Political Specialization*

Bernardo Guimaraes†
Sao Paulo School of Economics – FGV

Kevin D. Sheedy‡
London School of Economics

February 2017

Abstract

This paper presents a theory of political specialization in which some countries uphold the rule of law while others consciously choose not to do so, even though they are ex ante identical. This is borne out of two key insights: for incumbents in each country, (i) the first steps to the rule of law have the greatest private cost, and (ii) steps taken by some countries in the direction of the rule of law make it less attractive for others to follow the same path. The world equilibrium features a symbiotic relationship between despotic and rule-of-law economies: by producing technology-intensive goods that require protection of property rights, rule-of-law economies raise the relative price of natural resources and increase incentives for despotism in other countries; while the choice of despotism entails a positive externality because cheap oil makes the rule of law more attractive elsewhere in the world.

JEL classifications: D74; F43; O43; P48.

Keywords: rule of law; power sharing; international trade; resource curse; development.

---

*We thank Alessandro Belmonte, Francesco Caselli, Cristina Terra, and seminar participants at American University (Washington D.C.), the Anglo-French-Italian Macroeconomics Workshop, Econometric Society American Meeting 2016 (Philadelphia), Econometric Society European Meeting 2016 (Geneva), ESSIM 2016 (Helsinki), FGV-Rio, HSE-ICEF Moscow, 1st International REAP Meeting, KBTU-ISE Almaty, London School of Economics, PUC-Rio, RES 2016 (Sussex), SAET 2016 (Rio de Janeiro), Sao Paulo School of Economics – FGV, SED 2016 (Toulouse), and the Stony Brook Workshop on Political Economy for their comments.

†Sao Paulo School of Economics – FGV. Email: bernardo.guimaraes@fgv.br

‡LSE, CEPR, and CfM. Email: k.d.sheedy@lse.ac.uk

1 Introduction

It is a commonplace to claim that the world has become smaller and there are fewer and fewer differences between formerly exotic places and the West. An unprecedented flow of goods and ideas has allowed emulation of faraway countries, resulting in a large increase in conformity across the globe. However, politics seem to be immune to this trend. While the rule of law can be taken for granted in large parts of the world, authoritarianism prevails in far too many places, in spite of its well-known negative consequences.

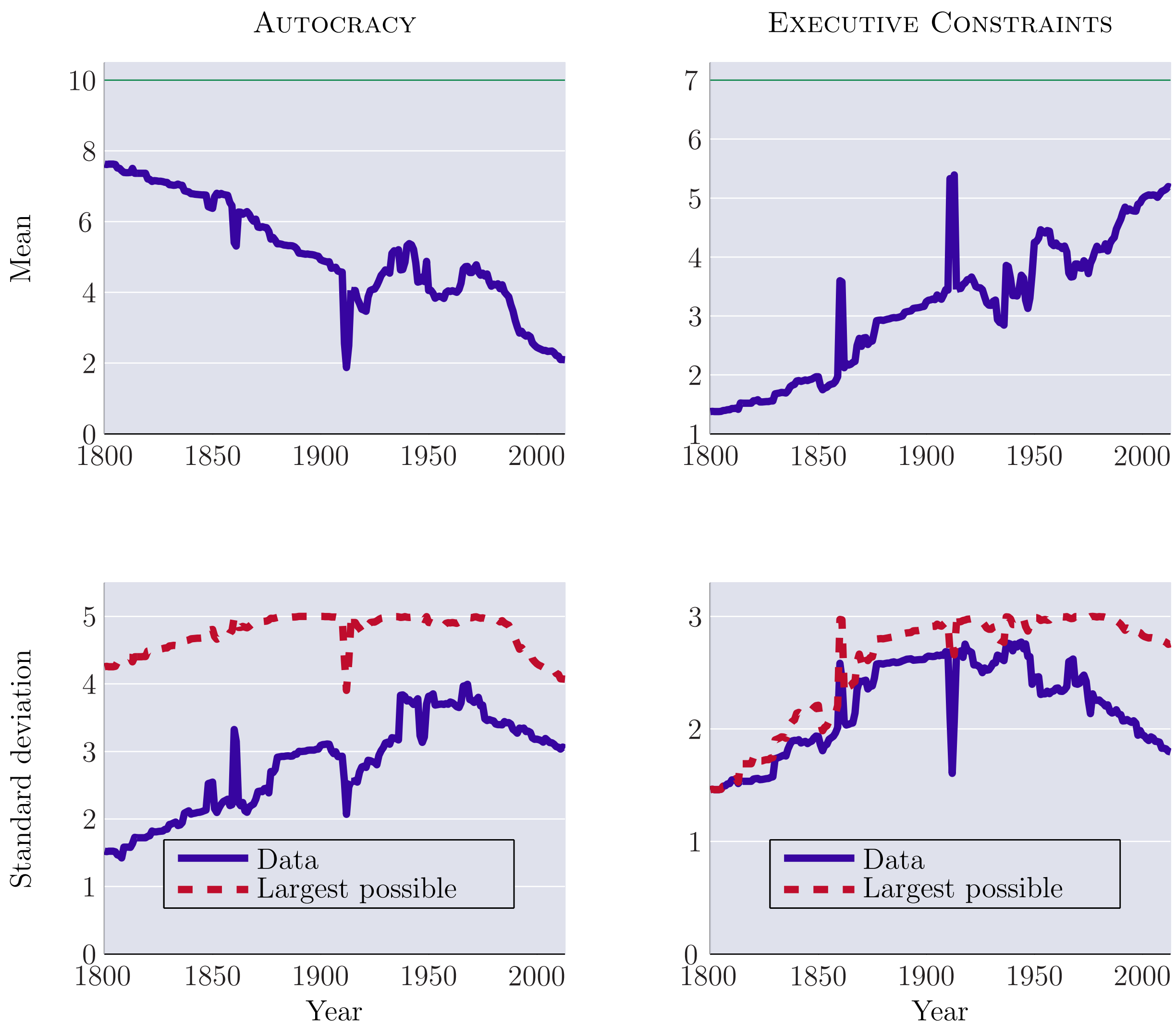
Figure 1 summarizes some key trends in governance around the world during the last two centuries. The charts show time series for two indices from the Polity IV database: 'autocracy' (an index from 0 to 10) on the left and 'executive constraints' (an index from 1 to 7) on the right. The first variable measures the extent to which political power is concentrated in a narrow group, and the second measures checks and balances on executive power. The top panels display population-weighted cross-country means of the two Polity scores for each year. Clearly, the world as a whole now features much less autocratic government and substantially more constraints on the executive than two centuries ago.

Interestingly, Figure 1 highlights that this change has occurred largely on the extensive margin. It is not that there has been a more or less uniform decline in autocracy across countries. Instead, some countries have substantially reduced or eliminated autocratic government, while others have seen little improvement. This is illustrated in the bottom panels of the figure which plot the population-weighted cross-country standard deviations of the two Polity scores. To judge the size of these standard deviations, consider a benchmark standard deviation where all countries are divided into two blocs with the minimum and maximum scores respectively, and the fractions of countries in each bloc are set so that the average score is equal to the observed global mean. The dashed lines in the bottom panels show these maximum possible standard deviations in each particular year. For both Polity scores, the standard deviation has remained closer to the limit of full divergence rather than moving towards zero. Instead of political convergence, a persistent pattern of political specialization is observed.

This paper presents a theory of political specialization to understand how an increasingly interconnected world can nonetheless sustain diametrically opposed systems of government. According to the theory, some countries will uphold the rule of law with commitment to property rights, while others will consciously choose not to do so. Moreover, the theory implies that political specialization is to be expected even if all countries were ex ante identical. This political specialization is borne out of two key insights of the theory: (i) there is a diminishing marginal benefit of good government at the world level, but not at the country level; and (ii) there is a diminishing marginal cost of good government at the country level. 'Good government' is taken to mean the extent to which individuals appropriate the benefits of their own production.

The first key insight is due to the impact of good government on economic activity varying between different types of goods. Some production can occur even in despotic countries where individuals have no protection against expropriation, for example, extraction of natural resources.

Figure 1: *Means and standard deviations of political regime characteristics in the world*



Notes: Annual data (1800–2013) on 'Institutionalized Autocracy' (AUTOC, score between 0 and 10) and 'Executive Constraints' (XCONST, score between 1 and 7), with the means and standard deviations calculated from the population-weighted cross-section of countries. The spikes in the graphs are due to missing observations for large countries (mainly China) in some years. The series labelled 'largest possible' in the lower panels are the maximum possible standard deviations for distributions where the mean matches the data (corresponding to distributions with two point masses at the extreme scores).
Source: Polity IV Project, Center for Systemic Peace (http://www.systemicpeace.org/inscrdata.html).

However, production of goods that require long-term investments in physical capital or research and development ('rule-of-law intensive goods' hereafter) relies on investors expecting property rights to be enforced. Hence, at the world level, the marginal benefit of good government is declining owing to the diminishing marginal rate of substitution between rule-of-law intensive goods and other goods less sensitive to a country's political system. However, at the country level, access to world markets means that the marginal benefit of good government is constant for a small open economy that does not affect world prices.

The second key insight is that the marginal cost of good government is decreasing at the level of an individual country. What is meant by the marginal cost of good government is the marginal loss of rents received by those who hold power as a consequence of marginally better government. The idea is that while improvements in governance such as the rule of law will increase economic activity, they will also curtail the rents that incumbents are able to extract. This cost is not a resource cost: it is how much those in power stand to lose from better government in terms of the distribution of income, which they will set against the marginal benefit when deciding whether to resist it. Those in power in economies where the rule of law is pervasive will receive only little in rents, while those in autocracies will extract a substantial amount. Consequently, better government is 'cheap' at the margin to those in power in countries where the quality of governance is already high because they have only small rents to lose, while better government is 'expensive' at the margin to those in power in autocracies because they have a lot to lose.

Combined, these two insights lead to political specialization. Owing to the diminishing marginal cost of good government combined with the constant marginal benefit at the level of an individual country, countries will either fail to provide any security to investors, or will have full protection of property rights. Starting from autocracy, the first steps to better government cost incumbents more in terms of lost rents than they gain from increased economic activity. However, additional steps in that direction lead to further improvements in economic activity and progressively smaller losses of rents. If a tipping point is reached where the subsequent gains outweigh the initial losses then a country can move from autocracy to the rule of law, but if not then it is in the interests of incumbents not to take the first steps.

While this logic pushes individual countries to one of the extremes, the same reasoning does not apply to the world as a whole. The diminishing marginal benefit of good government at the world level means that the relative price of rule-of-law intensive goods is lower when good government is more widespread around the world. As the price of rule-of-law intensive goods falls in world markets, incumbents' calculations of the gains and losses from the rule of law tip in favour of autocracy. This leads to strategic substitutability in the choice of good government among countries, and implies there will be a distribution of political regimes around the world in equilibrium.

The world equilibrium features a symbiotic relationship between rule-of-law economies and despotic regimes. By producing goods requiring protection of property rights, rule-of-law economies raise the relative price of other goods such as natural resources and thus increase incentives for despotism in other countries. Conversely, despotic regimes generate a positive externality in other countries because cheap oil makes the rule of law more attractive elsewhere in the world.

In equilibrium, those in power in ex-ante identical countries are indifferent between the choice of the rule of law or despotism, and they always gain from trade with other countries. However, the ensuing political specialization leads to economic divergence as well. The economies that adopt the rule of law become substantially richer than the despotic regimes which produce no rule-of-law intensive goods. International trade thus benefits some countries but harms others.

The theory provides a way of reconciling the claim that corruption, rent-seeking, and insecure property rights create significant barriers to development in some countries with the fact that history is replete with examples of other countries having overcome precisely these challenges. Without a theory of political specialization, one must rely on exogenous differences in countries' 'political technologies' to understand why bad regimes persist in a world where trade works to equalize the marginal benefits of good government across countries. Such an approach is unappealing in light of the widespread imitation of other successful ideas and technologies around the world. Instead, the theory of political specialization presented here provides an account of how the same political frictions can give rise to both winners and losers — and why not every country can be a winner.

While the theory predicts that political divergence arises even in a world of ex-ante identical economies, the assumption of ex-ante identical countries is not essential to the argument. When ex-ante heterogeneity across countries is considered, comparative advantage determines which countries specialize in which political system, with good government emerging in countries with a comparative advantage in rule-of-law intensive goods. A 'natural resource curse' arises owing to the effects of a comparative advantage in natural resources on the incentives for those in power to resist improvements in governance that would restrict their ability to extract rents.

The theory has some important lessons on how the problem of despotic regimes should be addressed. One prediction is that exogenous improvements in one country's political institutions, perhaps brought about by well-intentioned international pressure or intervention, will tend to be counteracted by an increase in incentives for despotism in other countries. However, this does not preclude a role for international policy because the total number of despotic regimes in the world is affected by the relative price of rule-of-law intensive goods, which is in turn influenced by patterns of demand. The theory thus suggests that subsidies to rule-of-law intensive goods, for example, channelling resources to the study and development of technology-intensive alternative fuels, would be more effective in curbing despotism than efforts directed to affect the political systems of particular countries.

Raising the relative price of rule-of-law intensive goods in world markets requires a degree of co-operation among countries. Sadly, it is easier to find examples of international cooperation intended to lower this relative price. A cartel of countries with large endowments of natural resources that exploits its market power to push up the price of natural resources effectively imposes a tariff on rule-of-law intensive goods, which results in fewer countries having the rule of law in equilibrium.

In modelling the ideas discussed above, although the diminishing marginal benefit of rule-of-law intensive goods is a natural feature of any economic model, the diminishing marginal cost of good government ought not to be a primitive in any formal model of politics. In order to understand the nature and the properties of the costs of good government, it is essential to consider how questions of distribution are resolved through the use of political power. Following from this, it is appropriate to study the adoption of political institutions using a model where the distribution of power and resources is jointly determined. Upholding the rule of law requires that those who hold power are able to commit to pre-established rules and follow them even when that is not in their own interests ex post. But how can such commitment be achieved and what costs are entailed?

An important implication of the model used in this paper is that power sharing enables commitment to rules that would otherwise be time inconsistent. Power sharing supports the rule of law because it increases the number of potential losers from changes to the status quo, and thus raises the number of people willing to defend the status quo against threats from both inside and outside the group in power. However, sharing power requires sharing the rents associated with holding positions of power, which goes against the interests of incumbents and drives a wedge between the social return to investments and the return perceived by those in power.

For incumbents, the cost of improving governance through sharing power is related to the difference between incumbents' own incomes and the incomes of those outside the group in power. This difference is lower in economies with more power sharing because an additional member of the ruling group is less important to the overall group of incumbents when power is already shared more widely. That is why the first steps to the rule of law have the greatest private cost to incumbents, or in other words, why the marginal cost of good government is decreasing.

Since the mechanism leading to political specialization relies on international trade, an implication of the model is that the possibility of trade drives countries apart, economically and politically. Williamson (2011) shows that economic divergence did indeed follow the first wave of globalization in the early 19$^{\text{th}}$ century, when the third world 'fell behind'. While those countries with a comparative advantage in primary goods grew more than in previous centuries, the gap between rich and poor countries widened. Here, using data from the Polity IV Project, this paper shows that economic divergence was accompanied by political divergence. At the start of the 19$^{\text{th}}$ century, both industrialization and constraints on the executive (a measure of power sharing) were rare all around the world. In subsequent decades, industrialization and constraints on the executive appeared in many European countries but, consistent with the theory, neither showed up in the periphery.

The plan of the paper is as follows. Section 1.1 discusses the related literature. Section 2 describes the economic model and develops the theory of political specialization with the diminishing marginal cost of good government taken as an assumption. Section 3 derives the policy implications of the theory and extends the analysis to consider ex-ante heterogeneity between countries and the possibility of cooperation among countries. Section 4 sets up the full political model, describes the assumptions on the power struggle, and characterizes the equilibrium allocation of power and resources. Section 5 documents the divergence in power sharing following the first wave of globalization in the 19$^{\text{th}}$ century. Section 6 concludes.

1.1 Related literature

The full political model of section 4 builds on the framework proposed by Guimaraes and Sheedy (2015). That paper develops a model where an allocation of power and resources is established in the interests of incumbents. An allocation needs to survive the threat of rebellions from both inside and outside the group in power. The mechanism for contesting an allocation is the same no matter what that allocation prescribes, and there are no special individuals in the model: everyone is ex ante identical. The model assumes no exogenous technology to protect property rights, instead, power sharing allows for the rule of law to emerge endogenously. Here, the environment is extended to a world with two goods and many economies, and the equilibrium allocation of power and resources is characterized, taking account of interactions between countries' political systems. The two main theoretical results of this paper are the diminishing marginal cost of good government and the strategic substitutability in political regimes when there is international trade.[1]

There is now a substantial body of research showing that institutional quality has an important role in explaining international trade, as surveyed by Nunn and Trefler (2014). This is consistent with the reason why countries trade with each other in the theory here. Much of that work takes institutions as given. The literature studies how institutions affect trade flows (for example, Anderson and Marcouiller, 2002), and the pattern of comparative advantage (Levchenko, 2007, Nunn, 2007) and its dynamic effects (Araujo, Mion and Ornelas, 2016).

The paper is also related to a literature on the effects of trade on political institutions. Milgrom, North and Weingast (1990), Greif (1993), and Greif, Milgrom and Weingast (1994) combine historical analysis and game theory to understand how institutions in medieval times allowed merchants to solve the commitment problems that arise in large-scale (international) trade.[2] Acemoglu, Johnson and Robinson (2005) and Puga and Trefler (2014) study how international trade induced institutional change by enriching and empowering merchant groups. Acemoglu, Johnson and Robinson (2005) argue that the Atlantic trade led to better institutions in European countries with good initial conditions, while Puga and Trefler (2014) show how empowering merchants in Venice led to important institutional innovations up to the 13$^{\text{th}}$ century, but also to political closure and reduced competition thereafter. In a similar vein, special-interest groups play a key role in Levchenko's (2013) analysis of the impact of international trade on institutional quality.[3] This paper takes a different perspective. It looks at the world economy as a whole and studies how international trade affects the distribution of good government around the world in a model with ex-ante identical individuals and ex-ante identical countries. The model predicts that the world equilibrium is asymmetric, giving rise to winners and losers, a point that is absent from this literature.

In this sense, the paper is related to Acemoglu, Robinson and Verdier (2015), who study specialization in economic systems and also find an asymmetric world equilibrium. However, the question there is a very different one: understanding why different types of capitalism can co-exist, in particular, why we cannot all be like Scandinavians as opposed to Americans. Here, the question is why examples of good government in some countries co-exist with examples of abject failure in others — why some of us must be Venezuelans. In consequence, the model here is completely different from Acemoglu, Robinson and Verdier (2015). For example, the power struggle plays a central role here but is absent from their analysis, while here the paper abstracts from changes in the world technological frontier, which is central to their paper.

As in the 'new trade theory' models of Krugman (1979, 1980), the model here predicts a substantial amount of trade between ex-ante identical economies. Those papers assume production technologies with increasing returns, so countries specialize in different varieties of goods to exploit economies of scale, and trade benefits all countries. Here, there is a form of increasing returns to

[1]There are alternative models related to power sharing (Acemoglu and Robinson, 2000, Jack and Lagunoff, 2006, Bai and Lagunoff, 2011), but power sharing in those papers is an extension of the democratic franchise. In contrast, power sharing in Guimaraes and Sheedy (2015) is connected to the emergence of the rule of law.

[2]These ideas are developed further in Greif (2006).

[3]See also Grossman and Helpman (1994) and Do and Levchenko (2009).

producing rule-of-law intensive goods (from the point of view of those in power) because a diminishing marginal cost of good government arises endogenously from the political environment. That leads to political specialization, and trade benefits those in power anywhere in the world irrespective of political system — but not all countries gain from trade.

The paper is also related to the large literature on the 'natural resource curse' working through political institutions.[4] Robinson, Torvik and Verdier (2006) and Mehlum, Moene and Torvik (2006) study the role of institutions in understanding the natural resource curse. While in those papers the curse is a consequence of bad institutions, here the key institutional variable (power sharing) is endogenously determined and is a consequence of having a comparative advantage in natural resources. There are models in which natural resources distort rulers' choices (for example, Acemoglu, Verdier and Robinson, 2004, Caselli and Cunningham, 2009, Caselli and Tesei, 2016), but the distinguishing and important implication of this paper is that for the world as a whole, the number of despotic regimes depends on the demand for natural resources, not on the supply.

There is also a large body of work based around the idea that trade hurts economies that specialize in primary goods. One possibility is that some sectors give rise to positive externalities on the whole economy (or within an industry) through knowledge creation.[5] In this paper, trade harms those economies that fail to establish institutions conductive to development, but not because of any intrinsic disadvantage of producing primary goods.[6]

Last, the paper is broadly related to discussions of democratization in the social sciences.[7] Following the demise of the Soviet Union and the end of the cold war in the early 1990s, Fukuyama (1992) famously predicted the 'end of history', arguing that the days of autocratic regimes were numbered. Reality, however, has not been so kind, and Fukuyama has since acknowledged that autocratic regimes have been stubbornly persistent (Fukuyama, 2011).[8] More systematically, using the methods developed in the literature on testing for convergence in levels of GDP per person across countries, the lack of cross-country convergence in Polity scores has been noted in work by Goorha (2007).

2 A theory of political specialization

This section develops the theory of political specialization in a simple model that takes the diminishing marginal cost of good government as an assumption.

The world comprises a measure-one continuum of countries each containing a measure-one continuum of individuals. Individuals within a country are indexed by $\imath \in [0,1]$, countries by $\jmath \in [0,1]$.

[4]For a discussion of the empirical evidence on the natural resource curse, see Sachs and Warner (2001) and Van der Ploeg (2011).

[5]See, for example, Krugman (1987), Rodrik (1996), and Melitz (2005).

[6]There is also a large literature in sociology that attempts to explain underdevelopment as the result of rich countries exploiting poor ones, so-called 'dependency theory'. See, for example, Cardoso and Faletto (1979).

[7]Huntington (1993) is an influential example.

[8]There is much work in political science on the survival of autocracies (for example, Gandhi and Przeworski, 2007), but that literature focuses on individual countries, while this paper studies the political equilibrium of the world as a whole.

2.1 The economy

There are two goods in the world, an *endowment* good (E) and an *investment* good (I), the names referring to how the goods are obtained. The only use of both goods is consumption. All individuals throughout the world have preferences over consumption of the two goods represented by the consumption aggregator:

$$C = \frac{c_{\mathrm{E}}^{1-\alpha} c_{\mathrm{I}}^{\alpha}}{(1-\alpha)^{1-\alpha}\alpha^{\alpha}}, \qquad [2.1]$$

where c_{E} and c_{I} are respectively consumption of the endowment and investment goods. The parameter α, satisfying $0 < \alpha < 1$, indicates the relative importance of the investment good.

Suppose each individual $\imath \in [0, 1]$ is able to choose consumption levels $c_{\mathrm{E}}(\imath)$ and $c_{\mathrm{I}}(\imath)$ to maximize $C(\imath)$ subject to the budget constraint:

$$c_{\mathrm{E}}(\imath) + \pi c_{\mathrm{I}}(\imath) = Y(\imath), \qquad [2.2]$$

where π is the relative price of the investment good in terms of the endowment good. Individual $\imath$'s disposable income is $Y(\imath)$ (in terms of the endowment good as numeraire), which takes account of the taxes and transfers specified by the political system.[9] Maximizing the consumption aggregator (2.1) subject to (2.2) implies the first-order condition:

$$\frac{c_{\mathrm{I}}(\imath)}{c_{\mathrm{E}}(\imath)} = \frac{\alpha}{(1-\alpha)\pi}, \qquad [2.3]$$

and substituting this into the budget constraint (2.2) leads to the demand functions:

$$c_{\mathrm{E}}(\imath) = (1-\alpha)Y(\imath), \quad c_{\mathrm{I}}(\imath) = \frac{\alpha Y(\imath)}{\pi}, \quad \text{and} \quad C(\imath) = \frac{Y(\imath)}{\pi^{\alpha}}, \qquad [2.4]$$

where the final equation is the maximized value of the consumption aggregator (2.1).

When the countries of the world are in contact with each other, the endowment and investment goods can be exchanged in perfectly competitive world markets (but individuals cannot move between countries). The relative price of the investment good in terms of the endowment good in those markets is denoted by π^*, which all countries take as given. A country must satisfy its international budget constraint:

$$x_{\mathrm{E}} + \pi^* x_{\mathrm{I}} = 0, \qquad [2.5]$$

where x_{E} and x_{I} respectively denote the country's net exports of the endowment and investment goods.

In a given country, all individuals exogenously receive an amount q of the endowment good. Output K of the investment good (the capital stock) is endogenous and is described below. Conditional on K, the country's resource constraints are:

$$c_{\mathrm{E}} + x_{\mathrm{E}} = \mathrm{q}, \quad \text{and} \quad c_{\mathrm{I}} + x_{\mathrm{I}} = K, \quad \text{where} \quad c_{\mathrm{E}} = \int_0^1 c_{\mathrm{E}}(\imath)\mathrm{d}\imath \; \text{and} \; c_{\mathrm{I}} = \int_0^1 c_{\mathrm{I}}(\imath)\mathrm{d}\imath, \qquad [2.6]$$

[9] The budget equation implicitly assumes that no taxes or subsidies affect the relative price of the two goods. It is shown in the model of section 4 that those in power would not want to distort relative prices.

with $c_{\mathrm{E}}(\imath)$ and $c_{\mathrm{I}}(\imath)$ denoting the (non-negative) quantities of the endowment and investment goods consumed by individual $\imath \in [0, 1]$, and c_{E} and c_{I} total consumption of the two goods.

With individuals choosing consumption of the two goods in perfectly competitive markets subject to incomes $Y(\imath)$, the resource constraints (2.6) are market clearing conditions. Given the capital stock K and net exports x_{E} and x_{I}, domestic markets clear at the following relative price of the investment good:

$$\tilde{\pi} = \frac{\alpha(\mathrm{q} - x_{\mathrm{E}})}{(1 - \alpha)(K - x_{\mathrm{I}})}, \qquad [2.7]$$

with individual incomes $Y(\imath)$ summing to national income Y (in units of the endowment good):

$$\int_0^1 Y(\imath)\mathrm{d}\imath = Y, \quad \text{where } Y = \mathrm{q} + \tilde{\pi}K + (\pi^* - \tilde{\pi})x_{\mathrm{I}}. \qquad [2.8]$$

These claims follow from (2.2), (2.3), and (2.4). The final term in the expression for national income above accounts for the possibility that domestic prices could diverge from international prices. Equations (2.4), (2.7), and (2.8) imply:

$$\int_0^1 C(\imath)\mathrm{d}\imath = C, \quad \text{where } C = \frac{Y}{\tilde{\pi}^{\alpha}} = \frac{(\mathrm{q} - x_{\mathrm{E}})^{1-\alpha}(K - x_{\mathrm{I}})^{\alpha}}{(1 - \alpha)^{1-\alpha}\alpha^{\alpha}}, \qquad [2.9]$$

with C being real GDP, equal to the sum of all individuals' maximized levels of consumption $C(\imath)$.

In each country, a positive fraction μ of individuals receives investment opportunities at random. Each investment opportunity leads to the production of one unit (a normalization) of the investment good if it is undertaken. Letting s denote the endogenous fraction of investment opportunities undertaken, the capital stock is:

$$K = \mu s. \qquad [2.10]$$

Taking an investment opportunity is rational only if individuals expect property rights will be protected when the investment comes to fruition. It is assumed that all investment would be optimal if property rights were fully protected, so that s is equal to the fraction of those receiving investment opportunities who expect their property rights to be protected. The strength of property rights in a country, as measured by $s \in [0, 1]$, is an endogenous variable determined by the country's political system. In what follows, the term 'good government' is used as a shorthand for the strength of property rights s.

2.2 The political friction

A country's political system affects both 'good government' s and the distribution $\{Y(\imath)\}$ of national income Y among individuals. A basic assumption is that institutions providing credible protection of property rights also affect the share of income taken by those individuals in power (referred to as the *incumbents*). The idea is that institutions which uphold the rule of law, which are necessary to guarantee individuals' property rights against those in power, are also likely to place more general limits on the amount of resources incumbents can appropriate for themselves. Mathematically, this political friction is represented by the equation:

$$Y_\mathrm{p} = \phi(s)Y, \quad \text{where } \phi'(s) < 0, \tag{2.11}$$

with Y_p denoting the income of individuals in power. The share of national income appropriated by individuals in power is $\phi(s)$, which is assumed to be decreasing in the strength of property rights s.

The function $\phi(s)$ embeds the effect of all political constraints (de facto, not only de jure) on the incomes of those in power, conditional on operating in a political system able to provide sufficiently strong property rights for a fraction s of investors to take on investment opportunities.[10] Here, the negative relationship (2.11) between s and the share of income appropriated by incumbents is simply taken as an assumption. In section 4, equation (2.11) arises as a result from a model based on more primitive political frictions. Subject to the constraint (2.11), those in power choose the quality of governance s to maximize their own consumption payoff C_p.

Using (2.4), (2.9), and (2.11), the payoff of those in power and the effect of a marginal improvement in governance on the payoff are:

$$C_\mathrm{p} = \phi(s)C, \quad \text{and} \quad \frac{\partial C_\mathrm{p}}{\partial s} = \phi(s)\left(\frac{\partial C}{\partial s} - \gamma(s)C_\mathrm{p}\right), \tag{2.12}$$

where the function $\gamma(s)$ is defined as follows:

$$\gamma(s) \equiv -\frac{\phi'(s)}{\phi(s)^2}. \tag{2.13}$$

According to (2.12), an increase in s has two effects on the payoff of those in power. First, better government leads to a larger K (from 2.10) and an increase in real GDP C available for consumption (from 2.9). The term $\partial C/\partial s$ is referred to as the *marginal benefit of good government*, which is multiplied in (2.12) by the share $\phi(s)$ received by those in power.

However, as the quality of government improves, the share $\phi(s)$ of national income appropriated by those in power will decline (see 2.11). This is the second effect of s on C_p in (2.12). Using the function $\gamma(s)$ defined in (2.13), the marginal loss to those in power from higher s reducing $\phi(s)$ is expressed in terms comparable to the marginal benefit of good government (i.e. scaled to reflect incumbents' share of the total pie), which means that $\gamma(s)C_\mathrm{p}$ is subtracted from $\partial C/\partial s$. The value of $\gamma(s) > 0$ gives the marginal loss as a fraction of the consumption C_p received by those in power, which is referred to as the *marginal cost of good government*. Note that the cost of good government is not a resource cost, and hence not a social cost (it is not deducted in the expression for real GDP in 2.9). It is a private cost borne by incumbents due to a less favourable distribution of income when property rights are strengthened.

The key assumption underpinning the theory of political specialization is that the marginal cost of good government is diminishing:[11]

$$\gamma'(s) < 0. \tag{2.14}$$

[10] These constraints include avoiding being overthrown by a popular uprising or a coup d'état, or being impeached or voted out of office.

[11] Note that by assuming a particular marginal cost function $\gamma(s)$, the differential equation in (4.12) defining the

This marginal cost is the marginal loss of rents extracted by those in power when the quality of governance improves. The idea is that the first steps to better government are very costly for those in power because they have substantial rents to lose when governance is initially poor. Where government is already very good, those in power receive relatively little in rents, and thus do not have much to lose by improving it further. Here, (2.14) is taken as an assumption, but in section 4, it is derived as a result from a model based on more primitive political frictions.

The behaviour of the marginal cost of good government is important in determining the curvature of the objective function of those in power. Differentiating (2.12) with respect to s again and evaluating at a critical point:

$$\left.\frac{\partial^2 C_{\mathrm{p}}}{\partial s^2}\right|_{\frac{\partial C_{\mathrm{p}}}{\partial s}=0} = \phi(s)\left(\frac{\partial^2 C_{\mathrm{p}}}{\partial s^2} - \gamma'(s) C_{\mathrm{p}}\right). \qquad [2.15]$$

The sign of this second derivative determines whether the incumbent payoff C_{p} is globally quasi-concave or quasi-convex in good government s. The formula above shows that this depends on the sign of $\partial^2 C/\partial s^2$, that is, whether the marginal benefit of good government $\partial C/\partial s$ is increasing or decreasing, and on the sign of $\gamma'(s)$, that is, whether the marginal cost of good government $\gamma(s)$ is increasing or decreasing. The assumption (2.14) of a diminishing marginal cost of good government is one reason why C_{p} might be a quasi-convex function of s.

Finally, for the political frictions described above ever to be relevant, it is necessary that the marginal cost of good government at the best quality of government is not too small:

$$\gamma(1) > \frac{\alpha}{\phi(1)}. \qquad [2.16]$$

This condition is assumed in what follows.

2.3 Benchmark case: An economy in autarky

First consider the case of an economy not in contact with other countries (autarky). With no access to international markets, net exports of each good are equal to zero ($x_{\mathrm{E}} = 0$ and $x_{\mathrm{I}} = 0$, which replace the international budget constraint 2.5). Using (2.9) and (2.10), real GDP is given by:

$$C = \frac{\mathrm{q}^{1-\alpha}\mu^{\alpha}s^{\alpha}}{(1-\alpha)^{1-\alpha}\alpha^{\alpha}}. \qquad [2.17]$$

The marginal benefit of good government is:

$$\frac{\partial C}{\partial s} = \mu\hat{\pi}^{1-\alpha}, \quad \text{where } \hat{\pi} = \frac{\alpha \mathrm{q}}{(1-\alpha)\mu s}, \qquad [2.18]$$

with $\hat{\pi}$ being the market-clearing relative price (see 2.7) of the investment good in the absence of international trade ($\hat{}$ is used to denote a variable in the case of autarky). Improvements in

relationship between $\gamma(s)$ and $\phi(s)$ can be solved to deduce the implied function $\phi(s)$ in equation (2.11):

$$\phi(s) = \frac{\phi(0)}{1 + \phi(0)\int_{z=0}^{s}\gamma(z)\mathrm{d}z},$$

where $\phi(0)$ is a given positive constant. This is decreasing in s if $\gamma(s)$ is positive.

governance allow more of the investment good to be produced, but have no effect on the endowment good. The marginal benefit of good government therefore depends positively on the equilibrium relative price $\hat{\pi}$ of the investment good. However, in autarky, greater output of the investment good reduces its equilibrium price, so the marginal benefit of good government is diminishing.

Proposition 1 *Let $\varepsilon(s)$ denote the elasticity (in absolute value) of the marginal cost of good government $\gamma(s)$. Assuming that*

$$\varepsilon(s) < 1 - \alpha, \quad \text{where } \varepsilon(s) \equiv -\frac{s\gamma'(s)}{\gamma(s)}, \qquad [2.19]$$

for an economy in autarky:

(i) The payoff C_{p} of those in power is a strictly quasi-concave function of s.

(ii) The payoff-maximizing value of s is given by the unique solution of the following equation:

$$\alpha = \hat{s}\gamma(\hat{s})\phi(\hat{s}), \qquad [2.20]$$

which satisfies $0 < \hat{s} < 1$.

PROOF See appendix A.1 ■

Note the elasticity (in absolute value) of the marginal benefit of good government from (2.18) is:

$$-s\frac{\partial^2 C}{\partial s^2} \bigg/ \frac{\partial C}{\partial s} = 1 - \alpha,$$

which means that the marginal benefit declines by $1-\alpha$ percent for every one-percent improvement in the quality of governance. The rate of decline reflects the weight $1-\alpha$ attached to the endowment good, for which the level of output does not depend on good government. As shown by (2.15), the diminishing marginal benefit of good government is one reason why the objective function of those in power might be quasi-concave. This could be offset by a declining marginal cost of good government, but under the assumption that the marginal cost does not decline at a faster rate than the marginal benefit (condition 2.19), the objective function is globally quasi-concave. Note that (2.19) holds automatically if the marginal cost of good government were constant or increasing.

When the objective function is globally quasi-concave, there is an interior solution for s reflecting the optimal trade-off at the margin between the benefits and the costs of good government from the perspective of those in power. The first improvements in governance have an extremely large marginal benefit because the investment good is scarce. As government improves, the scarcity of the investment good is reduced, implying a lower marginal benefit from further improvements. Eventually a point is reached where the marginal benefit equals the marginal cost, and this is the optimal choice of governance for those in power.

The equilibrium quality of governance $\hat{s}$ depends only on the parameter α and the function $\phi(s)$ (from which the marginal cost function $\gamma(s)$ is derived). The size of the endowment q and the quantity of investment opportunities μ have no effect on governance because changes in quantities would lead to opposite changes in prices, and total values are what matter to incumbents.

2.4 Government in a small open economy

Now suppose a country can trade investment and endowment goods in international markets at relative price π^*. Those in power choose good government s as before to maximize their payoff C_{p} subject to the political constraint (2.11). They also choose the country's net exports x_{E} and x_{I} subject to the international budget constraint (2.5), taking the world price π^* as given.

The choice of quantities of net exports by those in power can be interpreted equivalently, and more naturally, as the choice of a tariff or subsidy on imports or exports of the investment good that drives a wedge between the domestic market-clearing price $\tilde{\pi}$ and the world price π^*:

$$\tilde{\pi} = (1+\tau)\pi^*, \qquad [2.21]$$

where τ denotes the tariff (if positive, or subsidy, if negative).[12] Individuals can buy and sell goods in international markets subject to the tariff or subsidy.

Given a tariff τ in (2.21) and the domestic supplies q and K of the endowment and investment goods, the international budget constraint (2.5) and market clearing (2.7) imply net exports are:

$$x_{\mathrm{E}} = \frac{\alpha \mathrm{q} - (1-\alpha)(1+\tau)\pi^* K}{1+(1-\alpha)\tau}, \quad \text{and} \quad x_{\mathrm{I}} = \frac{(1-\alpha)(1+\tau)\pi^* K - \alpha \mathrm{q}}{(1+(1-\alpha)\tau)\pi^*}. \qquad [2.22]$$

These satisfy the international budget constraint (2.5) and take on the full range of values consistent with the resource constraints (2.6) and non-negative consumption as τ varies over its maximum range $(-1 < \tau < \infty)$. The choice of x_{E} and x_{I} by those in power is therefore equivalent to a choice of τ.

Substituting net exports from (2.22) into (2.9) implies real GDP is given by:

$$C = \frac{(1+\tau)^{1-\alpha}}{1+(1-\alpha)\tau} \frac{\mathrm{q} + \pi^* K}{\pi^{*\alpha}}. \qquad [2.23]$$

The second term above is the real value of the country's output calculated using international prices (see 2.8 and 2.9), with the first term representing the distortions caused by any tariffs or subsidies. The first term is a strictly quasi-concave function of τ with a critical point at $\tau = 0$. Those in power want to maximize $C_{\mathrm{p}} = \phi(s)C$, and since $\phi(s)$ depends only on s and not directly on net exports, it follows that $\tau = 0$ is chosen to minimize distortions.[13] Free trade is therefore always in the interests of those in power.

With no restrictions on the flow of goods between domestic and foreign markets ($\tau = 0$), real GDP and the marginal benefit of good government are obtained from (2.10) and (2.23):

$$C = \frac{\mathrm{q} + \mu \pi^* s}{\pi^{*\alpha}}, \quad \text{and} \quad \frac{\partial C}{\partial s} = \mu {\pi^*}^{1-\alpha}. \qquad [2.24]$$

Since a small open economy takes the world price π^* as given, the marginal benefit of good government is independent of s and is equalized across countries through trade. With $\partial^2 C/\partial s^2 = 0$ and the assumption (2.13) of a diminishing marginal cost of good government, equation (2.15) implies the payoff of those in power is a quasi-convex function of the quality of governance s.

[12] Since $\tilde{\pi}$ and π^* are relative prices in terms of the endowment good, the effects of a tariff on the endowment good are equivalent here to subsidizing the investment good, and vice versa.

[13] Note that the independence of $\phi(s)$ from net exports and τ is an assumption here, but this property is a result of the model of section 4 based on more primitive political frictions.

Proposition 2 (Political specialization) *For a small open economy:*

(i) Free trade ($\tau = 0$ in 2.21) is optimal for those in power.

(ii) The payoff C_{p} of those in power is a strictly quasi-convex function of the quality of governance s. Hence, the payoff-maximizing choice is either $s = 0$ or $s = 1$.

(iii) The choice of $s = 1$ maximizes the payoff of those in power if:

$$\frac{\mu\pi^*}{\mathrm{q}} \geq \Phi, \quad \textit{where} \;\; \Phi \equiv \phi(0)\int_0^1 \gamma(s)\mathrm{d}s. \qquad [2.25]$$

(iv) Irrespective of whether $s = 0$ or $s = 1$ is chosen, those in power are strictly better off compared to the case of autarky (for any world price π^*). Countries where $s = 1$ is chosen also have higher real GDP than under autarky.

Proof See appendix A.2. ■

As the objective function of those in power is a quasi-convex function of the strength of property rights s, those in power will either choose to have property rights so weak that there is no investment ($s = 0$) or property rights sufficiently strong so that all profitable investment opportunities are taken ($s = 1$). These two extremes of governance are referred to as *despotism* and the *rule of law* respectively.

Why do those in power favour one of the two extremes of governance? The first steps to the rule of law have the greatest private cost to those in power (the assumption 2.14), but this was also true in the case of autarky. Differently from autarky, the marginal benefit of progress towards the rule of law is constant in a small open economy. The country can import investment goods that it cannot produce if the rule of law is too weak, or alternatively, it can export a surplus of investment goods to the rest of the world if its domestic market is too small. Consequently, the marginal benefit of good government does not decline as governance improves. This means that it is never optimal to choose a level of governance where the marginal benefit equals the marginal cost: this is the point at which marginal improvements in governance begin to pay off for those in power.

If the subsequent net gains from better government do not outweigh the initial net losses, it is optimal for those in power to choose despotism ($s = 0$), while the rule of law ($s = 1$) is optimal if overall the gains outweigh the losses. The total benefit of moving from a despotic regime to the rule of law is adding $\mu\tilde{\pi}$ to national income. The total cost (to incumbents) sums the marginal losses of rents, equivalent to a multiple Φ from (2.25) of the initial national income q. This explains the condition for $s = 1$ to be payoff maximizing in (2.25).

