$  will gain 3 and  $E$  0, which is also the payoff of  $R$ . If  $K$  also transgresses, the payoffs will be  $-3-\rho$ ,  $8-\upsilon$ ,  $7-\kappa$  and  $-3-\varepsilon$ . The losses represented by Greek letters reflect  $R$ 's change of the rules of the game, which could lead to the loss of the payments system and access to cheap imports. The change of the rules of the game would follow from  $R$ 's loss when two types transgress.  $U$ 's and  $K$ 's gains from the inflation taxation of the common market, that is,  $R$ , are partly offset by the loss of the payments system. The point is that it is possible to gain by defecting if you issue large enough credits, given that the others cooperate. If  $E$  transgresses while  $U$  or  $K$  cooperates, the result will be similar, albeit with different payoffs. If all three types transgress, payoffs will be as in the lowest quadruple.  $U$ ,  $K$  and  $E$  will have suffered from excessive inflation stemming from their excessive issuing of credits and will suffer from a change of the payments system. On the other hand, they will have been able to tax  $R$ , which more than offsets the losses for  $U$  and  $K$  for reasonable values of  $\upsilon$  and  $\kappa$ . For  $E$ ,  $\varepsilon$  has to be less than or equal to zero for the benefits of membership to outweigh the drawbacks in case all types transgress. It would otherwise have been preferable to play defect. For  $U$  and  $K$ , to play defect is clearly inferior to both cooperating and deceiving, since this would mean giving up the payments system and the possibility to continue importing mainly energy at low prices as well as the possibility to levy an inflation tax on, above all,  $R$ .

<sup>45</sup> For the equilibrium to be a Nash equilibrium, it is required that the beliefs should be correct.

<sup>46</sup> As mentioned before, this does not mean that the countries as a whole gained. If we abstract from interest groups, there is another difficulty in deciding whether the non-Russian FSRs gained from membership. It has to do with the valuation of, for instance, oil. It would not be enough to take the difference between the world market price and the price a particular country paid for its oil, if such figures could be obtained. The reason is that the oil was most likely inefficiently used, since prices were artificial and enterprises were not private profit-seekers. And the value for society of wasted oil depends on how badly it is wasted. It could very well be negative.

defect was weakly dominated for  $E$  for a small enough  $\varepsilon$ .<sup>47</sup>

Ukraine was also determined to leave the ruble zone, but waited until it was expelled after having issued excessive credits in order to make short-term gains. In order to gain in the end by issuing excessive credits and possibly causing the country's exclusion, Table 4 shows that it is necessary that the short-term gains to that type of country from the issued credits should outweigh the costs of exclusion, in Ukraine's case measured by  $\psi$ . In other words, if an individual state had managed to issue enough credits before it was excluded or the ruble zone was dissolved, it would have gained.

It appears as if Ukraine had the ambition to maximize short-term gains. In an attempt to solve a rapidly growing problem with inter-enterprise arrears in the spring of 1992, Ukraine doubled the amount of credits extended to enterprises, knowing that most of the costs in terms of inflation could be shifted over to other members of the ruble zone, mainly Russia. This triggered Russia's change of the mechanisms of the ruble zone, as described below. But, as is clear from Johnson and Ustenko (1993), the Ukrainian monetary policy of issuing credits to nearly anyone who asked harmed Ukraine the most in a somewhat longer perspective.

Russia's behavior remains to be explained. The underlying assumption of rationality, together with the assumption of common knowledge, leads to the conclusion that Russia should never have set up such a game. In other words, it should not have admitted the other FSRs into the currency union, since it should have been able to foresee the results. It therefore appears as if Russia acted irrationally. As we will see in Chapter 6, however, the reformers did not have the knowledge of how to execute a successful currency reform during the period when the political conditions would have allowed such a radical reform.

Once they had gained this knowledge, the political conditions had changed. Most importantly, interest groups, above all state enterprise managers, had established structures that enabled them to cooperate and were therefore both unwilling to allow reforms lowering the effectiveness of rent seeking and strong enough to postpone such reforms. Given the relatively low level of development of democratic institutions in Russia, this meant that the economic reform space was quite limited. We will return to these issues, but will first show why the limits for deficits on the correspondent accounts set up by Russia in July 1992 failed to impose fiscal and monetary restraint on Russia's partner countries. An interpretation of Russia's decision will also be offered.

## **4.2 The Restricted Ruble Zone**

In the spring of 1992, the fallacies and costs of the unrestricted ruble zone became increasingly clear. The insight was spreading among reformers that if stabilization was to

---

<sup>47</sup> When Estonia left the ruble zone, it initially refrained from using the collected ruble notes at the expense of the remaining member countries. Later on, it chose to sell its notes to Chechnya because of discontent with the way Russia fulfilled its agreed obligations.

succeed, either the ruble would have to be nationalized, or effective restrictions would have to be imposed on trade deficits that member countries could run with Russia. At the same time, critique against the economic reforms escalated rapidly. The leading reformers had to give in, both in terms of replacing reformers with conservatives in high positions and in terms of policies. The most important, and damaging, policy change was that they agreed to issue subsidized credits to state enterprises. Thus urgent decisions on monetary reform had to be made in a hostile environment. The reformers chose to change the mechanisms of the ruble zone rather than dissolve it. This set the stage for a capricious and costly unraveling of the monetary area.

The restricted ruble zone did not improve incentives for fiscal and monetary restraint among member countries, for reasons developed in subsections B and C. In subsection A, the political changes and the new mechanisms of the currency union will be treated. As we will see, the nature of the game changed, but not enough to improve the macroeconomic outcome.

#### *A. Foundations for a New Payments System*

Boris Yeltsin tried to alleviate the escalating critique against the reforms by making concessions to opponents of the government, including the parliament and state enterprise managers (Åslund, 1995, p. 94). Yegor Gaidar realized that his position was weak as the Sixth Congress of People's Deputies was about to open on April 6, 1992. He managed to hold on to his position as acting prime minister by making compromises, but had to continue compromising after the Congress. In April, Gaidar gave up his position as minister of finance to an old apparatchik without ambitions for macroeconomic stabilization. In May, he did not challenge Yeltsin's decision to sack the minister of fuel and energy whose attempts to liberalize energy prices had provoked unanimous criticism from state enterprise managers. The energy portfolio was given to Viktor Chernomyrdin, a former Soviet minister of the gas industry, who joined the government together with two other high-rank state enterprise managers. They were all appointed deputy prime ministers, and the reformist hegemony in government had been broken.

Another setback for the reformers was that Viktor Gerashchenko, the last chairman of Gosbank, replaced the moderately reformist Georgy Matyukhin as chairman of the CBR.<sup>48</sup> Gerashchenko was supported by state enterprise managers. According to Åslund (1995, p. 98), he rewarded members of the old elite for their support by giving anybody who asked cheap credits, and kept the refinance rate low despite high inflation rates. Since the CBR was subordinate to the anti-reformist parliament and independent of the government, little could be done to avoid high or even hyperinflation from developing. Moreover, the divided government started to run large budget deficits.

In this environment, the reformers had to come up with a strategy that could end

---

<sup>48</sup> Gaidar was given a choice by Yeltsin between the reformist Boris Fedorov and Gerashchenko, but chose the latter since he believed that the Supreme Soviet would never approve Fedorov.

Russian subsidization of other FSRs through the ruble zone. The resistance was strong from the beneficiaries of subsidies, above all Russian state enterprise managers. The previously mentioned decision by Ukraine in June 1992 to clear inter-enterprise arrears by extending massive credits served as a catalyst in pushing through a change. The credit expansion approximately doubled Ukraine's money supply (Granville, 1995, p. 72).

The Russian government responded by limiting the available credits—which from that moment were called technical—on the correspondent accounts of other FSRs on July 1, 1992. The purpose of this decision was, according to Granville (*ibid.*), to limit the inflationary effect of Ukraine's measure and to ensure that no other FSR would follow Ukraine's example. The technical credits were subject to negotiations, and the interest rate was the refinance rate of the CBR. With the new limits, the automatic clearing of bank deposits in these countries into bank deposits in Russian banks was ended. Hence, an enterprise in Kazakhstan could no longer automatically pay a Russian enterprise with rubles held in a commercial bank in Kazakhstan. The Kazakhstan rubles would only be settled in Russian rubles if the Central Bank of Kazakhstan (CBK) held enough deposits at the CBR, or if the CBR explicitly granted the CBK credit or an overdraft for the settlement (Sachs, 1994).

An attempt was also made to decentralize the payments system, thereby making the CBR less vulnerable to pressure to finance deficits. The CBR encouraged commercial banks in the FSU to clear payments on their own by establishing correspondent accounts with each other. This policy was reflected in Russia's agreement with Estonia following the introduction of the Estonian kroon on June 20, 1992. The agreement allowed banks in Russia, which were allowed to open correspondent accounts under the general foreign exchange license, to open correspondent accounts in Estonia (Granville, 1995). Hence, provisions had been made for a payments system that could handle payments between Russia and countries that had left the ruble zone.

The advantage of the restricted system was, according to Sachs (1994), that the other FSRs could no longer automatically issue ruble credits for the purchase of Russian goods. The major drawback was that as, for instance, Kazakhstan had run out of credits on its correspondent account with Russia and could not get an overdraft, a Kazakhstan enterprise could not use its ruble deposits in Kazakhstan to buy goods in Russia since the CBR would not clear the transaction. Thus, a shortage of credits led automatically to inconvertibility of one ruble for another. This made it extremely difficult for the CBR and the Russian government to resist demands for new credits. Hence, under pressure from the Russian industrial lobby, Russia revised the limits for technical credits several times during the second half of 1992. Total limits for the second half of 1992 were set at Rbs 215 bn in July. At the end of the year, however, they had reached Rbs 942 bn.<sup>49</sup> The automatic technical credits during the first half of 1992 reached Rbs 316 bn, indicating that the limits were almost automatically increased during the second half of the year.

---

<sup>49</sup> These preliminary figures are from Granville and Lushin (1993), who relate to the CBR and Commerciant. They do not coincide with figures in Table 2, but are shown in order to make a comparison between the first

Thus, this mechanism provided precious weak incentives for the non-Russian members of the ruble zone to tighten their fiscal and monetary policies relative to the first half of 1992.

In the following subsections, we will seek answers to several questions provoked by this development. First, was the inflationary outcome inevitable, given the new rules of the game and the political setting within which the game was played? If so, why did the reformers change the rules of the game in the way they did? Two main alternatives can be identified. The first is that the punishment mechanism inherent in the restrictions was designed with the purpose of lowering inflation by forcing member states to pursue more restrictive monetary policies. The second is that the restrictions were merely intended as a step towards a complete dissolution of the currency union. If the latter was the case, why was the Russian ruble not nationalized immediately? In order to answer these questions, the game theoretic analysis of the previous section will be extended to incorporate the new features of the payments mechanism as well as the uncertainty about the goals of Russia brought about by the political changes. The analysis also provides a basis for an examination of the consequences of the restricted ruble zone for non-Russian FSRs that did not leave the currency union voluntarily at an early stage.

### *B. Bilateral Games with Renegotiation*

The ruble zone without explicit restrictions on bilateral deficits was analyzed as a multilateral game. With explicit restrictions on technical credits offered by Russia to the other members, the corresponding game changes into three bilateral games with  $R$  playing  $U$ ,  $K$  and  $E$  simultaneously in one game each. The fact that technical credits were limited but could be renegotiated when the limit had been reached makes it a repeated game. The uncertainty about the political situation in Russia, that is, whether reformers or conservatives will play the dominant role at a particular moment in time, transforms the game into one of incomplete information about the valuations of  $R$  in future periods.

To handle the problem of incomplete information, we assume that countries considering membership share a prior distribution over probabilities that  $R$  would be of a reformist,  $R_r$ , or conservative,  $R_c$ , type. The probability of  $R_r$  is set to  $\rho$ , and that of  $R_c$  to  $(1-\rho)$ .  $\rho$  is assumed to be of common knowledge. The type of  $R$  is decided by a move of a fifth player,  $N$ , for Nature, in each future period prior to the moves of the other players.

Following the same procedure as in Section 4.1, estimated relative valuations of aspects of membership in the restricted ruble zone for different types of countries are given in Table 5. Valuations differ from those of the unrestricted game since  $R$  can now be of two types and since the environment has changed. For instance, the provisions for decentralized payments systems made by the CBR lower the valuation of access to the centralized payments system of the ruble zone for  $R_r$  and  $E$ , the types aiming for, and having made preparations for, a decentralized payments system. For  $U$  and  $K$ , the

---

and second halves of 1992 possible.

valuation remains at 2. This is also the valuation of  $R_c$  since a centralized payments system suits the interests of conservatives better.

**Table 5. Valuations and Payoffs in the Bilateral Restricted Games**

	Type of country:				
	$R_r$	$R_c$	$U$	$K$	$E$
<b>Valuation of:</b>					
Payments system	1	2	2	2	0
Keeping old ties	0	2	0/2	0/3	0/1
Inflation taxation:					
a) gain/loss from unilateral inflation taxation increasing share of common cake by 10% <sup>1</sup>	—	1	5	4	0
b) loss if partner country transgresses	-3	-1	-5	-6	-10
Loss of monetary control to $R$	—	—	-1	0	-2
<b>Total payoff if:</b>					
Bilateral cooperation $R$ - others	1	4	1/3	2/5	-2/-1
The country transgresses, partner cooperates	—	5	6/8	6/9	-2/-1
The country cooperates, partner transgresses <sup>2</sup>	-2	3	1/-2	2/-1	-2/-11
Both transgress <sup>2</sup>	-2	4	6/3	6/3	-2/-11
The country expels ( $R$ ) or defects ( $U$ , $K$ and $E$ )	0	0	0	0	0

<sup>1</sup>For  $R_c$  one third of this figure

<sup>2</sup> $R_r$  is assumed not to transgress

The valuation of keeping old ties also differs between  $R_r$  and  $R_c$ . This aspect is essentially connected to the continued Russian export of energy and other resources at low prices.  $R_r$  wanted trade on market conditions, which required liberalization of prices on these products also.  $R_c$ , on the other hand, wanted to satisfy state enterprise managers reluctant to change their behavior, or simply seeking rents, by maintaining Soviet-type trade. Hence, payoffs are set to 0 and 2, respectively. The valuations of  $U$ ,  $K$  and  $E$  depend on the prices of resources. If  $R$  is reformist, there is no subsidization and the valuation is set to 0. If  $R$  is conservative, the valuation depends on how dependent a particular type is on continued Soviet-type trade, as shown in the table.

As regards inflation taxation,  $R_r$  would not transgress because of its ambition to stabilize, while  $R_c$  could actually gain from unilateral inflation taxation of its partner in each of the bilateral games. This reflects the fact that conservative leaders maximizing rents for their supporters are likely to rely on seigniorage for financing some or even most of the rents. Such a government would be willing to raise seigniorage from citizens in  $U$ ,  $K$ , and  $E$  also. The 0 for  $E$  reflects the fact that stabilization was vital for countries

belonging to this group. The losses if the partner country of each bilateral game transgresses have been inflated for  $U$ ,  $K$  and  $E$  in comparison with the previous game. This reflects bilateralism: there is no other type to share the burden of the partner country's inflation tax with, since each game is solved separately.<sup>50</sup> The loss to  $R_r$  is greater than that of  $R_c$ .

The loss of monetary control to  $R$  is a new feature of the restricted game. Compliance with the restrictions set up by  $R$ , which would be required by  $R_r$ , entails less discretion for  $U$ ,  $K$  and  $E$ . Similarly, if  $N$  selects  $R_c$ ,  $U$ ,  $K$  and  $E$  would not be able to stabilize their economies. In other words, membership in the currency union would mean that inflation rates in member countries and the amount of inflation taxation of Russia would be determined by chance. The -2 for  $E$  reflects a strong desire to move away from Russia's influence, whereas the 0 for  $K$  is explained by the ambition of this group to stay close to Russia.

The payoffs in the lower part of Table 5 show that the strategy transgress dominates over cooperate for  $U$  and  $K$ , and weakly dominates over cooperate for  $R_c$ . Once the strategy cooperate has been eliminated, remaining in the currency union and playing transgress dominates over the strategy defect for  $U$  and  $K$ . Hence, no matter what  $p$  is,  $U$  and  $K$  will remain and transgress. For type  $E$ , defect is a dominant strategy. For  $R_c$ , expel is a dominated strategy, but for  $R_r$ , expel is a dominant strategy after the elimination of bilateral cooperation.<sup>51</sup> Hence, it seems as if the type that set up the restrictions and made continued play of the game possible would have been better off by expelling the others.

In other words, the Nash equilibrium of the one-shot game is for  $R_r$  to expel  $U$ ,  $K$  and  $E$ . But this was not the choice of the Russian government. As suggested above, alternative explanations as to why the reformers would set up such a system are, first, that they hoped to support bilateral cooperation through a punishment strategy, and second, that the new mechanism was a step towards dissolution of the ruble zone. We now go on to see why a punishment strategy could not have secured a cooperative outcome, and accordingly that the former alternative could be dismissed.

### *C. Absence of Renegotiation Proofness*

At first glance, there seems to have been a case for a punishment strategy as a way of establishing cooperation. A cooperative equilibrium, that is, a well-functioning currency union, would be preferred over national currencies by all types except  $E$ . If the threat is that countries that transgress are expelled, cooperation could be secured if the long-run disadvantage of transgressing outweighs the short-run benefits of levying inflation taxes on Russia.

<sup>50</sup> Notice the relation between values for a) and b) under inflation taxation for  $U$ ,  $K$  and  $E$ . Since  $U$  does not care about inflation, the value in b) is the same as that in a) but with opposite signs. For  $K$  and  $E$ , the negative value of the resulting inflation has in each case been subtracted from 5 and -5, the valuations of the real flows.

<sup>51</sup> Since cooperation is a dominated strategy for  $U$ ,  $K$  and  $E$ , it can be eliminated.

Four main factors determine the likelihood that cooperation can be achieved by means of a trigger mechanism. The first is given by the relations between payoffs from transgressing and cooperating. The second is the length of the relationship. The longer it is, the larger is the weight of the punishment phase, and the less profitable it is to transgress. The third is the patience of players. The less they discount future payoffs, the more likely it is that they will cooperate. Likewise, the faster and more accurately transgression can be detected, the easier it is to support cooperation since the punishment phase is more likely to start early.<sup>52</sup>

By comparing the situation in the ruble zone with this list, we see that the prospects for supporting cooperation were bleak. In our numerical example, the relative sizes of payoffs imply that discount factors need to be rather high to sustain cooperation.<sup>53</sup> In the highly unstable environment in, for instance, Ukraine, this is not likely to have been the case. As regards the speed and accuracy with which transgression could be detected, the balances on the correspondent accounts were checked on a daily basis. This does not, however, mean that the CBR could get a clear picture of the current monetary and fiscal policies of member countries since balances were affected with a lag.<sup>54</sup> Furthermore, factors such as terms of trade and productivity shocks were also important for the demand and supply, and therefore for trade balances. Hence, the CBR could not be certain that a country had transgressed in terms of issuing excessive credits just because it had exceeded the limit for its trade deficit.

Finally, the length of the relationship was uncertain, at least for a relationship of this form. There was no guarantee that a country which cooperated would be allowed to remain in the ruble zone indefinitely. Nor was there a credible commitment from Russia's side that a transgressor would be punished indefinitely. Indeed, the possibility that conservative politicians could dominate Russia made indefinite punishment, that is, lasting exclusion from the ruble zone, less likely. This, in turn, made it even more likely that the other members would transgress. As we have already seen, the outcome of the game was that countries which remained in the ruble zone continued to issue excessive credits if they were of type *U* or *K* and continuously exhausted their credit limits. Russia, in turn, repeatedly renegotiated credit limits.

The fact that limits on technical credits were repeatedly increased shows that the revised system was not renegotiation proof.<sup>55</sup> If we concentrate on *R*, in the model, renegotiation of exhausted credit limits made it possible for *R*, to be better off in the subgame consisting of that particular period, since it would be possible to reach a cooperative equilibrium. But the possibility of renegotiation makes the punishment

---

<sup>52</sup> For details, see Fudenberg and Tirole (1991, ch. 4 and 5).

<sup>53</sup> *U* and *K* can gain at least 6 in the first period by transgressing and then receive 0 in future periods instead of 2 in each period. If we assume an infinitely repeated game, the inequality that has to be fulfilled in order to sustain cooperation is  $6 \leq 2 \sum_{n=0}^{\infty} \delta^n$ , where  $\delta$  is the discount factor. This yields  $\delta \geq 2/3$ .

<sup>54</sup> As mentioned above, inflation gained pace three to four months after a credit expansion in Russia. There is little reason to expect that lags would differ considerably in the other member countries.

<sup>55</sup> See Fudenberg and Tirole (1991, p. 176 f.) for a definition and discussion of this concept.



strategy ineffective since players can repeatedly transgress and then be readmitted.<sup>56</sup> In other words, the punishment equilibrium is not subgame perfect.<sup>57</sup> *U* and *K* will use the possibility to renegotiate to the detriment of the overall outcome. They will continue to issue excessive credits, and hope for increased technical credits. But this should be realized by *R*. Hence, a reformist leadership should not have set up such a system, at least not if the goal was to secure cooperation.

This result is further strengthened if we allow for a conservative type of *R*. With payoffs as in Table 5, *R<sub>c</sub>* is not worse off in a period where players transgress than during cooperation. Hence, it has no incentive to try to maintain a cooperative equilibrium and will therefore not expel a country that has transgressed. For the same reason, it will certainly readmit, or renegotiate the credit limits of, a country that has been expelled in a previous period by *R*. The structure of the decisions made by *R* and *U* in this repeated game is shown in Figure 4 in the Appendix.

Since renewed technical credits would be a feature of the system, the optimal strategy of non-Russian member states other than the Baltic countries was to continue to transgress. They should, in other words, go on issuing excessive credits and therefore continue to run deficits on the correspondent accounts.<sup>58</sup> As regards the state trading system, the non-Russian governments had two alternatives. The first was to force enterprises on their own soil to stick to trade agreements and deliver the goods—and make sure that they received payment, that is, to require payment in advance. In that case, the trade deficit would be within credit limits. This alternative corresponds to cooperation and can therefore be eliminated. The other option was to do nothing. The country would benefit from exports to third countries or from using the products on its own, while at the same time being able to import goods from Russia and paying with additional technical credits extended by the CBR.

The possibility to repeatedly transgress cemented imbalances in the centralized trade, while favorable conditions for market-based trade could not develop because of high inflation. The uncertainty about Russia's type made transgress even more likely. In our example, if *U* or *K* chose to cooperate, there was a risk that *R<sub>c</sub>* would be in power in the next period. Since *R<sub>c</sub>* would transgress, *U* and *K* would actually lose from membership if they did not transgress. With *R<sub>c</sub>* in power, credits would float freely and members who pursued relatively moderate policies would lose. The bilateral character of the restricted ruble zone would then have been replaced by multilateral attempts at free riding.<sup>59</sup>

---

<sup>56</sup> The remaining possibility to end up at a cooperative equilibrium was to have an outside agent enforcing these strategies. The IMF offered to do so, but given the institutional constraints and political mistrust, this strategy was unrealistic.

<sup>57</sup> For the equilibrium to be subgame perfect, there must not be an identifiable last period. If there is, transgress will be the dominant strategy in the last period and therefore in the second last and so on.

<sup>58</sup> Note that the assumed discrete nature of the decisions on credit issuing in the former system has in this system become explicit by the limits set for technical credits.

<sup>59</sup> In that case, hyperinflation would have developed. One of two things could then have happened. The first would be that reformers in Russia or other member countries had strengthened their positions, possibly enough to reform the ruble zone mechanism further or, in the case of non-Russian members, to introduce independent currencies. The alternative scenario would be a reversal of economic and democratic reforms

The game theoretic analysis of the restricted ruble zone shows that any system with the same main features would give weak incentives to restrict monetary and fiscal policies in order to keep deficits within limits. Hence, the previously mentioned inconvertibility of rubles as credit limits had been reached was not the reason why the system broke down. It was a direct consequence of the inherent weakness of a system with several central banks, all with the ability to issue credits valid as payment within the monetary area and renegotiable limits on the bilateral balance of payments deficits. Likewise, the fact that state trading was maintained did not affect the qualitative aspects of the outcome. No matter how the specific mechanisms had been designed, the ruble zone would have had to be dissolved because of the general institutional weakness of the centralized system.

We now have an explanation as to why inflation continued to rage within the ruble zone after restrictions had been imposed. We have also seen that a rational Russian government could not have hoped to secure cooperation in the restricted ruble zone. The new lines of credits that it set up in June 1992 were certain to be exhausted, and any further credits on the same conditions would also be exhausted. This result leaves us with the second alternative mentioned above as an explanation as to why the government set up restrictions rather than dissolved the ruble zone immediately. In other words, the restrictions were set up as a mechanism to make exclusion of other members possible. The way it would work would be to make the costs to Russia of maintaining the ruble zone clear, so that the reformers could gain enough support to nationalize the ruble.