The first panel of Figure 2 below shows the consumption of an incumbent in autarky and in a small open economy as functions of the quality of governance s. The open economy case is shown for both for an arbitrary world price π^* and a world price equal to the autarky market-clearing price $\hat{\pi}$. For an open economy that happens to face the autarky price in world markets, net exports would be zero, hence incumbents' payoff (and everything else) would be the same as in autarky. In both the open economy and in autarky, the marginal cost of good government (as perceived by incumbents) equals its marginal benefit. In autarky, that is the point where the incumbent payoff is maximized, but in an open economy where the world price is taken as given, that is the point at which improvements in governance begin to raise the incumbent payoff.

It can be seen from Figure 2 that incumbents will gain from international trade. Trade allows those in power to exploit gains from political specialization that arise from the diminishing marginal cost of good government. Since incumbents' share of real GDP declines with good government, a country where incumbents choose $s = 1$ must also have a higher real GDP than in autarky.

2.5 The world equilibrium

At the world level, net exports of both endowment and investment goods must sum to zero:

$$\int_0^1 x_{\mathrm{E}}(\jmath)\mathrm{d}\jmath = 0, \quad \text{and} \quad \int_0^1 x_{\mathrm{I}}(\jmath)\mathrm{d}\jmath = 0. \tag{2.26}$$

The world relative price π^* adjusts to ensure that the market clearing conditions in (2.26) hold, noting that either of the market clearing conditions implies the other because countries must satisfy their international budget constraints (2.5) at all prices.

For small open economies, Proposition 2 shows that it is not in the interests of those in power to impose tariffs or subsidies ($\tau = 0$). Hence, using (2.22), a country $\jmath$'s net exports are given by:

$$x_{\mathrm{E}}(\jmath) = \alpha \mathrm{q}(\jmath) - (1-\alpha)\pi^* K(\jmath), \quad \text{and} \ \ x_{\mathrm{I}}(\jmath) = (1-\alpha)K(\jmath) - \alpha \mathrm{q}(\jmath)/\pi^*, \tag{2.27}$$

where $\mathrm{q}(\jmath)$ and $K(\jmath)$ are the supplies of the endowment good and investment good in country $\jmath$. The key result of Proposition 2 is that each country will choose either $s = 0$ or $s = 1$, implying $K(\jmath) = 0$ or $K(\jmath) = \mu$ respectively (see 2.10). Let ω denote the fraction of countries around the world choosing $s = 1$. By integrating over countries in (2.27), world market clearing (2.26) is obtained at price:

$$\pi^* = \frac{\alpha \mathrm{q}^*}{(1-\alpha)K^*}, \quad \text{where} \ \ K^* = \mu\omega \ \ \text{and} \ \ \mathrm{q}^* = \int_0^1 \mathrm{q}(\jmath)\mathrm{d}\jmath, \tag{2.28}$$

with q^* and K^* denoting the world supplies of the two goods.

The baseline assumption in what follows is that there are no differences ex ante between countries. This means all countries share a common supply $\mathrm{q}(\jmath) = \mathrm{q} = \mathrm{q}^*$ of the endowment good.

Proposition 3 *The world equilibrium with ex-ante identical countries has the following features:*

(i) Strategic substitutability: The world market clearing price (2.28) *implies that condition* (2.25) *for the optimality of $s = 1$ is equivalent to the fraction ω of other countries with $s = 1$ being sufficiently low:*

$$\omega \leq \frac{\alpha}{(1-\alpha)\Phi}. \tag{2.29}$$

(ii) The equilibrium fraction of economies with $s = 1$ and the equilibrium world price are:

$$\tilde{\omega} = \frac{\alpha}{(1-\alpha)\Phi}, \quad \textit{and} \ \ \tilde{\pi}^* = \frac{\mathrm{q}\Phi}{\mu}, \tag{2.30}$$

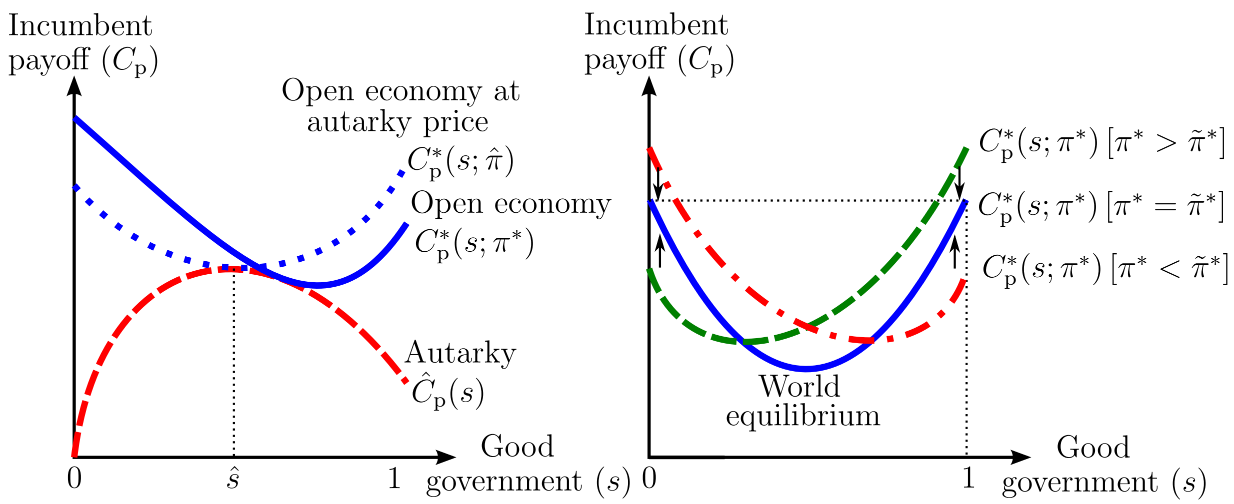
and it is always the case that $0 < \tilde{\omega} < 1$. Those in power receive the same payoff irrespective of whether $s = 0$ or $s = 1$ is chosen, while countries with $s = 1$ have higher real GDP than those with $s = 0$.

Proof See appendix A.3. ■

While the logic of Proposition 2 pushes individual countries to the extremes of governance, the same reasoning does not apply to the world as a whole. At the global level, prices depend on how much of the investment good is produced and hence on the number of economies with the rule of law. If more economies adopt the rule of law, the price of the investment good π^* falls, and thus the marginal benefit of good government is diminishing at the world level (see 2.24). This means that choices of political systems are strategic substitutes across countries: an increase in the global prevalence of the rule of law tilts the balance in favour of despotism for others.

In equilibrium, the world price adjusts in order to equate incumbent payoffs under the two extreme political systems, as illustrated in the right panel of Figure 2. If the rule of law were preferred by incumbents and adopted everywhere, the price π^* would fall, which raises incumbents' payoff from despotism until a point of indifference is reached. The world equilibrium thus features a bimodal distribution of political systems, with the extent of the rule of law limited by the size of the global market for rule-of-law intensive goods (investment goods in the model). Even without a predisposition to despotism, some economies will end up with a despotic political system, while others will end up with the rule of law even in the absence of any cultural or technological advantage.

Figure 2: *Comparison of autarky and open economies, and the world equilibrium*



While incumbents are indifferent between $s = 0$ and $s = 1$, these economies will be very different. The rule-of-law economies will produce $\mathrm{q} + \mu\pi^*$, while the despotic economies will only produce q. By engendering political specialization, international trade leads to economic divergence.

Underlying this political specialization is the symbiotic relationship between despots and incumbents in rule-of-law economies through international trade. The existence of the rule of law elsewhere in the world allows a despotic regime to import what its own political system precludes it from producing, allowing incumbents to focus on maximizing their rents. The existence of despotism elsewhere in the world allows an economy with the rule of law to capture a greater share of the global market for rule-of-law intensive goods, allowing incumbents to benefit from increasing returns due to the diminishing marginal cost of good government they face.

There is an analogy here with the 'new trade theory' models of Krugman (1979, 1980) where ex-ante identical economies specialize in producing different goods and thus trade with each in order to exploit increasing returns in production. Here, countries that are ex-ante identical economically and politically specialize in different political systems, which implies they become different ex post and thus trade with each other. Comparative advantage in institutionally-intensive goods would be seen to explain observed trade flows, consistent with the evidence in Nunn and Trefler (2014). However, comparative advantage is endogenous here, and there is a fundamental difference from 'new trade theory' because there are no increasing returns in production itself. This means the welfare implications of international trade are very different and the usual argument that there are gains from exploiting comparative advantage does not apply. Those countries that have low institutional quality in equilibrium actually lose by trading internationally.

The effect of international trade on global output is ambiguous. The fundamental friction is that good government reduces the rents those in power are able to extract, so incumbents have too little incentive to extend the reach of the rule of law. There are no distortions coming from trade itself, but the possibility of international trade affects how the consequences of the fundamental political friction are spread across countries. The usual reasons for international trade are absent from the model and so are the usual gains from trade (in the absence of political frictions, no trade would occur in equilibrium). This means that international trade is close to a zero-sum game because it has little impact on world output, but strong distributional consequences.

3 Policy implications and extensions of the analysis

3.1 Policy implications

The theory of political specialization proposed here has some strong implications for how the problem of despotic regimes ought to be addressed. Suppose that a benevolent global power uses its policy instruments with the goal of improving political and economic outcomes around the world. The first policy instrument considered is the use of military force to bring about regime change and directly impose a particular quality of governance s on a country, rather than s being chosen by the country's own rulers. Offering aid payments conditional on the adoption of good governance to persuade rulers to do what the benevolent power wants is another interpretation of this policy instrument (as carrot rather than stick). The second policy instrument considered gives the benevolent power the ability to insist a country imposes tariffs or subsidies on its trade with the rest of the world, that is, the benevolent power can directly set τ in a country, rather than this being chosen by its own rulers. Alternatively, this second instrument could be seen as a country or group of countries acting benevolently in choosing τ.

Proposition 4 *The two policy instruments have the following effects on the world equilibrium:*

(i) *If a fraction ω_0 of economies is forced to choose $s = 1$ then the equilibrium fraction $\tilde{\omega}$ of economies with $s = 1$ is unchanged as long as $\omega_0 \leq \tilde{\omega}$ for the initial $\tilde{\omega}$ (if $\omega_0 > \tilde{\omega}$ then only those economies that are forced to will have $s = 1$).*

(ii) *If a fraction σ of economies (all with $s = 1$) implement a subsidy ($\tau < 0$) on the investment good then this raises the equilibrium fraction of economies with $s = 1$. If all countries were to set a subsidy $\tau = -(1 - \alpha/(1 - \alpha)\Phi)$ then all countries would have $s = 1$ in equilibrium.*

PROOF See appendix A.4. ■

Surprisingly, direct intervention (even supposing it is feasible) turns out to have no effect whatsoever on the equilibrium fraction of countries with the rule of law (unless a point is reached where every country with good government has it imposed by external force). Owing to the strategic substitutability of political systems, an exogenous shift of a country from despotism to the rule of law must be counteracted in equilibrium by another country moving in the opposite direction. The key point here is that localized interventions are bound to fail owing to the general equilibrium effects on incumbents' incentives in other countries.

This negative result shifts the focus from what might be thought of as 'supply-side' policies to 'demand-side' policies. If a group of benevolent countries were to subsidize consumption of rule-of-law intensive goods then, all else equal, this would raise the world relative price of those goods and reduce the incentives for incumbents in other countries to choose a despotic regime (the effects would be analogous to an increase in the demand parameter α for rule-of-law intensive goods, see equation 2.30). This policy can in principle be effective at curbing despotism. In practice, though, it requires a sufficiently large number of countries to cooperate in implementing a policy that only benefits others and adds distortions to their own economies. If all countries were to subsidize rule-of-law intensive goods sufficiently then it would be possible to have an equilibrium where all countries have the rule of law without any distortions to the global allocation of resources. However, rulers of individual countries would still have an incentive to deviate from this policy.

3.2 Ex-ante heterogeneity between countries

It has been assumed up to this point that all countries are ex ante identical when analysing the world equilibrium. The result that specialization arises without any ex-ante heterogeneity highlights the strength of the mechanism pushing countries towards either despotism or the rule of law, but leaves open the path any particular country would take.

This section presents a simple extension with one dimension of ex-ante heterogeneity that explains which countries become despotic and which uphold the rule of law. The assumption is that countries differ in their endowments $\mathrm{q}(\jmath)$.[14] Specifically, given the world supply q^* of the endowment good, the distribution of relative endowments $\varrho(\jmath) = \mathrm{q}(\jmath)/\mathrm{q}^*$ across countries has a continuous probability distribution with cumulative distribution function $F(\varrho)$. This distribution has positive support and a mean of one.

[14] The analysis is isomorphic to heterogeneity in political frictions (the function $\phi(s)$ in 2.11) across countries.

In autarky, Proposition 1 has already shown that heterogeneity in endowments would have no effect on either the quality of governance or output of the investment good across countries. Differences in the supply of the endowment good are offset by opposite differences in its relative price. This means that any effects of ex-ante heterogeneity in a world of open economies must be due to its effects on specialization and trade.

In an open economy, the fundamental reason for specialization, as reflected in the quasi-convexity of the incumbent payoff, remains unchanged (see Proposition 2). However, the selection of countries that will be despotic and those that will establish the rule of law is no longer arbitrary. The criterion in (2.25) for the optimality of $s = 1$ determines selection when there is ex-ante heterogeneity.

Proposition 5 *Given a distribution function $F(\varrho)$ of relative endowments across countries:*

(i) There is a threshold $\tilde{\varrho}$ such that those countries with $s = 1$ have low relative endowments $\varrho \leq \tilde{\varrho}$.

(ii) There exists a unique equilibrium fraction $\tilde{\omega}$ of countries with $s = 1$, which is the solution of the equation $\tilde{\omega} = F(\alpha/(1-\alpha)\Phi\tilde{\omega})$ and satisfies $0 < \tilde{\omega} < 1$.

(iii) The equilibrium $\tilde{\omega}$ lies between the equilibrium with homogeneous endowments $\omega_0 = \alpha/(1-\alpha)\Phi$ and the fraction $\omega^* = F(1)$ of countries with an endowment below the global mean.

Proof See appendix A.5. ■

Economies with relatively small supplies of the endowment good obtain the rule of law in equilibrium; economies with large quantities of the endowment are condemned to despotism.[15] As before, the world equilibrium features a mixture of despotic and rule-of-law economies. Figure 3 depicts the consumption of incumbents and the per-person average level of consumption in the cross-section of economies. The consumption of incumbents is strictly increasing in the endowment q, especially so for despots because the gradient reflects the share each incumbent receives, which is greater in despotic countries (see 2.11). Consumption per person is also increasing in q, controlling for the political system. However, there is a discrete step down at the threshold between the rule of law and despotism. Crucially, at least some and possibly all economies with a large endowment are poorer than those with endowments low enough to have the rule of law. The model thus gives rise to a natural resource curse.

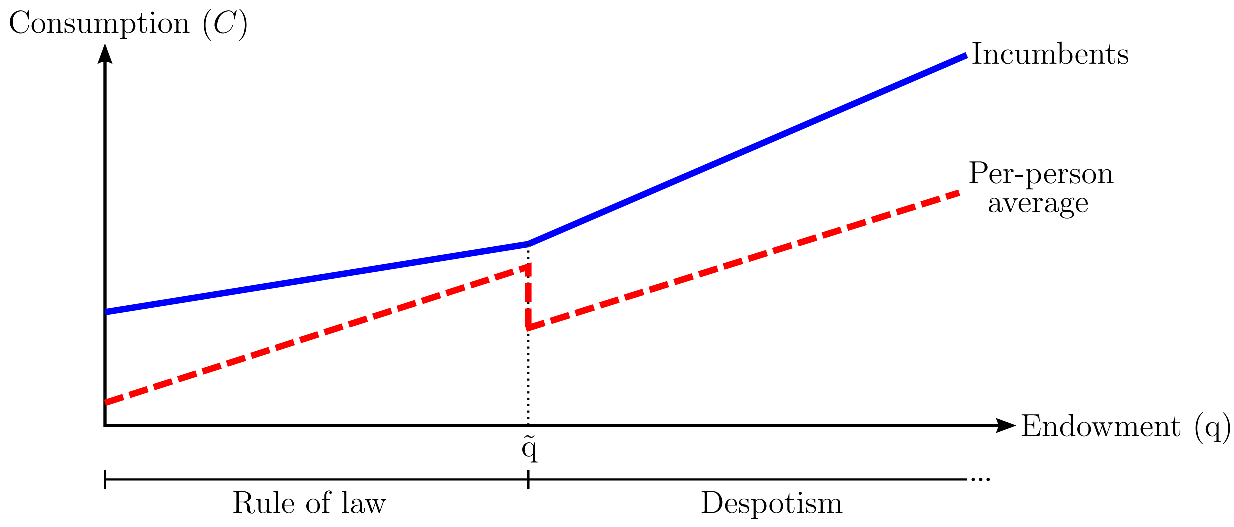
In general, the equilibrium fraction of rule-of-law economies could be larger or smaller than in the case of ex-ante identical countries. However, the final result of the proposition indicates that the rule of law will be more widespread when endowments are concentrated in a small group of countries.

3.3 International market power: Cartels

The analysis so far has only considered zero-measure countries that are price takers in world markets. This section considers a positive measure of countries with relatively large supplies of the endowment

[15] The output from an investment opportunity has been normalized to one unit of the investment good. This means q can also be interpreted as the quantity of the endowment good relative to the potential production of the investment good in a country.

Figure 3: *Ex-ante heterogeneity between countries*



good that can act together as a cartel.

To simplify the analysis, it is assumed that the cartel acts collectively, abstracting from its internal dynamics. This means that the cartel is essentially one large country with a government that maximizes the payoff of those in power as described earlier. Differently from before, the cartel knows that its choices of net exports will affect world prices. Formally, the cartel is a Stackelberg leader playing against an auctioneer who sets the world relative price, with all other countries being price takers in world markets. The cartel moves first in choosing net exports of the endowment good. Given this choice and the demand functions of the small open economies, the auctioneer chooses the world relative price to ensure that world markets clear.

There is a continuous distribution of endowments across countries with cumulative distribution function $F(\varrho)$ of the relative endowments $\varrho = \mathrm{q}/\mathrm{q}^*$. A regularity condition is imposed to ensure the world net demand for endowment goods is a well-behaved function: the distribution of log endowments must have a non-increasing reversed hazard rate.[16] Members of the cartel are drawn from those countries that would have $s = 0$ if they were small open economies.

Proposition 6 *As long as the cartel is not too large relative to the world and has a sufficiently large relative endowment, it remains optimal for the cartel to choose* $s = 0$.

(i) The cartel's pricing strategy is isomorphic to a tariff on imports of rule-of-law intensive goods ($\tau > 0$)*, implying* $\tilde{\pi} < \pi^*$. *The equilibrium world price* $\tilde{\pi}^*$ *of the investment good is lower than when the countries of the cartel act as small open economies.*

(ii) The equilibrium fraction $\tilde{\omega}$ *of countries with the rule of law is reduced by the presence of the cartel compared to a world of small open economies.*

Proof See appendix A.6. ■

[16]This requires that $\varrho F'(\varrho)/F(\varrho)$ is a weakly decreasing function of ϱ. The condition is satisfied by most common probability distributions, including the uniform, log Normal, exponential, and Pareto distributions.

As long as the cartel has a sufficiently large relative endowment, but is not too large relative to the world itself, it remains in the interests of those in power to have a despotic regime. In that case, countries in the cartel do not produce any of the investment good and export some of the endowment good. The cartel's pricing strategy is standard: in order to exploit its market power, the cartel exports less of the endowment good at a higher price. This can be implemented by a tariff on imports of rule-of-law intensive goods. Trade theory points out that tariffs might be optimal for large countries because part of the tax is effectively paid by foreigners. But from the perspective of the world as a whole, tariffs create inefficiencies by inhibiting some mutually beneficial exchanges.

In equilibrium at the world level, the cartel's actions make the endowment good more expensive (π^* is smaller). The analysis yields a novel implication: by effectively reducing the relative price of the rule-of-law intensive good, the presence of the cartel raises incentives for despotism, which leads to a smaller fraction of countries with the rule of law in equilibrium. A cartel of countries with large endowments is therefore the exact opposite of the policy implication for combating despotism from Proposition 4.

4 A model of the costs of good government

The theory of political specialization presented in section 2 was built around the assumption of a diminishing marginal cost of good government in protecting property rights (from the perspective of those who hold power), an idea we find intuitively plausible. However, we also take the view that matters such as the ability of those in power to create a system of government upholding the rule of law and the nature of the distributional consequences of such a system ought not to be assumptions, but should instead be analysed in terms of more fundamental political frictions. This section provides a model of the costs of good government in this spirit.

The model adds two features to the environment described in section 2. First, a struggle for political power whereby a group can establish an allocation of power and resources, but needs to avoid triggering rebellions that would see an alternative allocation established. Second, an incentive problem for investors whereby the technology for producing the investment good requires effort, but there is a time lag between making the effort and the good becoming available for consumption (the effort cost is sunk by this time). This time lag gives rise to a role for property rights and creates a commitment problem for those in power.

All individuals $\imath \in [0, 1]$ in a country are ex-ante identical. Individuals have the same preferences defined over consumption C, investment I, and *rebellion effort* R, and these preferences are represented by the utility function:

$$\mathscr{U} = \log C - I \log(1 + \theta) - \log(1 + R). \tag{4.1}$$

The utility function is logarithmic in consumption C, where C is the aggregator of the endowment and investment goods defined in (2.1).

As in section 2, a fraction μ of individuals receives investment opportunities at random, each allowing one unit of the investment good to be produced. Individuals must incur disutility $\log(1 +$

θ) if they take an investment opportunity (a binary choice, $I \in \{0,1\}$), where θ is a positive parameter.[17] This effort cost is sunk by the time the investment good is produced (any discounting of utility between investment and production is embedded in the parameter θ). Receiving an investment opportunity is private information, but whether or not it has been taken (the variable I) becomes common knowledge when the investment good is available for consumption.

Since an investment opportunity is private information, only the recipient can decide whether to take it, and this decision must maximize utility (4.1). Before individuals know whether they will receive an investment opportunity, their utility is given by the expected value of (4.1). The set of all individuals who invest ($I = 1$) is denoted by $\mathcal{I}$, and output of the investment good is given by equation (2.10), where $s = |\mathcal{I}|/\mu$ denotes the fraction of investment opportunities that are taken.

Individuals also receive disutility from any rebellion effort R, the role of which is explained below. The substantive implication of the functional form in (4.1) is that if an individual could obtain consumption C' instead of C by rebelling then the individual would be willing to exert no more than $R = (C' - C)/C$ units of rebellion effort. This means that the disutility from R units of rebellion effort is exactly compensated by the gain of a fraction R of consumption.

4.1 Allocations and rebellions

There is an *allocation* of power and resources, which will be determined endogenously. Allocations can be contested through *rebellions*, which lead to new allocations being established, a process referred to as the power struggle. The modelling here follows Guimaraes and Sheedy (2015).

An allocation specifies the set $\mathcal{P}$ of individuals currently in *power*, referred to as the incumbents. Each position of power confers an equal advantage on its holder in the event of any conflict, as described below. *Power sharing* $p = |\mathcal{P}|$ is defined as the measure of the group $\mathcal{P}$. The incumbent group $\mathcal{P}$ can have any size between 0% and 50% of the population ($0 \leq p \leq 1/2$). It is assumed that investment opportunities cannot be received by those individuals currently in power, but opportunities are otherwise randomly distributed among individuals.

An allocation also specifies how much individuals receive of each good, and how much of each good is exported or imported. Each individual's consumption of each good can depend on whether the individual is in power, and (for some or all individuals) whether an individual has taken an investment opportunity.[18] Formally, an allocation specifies a set $\mathcal{D}$ of individuals whose consumption is contingent on taking an investment opportunity (with $\mathcal{D} \cap \mathcal{P} = \emptyset$, as those in power do not receive investment opportunities). The set $\mathcal{K}$ of *capitalists* comprises those who have a consumption allocation that is contingent on investing and who have received and taken an investment opportunity ($\mathcal{K} = \mathcal{D} \cap \mathcal{I}$). The set $\mathcal{W}$ of *workers* comprises those individuals who are neither incumbents nor capitalists ($\mathcal{W} = [0,1] \backslash (\mathcal{P} \cup \mathcal{K})$). The amounts of the endowment good and the investment good

[17] The parameter θ can be interpreted as the cost of taking an investment opportunity expressed as a fraction of consumption (see equation 4.2 below).

[18] It would be possible to extend the analysis so that an allocation would specify a fully individual-specific consumption allocation, but that would add considerable complexity without necessarily affecting the results. In a related setting, Guimaraes and Sheedy (2015) allow for fully individual-specific consumption allocations, but find that in equilibrium, consumption is only contingent on being in power or investing.

consumed by each incumbent are denoted by c_{pE} and c_{pI}, for each capitalist by c_{kE} and c_{kI}, and for each worker by c_{wE} and c_{wI}. Note that an allocation directly specifies individuals' quantities of consumption of each good, which is different from section 2 where individuals were assumed to be able to exchange goods in competitive markets. Since the objective here is to study how political power interacts with private property, free exchange is not built in as an assumption.

Investment opportunities are private information, so an allocation cannot directly compel individuals to produce capital. Instead, by varying the investment-contingent consumption allocation, an allocation can determine the amount of investment through its effect on the fraction of individuals for whom investing is incentive compatible.[19] If investing leads to an individual receiving consumption C_{k} while not investing leads to consumption C_{w} (in terms of the aggregator 2.1) then the utility function (4.1) implies that investing is incentive compatible if:

$$C_{\mathrm{k}} \geq (1+\theta)C_{\mathrm{w}}. \tag{4.2}$$

The prevailing allocation implies the fraction of investment opportunities that are taken is:

$$s = \begin{cases} |\mathcal{D}|/(1-p) & \text{if (4.2) holds} \\ 0 & \text{otherwise} \end{cases}, \tag{4.3}$$

where the formula makes use of the random assignment of investment opportunities.

An allocation is formally defined as a collection $\mathscr{A} = \{\mathcal{P}, \mathcal{D}, c_{\mathrm{pE}}, c_{\mathrm{pI}}, c_{\mathrm{kE}}, c_{\mathrm{kI}}, c_{\mathrm{wE}}, c_{\mathrm{wI}}, x_{\mathrm{E}}, x_{\mathrm{I}}\}$, where the measure of $\mathcal{P}$ (power sharing p) is no more than 1/2, $\mathcal{D} \cap \mathcal{P} = \emptyset$, net exports x_{E} and x_{I} must satisfy (2.5) given the world price π^*, and the non-negative consumption allocation must satisfy (2.6) given the capital stock K from (2.10). The fraction s of investment opportunities taken is determined by the prevailing allocation according to (4.3).

The timing of events is depicted in Figure 4. An allocation is established, followed by opportunities for rebellions, with a new allocation established if a rebellion succeeds. Once there are no rebellions, investment opportunities are received. After individuals make investments, there are again opportunities for rebellions, with a new allocation established if a rebellion succeeds. Finally, endowments are received, and any investment goods that have been produced become available for consumption. The prevailing consumption allocation is then implemented.

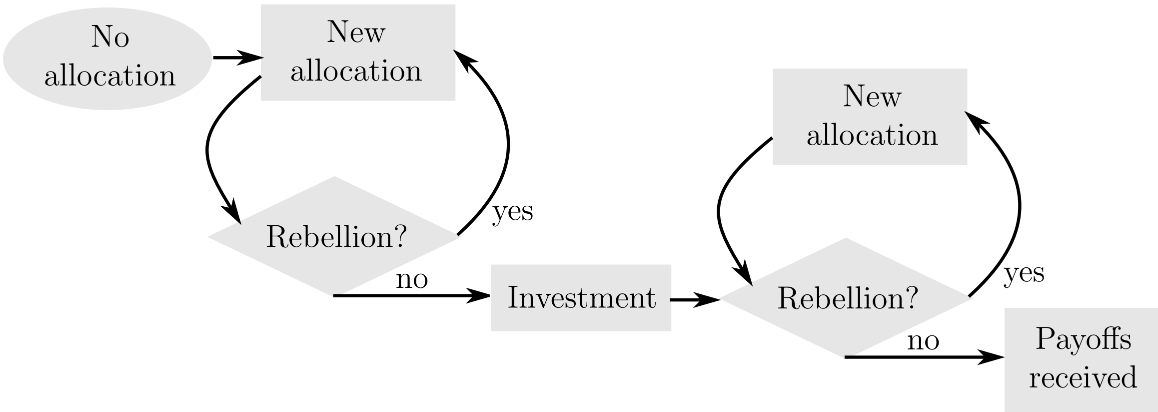
Note that utility from consumption is received at the final stage of the sequence of events in Figure 4, and there is no discounting of payoffs based on the number of rebellions that may have occurred. Any disutility from rebellion effort is additively separable between different rebellions.[20] At any stage of Figure 4, continuation payoffs are independent of any earlier investment effort or rebellion effort (these effort costs are sunk). For new allocations established after investment opportunities have been received, the fraction s and the capital stock K in (2.10) are state variables.

Rebellions are the only mechanism for changing an established allocation. A rebellion is described

[19] With heterogeneity in the effort cost of investing between different individuals, it would be possible for an allocation to determine the number of individuals for whom investing is incentive compatible through an investment-contingent consumption allocation that applies to everyone. See Guimaraes and Sheedy (2015) for an example of how this could be done.

[20] In the utility function (4.1), an individual who exerts rebellion effort in several rebellions receives disutility $\sum_{\ell} \log(1 + R_{\ell})$, where R_{ℓ} denotes effort in rebellion ℓ.

Figure 4: *Sequence of events*



by a *rebel faction* $\mathcal{R}$, an amount of (non-negative) rebellion effort $R(\imath)$ for each individual $\imath \in \mathcal{R}$ who belongs to the rebel faction, and a *loyal faction* $\mathcal{L}$ that defends the current allocation. A rebel faction can comprise those outside or inside the group currently in power, or a mixture of both. The loyal faction is drawn from those currently in power who do not join the rebel faction. Formally, a rebellion is a collection $\{\mathcal{L}, \mathcal{R}, R(\imath)\}$, where the sets $\mathcal{L}$ and $\mathcal{R}$ satisfy $\mathcal{L} \subseteq \mathcal{P}$ and $\mathcal{L} \cap \mathcal{R} = \emptyset$.

A rebellion succeeds if

$$\int_{\mathcal{R}} R(\imath)\mathrm{d}\imath > \int_{\mathcal{L}} \delta \mathrm{d}\imath, \qquad [4.4]$$

which requires that the strength of the rebel faction exceeds the strength of the loyal faction. Each faction's strength is the integral of the strengths of its members. The strength of individual $\imath \in \mathcal{R}$ in the rebel faction is the amount of rebellion effort $R(\imath)$ he exerts. Each individual $\imath \in \mathcal{L}$ in the loyal faction has strength measured by a positive parameter δ (the power parameter), which is obtained at no utility cost to these individuals.

The investment opportunity parameters μ and θ are assumed to satisfy the following bounds which depend on the power parameter δ:

$$\mu \leq \frac{\delta}{2(2+\delta)}, \quad \text{and} \quad \theta \leq \delta \min\left\{\frac{2(1+\delta+\delta^2)}{2+\delta^2}, \frac{4+5\delta+2\delta^2}{2(1+\delta)^2}\right\}. \qquad [4.5]$$

The interpretation of these restrictions is discussed later.

4.2 Equilibrium definition

In what follows, let $U(\imath)$ denote individual $\imath$'s continuation utility under a particular allocation (that is, utility excluding any sunk effort costs, and supposing the allocation prevails with no further rebellion effort exerted). The notation $'$ is used to signify any aspect of a new allocation that would be established following a rebellion.

An equilibrium allocation must be optimal in the sense of maximizing the payoff of incumbents, taking into account the threat of rebellions. Any rebellions must be rational in the sense defined below.

Definition 1 *A rebellion* $\{\mathcal{L}, \mathcal{R}, R(\imath)\}$ *against the current allocation* $\mathscr{A}$ *is rational given the subsequent allocation* $\mathscr{A}' = \{\mathcal{P}', \mathcal{D}', c'_{\mathrm{pE}}, c'_{\mathrm{pI}}, c'_{\mathrm{kE}}, c'_{\mathrm{kI}}, c'_{\mathrm{wE}}, c'_{\mathrm{wI}}, x'_{\mathrm{E}}, x'_{\mathrm{I}}\}$ *if:*

(i) All individuals in the rebel faction $\mathcal{R}$ *receive a position of power under the subsequent allocation yielding a payoff no lower than what the individual would receive under the current allocation, and the disutility of each individual's rebellion effort* $R(\imath)$ *does not exceed his utility gain from rebelling:*

$$\mathcal{R} = \{\imath \in \mathcal{P}' \mid U'(\imath) \geq U(\imath)\}, \quad \text{and} \;\; \log(1 + R(\imath)) \leq U'(\imath) - U(\imath). \qquad [4.6a]$$

(ii) The loyal faction $\mathcal{L}$ *(drawn from those currently holding positions of power who do not rebel) comprises those who would be worse off under the subsequent allocation:*

$$\mathcal{L} = \{\imath \in \mathcal{P} \backslash \mathcal{R} \mid U(\imath) > U'(\imath)\}. \qquad [4.6b]$$

(iii) Condition (4.4) *for a successful rebellion holds.* □

In a rational rebellion, the rebel faction $\mathcal{R}$ includes only individuals who would be in power under the subsequent allocation, which is an assumption designed to capture the incentive problems in inducing individuals to exert effort.[21] The maximum amount of rebellion effort exerted by an individual in the rebel faction has disutility equal to his utility gain from changing the allocation (see the utility function 4.1). Analogously, an individual in power will join the loyal faction $\mathcal{L}$ to defend the current allocation if this is in his own interest. With no uncertainty about the outcome, only those rebellions that would succeed can be rational.

The requirements for an allocation to be an equilibrium of the power struggle are now stated.

Definition 2 *An allocation* $\mathscr{A} = \{\mathcal{P}, \mathcal{D}, c_{\mathrm{pE}}, c_{\mathrm{pI}}, c_{\mathrm{kE}}, c_{\mathrm{kI}}, c_{\mathrm{wE}}, c_{\mathrm{wI}}, x_{\mathrm{E}}, x_{\mathrm{I}}\}$ *is an equilibrium of a stage of the power struggle in Figure 4 if the following conditions are satisfied:*

(i) ***Optimality for incumbents****: The allocation* $\mathscr{A}$ *maximizes the utility of incumbents when it is established, subject to:*

(ii) ***Rationality of rebels****: A rational rebellion occurs if according to Definition 1 there exists any rational rebellion against the current allocation* $\mathscr{A}$ *for some subsequent allocation* $\mathscr{A}'$*, subject to:*

(iii) ***Threats of rebellion are credible****: Any allocation* $\mathscr{A}'$ *established following a rebellion is itself an equilibrium of that stage of the power struggle.*

(iv) ***Independence of irrelevant history****: Allocations* $\mathscr{A}$ *and* $\mathscr{A}'$ *established at any two stages of the power struggle with the same continuation timeline and the same fundamental (payoff-relevant) state variables are identical up to a permutation of identities.* □

Definition 2 states that an equilibrium allocation must be established in the interests of incumbents (first condition) subject to avoiding any opportunity for rational rebellion (second condition), where the range of possible rational rebellions is itself limited by the set of equilibrium allocations that could

[21] See Guimaraes and Sheedy (2015) for a discussion of the rebellion mechanism used here.

be established following a rebellion (third condition).[22] Essentially, the third equilibrium condition precludes the rebels making a binding commitment to an allocation that is not in their interests ex post — for example, an allocation that would give rebels an incentive to exert more effort now, but which would not be optimal once the rebellion is over. Of those allocations satisfying the first three conditions, any that depend (apart from individual identities) on payoff-irrelevant histories are then deleted to leave a set of equilibrium allocations (fourth condition). At the pre-investment stage in Figure 4, there are no fundamental state variables, therefore equilibrium allocations in any round of the power struggle must be the same apart from changes in the identities of those in power. At the post-investment stage in Figure 4, the capital stock will be a fundamental state variable.

4.3 Properties of the equilibrium allocation

Allocations that trigger rebellions are not in the interests of incumbents. Hence, allocations are chosen so that rebellions do not occur in equilibrium, but the threat of rebellions places limits on the set of allocations that would survive the struggle for power. Effectively, incumbents face a myriad of 'no-rebellion constraints' from different subsets of the population, both before and after investment decisions have been made. The proposition below demonstrates how these no-rebellion constraints imply linkages in equilibrium between the fraction s of investment opportunities undertaken, the amount of power sharing p, and the payoffs received by incumbents and workers.

Proposition 7 *The equilibrium allocation must have the following properties:*

(i) The allocation of consumption across individuals is consistent with perfectly competitive markets, that is, equations (2.2), (2.3), (2.4), (2.7), (2.8), and (2.9) hold for some levels of disposable income Y_{p}, Y_{k}, and Y_{w} for incumbents, capitalists, and workers.

(ii) Net exports must maximize real GDP subject to the international budget constraint (2.5).

(iii) The incentive compatibility constraint (4.2) for investors is binding: $C_{\mathrm{k}} = (1+\theta)C_{\mathrm{w}}$.

(iv) Any rebellion at the post-investment stage would lead to full expropriation of investors' capital. This entails a positive relationship $s = \lambda(p)$ between power sharing p and the fraction s of investment opportunities that are undertaken:

$$\lambda(p) = \begin{cases} \frac{\delta(p-p^{\dagger})}{\mu\theta}\left(1+\frac{\delta p}{\delta p+(1-\delta)p^{\dagger}}\right) & \text{if } \delta < 1/2 \\ \frac{\delta(p-p^{\dagger})}{\mu\theta}\left(1+\frac{p}{2\delta p+p^{\dagger}}\right) & \text{if } \delta \geq 1/2 \end{cases}, \tag{4.7}$$

where $\lambda(p^{\dagger}) = 0$ and $\lambda(\bar{p}) = 1$ for $p^{\dagger}$ and $\bar{p}$ with $p^{\dagger} < \bar{p} \leq 1/2$:

$$p^{\dagger} = \frac{1}{2+\delta}, \quad \text{and} \quad \bar{p} = \begin{cases} \frac{\delta(\delta+(2+\delta)\mu\theta)+\sqrt{(\delta(\delta+(2+\delta)\mu\theta))^2+\delta(1+2\delta)(\delta+(2+\delta)\mu\theta)}}{\delta(1+2\delta)(2+\delta)} & \text{if } \delta < 1/2 \\ \frac{3\delta-1+(2+\delta)\mu\theta+\sqrt{(3\delta-1+(2+\delta)\mu\theta)^2+8(1-\delta)(\delta+(2+\delta)\mu\theta)}}{4\delta(2+\delta)} & \text{if } \delta \geq 1/2 \end{cases}.$$

(v) The consumption levels of workers and incumbents as fractions of real GDP are given by $C_{\mathrm{w}} = \psi_{\mathrm{w}}(p)C$ and $C_{\mathrm{p}} = \psi_{\mathrm{p}}(p)C$, where the shares $\psi_{\mathrm{w}}(p)$ and $\psi_{\mathrm{p}}(p)$ are negatively related to

[22] The equilibrium concept is related to the notion of blocking in coalitions as studied in Ray (2007), but there are two important differences here. First, an allocation must be optimal for incumbents. Second, changing the allocation requires costly rebellion effort.

power sharing p:

$$\psi_{\mathrm{w}}(p) = \frac{1}{2(\delta p + p^{\dagger})}, \quad \textit{and} \quad \psi_{\mathrm{p}}(p) = \begin{cases} \frac{1}{2(\delta p + (1-\delta)p^{\dagger})} & \textit{if } \delta < 1/2 \\ \frac{1+2\delta}{2(2\delta p + p^{\dagger})} & \textit{if } \delta \geq 1/2 \end{cases}. \qquad [4.8]$$

(vi) *The quality of governance s must maximize incumbents' consumption C_{p} subject to (4.7) and (4.8). The first derivative of C_{p} with respect to s can be written in terms of a strictly positive and strictly decreasing function $\chi(p)$ as follows:*

$$\frac{\partial C_{\mathrm{p}}}{\partial s} = \frac{\mu \psi_{\mathrm{p}}(p)}{\tilde{\pi}^{\alpha}} \left(\tilde{\pi} - (1 + \chi(p))\theta Y_{\mathrm{w}} \right). \qquad [4.9]$$

Proof See appendix A.7. ■

The first two results are that the equilibrium allocation of resources can be implemented through free trade both domestically and internationally, starting from some distribution of national income (after taxes and transfers). Domestic free trade was implicitly assumed in the model of section 2 but here arises as a result. International free trade was shown earlier to be chosen by incumbents subject to the ad hoc political friction (2.11), and this is confirmed here starting from more basic political frictions. The reason the equilibrium features free trade is that no-rebellion constraints effectively place lower bounds on individuals' payoffs, so Pareto-improving reallocations of goods are in the interests of incumbents. These are precisely the exchanges brought about by free markets, and crucially, these exchanges do not affect any aspect of the power struggle.

Production of the investment good requires offering higher consumption to investors to compensate for the cost of their effort. The third result of Proposition 7 is that the incentive constraint (4.2) binds for investors — intuitively, incumbents have no interest in giving investors any more than the minimum reward necessary. However, once the effort cost of investment is sunk at the post-investment stage, any rewards to investors (property rights) specified by the allocation raise incentives for rebellion by workers and incumbents.[23] A successful rebellion at the post-investment stage would lead to a new allocation in which investors' capital is expropriated. An allocation specifying incentives for investors must therefore take account of binding no-rebellion constraints at the post-investment stage, while no-rebellion constraints at the pre-investment stage are slack.

The equilibrium allocation established before investment occurs is the best possible solution from the point of view of incumbents to the problem of ensuring the incentives needed for investment are able to survive the power struggle. The result (the fourth statement in Proposition 7) is a relationship between power sharing p and the fraction s of investment opportunities undertaken. The requirements of equilibrium will determine the value of s as explained later, but conditional on a value of s, power sharing p must satisfy $s = \lambda(p)$ as given in (4.7). For high values of the power parameter ($\delta \geq 1/2$), the only binding threat of rebellion is from workers at the post-investment stage. For low values of the power parameter ($\delta < 1/2$), incumbents also face a binding threat of a coup d'état from within their own ranks at the post-investment stage. This means the expressions

[23]Investors themselves never want to rebel as long as there are not too many of them, which is guaranteed by the first parameter restriction in (4.5).

for $\lambda(p)$ (and thus all subsequent results) are different in these two cases, but the qualitative features are the same and no important result depends on the value of the power parameter δ.

The function $\lambda(p)$ is strictly increasing, so there is a positive relationship between power sharing and investment. To understand this, observe that there is a level of power sharing $p^\dagger$ such that no investment will occur in equilibrium ($s = 0$). This level of power sharing coincides with what would be the equilibrium level of power sharing after a rebellion at the post-investment stage. If $p = p^\dagger$, a rebellion at the post-investment stage by all incumbents requires no effort from them to succeed because everyone who holds power takes part in the rebellion (power sharing after the rebellion will be $p^\dagger$, so there is space for all current incumbents in the rebel faction). This means that incumbents essentially have full discretion to change the allocation of consumption after investment decisions have been made. As the promise of an incentive-compatible payoff to investors raises the gains from rebellion for both workers and incumbents, an allocation with $s > 0$ can only survive if there is also an increase in the cost of rebellion for both groups.

It turns out that an allocation with $s > 0$ can survive threats of rebellion if and only if there is an increase in power sharing p to a level above $p^\dagger$. A larger incumbent group means that outsiders must put in more effort to launch a successful rebellion. Crucially, $p > p^\dagger$ also means that insiders face a cost of launching a rebellion because other incumbents have an incentive to resist it — knowing that they would lose power as the equilibrium size of the post-rebellion incumbent group is only $p^\dagger$. Power sharing therefore enables commitment to rules that would otherwise be time inconsistent.

Since investment opportunities are taken if and only if investors expect their property rights to be protected, the variable s can be interpreted as a measure of good government in the sense that term was used in section 2. Here, power sharing is the means through which the quality of governance is improved. With a low level of power sharing $p = p^\dagger$, the country is despotic in that no investor has secure property rights, and zero investment is the result ($s = 0$). As power sharing increases, it becomes possible for more investors to receive rewards for past efforts that survive threats of rebellion, and the economy moves closer to the rule of law. At a high level of power sharing $\bar{p}$, the rule of law is attained and everyone has secure property rights, which means all investment opportunities are taken ($s = 1$).[24]

The fifth result of Proposition 7 uses the no-rebellion constraints faced by incumbents to obtain the shares $\psi_{\mathrm{w}}(p)$ and $\psi_{\mathrm{p}}(p)$ of GDP that must be given to workers and incumbents. These shares are decreasing in p because an increase in power sharing raises the amount of effort required for a successful rebellion, and thus allows for an equilibrium where workers and incumbents have more to gain from rebellion. Smaller shares for workers and incumbents free up resources to expand the group of individuals who are offered incentives to invest.[25]

[24] The second parameter restriction in (4.5) requires that the effort cost θ is not too large, which ensures the amount of power sharing $\bar{p}$ needed for the rule of law does not exceed the maximum possible amount $1/2$.

[25] The equilibrium allocation of resources can be implemented by the following taxes T_{w} and T_{k} on workers and capitalists (specified in terms of the endowment good):

$$T_{\mathrm{w}} = (1 - \psi_{\mathrm{w}}(p))\mathrm{q} - \mu s \psi_{\mathrm{w}}(p)\tilde{\pi}, \quad \text{and} \quad T_{\mathrm{k}} = (1 - (1+\theta)\psi_{\mathrm{w}}(p))\mathrm{q} + (1 - \mu(1+\theta)s\psi_{\mathrm{w}}(p))\tilde{\pi},$$

and an additional 100% tax ($T_{\mathrm{I}} = \tilde{\pi}$) on any individual who took an investment opportunity but who does not

The sixth result of the proposition highlights the source of inefficiency in the economy. Efficiency in production requires that output of the investment good is increased if its market value $\tilde{\pi}$ in terms of the endowments good is larger than the marginal effort cost θY_{w} of producing it. However, when deciding whether to increase investment by improving the quality of governance, incumbents factor in an additional cost represented by the positive term $\chi(p)$ in (4.9). This political wedge reflects the distributional consequences of the power sharing needed to support credible protection of investors' property rights, which is not a social cost. Note that the distortion does not arise from restrictions on feasible tax instruments (incumbents capture all surplus from investment), but from the maximizing behavior of incumbents subject to the threat of rebellion. Production of the investment good affects incentives for rebellions because of the need for property rights. If there were no such interactions, incumbents would choose *laissez faire*, as they do for exchange of goods domestically and internationally.

While the political wedge $\chi(p)$ is always positive, it declines as power sharing increases. Therefore, the political distortion is smaller at the margin in countries where power is shared more broadly. This feature will be crucial in understanding the results that follow.

4.4 The decreasing marginal cost of good government

The key theoretical result of the model is that the function $\lambda(p)$ is not only increasing, but is also strictly convex: a given increase in power sharing is associated with a larger rise in investment when power is already widely shared.[26] In other words, there are increasing returns to improving the quality of governance by sharing power more broadly. It follows that, as assumed in the model of section 2, the marginal cost of good government is decreasing.

To understand the source of the increasing returns to power sharing, note that with p incumbents, μs capitalists, and $1-p-\mu s$ workers, the resource constraint (2.8) can be written as $pC_{\mathrm{p}}+\mu s C_{\mathrm{k}}+(1-p-\mu s)C_{\mathrm{w}}=C$. Using the binding incentive constraint (4.2) and the expression for workers' consumption in (4.8), the resource constraint equation can be rearranged as follows:

$$s=\frac{1}{\mu\theta}\left(\frac{1}{\psi_{\mathrm{w}}(p)}-1-p\beta(p)\right), \qquad [4.10]$$

where the function $\psi_{\mathrm{w}}(p)$ gives workers' consumption relative to the average (see 4.8) and the variable $\beta=(C_{\mathrm{p}}-C_{\mathrm{w}})/C_{\mathrm{w}}$ denotes the rents received by an incumbent as a fraction of a worker's income. Using the expressions for incumbent and worker consumption in (4.8), incumbent rents can be written as a function of power sharing: $\beta(p)=\psi_{\mathrm{p}}(p)/\psi_{\mathrm{w}}(p)-1$. Conditional on power sharing p, the functions $\psi_{\mathrm{w}}(p)$ and $\beta(p)$ give the values of workers' relative consumption and incumbents' rents that are consistent with maximization by incumbents subject to the no-rebellion constraints.

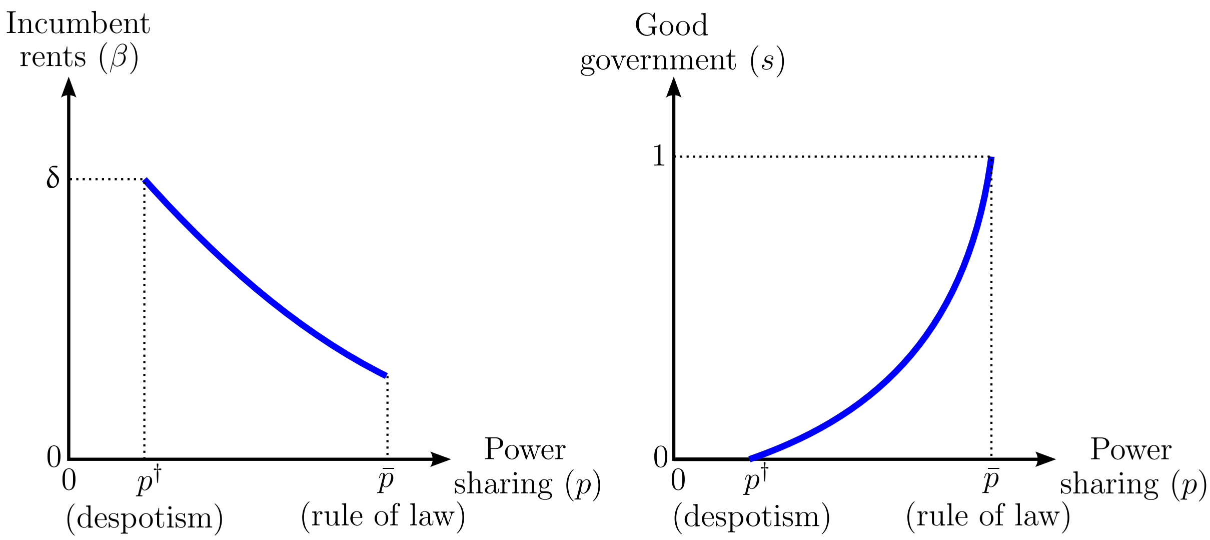
belong to the set $\mathcal{D}$ of individuals with secure property rights (which could be interpreted as those with a licence to produce the investment good). All tax revenue is distributed equally among incumbents. Individuals can then spend their post-tax incomes in the markets for consumption goods. Although capitalists have secure property rights that provide incentives to invest, their output is still taxed so that they receive only the minimum amount of consumption required for incentive compatibility.

[26] The positive relation between p and s arises in a related setting in Guimaraes and Sheedy (2015); the convexity result is novel.

This means that the equation in (4.10) defines the function $\lambda(p)$ from (4.7).

Owing to the nature of the rebellion mechanism and the effort cost of rebellion in (4.1), the term $1/\psi_{\mathrm{w}}(p)$ turns out to be linear in p (see equation 4.8), hence the convexity of $\lambda(p)$ is due entirely to the behaviour of incumbents' total rents $p\beta(p)$. Inspection of the expressions for $\psi_{\mathrm{w}}(p)$ and $\psi_{\mathrm{p}}(p)$ in (4.8) reveals that an increase in p leads to a decrease in workers' relative consumption $\psi_{\mathrm{w}}(p)$, but an even larger reduction in incumbents' relative consumption $\psi_{\mathrm{p}}(p)$, so rents $\beta(p)$ are decreasing in p. Consequently, total rents $p\beta(p)$ rise less than proportionally with p, which implies $\lambda(p)$ is strictly convex. The increasing returns to sharing power are thus explained by rents β being diluted as power sharing p increases, as depicted in Figure 5 below.

Figure 5: *Power sharing, incumbent rents, and good government*



The explanation for declining rents is that an additional incumbent is less important in defending an allocation when there are already many individuals in power. For example, if there is relatively little power sharing, a coup d'état could succeed with relatively little rebellion effort, which means that those in power would find it very worthwhile to recruit additional incumbents. Alternatively, when the power base is narrow, the marginal gain in extracting rents from workers of an additional incumbent is large. Both of these arguments point to a negative relationship between rents and power sharing, but which one is dominant turns out to depend on δ. The former is dominant when δ is low and the latter when δ is high.

In the case $\delta < 1/2$, the negative relationship between p and β comes from the relative effects of an increase in p on the no-rebellion constraints for workers and incumbents. An increase in power sharing p makes rebellions more costly for both incumbents and workers, and hence decreases both ψ_{p} and ψ_{w}. However, the effect on rebellions launched by incumbents is relatively more important, and especially so when p is small. When p is closer to $p^{\dagger}$, very few individuals would defend an allocation against a coup d'état, and thus incumbents would pose a substantially greater threat to the allocation. It follows that the ratio $\psi_{\mathrm{p}}/\psi_{\mathrm{w}}$ would need to be very large. As p increases, the threat posed by incumbents would become relatively less important, leading to smaller rents β.

In the case $\delta \geq 1/2$, incumbents are willing to accept more power sharing than required to avoid a coup d'état. In other words, the no-rebellion constraint for incumbents is not binding. Power sharing is thus extended up to the point where its marginal benefit in squeezing non-incumbent incomes (without triggering rebellions) is equal to its marginal cost. From the point of view of incumbents, the marginal cost of an additional member of the incumbent group is the difference between incumbent and worker payoffs, which is measured by rents β. Since the marginal extraction benefit from power sharing is decreasing in p, rents and power sharing must be negatively related.

Putting together the results of Proposition 7, there is a positive relationship between good government and power sharing, and a negative relationship between power sharing and the share of total income received by each incumbent. These two results imply that good government and the income share of incumbents are negatively related as assumed in the simple political friction (2.11). Crucially, the convexity of the relationship between power sharing and good government implies the marginal cost of good government is diminishing (2.14 holds). The proposition below confirms that the assumptions of the simple model from section 2 are results of the full model developed here.

Proposition 8 *The quality of governance s in the equilibrium allocation is equivalent to what is obtained in the simpler model of section 2, where:*

(i) The political friction (2.11) holds for a decreasing function $\phi(s)$ given by:

$$\phi(s) = \psi_{\mathrm{p}}(\lambda^{-1}(s)), \tag{4.11}$$

and the implied marginal cost of good government $\gamma(s)$ from (2.13) and the total cost Φ of the rule of law from (2.25) are:

$$\gamma(s) = \frac{2\delta \min\left\{1, \frac{2}{1+2\delta}\right\}}{\lambda'(\lambda^{-1}(s))}, \quad \text{and} \quad \Phi \equiv \phi(0)\int_0^1 \gamma(s)\mathrm{d}s = \delta(2+\delta)\min\left\{1, \frac{2}{1+2\delta}\right\}(\bar{p} - p^{\dagger}). \tag{4.12}$$

(ii) The marginal cost of good government $\gamma(s)$ is decreasing (2.14 holds) because $\lambda(p)$ is a strictly convex function. The curvature of $\lambda(p)$ depends only on the behaviour of incumbent rents $\beta(p) = (C_{\mathrm{p}} - C_{\mathrm{w}})/C_{\mathrm{w}}$, which are strictly positive, but strictly decreasing in power sharing p.

(iii) The political friction is relevant (2.16 holds) when:

$$\alpha < \bar{\alpha}, \quad \text{where } \bar{\alpha} = \begin{cases} \left(1 + \frac{(1-\delta)^2 + \delta(2+\delta)(3-\delta)\bar{p}}{\mu\theta(2+\delta)(1-\delta+\delta(2+\delta)\bar{p})}\right)^{-1} & \text{if } \delta < 1/2 \\ \left(1 + \frac{\delta + (2+\delta)(1+2\delta+2\delta^2)\bar{p}}{\mu\theta(2+\delta)(1+2\delta(2+\delta)\bar{p})}\right)^{-1} & \text{if } \delta \geq 1/2 \end{cases}.$$

(iv) The results of Proposition 1 hold in the case of autarky (the incumbent payoff is a quasi-concave function of s).

Proof See appendix A.8. ∎

4.5 Features of the world equilibrium

The additional structure provided by the full model of section 4 means that it is possible to make some statements about efficiency and welfare in the world equilibrium (with ex-ante identical countries) analysed earlier in section 2.5.

Proposition 9 *The world equilibrium has the following features:*

(i) Workers and investors are strictly better off in rule-of-law economies than in despotic regimes.

(ii) Workers and investors in rule-of-law economies are better off than under autarky. Opening up to trade is a Pareto improvement for a country that establishes the rule of law.

(iii) Rule-of-law economies are Pareto efficient; despotic regimes are Pareto inefficient. Under autarky, all economies would be Pareto inefficient.

(iv) Average power sharing across open economies $p^* = (1-\tilde{\omega})p^\dagger + \tilde{\omega}\bar{p}$ is strictly lower than the common level of power sharing $\hat{p}$ under autarky, with p^* and $\hat{p}$ equal to:

$$p^* = \frac{1}{2+\delta}\left(1 + \frac{\alpha}{(1-\alpha)\delta \min\left\{1, \frac{2}{1+2\delta}\right\}}\right), \quad \textit{and}$$

$$\hat{p} = \begin{cases} \frac{3\delta-1+4(1-\delta)\alpha+\sqrt{(3\delta-1+4(1-\delta)\alpha)^2+8(1-\delta)(\delta+(1-2\delta)\alpha)(1-\alpha)}}{4\delta(2+\delta)(1-\alpha)} & \textit{if } \delta < 1/2 \\ \frac{2\delta^2+(1+2\delta)\alpha+\sqrt{(2\delta^2+(1+2\delta)\alpha)^2+4\delta^2(1+2\delta)(1-\alpha)}}{2\delta(1+2\delta)(2+\delta)(1-\alpha)} & \textit{if } \delta \geq 1/2 \end{cases}.$$

Proof See appendix A.9. ∎

Workers in rule-of-law economies are strictly better off than workers under despotism because incumbents receive the same payoff in both types of economies and incumbents' rents as a proportion of workers' consumption decline (the ratio $C_\mathrm{p}/C_\mathrm{w}$ is lower) with higher s and p. Since investors obtain no surplus, in payoff terms they receive exactly the same amount as workers.[27] The additional output in rule-of-law economies pays for higher consumption for investors (to compensate them for their efforts), higher consumption for incumbents and workers, and a larger number of incumbents (who receive more than workers).

Workers and investors in rule-of-law economies are also strictly better off than they would be in autarky because $C_\mathrm{p}/C_\mathrm{w}$ is decreasing in s and incumbents gain from trade. Together with there being more incumbents in rule-of-law economies (who receive more than workers), it follows that opening up to trade is Pareto improving for rule-of-law economies. The decline in $C_\mathrm{p}/C_\mathrm{w}$ actually implies workers and investors in economies with the rule of law capture a relatively greater share of the gains from trade than incumbents.

An economy with the rule of law is Pareto efficient while despotic regimes have inefficiently low production.[28] It follows that a despotic regime has a positive pecuniary externality on the rest of the world by providing cheap endowment goods. This has implications for efficiency in other countries because it makes the rule of law more attractive. Conversely, economies with the rule

[27] In a related setting, Guimaraes and Sheedy (2015) assume that investors have private information about their effort cost, which means it is not possible to extract all their surplus from investment through taxes. As a result, investors are better off than workers.

[28] Since workers are better off under the rule of law, the effort cost of producing the investment good at the margin is lower in a despotic country than in a rule-of-law economy. The absence of investment in despotic economies is therefore not due to the high cost of investment — actually, incumbents would be able to extract more from a marginal investor in a despotic economy than in a rule-of-law economy if they were only able to commit not to take everything ex post.

of law impose a negative externality on other countries by increasing incentives for incumbents to choose despotism.

The last statement of Proposition 9 is that there are fewer individuals in power on average in a world of open economies compared to autarky. While increasing returns to power sharing would allow more to be produced through specialization even if there were the same number of people in power across the world, it is both possible and in the interests of incumbents everywhere to have fewer people in power worldwide.

5 Trade and political divergence

The 19th century witnessed a large boom in world trade and a sharp increase in the spread of industrialization beyond the pioneer countries. Among the reasons for the dramatic rise in trade volumes, Williamson (2011) points to the world transport revolution: the substantial decline in transportation costs due to technological progress (for example, railroads and steamships).[29] The transport revolution was part of a broader event, the industrial revolution, which led to a huge increase in productivity.

In the context of the economic model of section 2, the new industrial technologies can be seen as the ones used to produce the investment good. The relative importance of the investment good to overall economic activity is measured by the parameter α. Until the 18th century, α can be seen as very small because there was not much that could be efficiently produced using the prevailing industrial technologies. In the 19th century, α would go from close to zero to a sizeable number. Naturally, the model predicts this would trigger an increase in industrialization at the world level (an increase in output of the investment good, as represented by the variable s). Furthermore, the reduction in the cost of international trade meant that economies became closer to the limiting case of frictionless international trade than the limiting case of autarky.

The full model of section 4 yields two non-trivial implications: (i) industrialization (an increase in s) would occur in some countries, but not in others; and (ii) economic specialization would be accompanied by political specialization, as industrializing countries would experience an increase in power sharing (represented by the variable p).

The data confirm the first implication. Williamson (2011) shows that industrialization was restricted to a few (mainly European) countries. In many countries, there was no industrialization at all — and even some deindustrialization — for a substantial part of the 19th century. The index of per-capita levels of industrialization from Bairoch (1991), as used in the analysis of Williamson (2011), rises dramatically in the 'European core' between 1800 and 1913, but is almost flat in Brazil and Mexico and is actually decreasing in China and India.

The second implication — that industrialization and power sharing go hand in hand — is a key message of the model. This is tested by studying the evolution of power sharing in the different

[29]It also became safer to ship goods overseas owing to a reduction in piracy and wars. In the political sphere, an increasing number of countries moved towards pro-trade policies. All these factors helped to foster trade in the early 19th century.

groups of countries considered by Bairoch (1991) and Williamson (2011). The variable 'executive constraints' from the Polity IV database is used as a good proxy for power sharing because constraints on those in power can only be imposed by other people who also hold power.[30] The minimum score of 1 indicates "unlimited authority: no regular limitations on the executive's actions", while the maximum score of 7 indicates "executive parity or subordination".[31]

The evolution of scores for executive constraints from 1800 to 1913 is studied for Britain (the cradle of the industrial revolution) and three groups of countries: the 'European core', Latin America, and Asia. The 'European core' comprises the countries considered by Bairoch (1991) (except for Britain) and also including the Netherlands, an industrial pioneer for which Polity IV data are available. For Latin America, all countries are included for which there are data from 1830 or earlier.[32] The group of Asian countries comprises Afghanistan, China, Iran, Thailand, and Turkey. Figure 6 plots the time series of executive constraints in each group (taking simple averages within groups). A more detailed description of the data is given in Appendix B.

The pre-industrial world featured very little power sharing. In 1800, most countries had the lowest possible executive constraints score (1). Britain, where the process of industrialization was well under way, had the maximum possible score (7).[33] Around 1830, the European core and Latin America still looked very similar, with executive constraints scores of around 2.25 and 2 respectively.

From 1830 onwards, Figure 6 shows a striking pattern of political divergence. The average executive constraints score is always above 4 in the European core from 1861 and always above 5 from 1876. In sharp contrast, the score in Latin America only goes above 2 in 1867 and only reaches 3 in 1911. In Asia, the executive constraints score is 1 for almost the whole century. By the late 19$^{\text{th}}$ century, a considerable degree of power sharing is the norm in the European core, while there are very few constraints on the executive in Latin America and especially so in Asia.[34]

The case of Japan is an interesting one that well illustrates the link between economic and political progress. Unlike the other countries of Asia, Japan industrializes in the 19$^{\text{th}}$ century (Williamson, 2011), and for this reason it was not included in the group of Asian countries above. Politically, Japan goes from unlimited authority (executive constraints score 1) up until 1857 to executive parity or subordination (score 7) in 1868 with the Meiji Restoration (there is missing data in between). Consistent with the theory proposed in this paper, historians consider that the rapid

[30]According to the Polity IV Dataset Users' Manual (Marshall, Gurr and Jaggers, 2016, p. 24), this variable "refers to the extent of institutionalized constraints on the decision making powers of chief executives, whether individuals or collectivities. Such limitations may be imposed by any accountability groups. In Western democracies these are usually legislatures. Other kinds of accountability groups are the ruling party in a one-party state; councils of nobles or powerful advisors in monarchies; the military in coup-prone polities; and in many states a strong, independent judiciary. The concern is therefore with the checks and balances between the various parts of the decision-making process."

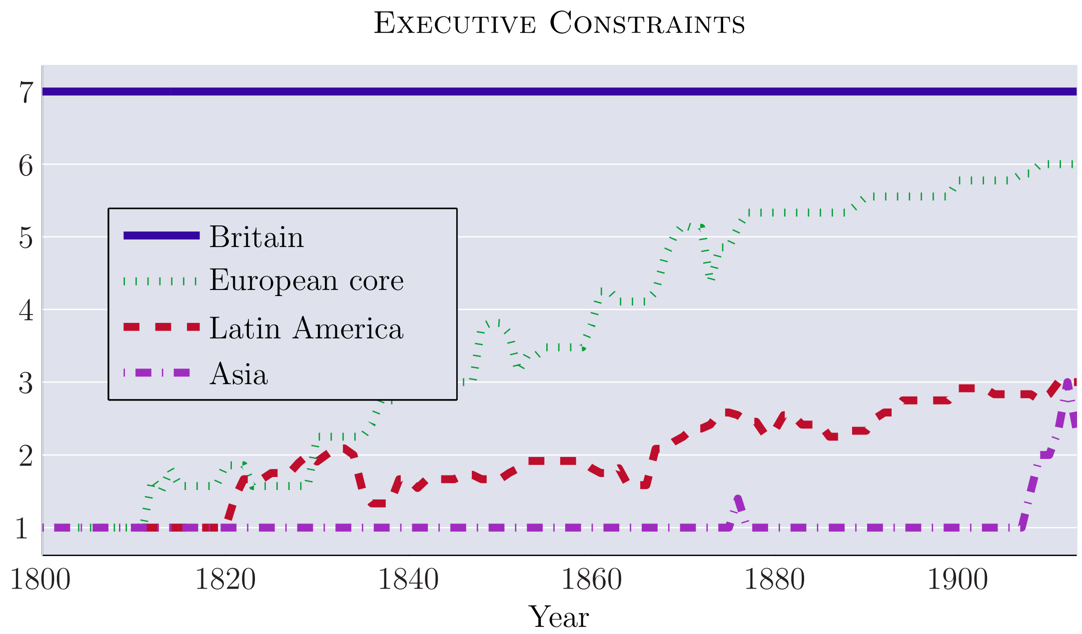
[31]A score of 3 indicates some real but limited restraints on the executive, and a score of 5 indicates that the executive is subject to substantial constraints by accountability groups.

[32]Polity IV data are available only for independent countries. Most Latin American countries became independent between 1818 and 1830.

[33]In 1800, the only countries with an executive constraints score greater than the minimum (1) were Britain (score of 7), the US (7), Korea (7), and Wuerttemburg (3).

[34]Moreover, within the European core, the transition from unlimited authority to executive parity occurs earlier in the industrial pioneers. For example, the Netherlands reaches a score of 6 in 1849 and Belgium reaches 7 in 1853. In contrast, Italy makes the transition from 3 to 5 only in 1900 and Austria never goes beyond 3.

Figure 6: *The evolution of executive constraints, 1800–1913*



Notes: Simple averages of 'executive constraints' for Britain (since 1800); the 'European core': Austria (since 1800), Belgium (1830), France (1800), Germany (1800), Italy (1815), Netherlands (1815), Spain (1800), Sweden (1800), and Switzerland (1848); Latin America, comprising Argentina (since 1825), Bolivia (1825), Brazil (1824), Chile (1818), Colombia (1821), Ecuador (1830), Haiti (1820), Mexico (1822), Paraguay (1811), Peru(1821), Uruguay (1830), and Venezuela (1830); and Asia, comprising Afghanistan (1800), China (1800), Iran (1800), Thailand (1800), and Turkey (1800).
Source: Polity IV Project, Center for Systemic Peace (http://www.systemicpeace.org/inscrdata.html).

industrialization of Japan began in 1868.

This pattern of both economic and political divergence also holds for some countries that were relatively prosperous in the 18th century but which lost ground with the first wave of globalization. Russia is a case in point.[35] During the 18th and early 19th centuries, Russia was a powerful European country that went through modernizing reforms both in the reign of Peter the Great and in the age of the Russian Enlightenment.[36] However, in the early 19th century, Russia experienced strong positive terms-of-trade shocks and subsequently fell behind. By the end of the century, Russia was one of the poorest countries in Europe (Nafziger, 2008). In contrast to the expansion of power sharing in other European countries at the time, and consistent with the theory of political specialization proposed here, Russia went through the whole 19th century without any kind of elected parliament and the lowest possible score for executive constraints.[37]

[35] Nafziger (2008) claims that "understanding what inhibited Russian economic development in the nineteenth century is an important task for economic historians."

[36] The expansion of the Russian empire was a sign of its power and development at the time. In the early 19th century, the Russians colonized Alaska and even founded settlements in California. Among other notable Russian sea exploration voyages, in 1820, a Russian expedition discovered the continent of Antarctica.

[37] Nicholas I ruled between 1825 and 1855, the time when Russian exports were becoming more expensive. He resisted any kind of power sharing and concentrated his existing powers even more, crushing demonstrations demanding power sharing and abolishing several areas of local autonomy (Bessarabia, Poland, and the Jewish Qahal).

To sum up, during the 19th century the rule of law was becoming established in the more economically advanced European countries as the absolute power of monarchs was eroded by increasingly powerful parliaments. However, countries in the periphery were experiencing none of this political change, despite a large increase in trade, which presumably led to greater exposure to foreign ideas.

6 Concluding remarks

For social scientists grappling with the welter of autocratic regimes around the world, one particular fact is noteworthy: the stubborn resistance to adopting the rule of law in spite of its proven success elsewhere. An important policy question is what can be done to bring about positive political change. The literature in political science has focused on country-specific factors that are seen as barriers to change such as culture and history. This paper highlights the importance of thinking about the problem in general equilibrium at the world level.

The adoption of the rule of law increases output of goods that require strong protection of property rights and thus reduces their relative price. This increases incentives for those in power in other countries to choose autocratic government. Consequently, the proportion of countries in the world with the rule of law is limited by the demand for goods that depend on the rule of law for their production. The analysis suggests those living under autocracy would be best helped by subsidies to goods that are substitutes for their own exports.

This paper also has implications for the policy debate about the merits of international trade. In spite of its well-known economic gains, international trade has its critics. Much of the opposition to free trade comes from a sense that some economies end up specializing in the wrong kinds of goods — primary goods — which is detrimental to development. The theory in this paper is consistent with a negative correlation between specializing in primary goods and economic performance. However, the policy prescription is not to limit trade. This is because the distortion is not found in trade itself, but in the distributional consequences of the power sharing needed for good government.

References

Acemoglu, D., Johnson, S. and Robinson, J. (2005), "The rise of Europe: Atlantic trade, institutional change, and economic growth", *American Economic Review*, **95(3)**:546–579 (June). 6

Acemoglu, D. and Robinson, J. A. (2000), "Why did the West extend the franchise? Democracy, inequality, and growth in historical perspective", *Quarterly Journal of Economics*, **115(4)**:1167–1199 (November). 6

Acemoglu, D., Robinson, J. A. and Verdier, T. (2015), "Asymmetric growth and institutions in an interdependent world", *Journal of Political Economy*, forthcoming. 6

Acemoglu, D., Verdier, T. and Robinson, J. A. (2004), "Kleptocracy and divide-and-rule: A model of personal rule", *Journal of the European Economic Association*, **2(2–3)**:162–192 (April–May). 7

Anderson, J. E. and Marcouiller, D. (2002), "Insecurity and the pattern of trade: An empirical investigation", *Review of Economics and Statistics*, **84(2)**:342–352 (May). 6

Araujo, L., Mion, G. and Ornelas, E. (2016), "Institutions and export dynamics", *Journal of International Economics*, **98**:2–20 (January). 6

Bai, J. H. and Lagunoff, R. (2011), "On the Faustian dynamics of policy and political power", *Review of Economic Studies*, **78(1)**:17–48. 6

Bairoch, P. (1991), "How and not why? Economic inequalities between 1800 and 1913: Some background figures", in J. Batou (ed.), *Between Development and Underdevelopment: The Precocious Attempts at Industrialization of the Periphery 1800–1870*, Librairie Droz, chap. 1, pp. 1–42. 33, 34

Cardoso, F. H. and Faletto, E. (1979), *Dependency and Development in Latin America*, University of California Press. 7

Caselli, F. and Cunningham, T. (2009), "Leader behaviour and the natural resource curse", *Oxford Economic Papers*, **61(4)**:628–650 (October). 7

Caselli, F. and Tesei, A. (2016), "Resource windfalls, political regimes, and political stability", *Review of Economics and Statistics*, **98(3)**:573–590 (July). 7

Do, Q.-T. and Levchenko, A. A. (2009), "Trade, inequality, and the political economy of institutions", *Journal of Economic Theory*, **144(4)**:1489–1520 (July). 6

Fukuyama, F. (1992), *The End of History and the Last Man*, Free Press. 7

——— (2011), *The Origins of Political Order: From Prehuman Times to the French Revolution*, Farrar, Straus and Giroux. 7

Gandhi, J. and Przeworski, A. (2007), "Authoritarian institutions and the survival of autocrats", *Comparative Political Studies*, **40(11)**:1279–1301 (November). 7

Goorha, P. (2007), "Polarized futures", *Kyklos*, **60(1)**:65–75 (February). 7

Greif, A. (1993), "Contract enforceability and economic institutions in early trade: The Maghribi traders' coalition", *American Economic Review*, **83(3)**:525–548 (June). 6

——— (2006), *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*, Cambridge University Press. 6

Greif, A., Milgrom, P. and Weingast, B. R. (1994), "Coordination, commitment, and enforcement: The case of the merchant guild", *Journal of Political Economy*, **102(4)**:745–776 (August). 6

Grossman, G. M. and Helpman, E. (1994), "Protection for sale", *American Economic Review*, **84(4)**:833–850 (September). 6

Guimaraes, B. and Sheedy, K. D. (2015), "Guarding the guardians", *Economic Journal*, forthcoming. 5, 6, 22, 23, 25, 29, 32

Huntington, S. (1993), *The Third Wave: Democratization in the Late Twentieth Century*, University of Oklahoma Press. 7

Jack, W. and Lagunoff, R. (2006), "Dynamic enfranchisement", *Journal of Public Economics*, **90(4–5)**:551–572 (May). 6

Krugman, P. R. (1979), "Increasing returns, monopolistic competition, and international trade", *Journal of International Economics*, **9(4)**:469–479 (November). 6, 17

——— (1980), "Scale economies, product differentiation, and the pattern of trade", *American Economic Review*, **70(5)**:950–959 (December). 6, 17

——— (1987), "The narrow moving band, the Dutch disease, and the competitive consequences of Mrs. Thatcher: Notes on trade in the presence of dynamic scale economies", *Journal of Development Economics*, **27(1–2)**:41–55 (October). 7

Levchenko, A. A. (2007), "Institutional quality and international trade", *Review of Economic Studies*, **74(3)**:791–819 (July). 6

——— (2013), "International trade and institutional change", *Journal of Law, Economics, & Organization*, **29(5)**:1145–1181 (October). 6

Marshall, M. G., Gurr, T. R. and Jaggers, K. (2016), *Polity IV Project Dataset Users' Manual*, Center for Systemic Peace. 34

Mehlum, H., Moene, K. and Torvik, R. (2006), "Institutions and the resource curse", *Economic Journal*, **116(508)**:1–20 (January). 7

Melitz, M. J. (2005), "When and how should infant industries be protected?", *Journal of International Economics*, **66(1)**:177–196 (May). 7

Milgrom, P. R., North, D. C. and Weingast, B. R. (1990), "The role of institutions in the revival of trade: The Law Merchant, private judges, and the Champagne fairs", *Economics & Politics*, **2(1)**:1–23 (March). 6

Nafziger, S. (2008), "Communal institutions, resource allocation, and Russian economic development: 1861–1905", *Journal of Economic History*, **68(2)**:570–575 (June). 35

Nunn, N. (2007), "Relationship-specificity, incomplete contracts, and the pattern of trade", *Quarterly Journal of Economics*, **122(2)**:569–600 (May). 6

Nunn, N. and Trefler, D. (2014), "Domestic institutions as a source of comparative advantage", in E. Helpman, K. Rogoff and G. Gopinath (eds.), *Handbook of International Economics*, vol. 4, Elsevier. 6, 17

Puga, D. and Trefler, D. (2014), "International trade and institutional change: Medieval Venice's response to globalization", *Quarterly Journal of Economics*, **129(2)**:753–821 (May). 6

Ray, D. (2007), *A Game-Theoretic Perspective on Coalition Formation*, Oxford University Press. 26

Robinson, J. A., Torvik, R. and Verdier, T. (2006), "Political foundations of the resource curse", *Journal of Development Economics*, **79(2)**:447–468 (April). 7

Rodrik, D. (1996), "Coordination failures and government policy: A model with applications to East Asia and Eastern Europe", *Journal of International Economics*, **40(1–2)**:1–22 (February). 7

Sachs, J. D. and Warner, A. M. (2001), "The curse of natural resources", *European Economic Review*, **45(4–6)**:827–838 (May). 7

Van der Ploeg, F. (2011), "Natural resources: Curse or blessing?", *Journal of Economic Literature*, **49(2)**:366–420 (June). 7

Williamson, J. G. (2011), *Trade and Poverty: When the Third World Fell Behind*, MIT Press. 5, 33, 34

A Technical appendix

A.1 Proof of Proposition 1

(i) The derivative of total consumption C from (2.17) with respect to s is given in (2.18), which can be written as follows:

$$\frac{\partial C}{\partial s} = \frac{\alpha \mathrm{q}^{1-\alpha}\mu^{\alpha}}{(1-\alpha)^{1-\alpha}\alpha^{\alpha}s^{1-\alpha}} = \frac{\alpha C}{s}. \tag{A.1.1}$$

Together with (2.12), the derivative of the incumbent payoff C_{p} with respect to s is:

$$\frac{\partial C_{\mathrm{p}}}{\partial s} = \phi(s)\left(\frac{\partial C}{\partial s} - \gamma(s)\phi(s)C\right) = \frac{\phi(s)C}{s}(\alpha - s\gamma(s)\phi(s)). \tag{A.1.2}$$

Equations (A.1.1) and (A.1.2) show that $\partial C_{\mathrm{p}}/\partial s$ becomes arbitrarily large as s approaches zero. Since $s > 0$ implies $C > 0$ according to (2.17), equation (A.1.2) shows that the only values of s that can satisfy the first-order condition $\partial C_{\mathrm{p}}/\partial s = 0$ are solutions of the equation (2.20).

Differentiating (A.1.1) again with respect to s:

$$\frac{\partial^2 C}{\partial s^2} = -\frac{\alpha(1-\alpha)\mathrm{q}^{1-\alpha}\mu^{\alpha}}{(1-\alpha)^{1-\alpha}\alpha^{\alpha}s^{2-\alpha}} = -\frac{\alpha(1-\alpha)C}{s^2},$$

where the second equality uses (2.17). Together with $\gamma'(s) = -\varepsilon(s)\gamma(s)/s$ implied by the definition of the marginal cost elasticity $\varepsilon(s)$ from (2.19), equation (2.15) for the second derivative of the incumbent payoff evaluated at a point where the first-order condition holds is:

$$\left.\frac{\partial^2 C_{\mathrm{p}}}{\partial s^2}\right|_{\frac{\partial C_{\mathrm{p}}}{\partial s}=0} = -\phi(s)\left(\frac{\alpha(1-\alpha)C}{s^2} - \frac{\varepsilon(s)\gamma(s)C_{\mathrm{p}}}{s}\right) = -\frac{\phi(s)C}{s^2}(\alpha(1-\alpha) - s\varepsilon(s)\gamma(s)\phi(s)),$$

where the second equality uses (2.12). Since it has already been established that any value of s with $\partial C_{\mathrm{p}}/\partial s = 0$ must satisfy (2.20), the term in $s\gamma(s)\phi(s)$ above can be replaced by α to deduce:

$$\left.\frac{\partial^2 C_{\mathrm{p}}}{\partial s^2}\right|_{\frac{\partial C_{\mathrm{p}}}{\partial s}=0} = -\frac{\alpha\phi(s)C}{s^2}((1-\alpha) - \varepsilon(s)).$$

Under the assumption in (2.19), the second derivative is strictly negative at any value of s satisfying the first-order condition. This demonstrates that C_{p} is a strictly quasi-concave function of s.

(ii) The first derivative of C_{p} with respect to s in (A.1.2) can be written as follows:

$$\frac{\partial C_{\mathrm{p}}}{\partial s} = \frac{\phi(s)^2 C}{s}\mathcal{A}(s), \quad \text{where } \mathcal{A}(s) \equiv \frac{\alpha}{\phi(s)} - s\gamma(s). \tag{A.1.3}$$

Since $\phi(s) > 0$ for all s, the equation (2.20) is equivalent to $\mathcal{A}(s) = 0$. It is also known that the first-order condition $\partial C_{\mathrm{p}}/\partial s = 0$ can only be satisfied where (2.20) holds. Observe that $\mathcal{A}(0) = \alpha/\phi(0) > 0$ and:

$$\mathcal{A}(1) = \frac{\alpha}{\phi(1)} - \gamma(1) < 0,$$

under the restriction assumed in (2.16). This implies the first derivative (A.1.3) is strictly negative at $s = 1$, so $s = 1$ cannot be optimal. With neither $s = 0$ nor $s = 1$ optimal and the objective function being quasi-concave, the optimum is an interior solution $0 < s < 1$ satisfying the first-order condition. Differentiating the function $\mathcal{A}(s)$ from (A.1.3), note that:

$$\mathcal{A}'(s) = -\frac{\alpha\phi'(s)}{\phi(s)^2} - \gamma(s) - s\gamma'(s) = \alpha\gamma(s) - \gamma(s) + \varepsilon(s)\gamma(s),$$

using the definitions of $\gamma(s)$ and $\varepsilon(s)$ from (2.13) and (2.19). It follows that $\mathcal{A}'(s) = -\gamma(s)((1-\alpha)-\varepsilon(s)) < 0$ for all s under the assumption in (2.19). Since $\mathcal{A}(s)$ is a strictly decreasing function with $\mathcal{A}(0) > 0 > \mathcal{A}(1)$, there exists a unique solution where $\mathcal{A}(\hat{s}) = 0$ with $0 < \hat{s} < 1$. Therefore, the optimum is the unique solution of the equation (2.20). This completes the proof.