---

and perhaps even resurrection of the Soviet Union. In other words, Russia would have been back to a situation reminiscent of that in the fall of 1991.

## **5. POLICIES OF MEMBER COUNTRIES**

### **5.1 Renegotiation of Credit Limits**

The game theoretic analysis again yields a close fit with actual developments in the ruble zone. The case of Ukraine serves well as an example. Payments between Russia and Ukraine were cut shortly after the introduction of the new system since Ukraine exhausted the new credit line amounting to Rbs 10 bn that was set up by the Russian government in only one week (Conway, 1995, p. 32). However, the directors of Russia's and Ukraine's central banks agreed on Rbs 100 bn in new credits on September 10, 1992, which made it possible to continue payments. In this renegotiation, Viktor Gerashchenko represented Russia. On September 21, the Russian government broke the agreement. (Bornefalk, 1992).

By the logic suggested above, we could argue that the reformist faction of the Russian government succeeded in pushing through this decision since they had increased the transparency of the payments system. Since the moment that agreement was broken, Ukraine has essentially stayed outside the ruble zone. Ukraine was excluded from the ruble zone because of its clear and repeated signals that it did not want to remain under Russia's wings and did not intend to follow agreements. In other words, Ukraine's disastrous monetary policies hurt its reputation to such an extent that Russia saw no alternative other than to exclude it.

The reformist faction was not, however, strong enough to stop the issuing of technical credits to Ukraine. The total amount Ukraine received for the second half of 1992 was Rbs 295 bn, substantially more than the previously fixed limits for the whole ruble zone of Rbs 215 bn. For the first half of this period, Ukraine received Rbs 168 bn, second only to Kazakhstan with Rbs 219 bn. In any event, the fact that even Ukraine's credit limits were repeatedly increased shows that the punishment of being excluded from trade and trade credits envisaged by the new system was not credible, since it was not renegotiation proof. In addition to renewed credits, Ukraine could continue to import oil and gas under favorable conditions, such as not having to pay for deliveries. Hence, the restricted ruble zone gave small, if any, incentives for non-Russian member countries to restrict monetary and fiscal policies.

### **5.2 Introduction of Currency Boards**

The trigger mechanism and the arrangements to facilitate decentralized payments between Russia and the other FSRs should have been interpreted as elements of a strategy to overcome resistance against breaking up the ruble zone by proceeding gradually. Despite this, the IMF continued to oppose the introduction of national currencies. Estonia had already decided to leave the ruble zone as the new mechanism was presented. Just prior to

the introduction of the kroon, the IMF decided to give Estonia its official support in order not to destroy the prospects for a successful monetary reform.

Latvia and Lithuania had to face considerable pressure from the IMF when they were considering whether to follow Estonia's initiative. The IMF even resisted the initiative of Einars Repse, governor of the Bank of Latvia, to form a currency exchange in Riga where the various account rubles could be traded. The currency exchange was needed to facilitate payments between member states of the ruble zone, since account rubles issued in different member states essentially became separate currencies shortly after the introduction of limits on technical credits. The payoffs of type *E* in the game theoretic analysis suggest that these countries suffered considerable losses from staying, mainly because of delayed stabilization.

The fact that Estonia, and later Lithuania, chose to introduce a currency board to support its new currency is noteworthy. An analysis of constitutional choice under different political conditions by Bornefalk (2000c) could help us understand why these countries made this decision. Again concentrating on Estonia, the adoption of the currency board coincided in time with the adoption of the new constitution, which, as a central feature, imposes a legal obligation on the government to refrain from budget deficits. These constitutional choices can be interpreted as measures aimed at reducing the pressure on the government and the parliament from interest groups seeking favors, by reducing the ability of politicians to extend such favors. This, in turn, would facilitate macroeconomic stabilization and make interest groups more interested in using their resources for productive purposes.

The effect of this type of reform on the extent to which activities aimed at affecting the distribution of income take place can be studied by means of the model of Bornefalk (2000c). It is shown that in countries where implicit restrictions against redistribution are strong enough from the outset, strengthened explicit, or constitutional, restrictions against redistribution improve economic performance in terms of the share of their resources that interest groups devote to production.<sup>60</sup> Unless the constitutional assembly is overly biased against organized groups, it will adopt a constitution that strengthens the explicit restrictions against redistribution. The adoption of a currency board is clearly an example of this kind of reform, since the obligation to exchange the reserve currency for the domestic currency at the fixed exchange rate can only be withdrawn by means of a new parliamentary decision. Given the fairly well developed democratic institutions and the relatively weak special interest groups in Estonia, it is likely that Estonia fulfilled these criteria for a successful adoption of a currency board.<sup>61</sup>

It is not surprising that, for instance, Russia and Ukraine did not choose to introduce

---

<sup>60</sup> In the terminology of Bornefalk (2000c), implicit restrictions against redistribution determine how costly it is to influence the voting of individual policy makers, whereas explicit restrictions determine the number of policy makers who need to be influenced to affect a decision. The maturity of political parties and the freedom and impact of mass media are two major determinants of the implicit restrictions.

<sup>61</sup> Political institutions have received little attention in the literature on currency boards. See Williamson (1995) for a list of other determinants of the effects of a currency board.

currency boards since this would run counter to the interests of powerful interest groups. The analysis of Bornefalk (2000c) does, however, indicate that an attempt to introduce a currency board could actually have increased appropriative activities, thereby worsening the economic situation. This is shown to be the consequence of strengthened explicit restrictions against redistribution when the implicit restrictions against redistribution are not strong enough. Such an attempt could, for instance, have led to massive protests aimed at blocking the introduction.

There are other factors that make currency boards more suitable for the Baltic countries than for Russia and Ukraine. As pointed out by Williamson (1995, p. 23), the large size, which translates into a low degree of openness, of the latter countries makes a currency board more costly and less beneficial than it would be for small, open countries.

### **5.3 Emergence of Hyperinflation**

The mechanism designed to put an end to the ruble zone, together with the uncertainty about the future political situation in Russia and the actions of the IMF, contributed to pushing the non-Russian FSRs that had not left the ruble zone into hyperinflation. Since the IMF blocked currency reform in countries that lacked sufficient know-how and national consent to introduce a national currency on their own, the best policy these countries could pursue was to remain in the ruble zone and play the dominant strategy, that is, to transgress.

Decision-makers in these countries probably realized the risk that reformers in Russia would strengthen their position and once again change the mechanisms of the ruble zone in a more restrictive direction. They should also have realized that hyperinflation threatened to develop, which would result in falling seigniorage and therefore rents. Hyperinflation could have developed either because a reformist Russia might eventually go so far as to exclude them, or because a conservative Russia would join the ruble rally. More reform-minded countries such as Latvia and Lithuania managed to escape hyperinflation by balancing their budgets, staying outside the ruble rally, and leaving as soon as possible.

Table 6 reports monthly inflation rates for countries that were expelled from the ruble zone. There is a clear connection between the introduction of national currencies and the emergence of hyperinflation in these countries. The squares in the table mark the introduction of national currencies not parallel to the Russian ruble. Defining hyperinflation as a monthly rate of inflation exceeding 50 percent, we see that all countries that were expelled from the ruble zone experienced hyperinflation within three months after their exclusion.<sup>62</sup>

The model of Cagan (1956) can be used to explain why inflation is bound to increase in a country that is excluded from a highly inflationary monetary union after having shifted over inflation taxes on other member countries. When the country is

---

<sup>62</sup> Data for Armenia are missing.

excluded, the rate of money growth will jump further up, unless monetary creation is reduced enough to neutralize the fall in the inflation tax base that results from the exclusion. Increasing money growth will increase inflation, which will lower real money holdings in the economy, that is, the remaining inflation tax base. This, in turn, will require an even more rapid rate of money growth to keep up seigniorage, which we assume is the objective. Eventually, this spiral will end in hyperinflation.

Inertia of the kind derived by Alesina and Drazen (1991) could explain why a sufficient reduction of the monetary expansion is unlikely to come about at an early stage. When different groups try to shift the burden of stabilization on to each other, stabilization becomes a “war of attrition” and is stalled until one group concedes and bears a disproportionate share of the burden. The weakness of the state and strength of industrial pressure groups in the FSRs make this a plausible explanation in our case.

**Table 6. Introduction of National Currencies<sup>1</sup> and Emergence of Hyperinflation<sup>2</sup>**

1992-93	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Ukraine '92	249	20	12	7	13	19	20	15	14	19	20	30
'93	100	27	19	19	25	79	21	22	80	55	45	91
1993	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Belarus	29	34	36	24	16	19	18	22	35	45	46	49
Moldova	37	28	25	20	18	19	34	35	64	38	38	59
Kazakhstan	33	32	33	21	16	18	22	29	29	38	56	34
Turkmenist.	15	141	15	28	14	56	15	16	22	28	429	26
Uzbekistan	30	20	7	15	10	53	17	24	7	30	60	57
Azerbaijan <sup>3</sup>	42	15	1	5	19	10	17	9	14	17	49	59
Georgia	36	30	17	17	40	35	29	19	53	102	115	96

Source: Karlsson (1994, p. 3).

1. Squares mark the month of introduction of a national currency.

2. Inflation measured as monthly percentage change in CPI or retail prices.

3. Introduced a national currency in January 1994.

In addition to the effects of inertia, dependence on products that are cheaper within the monetary union might also push a country into hyperinflation as it is excluded. A country that used to benefit from subsidies in the form of, for instance, cheap imports of energy will have to adapt to that loss. If it tries to adapt by extracting more seigniorage from its own economy, seigniorage needs will surely become unsustainable. This latter mechanism was pointed out by Goldberg et al. (1994, footnote no. 29).

## 5.4 Concluding Remarks

Our analysis has shown that the behavior of non-Russian member states can be explained without resorting to explanations based on lack of knowledge or irrationality. The relatively low level of development reached by democratic institutions in all but the Baltic countries, and the degree of cooperation between different groups within individual countries, suggest that the high levels of credit expansion and inflation in these countries were parts of a policy that benefited interest groups.

The fallacies of the ruble zone mechanism, the relative size of the economies of different member countries, and the poor level of development of democratic institutions in these countries made high inflation unavoidable. Extraction and distribution of seigniorage to organized groups through the ruble zone was an effective way of achieving the desired level of redistribution. In the Baltic countries, on the other hand, these factors made politicians little interested in achieving redistribution. They therefore refrained from extending credits to firms, which in turn could not enrich themselves at the expense of unorganized groups throughout the ruble zone.

However, it is still difficult to understand the behavior of Russian reformers. The game theoretic analysis suggests that they should have expelled the non-Russian member countries immediately in June 1992 instead of choosing a gradual and costly strategy. An examination of the factors pointed out in Bornefalk (2000b) does, however, indicate that the available reform space was not wide enough for a radical reform, which placed severe restrictions on the ability of interest groups to acquire rents through lax monetary policies. In the next chapter, we will see how the reformers managed to delimit the currency area by utilizing the available reform space to implement the first steps of the gradual monetary reform, and at the same time strove to increase the reform space. They did this by strengthening the democratic institutions, which made further reforms possible. The reforms, in turn, divided the interests of powerful industries, which eventually made a complete delimitation possible.

## **6. THE POLITICAL ECONOMY OF RUSSIA'S MONETARY REFORM**

We now change our focus from the strategic interaction between member states of the ruble zone to the struggle between reformers and conservatives within Russia over monetary reform. Section 6.1 takes a closer look at why state enterprise managers, the most influential interest group, were interested in maintaining the ruble zone. It also analyzes how the industrial lobby could have such an influence on the process. In Section 6.2, we study how the government managed to overcome resistance against a break-up by implementing a gradual strategy. The account also shows how the available economic reform space was affected by changes in political conditions. Section 6.3 discusses whether it would have been possible to undertake a more radical dissolution of the ruble zone. Based on the constraints faced by the reformers, Section 6.4 concludes that the Russian monetary reform was quite an achievement despite the huge costs brought about by its delay.

### **6.1 Strength and Interests of State Enterprise Managers**

State enterprise managers, particularly in heavy industry, opposed a radical change of economic system for three major reasons. The first was that the reformers' ultimate aim was to transform Russia from an economy aiming at maximizing its military strength to an economy where consumer goods and services had higher priority and where private small and medium sized enterprises produced a large share of output. As part of this goal, Russia would give up the remnants of its empire, that is, the non-Russian FSRs, and seek to become integrated in the world economy. Given the industrial structure and trade patterns as described in Chapter 2, this would necessarily hurt heavy industry as a group.