A.2 Proof of Proposition 2

(i) Suppose that a tariff (or subsidy) τ is imposed that drives a wedge between the domestic market-clearing price $\tilde{\pi}$ and the world price π^* as in equation (2.21). Substituting the expression for $\tilde{\pi}$ from (2.7):

$$\frac{\alpha(\mathrm{q}-x_{\mathrm{E}})}{(1-\alpha)(K-x_{\mathrm{I}})} = (1+\tau)\pi^*.$$

Multiplying both sides by $(1-\alpha)(K-x_{\mathrm{I}})$ and using $-\pi^* x_{\mathrm{I}} = x_{\mathrm{E}}$ implied by the international budget constraint (2.5):

$$\alpha\mathrm{q} - \alpha x_{\mathrm{E}} = (1-\alpha)(1+\tau)\pi^* K + (1-\alpha)(1+\tau)x_{\mathrm{E}}.$$

Solving this equation for x_{E} yields the expression given in (2.22). The expression for x_{I} then follows using $x_{\mathrm{I}} = -x_{\mathrm{E}}/\pi^*$ according to (2.5). The implied levels of consumption of each of the goods are obtained by substituting (2.22) into the resource constraints (2.6):

$$c_{\mathrm{E}} = \mathrm{q} - x_{\mathrm{E}} = \frac{(1-\alpha)(1+\tau)(\mathrm{q}+\pi^* K)}{1+(1-\alpha)\tau}, \quad \text{and} \quad c_{\mathrm{I}} = K - x_{\mathrm{I}} = \frac{\alpha(\mathrm{q}+\pi^* K)}{(1+(1-\alpha)\tau)\pi^*}.$$

Substituting these into the expression for aggregate consumption C from (2.9):

$$C = \Delta(\tau)\frac{\mathrm{q}+\pi^* K}{\pi^{*\alpha}}, \quad \text{where} \quad \Delta(\tau) \equiv \frac{(1+\tau)^{1-\alpha}}{1+(1-\alpha)\tau}, \tag{A.2.1}$$

which confirms the formula given in (2.23).

Note from the definition in (A.2.1) that:

$$\Delta(0) = 1, \quad \Delta'(\tau) = -\frac{\alpha(1-\alpha)(1+\tau)^{-\alpha}}{(1+(1-\alpha)\tau)^2}\tau,$$

$$\text{and} \quad \Delta''(\tau) = -\frac{\alpha(1-\alpha)(1+\tau)^{-\alpha}}{(1+(1-\alpha)\tau)^2} - \tau\frac{\partial}{\partial\tau}\frac{\alpha(1-\alpha)(1+\tau)^{-\alpha}}{(1+(1-\alpha)\tau)^2}.$$

It follows that $\Delta'(\tau) = 0$ only if $\tau = 0$, and $\Delta''(0) < 0$ when $\tau = 0$. This implies $\Delta(\tau)$ is a strictly quasi-concave function of τ, and is maximized at $\tau = 0$ where $\Delta(0) = 1$. Since those in power choose τ to maximize $C_{\mathrm{p}} = \phi(s)C$, it can be seen from (A.2.1) that the optimal value of τ is zero, confirming that free trade will be chosen.

(ii) Substituting $\tau = 0$ into (2.23) yields the expressions for total consumption and the marginal benefit of good government $\partial C/\partial s$ in (2.24). Since π^* is taken as given by a small open economy, it follows that C is linear in s, and thus $\partial^2 C/\partial s^2 = 0$. Together with (2.15), this implies:

$$\left.\frac{\partial^2 C_{\mathrm{p}}}{\partial s^2}\right|_{\frac{\partial C_{\mathrm{p}}}{\partial s}=0} = -\phi(s)\gamma'(s)C_{\mathrm{p}},$$

which is strictly positive under assumption (2.14). Since the second derivative of C_{p} is positive where the first derivative is zero, C_{p} is a strictly quasi-convex function of s. The maximum value of C_{p} is therefore found either at $s = 0$ or $s = 1$.

(iii) Using equations (2.12) and (2.24), the consumption received by those in power is:

$$C_{\mathrm{p}} = \phi(s)\frac{\mathrm{q}+\mu\pi^* s}{\pi^{*\alpha}},$$

and this objective function is strictly quasi-convex in s. It follows that $s = 1$ is optimal if:

$$\phi(1)\frac{\mathrm{q}+\mu\pi^*}{\pi^{*\alpha}} \geq \phi(0)\frac{\mathrm{q}}{\pi^{*\alpha}},$$

which is equivalent to:

$$\frac{\mu\pi^*}{\mathrm{q}} \geq \frac{\phi(0)-\phi(1)}{\phi(1)}. \tag{A.2.2}$$

Now define a function $\mathit{\Gamma}(s)$:

$$\mathit{\Gamma}(s) \equiv \frac{1}{\phi(s)}, \quad \text{and hence} \quad \mathit{\Gamma}'(s) = -\frac{\phi'(s)}{\phi(s)^2} = \gamma(s),$$

which uses the definition of $\gamma(s)$ in (2.13). The differential equation $\mathit{\Gamma}'(s) = \gamma(s)$ has the following general solution:

$$\mathit{\Gamma}(s) = \Gamma_0 + \int_{z=0}^{s} \gamma(z)\mathrm{d}z,$$

where Γ_0 is an arbitrary constant. Since $1/\phi(0) = \mathit{\Gamma}(0) = \Gamma_0$, the function $\phi(s) = 1/\mathit{\Gamma}(s)$ can be written as follows:

$$\phi(s) = \frac{\phi(0)}{1 + \phi(0)\int_{z=0}^{s} \gamma(z)\mathrm{d}z}.$$

It follows that the right-hand side of (A.2.2) can be expressed as:

$$\frac{\phi(0) - \phi(1)}{\phi(1)} = \frac{\phi(0) - \frac{\phi(0)}{1+\phi(0)\int_{s=0}^{1}\gamma(s)\mathrm{d}s}}{\frac{\phi(0)}{1+\phi(0)\int_{s=0}^{1}\gamma(s)\mathrm{d}s}} = \phi(0)\int_{s=0}^{1}\gamma(s)\mathrm{d}s \equiv \Phi, \tag{A.2.3}$$

which confirms the condition given in (2.25).

(iv) Take an arbitrary world price π^*. For each $s \in [0, 1]$, let the functions $\hat{C}(s)$ and $C^*(s)$ respectively denote the level of real GDP in autarky and with free trade ($\tau = 0$) in an open economy:

$$\hat{C}(s) = \frac{\mathrm{q}^{1-\alpha}\mu^{\alpha}s^{\alpha}}{(1-\alpha)^{1-\alpha}\alpha^{\alpha}}, \quad \text{and} \quad C^*(s) = \frac{\mathrm{q} + \mu\pi^* s}{\pi^{*\alpha}},$$

which are obtained from equations (2.17) and (2.24). Note that in an open economy with s, it would always be possible to obtain the same consumption outcomes as autarky (with the same s) by setting a tariff that results in net exports of zero. Using (2.22), the required tariff $\hat{\tau}$ is:

$$\hat{\tau} = \frac{\alpha\mathrm{q}}{(1-\alpha)\mu\pi^*\hat{s}} - 1 = \frac{\hat{\pi}}{\pi^*} - 1, \tag{A.2.4}$$

which can be written in terms of the autarky price $\hat{\pi}$ using (2.18). With $x_\mathrm{E} = 0$ and $x_\mathrm{I} = 0$ and the same s and thus same K, equation (2.9) implies that real GDP would be equal to its autarky value: $C = \hat{C}(s)$. Real GDP can also be compared to the free-trade open-economy level $C^*(s)$, with (A.2.1) implying $C = \mathit{\Delta}(\hat{\tau})C^*(s)$. It follows that $\hat{C}(s) = \mathit{\Delta}(\hat{\tau})C^*(s)$, and hence by using (2.12):

$$\hat{C}_\mathrm{p}(s) = \mathit{\Delta}(\hat{\tau})C^*_\mathrm{p}(s),$$

where $\hat{C}_\mathrm{p}(s) = \phi(s)\hat{C}(s)$ and $C^*_\mathrm{p}(s) = \phi(s)C^*(s)$ are the consumption levels of those in power respectively under autarky and with free trade in an open economy. Since $\mathit{\Delta}(\tau) \leq 1$ for all τ, this implies $\hat{C}_\mathrm{p}(s) \leq C^*_\mathrm{p}(s)$ for all $0 \leq s \leq 1$. If $\pi^* = \hat{\pi}$ for some particular value of s, (A.2.4) implies $\hat{\tau} = 0$ and $\mathit{\Delta}(\hat{\tau}) = 1$, in which case $\hat{C}_\mathrm{p}(s) = C^*_\mathrm{p}(s)$.

Given the strict quasi-convexity of $C^*_\mathrm{p}(s)$, the maximized level of consumption is:

$$C^*_\mathrm{p} = \max\{C^*_\mathrm{p}(0), C^*_\mathrm{p}(1)\} > C^*_\mathrm{p}(\hat{s}),$$

where $0 < \hat{s} < 1$ is equilibrium value of $\hat{s}$ under autarky as characterized in Proposition 1, with $\hat{C}_\mathrm{p} = \hat{C}_\mathrm{p}(\hat{s})$ being the autarky consumption of those in power. Together with $C^*_\mathrm{p}(\hat{s}) \geq \hat{C}_\mathrm{p}(\hat{s})$, it follows that $C^*_\mathrm{p} > \hat{C}_\mathrm{p}$. Those in power always strictly gain from the ability to trade with the rest of the world irrespective of world prices.

Let $\bar{C}$ and $\bar{C}_\mathrm{p}$ denote real GDP and the consumption of those in power in the case of an open economy with $s = 1$. If $s = 1$ is chosen, it must be the case that $\bar{C}_\mathrm{p} > \hat{C}_\mathrm{p}$. Since $\bar{C}_\mathrm{p} = \phi(1)\bar{C}$ and $\hat{C}_\mathrm{p} = \phi(\hat{s})\hat{C}$:

$$\bar{C} = \frac{\bar{C}_\mathrm{p}}{\phi(1)} > \frac{\hat{C}_\mathrm{p}}{\phi(1)} = \frac{\phi(\hat{s})}{\phi(1)}\hat{C} > \hat{C},$$

because $\phi(\hat{s}) > \phi(1)$. The real value of the economy's output is increased by trade if those in power choose $s = 1$. This completes the proof.

A.3 Proof of Proposition 3

(i) The equilibrium world price is (2.28), and endowments are equal across countries, so $\mathrm{q} = \mathrm{q}^*$. Given a fraction ω of countries where $s = 1$, the world supply of investment goods is $K^* = \mu\omega$, which implies:

$$\frac{\mathrm{q}\pi^*}{\mu} = \frac{\alpha}{(1-\alpha)\omega}.$$

Substituting this into (2.25) shows that the condition for $s = 1$ to be optimal for those in power is equivalent to:

$$\frac{\alpha}{(1-\alpha)\omega} \geq \Phi,$$

and rearranging this inequality confirms (2.29).

(ii) Under assumption (2.14), $\gamma(s)$ is a decreasing function of s, which implies a bound on the integral below:

$$\int_0^1 \gamma(s)\mathrm{d}s \geq \gamma(1).$$

Together with the assumption (2.16), this leads to:

$$\int_0^1 \gamma(s)\mathrm{d}s > \frac{\alpha}{\phi(1)}.$$

Multiplying both sides by $\phi(0)$ and using the definition of Φ in (2.25):

$$\Phi > \alpha\frac{\phi(0)}{\phi(1)}.$$

Equation (A.2.3) implies that $\phi(0)/\phi(1) = 1 + \Phi$, so the inequality above becomes $\Phi > \alpha(1 + \Phi)$, which can be rearranged to deduce:

$$\Phi > \frac{\alpha}{1-\alpha}. \qquad \text{[A.3.1]}$$

Now consider the world equilibrium value of ω. The value of s in each country must be an equilibrium given the world price π^* (s is either 0 or 1), and world markets must clear given the fraction ω of countries with $s = 1$. There cannot be an equilibrium with $\omega = 0$ because (2.29) implies incumbents in all countries would have an incentive to choose $s = 1$, resulting in $\omega = 1$. Similarly, there cannot be an equilibrium with $\omega = 1$. The condition in (A.3.1) implies $\alpha/((1-\alpha)\Phi) < 1$, so $\omega > \alpha/((1-\alpha)\Phi)$, indicating incumbents in all countries have incentives to choose $s = 0$ that would lead to $\omega = 0$. Finally, consider the possibility of an equilibrium with $0 < \omega < 1$. This requires that incumbents in some countries choose $s = 0$ and others choose $s = 1$. Since incumbents in all ex-ante identical countries share the same optimality condition (2.29) for $s = 1$, this condition must hold with equality:

$$\omega = \frac{\alpha}{(1-\alpha)\Phi},$$

which confirms the result for $\tilde{\omega}$ given in (2.30). It can be seen from (A.3.1) that $0 < \tilde{\omega} < 1$. The equilibrium world price $\tilde{\pi}^*$ in (2.30) follows from substituting the expression for $\tilde{\omega}$ into (2.28) with $\mathrm{q}^* = \mathrm{q}$. This is the unique world equilibrium.

Since the optimality condition (2.25) for $s = 1$ holds with equality, incumbents receive the same payoff from $s = 0$ and $s = 1$ in equilibrium. Using (2.23), the equilibrium levels of real GDP for $s = 0$ and $s = 1$ countries are respectively $\mathrm{q}/(\tilde{\pi}^*)^\alpha$ and $(\mathrm{q} + \mu\tilde{\pi}^*)/(\tilde{\pi}^*)^\alpha$, with the latter clearly being larger than the former. This completes the proof.

A.4 Proof of Proposition 4

(i) A fraction ω_0 of countries is compelled to choose $s = 1$. In the remaining fraction $1 - \omega_0$ of countries, the value of s is chosen to maximize the payoff of those in power. For these countries the logic of Proposition 3 continues to apply, with either $s = 0$ or $s = 1$ being optimal. It follows that the fraction ω of countries in the world with $s = 1$ must satisfy $\omega \geq \omega_0$. For a particular value of ω, Proposition 3 shows that $s = 1$ is optimal if (2.29) holds.

First consider the case where $\omega_0 \leq \tilde{\omega}$, where $\tilde{\omega}$ is the equilibrium fraction of countries with $s = 1$ in the absence of any direct intervention, which is $\tilde{\omega} = \alpha/(1-\alpha)\Phi < 1$ as given in (2.30). If there were an equilibrium with $\omega_0 \leq \omega < \tilde{\omega}$ then it follows from (2.29) that $s = 1$ is optimal, but this would imply $\omega = 1$ because all countries would have $s = 1$, so this cannot be an equilibrium. If there were an equilibrium with $\omega > \tilde{\omega} \geq \omega_0$ then it follows from (2.29) that $s = 0$ is optimal, implying $\omega = \omega_0$ (because only those countries compelled to would have $s = 1$), which also cannot be an equilibrium. With $\omega = \tilde{\omega} \geq \omega_0$, those in power are indifferent between $s = 0$ and $s = 1$, which implies it is possible to have any $\omega \geq \omega_0$. It follows that $\omega = \tilde{\omega}$ is the unique equilibrium. Imposing the rule of law on a fraction $\omega_0 \leq \tilde{\omega}$ of countries has no effect on the equilibrium fraction of countries with the rule of law.

Now consider the case $\omega_0 > \tilde{\omega} = \alpha/(1-\alpha)\Phi$. With the requirement $\omega \geq \omega_0 > \tilde{\omega}$, it follows from (2.29) that all countries with a choice will have $s = 0$. The unique equilibrium is $\omega = \omega_0$, so the only countries with the rule of law are those directly compelled to have it.

(ii) Suppose a fraction σ of countries impose a subsidy $\tau < 0$. Suppose these countries are drawn exclusively from those with $s = 1$, which requires $\sigma \leq \omega$. With $K = \mu$ in all these countries, equation (2.27) implies each has the following net exports of the endowment good:

$$x_{\mathrm{E}} = \frac{\alpha q - (1-\alpha)\mu(1+\tau)\pi^*}{1 + (1-\alpha)\tau}. \tag{A.4.1}$$

For the remaining fraction $1 - \sigma$ of countries with no tariff or subsidy ($\tau = 0$), net exports are $x_{\mathrm{E}} = \alpha q - (1-\alpha)\mu\pi^* s$. A measure $1 - \omega$ of these countries have $s = 0$ and $x_{\mathrm{E}} = \alpha q$, while a measure $\omega - \sigma$ have $s = 1$ and $x_{\mathrm{E}} = \alpha q - (1-\alpha)\mu\pi^*$. Using these observations together with (A.4.1), world market clearing (2.26) requires:

$$\sigma\frac{\alpha q - (1-\alpha)\mu(1+\tau)\pi^*}{1 + (1-\alpha)\tau} + (1-\omega)\alpha q + (\omega - \sigma)(\alpha q - (1-\alpha)\mu\pi^*) = 0.$$

Collecting terms proportional to π^* on one side:

$$(1-\alpha)\mu\left(\omega - \sigma + \frac{\sigma(1+\tau)}{1 + (1-\alpha)\tau}\right)\pi^* = \alpha q\left(1 - \sigma + \frac{\sigma}{1 + (1-\alpha)\tau}\right),$$

and by multiplying both sides by the positive number $1 + (1-\alpha)\tau$ and rearranging, the equilibrium world price is:

$$\pi^* = \frac{\alpha q}{(1-\alpha)\mu}\left(\frac{1 + (1-\alpha)(1-\sigma)\tau}{\alpha\sigma\tau + (1 + (1-\alpha)\tau)\omega}\right). \tag{A.4.2}$$

The condition for optimality of $s = 1$ is given in (2.25), and this applies even for those countries with $\tau < 0$ (see 2.23 and note that τ has a multiplicative effect on the real value of a country's output at world prices). Using the equilibrium world price from (A.4.2), the optimality of $s = 1$ is equivalent to:

$$\frac{\alpha\sigma\tau + (1 + (1-\alpha)\tau)\omega}{1 + (1-\alpha)(1-\sigma)\tau} \leq \frac{\alpha}{(1-\alpha)\Phi}, \tag{A.4.3}$$

noting that $1 + (1-\alpha)(1-\sigma)\tau$ must be a positive number.

There cannot be an equilibrium with $\omega = 0$ because this would imply the left-hand side of (A.4.3) is strictly negative (since $\tau < 0$), which means that $s = 1$ would be optimal for those in power in all countries. The left-hand side of (A.4.3) is strictly increasing in ω, so if there is a solution $\tilde{\omega}$ where (A.4.3) holds with equality and $0 < \tilde{\omega} \leq 1$ then $\tilde{\omega}$ is the unique equilibrium. If there is no solution in the unit interval then $\tilde{\omega} = 1$ is the unique equilibrium.

Let $\omega_0 = \alpha/(1-\alpha)\Phi$ denote the equilibrium value of ω in the absence of any subsidies (see equation 2.30). If (A.4.3) is to hold with equality, ω must satisfy:

$$\frac{\alpha\sigma\tau + (1+(1-\alpha)\tau)\omega}{1+(1-\alpha)(1-\sigma)\tau} = \omega_0.$$

This linear equation has a unique solution for $\tilde{\omega}$:

$$\tilde{\omega} = \omega_0 - \sigma\frac{\tau(\alpha + (1-\alpha)\omega_0)}{1+(1-\alpha)\tau}. \tag{A.4.4}$$

Note that:

$$\omega_0 - \frac{\tau(\alpha + (1-\alpha)\omega_0)}{1+(1-\alpha)\tau} = \frac{\omega_0 - \alpha\tau}{1+(1-\alpha)\tau} = 1 - \frac{1+\tau-\frac{\alpha}{(1-\alpha)\Phi}}{1+(1-\alpha)\tau}, \tag{A.4.5}$$

which uses $\omega_0 = \alpha/(1-\alpha)\Phi$.

Consider the case where $\sigma < 1$ and $\tau > -(1-\alpha/(1-\alpha)\Phi)$, with larger subsidies considered next. The bound on τ implies the expression in (A.4.5) is strictly less than one, and $\sigma < 1$ implies $\tilde{\omega}$ in (A.4.4) is lower than that expression. It follows that $\tilde{\omega} < 1$ and that this is the unique equilibrium. Equation (A.4.4) implies $\omega > \omega_0$ because $\tau < 0$, which demonstrates that the subsidy raises the equilibrium faction $\tilde{\omega}$ of countries with $s = 1$.

Now consider the case where $\sigma = 1$. It can be seen from (A.4.4) that $\tilde{\omega}$ is equal to the expression given in (A.4.5). If the subsidy is set so that $\tau = -(1-\alpha/(1-\alpha)\Phi)$ then the expression in (A.4.5) is equal to one, and thus $\tilde{\omega} = 1$ is the unique equilibrium with all countries having $s = 1$. This completes the proof.

A.5 Proof of Proposition 5

(i) The finding of Proposition 2 that the optimal choice of those in power is either $s = 0$ or $s = 1$ still applies here. Conditional on the fraction ω of countries with $s = 1$, the equilibrium world price π^* from (2.26) implies:

$$\frac{\mu\pi^*}{\mathrm{q}} = \frac{\alpha\mathrm{q}^*}{(1-\alpha)\mathrm{q}\omega},$$

where q is an arbitrary country-specific endowment and q^* is the global mean endowment. The condition (2.25) for the optimality of $s = 1$ also applies here, and using the equation above shows that it is equivalent to:

$$\frac{\alpha\mathrm{q}^*}{(1-\alpha)\mathrm{q}\omega} \geq \Phi, \quad \text{and hence } \omega\varrho \leq \frac{\alpha}{(1-\alpha)\Phi},$$

where $\varrho = \mathrm{q}/\mathrm{q}^*$ is the relative endowment and Φ is the constant defined in (2.25). Given the value of the variable ω, this is equivalent to the threshold condition:

$$\varrho \leq \tilde{\varrho}, \quad \text{where } \tilde{\varrho} = \frac{\alpha}{(1-\alpha)\Phi\omega}, \tag{A.5.1}$$

confirming the claim in the proposition.

(ii) Since $\varrho < \tilde{\varrho}$ is necessary for $s = 1$, the fraction of countries with $s = 1$ must satisfy $\tilde{\omega} = F(\tilde{\varrho})$. Substituting this into (A.5.1) leads to the following equation for the equilibrium $\tilde{\omega}$:

$$\tilde{\omega} = F\left(\frac{\alpha}{(1-\alpha)\Phi\tilde{\omega}}\right).$$

This equation can be stated as $H(\tilde{\omega}) = 0$, where $H(\omega)$ is defined by:

$$H(\omega) = \omega - F\left(\frac{\alpha}{(1-\alpha)\Phi\omega}\right). \tag{A.5.2}$$

The cumulative distribution function $F(\varrho)$ is weakly increasing in ϱ, which implies that the function $H(\omega)$ defined above is strictly increasing in ω. Any solution of the equation $H(\tilde{\omega}) = 0$ must therefore be unique. A property of the cumulative distribution function is $F(\infty) = 1$, which leads to $H(0) = -1$.

Proposition 3 establishes that $0 < \alpha/(1-\alpha)\Phi < 1$. Since ϱ has mean one, the probability of having ϱ strictly less than the mean must be strictly lower than one, and hence $F(\alpha/(1-\alpha)\Phi) < 1$. This implies that $H(1) = 1 - F(\alpha/(1-\alpha)\Phi) > 0$. Since ϱ has a continuous distribution, the function $H(\omega)$ defined in (A.5.2) must be continuous. With $H(0) < 0$ and $H(1) > 0$, the intermediate value theorem implies there exists a $\tilde{\omega}$ such that $H(\tilde{\omega}) = 0$ satisfying $0 < \tilde{\omega} < 1$. This confirms the claim.

(iii) Let $\omega_0 = \alpha/(1-\alpha)\Phi$ denote the equilibrium value of ω with homogeneous endowments across countries (see Proposition 3), and let $\omega^* = F(1)$ denote the fraction of countries with an endowment less than the global mean. With these definitions and (A.5.2):

$$H(\omega_0) = \omega_0 - F\left(\frac{\omega_0}{\omega_0}\right) = \omega_0 - F(1) = \omega_0 - \omega^*, \quad \text{and} \quad H(\omega^*) = \omega^* - F\left(\frac{\omega_0}{\omega^*}\right) = F(1) - F\left(\frac{\omega_0}{\omega^*}\right).$$

Since ϱ has a continuous distribution with mean equal to one, there must be some probability mass strictly below and strictly above the mean, implying $F(\omega_0/\omega^*) < F(1)$ if $\omega_0/\omega^* < 1$ and $F(\omega_0/\omega^*) > F(1)$ if $\omega_0/\omega^* > 1$. Together with the expression for $H(\omega_0)$ above, this demonstrates that $H(\omega)$ changes sign between ω_0 and ω^* (irrespective of the ordering of the terms). The unique solution $\tilde{\omega}$ must therefore lie between ω_0 and ω^*. This completes the proof.

A.6 Proof of Proposition 6

Assume that the cartel comprises a strictly positive fraction ξ of countries ($\jmath \in \Xi$) drawn from those with $\varrho > \tilde{\varrho}^\circ$ (where $\tilde{\varrho}^\circ$ is the equilibrium threshold for the rule of law when all countries are small open economies, as described in Proposition 5). The other fraction $1-\xi$ of countries ($\jmath \notin \Xi$) remain as small open economies that do not individually affect the world relative price π^*. The cartel is modelled as one large country that pools the resources of its members, and acts to maximize the consumption of those in power $\hat{C}_\mathrm{p} = \phi(\hat{s})\hat{C}$ (as in 2.12), where $\hat{\cdot}$ denotes the value of a variable in the cartel.

The cartel moves first and chooses net exports $\hat{x}_\mathrm{E}$ of the endowment good. Given this choice and the demand function of the small open economies, the auctioneer chooses the world price π^* to ensure that world markets clear (2.26). This equilibrium price will depend on the cartel's choice of net exports, so $\pi^*(\hat{x}_\mathrm{E})$ is a function of $\hat{x}_\mathrm{E}$. The cartel does not have an independent choice of net exports $\hat{x}_\mathrm{I}$ of the investment good because it must satisfy its international budget constraint (2.5):

$$\hat{x}_\mathrm{E} + \pi^*(\hat{x}_\mathrm{E})\hat{x}_\mathrm{I} = 0, \quad \text{hence} \ \ \hat{x}_\mathrm{I} = -\frac{\hat{x}_\mathrm{E}}{\pi^*(\hat{x}_\mathrm{E})}. \qquad \text{[A.6.1]}$$

The impact of the cartel on the world equilibrium price

Now consider the world equilibrium price conditional on the cartel's choice of $\hat{x}_\mathrm{E}$. A small open economy $\jmath \in [0,1]\backslash\Xi$ does not impose any tariffs ($\tau(\jmath) = 0$) and thus chooses net exports $x_\mathrm{E}(\jmath) = \alpha \mathrm{q}(\jmath) - (1-\alpha)\mu\pi^* s(\jmath)$ (see Proposition 2 and equation 2.27). For small open economies, there is specialization determined by a country's relative endowment ϱ. If ϱ is below a threshold $\tilde{\varrho}$ then $s = 1$ is optimal, otherwise $s = 0$ (see Proposition 5). Market clearing (2.26) in the world market for the endowment good requires:

$$\xi\hat{x}_\mathrm{E} + \alpha\int_{[0,1]\backslash\Xi} \mathrm{q}(\jmath)\mathrm{d}\jmath - (1-\alpha)\mu\pi^*\int_{[0,1]\backslash\Xi} \mathbb{1}[\mathrm{q}(\jmath) \leq \mathrm{q}^*\tilde{\varrho}]\mathrm{d}\jmath = 0, \qquad \text{[A.6.2]}$$

where the threshold $\tilde{\varrho}$ below which $s = 1$ is chosen depends on the price π^* (see 2.25 in Proposition 2 and use the definition $\varrho(\jmath) = \mathrm{q}(\jmath)/\mathrm{q}^*$):

$$\tilde{\varrho} = \frac{\mu\pi^*}{\mathrm{q}^*\Phi}. \qquad \text{[A.6.3]}$$

The constant Φ is defined in (2.25). Note that:

$$\int_{[0,1]\backslash\Xi} \mathrm{q}(\jmath)\mathrm{d}\jmath = \int_{[0,1]} \mathrm{q}(\jmath)\mathrm{d}\jmath - \int_{\Xi} \mathrm{q}(\jmath)\mathrm{d}\jmath = (1-\xi\hat{\varrho})\mathrm{q}^*,$$

where $\hat{\varrho}$ denotes the average endowment of cartel members relative to the global mean (since the global mean of ϱ is equal to 1 by definition, it must be the case that $1 - \xi\hat{\varrho} > 0$). Substituting this into (A.6.2), dividing both sides by αq^*, and replacing π^* by a multiple of $\tilde{\varrho}$ using (A.6.3):

$$(1 - \xi\hat{\varrho}) + \xi\hat{\varrho}\frac{\hat{x}_{\mathrm{E}}}{\alpha\hat{q}} = \frac{(1-\alpha)\Phi}{\alpha}\tilde{\varrho}\int_{[0,1]\backslash\Xi} \mathbb{1}[\varrho(\jmath) \leq \tilde{\varrho}]\mathrm{d}\jmath. \qquad \text{[A.6.4]}$$

The right-hand side is strictly increasing in $\tilde{\varrho}$, hence $\tilde{\varrho}(\hat{x}_{\mathrm{E}})$ is a strictly increasing function of $\hat{x}_{\mathrm{E}}$. Note that $\hat{x}_{\mathrm{E}} = \alpha\hat{q}$ corresponds to the net exports from the countries of the cartel were they to act as small open economies (which would choose $\hat{s} = 0$), and hence $\tilde{\varrho}(\alpha\hat{q}) = \tilde{\varrho}^\circ$. It is conjectured that $\hat{x}_{\mathrm{E}} < \alpha\hat{q}$, which implies $\tilde{\varrho} > \tilde{\varrho}^\circ$.

Given the solution $\tilde{\varrho}(\hat{x}_{\mathrm{E}})$ derived from (A.6.4), the world price is obtained from (A.6.3):

$$\pi^*(\hat{x}_{\mathrm{E}}) = \frac{q^*\Phi}{\mu}\tilde{\varrho}(\hat{x}_{\mathrm{E}}), \qquad \text{[A.6.5]}$$

which is therefore a strictly increasing function of $\hat{x}_{\mathrm{E}}$. It can be seen from (A.6.4) and (A.6.5) that $\tilde{\varrho}(\hat{x}_{\mathrm{E}})$ and $\pi^*(\hat{x}_{\mathrm{E}})$ depend only on $\hat{x}_{\mathrm{E}}$ and parameters, not on other endogenous variables. Let $\eta(\hat{x}_{\mathrm{E}})$ denote the elasticity of $\pi^*(\hat{x}_{\mathrm{E}})$ with respect to the cartel's net exports (the reciprocal of the price elasticity of the demand curve effectively faced by the cartel):

$$\eta(\hat{x}_{\mathrm{E}}) = \frac{\hat{x}_{\mathrm{E}}\pi^{*\prime}(\hat{x}_{\mathrm{E}})}{\pi^*(\hat{x}_{\mathrm{E}})}, \qquad \text{[A.6.6]}$$

which is also a function of $\hat{x}_{\mathrm{E}}$ and parameters only. Since $\pi^*(\hat{x}_{\mathrm{E}})$ is positive and strictly increasing in $\hat{x}_{\mathrm{E}}$, the elasticity $\eta(\hat{x}_{\mathrm{E}})$ is positive for $\hat{x}_{\mathrm{E}} > 0$ and negative for $\hat{x}_{\mathrm{E}} < 0$. Given the conjecture $\tilde{\varrho} < \tilde{\varrho}^\circ$ and $\varrho(\jmath) > \tilde{\varrho}^\circ$ for all $\jmath \in \Xi$:

$$\int_{[0,1]\backslash\Xi} \mathbb{1}[\varrho(\jmath) \leq \tilde{\varrho}]\mathrm{d}\jmath = \int_{[0,1]} \mathbb{1}[\varrho(\jmath) \leq \tilde{\varrho}]\mathrm{d}\jmath = F(\tilde{\varrho}),$$

where $F(\varrho)$ is the cumulative distribution function. Equation (A.6.4) for the equilibrium threshold $\tilde{\varrho}$ is therefore:

$$(1 - \xi\hat{\varrho}) + \xi\hat{\varrho}\frac{\hat{x}_{\mathrm{E}}}{\alpha\hat{q}} = \frac{(1-\alpha)\Phi}{\alpha}\tilde{\varrho}F(\tilde{\varrho}). \qquad \text{[A.6.7]}$$

Given the proportionality between $\tilde{\varrho}(\hat{x}_{\mathrm{E}})$ and $\pi^{*\prime}(\hat{x}_{\mathrm{E}})$ in (A.6.5), the elasticity $\eta(\hat{x}_{\mathrm{E}})$ defined in (A.6.6) is also the elasticity of $\tilde{\varrho}(\hat{x}_{\mathrm{E}})$ with respect to $\hat{x}_{\mathrm{E}}$:

$$\eta(\hat{x}_{\mathrm{E}}) = \frac{\hat{x}_{\mathrm{E}}\tilde{\varrho}'(\hat{x}_{\mathrm{E}})}{\tilde{\varrho}(\hat{x}_{\mathrm{E}})}.$$

Note that:

$$\frac{\partial \log(\tilde{\varrho}F(\tilde{\varrho}))}{\partial \log\tilde{\varrho}} = 1 + h(\tilde{\varrho}), \quad \text{where } h(\varrho) \equiv \frac{\varrho F'(\varrho)}{F(\varrho)}.$$

The function $h(\varrho)$ is the reversed hazard rate of the log endowment probability distribution, which is assumed to be a weakly decreasing function. It is always non-negative since $F'(\varrho) \geq 0$. Using this equation together with (A.6.7), the inverse price elasticity of net exports is equal to:

$$\eta(\hat{x}_{\mathrm{E}}) = \left(\frac{1}{1 + h(\tilde{\varrho}(\hat{x}_{\mathrm{E}}))}\right)\left(\frac{\xi\hat{\varrho}\frac{\hat{x}_{\mathrm{E}}}{\alpha\hat{q}}}{(1-\xi\hat{\varrho}) + \xi\hat{\varrho}\frac{\hat{x}_{\mathrm{E}}}{\alpha\hat{q}}}\right). \qquad \text{[A.6.8]}$$

The first term on the right-hand side is bounded by 0 and 1, and is weakly increasing in $\hat{x}_{\mathrm{E}}$ since $\tilde{\varrho}(\hat{x}_{\mathrm{E}})$ is an increasing function and $h(\varrho)$ is a decreasing function. As $1 - \xi\hat{\varrho} > 0$, the second term is strictly increasing in $\hat{x}_{\mathrm{E}}$. It is positive (but strictly less than one) when $\hat{x}_{\mathrm{E}} > 0$, and negative when $\hat{x}_{\mathrm{E}} < 0$ (the denominator cannot be negative given A.6.4). It follows that $\eta(\hat{x}_{\mathrm{E}}) < 1$ in all cases, and $\eta(\hat{x}_{\mathrm{E}})$ is a strictly increasing function where $\hat{x}_{\mathrm{E}} \geq 0$.

The cartel's choice of net exports

The cartel has a choice of $\hat{x}_{\mathrm{E}}$ and $\hat{s}$, taking the function $\pi^*(\hat{x}_{\mathrm{E}})$ as given. Taking account of the demand curve (derived from A.6.5), the international budget constraint (A.6.1) and the resource constraints (embedded in 2.9), the objective function of the cartel is:

$$\hat{C}_{\mathrm{p}} = \phi(\hat{s})\frac{(\hat{\mathrm{q}} - \hat{x}_{\mathrm{E}})^{1-\alpha}\left(\mu\hat{s} + \frac{\hat{x}_{\mathrm{E}}}{\pi^*(\hat{x}_{\mathrm{E}})}\right)^{\alpha}}{(1-\alpha)^{1-\alpha}\alpha^{\alpha}}. \qquad \text{[A.6.9]}$$

The partial derivative with respect to net exports $\hat{x}_{\mathrm{E}}$ is:

$$\frac{\partial\hat{C}_{\mathrm{p}}}{\partial\hat{x}_{\mathrm{E}}} = \phi(\hat{s})\left(\left(\frac{1}{\pi^*(\hat{x}_{\mathrm{E}})} - \frac{\hat{x}_{\mathrm{E}}\pi^{*\prime}(\hat{x}_{\mathrm{E}})}{(\pi^*(\hat{x}_{\mathrm{E}}))^2}\right)\left(\frac{\alpha(\hat{\mathrm{q}} - \hat{x}_{\mathrm{E}})}{(1-\alpha)\left(\mu\hat{s} + \frac{\hat{x}_{\mathrm{E}}}{\pi^*(\hat{x}_{\mathrm{E}})}\right)}\right)^{1-\alpha} - \left(\frac{\alpha(\hat{\mathrm{q}} - \hat{x}_{\mathrm{E}})}{(1-\alpha)\left(\mu\hat{s} + \frac{\hat{x}_{\mathrm{E}}}{\pi^*(\hat{x}_{\mathrm{E}})}\right)}\right)^{-\alpha}\right) = \phi(\hat{s})\left(\frac{1-\eta(\hat{x}_{\mathrm{E}})}{\pi^*(\hat{x}_{\mathrm{E}})}\hat{\pi}^{1-\alpha} - \hat{\pi}^{-\alpha}\right),$$

where the second equality uses the definition of the elasticity $\eta(\hat{x}_{\mathrm{E}})$ from (A.6.6) and the following expression for the domestic relative price $\hat{\pi}$ inside the cartel (derived from 2.7 and A.6.1):

$$\hat{\pi} = \frac{\alpha(\hat{\mathrm{q}} - \hat{x}_{\mathrm{E}})}{(1-\alpha)\left(\mu\hat{s} + \frac{\hat{x}_{\mathrm{E}}}{\pi^*(\hat{x}_{\mathrm{E}})}\right)}. \qquad \text{[A.6.10]}$$

The levels of domestic consumption implied by the choice of net exports must be non-negative. Since $\hat{c}_{\mathrm{E}} = \hat{\mathrm{q}} - \hat{x}_{\mathrm{E}}$ (see 2.6), this requires $\hat{x}_{\mathrm{E}} \leq \hat{\mathrm{q}}$. With $\hat{c}_{\mathrm{I}} = \mu\hat{s} + (\hat{x}_{\mathrm{E}}/\pi^*(\hat{x}_{\mathrm{E}}))$ given $\hat{s}$, it must be the case that $\hat{x}_{\mathrm{E}}/\pi^*(\hat{x}_{\mathrm{E}}) \geq -\mu\hat{s}$. As $\eta(\hat{x}_{\mathrm{E}}) < 1$, the ratio $\hat{x}_{\mathrm{E}}/\pi^*(\hat{x}_{\mathrm{E}})$ is strictly increasing in $\hat{x}_{\mathrm{E}}$. The requirement that $\hat{x}_{\mathrm{E}}/\pi^*(\hat{x}_{\mathrm{E}}) \geq -\mu\hat{s}$ is therefore equivalent to a finite lower bound for $\hat{x}_{\mathrm{E}}$ conditional on the choice of $\hat{s}$. At this lower bound, equation (A.6.10) shows that $\hat{\pi}$ becomes arbitrarily large, implying $\partial\hat{C}_{\mathrm{p}}/\partial\hat{x}_{\mathrm{E}}$ tends to infinity (since $0 < \alpha < 1$ and $\eta(\hat{x}_{\mathrm{E}}) < 1$). Similarly, at the upper bound for $\hat{x}_{\mathrm{E}}$, the domestic relative price $\hat{\pi}$ tends to zero, implying $\partial\hat{C}_{\mathrm{p}}/\partial\hat{x}_{\mathrm{E}}$ approaches minus infinity. This argument establishes that there is never a corner solution for the optimal value of $\hat{x}_{\mathrm{E}}$ conditional on a particular value of $\hat{s}$. With an interior solution, the necessary condition for optimality is $\partial\hat{C}_{\mathrm{p}}/\partial\hat{x}_{\mathrm{E}} = 0$. Noting that:

$$\frac{\partial\hat{C}_{\mathrm{p}}}{\partial\hat{x}_{\mathrm{E}}} = \frac{\phi(\hat{s})(1-\eta(\hat{x}_{\mathrm{E}}))}{\pi^*(\hat{x}_{\mathrm{E}})\hat{\pi}^{\alpha}}(\hat{\pi} - (1+\kappa(\hat{x}_{\mathrm{E}}))\pi^*(\hat{x}_{\mathrm{E}})), \quad \text{where } \kappa(\hat{x}_{\mathrm{E}}) \equiv \frac{\eta(\hat{x}_{\mathrm{E}})}{1-\eta(\hat{x}_{\mathrm{E}})}, \qquad \text{[A.6.11]}$$

the optimal choice of $\hat{x}_{\mathrm{E}}$ conditional on $\hat{s}$ requires:

$$\hat{\pi} = (1+\kappa(\hat{x}_{\mathrm{E}}))\pi^*(\hat{x}_{\mathrm{E}}). \qquad \text{[A.6.12]}$$

The value of $\kappa(\hat{x}_{\mathrm{E}})$ given in (A.6.11) is well defined because $\eta(\hat{x}_{\mathrm{E}}) < 1$. It is positive and strictly increasing in $\hat{x}_{\mathrm{E}}$ when $\hat{x}_{\mathrm{E}} > 0$ (since $\eta(\hat{x}_{\mathrm{E}})$ is positive and strictly increasing in this case), and negative when $\hat{x}_{\mathrm{E}} < 0$ (but always greater than -1).