The second reason to resist economic reform has to do with uncertainty regarding the distribution of gains and losses from the change of economic system. While it was reasonably clear which sectors had much to lose and which stood to gain, there was uncertainty regarding gains and losses for enterprises within different sectors and even more so for managers. Fernandez and Rodrik (1991) have shown how this kind of uncertainty creates a bias towards the status quo and therefore against economic reform.

The third reason for state enterprise managers to resist economic reform has to do with the political and economic conditions treated in Section 3.2. There was a growing lack of balance between the influence of the industrial lobby and that of unorganized groups during 1992. State enterprise managers were well organized and well connected, and managed to establish a high degree of cooperation with each other. Democratic institutions, on the other hand, were little developed. As a consequence, unorganized groups had little influence over decision makers. This made state enterprise managers relatively more interested in achieving redistribution from the rest of the economy than in



contributing to production. They therefore opposed reforms that would increase the protection of property rights and reduce the ability of politicians to grant subsidies.<sup>63</sup>

The ruble zone gave abundant opportunities for rent seeking. The opaqueness of the payments system made it difficult for mass media and the public to observe the extent of redistributive activities.<sup>64</sup> This made appropriative effectiveness particularly high in this area. The opportunities for extracting rents were the fundamental reason why state enterprise managers were in favor of delaying the dissolution of the ruble zone, and thereby the termination of Russia's subsidization of bilateral balance of payments deficits. To delay, and possibly even block, a break-up of the ruble zone was also an effective means to overcome the need for rapid adaptation to market conditions, since it allowed Russian enterprises to export products to enterprises in other FSRs which otherwise would often have been difficult to sell. Maintaining the ruble zone for the time being also gave state enterprise managers time to try to determine whether or not they would belong to the beneficiaries of the proposed reform.

State enterprise managers were quite successful in delaying the dissolution of the ruble zone. Olson (1994) argues that the power of the industrial lobby during the transition period was an inheritance from Soviet times. Josef Stalin's totalitarian rule gave him an encompassing stake in society since he could use most of an increase in the economy's productivity for his own goals. But when Stalin's purges ended, enterprise managers could take advantage of their superior information about the productive processes they managed, with the goal of making plan fulfillment easier. If all factory managers in an industry cooperated in lowering their superior's expectations of productivity, they would all gain. This could be achieved by avoiding competition between different enterprises. In fact, the gigantic bureaucratic establishments that developed were ideal in this sense.

Eventually, the collusion created powerful insider lobbies at the expense of the center. The vested interests had no encompassing stake in society and therefore did not care about society's economic development as a whole. This situation carried over to the new economic and political conditions, although the new democratic governments were stronger than the last communist governments. After the introduction of democratic rights, however, vested interests could lobby more openly, and workers—who were still under considerable influence from their managers—could go on strike more easily. Moreover, the reformers were heavily dependent on exports, particularly from the energy sector, to limit the balance of payments deficit, and tax receipts, either in money or in kind, to keep

---

<sup>63</sup> State enterprise managers also opposed a rapid stabilization. The explanation for this coincides, to a great extent, with the explanation for their resistance against economic reform in general. Political and economic conditions determining the extent to which interest groups choose to engage in appropriative activities also affect the rate of inflation, if monetary conditions are such that rents are easily extracted by means of subsidized credits and financing of subsidies to these groups by printing money. The mechanism pointed out by Alesina and Drazen (1991), and discussed in Section 5.3, is also important in understanding why interest groups might prefer to delay stabilization.

<sup>64</sup> See Åslund (1995, p. 124) for a description of how the centralized payments system of the ruble zone made monitoring of the balance of payments with other member countries impossible and how payments were delayed.

down the budget deficit.

The relatively low level of development reached by democratic institutions and the power of state enterprise managers meant that appropriative effectiveness was relatively high in Russia following the break-up of the Soviet Union.<sup>65</sup> According to the analysis of Bornefalk (2000b), under such conditions organized groups will use a large part of their resources to affect the distribution of income. The extent of redistribution will therefore also be relatively large. As a consequence, total production will fall, and the distribution of income will be skewed to the advantage of organized groups. This pattern fits well with the development in post-soviet Russia. Massive rent seeking contributed to the drastic fall in production and the increasing differences in income between different groups. A large share of these rents emanated from the ruble zone mechanism.

But if the ruble zone brought such harm to the Russian economy, how could state enterprise managers support it? Bornefalk (ibid.) shows that as long as organized groups are able to cooperate with each other, they will never lose by engaging in appropriative activities. They will gain when appropriative effectiveness is high enough. The explanation for this is that in the cooperative case they are able to incorporate negative effects of their behavior on total production. Thus the monetary mechanism of the ruble zone suited the interests of state enterprise managers quite well.

When interest groups are unable to cooperate, they will use an even larger share of their resources for appropriative activities, *ceteris paribus*. However, there is now a considerable risk that they will suffer from engaging in appropriative activities, compared with a situation in which they could credibly commit to refraining from such activities. This will come about if the share of total resources that they control is relatively large and the degree of appropriative effectiveness is relatively low. Under such conditions, they will prefer a system that supports productive rather than appropriative activities. Hence, they would be more interested in a payments system that facilitated exchange and access to new markets. This suggests that it could have been possible to undertake a radical dissolution of the ruble zone before state enterprise managers had managed to resurrect a high degree of cooperation in the spring of 1992.<sup>66</sup>

---

<sup>65</sup> It is, however, unlikely that appropriative effectiveness was as high as it had been during the Soviet period, when the economic system was geared towards extracting resources from the civilian sector for use in heavy industry, with the explicit aim to maximize military strength.

<sup>66</sup> Note that the share of resources controlled by interest groups affects their attitude towards reform only in the non-cooperative case in the analysis of Bornefalk (2000b). Olson (1965) does not distinguish between the cooperative and non-cooperative cases. According to his analysis, a group that controls a comparably small share of society's resources is more likely to act at the expense of society as a whole. The reason is that such a group is more likely to be able to increase its own share in the economy enough to offset the effect of the overall decline. Bornefalk (ibid.) also distinguishes between the share of initial resources controlled by a certain group and the share of total production secured by this group.

## 6.2 The Struggle over Russian Monetary Reform

### *A. Resistance to Reform*

State enterprise managers were represented by a group of politicians referred to as “industrialists.” The most influential of these was Arkady Volsky. The industrialists lobbied for soft credit policies, subsidies for industry, and a general retreat from the attempt at shock therapy launched on January 2, 1992 (Hanson and Teague, 1992). By joining forces with conservative politicians interested in maintaining Russia’s influence over the FSRs, the industrialists increased aggregate uncertainty about the goal of transition.<sup>67</sup> Was Russia going to try to keep some of its empire, and if so, how much would that be allowed to cost? To keep up Soviet-type trade and production by maintaining the ruble zone was an effective way of securing Russia’s influence in its neighboring countries. It was, however, costly since this entailed continued subsidization of the other FSRs.

The industrialists tried to circumvent stabilization in general and a dissolution of the ruble zone in particular by using the parliament and the Congress of People’s Deputies, where representatives from industry and other conservatives had a clear majority. To try to secure control over the CBR was an important part of this strategy. On November 22, 1991, the parliament stopped an attempt by the government to take control over the CBR (Åslund, 1995, p. 97). This was before the radical reforms were launched, and was a major impediment to stabilization since the CBR remained under pressure to keep up production by means of monetary injections, as many members of parliament had a personal interest in this.

When the industrialists had gained enough strength, they went for an outright confrontation with the government in the spring of 1992. As mentioned in Section 4.2, the industrialists managed to replace a number of reformers in the government with representatives of state enterprise managers. This meant that the reformers lost what was left of their agenda setting power. Since the industrialists also succeeded in replacing the moderate Matyukhin with Gerashchenko as head of the CBR, they had managed to block stabilization for the time being. The reformers did, however, manage to impose restrictions on the trade credits available to other member states.

The new priorities of the government and the ambitions of the CBR under Gerashchenko had a crucial impact on monetary policy in the restricted ruble zone. Indeed, Gerashchenko regarded himself as a central banker for the entire ruble area and aimed at supporting production. He also argued that issuing cheap credits to enterprises could cure inflation, since the credits would promote production and thereby limit the gap between supply and demand. Moreover, Gerashchenko expressed more concern over unemployment, which was still quite low, than over inflation, which was rampant. These

---

<sup>67</sup> In the terminology of Dewatripont and Roland (1996), aggregate uncertainty of outcomes refers to uncertainty about the goal of transition or, if there is a clear goal, how to get there.

statements are hard to interpret in any other way than that Gerashchenko provided industrialists with a veil behind which they could maximize rents. Since there was no legislation preventing the CBR from increasing the limits for technical credits, the ruble zone was an excellent tool to reach this objective.

Hence, in mid 1992, state enterprise managers were well organized, had great incentives to postpone the dissolution of the ruble zone, and had established powerful coalitions against reform in all major state bodies. The following subsection shows how the government in the end managed to overcome the resistance against a break-up of the ruble zone by simultaneously undertaking political reforms, thereby widening the available economic reform space.

### *B. Nationalization of the Russian Ruble*

The Congress of People's Deputies in December 1992 led to further changes among top decision makers that initially seemed to strengthen the proponents of the ruble zone even more. Yegor Gaidar had to go and Viktor Chernomyrdin was promoted to prime minister. Chernomyrdin had strong support from the industrial lobby, and was a supporter of the ruble zone. After a rapid rise of inflation following Gaidar's dismissal, the reformers managed to persuade Yeltsin and the new prime minister to appoint Boris Fedorov as deputy prime minister and minister of finance (Åslund, 1995, p. 193). Fedorov became the main opponent of the ruble zone, and proved to be a quite effective reformer.

In the spring of 1993, Russian policies on trade credits began to tighten considerably as part of an ambitious attempt at macroeconomic stabilization. The reformers managed to put an end to the subsidization of other FSRs through cheap credits by persuading the parliament to join the government in abolishing technical credits, by a decision on April 20, 1993. Instead, tied credits for the purchase of specific Russian goods were introduced, and all previous credits to the FSRs were transformed into state debts with an interest rate tied to LIBOR (London Inter-Bank Offering Rate). The detrimental role of the CBR was ended by making these credits available only with the approval of the Ministry of Finance. The account rubles of the FSRs had thereby been effectively separated.

The reformers used the result of a referendum, held on April 25, 1993, to strengthen their position relative to the CBR and conservatives in the government. The referendum approved economic reform policies, which enabled the reformers to push through a joint declaration of the government and the CBR that established quarterly credit ceilings. The CBR was also forced to raise the highly negative refinance rate. These conditions opened up the way for an agreement with the IMF on a Systemic Transformation Facility.

The failure of the CBR to stop the fall in GDP by generous credit emissions also contributed to strengthening the reformers. Rather than bringing any advantages, the costs of the resulting inflation became clear to a larger number of decision makers. Inflation also came to be identified as the main problem in opinion polls. Thus the malfunctioning

of the ruble zone with its huge transaction costs<sup>68</sup> influenced the process leading to the overall tightening of monetary and fiscal policies, as argued by Bornefalk (1995). The increased transparency brought about by the limits for technical credits in the currency reform of July 1992 played an important part in this development. Not only did decision makers become more aware of the costs incurred through the ruble zone, but the greater transparency also made it more difficult for them to extend subsidies to industrialists.

After technical credits had been abolished, the CBR managed to increase the deliveries of bank notes to other FSRs sharply, to keep up trade. Therefore, the remaining task to achieve a national Russian currency was the separation of Russia's cash rubles from those of other FSRs. This was accomplished through the monetary reform of July 24, 1993, when pre-1993 bank notes—the only ones available in other FSRs—were withdrawn from circulation in Russia.<sup>69</sup> The currency reform was executed in a way that was reminiscent of failed attempts at currency reform towards the end of the Soviet Union. The reform was undertaken by the CBR with approval from Yeltsin and Chernomyrdin, while the reformers were not informed. The conservative parliament strongly criticized the unpopular decision and blamed the government and presidency. According to Åslund (1995, p. 130), the conservatives were therefore strengthened in the aftermath of the currency reform.

Though damaging to Russia and the credibility of the reformers, the invalidation of old ruble notes in Russia was most damaging for the countries that still used them in their own territories. After the withdrawal of Soviet bank notes, the FSRs that had not introduced their own currencies had to choose between doing so or opening negotiations with Russia for a “new-style ruble zone” based on the Russian ruble.<sup>70</sup> Azerbaijan, Georgia, and Moldova, which used parallel coupons, chose the former alternative, as did

---

<sup>68</sup> Dabrowski (1995) argues that low transaction costs were the only advantage of the ruble zone. He includes in his evaluation costs of the exchange operation as such, the exchange rate risk which arises with flexible exchange rates, and costs due to separate currencies not being convertible. But this interpretation of transaction costs is applicable only in a limited sense, namely for transactions that did actually take place, and even then only when payments were cleared without long delays.

Transaction costs as perceived by Coase (1937) are much broader. They include search and information costs, bargaining and decision costs, and policing and enforcement costs. It is easy to conceive how the opaqueness of the ruble zone boosted transaction costs in this sense. For instance, the long arrears and centralization of payments made detection of non-payments more difficult and enforcement therefore more costly. Also, the long arrears and the high inflation led to great uncertainty about the value of a potential transaction. The inconvertibility of various account rubles that emerged in the restricted ruble zone was another source of transaction costs. If we include the value of lost opportunities from transactions that were never undertaken because of uncertainty or low liquidity, and if we take into consideration that the ruble zone contributed to the high levels of inflation, it is clear that transaction costs emanating from the ruble zone were huge. These truly excessive transaction costs hindered the development of market-based trade, and also hurt state controlled trade. They thus made the fall in trade unnecessarily great.

One can also argue that Russia's unintended subsidization of the other members should be included in transaction costs, which thereby become exorbitant. The reason is that they stemmed from a lack of information on the other member countries' future behavior. The costs of the search, contracting and enforcement necessary to set up a functioning currency union under the circumstances that prevailed in the FSU were practically insurmountable. Russia did not take on more than a fraction of these costs, and had to pay for that neglect.

<sup>69</sup> See Fedorov and Kazmin (1994) for details.

<sup>70</sup> For a detailed account of negotiations on the new style ruble zone, see Granville (1995, p.75 ff).

Turkmenistan. For reasons discussed in the previous chapter, they were immediately struck by hyperinflation.<sup>71</sup>

Countries more interested in continued cooperation with Russia, that is, Kazakhstan, Uzbekistan, Tajikistan, Belarus, and Armenia, settled for the latter option and signed a far-reaching agreement with Russia on September 7, 1993. The agreement covered the rules for the transition period to a monetary union. According to Granville (1995), it appears to have been a result of a compromise between Russian reformers and conservatives, reflecting a stalemate in the power struggle during this period. Thus, while requirements for prospective members were fairly far-reaching in several areas, the agreement did not specify how future trade deficits with Russia should be financed. The principles of cash and credit emission were also extremely vague, as pointed out by Granville (1995, pp. 75-76).

Boris Yeltsin decided to dissolve the parliament, which had become increasingly dominated by communists aiming to reverse the economic transition, on September 21, 1993. This meant that state enterprise managers lost their main channel for influence, and led to a dramatic strengthening of the reformers' power. With the parliament set aside, Boris Fedorov managed to gather enough support to take the final step away from disorderly monetary arrangements with the other FSRs. The government could therefore sharpen the conditions for future membership to such an extent that it became unattractive, if at all possible, to become a member. But, as noted by Granville (*ibid.* p. 76), the prospective members did not give up their hope of continued subsidies. The Kazakh and Uzbek governments put pressure on Russia to distribute cash in new ruble bank notes. The response of the Russian reformers was to issue tough conditions that had to be fulfilled to receive cash. Since the conditions would exclude the possibility of receiving subsidies from Russia, and since there were substantial punishments tied to any failure to meet the conditions, Kazakhstan and Uzbekistan decided to introduce their own currencies on November 15, 1993. These countries also immediately experienced hyperinflation.

The influence of the reformers fell after their relative failure in the December 1993 parliamentary elections. This gave a last opportunity to Gerashchenko to extend the ruble zone beyond Russia's borders. On January 5, 1994, the CBR director signed an agreement with the chairman of the National Bank of Belarus to admit Belarus on looser conditions than those imposed by the Russian Ministry of Finance. In particular, credits would be given first, and conditions set later (*ibid.* p. 77). However, the reformers were still strong enough to repudiate the agreement.

A final attempt by the conservatives to extend the ruble zone was made on April 12, 1994. On that date, Chernomyrdin signed a monetary union treaty with the Belarusian prime minister Vyacheslav Kebich. As part of the treaty, a customs union was implemented on May 1, 1994. Russian export duties on oil and gas were to be removed, and Russia would not have to pay transit fees for oil and gas pipelines. At the end of 1994,

---

<sup>71</sup> See Table 6.

Russia was also supposed to export energy to Belarus at domestic Russian prices. On the monetary area, Belarusian “zaichiks” were to be exchanged at a rate of 1:1 up to a ceiling. Since the Russian ruble was valued higher than the zaichik, this would be a sheer gift from Russia amounting to about 10 percent of the April 1994 monetary stock (ibid. p. 78).

The treaty required Belarus to give up the independence of its central bank, which divided the Belarusian leadership. Consequently, ratification was postponed until presidential elections had been held in June 1994. The victory of Alexander Lukashenko over Kebich put an end to the interest on the Belarusian side. On the Russian side, the interest in a costly monetary union finally faded when Chernomyrdin, the main proponent of the ruble zone along with Gerashchenko, became less enthusiastic at the very end of the negotiations with Belarus (Dabrowski, 1995, p. 25).

### *C. Development of a New Payments System*

Apart from attempts to resurrect the ruble zone, considerable effort was expended to improve the centralized payments system throughout the gradual dissolution of the currency union. In early 1993, a matrix of bilateral correspondent accounts was almost in place (Eichengreen, 1993, p. 316). However, the payments mechanism did not suffice to finance trade between the FSRs multilaterally because of lack of trust. The January 1993 proposal to found an Interstate Bank for multilateral clearing of trade between the FSRs was a new attempt. According to Eichengreen, it suffered from inadequate credibility of sanctions to be imposed on countries exceeding their credit ceilings. Furthermore, it required a willingness and ability on the part of Russia to take charge of it, and this was absent.<sup>72</sup>

The vast flow of technical credits in the fall of 1992 could not secure effective payments. Trade links therefore broke up further, and a decentralized payments system had to be developed. The pillars of a market-based payments system began to emerge after the Russian currency reform of July 1992. As mentioned in Section 5.2, Einars Repse took a seminal step in August 1992 by opening a currency exchange in Riga where account rubles from different FSRs could be traded. Their value in relation to the Russian account ruble depended on the excess demand for Russian rubles in different FSRs, which in turn was largely determined by the fiscal and monetary policies pursued in those countries. Repse’s initiative was soon followed in other FSRs, but the capacity of currency exchanges was small and access was limited. Moreover, interstate correspondent banking, introduced by the agreement between Russia and Estonia following Estonia’s currency reform, was not generally permitted until 1993. But then, since exchange markets had started to function reasonably well, a fairly effective payments system had

---

<sup>72</sup> According to Åslund (1995, p.127), the Russian reformers resisted its creation since they believed it would diffuse responsibility and diminish transparency in monetary policy, which would increase pressure for more technical credits from Russia. The other FSRs were not willing to integrate their policies with Russia, which would have been required to make the Interstate Bank work. Still, they did not reject the idea immediately since they wanted to maintain the flow of subsidies from Russia as long as possible. The main supporters of

finally been created (IMF, 1994:1).

### 6.3 Radical versus Gradual Dissolution

In late 1991 and early 1992, the advisors of the radical Russian reformers argued for a swift nationalization of the Russian ruble. But the early proposals to dissolve the ruble zone did not acknowledge the complexities of the monetary situation in the FSU, in particular the complementarities between financial and trade related aspects of the break-up. The advisors also failed to realize the degree to which the ruble zone actually threatened Russia's stabilization: it was believed that the stabilization program launched in January 1992 was likely to succeed, although the monetary relations with the FSRs had not been settled. Marek Dabrowski, who belonged to the team of foreign advisors, formulates this as follows (Dabrowski, 1995, pp. 26-27):

... in the first months of the Russian transformation (that is, the end of 1991 and beginning of 1992), the threat associated with this type of monetary system was not very noticeable for the Russian reformers or for the IMF and a number of foreign experts. It became clear only in the spring of 1992.

According to Dabrowski (1995, footnote 16), a memo written by Jeffrey Sachs and David Lipton dated May 1, 1992, was one of the first warnings to the Russian government of the magnitude and complexity of the problem.<sup>73</sup> The memo explains why the breakdown of the Soviet monetary and financial systems, and the destabilizing nature of the ruble zone mechanism that replaced them in the beginning of 1992, made monetary stabilization particularly complex in the former Soviet Union. Sachs and Lipton also argued that there was a clear threat that hyperinflation would develop. This, in turn, would have halted the development of a monetary and financial system that could support an increasing role for market relations.

The radical reformers could not take the risk of implementing radical reforms that were not reasonably coherent. The environment they had to operate in would not have allowed any major mistakes. As we saw in Section 6.2, they had failed to gain control over the CBR and, more generally, to secure support from the parliament for stabilization. The parliament was more attentive to the demands from industrialists than to the economy as a whole. In other words, they faced few restrictions against redistribution and low incentives to promote economic development. Reasons for this are easy to identify. First of all, no elections had been held after the fall of communism in August 1991 to secure public support for the transition to a market economy. Another explanation, offered by Åslund (1995, p. 88), is that the government simply did not try to involve the parliament in the reform process. Parliamentarians therefore felt excluded and started to protest.

A more fundamental explanation, emphasized in Bornefalk (2000b), is that

---

the Interstate Bank were the IMF and the CBR, who saw it as a means to maintain the ruble zone.

<sup>73</sup> The memo was later published as Sachs and Lipton (1993).



democratic political parties had had little time to develop. Most members of the parliament and Congress of People's Deputies represented themselves. Hence, there was little party discipline. But even the political parties that did exist had built up little credibility for pursuing policies that favored voters at large. Many deputies also realized that they would have small chances to be reelected in free elections and therefore had weak incentives to invest in their reputation. For these reasons, they had few incentives to resist pressure and bribes from interest groups. Hence, a radical and ill-prepared dissolution of the ruble zone would have been quite risky to undertake.

The previously mentioned memo by Sachs and Lipton also contained a detailed proposal for how a market-based monetary system could be created. Once the reformers had received this concrete piece of advice, they proceeded with considerable determination. The window of opportunity had, however, been shut by the political changes brought about by the industrialists, at least to the extent that it had ever been open. The limited available reform space meant that an attempt to initiate a radical monetary reform would have unleashed overwhelming resistance. At the same time, the reform had to be far-reaching enough to start the process towards a dissolution of the monetary area. The introduction of limits on technical credits in July 1992 fulfilled these conditions. This was the first step in Russia's gradual monetary reform.

The introduction of limits on technical credits did not immediately reduce the costs of the ruble zone, that is, inflation and costly subsidies from Russia to other FSRs. It did, however, increase transparency, which was necessary to gain enough support for an eventual nationalization of the Russian ruble. As argued in Section 4.2, the ambitions of the CBR under Viktor Gerashchenko together with the power of the industrial lobby led to the limits not being renegotiation proof. This, in turn, led to continued issuing of excessive credits by the central banks of those FSRs that were not prepared or willing to undertake successful currency reforms backed by forceful stabilization policies. Inflation therefore continued to rage throughout the ruble zone. State enterprise managers benefited from this, while the economies as a whole suffered.

An eventual separation of account rubles was unavoidable, but this could not occur until the Russian reformers had strengthened their position. The process became prolonged and costly because of the poorly developed democratic institutions in Russia. But the introduction of limits on technical credits could also have initiated a rapid break-up if the political situation had swung in favor of the reformers. Had the director of the CBR been more interested in stabilization and less influenced by the parliament, or if the parliament itself had been more interested in stabilization, the system with limits on technical credits would have been fairly restrictive. As soon as a non-Russian member state had run out of credits on its correspondent account, the CBR would have had an excuse not to settle payments, which would have forced the country out of the ruble zone. It would therefore have had to introduce its own currency and preferably make it convertible.