Now consider how an arbitrary choice of $\hat{s}$ affects the optimal value of $\hat{x}_{\mathrm{E}}$. Combining equations (A.6.10) and (A.6.12):

$$(1-\alpha)(1+\kappa(\hat{x}_{\mathrm{E}}))\pi^*(\hat{x}_{\mathrm{E}})\left(\mu\hat{s} + \frac{\hat{x}_{\mathrm{E}}}{\pi^*(\hat{x}_{\mathrm{E}})}\right) = \alpha(\hat{\mathrm{q}} - \hat{x}_{\mathrm{E}}),$$

which can be rearranged to deduce:

$$\hat{x}_{\mathrm{E}} = \frac{\alpha\hat{\mathrm{q}} - (1-\alpha)\mu(1+\kappa(\hat{x}_{\mathrm{E}}))\pi^*(\hat{x}_{\mathrm{E}})\hat{s}}{1+(1-\alpha)\kappa(\hat{x}_{\mathrm{E}})}. \qquad \text{[A.6.13]}$$

Since $\kappa(\hat{x}_{\mathrm{E}}) > -1$, the denominator of the expression above is strictly positive. The equation can be rearranged to give $\hat{s}$ as a function of $\hat{x}_{\mathrm{E}}$:

$$\hat{s} = \frac{\alpha\hat{\mathrm{q}} - (\alpha + (1-\alpha)(1+\kappa(\hat{x}_{\mathrm{E}})))\hat{x}_{\mathrm{E}}}{(1-\alpha)\mu(1+\kappa(\hat{x}_{\mathrm{E}}))\pi^*(\hat{x}_{\mathrm{E}})}, \qquad \text{[A.6.14]}$$

where the denominator is strictly positive (owing to $\kappa(\hat{x}_{\mathrm{E}}) > -1$). Note that for each value of $\hat{x}_{\mathrm{E}}$, the above gives the unique value of $\hat{s}$ consistent with equation (A.6.13). At the maximum value of $\hat{x}_{\mathrm{E}}$ consistent with non-negative consumption levels of the two goods ($\hat{x}_{\mathrm{E}} = \hat{\mathrm{q}}$) the value of $\hat{s}$ implied by (A.6.14) would be negative. Similarly, at the minimum possible value of $\hat{x}_{\mathrm{E}}$ consistent with non-negative consumption ($\hat{x}_{\mathrm{E}} = -\mu\pi^*(\hat{x}_{\mathrm{E}})$, given maximum production of the investment good) the value of $\hat{s}$ implied by (A.6.14) would be greater than one. Since $\hat{s}$ is a well-defined continuous function of $\hat{x}_{\mathrm{E}}$, it follows that for each economically meaningful value of $\hat{s} \in [0, 1]$, there exists at least one value of $\hat{x}_{\mathrm{E}}$ consistent with equation (A.6.13).

Take a value of $\hat{s}$ for which there exists a solution of (A.6.13) with $\hat{x}_{\mathrm{E}} < 0$. Considering an alternative value of $\hat{x}_{\mathrm{E}} > 0$ would raise both terms in the denominator of (A.6.14) (since $\pi^*(\hat{x}_{\mathrm{E}})$ is strictly increasing, and $\kappa(\hat{x}_{\mathrm{E}})$ switches from negative to positive). As this change also lowers the denominator in (A.6.14) (because $\hat{x}_{\mathrm{E}}$ switches from negative to positive), the change to $\hat{x}_{\mathrm{E}}$ cannot be consistent with the same value of $\hat{s}$. Therefore, there cannot be another solution (given $\hat{s}$) of (A.6.13) with $\hat{x}_{\mathrm{E}} > 0$ when there is already a solution with $\hat{x}_{\mathrm{E}} < 0$. Now suppose for a given $\hat{s}$ there exists a solution of (A.6.13) with $\hat{x}_{\mathrm{E}} > 0$. If there were also a solution with $\hat{x}_{\mathrm{E}} < 0$ then the logic above would lead to a contradiction. Furthermore, the solution for $\hat{x}_{\mathrm{E}}$ must be unique in this case. Since attention can be restricted to $\hat{x}_{\mathrm{E}} > 0$, it is known that $\kappa(\hat{x}_{\mathrm{E}})$ is a strictly increasing function in this range, as well as $\pi^*(\hat{x}_{\mathrm{E}})$. Using (A.6.14), any change in $\hat{x}_{\mathrm{E}}$ would unambiguously require $\hat{s}$ to move in the opposite direction.

It can be seen from (A.6.13) that the (unique) solution for $\hat{x}_{\mathrm{E}}$ is strictly positive. Since there is a negative relationship between $\hat{s}$ and $\hat{x}_{\mathrm{E}}$ implied by (A.6.14) when $\hat{x}_{\mathrm{E}} > 0$, it follows that the implied value of $\hat{x}_{\mathrm{E}}$ is positive from $\hat{s} = 0$ up to some positive value of $\hat{s}$, and strictly decreasing in $\hat{x}_{\mathrm{E}}$. It may be the case that for sufficiently high $\hat{s}$, the required value of $\hat{x}_{\mathrm{E}}$ is negative, but it is never the case that an even higher value of $\hat{s}$ would be associated with a positive value of $\hat{x}_{\mathrm{E}}$. However, beyond that, it is not possible to make any further claims about the relationship between $\hat{s}$ and $\hat{x}_{\mathrm{E}}$ once $\hat{x}_{\mathrm{E}}$ becomes negative.

In the case where $\hat{s}$ leads to a positive value of $\hat{x}_{\mathrm{E}}$, $\hat{C}_{\mathrm{p}}$ must be a quasi-concave function of $\hat{x}_{\mathrm{E}}$ (given $\hat{s}$), implying that the optimality condition (A.6.12) for $\hat{x}_{\mathrm{E}}$ is sufficient as well as necessary. The second derivative of $\hat{C}_{\mathrm{p}}$ with respect to $\hat{x}_{\mathrm{E}}$, evaluated at a point where $\partial\hat{C}_{\mathrm{p}}/\partial\hat{x}_{\mathrm{E}} = 0$, is:

$$\left.\frac{\partial^2 \hat{C}_{\mathrm{p}}}{\partial \hat{x}_{\mathrm{E}}^2}\right|_{\frac{\partial \hat{C}_{\mathrm{p}}}{\partial \hat{x}_{\mathrm{E}}}=0} = -\frac{\phi(\hat{s})(1-\eta(\hat{x}_{\mathrm{E}}))}{\pi^*(\hat{x}_{\mathrm{E}})\hat{\pi}^{\alpha}}\left((1+\kappa(\hat{x}_{\mathrm{E}}))\pi^{*\prime}(\hat{x}_{\mathrm{E}}) + \pi^*(\hat{x}_{\mathrm{E}})\kappa'(\hat{x}_{\mathrm{E}}) - \frac{\partial\hat{\pi}}{\partial\hat{x}_{\mathrm{E}}}\right). \qquad \text{[A.6.15]}$$

Using (A.6.10), the final term in parentheses is:

$$\frac{\partial\hat{\pi}}{\partial\hat{x}_{\mathrm{E}}} = -\frac{\alpha}{(1-\alpha)\hat{c}_{\mathrm{I}}}\left(1 + \frac{(\hat{\mathrm{q}} - \hat{x}_{\mathrm{E}})\left(\frac{1-\eta(\hat{x}_{\mathrm{E}})}{\pi^*(\hat{x}_{\mathrm{E}})}\right)}{\mu\hat{s} + \frac{\hat{x}_{\mathrm{E}}}{\pi^*(\hat{x}_{\mathrm{E}})}}\right),$$

where $\hat{c}_{\mathrm{I}} = \mu\hat{s} + \hat{x}_{\mathrm{E}}/\pi^*(\hat{x}_{\mathrm{E}}) > 0$, and the definition of $\eta(\hat{x}_{\mathrm{E}})$ from (A.6.6) has been used. At a point where $\partial\hat{C}_{\mathrm{p}}/\partial\hat{x}_{\mathrm{E}} = 0$, (A.6.11) implies $(1-\eta(\hat{x}_{\mathrm{E}}))/\pi^*(\hat{x}_{\mathrm{E}}) = 1/((1+\kappa(\hat{x}_{\mathrm{E}}))\pi^*(\hat{x}_{\mathrm{E}})) = 1/\hat{\pi}$. By using (A.6.10) again, it follows that:

$$-\left.\frac{\partial\hat{\pi}}{\partial\hat{x}_{\mathrm{E}}}\right|_{\frac{\partial \hat{C}_{\mathrm{p}}}{\partial \hat{x}_{\mathrm{E}}}=0} = \frac{\alpha}{(1-\alpha)\hat{c}_{\mathrm{I}}}\left(1 + \frac{1}{\hat{\pi}}\frac{\hat{\mathrm{q}} - \hat{x}_{\mathrm{E}}}{\mu\hat{s} + \frac{\hat{x}_{\mathrm{E}}}{\pi^*(\hat{x}_{\mathrm{E}})}}\right) = \frac{\alpha}{(1-\alpha)\hat{c}_{\mathrm{I}}}\left(1 + \frac{1-\alpha}{\alpha}\right) = \frac{1}{(1-\alpha)\hat{c}_{\mathrm{I}}},$$

which is strictly positive. Together with $\kappa'(\hat{x}_{\mathrm{E}}) > 0$ when $\hat{x}_{\mathrm{E}} > 0$, and $1 + \kappa(\hat{x}_{\mathrm{E}}) > 0$ and $\pi^{*\prime}(\hat{x}_{\mathrm{E}}) > 0$, the second derivative in (A.6.15) is seen to be strictly negative (recall $\eta(\hat{x}_{\mathrm{E}}) < 1$). This establishes that $\hat{C}_{\mathrm{p}}$ is a quasi-concave function of $\hat{x}_{\mathrm{E}}$ where $\hat{x}_{\mathrm{E}} > 0$.

The choice of the quality of government in the cartel

Now define a function $G(\hat{s})$ that represents log consumption $\log\hat{C}$ for an arbitrary value of $\hat{s}$ (see A.6.9, noting that $\hat{C}_{\mathrm{p}} = \phi(\hat{s})\hat{C}$), but with maximization over net exports $\hat{x}_{\mathrm{E}}$:

$$G(\hat{s}) = \max_{\hat{x}_{\mathrm{E}}}\left\{(1-\alpha)\log(\hat{\mathrm{q}} - \hat{x}_{\mathrm{E}}) + \alpha\log\left(\mu\hat{s} + \frac{\hat{x}_{\mathrm{E}}}{\pi^*(\hat{x}_{\mathrm{E}})}\right) - \log(1-\alpha)^{1-\alpha}\alpha^{\alpha}\right\}. \qquad \text{[A.6.16]}$$

By the envelope theorem, the derivative of $G(\hat{s})$ with respect to $\hat{s}$ can be obtained by holding $\hat{x}_{\mathrm{E}}$ constant:

$$G'(\hat{s}) = \frac{\alpha\mu}{\mu\hat{s} + \frac{\hat{x}_{\mathrm{E}}}{\pi^*(\hat{x}_{\mathrm{E}})}} = \frac{\alpha\mu\pi^*(\hat{x}_{\mathrm{E}})}{\hat{x}_{\mathrm{E}} + \mu\pi^*(\hat{x}_{\mathrm{E}})\hat{s}}. \tag{A.6.17}$$

Since $\phi(\hat{s})$ is not a function of $\hat{x}_{\mathrm{E}}$, the value of $\hat{x}_{\mathrm{E}}$ that maximizes $G(\hat{s})$ for a given $\hat{s}$ is the same as the one that maximizes $\hat{C}_{\mathrm{p}} = \phi(\hat{s})\hat{C}$. The first-order condition for the optimal $\hat{x}_{\mathrm{E}}$ is therefore (A.6.12), which implies equation (A.6.13) as before. That equation can be used to deduce:

$$\hat{x}_{\mathrm{E}} + \mu\pi^*(\hat{x}_{\mathrm{E}})\hat{s} = \frac{\alpha\hat{q} - (1-\alpha)(1+\kappa(\hat{x}_{\mathrm{E}}))\mu\pi^*(\hat{x}_{\mathrm{E}})\hat{s} + \mu\pi^*(\hat{x}_{\mathrm{E}})\hat{s} + (1-\alpha)\kappa(\hat{x}_{\mathrm{E}})\mu\pi^*(\hat{x}_{\mathrm{E}})\hat{s}}{1 + (1-\alpha)\kappa(\hat{x}_{\mathrm{E}})}$$
$$= \frac{\alpha(\hat{q} + \mu\pi^*(\hat{x}_{\mathrm{E}})\hat{s})}{1 + (1-\alpha)\kappa(\hat{x}_{\mathrm{E}})},$$

and substituting this into (A.6.17) implies that the derivative $G'(\hat{s})$ is given by the following for all $\hat{s}$:

$$G'(\hat{s}) = (1 + (1-\alpha)\kappa(\hat{x}_{\mathrm{E}}))\frac{\mu\pi^*(\hat{x}_{\mathrm{E}})}{\hat{q} + \mu\pi^*(\hat{x}_{\mathrm{E}})\hat{s}}, \tag{A.6.18}$$

where $\hat{x}_{\mathrm{E}}$ is given by equation (A.6.13) conditional on the value of $\hat{s}$.

Let π^* denote the value of $\pi^*(\hat{x}_{\mathrm{E}})$ for the optimal $\hat{x}_{\mathrm{E}}$ when $\hat{s} = 0$, and $\hat{\kappa}$ be the corresponding value of $\kappa(\hat{x}_{\mathrm{E}})$. With $\hat{s} = 0$ implying $\hat{x}_{\mathrm{E}} > 0$, it must be the case that $\hat{\kappa} > 0$. Since $\hat{s} > 0$ never features an optimizing value of $\hat{x}_{\mathrm{E}}$ greater than when $\hat{s} = 0$, and as $\pi^*(\hat{x}_{\mathrm{E}})$ is a strictly increasing function, it follows that $\pi^*(\hat{x}_{\mathrm{E}}) \leq \pi^*$ for all $\hat{s} \in [0,1]$. Similarly, since $\kappa(\hat{x}_{\mathrm{E}})$ is a strictly increasing function while $\hat{x}_{\mathrm{E}} > 0$ and strictly negative when $\hat{x}_{\mathrm{E}} < 0$, it follows that $\kappa(\hat{x}_{\mathrm{E}}) \leq \hat{\kappa}$ for all $\hat{s} \in [0,1]$. Finally, noting that the right-hand side of (A.6.18) is increasing in $\kappa(\hat{x}_{\mathrm{E}})$, and $\mu\pi^*(\hat{x}_{\mathrm{E}})/(\hat{q} + \mu\pi^*\hat{s}(\hat{x}_{\mathrm{E}}))$ is increasing in π^*, it follows that:

$$G'(\hat{s}) \leq (1 + (1-\alpha)\hat{\kappa})\frac{\mu\pi^*}{\hat{q} + \mu\pi^*\hat{s}}, \tag{A.6.19}$$

for all $\hat{s} \in [0,1]$. Given the integral below:

$$\int_{s=0}^{\hat{s}} \frac{\mu\pi^*}{\hat{q} + \mu\pi^* s}\mathrm{d}s = \log(\hat{q} + \mu\pi^*\hat{s}) - \log\hat{q} = \log\left(1 + \frac{\mu\pi^*\hat{s}}{\hat{q}}\right),$$

by integrating both sides of (A.6.19) from 0 to $\hat{s}$, it follows that:

$$G(\hat{s}) - G(0) \leq (1 + (1-\alpha)\hat{\kappa})\log\left(1 + \frac{\mu\pi^*\hat{s}}{\hat{q}}\right). \tag{A.6.20}$$

Noting that $\log\hat{C}_{\mathrm{p}} = \log\phi(\hat{s}) + G(\hat{s})$, $\hat{s} = 0$ is the optimal choice if:

$$G(\hat{s}) - G(0) \leq \log\phi(0) - \log\phi(\hat{s}), \quad \text{for all } \hat{s} \in [0,1].$$

From (A.6.20), a sufficient condition for this is:

$$(1 + (1-\alpha)\hat{\kappa})\log\left(1 + \frac{\mu\pi^*\hat{s}}{\hat{q}}\right) \leq \log\phi(0) - \log\phi(\hat{s}), \tag{A.6.21}$$

for all $\hat{s} \in [0,1]$.

The analysis of small open economies has shown that the objective function $\phi(s)(\mathrm{q} + \mu\pi^* s)$ is a quasi-convex function of s, and that there exists a threshold $\tilde{\mathrm{q}}$ such that for any $\mathrm{q} \geq \tilde{\mathrm{q}}$, $s = 0$ maximizes the objective function (see Proposition 2 and Proposition 5). This implies:

$$\log\left(1 + \frac{\mu\pi^*}{\tilde{q}}\hat{s}\right) \leq \log\phi(0) - \log\phi(\hat{s}), \quad \text{for all } \hat{s} \in [0,1], \quad \text{where } \tilde{q} = \frac{\mu\pi^*}{\Phi}. \tag{A.6.22}$$

Now consider the following inequality:

$$1 + \Phi\frac{\tilde{\varrho}}{\hat{\varrho}}s \leq (1 + \Phi s)^{\frac{1}{1+(1-\alpha)\hat{\kappa}}}. \tag{A.6.23}$$

The left- and right-hand sides are equal when $s = 0$. Note that the left-hand side is linear in s, while the right-hand side is concave function of s because $1 + (1-\alpha)\hat{\kappa} > 1$. Hence, the inequality is satisfied for all

$s \in [0,1]$ if and only if it is satisfied for $s = 1$:

$$1 + \Phi\frac{\tilde{\varrho}}{\hat{\varrho}} \leq (1+\Phi)^{\frac{1}{1+(1-\alpha)\hat{\kappa}}},$$

which can be rearranged as follows:

$$\hat{\varrho} \geq \left(\frac{\Phi}{(1+\Phi)^{\frac{1}{1+(1-\alpha)\hat{\kappa}}} - 1}\right)\tilde{\varrho}. \qquad \text{[A.6.24]}$$

Lastly, by taking logarithms of both sides of (A.6.23) and using the definition of $\tilde{\mathrm{q}}$ from (A.6.22) and the definitions $\hat{\varrho} = \hat{\mathrm{q}}/\mathrm{q}^*$ and $\tilde{\varrho} = \tilde{\mathrm{q}}/\mathrm{q}^*$, (A.6.23) is seen to be equivalent to:

$$(1+(1-\alpha)\hat{\kappa})\log\left(1+\frac{\mu\pi^*}{\hat{\mathrm{q}}}s\right) \leq \log\left(1+\frac{\mu\pi^*}{\tilde{\mathrm{q}}}s\right).$$

Therefore, if (A.6.24) is satisfied, this implies (A.6.3) holds for all $s \in [0,1]$, which is equivalent to the above inequality, which together with (A.6.22) implies (A.6.21) holds for all $\hat{s} \in [0,1]$. This is a sufficient condition for $\hat{s} = 0$ to be the optimal choice for the cartel.

The equilibrium consequences of the cartel for other countries

Since $\hat{s} > 0$ implies a value of $\kappa(\hat{x}_{\mathrm{E}})$ equal to $\hat{\kappa} > 0$, equation (A.6.13) shows that the optimal value of net exports is such that $\hat{x}_{\mathrm{E}} < \alpha\hat{\mathrm{q}}$. This is less than what the countries of the cartel would choose if they were small open economies ($\hat{x}^\circ = \alpha\hat{\mathrm{q}}$), and since $\tilde{\varrho}$ is strictly increasing in $\hat{x}_{\mathrm{E}}$, this confirms the earlier conjecture that $\tilde{\varrho} < \tilde{\varrho}^\circ$.

As equation (A.6.12) holds with $\kappa(\hat{x}_{\mathrm{E}}) = \hat{\kappa} > 0$, the cartel's pricing strategy is equivalent to a tariff $\hat{\tau} = \hat{\kappa}$ (compare equations 2.21 and A.6.12), where the tariff is equal to the cartel's optimal markup. The world relative price π^* of the investment good is proportional to $\tilde{\varrho}$ (see A.6.5), and since $\tilde{\varrho}$ falls, so must π^*. This means the formation of the cartel leads to a higher relative price of the endowment good in equilibrium, confirming claim (i) in the proposition. The fraction of countries choosing $s = 1$ (those satisfying $\varrho \leq \tilde{\varrho}$) is $\omega = F(\tilde{\varrho})$, and the reduction in $\tilde{\varrho}$ implies that the formation of the cartel strictly reduces ω in equilibrium. This confirms claim (ii) in the proposition.

Conditions under which the cartel does not produce any of the investment good

Since $\tilde{\varrho} < \tilde{\varrho}^\circ$, it can be seen from (A.6.24) that an easier-to-verify sufficient condition for the cartel to choose $\hat{s} = 0$ is:

$$\hat{\varrho} \geq \left(\frac{\Phi}{(1+\Phi)^{\frac{1}{1+(1-\alpha)\hat{\kappa}}} - 1}\right)\tilde{\varrho}^\circ. \qquad \text{[A.6.25]}$$

It is also possible to state a sufficient condition purely in terms of exogenous parameters. With $0 < \hat{x}_{\mathrm{E}} < \alpha\hat{\mathrm{q}}$, note that (A.6.8) implies $\eta(\hat{x}_{\mathrm{E}}) \leq \xi\hat{\varrho}$. Since $\kappa(\hat{x}_{\mathrm{E}})$ is increasing in $\hat{x}_{\mathrm{E}}$ according to (A.6.11), it follows that:

$$\hat{\kappa} \leq \frac{\xi\hat{\varrho}}{1-\xi\hat{\varrho}}.$$

Since the denominator of (A.6.25) is decreasing in $\hat{\kappa}$, the inequality above demonstrates that:

$$\hat{\varrho} \geq \left(\frac{\Phi}{(1+\Phi)^{\frac{1}{1+(1-\alpha)\frac{\xi\hat{\varrho}}{1-\xi\hat{\varrho}}}} - 1}\right)\tilde{\varrho}^\circ, \qquad \text{[A.6.26]}$$

is a sufficient condition for (A.6.25) and the optimality of $\hat{s} = 0$. The term in parentheses is increasing in the cartel size ξ and tends to 1 as ξ tends to zero. It follows that (A.6.26) is satisfied for a combination of ξ being not too large and the relative endowment $\hat{\varrho}$ being sufficiently above the average. This completes the proof.

A.7 Proof of Proposition 7

An allocation specifies the individuals in power $\mathcal{P}$ (a set of measure $p = |\mathcal{P}|$), and those individuals $\mathcal{D}$ for whom the consumption allocation is contingent on investing (a set of measure $d = |\mathcal{D}|$). The remaining individuals belong to a set $\mathcal{N} = [0,1]\backslash(\mathcal{P} \cup \mathcal{D})$, noting that $\mathcal{P} \cap \mathcal{D} = \emptyset$ as those in power do not receive investment opportunities. Since there are $1-p$ non-incumbents in total, and μ investment opportunities distributed at random among them, each non-incumbent has probability $\nu = \mu/(1-p)$ of receiving an investment opportunity (this is a well-defined probability because $1-p \geq 1/2$ and $\mu < 1/2$ according to 4.5).

If this allocation prevails (that is, if it is not replaced by a rebellion) there will be three groups of individuals (payoffs the same within groups) after investment decisions have been made: incumbents $\mathcal{P}$, capitalists $\mathcal{K} = \mathcal{D} \cap \mathcal{I}$, and workers $\mathcal{W} = [0,1]\backslash(\mathcal{P} \cup \mathcal{K})$. As the allocation is the prevailing one from the pre-investment stage, the set of investors $\mathcal{I}$ must be a subset of $\mathcal{D}$ otherwise investing would not have been incentive compatible (see 4.3), hence $\mathcal{K} = \mathcal{I}$. Since each investment opportunity leads to one unit of the investment good, $|\mathcal{I}| = K = \mu s$ (see 2.10), where s is the fraction of investment opportunities that are taken, there are $k = \mu s$ capitalists, and $w = 1 - p - k$ workers. Using the utility function (4.1), the continuation payoffs (ignoring sunk costs) of these individuals at the post-investment stage under the current allocation are:

$$U_{\mathrm{p}} = \log C_{\mathrm{p}}, \quad U_{\mathrm{k}} = \log C_{\mathrm{k}}, \quad \text{and} \ \ U_{\mathrm{w}} = \log C_{\mathrm{w}}. \tag{A.7.1}$$

There are also three groups of individuals (payoffs the same within groups) at the pre-investment stage: incumbents $\mathcal{P}$, those whose consumption allocation is contingent on taking an investment opportunity (those in $\mathcal{D}$), and those who are neither incumbents nor have a consumption allocation contingent on investing ($\mathcal{N}$). Those in the group $\mathcal{N}$ have no incentive to invest even if they receive an opportunity, and thus will become workers. If the incentive constraint (4.2) is satisfied, those in $\mathcal{D}$ who receive an opportunity (probability ν) will invest and become capitalists subsequently. Those who do not receive an opportunity (probability $1-\nu$) become workers. The measure of investors is therefore $\mu s = \nu d$, which implies the group $\mathcal{D}$ has size $d = (1-p)s$ using $\nu = \mu/(1-p)$, and the group $\mathcal{N}$ has size $(1-p)(1-s)$. If the incentive constraint (4.2) is satisfied then the allocation determines s (see 4.3) through the number of individuals included in $\mathcal{D}$, and $s = 0$ otherwise. The continuation payoffs at the pre-investment stage are:

$$U_{\mathrm{p}} = \log C_{\mathrm{p}}, \quad U_{\mathrm{d}} = (1-\nu)\log C_{\mathrm{w}} + \nu\left(\log C_{\mathrm{k}} - \log(1+\theta)\right), \quad \text{and} \ \ U_{\mathrm{n}} = \log C_{\mathrm{w}}. \tag{A.7.2}$$

In what follows, let $U(\imath)$ denote the utility of individual $\imath$ under the current allocation.

Now consider the conditions that must be satisfied for there to be no rational rebellion (Definition 1) against the allocation. There are opportunities for rebellion before and after investment decisions are made. Starting from rebellions at the pre-investment stage, let $\mathcal{P}'$ and $U'(\imath)$ denote a given set of individuals in power and individuals' levels of utility under a new allocation established after a rebellion ($'$ denotes an aspect of an allocation established following a rebellion at the pre-investment stage). The absence of a rational rebellion requires that (4.4) cannot hold for any $\{\mathcal{L}, \mathcal{R}, R(\imath)\}$. Since rebellion effort $R(\imath)$ must be non-negative (see 4.6a) and the loyal faction $\mathcal{L}$ excludes those belonging to the rebel faction $\mathcal{R}$ (see 4.6b), the absence of any rational rebellion is equivalent to establishing (4.4) does not hold when $\mathcal{R}$ includes all individuals in $\mathcal{P}'$ with $U'(\imath) \geq U(\imath)$, that is, $U'_{\mathrm{p}} \geq U(\imath)$, and when these individuals exert the maximum amount of rebellion effort $R(\imath) = \exp\{U'(\imath) - U(\imath)\} - 1$ (see 4.6a). The strength of the rebel faction in this case is:

$$\int_{\mathcal{R}} R(\imath)\mathrm{d}\imath = \int_{\mathcal{R}} \left(\exp\{U'_{\mathrm{p}} - U(\imath)\} - 1\right)\mathrm{d}\imath = \int_{\mathcal{P}'} \max\{\exp\{U'_{\mathrm{p}} - U(\imath)\} - 1, 0\}\mathrm{d}\imath. \tag{A.7.3}$$

The corresponding loyal faction $\mathcal{L}$ is determined by (4.6b). Any individual $\imath \in \mathcal{P}$ with $U_{\mathrm{p}} > U'(\imath)$ must belong to $\mathcal{L}$ because the requirement $U'(\imath) \geq U(\imath)$ for belonging to the rebel faction cannot hold (see 4.6a). It follows that $\mathcal{L}$ comprises all individuals in $\mathcal{P}$ with $U_{\mathrm{p}} > U'(\imath)$, and thus the strength of the loyal faction is:

$$\int_{\mathcal{L}} \delta\mathrm{d}\imath = \delta\int_{\mathcal{P}} \mathbb{1}[U_{\mathrm{p}} > U'(\imath)]\mathrm{d}\imath, \tag{A.7.4}$$

where $\mathbb{1}[\cdot]$ is the indicator function. Given $\mathcal{P}'$ and $U'(\imath)$, the absence of any rational rebellion at the pre-investment stage is therefore equivalent to:

$$\int_{\mathcal{P}'} \max\{\exp\{U'_{\mathrm{p}} - U(\imath)\} - 1, 0\}\mathrm{d}\imath \leq \delta \int_{\mathcal{P}} \mathbb{1}[U_{\mathrm{p}} > U'(\imath)]\mathrm{d}\imath. \tag{A.7.5a}$$

Considering the post-investment stage, let $\mathcal{P}^{\ddagger}$ and $U^{\ddagger}(\imath)$ denote a given set of individuals in power and individuals' levels of utility under a new allocation established after a rebellion (‡ denotes an aspect of an allocation established following a rebellion at the post-investment stage). Following the same reasoning that led to (A.7.5a), the absence of any rational rebellion at the post-investment stage is equivalent to:

$$\int_{\mathcal{P}^{\ddagger}} \max\{\exp\{U^{\ddagger}_{\mathrm{p}} - U(\imath)\} - 1, 0\}\mathrm{d}\imath \leq \delta \int_{\mathcal{P}} \mathbb{1}[U_{\mathrm{p}} > U^{\ddagger}(\imath)]\mathrm{d}\imath. \tag{A.7.5b}$$

It is supposed for now that incumbents will receive more utility than if they were to lose power through a rebellion at any stage of the power struggle, and that incumbents obtain a higher level of utility by avoiding all rebellions (these two statements are not the same because incumbents can themselves participate in rebellions). The second conjecture means that the conditions (A.7.5a) and (A.7.5b) for the absence of rebellion are treated as constraints when maximizing incumbents' payoff, referred to as 'no-rebellion constraints'. The no-rebellion constraints in (A.7.5) can also be simplified under the first conjecture because it means the condition in (4.6b) for participation in the loyal faction is automatically satisfied for any incumbent who loses power through a rebellion (so $\mathcal{L}$ is the set $\mathcal{P}$ excluding those in the rebel faction $\mathcal{R}$). The pre- and post-investment no-rebellion constraints are therefore:

$$\int_{\mathcal{P}'} \max\{\exp\{U'_{\mathrm{p}} - U(\imath)\} - 1, 0\}\mathrm{d}\imath \leq \delta \left(\int_{\mathcal{P}} \mathrm{d}\imath - \mathbb{1}[U_{\mathrm{p}} \leq U'_{\mathrm{p}}] \int_{\mathcal{P}\cap\mathcal{P}'} \mathrm{d}\imath\right); \tag{A.7.6a}$$

$$\int_{\mathcal{P}^{\ddagger}} \max\{\exp\{U^{\ddagger}_{\mathrm{p}} - U(\imath)\} - 1, 0\}\mathrm{d}\imath \leq \delta \left(\int_{\mathcal{P}} \mathrm{d}\imath - \mathbb{1}[U_{\mathrm{p}} \leq U^{\ddagger}_{\mathrm{p}}] \int_{\mathcal{P}\cap\mathcal{P}^{\ddagger}} \mathrm{d}\imath\right), \tag{A.7.6b}$$

where $\mathcal{P}'$ and U'_{p} is the set of those in power and their payoff following a pre-investment rebellion, and $p^{\ddagger}$ and $U^{\ddagger}_{\mathrm{p}}$ are the post-investment equivalents. Note that since the constraints in (A.7.6) are looser than the originals in (A.7.5), it suffices to work with (A.7.6) first and then verify the conjecture that incumbents lose from rebellions.

The equilibrium conditions (Definition 2) restrict any rational rebellion to be followed by a new allocation that is an equilibrium of the power struggle. This means (A.7.6a) must be checked for all $\mathcal{P}'$ and U'_{p} corresponding to equilibrium allocations at the pre-investment stage, and (A.7.6b) for all $\mathcal{P}^{\ddagger}$ and $U^{\ddagger}_{\mathrm{p}}$ corresponding to equilibrium allocations established at the post-investment stage. Given the independence of irrelevant history condition in the equilibrium definition (Definition 2), this means there will be some definite values of U'_{p} and $U^{\ddagger}_{\mathrm{p}}$ and some definite sizes of the sets $\mathcal{P}'$ and $\mathcal{P}^{\ddagger}$ consistent with equilibrium, where these may depend on relevant state variables. But since all individuals are ex ante identical, the equilibrium conditions place no restrictions on the identities of incumbents and non-incumbents: any permutation of the individual identities in an equilibrium allocation is also an equilibrium. No-rebellion constraints are therefore derived by considering all possible compositions of the subsequent incumbent group subject to $|\mathcal{P}'| = p'$ and $|\mathcal{P}^{\ddagger}| = p^{\ddagger}$ for the equilibrium incumbent group sizes p' and $p^{\ddagger}$ at the pre-investment and post-investment stages, respectively. There are no relevant state variables at the pre-investment stage, so independence of irrelevant history simply requires that $p = p'$ and $U_{\mathrm{p}} = U'_{\mathrm{p}}$ hold in equilibrium. At the post-investment stage, the capital stock K is a relevant state variable, and $p^{\ddagger}$ and $U^{\ddagger}_{\mathrm{p}}$ may depend on K.

Since individuals in a group all receive the same continuation payoffs (A.7.1) and (A.7.2) under the prevailing allocation, the collection of no-rebellion constraints (A.7.6) can be stated more concisely in terms of the fraction of the post-rebellion incumbents that would be drawn from each current group. For a rebellion at the pre-investment stage, let ζ_{n}, ζ_{d}, and ζ_{p} denote respectively the fractions of the size p' post-rebellion incumbent group drawn from those in $\mathcal{N}$, $\mathcal{D}$, and $\mathcal{P}$. These proportions must be non-negative and sum to one, as well as satisfying the natural restrictions $\zeta_{\mathrm{n}} \leq (1-p)(1-s)/p'$, $\zeta_{\mathrm{d}} \leq (1-p)s/p'$, and $\zeta_{\mathrm{p}} \leq p/p'$. Similarly, for a rebellion at the post-investment stage, let ζ_{w}, ζ_{k}, and ζ_{p} denote respectively the fractions of the size $p^{\ddagger}$ post-rebellion incumbent group drawn from current workers, capitalists, and incumbents. The natural restrictions on these fractions are $\zeta_{\mathrm{w}} \leq (1-p-k)/p^{\ddagger}$, $\zeta_{\mathrm{k}} \leq k/p^{\ddagger}$, and $\zeta_{\mathrm{p}} \leq p/p^{\ddagger}$.

Using this notation, the pre-investment no-rebellion constraints (A.7.6a) can be stated as:

$$\zeta_{\mathrm{n}} p' \max\{\exp\{U'_{\mathrm{p}} - U_{\mathrm{n}}\} - 1, 0\} + \zeta_{\mathrm{d}} p' \max\{\exp\{U'_{\mathrm{p}} - U_{\mathrm{d}}\} - 1, 0\} \\ + \zeta_{\mathrm{p}} p' \mathbb{1}[U_{\mathrm{p}} \leq U'_{\mathrm{p}}] \left(\exp\{U'_{\mathrm{p}} - U_{\mathrm{p}}\} - 1 + \delta\right) \leq \delta p, \quad \text{[A.7.7a]}$$

and the post-investment no-rebellion constraints (A.7.6b) as:

$$\zeta_{\mathrm{w}} p^{\ddagger} \max\{\exp\{U^{\ddagger}_{\mathrm{p}} - U_{\mathrm{w}}\} - 1, 0\} + \zeta_{\mathrm{k}} p^{\ddagger} \max\{\exp\{U^{\ddagger}_{\mathrm{p}} - U_{\mathrm{k}}\} - 1, 0\} \\ + \zeta_{\mathrm{p}} p^{\ddagger} \mathbb{1}[U_{\mathrm{p}} \leq U^{\ddagger}_{\mathrm{p}}] \left(\exp\{U^{\ddagger}_{\mathrm{p}} - U_{\mathrm{p}}\} - 1 + \delta\right) \leq \delta p. \quad \text{[A.7.7b]}$$

Following from this, the problem of finding an allocation to maximize the incumbent payoff subject to avoiding rebellions can be stated without reference to individual identities. The equilibrium conditions will determine power sharing $p = |\mathcal{P}|$ rather than the identities of the individuals in $\mathcal{P}$. Likewise, the equilibrium conditions will determine the fraction s of investment opportunities that are taken through $d = |\mathcal{D}|$ and $s = d/(1-p)$ (if 4.2 is satisfied, and $s = 0$ otherwise).

Free markets and no taxes that distort the allocation of consumption between goods

The utilities under the current allocation in (A.7.1) and (A.7.2) depend only on consumption (and power sharing p through ν). Using the consumption aggregator in (2.1):

$$\log C_{\mathrm{p}} = (1-\alpha)\log c_{\mathrm{pE}} + \alpha \log c_{\mathrm{pI}} - \log\left((1-\alpha)^{1-\alpha}\alpha^{\alpha}\right); \quad \text{[A.7.8a]}$$

$$\log C_{\mathrm{k}} = (1-\alpha)\log c_{\mathrm{kE}} + \alpha \log c_{\mathrm{kI}} - \log\left((1-\alpha)^{1-\alpha}\alpha^{\alpha}\right); \text{ and} \quad \text{[A.7.8b]}$$

$$\log C_{\mathrm{w}} = (1-\alpha)\log c_{\mathrm{wE}} + \alpha \log c_{\mathrm{wI}} - \log\left((1-\alpha)^{1-\alpha}\alpha^{\alpha}\right). \quad \text{[A.7.8c]}$$

Given the sizes p, k, and $1-p-k$ of the groups of incumbents, capitalists, and workers, the resource constraints in (2.6) can be written as follows:

$$p c_{\mathrm{pE}} + k c_{\mathrm{kE}} + (1-p-k) c_{\mathrm{wE}} = \mathrm{q} - x_{\mathrm{E}}, \quad \text{and} \quad p c_{\mathrm{pI}} + k c_{\mathrm{kI}} + (1-p-k) c_{\mathrm{wI}} = K - x_{\mathrm{I}}. \quad \text{[A.7.9]}$$

An equilibrium allocation maximizes the incumbent payoff U_{p} subject to the no-rebellion constraints (A.7.7), the resource constraints (A.7.9), the international budget constraint (2.5), and the capital stock (2.10). Taking as given the sizes of the different groups, the no-rebellion constraints (A.7.7) depend only on the continuation utilities and K through $p^{\ddagger}$ or $U^{\ddagger}_{\mathrm{p}}$. Given p, s, x_{E}, and x_{I}, consider the allocation of goods consistent with the resource constraints (A.7.9) that maximizes U_{p} subject to providing payoffs U_{k} and U_{w} above thresholds sufficient to satisfy the no-rebellion constraints. The first-order conditions for the consumption allocation $\{c_{\mathrm{pE}}, c_{\mathrm{pI}}, c_{\mathrm{kE}}, c_{\mathrm{kI}}, c_{\mathrm{wE}}, c_{\mathrm{wI}}\}$ are:

$$\frac{\alpha / c_{\mathrm{pI}}}{(1-\alpha)/c_{\mathrm{pE}}} = \frac{\alpha / c_{\mathrm{kI}}}{(1-\alpha)/c_{\mathrm{kE}}} = \frac{\alpha / c_{\mathrm{wI}}}{(1-\alpha)/c_{\mathrm{wE}}}.$$

This requires equating the marginal rates of substitution between goods across all individuals, and given the consumption basket (2.1), this means equating the consumption ratios of the two goods $c_{\mathrm{I}}(\imath)/c_{\mathrm{E}}(\imath)$ across all individuals $\imath$. Using the resource constraints (A.7.9), this common ratio is:

$$\frac{c_{\mathrm{pI}}}{c_{\mathrm{pE}}} = \frac{c_{\mathrm{kI}}}{c_{\mathrm{kE}}} = \frac{c_{\mathrm{wI}}}{c_{\mathrm{wE}}} = \frac{K - x_{\mathrm{I}}}{\mathrm{q} - x_{\mathrm{E}}}. \quad \text{[A.7.10]}$$

If there were a market where individuals could exchange the two goods (with relative price π of the investment good) subject to individual disposable incomes $Y(\imath)$ (in terms of the endowment good) then the demand functions would be those given in (2.4) (derived from maximizing 2.1 subject to 2.2). Those demand functions imply a common consumption ratio equal to (2.3), and since markets would clear at price $\pi = \tilde{\pi}$ from (2.7), the consumption ratio would be exactly the same as the one in equation (A.7.10). Now let Y_{p}, Y_{k}, and Y_{w} denote the disposable incomes (in terms of the endowment good) of incumbents, capitalists, and workers, and suppose these satisfy the following aggregate budget constraint:

$$p Y_{\mathrm{p}} + k Y_{\mathrm{k}} + (1-p-k) Y_{\mathrm{w}} = Y, \quad \text{[A.7.11]}$$

where national income Y is given in (2.8). National income can be written as $Y = (\mathrm{q} - x_{\mathrm{E}}) + \tilde{\pi}(K - x_{\mathrm{I}})$

using (2.5), and observe that the market-clearing price $\tilde{\pi}$ from (2.7) implies:

$$Y = \frac{\mathrm{q} - x_{\mathrm{E}}}{1-\alpha}, \quad \text{and} \quad Y = \frac{\tilde{\pi}(K - x_{\mathrm{I}})}{\alpha}. \tag{A.7.12}$$

It follows that both resource constraints in (A.7.9) are satisfied given incomes consistent with (A.7.11), individual demands (2.4), and the market-clearing price (2.7). This means that markets can deliver exactly the same consumption allocation as the equilibrium allocation given some individual disposable incomes that maximize incumbent payoffs subject to the no-rebellion constraints. The equilibrium allocation effectively allows such markets to be established and does not interfere with their operation, confirming the claim in the proposition. The aggregate budget constraint (A.7.11) can also be expressed in terms of consumption using (2.8) and (2.9)

$$pC_{\mathrm{p}} + kC_{\mathrm{k}} + (1-p-k)C_{\mathrm{w}} = C, \tag{A.7.13}$$

where C is aggregate real GDP from (2.9). This confirms part (i) of the proposition.

Free international trade

Conditional on p and s (and hence K), the utilities (A.7.1) and (A.7.2) under the current allocation depend only on the consumption levels in (A.7.13), which stands in for the resource constraints (A.7.9) given (2.9). Maximizing C_{p} subject to the no-rebellion constraints (A.7.7) therefore requires net exports x_{E} and x_{I} are chosen to maximize C in (2.9). It is shown in section 2.4 that this is equivalent to free trade, confirming part (ii) of the proposition.

The equilibrium allocation following a post-investment rebellion

If there were a rebellion at the post-investment stage, the following result characterizes the subsequent equilibrium level of power sharing $p^{\ddagger}$ and the incumbent payoff $C_{\mathrm{p}}^{\ddagger}$.

Lemma 1 *An equilibrium allocation established at the post-investment stage of the sequence of events in Figure 4 has the following features:*

(i) There is free exchange of goods domestically and internationally subject to given levels of disposable income.

(ii) Full expropriation of capital: The income levels of all non-incumbents are equalized irrespective of whether individuals have produced capital, that is, $\mathcal{D} = \emptyset$ (and hence $\mathcal{K} = \emptyset$), or equivalently, $Y_{\mathrm{k}} = Y_{\mathrm{w}}$.

(iii) The equilibrium level of power sharing and the payoffs of incumbents and workers are:

$$p^{\ddagger} = \frac{1}{2+\delta}, \quad C_{\mathrm{p}}^{\ddagger} = \frac{2+\delta}{2}C, \quad \text{and} \quad C_{\mathrm{w}}^{\ddagger} = \frac{2+\delta}{2(1+\delta)}C, \tag{A.7.14}$$

where C is real GDP from (2.9).

Proof See appendix A.10. ■

This result confirms the full expropriation of capital that would follow such a rebellion, as claimed in part (iv) of the proposition. In what follows, the case of $s > 0$ is considered. If $s = 0$, it turns out that the equilibrium allocation coincides with that characterized in Lemma 1.

The no-rebellion constraints (A.7.7) can be written in terms of consumption using the expressions for the continuation payoffs in (A.7.1) and (A.7.2):

$$\zeta_{\mathrm{n}}p' \max\left\{\frac{C'_{\mathrm{p}}}{C_{\mathrm{w}}} - 1, 0\right\} + \zeta_{\mathrm{d}}p' \max\left\{\frac{C'_{\mathrm{p}}}{C_{\mathrm{w}}^{1-\nu}\left(\frac{C_{\mathrm{k}}}{1+\theta}\right)^{\nu}} - 1, 0\right\} + \zeta_{\mathrm{p}}p'\mathbb{1}[C_{\mathrm{p}} \leq C'_{\mathrm{p}}]\left(\frac{C'_{\mathrm{p}}}{C_{\mathrm{p}}} - 1 + \delta\right) \leq \delta p; \tag{A.7.15a}$$

$$\zeta_{\mathrm{w}}p^{\ddagger} \max\left\{\frac{C_{\mathrm{p}}^{\ddagger}}{C_{\mathrm{w}}} - 1, 0\right\} + \zeta_{\mathrm{k}}p^{\ddagger} \max\left\{\frac{C_{\mathrm{p}}^{\ddagger}}{C_{\mathrm{k}}} - 1, 0\right\} + \zeta_{\mathrm{p}}p^{\ddagger}\mathbb{1}[C_{\mathrm{p}} \leq C_{\mathrm{p}}^{\ddagger}]\left(\frac{C_{\mathrm{p}}^{\ddagger}}{C_{\mathrm{p}}} - 1 + \delta\right) \leq \delta p, \tag{A.7.15b}$$

which must hold for all valid ζ_{n}, ζ_{d}, and ζ_{p}, and for all valid ζ_{w}, ζ_{k}, and ζ_{p} at the pre- and post-investment stages.

The time inconsistency problem

The equilibrium conditions require that the allocation maximizes the incumbent payoff U_{p}. This is subject to the resource constraint (A.7.13), the incentive constraint (4.2), and the no-rebellion constraints (A.7.15). Given s, and hence K, the equilibrium allocation established following a rebellion at the post-investment stage would maximize the incumbent payoff C_{p} subject only to the resource constraint (A.7.13) and the post-investment no-rebellion constraints (A.7.15b). The solution of that constrained maximization problem is characterized in Lemma 1. In that problem, it is optimal to give all non-incumbents the same level of consumption. Here, the equilibrium allocation must also satisfy the incentive constraint (4.2), which requires granting a positive measure μs of non-incumbents a consumption level strictly greater than workers (since $\theta > 0$). Given that the resource constraint is the same in both cases, satisfying the incentive constraint strictly reduces the payoff of incumbents. Furthermore, the pre-investment no-rebellion constraints (A.7.15a) must also be satisfied. With these additional constraints, the incumbent payoff C_{p} must be strictly lower than the constrained-maximum payoff $C_{\mathrm{p}}^{\ddagger}$ following a rebellion at the post-investment stage.

Since $C_{\mathrm{p}}^{\ddagger} > C_{\mathrm{p}}$, members of the incumbent group have an incentive to rebel against an allocation after investment has occurred, even though the allocation was established to maximize their payoff at the pre-investment stage. An allocation that gives credible incentives for investors is therefore subject to a time-inconsistency problem.

There must be more power sharing

Suppose that the allocation specifies power sharing p no more than the equilibrium level of power sharing $p^{\ddagger}$ following a rebellion at the post-investment stage. With $p \leq p^{\ddagger}$, it is feasible to have a post-investment rebellion with $\zeta_{\mathrm{p}} = p/p^{\ddagger}$ and values of ζ_{k}, and ζ_{w} such that $\zeta_{\mathrm{p}} + \zeta_{\mathrm{k}} + \zeta_{\mathrm{w}} = 1$ in (A.7.15b). Since $\zeta_p p^{\ddagger} = p$ and $C_{\mathrm{p}} < C_{\mathrm{p}}^{\ddagger}$, the left-hand side is strictly greater than the right-hand side, violating the no-rebellion constraint. The equilibrium allocation must therefore feature $p > p^{\ddagger}$. This means an increase in power sharing relative to what would be optimal ex post for incumbents.

No capitalists in a rebel faction

The natural limit on the fraction of workers ζ_{w} included in the incumbent group following a rebellion at the post-investment stage is $\zeta_{\mathrm{w}} \leq (1-p-k)/p^{\ddagger}$. Since $p \leq 1/2$ and $k \leq \mu$, a sufficient condition for $\zeta_{\mathrm{w}} = 1$ to be feasible is:

$$1 - \frac{1}{2} - \mu \geq p^{\ddagger}, \quad \text{or equivalently} \quad \mu \leq \frac{1}{2} - \frac{1}{2+\delta} = \frac{\delta}{2(2+\delta)},$$

which uses the expression for $p^{\ddagger}$ in (A.7.14). This condition holds under the parameter restriction on μ in (4.5).

The incentive constraint (4.2) requires $C_{\mathrm{k}} > C_{\mathrm{w}}$, which implies $\max\{C_{\mathrm{p}}^{\ddagger}/C_{\mathrm{k}} - 1, 0\} \leq \max\{C_{\mathrm{p}}^{\ddagger}/C_{\mathrm{w}} - 1, 0\}$. Since $\zeta_{\mathrm{w}} = 1$ is feasible in the post-investment no-rebellion constraint (A.7.15b), it follows that ζ_{k} can be set to zero because if the no-rebellion constraint holds for all feasible values of ζ_{w} and ζ_{p} given $\zeta_{\mathrm{k}} = 0$, it must hold any positive value of ζ_{k} as well. The most dangerous rebel faction does not include capitalists because workers have a greater incentive to change the allocation, and there is no shortage of workers in filling the places in the incumbent group following a rebellion.

With $\zeta_{\mathrm{k}} = 0$ and $C_{\mathrm{p}} < C_{\mathrm{p}}^{\ddagger}$, the set of relevant post-investment no-rebellion constraints can be written as follows:

$$(1-\zeta)p^{\ddagger} \max\left\{\frac{C_{\mathrm{p}}^{\ddagger}}{C_{\mathrm{w}}} - 1, 0\right\} + \zeta p^{\ddagger}\left(\frac{C_{\mathrm{p}}^{\ddagger}}{C_{\mathrm{p}}} - 1 + \delta\right) \leq \delta p, \qquad \text{[A.7.16]}$$

which must hold for all $\zeta \in [0,1]$ (where $\zeta = \zeta_{\mathrm{p}}$, and all values of ζ in the unit interval are feasible, which follows from $p > p^{\ddagger}$ and $1 - p - k \geq p^{\ddagger}$).

A no-rebellion constraint must bind for a rebel faction comprising a positive measure of non-incumbents

Given the equilibrium conditions (Definition 2), the equilibrium allocation will feature $U_{\mathrm{p}} = U'_{\mathrm{p}}$, $U_{\mathrm{d}} = U'_{\mathrm{d}}$, and $U_{\mathrm{n}} = U'_{\mathrm{n}}$, and if it is in the interests of incumbents who lose power at the pre-investment stage to defend the allocation ($U_{\mathrm{p}} > U'_{\mathrm{n}}$) then $U'_{\mathrm{p}} > U_{\mathrm{n}}$, and hence $C'_{\mathrm{p}} > C_{\mathrm{w}}$ (non-incumbents would be willing to exert some positive rebellion effort in a rebel faction). Since the incentive constraint (4.2) requires $C_{\mathrm{k}}/(1+\theta) \geq C_{\mathrm{w}}$, and as C_{k} does not appear in the post-investment no-rebellion constraints (A.7.16), it follows that if $C_{\mathrm{w}}^{1-\nu}(C_{\mathrm{k}}/(1+\theta))^{\nu} > C'_{\mathrm{p}}$ then C_{k} can be reduced by some positive amount without violating any constraint. This would allow the incumbent payoff to be increased, so an optimal allocation must feature $C_{\mathrm{w}}^{1-\nu}(C_{\mathrm{k}}/(1+\theta))^{\nu} \leq C'_{\mathrm{p}}$. It follows that the pre-investment no-rebellion constraints can be stated as:

$$(1-\zeta)p'\left(\left(\frac{1-\varkappa}{C_{\mathrm{w}}} + \frac{\varkappa}{C_{\mathrm{w}}^{1-\nu}\left(\frac{C_{\mathrm{k}}}{1+\theta}\right)^{\nu}}\right)C'_{\mathrm{p}} - 1\right) + \zeta p' \mathbb{1}[C_{\mathrm{p}} \leq C'_{\mathrm{p}}]\left(\frac{C'_{\mathrm{p}}}{C_{\mathrm{p}}} - 1 + \delta\right) \leq \delta p, \qquad \text{[A.7.17a]}$$

where the notation $\zeta = \zeta_{\mathrm{p}}$ and $\varkappa = \zeta_{\mathrm{d}}/(\zeta_{\mathrm{n}} + \zeta_{\mathrm{d}})$ is used, and the constraints must hold for all feasible values of ζ and $\varkappa$. Since $C_{\mathrm{w}} \leq C_{\mathrm{p}} < C^{\ddagger}_{\mathrm{p}}$, the post-investment no-rebellion constraints (A.7.16) can also be written in a simpler form:

$$(1-\zeta)p^{\ddagger}\left(\frac{C^{\ddagger}_{\mathrm{p}}}{C_{\mathrm{w}}} - 1\right) + \zeta p^{\ddagger}\left(\frac{C^{\ddagger}_{\mathrm{p}}}{C_{\mathrm{p}}} - 1 + \delta\right) \leq \delta p, \qquad \text{[A.7.17b]}$$

which must hold for all $\zeta \in [0,1]$. One of the constraints (A.7.17a) or (A.7.17b) must bind for some $\zeta < 1$. Otherwise the terms multiplying $1-\zeta$ in these constraints are both too low, allowing C_{w} to be reduced by some positive amount without violating any no-rebellion constraint (reducing C_{w} does not violate the incentive constraint 4.2). This would allow the payoff of incumbents to be increased.

The incentive compatibility constraint is binding

Now suppose the incentive compatibility constraint (4.2) is slack, that is, $C_{\mathrm{k}} > (1+\theta)C_{\mathrm{w}}$. Consider a redistribution of contingent consumption between workers and capitalists such that workers will receive C_{b} and capitalists will receive $(1+\theta)C_{\mathrm{b}}$, which means that the incentive constraint becomes binding. Since there will be k capitalists and $1-p-k$ workers, the resource constraint requires:

$$C_{\mathrm{b}} = \frac{(1-p-k)C_{\mathrm{w}} + kC_{\mathrm{k}}}{(1-p-k) + k(1+\theta)} = \left(\frac{1-p-k}{1-p+k\theta}\right)C_{\mathrm{w}} + \left(\frac{k(1+\theta)}{1-p+k\theta}\right)\frac{C_{\mathrm{k}}}{1+\theta}. \qquad \text{[A.7.18]}$$

This shows that C_{b} is a weighted average of C_{w} and $C_{\mathrm{k}}/(1+\theta)$, where $C_{\mathrm{k}}/(1+\theta) > C_{\mathrm{w}}$. Using Jensen's inequality, the convexity of the function $1/C$ implies:

$$\frac{1}{C_{\mathrm{b}}} < \left(\frac{1-p-k}{1-p+k\theta}\right)\frac{1}{C_{\mathrm{w}}} + \left(\frac{k(1+\theta)}{1-p+k\theta}\right)\frac{1+\theta}{C_{\mathrm{k}}}. \qquad \text{[A.7.19]}$$

Since $1/C_{\mathrm{w}} > (1+\theta)/C_{\mathrm{k}}$ and $(1-p-k)/(1-p+k\theta) < (1-p-k)/(1-p)$, it follows that:

$$\left(\frac{1-p-k}{1-p+k\theta}\right)\frac{1}{C_{\mathrm{w}}} + \left(\frac{k(1+\theta)}{1-p+k\theta}\right)\frac{1+\theta}{C_{\mathrm{k}}} < \left(\frac{1-p-k}{1-p}\right)\frac{1}{C_{\mathrm{w}}} + \left(\frac{k}{1-p}\right)\frac{1+\theta}{C_{\mathrm{k}}}. \qquad \text{[A.7.20]}$$

Finally, by putting together (A.7.19) and (A.7.20) and noting $C_{\mathrm{w}}^{1-\nu}(C_{\mathrm{k}}/(1+\theta))^{\nu} < C_{\mathrm{k}}/(1+\theta)$:

$$\frac{1}{C_{\mathrm{b}}} < \frac{1-\left(\frac{k}{1-p}\right)}{C_{\mathrm{w}}} + \frac{\left(\frac{k}{1-p}\right)}{C_{\mathrm{w}}^{1-\nu}\left(\frac{C_{\mathrm{k}}}{1+\theta}\right)^{\nu}}. \qquad \text{[A.7.21]}$$

Observe that with $C_{\mathrm{w}} < C_{\mathrm{k}}/(1+\theta)$, the value of $\varkappa$ associated with the highest level of the left-hand side of (A.7.17a) (given a ζ) is the lowest feasible value of $\varkappa$. This value of $\varkappa$ must be such that $\varkappa \leq k/(1-p)$ since there are k capitalists and $1-p$ non-incumbents. By using (A.7.21), it must be the case that:

$$\frac{1}{C_{\mathrm{b}}} < \frac{1-\varkappa}{C_{\mathrm{w}}} + \frac{\varkappa}{C_{\mathrm{w}}^{1-\nu}\left(\frac{C_{\mathrm{k}}}{1+\theta}\right)^{\nu}},$$

and it follows that the redistribution specified in (A.7.18) slackens the pre-investment no-rebellion constraint (A.7.17a) if this is binding for some $\zeta < 1$. This would allow the payoff of incumbents to be increased. If (A.7.17a) is not binding for any $\zeta < 1$ then C_k does not appear in any relevant no-rebellion constraint (it is absent from A.7.17b), and so C_k can be reduced until the incentive constraint (4.2) is binding. Therefore, in either case, the equilibrium allocation must necessarily have:

$$C_\mathrm{k} = (1+\theta)C_\mathrm{w}. \tag{A.7.22}$$

This confirms claim (iii) in the proposition.

There are only two independent no-rebellion constraints at the post-investment stage

With $C_\mathrm{k}/(1+\theta) = C_\mathrm{w}$, the pre-investment no-rebellion constraints simplify to:

$$(1-\zeta)p'\left(\frac{C'_\mathrm{p}}{C_\mathrm{w}} - 1\right) + \zeta p' \mathbb{1}[C_\mathrm{p} \leq C'_\mathrm{p}]\left(\frac{C'_\mathrm{p}}{C_\mathrm{p}} - 1 + \delta\right) \leq \delta p, \tag{A.7.23a}$$

which must hold for all feasible $\zeta \in [0, \min\{p/p', 1\}]$ (note that $\zeta = 0$ is feasible because $p \leq 1/2$ and $p' \leq 1/2$). The set of post-investment no-rebellion constraints (A.7.17b) is linear in ζ and must be verified for all ζ in the fixed interval $[0, 1]$ (all of which are feasible). If (A.7.17b) holds for all $\zeta \in [0, 1]$ then it must hold in particular at $\zeta = 0$ and $\zeta = 1$:

$$p^\ddagger\left(\frac{C^\ddagger_\mathrm{p}}{C_\mathrm{w}} - 1\right) \leq \delta p; \tag{A.7.23b}$$

$$p^\ddagger\left(\frac{C^\ddagger_\mathrm{p}}{C_\mathrm{p}} - 1 + \delta\right) \leq \delta p. \tag{A.7.23c}$$

Furthermore, if both of the individual no-rebellion constraints (A.7.23b) and (A.7.23c) hold then taking a linear combination with weights $1-\zeta$ and ζ implies that (A.7.17b) holds for any $\zeta \in [0, 1]$. The relevant set of no-rebellion constraints has therefore been reduced to (A.7.23a), (A.7.23b) and (A.7.23c).

Only a single no-rebellion constraint involving non-incumbents is relevant at the pre-investment stage

Consider a simpler constrained maximization problem for the incumbent payoff $U_\mathrm{p} = \log C_\mathrm{p}$ where the pre-investment no-rebellion constraint is only required to hold for $\zeta = 0$:

$$p'\left(\frac{C'_\mathrm{p}}{C_\mathrm{w}} - 1\right) \leq \delta p. \tag{A.7.24}$$

All other relevant no-rebellion constraints must hold (A.7.23b and A.7.23c), as must the resource constraint (A.7.13) and the (binding) incentive constraint (A.7.22). Equivalents of the equilibrium conditions in Definition 2 are required to hold for this simpler problem.

Let $\{\dot{p}, \dot{C}_\mathrm{p}, \dot{C}_\mathrm{w}\}$ denote an equilibrium of this simpler problem. Since there are no changes to fundamental state variables following any pre-investment rebellion, independence of irrelevant history means $p' = \dot{p}$ and $C'_\mathrm{p} = \dot{C}_\mathrm{p}$. Since (A.7.24) must hold, it follows that $\dot{C}_\mathrm{p}/\dot{C}_\mathrm{w} \leq 1 + \delta$.

Now consider the original constrained maximization problem, taking $p' = \dot{p}$ and $C'_\mathrm{p} = \dot{C}_\mathrm{p}$ for an equilibrium of the simpler problem. When $p = \dot{p}$, any value of ζ between zero and one is feasible, and an allocation must therefore satisfy the no-rebellion constraint (A.7.23a) for all $\zeta \in [0, 1]$. If $p = \dot{p}$, $C_\mathrm{w} = \dot{C}_\mathrm{w}$, and $C_\mathrm{p} = \dot{C}_\mathrm{p}$ are to satisfy these constraints then the following must hold:

$$(1-\zeta)\dot{p}\left(\frac{\dot{C}_\mathrm{p}}{\dot{C}_\mathrm{w}} - 1\right) + \delta\zeta\dot{p} \leq \delta\dot{p},$$

which does indeed follow for any $\zeta \in [0, 1]$ given that $\dot{C}_\mathrm{p}/\dot{C}_\mathrm{w} \leq 1 + \delta$. Since the allocation satisfies this and all other no-rebellion constraints (as it is an equilibrium of the simpler problem), and all other resource and incentive constraints, and because it maximizes the incumbent payoff subject to the weaker no-rebellion constraint (A.7.24), it follows that it is optimal subject to the full set of no-rebellion constraints (A.7.23a). The allocation is thus an equilibrium of the original problem.

Now consider the converse. Take an allocation with $\{\dot{p}, \dot{C}_{\mathrm{p}}, \dot{C}_{\mathrm{w}}\}$ that is an equilibrium of the original problem. When $p' = \dot{p}$ and $C'_{\mathrm{p}} = \dot{C}_{\mathrm{p}}$, this allocation clearly satisfies the simpler no-rebellion constraint (A.7.24) (a special case of A.7.23a for $\zeta = 0$, which is always feasible). But now suppose that it is not optimal subject only to (A.7.24) instead of (A.7.23a) (with $p' = \dot{p}$ and $C'_{\mathrm{p}} = \dot{C}_{\mathrm{p}}$) for all feasible ζ values. This means there is an alternative allocation with $\{p, C_{\mathrm{p}}, C_{\mathrm{w}}\}$ satisfying (A.7.24) (with $p' = \dot{p}$ and $C'_{\mathrm{p}} = \dot{C}_{\mathrm{p}}$) that yields a higher incumbent consumption level $C_{\mathrm{p}} > \dot{C}_{\mathrm{p}}$. Now take any $\zeta \in [0, 1]$ and observe that:

$$(1-\zeta)\dot{p}\left(\frac{\dot{C}_{\mathrm{p}}}{C_{\mathrm{w}}} - 1\right) + \zeta\dot{p}\mathbb{1}[C_{\mathrm{p}} \le \dot{C}_{\mathrm{p}}]\left(\frac{\dot{C}_{\mathrm{p}}}{C_{\mathrm{p}}} - 1 + \delta\right) = (1-\zeta)\dot{p}\left(\frac{\dot{C}_{\mathrm{p}}}{C_{\mathrm{w}}} - 1\right) \le (1-\zeta)\delta p \le \delta p,$$

where the first equality follows from $C_{\mathrm{p}} > \dot{C}_{\mathrm{p}}$, the first inequality follows from the alternative allocation satisfying (A.7.24) (with $p' = \dot{p}$ and $C'_{\mathrm{p}} = \dot{C}_{\mathrm{p}}$), and the second inequality follows from $0 \le \zeta \le 1$. This shows that the alternative allocation satisfies (A.7.23a) for all feasible ζ (a subset of $[0, 1]$), and thus satisfies all the constraints of the original problem, yet yields a higher payoff than the equilibrium allocation, contradicting the optimality requirement of the allocation with $\{\dot{p}, \dot{C}_{\mathrm{p}}, \dot{C}_{\mathrm{w}}\}$. This contradicts the optimality of the allocation when the constraints (A.7.23a) are weakened to (A.7.24) (and other constraints are the same). The allocation is thus an equilibrium of the simpler problem.

This logic shows that the set of equilibrium allocations is not affected by imposing only (A.7.24) instead of (A.7.23a). Therefore, the only relevant no-rebellion constraints are (A.7.23b), (A.7.23c), and (A.7.24).

The resource constraint (A.7.13) implies the following expression for incumbent consumption C_{p} in terms of the total resources C available for consumption, and the consumption levels C_{w} and C_{k} of workers and capitalists:

$$C_{\mathrm{p}} = \frac{C - (1 - p - \mu s)C_{\mathrm{w}} - \mu s C_{\mathrm{k}}}{p},$$

where C is given in (A.7.13), and the measure of capitalists is $k = \mu s$. Eliminating C_{k} using the binding incentive compatibility constraint (A.7.22):

$$C_{\mathrm{p}} = \frac{C - (1 - p + \mu\theta s)C_{\mathrm{w}}}{p}.$$

Let $\psi_{\mathrm{p}} \equiv C_{\mathrm{p}}/C$ and $\psi_{\mathrm{w}} \equiv C_{\mathrm{w}}/C$ denote the per-person consumption shares of incumbents and workers of total resources C. Dividing both sides of the expression for C_{p} by C:

$$\psi_{\mathrm{p}} = \frac{1 - (1 - p + \mu\theta s)\psi_{\mathrm{w}}}{p}. \tag{A.7.25}$$

The value of s is taken as given at this stage, which means the capital stock K and total resources C are taken as given for now (see 2.9 and 2.10). The optimality of free exchange and free international trade has already been shown, which means that the price $\tilde{\pi}$ is also taken as given. The optimality condition (Definition 2) then requires that ψ_{p} (equation A.7.25) be maximized subject to the no-rebellion constraints. All other constraints have already been accounted for in (A.7.25). The no-rebellion constraints are (A.7.24), (A.7.23b), and (A.7.23c), which can be written as follows:

$$C_{\mathrm{w}} \ge \frac{C'_{\mathrm{p}}}{1 + \delta\frac{p}{p'}}, \quad C_{\mathrm{w}} \ge \frac{C^{\ddagger}_{\mathrm{p}}}{1 + \delta\frac{p}{p^{\ddagger}}}, \quad \text{and } C_{\mathrm{p}} \ge \frac{C^{\ddagger}_{\mathrm{p}}}{1 + \delta\frac{p - p^{\ddagger}}{p^{\ddagger}}}.$$

The capital stock K is predetermined at the post-investment stage, and it has been shown here and in Lemma 1 that free exchange domestically and free international trade are necessary features of any equilibrium allocation. In equilibrium, independence of irrelevant history requires $s = s'$ (Definition 2), which means that there would be no change to the total amount of resources C available for consumption after any rebellion ($C = C' = C^{\ddagger}$). Dividing both sides of all the no-rebellion constraints above by this common level of total output leads to the following set of no-rebellion constraints written in terms of ψ'_{p}

and $\psi_{\mathrm{p}}^{\ddagger}$:

$$\psi_{\mathrm{w}} \geq \frac{\psi_{\mathrm{p}}'}{1+\delta\frac{p}{p'}}; \tag{A.7.26a}$$

$$\psi_{\mathrm{w}} \geq \frac{\psi_{\mathrm{p}}^{\ddagger}}{1+\delta\frac{p}{p^{\ddagger}}}; \tag{A.7.26b}$$

$$\psi_{\mathrm{p}} \geq \frac{\psi_{\mathrm{p}}^{\ddagger}}{1+\delta\frac{p-p^{\ddagger}}{p^{\ddagger}}}. \tag{A.7.26c}$$

The values of $p^{\ddagger}$, $\psi_{\mathrm{p}}^{\ddagger}$, and $\psi_{\mathrm{w}}^{\ddagger}$ for any equilibrium allocation established at the post-investment are obtained from (A.7.14):

$$p^{\ddagger} = \frac{1}{2+\delta}, \quad \psi_{\mathrm{p}}^{\ddagger} = \frac{2+\delta}{2}, \quad \text{and} \quad \psi_{\mathrm{w}}^{\ddagger} = \frac{2+\delta}{2(1+\delta)}. \tag{A.7.27}$$

The values of p', ψ_{p}', and ψ_{w}' must satisfy independence of irrelevant history in equilibrium: $p'=p$, $\psi_{\mathrm{p}}'=\psi_{\mathrm{p}}$, and $\psi_{\mathrm{w}}'=\psi_{\mathrm{w}}$. There is also a size constraint on the incumbent group, namely $0 \leq p \leq 1/2$, and it remains to be confirmed that incumbents who would lose power in a rebellion are willing to belong to the loyal faction, that is, $\psi_{\mathrm{p}} > \psi_{\mathrm{w}}'$ and $\psi_{\mathrm{p}} > \psi_{\mathrm{w}}^{\ddagger}$.

It has already been established that one of (A.7.26a) or (A.7.26b) must be binding, that $\psi_{\mathrm{p}} < \psi_{\mathrm{p}}^{\ddagger}$, and that $p > p^{\ddagger}$ for any equilibrium allocation.

The pre-investment no-rebellion constraint for workers cannot be the only binding constraint

Suppose for contradiction that (A.7.26a) is the only binding no-rebellion constraint, where ψ_{p}' and p' are taken as given. Substituting the binding constraint (A.7.26a) into (A.7.25), the optimality condition requires that p maximize ψ_{p} (with s taken as given):

$$\psi_{\mathrm{p}} = \frac{1-(1-p+\mu\theta s)\frac{\psi_{\mathrm{p}}'}{1+\delta\frac{p}{p'}}}{p}. \tag{A.7.28}$$

The first derivative with respect to p is:

$$\frac{\partial\psi_{\mathrm{p}}}{\partial p} = \frac{1}{p}\left(\left(1+\frac{\frac{\delta(1-p+\mu\theta s)}{p'}}{1+\delta\frac{p}{p'}}\right)\frac{\psi_{\mathrm{p}}'}{1+\delta\frac{p}{p'}} - \frac{1-(1-p+\mu\theta s)\frac{\psi_{\mathrm{p}}'}{1+\delta\frac{p}{p'}}}{p}\right), \tag{A.7.29}$$

and the second derivative evaluated at a point where the first derivative is zero is:

$$\left.\frac{\partial^2\psi_{\mathrm{p}}}{\partial p^2}\right|_{\frac{\partial\psi_{\mathrm{p}}}{\partial p}=0} = -\frac{2\delta}{pp'}\left(1+\frac{\frac{\delta(1-p+\mu\theta s)}{p'}}{1+\delta\frac{p}{p'}}\right)\frac{\psi_{\mathrm{p}}'}{\left(1+\delta\frac{p}{p'}\right)^2}.$$

This is unambiguously negative, which demonstrates that ψ_{p} is a quasi-concave function of p. Using (A.7.29), the optimal value of p therefore satisfies (assuming $0 \leq p \leq 1/2$):

$$\frac{1-(1-p+\mu\theta s)\frac{\psi_{\mathrm{p}}'}{1+\delta\frac{p}{p'}}}{p} = \frac{\psi_{\mathrm{p}}'}{1+\delta\frac{p}{p'}}\left(1+\frac{(1-p+\mu\theta s)\frac{\delta}{p'}}{1+\delta\frac{p}{p'}}\right).$$

The equilibrium conditions $p'=p$ and $\psi_{\mathrm{p}}'=\psi_{\mathrm{p}}$ must hold, so this equation reduces to:

$$\left(\frac{\delta(1-p+\mu\theta s)}{(1+\delta)^2 p} + \frac{1}{1+\delta} - 1\right)\psi_{\mathrm{p}} = 0,$$

and since $\psi_{\mathrm{p}} > 0$, the equation further simplifies to:

$$\frac{\delta(1-p+\mu\theta s)}{(1+\delta)p} = \delta.$$

After multiplying both sides by $p > 0$, this is a linear equation in p and can be solved as follows:

$$p = \frac{1 + \mu\theta s}{2 + \delta}, \quad \text{and hence} \;\; p - p^{\ddagger} = \frac{\mu\theta s}{2 + \delta}, \tag{A.7.30}$$

where the second equation uses the expression for $p^{\ddagger}$ in (A.7.14). Observe that $p > p^{\ddagger}$ for any $s > 0$. It is assumed here that $p \leq 1/2$; the case where the constraint $p \leq 1/2$ might bind is considered later.

Using the equilibrium value of p in (A.7.30) and $p = p'$ and $\psi_{\mathrm{p}} = \psi'_{\mathrm{p}}$, the implied equilibrium income share ψ_{p} can be obtained from (A.7.28):

$$p\psi_{\mathrm{p}} = 1 - \left(1 - p + (2 + \delta)(p - p^{\ddagger})\right) \frac{\psi_{\mathrm{p}}}{1 + \delta},$$

which can be simplified by noting $p^{\ddagger} = 1/(2 + \delta)$:

$$(1 + \delta)p\psi_{\mathrm{p}} = (1 + \delta) - (1 + \delta)p\psi_{\mathrm{p}}.$$

The per-person income shares of incumbents and workers are therefore:

$$\psi_{\mathrm{p}} = \frac{1}{2p}, \quad \text{and} \;\; \psi_{\mathrm{w}} = \frac{1}{2(1 + \delta)p}, \tag{A.7.31}$$

where the latter follows from the former together with the binding no-rebellion constraint (A.7.26a), which implies $\psi_{\mathrm{w}} = \psi_{\mathrm{p}}/(1 + \delta)$ in equilibrium.

For this case to be an equilibrium it is necessary that the other no-rebellion constraints (A.7.26b) and (A.7.26c) are satisfied. By using (A.7.31), the constraint (A.7.26b) requires:

$$\frac{1}{2(1 + \delta)p} \geq \frac{1}{2(\delta p + p^{\ddagger})},$$

where the right-hand side follows from the expressions for $p^{\ddagger}$ and $\psi^{\ddagger}_{\mathrm{p}}$ (see A.7.27) by noting $p^{\ddagger}\psi^{\ddagger}_{\mathrm{p}} = 1/2$. Rearranging the inequality above reveals that it is equivalent to $p^{\ddagger} \geq p$. However, this is violated for any $s > 0$ because that would imply $p > p^{\ddagger}$ according to (A.7.30).

Now consider the case where the no-rebellion constraint (A.7.26a) binds in conjunction with the constraint $p \leq 1/2$ on power sharing. Since ψ_{p} in (A.7.28) is quasi-concave in p, the constraint $p \leq 1/2$ binds when the p value where the first derivative of ψ_{p} is zero is found to the right of $1/2$. Using (A.7.30), this occurs when

$$\frac{1}{2} < p^{\ddagger} + \frac{\mu\theta s}{2 + \delta} = \frac{1 + \mu\theta s}{2 + \delta}, \quad \text{or equivalently} \;\; \mu\theta s > \frac{\delta}{2}. \tag{A.7.32}$$

With $p = p' = 1/2$, the expression for ψ_{p} in (A.7.28) can be used to deduce the following in equilibrium ($\psi_{\mathrm{p}} = \psi'_{\mathrm{p}}$):

$$\frac{\psi_{\mathrm{p}}}{2} = 1 - \left(1 - \frac{1}{2} + \mu\theta s\right) \frac{\psi_{\mathrm{p}}}{1 + \delta},$$

and solving this linear equation in ψ_{p} and using $\psi_{\mathrm{w}} = \psi_{\mathrm{p}}/(1 + \delta)$ (when A.7.26a binds):

$$\psi_{\mathrm{p}} = \frac{1 + \delta}{1 + \delta/2 + \mu\theta s}, \quad \text{and} \;\; \psi_{\mathrm{w}} = \frac{1}{1 + \delta/2 + \mu\theta s}. \tag{A.7.33}$$

With $p = 1/2$, $p^{\ddagger} = 1/(2 + \delta)$, and $p^{\ddagger}\psi^{\ddagger}_{\mathrm{p}} = 1/2$, the post-investment no-rebellion constraint for workers (A.7.26b) requires:

$$\psi_{\mathrm{w}} \geq \frac{1}{2(\delta/2 + 1/(2 + \delta))} = \frac{1}{\delta + 1/(1 + \delta/2)}. \tag{A.7.34}$$

However, under the conditions shown in (A.7.32) when $p \leq 1/2$ is binding, it follows from the expression for ψ_{w} in (A.7.33) that:

$$\psi_{\mathrm{w}} < \frac{1}{1 + \delta} < \frac{1}{\delta + 1/(1 + \delta/2)},$$

which contradicts (A.7.34), so the no-rebellion constraint (A.7.26b) is violated when both (A.7.26a) and

$p \leq 1/2$ are binding. Hence, (A.7.26a) cannot be the only binding no-rebellion constraint, irrespective of whether $0 \leq p \leq 1/2$ is also binding or not.

The no-rebellion constraints for workers at the pre-investment stage and incumbents at the post-investment stage cannot bind together

Suppose no-rebellion constraints (A.7.26a) and (A.7.26c) are both binding. In equilibrium, with $p = p'$ and $\psi_{\mathrm{p}} = \psi'_{\mathrm{p}}$, having (A.7.26a) bind requires:

$$\psi_{\mathrm{w}} = \frac{\psi_{\mathrm{p}}}{1+\delta}, \tag{A.7.35}$$

and having (A.7.26c) bind requires:

$$\psi_{\mathrm{p}} = \frac{1}{2(\delta p + (1-\delta)p^{\ddagger})},$$

which is obtained using $p^{\ddagger}\psi_{\mathrm{p}}^{\ddagger} = 1/2$. Using this equation to substitute for ψ_{p} in (A.7.35) yields an expression for ψ_{w}:

$$\psi_{\mathrm{w}} = \frac{1}{2(1+\delta)(\delta p + (1-\delta)p^{\ddagger})}.$$

The no-rebellion constraint (A.7.26b) must also be satisfied. Given the expression for ψ_{w} above, and noting $p^{\ddagger}\psi_{\mathrm{p}}^{\ddagger} = 1/2$, this requires:

$$\frac{1}{2(1+\delta)(\delta p + (1-\delta)p^{\ddagger})} \geq \frac{1}{2(\delta p + p^{\ddagger})}.$$

Rearranging this inequality shows that it is equivalent to $\delta^2(p^{\ddagger} - p) \geq 0$. However, since $p > p^{\ddagger}$ is known to be required in any equilibrium with $s > 0$, the no-rebellion constraint (A.7.26b) cannot hold. Therefore, this configuration of binding no-rebelllion constraints is not possible in equilibrium.

The case where the no-rebellion constraint for workers at the post-investment stage is the only binding constraint

Suppose that (A.7.26b) is the only binding no-rebellion constraint. The per-person worker share ψ_{w} is thus given by:

$$\psi_{\mathrm{w}} = \frac{1}{2(\delta p + p^{\ddagger})}, \tag{A.7.36}$$

where this expression uses $p^{\ddagger}\psi_{\mathrm{p}}^{\ddagger} = 1/2$. Substituting the equation above into (A.7.25) yields an expression for the per-person incumbent share ψ_{p}:

$$\psi_{\mathrm{p}} = \frac{1 - \frac{(1-p+\mu\theta s)}{2(\delta p + p^{\ddagger})}}{p}, \tag{A.7.37}$$

and the optimality condition for an equilibrium allocation requires that p maximizes ψ_{p} (given a value of s). The first derivative of ψ_{p} with respect to p is:

$$\frac{\partial \psi_{\mathrm{p}}}{\partial p} = \frac{1}{p}\left(\left(1 + \frac{\delta(1-p+\mu\theta s)}{\delta p + p^{\ddagger}}\right)\frac{1}{2(\delta p + p^{\ddagger})} - \frac{1 - \frac{(1-p+\mu\theta s)}{2(\delta p + p^{\ddagger})}}{p}\right), \tag{A.7.38}$$

which can be used to find the second derivative and to evaluate it at a point where the first derivative is zero:

$$\left.\frac{\partial^2 \psi_{\mathrm{p}}}{\partial p^2}\right|_{\frac{\partial \psi_{\mathrm{p}}}{\partial p} = 0} = -\frac{\delta}{p(\delta p + p^{\ddagger})^2}\left(1 + \frac{\delta(1-p+\mu\theta s)}{\delta p + p^{\ddagger}}\right).$$

This is unambiguously negative, demonstrating that ψ_{p} is a quasi-concave function of p (given s), so the first-order condition is necessary and sufficient for a maximum. Setting the first derivative in (A.7.38) to zero and rearranging terms to have common denominators on both sides:

$$\frac{1}{2(\delta p + p^\ddagger)}\left(\frac{\delta p + p^\ddagger + \delta - \delta p + \delta\mu\theta s}{\delta p + p^\ddagger}\right) = \frac{(1+2\delta)p - (1-2p^\ddagger) - \mu\theta s}{2(\delta p + p^\ddagger)p}.$$

Cancelling common terms, using $1 - 2p^\ddagger = \delta p^\ddagger$, and simplifying leads to:

$$\frac{\delta + p^\ddagger + \delta\mu\theta s}{\delta p + p^\ddagger} = \frac{((1+2\delta)p - \delta p^\ddagger) - \mu\theta s}{p},$$

and therefore by multiplying both sides by $p(\delta p + p^\ddagger)$ and expanding the brackets:

$$(\delta + p^\ddagger + \delta\mu\theta s)p = \delta(1+2\delta)p^2 + (1+2\delta)p^\ddagger p - \delta^2 p^\ddagger p - \delta {p^\ddagger}^2 - (\delta p + p^\ddagger)\mu\theta s.$$

Collecting terms in $\mu\theta s$ on the left-hand side and simplifying:

$$(2\delta p + p^\ddagger)\mu\theta s = \delta(1+2\delta)p^2 + (1 + 2\delta - \delta^2 - 1 - \delta(2+\delta))p^\ddagger p - \delta {p^\ddagger}^2 = \delta((1+2\delta)p^2 - 2\delta p^\ddagger p - {p^\ddagger}^2),$$

where $\delta p = \delta(2+\delta)p^\ddagger p$ has been used. Note that the term on the right-hand side can be factorized:

$$(2\delta p + p^\ddagger)\mu\theta s = \delta(p - p^\ddagger)((1+2\delta)p + p^\ddagger),$$

and therefore the required relationship between s and p is given by:

$$s = \frac{\delta(p - p^\ddagger)}{\mu\theta}\left(1 + \frac{p}{2\delta p + p^\ddagger}\right). \qquad \text{[A.7.39]}$$

It can be seen that $s > 0$ is consistent with $p > p^\ddagger$. Under the parameter restrictions in (4.5), it will be shown below that the constraint $p \leq 1/2$ is satisfied for any $0 \leq s \leq 1$.