The gradual strategy pursued by the Russian reformers took advantage of

complementarities of reforms to secure nationalization of the Russian ruble.<sup>74</sup> By launching a reform of the payments system, pushing for further liberalization of foreign trade and energy prices, and starting a radical program for privatization of large enterprises, the reformers managed to gain support from parts of industry for the break-up of the ruble zone. Many managers simply realized that they had much to gain from trade at world market prices and the more flexible system of trade that developed in 1993 and after. The privatization process, spontaneous as well as controlled, contributed to this since it strengthened the property rights of managers over firms. This gave them control over a larger share of resources in the economy, which, according to the analysis of Bornefalk (2000b), made them more inclined to accept economic reforms facilitating profit maximization rather than rent-seeking. Even more important for the available economic reform space was, according to the same analysis, the division of interests of different industries.

The energy sector was particularly important in the process. Åslund (1995, p. 300 ff) argues that it was the most influential industrial sector. It was also well represented among top decision makers. In the beginning of the change of economic system, managers in the energy sector could gain enormous rents from continued price and trade controls. As reforms had gone far enough and managers had been able to acquire property rights over much of their enterprises, trade on market conditions became more attractive. Interestingly, Viktor Chernomyrdin entered the scene when the monetary reform was launched in the summer of 1992, which contributed to making it a prolonged process. The process did not end until Chernomyrdin opted for not readmitting Belarus into the ruble zone in the spring of 1994. The interpretation of Dabrowski (1995, footnote 15) is that he had finally realized the costs to Russia of doing so. An alternative explanation is that conditions had changed enough to make the energy sector—Chernomyrdin's main supporter—more interested in trade on market conditions.

Producers of non-competitive products, for instance agricultural machinery, perpetuated the dream of resurrecting the ruble zone. They therefore lobbied intensively in early 1994 for an economic union between Russia and Belarus. Because of the division of interests, however, the lobbying for a new costly monetary union was not powerful enough to dictate the development. In fact, the analysis of Bornefalk (2000b) suggests that the division of interests as such should have made interest groups more willing to accept economic reforms limiting their possibilities to extract rents.

The steps towards increased political support for stabilization and the increasingly differing interests of industrial sectors, both to some extent explained by the weaknesses of the ruble zone, led to a possibility of delimiting the ruble area, which in turn made the Russian ruble potentially stable. For the non-Russian FSRs, the introduction of national

---

<sup>74</sup> Dewatripont and Roland (1995) analyze under what conditions gradualism is optimal when complementary reforms are to be implemented. Assuming that a single reform can never be attractive per se, they find that complementarity of reforms is a necessary condition for gradualism to be optimal. In a different setting, Friedman and Johnson (1995) find that complementarity of reforms makes radical reforms optimal.

currencies made the importance of fiscal and monetary policies clear. This was a precondition for stabilization in the individual countries as well as for the FSU as a whole. Hence, the high transaction costs of the ruble zone were used by the radical reformers to bring about institutional change in the direction of well-delimited currency areas. In addition to the direct effect in terms of greater support from mass media and political organizations, there was an indirect effect. The increased transparency lowered the incentives of interest groups to engage in appropriative activities since it made production relatively more profitable. It also helped the government to strengthen its position through elections or referendums, which in turn contributed to the development of democratic institutions, for reasons discussed in Bornefalk (2000b). In this way, the government managed to impose reforms that had initially run counter to the interests of state enterprise managers.

#### **6.4 The Relative Success of the Break-Up**

Avinash Dixit's (1996) "transaction-cost politics" framework recognizes that economic policymaking is a process in real time combining rule-making and individual acts. Such a process can be analyzed as a dynamic game, in which conditions are uncertain and changing, and rules are at least partially made by the participants over the course of the game (*ibid.* p. 30). "Strategic moves" play a crucial role in games of this nature.<sup>75</sup> These are actions taken by participants to manipulate the operation of the subsequent game in order to favor their own interests. The strategic moves are commitments, which can be either unconditional or contingent on other participants' actions.

Dixit (1990) discusses political transaction costs that emanate from asymmetric information and time inconsistency.<sup>76</sup> Opportunism, for instance, arises when actions of agents are unobservable.<sup>77</sup> To secure efficiency, monitoring and incentive schemes can be used, at least in theory. But if agency relationships are too complicated, which they often are in politics and public administration, more blunt instruments such as commitments must be relied on.

Agency problems in the ruble zone were overwhelming because of the multitude of agents involved in trade and financial aspects of relations between member countries. Most importantly, there were severe common agency problems obstructing efficient policies of both the government and the CBR, particularly after the political changes in the spring of 1992. In other words, the reformers could not control the process since many decision makers were heavily influenced by industrial lobbies. Hence, there was no way to secure commitment to the restrictions on technical credits. Therefore, Russian enterprises continued exporting to other FSRs that were running large deficits, expecting that they would eventually receive payment through the actions of the CBR.

---

<sup>75</sup> This term was coined by Schelling (1960).

<sup>76</sup> The concept of political transaction costs is developed in North's (1990) transaction cost theory of politics.

<sup>77</sup> For a taxonomy of other political transaction costs, see Dixit (1996, p. 54 ff.)

Another way of overcoming moral hazard problems is through repetition, as in repeated prisoner's dilemma games. In the analysis of Section 4.2, the non-Russian central banks were agents whose actions were unobservable to the CBR and the Russian government. But, as was shown, there was no way to secure cooperation through that mechanism. Instead, the mechanism was interpreted as a strategic move that eventually enabled the reformers to secure a nationalization of the Russian ruble.

Some scholars recognize Russia's large-scale privatization as the success of Russia's systemic change in the 1990s.<sup>78</sup> However, the political transaction costs stalling the process towards well-delimited currency areas were greater than those hindering large-scale privatization. At the same time, the economic transaction costs of the ruble zone were so great that the political transaction costs would sooner or later have to be overcome. This is not true of delayed privatization to the same extent, as can be seen from the experience of several countries that have undergone a change of economic system.

To start with the understanding of reforms, the question of state versus private ownership was easier to understand than whether a common currency or national currencies would serve better to facilitate economic integration. Large-scale privatizations had been launched or were being discussed in most countries of central and eastern Europe in 1991-92, while monetary disintegration only took place in former Yugoslavia during that period, and on top of that as a result of civil war. At the same time, countries in the European Union decided to deepen their integration by forming a monetary union. Hence, few policy makers understood that the issue of monetary relations with the FSRs was really one of unified control over a currency area. Privatization, on the other hand, had been discussed for a long time, and the Russian Supreme Soviet adopted a privatization law already in June 1991. The fact that the IMF supported the ruble zone, whereas international financial organizations stood in line to provide technical assistance on how to privatize, strengthened this difference.

As regards the possibility for the reformers to commit to overcoming problems of opportunism, that was easier in the case of privatization than in that of the break-up of the ruble zone. One reason for this is that both the parliament and the president supported the privatization process, partly because it was conceptually easier to understand, partly because it had been an issue for a longer time. There was also weaker resistance from interest groups since privatization did not necessarily mean that they could acquire less rents. Once it became clear that the government opted for a mass privatization scheme, however, resistance from lobby groups and their agents in parliament started to grow. In the end, state enterprise managers managed to secure control rights over a large share of Russian industry.

Problems due to multiple agency and common agency were also more serious in the case of the ruble zone. Most importantly, the radical reformers managed to gain control over the privatization process by creating the GKI, the state property committee, which soon had local branches over most of Russia. It had no predecessor, and Anatoly Chubais,

---

<sup>78</sup> See, for instance, Åslund (1995, p. 293).

one of the leading reformers, was appointed minister of privatization and chairman of the GKI. In the area of monetary and trade-related reform, the reformers had to build on existing institutions. The failure to secure control over the CBR was of fundamental importance, but can be explained by the influence of the parliament. The influence of Roskontrakt, the agency responsible for state controlled trade, was also detrimental. The fact that 14 other independent countries were involved inflated political transaction costs further.

Hence, political transaction costs were greater in relation to decisions on the ruble zone than to large-scale privatization. The list of criteria for judging major developments proposed by Åslund (1995, pp. 293-294) does not explicitly acknowledge this kind of costs.<sup>79</sup> To the extent that these costs are important, they should complement Åslund's criteria to make them suitable for comparing the relative success of large-scale privatization and monetary reform.

Some of Åslund's criteria do nevertheless apply. To take just one example, the extent to which international advice and experience was used does not differ considerably. The advice on monetary relations with other FSRs was simply poor until May 1992. Once the government received good advice, it proceeded with great determination. But, as we have seen, the window of opportunity had then been shut. To conclude, we find no support for the view that Russia's privatization was implemented in a better way than its monetary reform. On the contrary, it seems as if policy makers utilized the available reform space quite well in both cases, but that it was more limited in the case of monetary reform.

---

<sup>79</sup> The list includes the following criteria: strong and stable leadership in the field; support from the supreme political leader; development of a new administration; use of international advice and experiences; clear and sophisticated strategy; elaboration of an operative program; recognition of relevant interests without losing strategy; parliamentary mandate; public outreach; and swift implementation.

## 7. CONCLUSIONS

The disintegration of the ruble zone is a fascinating example of the political economy of a collapsing monetary union. Russia suffered exorbitant costs since it did not dissolve the ruble zone in early 1992 despite a prolonged political disintegration process. The non-Russian member countries cannot be blamed for Russia's misfortunes. Given the constraints they had to face, they clearly reacted rationally to the inadequacies of the mechanisms of the ruble zone imposed by Russia. Russia, on the other hand, must bear some of the responsibility for the hyperinflations that erupted in several FSRs as these were eventually thrown out of the ruble zone.

The dissolution of the ruble zone must nevertheless be considered an achievement for the Russian reformers, who had to overcome massive resistance from rent-seeking industrialists and conservative politicians wanting to keep some influence over the former Soviet empire. The introduction of limits on trade credits to other FSRs was a decisive step towards the dissolution. This decision was subtle enough to be pushed through, but important enough to get the process going.

The timing of the Russian currency reform was largely determined by internal political developments. Since a radical nationalization of the Russian currency was not launched before resistance against the change of economic system had grown strong, a gradual strategy had to be followed. This strategy utilized exceedingly costly monetary actions undertaken by other member countries to impose reforms. However, the gradual reform strategy with its implicit mechanism for exclusion was far from risk free. Had the conservatives been able to stall the process to a greater extent than they did, the ruble zone could have caused an even greater monetary disorder, possibly followed by a reversal of democratic and economic reforms.

The costly disintegration of the ruble zone is interesting from a more general perspective. The numerous suggestions from western economists on how a currency union, a payments union, or other more or less far-reaching types of monetary integration could be arranged all failed when faced with the political disintegration process that the FSRs were undergoing. The effects of the political disintegration process were exacerbated by differences in the level of development reached by political institutions in different member countries and in the relative size of the economies, since these factors led to differences in the preferred rates of monetary expansion between member countries. Differences in political institutions also meant that the available economic reform space differed between countries. The failure of the IMF and a number of monetary economists to realize the impossibility of establishing a functioning monetary union in the FSU suggests that political factors need to be given greater attention in analyses of monetary integration.

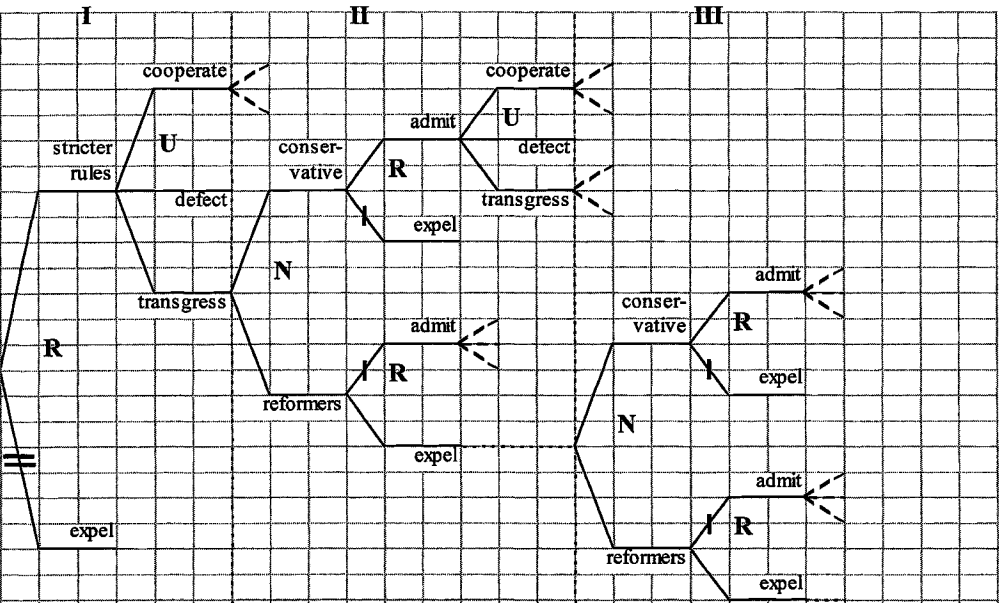
The second general conclusion is also related to political integration. Bordo and Jonung (1997) conclude that it is necessary that members of a currency union should

pursue political integration, and that they should surrender their monetary sovereignty to a central monetary authority which can guarantee stability of the common currency, if a currency union is to remain functional. For this to be realistic, however, we find that the potential member countries need to be sufficiently similar so that they will prefer similar monetary policies.