The next step is to derive an expression for the maximized per-person incumbent share ψ_p given the optimal p and s relationship in (A.7.39). Taking a common denominator of the expression for ψ_p in (A.7.37):

$$\psi_\text{p} = \frac{(1+2\delta)p - (1-2p^\ddagger) - \mu\theta s}{2(\delta p + p^\ddagger)p} = \frac{((1+2\delta)p - \delta p^\ddagger) - \mu\theta s}{2(\delta p + p^\ddagger)p},$$

which uses $1 - 2p^\ddagger = \delta p^\ddagger$. Substituting for $\mu\theta s$ using (A.7.39):

$$\psi_\text{p} = \frac{(1+2\delta)p - (1-2p^\ddagger) - \mu\theta s}{2(\delta p + p^\ddagger)p} = \frac{((1+2\delta)p - \delta p^\ddagger) - \frac{\delta(p-p^\ddagger)((1+2\delta)p+p^\ddagger)}{2\delta p + p^\ddagger}}{2(\delta p + p^\ddagger)p},$$

and rearranging to have a common denominator again:

$$\psi_\text{p} = \frac{((1+2\delta)p - \delta p^\ddagger)(2\delta p + p^\ddagger) - \delta(p - p^\ddagger)((1+2\delta)p + p^\ddagger)}{2(2\delta p + p^\ddagger)(\delta p + p^\ddagger)p}.$$

Multiplying out the brackets in the numerator leads to:

$$\psi_\text{p} = \frac{2\delta(1+2\delta)p^2 - 2\delta^2 p^\ddagger p + (1+2\delta)p^\ddagger p - \delta {p^\ddagger}^2 - \delta((1+2\delta)p^2 - (1+2\delta)p^\ddagger p + p^\ddagger p - {p^\ddagger}^2)}{2(2\delta p + p^\ddagger)(\delta p + p^\ddagger)p},$$

which can be simplified as follows:

$$\psi_\text{p} = \frac{\delta(1+2\delta)p^2 + (1+2\delta)p^\ddagger p}{2(2\delta p + p^\ddagger)(\delta p + p^\ddagger)p} = \frac{(1+2\delta)\left(\delta p + p^\ddagger\right)p}{2(2\delta p + p^\ddagger)(\delta p + p^\ddagger)p}.$$

Therefore, setting p to maximize ψ_p (given s) implies that:

$$\psi_\text{p} = \frac{1+2\delta}{2(2\delta p + p^\ddagger)}. \qquad \text{[A.7.40]}$$

The two other no-rebellion constraints must also be satisfied, namely (A.7.26a) for workers at the pre-investment stage, and (A.7.26c) for incumbents at the post-investment stage. Taking the first of these, in equilibrium with $p' = p$ and $\psi'_\text{p} = \psi_\text{p}$, (A.7.26a) requires $\psi_\text{w} \geq \psi_\text{p}/(1+\delta)$, and by using the expressions for

ψ_{w} and ψ_{p} in (A.7.36) and (A.7.40), this is equivalent to:

$$\frac{1}{2(\delta p+p^{\ddagger})}\geq\frac{1+2\delta}{2(1+\delta)(2\delta p+p^{\ddagger})}.$$

Rearranging the inequality above leads to $2\delta(1+\delta)p+(1+\delta)p^{\ddagger}\geq\delta(1+2\delta)p+(1+2\delta)p^{\ddagger}$, which simplifies to $\delta(p-p^{\ddagger})\geq 0$. This is satisfied because $p>p^{\ddagger}$ when $s>0$, so the no-rebellion constraint (A.7.26a) is therefore slack in this case. For the no-rebellion constraint (A.7.26c), using $p^{\ddagger}\psi_{\mathrm{p}}^{\ddagger}=1/2$ and the expression for ψ_{p} in (A.7.40), the constraint requires:

$$\frac{1+2\delta}{2(2\delta p+p^{\ddagger})}\geq\frac{1}{2(\delta p+(1-\delta)p^{\ddagger})}.$$

Rearranging shows that this inequality is equivalent to $\delta(1+2\delta)p+(1+2\delta)(1-\delta)p^{\ddagger}\geq 2\delta p+p^{\ddagger}$, which simplifies to $\delta(2\delta-1)(p-p^{\ddagger})\geq 0$. With $\delta>0$, this is in turn equivalent to $(2\delta-1)(p-p^{\ddagger})\geq 0$. Since $p>p^{\dagger}$ when $s>0$, (A.7.26c) holds in this case if and only if $\delta\geq 1/2$.

Assuming $\delta\geq 1/2$, it must also be confirmed that incumbents would be willing to defend the allocation against a rebellion that would result in them losing power. At the pre-investment stage, in equilibrium with $\psi_{\mathrm{w}}'=\psi_{\mathrm{w}}$ this requires $\psi_{\mathrm{p}}>\psi_{\mathrm{w}}$. Since it has been shown the no-rebellion constraint (A.7.26c) is satisfied when $\delta\geq 1/2$, observe that by using the expression for ψ_{w} in (A.7.36):

$$\psi_{\mathrm{p}}\geq\frac{1}{2(\delta p+(1-\delta)p^{\ddagger})}>\frac{1}{2(\delta p+p^{\ddagger})}=\psi_{\mathrm{w}},$$

where the middle inequality follows from $\delta p+(1-\delta)p^{\ddagger}<\delta p+p^{\ddagger}$. This confirms $\psi_{\mathrm{p}}>\psi_{\mathrm{w}}'$ at the pre-investment stage. At the post-investment stage, the requirement is $\psi_{\mathrm{p}}>\psi_{\mathrm{w}}^{\ddagger}$, and by using the expression for ψ_{p} in (A.7.40) and $\psi_{\mathrm{w}}^{\ddagger}=(2+\delta)/(2+2\delta)$ from (A.7.27), this is equivalent to:

$$\frac{1+2\delta}{2(2\delta p+p^{\ddagger})}>\frac{2+\delta}{2(1+\delta)}.$$

Using $(2+\delta)p^{\ddagger}=1$ and rearranging this inequality leads to $1+3\delta+2\delta^2>2\delta(2+\delta)p+1$, which simplifies to $p<(3+2\delta)/(4+2\delta)$. The right-hand side of the inequality is increasing in δ, so with $\delta\geq 1/2$, it is sufficient to verify it for the case of $\delta=1/2$. The requirement is $p<4/5$, which is necessarily satisfied because $p\leq 1/2$, hence $\psi_{\mathrm{p}}>\psi_{\mathrm{w}}^{\ddagger}$ is confirmed.

In summary, in the case $\delta\geq 1/2$, all constraints are satisfied when only the no-rebellion constraint (A.7.26b) binds. But if $\delta<1/2$, this configuration of binding constraint is not an equilibrium.

The case where both no-rebellion constraints at the post-investment stage are binding

Now suppose that both no-rebellion constraints (A.7.26b) and (A.7.26c) at the post-investment stage are binding. This means that the per-person worker and incumbent shares are:

$$\psi_{\mathrm{w}}=\frac{1}{2(\delta p+p^{\ddagger})},\quad\text{and}\quad\psi_{\mathrm{p}}=\frac{1}{2(\delta p+(1-\delta)p^{\ddagger})}.\qquad\text{[A.7.41]}$$

Using these to substitute for ψ_{w} and ψ_{p} in equation (A.7.25) leads to:

$$\frac{1}{2(\delta p+(1-\delta)p^{\ddagger})}=\frac{1-\frac{1-p+\mu\theta s}{2(\delta p+p^{\ddagger})}}{p},$$

and multiplying both sides by p and simplifying:

$$\frac{p}{2(\delta p+(1-\delta)p^{\ddagger})}=1-\frac{1-p+\mu\theta s}{2(\delta p+p^{\ddagger})}=\frac{(1+2\delta)p-(1-2p^{\ddagger})-\mu\theta s}{2(\delta p+p^{\ddagger})}=\frac{((1+2\delta)p-\delta p^{\ddagger})-\mu\theta s}{2(\delta p+p^{\ddagger})}.$$

This equation can be rearranged to yield an expression for $\mu\theta s$ in terms of p:

$$\mu\theta s=((1+2\delta)p-\delta p^{\ddagger})-\frac{(\delta p+p^{\ddagger})p}{\delta p+(1-\delta)p^{\ddagger}}=\frac{((1+2\delta)p-\delta p^{\ddagger})(\delta p+(1-\delta)p^{\ddagger})-(\delta p+p^{\ddagger})p}{\delta p+(1-\delta)p^{\ddagger}},$$

and expanding the brackets in the numerator leads to:

$$\mu\theta s = \frac{\delta(1+2\delta)p^2 - \delta^2 p^\ddagger p + (1-\delta)(1+2\delta)p^\ddagger p - \delta(1-\delta){p^\ddagger}^2 - \delta p^2 - p^\ddagger p}{\delta p + (1-\delta)p^\ddagger}.$$

This expression can be simplified as follows:

$$\mu\theta s = \frac{\delta\left(2\delta p^2 + (1-3\delta)p^\ddagger p - (1-\delta){p^\ddagger}^2\right)}{\delta p + (1-\delta)p^\ddagger},$$

and note that the numerator can be factorized, hence:

$$\mu\theta s = \frac{\delta(p-p^\ddagger)(2\delta p + (1-\delta)p^\ddagger)}{\delta p + (1-\delta)p^\ddagger}.$$

Therefore, the relationship between s and p in this case is given by the equation:

$$s = \frac{\delta(p-p^\ddagger)}{\mu\theta}\left(1 + \frac{\delta p}{\delta p + (1-\delta)p^\ddagger}\right). \qquad \text{[A.7.42]}$$

It can be seen that $s > 0$ is consistent with $p > p^\ddagger$. Under the parameter restrictions in (4.5) it will be shown below that the constraint $p \leq 1/2$ is satisfied for any $0 \leq s \leq 1$.

The remaining no-rebellion constraint (A.7.26a) for workers at the pre-investment stage must also hold. In equilibrium with $p' = p$ and $\psi'_{\mathrm{p}} = \psi_{\mathrm{p}}$, this requires $\psi_{\mathrm{w}} \geq \psi_{\mathrm{p}}/(1+\delta)$. Using the expressions for ψ_{p} and ψ_{w} in (A.7.41), this condition is equivalent to:

$$\frac{1}{2(\delta p + p^\ddagger)} \geq \frac{1}{2(1+\delta)(\delta p + (1-\delta)p^\ddagger)}.$$

Rearranging this inequality leads to $\delta(1+\delta)p + (1+\delta)(1-\delta)p^\ddagger \geq \delta p + p^\ddagger$, which simplifies to $\delta^2(p-p^\ddagger) \geq 0$. Given that $p > p^\ddagger$ whenever $s > 0$, it is confirmed that the no-rebellion constraint (A.7.26a) holds in this case.

It is known that when $\delta \geq 1/2$, the choice of p that maximizes ψ_{p} subject only to the no-rebellion constraint (A.7.26b) for workers at the post-investment also satisfies all other constraints. Therefore, both no-rebellion constraints (A.7.26b) and (A.7.26c) can only bind when $\delta \leq 1/2$.

Assuming $\delta \leq 1/2$, it must also be confirmed that incumbents are willing to defend the allocation against a rebellion that would see them lose power subsequently. At the pre-investment stage, in equilibrium with $\psi'_{\mathrm{w}} = \psi_{\mathrm{w}}$, this requires $\psi_{\mathrm{p}} > \psi_{\mathrm{w}}$. Using the expressions for ψ_{w} and ψ_{p} in (A.7.41), since $\delta p + (1-\delta)p^\ddagger < \delta p + p^\ddagger$, it immediately follows that:

$$\psi_{\mathrm{p}} = \frac{1}{2(\delta p + (1-\delta)p^\ddagger)} > \frac{1}{2(\delta p + p^\ddagger)} = \psi_{\mathrm{w}}.$$

At the post-investment stage, the requirement is $\psi_{\mathrm{p}} > \psi^\ddagger_{\mathrm{w}}$, and by using the expression for ψ_{p} from (A.7.41) and $\psi^\ddagger_{\mathrm{w}} = (2+\delta)/(2+2\delta)$ from (A.7.27), this is equivalent to:

$$\frac{1}{2(\delta p + (1-\delta)p^\ddagger)} > \frac{2+\delta}{2(1+\delta)}.$$

Using $(2+\delta)p^\ddagger = 1$, and rearranging this inequality leads to $1+\delta > \delta(2+\delta)p + (1-\delta)$, which simplifies to $p < 2/(2+\delta)$. The right-hand side of the inequality is decreasing in δ, and since $\delta \leq 1/2$ in this case, it suffices to verify this for $\delta = 1/2$. The requirement is $p < 4/5$, which is necessarily satisfied since $p \leq 1/2$, hence $\psi_{\mathrm{p}} > \psi^\ddagger_{\mathrm{w}}$ is confirmed.

In summary, when $\delta \leq 1/2$, all constraints are satisfied in this case and the minimum number of constraints is binding.

Power sharing and the rule of law

For any value of δ, it has been shown that the no-rebellion constraint (A.7.26b) for workers at the post-investment stage is always binding. When $\delta < 1/2$, the no-rebellion constraint (A.7.26c) for incumbents at the post-investment stage is also binding. All other no-rebellion constraints are redundant or slack.

In both cases $\delta < 1/2$ and $\delta \geq 1/2$, an equilibrium allocation must feature a relationship $s = \lambda(p)$ between power sharing p and the reach of the rule of law s given respectively by (A.7.42) or (A.7.39). These equations confirm the functional form of $\lambda(p)$ given in the proposition (see 4.7) with the definition $p^\dagger \equiv 1/(2+\delta) = p^\ddagger$. Note that in either case, $s = 0$ corresponds to $p = p^\dagger$.

Taking the function $\lambda(p)$ in the case where $\delta \leq 1/2$ (equation A.7.42):

$$\lambda(p) = \frac{\delta(p - p^\dagger)}{\mu\theta}\left(1 + \frac{\delta p}{\delta p + (1-\delta)p^\dagger}\right). \qquad [A.7.43]$$

The first derivative is

$$\begin{aligned}\lambda'(p) &= \frac{\delta}{\mu\theta}\left(1 + \frac{\delta p}{\delta p + (1-\delta)p^\dagger} + \frac{\delta(p-p^\dagger)}{\delta p + (1-\delta)p^\dagger} - \frac{\delta^2(p-p^\dagger)p}{(\delta p + (1-\delta)p^\dagger)^2}\right) \\ &= \frac{\delta\left((\delta p + (1-\delta)p^\dagger)^2 + \delta p(\delta p + (1-\delta)p^\dagger) + \delta(p-p^\dagger)(\delta p + (1-\delta p^\dagger) - \delta^2(p-p^\dagger)p\right)}{\mu\theta(\delta p + (1-\delta)p^\dagger)^2} \\ &= \frac{\delta\left(2\delta^2p^2 + (4\delta - 4\delta^2)p^\dagger p + (1 - 3\delta + 2\delta^2){p^\dagger}^2\right)}{\mu\theta(\delta p + (1-\delta)p^\dagger)^2} = \frac{\delta\left(2\delta^2p^2 + 4\delta(1-\delta)p^\dagger p + (1-\delta)(1-2\delta){p^\dagger}^2\right)}{\mu\theta(\delta p + (1-\delta)p^\dagger)^2}.\end{aligned} \qquad [A.7.44]$$

Since $\delta \leq 1/2$ in this case, the terms $(1-\delta)$ and $(1-2\delta)$ are non-negative, so the first derivative is strictly positive, confirming that $\lambda(p)$ is an increasing function.

With $\lambda(p)$ being an increasing function, the constraint $p \leq 1/2$ can be verified for all $s \in [0,1]$ by checking whether $\lambda(1/2) \geq 1$. Using (A.7.43) and $p^\dagger = 1/(2+\delta)$, this requires:

$$1 \leq \frac{\delta\left(\frac{1}{2} - \frac{1}{2+\delta}\right)}{\mu\theta}\left(1 + \frac{\frac{\delta}{2}}{\frac{\delta}{2} + \frac{(1-\delta)}{(2+\delta)}}\right),$$

which is equivalent to:

$$\mu\theta \leq \frac{\delta((2+\delta)-2)}{2(2+\delta)}\left(1 + \frac{\delta(2+\delta)}{\delta(2+\delta) + 2(1-\delta)}\right) = \frac{\delta^2(2 + 2\delta + 2\delta^2)}{2(2+\delta)(2+\delta^2)} = \frac{\delta}{2(2+\delta)}\frac{2\delta(1+\delta+\delta^2)}{2+\delta^2}.$$

It is clear that the parameter restrictions on μ and θ in (4.5) are sufficient for this to hold, confirming that $p \leq 1/2$ is never binding. Let $\bar{p}$ denote the value of p associated with $s = 1$ (the rule of law). Using (A.7.43), $\bar{p}$ is the solution (with $\bar{p} > p^\dagger$) of the equation:

$$\frac{\delta(\bar{p} - p^\dagger)}{\mu\theta}\left(1 + \frac{\delta\bar{p}}{\delta\bar{p} + (1-\delta)p^\dagger}\right) = 1, \qquad [A.7.45]$$

which is equivalent to:

$$\delta(\bar{p} - p^\dagger)(2\delta p + (1-\delta)p^\dagger) = \mu\theta(\delta\bar{p} + (1-\delta)p^\dagger).$$

Using $p^\dagger = 1/(2+\delta)$ and multiplying both sides of the equation by $(2+\delta)^2$ leads to the following quadratic in $\bar{p}$:

$$2\delta^2(2+\delta)^2\bar{p}^2 - \delta(2+\delta)(3\delta - 1 + (2+\delta)\mu\theta)\bar{p} - (1-\delta)(\delta + (2+\delta)\mu\theta) = 0.$$

Since $0 < \delta \leq 1/2$ in this case, the quadratic has only one positive root, hence the solution for $\bar{p}$ is:

$$\bar{p} = \frac{3\delta - 1 + (2+\delta)\mu\theta + \sqrt{(3\delta - 1 + (2+\delta)\mu\theta)^2 + 8(1-\delta)(\delta + (2+\delta)\mu\theta)}}{4\delta(2+\delta)},$$

which confirms the expression for $\bar{p}$ given in the proposition.

Now consider the case with $\delta \geq 1/2$. The relevant expression for $\lambda(p)$ in this case is from (A.7.39):

$$\lambda(p) = \frac{\delta(p - p^\dagger)}{\mu\theta}\left(1 + \frac{p}{2\delta p + p^\dagger}\right), \qquad [A.7.46]$$

which has first derivative:

$$\lambda'(p) = \frac{\delta}{\mu\theta}\left(1 + \frac{p}{2\delta p + p^\dagger} + \frac{p - p^\dagger}{2\delta p + p^\dagger} - \frac{2\delta(p - p^\dagger)p}{(2\delta p + p^\dagger)^2}\right)$$
$$= \frac{\delta\left((2\delta p + p^\dagger)^2 + p(2\delta p + p^\dagger) + (p - p^\dagger)(2\delta p + p^\dagger) - 2\delta(p - p^\dagger)p\right)}{\mu\theta(2\delta p + p^\dagger)^2}$$
$$= \frac{\delta\left((2\delta + 4\delta^2)p^2 + (2 + 4\delta)p^\dagger p\right)}{\mu\theta(2\delta p + p^\dagger)^2} = \frac{2\delta(1 + 2\delta)(\delta p + p^\dagger)p}{\mu\theta(2\delta p + p^\dagger)^2}. \quad \text{[A.7.47]}$$

This is positive, confirming that $\lambda(p)$ is an increasing function.

With $\lambda(p)$ being increasing in p, the constraint $p \leq 1/2$ can be verified for all $s \in [0, 1]$ by checking whether $\lambda(1/2) \geq 1$. Using (A.7.46) and $p^\dagger = 1/(2 + \delta)$, this requires:

$$1 \leq \frac{\delta\left(\frac{1}{2} - \frac{1}{2+\delta}\right)}{\mu\theta}\left(1 + \frac{\frac{1}{2}}{\frac{2\delta}{2} + \frac{1}{2+\delta}}\right),$$

which is equivalent to:

$$\mu\theta \leq \frac{\delta((2+\delta) - 2)}{2(2+\delta)}\left(1 + \frac{2+\delta}{2\delta(2+\delta) + 2}\right) = \frac{\delta^2(4 + 5\delta + 2\delta^2)}{4(2+\delta)(1 + 2\delta + \delta^2)} = \frac{\delta}{2(2+\delta)}\frac{\delta(4 + 5\delta + 2\delta^2)}{2(1+\delta)^2}.$$

It is clear that the parameter restrictions in (4.5) are sufficient for this to hold, confirming that $p \leq 1/2$ is never binding. Let $\bar{p}$ denote the value of p associated with $s = 1$ (the rule of law). Using (A.7.46), $\bar{p}$ is the solution (with $\bar{p} > p^\dagger$) of the equation:

$$\frac{\delta(\bar{p} - p^\dagger)}{\mu\theta}\left(1 + \frac{\bar{p}}{2\delta\bar{p} + p^\dagger}\right), \quad \text{[A.7.48]}$$

which is equivalent to:

$$\delta(\bar{p} - p^\dagger)((1 + 2\delta)\bar{p} + p^\dagger) = \mu\theta(2\delta\bar{p} + p^\dagger).$$

Using $p^\dagger = 1/(2+\delta)$ and multiplying both sides of the equation by $(2+\delta)^2$ leads to the following quadratic in $\bar{p}$:

$$\delta(1 + 2\delta)(2 + \delta)^2\bar{p}^2 - 2\delta(2+\delta)(\delta + (2+\delta)\mu\theta) - (\delta + (2+\delta)\mu\theta) = 0.$$

This quadratic equation has one positive root, so $\bar{p}$ is given by:

$$\bar{p} = \frac{\delta(\delta + (2+\delta)\mu\theta) + \sqrt{(\delta(\delta + (2+\delta)\mu\theta))^2 + \delta(1 + 2\delta)(\delta + (2+\delta)\mu\theta)}}{\delta(1 + 2\delta)(2 + \delta)},$$

which confirms the expression for $\bar{p}$ given in the proposition. All claims in part (iv) have now been demonstrated.

The income distribution

The no-rebellion constraint (A.7.26b) is binding for all parameter values, which confirms the formula given in (4.8). In the case $0 < \delta \leq 1/2$, both post-investment no-rebellion constraints are binding and the per-person incumbent share is given in (A.7.41). In the case $\delta \geq 1/2$, only the post-investment no-rebellion constraint for workers is binding, and the optimal choice of p subject to this constraint implies the incumbent share specified in (A.7.40). These two cases confirm the formula given in (4.8), and as both $\psi_{\mathrm{p}}(p)$ and $\psi_{\mathrm{w}}(p)$ are strictly decreasing in p, all the claims in part (v) are demonstrated.

The impact of the rule of law on incumbents' payoff

Consider a particular value of s consistent with (4.3). For an allocation with s to be an equilibrium, power sharing p must be $p = \lambda^{-1}(s)$ and the incumbent share must be $\psi_{\mathrm{p}}(\lambda^{-1}(s))$. With $U_{\mathrm{p}} = \log C_{\mathrm{p}}$ (from A.7.8), maximizing the incumbent payoff is equivalent to maximizing incumbent consumption C_{p} with respect to

s. After taking account of the constraints, this means s must maximize the following function:

$$C_{\mathrm{p}}(s) = \psi_{\mathrm{p}}(\lambda^{-1}(s))C,$$

where C is given in (2.9). Using $\partial C/\partial s = \mu\tilde{\pi}/\tilde{\pi}^{\alpha}$, the first derivative is:

$$C'_{\mathrm{p}}(s) = \frac{\mu\tilde{\pi}\psi_{\mathrm{p}}(\lambda^{-1}(s))}{\tilde{\pi}^{\alpha}} + \frac{\psi'_{\mathrm{p}}(\lambda^{-1}(s))C}{\lambda'(\lambda^{-1}(s))}.$$

Using $C = Y/\tilde{\pi}^{\alpha}$ (from 2.9) and $Y = Y_{\mathrm{w}}/\psi_{\mathrm{w}}(\lambda^{-1}(s))$ (from 4.8), the derivative above can be written as:

$$C'_{\mathrm{p}}(s) = \frac{\mu\psi_{\mathrm{p}}(\lambda^{-1}(s))}{\tilde{\pi}^{\alpha}}\left(\tilde{\pi} - \left(\frac{-\psi'_{\mathrm{p}}(\lambda^{-1}(s))}{\mu\theta\lambda'(\lambda^{-1}(s))\psi_{\mathrm{w}}(\lambda^{-1}(s))\psi_{\mathrm{p}}(\lambda^{-1}(s))}\right)\theta Y_{\mathrm{w}}\right),$$

which is equivalent to:

$$C'_{\mathrm{p}}(s) = \frac{\mu\psi_{\mathrm{p}}(p)}{\tilde{\pi}^{\alpha}}\left(\tilde{\pi} - (1+\chi(p))\theta Y_{\mathrm{w}}\right), \quad \text{where } \chi(p) = \left(\frac{-\psi'_{\mathrm{p}}(p)}{\mu\theta\lambda'(p)\psi_{\mathrm{w}}(p)\psi_{\mathrm{p}}(p)}\right) - 1. \qquad \text{[A.7.49]}$$

Thus, there exists a function $\chi(p)$ such that the expression for $C'_{\mathrm{p}}(s)$ given in (4.9) is valid.

Consider first the case where $\delta < 1/2$. Differentiating the expression for $\psi_{\mathrm{p}}(p)$ given in (4.8) shows that:

$$-\psi'_{\mathrm{p}}(p) = 2\delta\psi_{\mathrm{p}}(p)^2,$$

and substituting this into the definition of $\chi(p)$ from (A.7.49):

$$\chi(p) = \frac{2\delta\psi_{\mathrm{p}}(p)}{\mu\theta\lambda'(p)\psi_{\mathrm{w}}(p)} - 1.$$

Using the formulas for $\psi_{\mathrm{w}}(p)$ and $\psi_{\mathrm{p}}(p)$ from (4.8) and the expression for $\lambda'(p)$ from equation (A.7.44):

$$\chi(p) = \frac{\frac{2\delta}{2(\delta p+(1-\delta)p^{\dagger})}}{\frac{\delta\left(2\delta^2p^2+4\delta(1-\delta)p^{\dagger}p+(1-\delta)(1-2\delta){p^{\dagger}}^2\right)}{2(\delta p+p^{\dagger})(\delta p+(1-\delta)p^{\dagger})^2}} - 1 = \frac{2(\delta p+p^{\dagger})(\delta p+(1-\delta)p^{\dagger})}{2\delta^2p^2+4\delta(1-\delta)p^{\dagger}p+(1-\delta)(1-2\delta){p^{\dagger}}^2} - 1.$$

By expanding the brackets in the numerator and taking a common denominator, the following expression for $\chi(p)$ is obtained:

$$\begin{aligned}\chi(p) &= \frac{2\delta^2p^2+2\delta(2-\delta)p^{\dagger}p+2(1-\delta){p^{\dagger}}^2-2\delta^2p^2-4\delta(1-\delta)p^{\dagger}p-(1-\delta)(1-2\delta){p^{\dagger}}^2}{2\delta^2p^2+4\delta(1-\delta)p^{\dagger}p+(1-\delta)(1-2\delta){p^{\dagger}}^2} \\ &= \frac{2\delta^2p^{\dagger}p+(1-\delta)(1+2\delta){p^{\dagger}}^2}{2\delta^2p^2+4\delta(1-\delta)p^{\dagger}p+(1-\delta)(1-2\delta){p^{\dagger}}^2} = \frac{p^{\dagger}\left(2\delta^2p+(1-\delta)(1+2\delta)p^{\dagger}\right)}{2\delta^2p^2+4\delta(1-\delta)p^{\dagger}p+(1-\delta)(1-2\delta){p^{\dagger}}^2}.\end{aligned}$$

Since $\delta < 1/2$, the terms $1-\delta$ and $1-2\delta$ are strictly positive, hence the equation above confirms that $\chi(p)$ is strictly positive. Differentiating $\chi(p)$ with respect to p and simplifying:

$$\begin{aligned}\chi'(p) &= \frac{2\delta^2(2\delta^2p^2+4\delta(1-\delta)p^{\dagger}p+(1-\delta)(1-2\delta){p^{\dagger}}^2)-4\delta(\delta p+(1-\delta)p^{\dagger})(2\delta^2p+(1-\delta)(1+2\delta)p^{\dagger})}{{p^{\dagger}}^{-1}\left(2\delta^2p^2+4\delta(1-\delta)p^{\dagger}p+(1-\delta)(1-2\delta){p^{\dagger}}^2\right)^2} \\ &= \frac{-2\delta\left(2(2\delta^3p^2+(1-\delta)(\delta+4\delta^2)p^{\dagger}p+(1-\delta)(1+2\delta){p^{\dagger}}^2)-2\delta^3p^2-\delta(1-\delta)(4p^{\dagger}p+(1-2\delta){p^{\dagger}}^2)\right)}{{p^{\dagger}}^{-1}\left(2\delta^2p^2+4\delta(1-\delta)p^{\dagger}p+(1-\delta)(1-2\delta){p^{\dagger}}^2\right)^2} \\ &= \frac{-2\delta p^{\dagger}\left(2\delta^3p^2+2\delta(1-\delta)(1+2\delta)p^{\dagger}p+(1-\delta)(2(1-\delta)(1+2\delta)-\delta(1-2\delta)){p^{\dagger}}^2\right)}{\left(2\delta^2p^2+4\delta(1-\delta)p^{\dagger}p+(1-\delta)(1-2\delta){p^{\dagger}}^2\right)^2} \\ &= -\frac{2\delta p^{\dagger}\left(2\delta^3p^2+2\delta(1-\delta)(1+2\delta)p^{\dagger}p+(1-\delta)(2+\delta(1-2\delta)){p^{\dagger}}^2\right)}{\left(2\delta^2p^2+4\delta(1-\delta)p^{\dagger}p+(1-\delta)(1-2\delta){p^{\dagger}}^2\right)^2},\end{aligned}$$

which is unambiguously negative because $\delta < 1/2$ implies the terms $1-\delta$ and $1-2\delta$ are strictly positive. Hence, $\chi(p)$ is strictly positive and strictly decreasing when $\delta < 1/2$.

Now consider the case where $\delta \geq 1/2$. Differentiating the expression for $\psi_{\mathrm{p}}(p)$ from (4.8) shows that:

$$-\psi_{\mathrm{p}}'(p) = \frac{4\delta}{1+2\delta}\psi_{\mathrm{p}}(p)^2,$$

and substituting this into the definition of $\chi(p)$ from (A.7.49):

$$\chi(p) = \frac{4\delta\psi_{\mathrm{p}}(p)}{\mu\theta(1+2\delta)\lambda'(p)\psi_{\mathrm{w}}(p)} - 1.$$

Using the formulas for $\psi_{\mathrm{w}}(p)$ and $\psi_{\mathrm{p}}(p)$ from (4.8) and the expression for $\lambda'(p)$ from equation (A.7.47):

$$\chi(p) = \frac{\frac{4\delta(1+2\delta)}{2(2\delta p+p^{\dagger})}}{\frac{(1+2\delta)2\delta(1+2\delta)(\delta p+p^{\dagger})p}{2(\delta p+p^{\dagger})(2\delta p+p^{\dagger})^2}} - 1 = \frac{2(2\delta p+p^{\dagger})}{(1+2\delta)p} - 1.$$

By taking a common denominator, $\chi(p)$ can be written as follows:

$$\chi(p) = \frac{4\delta p + 2p^{\dagger} - p - 2\delta p}{(1+2\delta)p} = \frac{(2\delta-1)p + 2p^{\dagger}}{(1+2\delta)p} = \frac{(2\delta-1) + 2\frac{p^{\dagger}}{p}}{1+2\delta}.$$

It can be seen immediately that $\chi(p)$ is strictly positive and strictly decreasing when $\delta \geq 1/2$, which ensures the term $2\delta - 1$ is strictly positive. Therefore, it has been established that $\chi(p) > 0$ and $\chi'(p) < 0$ for all values of δ, which are the properties claimed in the proposition. This completes the proof.

A.8 Proof of Proposition 8

(i) The income distribution results from Proposition 7 imply that $Y_{\mathrm{p}} = \psi_{\mathrm{p}}(p)Y$, where p is power sharing (see 4.8). That proposition also links power sharing p to the quality of government s according to $s = \lambda(p)$. Since $\lambda(p)$ is a strictly increasing function and takes on all values between 0 and 1 for feasible values of p ($0 < p \leq 1/2$) it has a well-defined inverse $p = \lambda^{-1}(s)$. Substituting this into the equation for Y_{p}:

$$Y_{\mathrm{p}} = \psi_{\mathrm{p}}(\lambda^{-1}(s))Y,$$

which has exactly the same form as (2.11) with the function $\phi(s)$ being as given in (4.11). The first derivative of the function $\phi(s)$ is:

$$\phi'(s) = \frac{\psi_{\mathrm{p}}'(\lambda^{-1}(s))}{\lambda'(\lambda^{-1}(s))}, \qquad [\text{A.8.1}]$$

which is strictly negative because $\psi_{\mathrm{p}}'(p) < 0$ and $\lambda'(p) > 0$ according to Proposition 7. This confirms that the equilibrium allocation features an equivalent of the political friction (2.11) for some strictly decreasing function $\phi(s)$.

Using the definition of the marginal cost of good government $\gamma(s)$ from (2.13) and substituting from (A.8.1):

$$\gamma(s) = \frac{-\frac{\psi_{\mathrm{p}}'(\lambda^{-1}(s))}{\psi_{\mathrm{p}}(\lambda^{-1}(s))^2}}{\lambda'(\lambda^{-1}(s))}. \qquad [\text{A.8.2}]$$

In the case where $\delta < 1/2$, Proposition 7 shows that the function $\psi_{\mathrm{p}}(p)$ is given by:

$$\psi_{\mathrm{p}}(p) = \frac{1}{2(\delta p + (1-\delta)p^{\dagger})}, \qquad [\text{A.8.3}]$$

which has the following derivative with respect to p:

$$\psi_{\mathrm{p}}'(p) = -\frac{\delta}{2(\delta p + (1-\delta)p^{\dagger})^2} = -2\delta\psi_{\mathrm{p}}(p)^2. \qquad [\text{A.8.4}]$$

In the case where $\delta \geq 1/2$, Proposition 7 shows that the function $\psi_{\mathrm{p}}(p)$ is given by:

$$\psi_{\mathrm{p}}(p) = \frac{1+2\delta}{2(2\delta p + p^{\dagger})}, \qquad \text{[A.8.5]}$$

which has the following derivative with respect to p:

$$\psi'_{\mathrm{p}}(p) = -\frac{\delta(1+2\delta)}{(2\delta p + p^{\dagger})^2} = -\frac{4\delta}{1+2\delta}\psi_{\mathrm{p}}(p)^2. \qquad \text{[A.8.6]}$$

Since $\min\{1, 2/(1+2\delta)\}$ equals 1 when $\delta < 1/2$ and $2/(1+2\delta)$ when $\delta \geq 1/2$, equations (A.8.4) and (A.8.6) together show that:

$$-\frac{\psi'_{\mathrm{p}}(p)}{\psi_{\mathrm{p}}(p)^2} = 2\delta \min\left\{1, \frac{2}{1+2\delta}\right\}. \qquad \text{[A.8.7]}$$

Substituting this into (A.8.2) confirms the expression for $\gamma(s)$ given in (4.12).

It is shown in the proof of Proposition 2 (see A.2.3) that $\Phi = (\phi(0) - \phi(1))/\phi(1)$. Using (4.11) and the results of Proposition 7, $\phi(0) = \psi_{\mathrm{p}}(\lambda^{-1}(0)) = \psi_{\mathrm{p}}(p^{\dagger})$ and $\phi(1) = \psi_{\mathrm{p}}(\lambda^{-1}(1)) = \psi_{\mathrm{p}}(\bar{p})$, hence $\Phi = (\psi_{\mathrm{p}}(p^{\dagger}) - \psi_{\mathrm{p}}(\bar{p}))/\psi_{\mathrm{p}}(\bar{p})$. Using the expression for $\psi_{\mathrm{p}}(p)$ from (4.8), observe that $\psi_{\mathrm{p}}(p^{\dagger}) = (2+\delta)/2$. Take the case of $\delta < 1/2$:

$$\frac{\psi_{\mathrm{p}}(p^{\dagger}) - \psi_{\mathrm{p}}(\bar{p})}{\psi_{\mathrm{p}}(\bar{p})} = \frac{\frac{2+\delta}{2}}{\frac{1}{2(\delta\bar{p}+(1-\delta)p^{\dagger})}} - 1 = (2+\delta)(\delta\bar{p} + (1-\delta)p^{\dagger}) - (2+\delta)p^{\dagger} = \delta(2+\delta)(\bar{p} - p^{\dagger}), \qquad \text{[A.8.8a]}$$

which uses $p^{\dagger} = 1/(2+\delta)$ (see 4.7), and in the case of $\delta \geq 1/2$:

$$\frac{\psi_{\mathrm{p}}(p^{\dagger}) - \psi_{\mathrm{p}}(\bar{p})}{\psi_{\mathrm{p}}(\bar{p})} = \frac{\frac{2+\delta}{2}}{\frac{1+2\delta}{2(2\delta\bar{p}+p^{\dagger})}} - 1 = \frac{(2+\delta)(2\delta\bar{p} + p^{\dagger}) - (1+2\delta)(2+\delta)p^{\dagger}}{1+2\delta} = \frac{2\delta(2+\delta)}{1+2\delta}(\bar{p} - p^{\dagger}). \qquad \text{[A.8.8b]}$$

Since $\min\{1, 2/(1+2\delta)\}$ equals 1 when $\delta < 1/2$ and $2/(1+2\delta)$ when $\delta \geq 1/2$, it follows from (A.8.8a) and (A.8.8b) that a general formula for $(\psi_{\mathrm{p}}(p^{\dagger}) - \psi_{\mathrm{p}}(\bar{p}))/\psi_{\mathrm{p}}(\bar{p})$ valid for all parameter values is:

$$\frac{\psi_{\mathrm{p}}(p^{\dagger}) - \psi_{\mathrm{p}}(\bar{p})}{\psi_{\mathrm{p}}(\bar{p})} = \delta(2+\delta)\min\left\{1, \frac{2}{1+2\delta}\right\}(\bar{p} - p^{\dagger}). \qquad \text{[A.8.9]}$$

As this is equal to Φ, the claims in (4.12) are confirmed.

(ii) Differentiating the expression for the marginal cost of good government $\gamma(s)$ in (4.12):

$$\gamma'(s) = \frac{2\delta\min\left\{1, \frac{2}{1+2\delta}\right\}\lambda''(\lambda^{-1}(s))}{(\lambda'(\lambda^{-1}(s)))^3}.$$

It is known that $\lambda(p)$ is a strictly increasing function (Proposition 7), hence (2.14) is satisfied ($\gamma'(s) < 0$) if and only if $\lambda''(p) > 0$.

To demonstrate strict convexity of the function $\lambda(p)$, first consider the case $\delta < 1/2$. By using (A.7.44) and $(\delta p + (1-\delta)p^{\dagger})^2 = \delta^2 p^2 + 2\delta(1-\delta)p^{\dagger}p + (1-\delta)^2 {p^{\dagger}}^2$, the first derivative $\lambda'(p)$ can be written as follows:

$$\lambda'(p) = \frac{\delta\left(2(\delta p + (1-\delta)p^{\dagger})^2 - 2(1-\delta)^2{p^{\dagger}}^2 + (1-\delta)(1-2\delta){p^{\dagger}}^2\right)}{\mu\theta(\delta p + (1-\delta)p^{\dagger})^2}$$
$$= \frac{\delta\left(2(\delta p + (1-\delta)p^{\dagger})^2 - (1-\delta){p^{\dagger}}^2\right)}{\mu\theta(\delta p + (1-\delta)p^{\dagger})^2} = \frac{\delta}{\mu\theta}\left(2 - \frac{(1-\delta){p^{\dagger}}^2}{(\delta p + (1-\delta)p^{\dagger})^2}\right).$$

The second derivative of $\lambda(p)$ is thus:

$$\lambda''(p) = \frac{2\delta^2(1-\delta){p^{\dagger}}^2}{\mu\theta(\delta p + (1-\delta)p^{\dagger})^2},$$

which is always positive, confirming that $\lambda(p)$ is a convex function in this case. Now consider $\delta \geq 1/2$. Note that $(2\delta p + p^{\dagger})^2 = 4\delta(\delta p + p^{\dagger})p + {p^{\dagger}}^2$ and hence $2\delta(\delta p + p^{\dagger}) = (2\delta p + p^{\dagger})^2/2 - {p^{\dagger}}^2/2$. Substituting this into (A.7.47) shows that $\lambda'(p)$ can be written as follows:

$$\lambda'(p) = \frac{(1+2\delta)\left((2\delta p + p^{\dagger})^2 - {p^{\dagger}}^2\right)}{2\mu\theta(2\delta p + p^{\dagger})^2} = \frac{1+2\delta}{2\mu\theta}\left(1 - \frac{{p^{\dagger}}^2}{(2\delta p + p^{\dagger})^2}\right).$$

The second derivative of $\lambda(p)$ is thus:

$$\lambda''(p) = \frac{2\delta(1+2\delta){p^{\dagger}}^2}{\mu\theta(2\delta p + p^{\dagger})^3},$$

which is always positive. This confirms the claim that $\lambda(p)$ is a convex function in all cases, and thus that $\gamma(s)$ is strictly decreasing.

Proposition 7 demonstrates that incumbents and workers receive consumption $C_{\mathrm{p}} = \psi_{\mathrm{p}}(p)C$ and $C_{\mathrm{w}} = \psi_{\mathrm{w}}(p)C$, and that the incentive compatibility constraint for investors must bind, hence $C_{\mathrm{k}} = (1+\theta)C_{\mathrm{w}}$. It is shown in the proof of Proposition 7 (see A.7.25) that substituting these consumption levels into the resource constraint implies:

$$p\psi_{\mathrm{p}}(p) + (1 - p + \mu\theta s)\psi_{\mathrm{w}}(p) = 1. \qquad \text{[A.8.10]}$$

With $C_{\mathrm{p}} = \psi_{\mathrm{p}}(p)C$ and $C_{\mathrm{w}} = \psi_{\mathrm{w}}(p)C$, rents $\beta = (C_{\mathrm{p}} - C_{\mathrm{w}})/C_{\mathrm{w}}$ as a function of power sharing p are $\beta(p) = (\psi_{\mathrm{p}}(p) - \psi_{\mathrm{w}}(p))/\psi_{\mathrm{w}}(p)$. It follows that $\psi_{\mathrm{p}}(p) = (1+\beta(p))\psi_{\mathrm{w}}(p)$, which can be substituted into the resource constraint above to obtain:

$$p(1+\beta(p))\psi_{\mathrm{w}}(p) + (1 - p + \mu\theta s)\psi_{\mathrm{w}}(p) = 1.$$

Substituting the relationship $s = \lambda(p)$, dividing both sides by $\psi_{\mathrm{w}}(p)$ and using the formula for $\psi_{\mathrm{w}}(p)$ from (4.8):

$$p(1+\beta(p)) + 1 - p + \mu\theta\lambda(p) = 2(\delta p + p^{\dagger}).$$

This equation can be rearranged to give the following expression for $\lambda(p)$:

$$\lambda(p) = \frac{2(\delta p + p^{\dagger}) - 1 - B(p)}{\mu\theta}, \quad \text{with} \;\; B(p) = p\beta(p),$$

where $B(p)$ denotes total rents received by all individuals in power. This shows that a formula for the second derivative of $\lambda(p)$ is:

$$\lambda''(p) = -\frac{B''(p)}{\mu\theta},$$

and thus the curvature of the function $\lambda(p)$ depends only on the behaviour of rents $\beta(p)$ through the total rents function $B(p) = p\beta(p)$. The convexity of $\lambda(p)$ is therefore explained by the concavity of total rents $B(p)$. Since $B''(p) = 2\beta'(p) + p\beta''(p)$, the concavity of $B(p)$ comes either from rents $\beta(p)$ being decreasing in power sharing p, or having a sufficiently declining rate of increase.