We have also seen that a dominant country of a currency union has a lot to lose from political disintegration. It must therefore be able to act quickly once political disintegration threatens. This finding could possibly be of relevance for Germany as a member of the EMU.

# APPENDIX

The following figure shows the structure of the decisions that were taken by R and U in the model of the restricted rubble zone in Section IX. It is assumed that the decision on the new rules of the game was made by a reformist government, R<sub>r</sub>, in period I since this was the case in reality. The choice has been indicated in the figure: R<sub>r</sub> chose to impose stricter rules rather than to expel the other members. U then chose to transgress knowing that it was unlikely that it would be expelled indefinitely for reasons explained in subsection IX B. In period II, if N selects R<sub>c</sub>, U would receive new technical credits and be able to transgress again. If N selects R<sub>r</sub>, U risks being expelled. If it is, there will be a possibility to be readmitted in period III by R<sub>c</sub>. The game goes on in the same fashion until the rules are changed.



**Figure 4: A Repeated Game under Uncertainty and Renegotiation**

Note: Interrupted lines indicate that the consecutive branches follow the same logic as those that are shown.



## REFERENCES

- Admiraal, P. H., ed. *Economic Transition in Eastern Europe*. Blackwell, Oxford, 1993.
- Alesina, Alberto and Drazen, Allan. "Why Are Stabilizations Delayed?" *American Economic Review*, vol. 81, December, pp. 1170-1188, 1991.
- Åslund, Anders. *How Russia Became a Market Economy*. The Brookings Institution, Washington, D.C., 1995.
- Åslund, Anders. "The Nature of the Transformation Crisis in the Former Soviet Countries." in Siebert, ed., 1994.
- Åslund, Anders. "Key Dilemmas in the Russian Economic Reform." Mimeo, Stockholm School of Economics, 1993.
- Åslund, Anders, ed. *Russian Economic Reform at Risk*. Pinter Publishers, London, 1995.
- Åslund, Anders, ed. *Economic Transformation in Russia*. Pinter Publishers, London, 1994.
- Åslund, Anders and Layard, Richard, eds. *Changing the Economic System in Russia*. Pinter Publishers, London, 1993.
- Berglund, Sten, Hellén, Thomas, and Aarebrot, Frank, eds. *Handbook of Political Change in Eastern Europe*, draft, 1998.
- Blanchard, Olivier and Layard, Richard, eds. *Post-Communist Reform: Pain and Progress*. MIT Press, Cambridge and London, 1993.
- Boeva, Irina and Dolgopiatova, Tatiana. "State Enterprises during Transition: Forming of Survival Strategies." in Åslund, ed., 1994.
- Bordo, Michael D. and Jonung, Lars. "The History of Monetary Regimes - Some Lessons for Sweden and the EMU." *Swedish Economic Policy Review*, vol. 4, pp. 285-358, 1997.
- Bornefalk, Anders. "Social Conflict with Passive Groups." in *Essays on Social Conflict and Reform*, Ph.D. Dissertation, Stockholm School of Economics, 2000a.
- Bornefalk, Anders. "Democratization, Rent Seeking, and Economic Transition." in *Essays on Social Conflict and Reform*, Ph.D. Dissertation, Stockholm School of Economics, 2000b.
- Bornefalk, Anders. "Constitutional Constraints and Redistributive Activities." in *Essays on Social Conflict and Reform*, Ph.D. Dissertation, Stockholm School of Economics, 2000c.
- Bornefalk, Anders. "The Ruble Zone: A Case of Irrationality?" *Ukrainian Economic Review*, vol. 1, no. 1-2, 1995.
- Bornefalk, Anders. "Currencies and Payments Systems in the Former Soviet Union." in Swedish, *Östekonomisk Rapport*, vol. 4, no. 15, 1992.
- Cagan, Philip. "The Monetary Dynamics of Hyperinflation." In Milton Friedman, ed., *Studies in the Quantity Theory of Money*. University of Chicago Press, Chicago, 1956.
- Cheasty, Adrienne and Spencer, Grant. "The Ruble Area: A Breaking of Old Ties?" *Finance and Development*, vol. 30:2, 1993.

- Coase, Ronald H. "The Nature of the Firm." *Economica*, vol. 4, pp. 386-405, 1937.
- Conway, Patrick. "Rubles, Rubles, Everywhere—Currency Shortages and Financial Disintermediation in the Former Soviet Union." report of research under World Bank Research Support Budget RPO 678-08, 1993.
- Conway, Patrick. *Currency Proliferation: The Monetary Legacy of the Soviet Union*. Essays in International Finance No. 197, Princeton, New Jersey, June 1995.
- Dabrowski, Marek. *The Reasons of the Collapse of the Ruble Zone*. Studies and Analyses, no. 58, CASE, Warsaw, 1995.
- Dewatripont, Mathias and Roland, Gérard. "Transition As a Process of Large-Scale Institutional Change." *Economics of Transition*, vol. 4, no. 1, pp. 1-30, 1996.
- Dewatripont, Mathias and Roland, Gérard. "The Design of Reform Packages under Uncertainty." *American Economic Review*, vol. 85, December, pp. 1207-23, 1995.
- Dixit, Avinash K. *The Making of Economic Policy: A Transaction-Cost Politics Perspective*. MIT Press, Cambridge, MA, 1996.
- Dornbusch, Rudiger. "Payments Arrangements among the Republics." in Blanchard and Layard, 1993.
- Dornbusch, Rudiger. "Monetary Problems of Post-Communism: Lessons From the End of the Austro-Hungarian Empire." *Weltwirtschaftliches Archiv*, vol. 128, no. 3, 1992, pp. 391-424.
- Eichengreen, Barry. "European Payments Union." *Economic Policy*, October 1993.
- Ellingsen, Tore. "Payments in Kind." mimeo. Stockholm School of Economics, January 1998.
- Fedorov, Boris and Kazmin, Andrei. "1993: The First Experiences of the Russian Financial and Monetary Stabilization Policy." in Åslund, ed., 1994.
- Fernandez, Raquel and Rodrik, Dani. "Resistance to Reform: Status Quo Bias in the Presence of Individual-Specific Uncertainty." *American Economic Review*, vol. 81, December, pp. 1146-55, 1991.
- Freedom House. *Nations in Transit 1998*. Available: [Online]: [www.freedomhouse.org](http://www.freedomhouse.org), 11/17/1999.
- Friedman, Eric J. and Johnson, Simon. "Complementarities and Optimal Reform." SITE Working Paper no. 109, Stockholm School of Economics, 1995.
- Friedman, Milton. *Money Mischief: Episodes in Monetary History*. Harcourt, Brace and Jovanovich, New York, 1992.
- Fudenberg, Drew and Tirole, Jean. *Game Theory*. MIT Press, Cambridge, MA, 1991.
- Goldberg, Linda S, Ickes, Barry W., and Ryterman, Randi. "Departures from the Ruble Zone: The Implications of Adopting Independent Currencies." *The World Economy*, vol. 17, no. 3, May 1994.
- Gaidar, Yegor. "Inflationary Pressures and Economic Reform in the Soviet Union." in Admiraal, ed., (1993).
- Granville, Brigitte. "Farewell, Ruble Zone." In Åslund, ed., 1995.
- Granville, Brigitte and Lushin, Andrei. "The New Style Ruble Zone or the Old Soviet

- Union Revisited." Memo, Macroeconomic and Finance Unit under the Ministry of Finance of Russia, 1993.
- Hanson, Philip and Teague, Elizabeth. "The Industrialists and Russian Economic Reform." *RFE/RL Research Report*, vol. 1, no. 19, 1992.
- Hansson, Ardo H. "The Trouble with the Ruble: Monetary Reform in the Former Soviet Union." in Åslund and Layard, eds, 1993.
- International Monetary Fund. "Financial Relations among Countries of the Former Soviet Union." *IMF Economic Reviews*, no. 1, 1994.
- International Monetary Fund. "Trade Policy Reform in the Countries of the Former Soviet Union." *IMF Economic Reviews*, no. 2, 1994.
- Johnson, Simon and Ustenko, Oleg. "The Road to Hyperinflation: Economic Independence in Ukraine, 1991-93." paper presented at the conference on First Steps Toward Economic Independence: New States of the Postcommunist World, Stockholm Institute of East European Economics, August, 1993.
- Karlsson, Joakim. "Former Soviet Republics: Widely Differing Performance." *Östekonomisk Rapport*, vol. 6, no. 6, 1994.
- Kornai, János. *The Socialist System: The Political Economy of Communism*. Princeton University Press, Princeton, 1992.
- Lagerspetz, Mikko and Vogt, Henri. "Estonia." in Berglund, Hellén, and Aarebrot, eds, 1998.
- Leijonhufvud, Axel. "The Nature of the Depression in the Former Soviet Union." *New Left Review*, no. 199, May-June 1993.
- Lipton, David and Sachs, Jeffrey D. "Prospects for Russia's Economic Reforms." *Brookings Papers on Economic Activity*, 2:1992.
- Lushin, Andrei and Sarafanov, Mikhail. "Prices in Inter Republican Trade." Memo, Macro and Finance Unit, Moscow, 1993.
- Normak, Tatyana and Teplukhin, Pavel. "Trade Between States of the Former Soviet Union." Mimeo, Moscow, 1994.
- North, Douglass C. *Institutions, Institutional Change, and Economic Performance*. Cambridge University Press, Cambridge, 1990.
- Olson, Mancur. "Why Is Economic Performance Even Worse when Communism Is Abolished?" Paper presented at the Institute for the Economy in Transition, Moscow, March 1994.
- Olson, Mancur. *The Logic of Collective Action*. Harvard University Press, Cambridge, MA, 1965.
- Pasvolksy, Leo. *Economic Nationalism of the Danubian States*. George Allen & Unwin, London, 1928.
- Sachs, Jeffrey D. *Poland's Jump to a Market Economy*. MIT Press, Cambridge, MA, 1993.
- Sachs, Jeffrey D. "Prospects for Monetary Stabilization in Russia." in Åslund, ed., 1994.
- Sachs, Jeffrey D. and Lipton, David. "Remaining Steps To a Market-Based Economic

- System in Russia.” in Åslund and Layard, eds, 1993.
- Sargent, Thomas J. “The Ends of Four Big Inflations,” in Sargent, ed., 1986.
- Sargent, Thomas J., ed. *Rational Expectations and Inflation*. Harper & Row, 1986.
- Schelling, Thomas. *The Strategy of Conflict*. Harvard University Press, Cambridge, MA, 1960
- Siebert, Horst, ed. *Overcoming the Transformation Crisis, Lessons for the Successor States of the Soviet Union*. Institut für Weltwirtschaft an der Universität Kiel, J.C.B. Mohr (Paul Siebeck), Tübingen, 1994.
- Williamson, John. *What Role for Currency Boards?* Institute for International Economics, Washington D.C., 1995.

# **EFI**

## **The Economic Research Institute**

**Reports since 1995**

*Published in the language indicated by the title*

**1999**

- Adler, N.**, Managing Complex Product Development.
- Allgulin, M.**, Supervision and Monetary Incentives.
- Andersson, P.**, Experto Credite: Three Papers on Experienced Decision Makers.
- Ekman, G.**, Från text till batong – Om poliser, busar och svennar.
- Eliasson, A-C.**, Smooth Transitions in Macroeconomic Relationships.
- Flink, H., Gunnarsson, J., Wahlund, R.**, Svenska hushållens sparande och skuldsättning– ett konsumentbeteende-perspektiv
- Gunnarsson, J.**, Portfolio-Based Segmentation and Consumer Behavior: Empirical Evidence and Methodological Issues.
- Hamrefors, S.**, Spontaneous Environmental Scanning.
- Helgesson, C-F.**, Making a Natural Monopoly: The Configuration of a Techno-Economic Order in Swedish Telecommunications.
- Japanese Production Management in Sunrise or Sunset.** Christer Karlsson (editor).
- Jönsson, B., Jönsson, L., Kobelt, G.**, Modelling Disease Progression And the Effect of Treatment in Secondary Progressive MS. Research Report.
- Lindé, J.**, Essays on the Effects of Fiscal and Monetary Policy.
- Ljunggren, U.**, Indikatorer i grundskolan i Stockholms stad före stadsdels-nämndsreformen – en kartläggning.
- Ljunggren, U.**, En utvärdering av metoder för att mäta produktivitet och effektivitet i skolan – med tillämpning i Stockholms stads grundskolor
- Lundbergh, S.**, Modelling Economic High-Frequency Time Series.
- Mägi, A.**, Store Loyalty? An Empirical Study of Grocery Shopping.
- Mölleryd, B.G.**, Entrepreneurship in Technological Systems – the Development of Mobile Telephony in Sweden.
- Nilsson, K.**, Ledtider för ledningsinformation.
- Rognes, J.**, Telecommuting – Organisational Impact of Home Based – Telecommuting.
- Sandström, M.**, Evaluating the Benefits and Effectiveness of Public Policy.
- Sjöstrand, S-E., Sandberg, J., Tyrstrup, M., (red)** Osynlig Företagsledning.
- Skalin, J.**, Modelling Macroeconomic Time Series with Smooth Transition Autoregressions.
- Spagnolo, G.**, Essays on Managerial Incentives and Product-Market Competition.
- Strauss, T.**, Governance and Structural Adjustment Programs: Effects on Investment, Growth and Income Distribution.
- Svedberg Nilsson, K.**, Effektiva företag? En studie av hur privatiserade organisationer konstrueras.
- Söderström, U.**, Monetary Policy under Uncertainty.
- Werr, A.**, The Language of Change The Roles of Methods in the Work of Management Consultants.
- Wijkström, F.**, Svenskt organisationsliv – framväxten av en ideell sektor.

- Andersson, M.**, On Testing and Forecasting in Fractionally Integrated Time Series Models.
- Berg-Suurwee, U.**, Styrning av kultur- och fritidsförvaltning innan stadsdelsnämndsreformen
- Berg-Suurwee, U.**, Nyckeltal avseende kultur- och fritidsförvaltning innan stadsdelsnämndsreformen.
- Bergström, F.**, Essays on the Political Economy of Industrial Policy.
- Bild, M.**, Valuation of Takeovers.
- Charpentier, C., Samuelson, L.A.**, Effekter av en sjukvårdsreform – en analys av Stockholmsmodellen.
- Eriksson-Skoog, G.**, The Soft Budget Constraint: The Emergence, Persistence and Logic of an Institution. The Case of Tanzania 1967-1992.
- Gredenhoff, M.**, Bootstrap Inference in Time Series Econometrics.
- Ioannidis, D.**, I nationens tjänst? Strategisk handling i politisk miljö – en nationell teleoperatörs interorganisatoriska, strategiska utveckling.
- Johansson, S.**, Savings Investment, and Economic Reforms in Developing Countries.
- Levin, J.**, Essays in Company Valuation.
- Ljunggren, U.**, Styrning av grundskolan i Stockholms stad innan stadsdelsnämndsreformen.
- Mattsson, S.**, Från stat till marknad – effekter på nätverksrelationer vid en bolagiseringsreform.
- Nyberg, A.**, Innovation in Distribution Channels – An Evolutionary Approach.
- Olsson, P.**, Studies in Company Valuation.
- Reneby, J.**, Pricing Corporate Securities.
- Roszbach, K.**, Essays on Banking Credit and Interest Rates.
- Runsten, M.**, The Association Between Accounting Information and Stock Prices. Model development and empirical tests based on Swedish Data.
- Segendorff, B.**, Essays on Bargaining and Delegation.
- Sjöberg, L., Bagozzi, R., Ingvar, D.H.**, Will and Economic Behavior.
- Sjögren, A.**, Perspectives on Human Capital: Economic Growth, Occupational Choice and Intergenerational Mobility.
- Studier i kostnadsintäktsanalys**, red Jennergren, P.
- Söderholm, J.**, Målstyrning av decentraliserade organisationer. Styrning mot finansiella och icke-finansiella mål.
- Thorburn, K.**, Cash Auction Bankruptcy and Corporate Restructuring
- Wijkström, F.**, Different Faces of Civil Society.
- Zethraeus, N.**, Essays on Economic Evaluation in Health Care. Evaluation of Hormone Replacement Therapy and Uncertainty in Economic Evaluations.

- Alexius, A.**, Essays on Exchange Rates, Prices and Interest Rates.
- Andersson, B.**, Essays on the Swedish Electricity Market.
- Berggren, N.**, Essays in Constitutional Economics.
- Ericsson, J.**, Credit Risk in Corporate Securities and Derivatives. Valuation and Optimal Capital Structure Choice.
- Charpentier, C.**, Budgeteringens roller, aktörer och effekter. En studie av budgetprocesserna i en offentlig organisation.
- De Geer, H., Silfverberg, G.**, Citizens' Trust and Authorities' Choices, A Report from The Fourth International Conference on Ethics in the Public Service, Stockholm June 15-18, 1994.
- Friberg, R.**, Prices, Profits and Exchange Rates.

**Från optionsprissättning till konkurslagstiftning**, red. Bergström, C., Björk, T.  
**Hagerud, G.E.**, A New Non-Linear GARCH Model.  
**Haksar, A.**, Environmental Effects of Economywide Policies: Case Studies of Costa Rica and Sri Lanka.  
**He, C.**, Statistical Properties of Garch Processes.  
**Holmgren, M.**, Datorbaserat kontrollrum inom processindustrin; erfarenheter i ett tidsperspektiv.  
**Jennergren, P.**, Tutorial on the McKinsey Model for Valuation of Companies.  
**Lagerlöf, J.**, Essays on Political Economy, Information, and Welfare.  
**Lange, F., Wahlund, R.**, Planerade och oplanerade köp - Konsumenternas planering och köp av dagligvaror.  
**Löthgren, M.**, Essays on Efficiency and Productivity; Contributions on Bootstrap, DEA and Stochastic Frontier Models.  
**Nilsson, H.E.**, Using Conceptual Data Modelling in Management Accounting: A Case Study.  
**Sjöberg, L., Ramsberg, J.**, En analys av en samhällsekonomisk bedömning av ändrade säkerhetsföreskrifter rörande heta arbeten.  
**Säfvenblad, P.**, Price Formation in Multi-Asset Securities Markets.  
**Sällström, S.**, On the Dynamics of Price Quality.  
**Södergren, B.**, På väg mot en horisontell organisation? Erfarenheter från näringslivet av decentralisering och därefter.  
**Tambour, M.**, Essays on Performance Measurement in Health Care.  
**Thorén, B., Berg-Surwee, U.**, Områdesarbete i Östra Hökarängen - ett försök att studera effekter av decentralisering.  
**Zhang Gang.**, Chinese Rural Enterprises Between Plan and Market.  
**Åhlström, P.**, Sequences in the Process of Adopting Lean Production.  
**Åkesson, G.**, Företagsledning i strategiskt vakuum. Om aktörer och förändringsprocesser.  
**Åsbrink, S.**, Nonlinearities and Regime Shifts in Financial Time Series.

**1996**

**Advancing your Business. People and Information Systems in Concert.**  
red. Lundeberg, M., Sundgren, B.  
**Att föra verksamheten framåt. Människor och informationssystem i samverkan.**  
red. Lundeberg, M., Sundgren, B.  
**Andersson, P.**, Concurrence, Transition and Evolution - Perspectives of Industrial Marketing Change Processes.  
**Andersson, P.**, The Emergence and Change of Pharmacia Biotech 1959-1995. The Power of the Slow Flow and the Drama of Great Events.  
**Asplund, M.**, Essays in Industrial Economics.  
**Delmar, F.**, Entrepreneurial Behavior & Business Performance.  
**Edlund, L.**, The Marriage Market: How Do You Compare?  
**Gunnarsson, J.**, Three Studies of Financial Behavior. Research Report.  
**Hedborg, A.**, Studies of Framing, Judgment and Choice.  
**Holgersson, C., Höök, P.**, Ledarutveckling för kvinnor - Uppföljning av en satsning på Volvo - Research Report.  
**Holmberg, C.**, Stores and Consumers - Two Perspectives on Food Purchasing.  
**Håkansson, P., Wahlund, R.**, Varumärken. Från teori till praktik.  
**Karlsson, A.**, The Family Business as an Heirloom. Research Report.  
**Linghag, S.**, Man är handelsstudent. Research Report.  
**Molin, J.**, Essays on Corporate Finance and Governance.  
**Mägi, A.**, The French Food Retailing Industry - A Descriptive Study.

**Mölleryd, B.**, Så byggdes en världsindustri - Entreprenörskapets betydelse för svensk mobiltelefoni. Research Report.

**Nielsen, S.**, Omkostningskalkulation for avancerede produktions-omgivelser - en sammenligning af stokastiske og deterministiske omkostningskalkulationsmodeller. Research Report.

**Normark, P., Danielsson, L., Larsson, A., Lundblad, P.**, Kooperativa nyckeltal. Research Report.

**Sandin, R.**, Heterogeneity in Oligopoly: Theories and Tests. Research Report.

**Sandén-Håkansson, U.**, Från kampanjmål till mediemix - en studie av samarbete mellan annonsörer, reklambyråer och mediebyråer. Research Report.

**Stein, J., Söderlund, M.**, Framgång i arbetet, strategier för att utföra arbetsuppgifterna, arbetsuppgifternas karaktär och utbildningskvalitet. En empirisk studie av civil-ekonomer. Research Report.

**Strömberg, P., Thorburn, K.**, An Empirical Investigation of Swedish Corporations in Liquidation Bankruptcy. Research Report.

**Söderlund, M.**, Och ge oss den nöjda kunden. En studie av kundtillfredsställelse och dess orsaker och effekter. Research Report.

**Thodenius, B.**, Användningen av ledningsinformationssystem i Sverige: Lägesbild 1995. Research Report.

**Ulfsdotter, U.**, Internationalisering för expansion eller hemmamarknadsförsvar? De nordiska marknaderna för fruktyghurt 1982-1994. Research Report.

**Westelius, A.**, A Study of Patterns of Communication in Management Accounting and Control Projects. Research Report.

**Wijkström, F.**, Den svenska ideella sektorn och pengarna. Research Report.

**Örtendahl, M.**, Health and Time - A Problem of Discounting. Research Report.

## **1995**

**Becker, T.**, Essays on Stochastic Fiscal Policy, Public Debt and Private Consumption. Research Report.

**Blomberg, J.**, Ordning och kaos i projektsamarbete - en socialfenomenologisk upplösning av en organisationsteoretisk paradox. Research Report.

**Brodin, B., Lundkvist, L., Sjöstrand, S-E., Östman, L.**, Styrelsearbete i koncerner. Research Report.

**Brännström, T.**, Bias Approximation and Reduction in Vector Autoregressive Models. Research Report.

**Ekonomisk politik i omvandling.** red. Jonung, L. Research Report.

**Gunnarsson, J., Wahlund, R.**, Hushållens finansiella strategier. En explorativ studie. Research Report.

**Höök, P.**, Chefsutveckling ur könsperspektiv - Mentorskap och nätverk på Vattenfall - Research Report.

**Levin, J., Olsson, P.**, Looking Beyond the Horizon and Other Issues in Company Valuation. Research Report.

**Mägi, A.**, Customer Satisfaction in a Store Performance Framework. Research Report.

**Persson, P-G.**, Modeling the Impact of Sales Promotion on Store Profits. Research Report.

**Roman, L.**, Institutions in Transition. A Study of Vietnamese Banking. Research Report.

**Sandberg, J.**, Statistisk metod - dess vetenskapliga hemvist, grundläggande principer och möjligheter inom samhällsvetenskapen. Research Report.

**Sandberg, J.**, How Do We Justify Knowledge Produced by Interpretative Approaches? Research Report.

**Schuster, W.**, Redovisning av konvertibla skuldebrev och konvertibla vinstandelsbevis - klassificering och värdering. Research Report.

**Söderberg, K.**, Farmartjänsten - Ny Kooperation inom lantbruket. Research Report.

**Söderqvist, T.**, Benefit Estimation in the Case of Nonmarket Goods. Four Essays on Reductions of Health Risks Due to Residential Radon Radiation. Research Report.

**Thorén, B.**, Användning av information vid ekonomisk styrning - månadsrapporter och andra informationskällor. Research Report.