In the case of $\delta < 1/2$, an expression for rents $\beta(p)$ can be obtained from (4.8):

$$\beta(p) = \frac{\delta p + p^{\dagger}}{\delta p + (1-\delta)p^{\dagger}} - 1 = \frac{\delta p^{\dagger}}{\delta p + (1-\delta)p^{\dagger}},$$

which is positive and decreasing in p. In the case of $\delta \geq 1/2$, the equations in (4.8) imply that rents $\beta(p)$ are:

$$\beta(p) = \frac{(1+2\delta)(\delta p + p^{\dagger})}{2\delta p + p^{\dagger}} - 1 = \frac{\delta(1+2\delta)p + (1+2\delta)p^{\dagger} - 2\delta p - p^{\dagger}}{2\delta p + p^{\dagger}} = \frac{\delta(2\delta - 1)p + 2\delta p^{\dagger}}{2\delta p + p^{\dagger}}.$$

The derivative with respect to p is:

$$\beta'(p) = \frac{\delta(2\delta-1)(2\delta p + p^{\dagger}) - 2\delta(\delta(2\delta-1)p + 2\delta p^{\dagger})}{(2\delta p + p^{\dagger})^2} = -\frac{-\delta(1+2\delta)p^{\dagger}}{(2\delta p + p^{\dagger})^2},$$

which is negative, showing that rents $\beta(p)$ are decreasing in p in this case too.

(iii) Start by defining a function $\mathcal{B}(s)$:

$$\mathcal{B}(s) = \alpha - s\gamma(s)\phi(s), \tag{A.8.11}$$

where $\psi(s)$ and $\gamma(s)$ are as given in (4.11) and (4.12). Using the formulas in the equations for $\phi(s)$ and $\gamma(s)$, the function $\mathcal{B}(s)$ can be written as:

$$\mathcal{B}(s) = \alpha - 2\delta \min\left\{1, \frac{2}{1+2\delta}\right\} \frac{\lambda(\lambda^{-1}(s))\psi_{\mathrm{p}}(\lambda^{-1}(s))}{\lambda'(\lambda^{-1}(s))},$$

noting that $s = \lambda(\lambda^{-1}(s))$. Since s only appears in the above through $p = \lambda^{-1}(s)$, the expression for $\mathcal{B}(s)$ can be stated in terms of a function $\mathcal{M}(p)$ of p:

$$\mathcal{B}(s) = \mathcal{M}(\lambda^{-1}(s)), \quad \text{where } \mathcal{M}(p) = \alpha - 2\delta \min\left\{1, \frac{2}{1+2\delta}\right\} \frac{\lambda(p)\psi_{\mathrm{p}}(p)}{\lambda'(p)}. \tag{A.8.12}$$

First take the case where $\delta < 1/2$. Using the expression for $\lambda'(p)$ from equation (A.7.44) of the proof of Proposition 7 and the expression for $\psi_{\mathrm{p}}(p)$ from (4.8):

$$\frac{\lambda'(p)}{\psi_{\mathrm{p}}(p)} = \frac{2\delta\left(2\delta^2 p^2 + 4\delta(1-\delta)p^{\dagger}p + (1-\delta)(1-2\delta){p^{\dagger}}^2\right)}{\mu\theta(\delta p + (1-\delta)p^{\dagger})}. \tag{A.8.13}$$

The formula for $\lambda(p)$ from (4.7) in the case $\delta \leq 1/2$ is

$$\lambda(p) = \frac{\delta(p-p^{\dagger})}{\mu\theta}\left(\frac{2\delta p + (1-\delta)p^{\dagger}}{\delta p + (1-\delta)p^{\dagger}}\right),$$

which can be used together with (A.8.13) to deduce the following:

$$2\delta \min\left\{1, \frac{2}{1+2\delta}\right\} \frac{\lambda(p)\psi_{\mathrm{p}}(p)}{\lambda'(p)} = \frac{\delta(p-p^{\dagger})(2\delta p + (1-\delta)p^{\dagger})}{2\delta^2 p^2 + 4\delta(1-\delta)p^{\dagger}p + (1-\delta)(1-2\delta){p^{\dagger}}^2}. \tag{A.8.14}$$

When $\delta \geq 1/2$, the relevant expression for $\lambda'(p)$ is given in equation (A.7.47) of the proof of Proposition 7. Combined with (4.8), this implies:

$$\frac{\lambda'(p)}{\psi_{\mathrm{p}}(p)} = \frac{4\delta(\delta p + p^{\dagger})}{\mu\theta(2\delta p + p^{\dagger})}. \tag{A.8.15}$$

The function $\lambda(p)$ in the case $\delta \geq 1/2$ is given in (4.7):

$$\lambda(p) = \frac{\delta(p-p^{\dagger})}{\mu\theta}\left(\frac{(1+2\delta)p + p^{\dagger}}{2\delta p + p^{\dagger}}\right),$$

which together with (A.8.15) can be used to deduce the following:

$$2\delta \min\left\{1, \frac{2}{1+2\delta}\right\} \frac{\lambda(p)\psi_{\mathrm{p}}(p)}{\lambda'(p)} = \frac{\delta(p-p^{\dagger})((1+2\delta)p + p^{\dagger})}{(1+2\delta)(\delta p + p^{\dagger})p}. \tag{A.8.16}$$

Substituting equations (A.8.14) and (A.8.16) respectively for the cases of $\delta \leq 1/2$ and $\delta \geq 1/2$ into (A.8.12) leads to:

$$\begin{aligned}\mathcal{M}(p) &= \alpha - \frac{2\delta^2 p^2 + \delta(1-3\delta)p^{\dagger}p - \delta(1-\delta){p^{\dagger}}^2}{2\delta^2 p^2 + 4\delta(1-\delta)p^{\dagger}p + (1-\delta)(1-2\delta){p^{\dagger}}^2} \\ &= \frac{2\delta^2(\alpha-1)p^2 + \delta(3\delta - 1 + 4(1-\delta)\alpha)p^{\dagger}p + (1-\delta)(\delta + (1-2\delta)\alpha){p^{\dagger}}^2}{2\delta^2 p^2 + 4\delta(1-\delta)p^{\dagger}p + (1-\delta)(1-2\delta){p^{\dagger}}^2},\end{aligned} \tag{A.8.17a}$$

for the case $\delta < 1/2$, and for $\delta \geq 1/2$:

$$\begin{aligned}\mathcal{M}(p) &= \alpha - \frac{\delta(1+2\delta)p^2 - 2\delta^2 p^{\dagger}p - \delta{p^{\dagger}}^2}{\delta(1+2\delta)p^2 + (1+2\delta)p^{\dagger}p} \\ &= \frac{\delta(1+2\delta)(\alpha-1)p^2 + (2\delta^2 + (1+2\delta)\alpha)p^{\dagger}p + \delta{p^{\dagger}}^2}{(1+2\delta)(\delta p + p^{\dagger})p}.\end{aligned} \tag{A.8.17b}$$

Now define the following quadratic function of p:

$$\mathcal{Q}(p) \equiv \begin{cases} 2\delta^2(2+\delta)^2(1-\alpha)p^2 - \delta(2+\delta)(3\delta - 1 + 4(1-\delta)\alpha)p - (1-\delta)(\delta + (1-2\delta)\alpha) & \text{for } \ \delta < 1/2 \\ \delta(1+2\delta)(2+\delta)^2(1-\alpha)p^2 - (2+\delta)(2\delta^2 + (1+2\delta)\alpha)p - \delta & \text{for } \ \delta \geq 1/2 \end{cases}, \tag{A.8.18}$$

noting that it is equal to the negative of the numerator of the expressions for $\mathcal{M}(p)$ in (A.8.17). The denominator of the expressions for $\mathcal{M}(p)$ in (A.8.17) is seen to be strictly positive. Since $\bar{p} = \lambda^{-1}(1)$, it follows from these observation and (A.8.12) that $\mathcal{B}(1) < 0$ is equivalent to $\mathcal{Q}(\bar{p}) > 0$. With $\mathcal{B}(1) = \alpha - \gamma(1)\phi(1)$ (equation A.8.11), the condition (2.16) for the political frictions to be binding is $\mathcal{B}(1) < 0$.

In the case $\delta \leq 1/2$, it is shown in the proof of Proposition 7 that $\bar{p}$ satisfies the quadratic equation (A.7.45), and hence:

$$2\delta^2(2+\delta)^2\bar{p}^2 = \delta(2+\delta)(3\delta - 1 + (2+\delta)\mu\theta)\bar{p} + (1-\delta)(\delta + (2+\delta)\mu\theta).$$

Evaluating the quadratic $\mathcal{Q}(p)$ in (A.8.18) at $p = \bar{p}$ and using the equation above to deduce:

$$\begin{aligned} \mathcal{Q}(\bar{p}) = {} & (1-\alpha)\left(\delta(2+\delta)(3\delta - 1 + (2+\delta)\mu\theta)\bar{p} + (1-\delta)(\delta + (2+\delta)\mu\theta)\right) \\ & - \delta(2+\delta)(3\delta - 1 + 4(1-\delta)\alpha)\bar{p} - (1-\delta)(\delta + (1-2\delta)\alpha) \\ & = (1-\alpha)\mu\theta(2+\delta)(1 - \delta + \delta(2+\delta)\bar{p}) - \alpha((1-\delta)^2 + \delta(2+\delta)(3-\delta)), \end{aligned}$$

from which it follows that $\mathcal{Q}(\bar{p}) > 0$ is equivalent to:

$$\frac{\alpha}{1-\alpha} < \frac{\mu\theta(2+\delta)(1-\delta+\delta(2+\delta)\bar{p})}{(1-\delta)^2 + \delta(2+\delta)(3-\delta)\bar{p}}.$$

This inequality can be stated as $\alpha < \bar{\alpha}$, where $\bar{\alpha}$ is given by:

$$\bar{\alpha} = \left(1 + \frac{(1-\delta)^2 + \delta(2+\delta)(3-\delta)\bar{p}}{\mu\theta(2+\delta)(1-\delta+\delta(2+\delta)\bar{p})}\right)^{-1}, \tag{A.8.19a}$$

where this number lies strictly between 0 and 1. Similarly, in the case $\delta \geq 1/2$, the proof of Proposition 7 shows that $\bar{p}$ satisfies the quadratic equation (A.7.48), and hence:

$$\delta(1+2\delta)(2+\delta)^2\bar{p}^2 = 2\delta(2+\delta)(\delta + (2+\delta)\mu\theta)\bar{p} + (\delta + (2+\delta)\mu\theta).$$

Evaluating the quadratic $\mathcal{Q}(p)$ in (A.8.18) at $p = \bar{p}$ and using the equation above to deduce:

$$\begin{aligned} \mathcal{Q}(\bar{p}) = {} & (1-\alpha)\left(2\delta(2+\delta)(\delta + (2+\delta)\mu\theta)\bar{p} + (\delta + (2+\delta)\mu\theta)\right) - (2+\delta)(2\delta^2 + (1+2\delta)\alpha)\bar{p} - \delta \\ & = (1-\alpha)\mu\theta(2+\delta)(1 + 2\delta(2+\delta)\bar{p}) - \alpha(\delta + (2+\delta)(1 + 2\delta + 2\delta^2)\bar{p}), \end{aligned}$$

from which it follows that $\mathcal{Q}(\bar{p}) > 0$ is equivalent to:

$$\frac{\alpha}{1-\alpha} < \frac{\mu\theta(2+\delta)(1+2\delta(2+\delta)\bar{p})}{\delta + (2+\delta)(1+2\delta+2\delta^2)\bar{p}}.$$

This inequality can be stated as $\alpha < \bar{\alpha}$, where $\bar{\alpha}$ is given by:

$$\bar{\alpha} = \left(1 + \frac{\delta + (2+\delta)(1+2\delta+2\delta^2)\bar{p}}{\mu\theta(2+\delta)(1+2\delta(2+\delta)\bar{p})}\right)^{-1}, \tag{A.8.19b}$$

where this number lies strictly between 0 and 1. This confirms that the political friction is binding (2.16 holds) when $\alpha < \bar{\alpha}$, and the expression given for $\bar{\alpha}$ is confirmed by using (A.8.19a) and (A.8.19b).

(iv) Since (2.11) holds for $\phi(s) = \psi_{\mathrm{p}}(\lambda^{-1}(s))$, the formula derived in (A.1.2) from the proof of Proposition 1 is valid. Using the definition of the function $\mathcal{B}(s)$ from (A.8.11), this means that the derivative of the incumbent payoff $C_{\mathrm{p}}(s)$ with respect to s is:

$$C_{\mathrm{p}}'(s) = \frac{\phi(s)C}{s}\mathcal{B}(s). \tag{A.8.20}$$

The formula (2.17) for real GDP under autarky implies:

$$\frac{C}{s} = \frac{\mathrm{q}^{1-\alpha}\mu^{\alpha}}{(1-\alpha)^{1-\alpha}\alpha^{\alpha}s^{1-\alpha}},$$

which is strictly positive for all $s \in [0,1]$. Since $\phi(s)$ is also strictly positive for all valid s values, it follows from (A.8.20) that:

$$C_{\mathrm{p}}''(s) = \frac{\phi(s)C}{s}\mathcal{B}'(s), \quad \text{if } C_{\mathrm{p}}'(s) = 0. \qquad \text{[A.8.21]}$$

Confirming $C_{\mathrm{p}}(s)$ is a strictly quasi-concave function of s requires showing that $C_{\mathrm{p}}''(s) < 0$ where $C_{\mathrm{p}}'(s) = 0$. Given that $\phi(s)C/s$ is always strictly positive for $s \in [0,1]$, (A.8.20) and (A.8.21) demonstrate that this is equivalent to showing $\mathcal{B}'(s) < 0$ where $\mathcal{B}(s) = 0$.

Using the link between $\mathcal{B}(s)$ and $\mathcal{M}(p)$ in (A.8.12), the derivative of $\mathcal{B}(s)$ is:

$$\mathcal{B}'(s) = \frac{\mathcal{M}'(\lambda^{-1}(s))}{\lambda'(\lambda^{-1}(s))},$$

and thus $\mathcal{B}'(s) < 0$ where $\mathcal{B}(s) = 0$ is equivalent to $\mathcal{M}'(p) < 0$ where $\mathcal{M}(p) = 0$. The latter needs to be verified for p in the range between $p^{\dagger}$ and $\bar{p}$ given that $p^{\dagger} = \lambda^{-1}(0)$ and $\bar{p} = \lambda^{-1}(1)$. Using (A.8.17), the function $\mathcal{M}(p)$ can be expressed as the negative of a quadratic function $\mathcal{Q}(p)$ from (A.8.18) divided by a strictly positive and finite function of p. This means that $\mathcal{M}'(p) < 0$ where $\mathcal{M}(p) = 0$ is in turn equivalent to $\mathcal{Q}'(p) > 0$ where $\mathcal{Q}(p) = 0$, considering values of p between $p^{\dagger}$ and $\bar{p}$.

Inspection of equation (A.8.18) confirms that $\mathcal{Q}(0) < 0$ and $\mathcal{Q}''(p) > 0$ for all parameter values, so it follows that the quadratic equation $\mathcal{Q}(p) = 0$ has one positive and one negative root. The function $\mathcal{Q}(p)$ is negative for p values between 0 and the positive root, and positive for p values above the positive root, which means that the derivative $\mathcal{Q}'(p)$ is positive at the positive root. Since either the positive root lies between $p^{\dagger}$ and $\bar{p}$, or there is no root in this range, it follows that $\mathcal{Q}'(p) > 0$ where $\mathcal{Q}(p) = 0$ for $p \in [p^{\dagger}, \bar{p}]$. This confirms that $C_{\mathrm{p}}(s)$ is a strictly quasi-concave function of s, completing the proof.

A.9 Proof of Proposition 9

Given that the incentive constraint (4.2) is binding (Proposition 7), workers and investors have the same level of utility $U_{\mathrm{w}} = \log C_{\mathrm{w}} = \log C_{\mathrm{k}} - \log(1+\theta)$.

(i) For open economies, real GDP is larger with $s = 1$ than $s = 0$ (Proposition 3), that is, $\bar{C} > C^{\dagger}$. Using (4.8), the consumption of workers is $C_{\mathrm{w}} = \psi_{\mathrm{w}}(p)C$. Since $\psi_{\mathrm{w}}(p)$ is strictly decreasing in p and $\bar{p} > p^{\dagger}$ (Proposition 7), it follows that $\psi_{\mathrm{w}}(\bar{p}) > \psi_{\mathrm{w}}(p^{\dagger})$. Putting these observations together implies $\bar{C}_{\mathrm{w}} > C_{\mathrm{w}}^{\dagger}$, and hence both workers and investors are strictly better off in open economies with $s = 1$ rather than $s = 0$.

(ii) For open economies with $s = 1$, real GDP is higher than under autarky (Proposition 2), that is, $\bar{C} > \hat{C}$. Since $\psi_{\mathrm{w}}(p)$ is strictly decreasing in p and $\bar{p} > \hat{p}$ (Proposition 7), it follows that $\psi_{\mathrm{w}}(\bar{p}) > \psi_{\mathrm{w}}(\hat{p})$. Putting these observations together implies $\bar{C}_{\mathrm{w}} > \hat{C}_{\mathrm{w}}$, and hence both workers and investors are strictly better off in open economies with $s = 1$ than under autarky. As the number of incumbents increases ($\bar{p} > \hat{p}$) and as incumbents receive a higher payoff than workers and capitalists ($\psi_{\mathrm{p}}(p) > \psi_{\mathrm{w}}(p)$), opening up to trade raises the payoffs of all individuals in countries that end up with $s = 1$. This is a Pareto improvement.

(iii) The following result provides the conditions for an allocation to be Pareto efficient and compares these to the conditions for an equilibrium allocation.

Lemma 2 *Consider allocations $\mathscr{A}$ satisfying (2.5), (2.6), (2.10), and (4.3), with (4.2) binding.*

(i) The conditions for an allocation $\mathscr{A}$ to be Pareto efficient are:

$$\frac{c_{\mathrm{pI}}}{c_{\mathrm{pE}}} = \frac{c_{\mathrm{kI}}}{c_{\mathrm{kE}}} = \frac{c_{\mathrm{wI}}}{c_{\mathrm{wE}}}, \quad \frac{\alpha}{1-\alpha}\frac{\mathrm{q} - x_{\mathrm{E}}}{K - x_{\mathrm{I}}} = \pi^*, \quad \textit{and} \qquad \text{[A.9.1a]}$$

$$\alpha \leq \theta c_{\mathrm{wI}} \ \textit{if}\ s = 0, \quad \alpha = \theta c_{\mathrm{wI}} \ \textit{if}\ 0 < s < 1, \quad \textit{and}\ \alpha \geq \theta c_{\mathrm{wI}} \ \textit{if}\ s = 1. \qquad \text{[A.9.1b]}$$

(ii) All equilibrium allocations satisfy (A.9.1a)*, but not necessarily* (A.9.1b)*. An equilibrium allocation featuring $s = 1$ is Pareto efficient; one featuring $0 < s < 1$ is Pareto inefficient; and one featuring $s = 0$ is inefficient if $\tilde{\pi} > \theta Y_{\mathrm{w}}$. The incumbent payoff increases with s if $\tilde{\pi} > (1 + \chi(p))\theta Y_{\mathrm{w}}$, while the condition $\alpha > \theta c_{\mathrm{wI}}$ for it to be inefficient not to increase s is equivalent to $\tilde{\pi} > \theta Y_{\mathrm{w}}$.*

PROOF See appendix A.11. ■

Using the results of Lemma 2, for $\bar{s} = 1$ to be an equilibrium in some countries it must be the case that $\tilde{\pi}^* \geq (1+\chi(\bar{p}))\theta \bar{Y}_{\mathrm{w}}$. Together with $\chi(\bar{p}) > 0$, this implies $\tilde{\pi}^* > \theta \bar{Y}_{\mathrm{w}}$. Since the price of the investment good is the same everywhere in the world, the result $\bar{C}_{\mathrm{w}} > C^{\dagger}_{\mathrm{w}}$ shown above implies that $\bar{Y}_{\mathrm{w}} > Y^{\dagger}_{\mathrm{w}}$. Putting the two inequalities together yields $\tilde{\pi}^* > \theta Y^{\dagger}_{\mathrm{w}}$, and the results of Lemma 2 show that this means the allocation in the despotic country ($s^{\dagger} = 0$) is Pareto inefficient. Lemma 2 also directly shows that the allocation in the rule-of-law economy ($\bar{s} = 1$) is Pareto efficient, and that all countries would have a Pareto inefficient allocation under autarky ($0 < \hat{s} < 1$).

(iv) Consider first the case of autarky. Proposition 1 shows that the equilibrium in all countries is a level of good government satisfying $0 < \hat{s} < 1$, which is associated with a level of power sharing $\hat{p} = \lambda^{-1}(\hat{s})$. Using Proposition 7, this level of power sharing must satisfy $p^{\dagger} < \hat{p} < \bar{p}$. The value of $\hat{s}$ must maximize the incumbent payoff $C_{\mathrm{p}}(s)$, and the proof of quasi-concavity in Proposition 8 confirms that the first-order condition is necessary and sufficient, and that this first-order condition is equivalent to $\hat{p}$ being a root of the quadratic equation $\mathcal{Q}(p) = 0$, where $\mathcal{Q}(p)$ is defined in (A.8.18). The quadratic equation has a unique positive root, which must be equal to $\hat{p}$. Solving the equation in the case $\delta < 1/2$ leads to:

$$\hat{p} = \frac{3\delta - 1 + 4(1-\delta)\alpha + \sqrt{(3\delta - 1 + 4(1-\delta)\alpha)^2 + 8(1-\delta)(\delta + (1-2\delta)\alpha)(1-\alpha)}}{4\delta(2+\delta)(1-\alpha)},$$

and in the $\delta \geq 1/2$ case:

$$\hat{p} = \frac{2\delta^2 + (1+2\delta)\alpha + \sqrt{(2\delta^2 + (1+2\delta)\alpha)^2 + 4\delta^2(1+2\delta)(1-\alpha)}}{2\delta(1+2\delta)(2+\delta)(1-\alpha)}.$$

These equations confirm the expression for $\hat{p}$ given in the proposition.

Proposition 3 shows that with open economies, the world equilibium features a fraction $\tilde{\omega}$ of countries with $s = 1$ and a fraction $1 - \tilde{\omega}$ of countries with $s = 0$. The value of $\tilde{\omega}$ is given in (2.30) in terms of Φ, and Proposition 8 shows that Φ is equal to the expression given in (4.12). The fraction of countries with $s = 1$ is therefore:

$$\tilde{\omega} = \frac{\alpha}{(1-\alpha)\delta(2+\delta)\min\left\{1, \frac{2}{1+2\delta}\right\}(\bar{p} - p^{\dagger})},$$

which satisfies $0 < \tilde{\omega} < 1$. Countries with $s = 0$ have power sharing $p^{\dagger} = \lambda^{-1}(0)$, and those with $s = 1$ have power sharing $\bar{p} = \lambda^{-1}(1)$. The average amount of power sharing around the world is therefore $p^* = (1-\tilde{\omega})p^{\dagger} + \tilde{\omega}\bar{p}$, which must be greater than $p^{\dagger}$. The equation for p^* can be written as $p^* = p^{\dagger} + \tilde{\omega}(\bar{p} - p^{\dagger})$, and by using the formula for $\tilde{\omega}$ above:

$$p^* = p^{\dagger} + \frac{\alpha}{(1-\alpha)\delta(2+\delta)\min\left\{1, \frac{2}{1+2\delta}\right\}} = \frac{1}{2+\delta}\left(1 + \frac{\alpha}{(1-\alpha)\delta\min\left\{1, \frac{2}{1+2\delta}\right\}}\right), \qquad \text{[A.9.2]}$$

which uses $p^{\dagger} = 1/(2+\delta)$ from (4.7). This confirms the expression for p^* given in the proposition.

Note that the formula for p^* from (A.9.2) can be broken down into two cases for $\delta < 1/2$ and $\delta \geq 1/2$:

$$p^* = \begin{cases} \frac{1}{\delta(2+\delta)}\left(\delta + \frac{\alpha}{1-\alpha}\right) & \text{if } \delta < 1/2 \\ \frac{1}{2+\delta}\left(1 + \frac{(1+2\delta)\alpha}{2\delta(1-\alpha)}\right) & \text{if } \delta \geq 1/2 \end{cases}. \qquad \text{[A.9.3]}$$

First consider the case where $\delta < 1/2$. Using the formula for the quadratic $\mathcal{Q}(p)$ in (A.8.18) for this case together with (A.9.3):

$$\begin{aligned}\mathcal{Q}(p^*) &= 2(1-\alpha)\left(\delta+\frac{\alpha}{1-\alpha}\right)^2-(3\delta-1+4(1-\delta)\alpha)\left(\delta+\frac{\alpha}{1-\alpha}\right)-(1-\delta)(\delta+(1-2\delta)\alpha)\\ &= 2\delta^2-2\delta^2\alpha+4\delta\alpha+2\frac{\alpha^2}{1-\alpha}-(3\delta-1)\delta-4\delta(1-\delta)\alpha-(3\delta-1+4(1-\delta)\alpha)\frac{\alpha}{1-\alpha}-\delta(1-\delta)-(1-\delta)(1-2\delta)\alpha\\ &= -\alpha\left(2\delta^2-4\delta-\frac{2\alpha}{1-\alpha}+4\delta(1-\delta)+\frac{3\delta-1}{1-\alpha}+\frac{4(1-\delta)\alpha}{1-\alpha}+(1-\delta)(1-2\delta)\right)\\ &= -\alpha\left(1+\frac{(2-4\delta)\alpha}{1-\alpha}-3\delta+\frac{3\delta-1}{1-\alpha}\right)=-\alpha\left((3\delta-1)\frac{\alpha}{1-\alpha}+(2-4\delta)\frac{\alpha}{1-\alpha}\right)=-(1-\delta)\frac{\alpha^2}{1-\alpha}.\end{aligned} \tag{A.9.4a}$$

Now consider the case where $\delta \geq 1/2$. Again, evaluating the quadratic $\mathcal{Q}(p)$ in (A.8.18) at p^* from (A.9.3):

$$\begin{aligned}\mathcal{Q}(p^*) &= \delta(1+2\delta)(1-\alpha)\left(1+\frac{(1+2\delta)\alpha}{2\delta(1-\alpha)}\right)^2-(2\delta^2+(1+2\delta)\alpha)\left(1+\frac{(1+2\delta)\alpha}{2\delta(1-\alpha)}\right)-\delta\\ &= \delta(1+2\delta)-\delta(1+2\delta)\alpha+(1+2\delta)^2\alpha+\frac{(1+2\delta)^3\alpha^2}{4\delta(1-\alpha)}-2\delta^2-(1+2\delta)\alpha-\frac{\delta(1+2\delta)\alpha}{1-\alpha}-\frac{(1+2\delta)^2\alpha^2}{2\delta(1-\alpha)}-\delta\\ &= -(1+2\delta)\alpha\left(1+\delta-(1+2\delta)+\frac{\delta}{1-\alpha}+\frac{(1+2\delta)(2-(1+2\delta))\alpha}{4\delta(1-\alpha)}\right)\\ &= -(1+2\delta)\alpha\left(\frac{\delta}{1-\alpha}-\delta+\frac{(1+2\delta)(1-2\delta)\alpha}{4\delta(1-\alpha)}\right)=-\frac{(1+2\delta)\alpha}{4\delta}\left(\frac{4\delta^2\alpha}{1-\alpha}+\frac{(1+2\delta-2\delta-4\delta^2)\alpha}{1-\alpha}\right)\\ &= -\frac{(1+2\delta)\alpha^2}{4\delta(1-\alpha)}.\end{aligned} \tag{A.9.4b}$$

Observe from (A.9.4a) and (A.9.4b) that $\mathcal{Q}(p^*) < 0$ in both cases. The autarky level of power sharing $\hat{p}$ satisfies $\mathcal{Q}(\hat{p}) = 0$, and it is shown in the proof of Proposition 8 that $\mathcal{Q}(p)$ is negative to the left of the positive root $\hat{p}$ and positive to the right. This demonstrates that $p^* < \hat{p}$ in all cases, completing the proof.

A.10 Proof of Lemma 1

Consider an allocation established at the post-investment stage of Figure 4. At this point, investment decisions have been made, so the set of investors $\mathcal{I}$ (with measure $|\mathcal{I}| = K = \mu s$) is a state variable. The allocation specifies the individuals $\mathcal{P}$ in power (a set of measure $p = |\mathcal{P}|$), and those individuals $\mathcal{D}$ whose consumption depends on whether they took a past investment opportunity. The sets of capitalists and workers are then $\mathcal{K} = \mathcal{D} \cap \mathcal{I}$ and $\mathcal{W} = [0,1]\backslash(\mathcal{P} \cup \mathcal{K})$ respectively. Note that $\mathcal{K} \subseteq \mathcal{I}$, so the measure of capitalists $k = |\mathcal{K}|$ must satisfy $k \leq \mu s$ (see 2.10). The number of workers is $w = 1 - p - k$. Under the current allocation, incumbents, capitalists, and workers have continuation utilities U_p, U_k, and U_w (ignoring any sunk effort costs of past investment or rebellion) given in (A.7.8).

Let $U(\imath)$ denote the continuation utility of individual $\imath$ under the current allocation, and let $'$ denote an aspect of an allocation that would be established following a rebellion. Following the same steps in the derivation of the no-rebellion condition (A.7.6a) in the proof of Proposition 7, the condition for the absence of a rational rebellion (Definition 1) given subsequent incumbents $\mathcal{P}'$ and subsequent payoffs $U'(\imath)$ is:

$$\int_{\mathcal{P}'} \max\{\exp\{U'_\text{p} - U(\imath)\} - 1, 0\}\mathrm{d}\imath \leq \delta \int_{\mathcal{P}} \mathbb{1}[U_\text{p} > U'(\imath)]\mathrm{d}\imath. \tag{A.10.1}$$

Conjecturing that incumbents want to avoid rebellions and would be worse off if they were to lose power through a rebellion, the condition (A.10.1) becomes a no-rebellion constraint and the same argument from the proof of Proposition 7 shows that the constraint can be stated as follows:

$$\int_{\mathcal{P}'} \max\{\exp\{U'_\text{p} - U(\imath)\} - 1, 0\}\mathrm{d}\imath \leq \delta \left(\int_{\mathcal{P}} \mathrm{d}\imath - \mathbb{1}[U_\text{p} \leq U'_\text{p}] \int_{\mathcal{P}\cap\mathcal{P}'} \mathrm{d}\imath\right). \tag{A.10.2}$$

The same reasoning from the proof of Proposition 7 demonstrates that (A.10.1) is required to hold for all compositions of the subsequent incumbent group $\mathcal{P}'$ with measure equal to some definite $p' = |\mathcal{P}'|$. In equilibrium, p' must satisfy $p' = p$, where p is the equilibrium amount of power sharing, and U'_{p} must satisfy $U'_{\mathrm{p}} = U_{\mathrm{p}}$, where U_{p} is the equilibrium incumbent payoff. Since the capital stock K is a relevant state variable at the post-investment stage, both p' and U'_{p} may depend on K.

Following the same steps as in the derivation of (A.7.7b), the set of no-rebellion constraints (A.10.2) can be stated in terms of the fractions ζ_{p}, ζ_{k}, and ζ_{w} of the subsequent incumbent group drawn from current incumbents, capitalists, and workers respectively. These non-negative numbers must sum to one and satisfy the natural restrictions $\zeta_{\mathrm{p}} \leq p/p'$, $\zeta_{\mathrm{k}} \leq k/p'$, and $\zeta_{\mathrm{w}} \leq (1-p-k)/p'$. The no-rebellion constraints are:

$$\zeta_{\mathrm{w}} p' \max\{\exp\{U'_{\mathrm{p}} - U_{\mathrm{w}}\} - 1, 0\} + \zeta_{\mathrm{k}} p' \max\{\exp\{U'_{\mathrm{p}} - U_{\mathrm{k}}\} - 1, 0\} \\ + \zeta_{\mathrm{p}} p' \mathbb{1}[U_{\mathrm{p}} \leq U'_{\mathrm{p}}] \left(\exp\{U'_{\mathrm{p}} - U_{\mathrm{p}}\} - 1 + \delta\right) \leq \delta p, \quad \text{[A.10.3]}$$

for all feasible ζ_{p}, ζ_{k}, and ζ_{w}.

Free exchange domestically and internationally

The argument is exactly the same as the one developed in the proof of Proposition 7. This demonstrates that the equilibrium allocation of goods across incumbents, capitalists, and workers must be consistent with individuals making individual consumption choices in free markets subject to some levels of disposable income Y_{p}, Y_{k}, and Y_{w} that aggregate to national income Y in (2.8). These are associated with levels C_{p}, C_{k}, and C_{w} of the consumption basket (2.1) that aggregate to real GDP C in (2.9). Net exports x_{E} and x_{I} must also maximize real GDP C in equilibrium, which is consistent with free trade internationally. This means that the economy's resource constraints can be stated as:

$$pY_{\mathrm{p}} + kY_{\mathrm{k}} + (1-p-k)Y_{\mathrm{w}} = Y, \quad \text{or} \quad pC_{\mathrm{p}} + kC_{\mathrm{k}} + (1-p-k)C_{\mathrm{w}} = C. \quad \text{[A.10.4]}$$

A no-rebellion constraint must bind for a rebel faction comprising a positive measure of non-incumbents

Substituting continuation utilities $U(\imath) = \log C(\imath)$ into the set of no-rebellion constraints (A.10.3) leads to:

$$\zeta_{\mathrm{w}} p' \max\{C'_{\mathrm{p}}/C_{\mathrm{w}} - 1, 0\} + \zeta_{\mathrm{k}} p' \max\{C'_{\mathrm{p}}/C_{\mathrm{k}} - 1, 0\} + \zeta_{\mathrm{p}} p' \mathbb{1}[C_{\mathrm{p}} \leq C'_{\mathrm{p}})] \left(C'_{\mathrm{p}}/C_{\mathrm{p}} - 1 + \delta\right) \leq \delta p,$$

for all feasible weights ζ_{p}, ζ_{k}, and ζ_{w}. An allocation cannot satisfy the optimality condition for equilibrium if $C_{\mathrm{w}} > C'_{\mathrm{p}}$ or $C_{\mathrm{k}} > C'_{\mathrm{p}}$ otherwise strictly less consumption could be allocated to workers or capitalists, allowing incumbent consumption to be raised, yet still ensuring that the no-rebellion constraints above hold. Attention can therefore be restricted to allocations specifying $C_{\mathrm{w}} \leq C'_{\mathrm{p}}$ and $C_{\mathrm{k}} \leq C'_{\mathrm{p}}$, which means that the no-rebellion constraints can be stated as follows:

$$\zeta_{\mathrm{w}} p' \left(\frac{C'_{\mathrm{p}}}{C_{\mathrm{w}}} - 1\right) + \zeta_{\mathrm{k}} p' \left(\frac{C'_{\mathrm{p}}}{C_{\mathrm{k}}} - 1\right) + \zeta_{\mathrm{p}} p' \mathbb{1}[C_{\mathrm{p}} \leq C'_{\mathrm{p}}] \left(\frac{C'_{\mathrm{p}}}{C_{\mathrm{p}}} - 1 + \delta\right) \leq \delta p,$$

for all feasible ζ_{p}, ζ_{k}, and ζ_{w}. It is convenient to reformulate these as:

$$(1-\zeta)p' \left(\left(\frac{1-\varkappa}{C_{\mathrm{w}}} + \frac{\varkappa}{C_{\mathrm{k}}}\right) C'_{\mathrm{p}} - 1\right) + \zeta p' \mathbb{1}[C_{\mathrm{p}} \leq C'_{\mathrm{p}}] \left(\frac{C'_{\mathrm{p}}}{C_{\mathrm{p}}} - 1 + \delta\right) \leq \delta p, \quad \text{[A.10.5]}$$

which must hold for all feasible ζ and $\varkappa$, where $\zeta \equiv \zeta_{\mathrm{p}}$ is the fraction of current incumbents included in the post-rebellion incumbent group, and $\varkappa \equiv \zeta_{\mathrm{k}}/(\zeta_{\mathrm{w}} + \zeta_{\mathrm{k}})$ is the fraction of capitalists among current non-incumbents who are included in the post-rebellion incumbent group. Since both p and p' must be less than $1/2$, feasible values of ζ lie between 0 and $\min\{p/p', 1\}$.

The constraint (A.10.5) must bind for some $\zeta < 1$. If not, this would mean the term multiplying $1-\zeta$ is too high, allowing C_{w} or C_{k} to be reduced, which increases the incumbent payoff U_{p}, yet still satisfies all no-rebellion constraints.

Equalization of non-incumbent payoffs: full expropriation of capital

Consider a case where the capital stock K is positive and the allocation implies that the set of capitalists ($\mathcal{K} = \mathcal{D} \cap \mathcal{I}$) is non-empty ($k > 0$). Suppose the allocation were to assign different amounts of consumption to capitalists and workers ($C_{\mathrm{k}} \neq C_{\mathrm{w}}$). Note that (4.5) implies $\mu < 1/2$, so $k < 1/2$, and together with $p \leq 1/2$ this means there are always some workers ($w = 1 - p - k > 0$).

Consider first the case where $C_{\mathrm{k}} > C_{\mathrm{w}}$, which would mean the no-rebellion constraint (A.10.5) with the largest value of the left-hand side would be the one with the smallest possible value of $\varkappa$ (for each ζ). Now suppose that consumption is redistributed equally between workers and capitalists to give both consumption C_{kw} instead:

$$C_{\mathrm{kw}} = \frac{(1-p-k)C_{\mathrm{w}} + kC_{\mathrm{k}}}{1-p},$$

which is feasible given the resource constraint (A.10.4). Since $1/C$ is a convex function, Jensen's inequality implies that:

$$\frac{1}{C_{\mathrm{kw}}} < \left(\frac{1-p-k}{1-p}\right)\frac{1}{C_{\mathrm{w}}} + \left(\frac{k}{1-p}\right)\frac{1}{C_{\mathrm{k}}}.$$

Note that the smallest feasible value of $\varkappa$ (for a given ζ) is always smaller than $k/(1-p)$ because there are $(1-\zeta)p' \leq 1-p$ places for current non-incumbents in the post-rebellion incumbent group, and some workers are available to join this group. Since the smallest feasible $\varkappa$ is such that $\varkappa < k/(1-p)$ and as $1/C_{\mathrm{w}} > 1/C_{\mathrm{k}}$:

$$\left(1 - \frac{k}{1-p}\right)\frac{1}{C_{\mathrm{w}}} + \left(\frac{k}{1-p}\right)\frac{1}{C_{\mathrm{k}}} < \frac{1-\varkappa}{C_{\mathrm{w}}} + \frac{\varkappa}{C_{\mathrm{k}}}.$$

Putting these results together, it follows that:

$$\frac{1}{C_{\mathrm{kw}}} < \frac{1-\varkappa}{C_{\mathrm{w}}} + \frac{\varkappa}{C_{\mathrm{k}}}.$$

If the no-rebellion constraints hold for all values of ζ and $\varkappa$ initially, it follows that all no-rebellion constraints with $\zeta < 1$ are slackened by a redistribution that equalizes the consumption levels of capitalists and workers. Since one such constraint must be binding, consumption inequality between capitalists and workers is not consistent with an allocation being optimal. An exactly analogous argument holds in the case where $C_{\mathrm{w}} > C_{\mathrm{k}}$. This establishes that an equilibrium allocation must feature $C_{\mathrm{k}} = C_{\mathrm{w}}$ if $k > 0$.

Note that with $C_{\mathrm{k}} = C_{\mathrm{w}}$, all non-incumbents share a common level of consumption, so the labelling of individuals as capitalists or workers becomes irrelevant. Hence establishing an allocation with $k = 0$ (by setting $\mathcal{D} = \emptyset$) is optimal, confirming the claim in the proposition. The set of no-rebellion constraints reduces to:

$$(1-\zeta)p'\left(\frac{C'_{\mathrm{p}}}{C_{\mathrm{w}}} - 1\right) + \zeta p' \mathbb{1}[C_{\mathrm{p}} \leq C'_{\mathrm{p}}]\left(\frac{C'_{\mathrm{p}}}{C_{\mathrm{p}}} - 1 + \delta\right) \leq \delta p, \qquad \text{[A.10.6]}$$

for all feasible values of ζ, that is, all $\zeta \in [0, \min\{p/p', 1\}]$.

The equilibrium allocation can be characterized by considering only no-rebellion constraints for workers

The claim is that the equilibrium allocation can be found by solving a simpler problem where it is assumed that rebel factions can only include those not currently in power (formally, Definition 2 is unchanged, but Definition 1 is modified to require that $\mathcal{R}$ does not include any individuals in $\mathcal{P}$). In other words, this means taking account only of the no-rebellion constraint (A.10.6) when $\zeta = 0$:

$$p'\left(\frac{C'_{\mathrm{p}}}{C_{\mathrm{w}}} - 1\right) \leq \delta p. \qquad \text{[A.10.7]}$$

First, let $\{p^{\ddagger}, C^{\ddagger}_{\mathrm{p}}, C^{\ddagger}_{\mathrm{w}}\}$ denote values of these variables under the equilibrium allocation in the simple problem where avoiding rebellion means satisfying (A.10.7). With no change in the state variable K following rebellions, the equilibrium conditions require $p' = p^{\ddagger}$ and $C'_{\mathrm{p}} = C^{\ddagger}_{\mathrm{p}}$. Since (A.10.7) must hold, it follows that $C^{\ddagger}_{\mathrm{p}}/C^{\ddagger}_{\mathrm{w}} \leq 1 + \delta$.

Now consider the optimal allocation in the original problem, taking $p' = p^{\ddagger}$ and $C'_{\mathrm{p}} = C^{\ddagger}_{\mathrm{p}}$ as given. With $p = p^{\ddagger}$, the optimal allocation must satisfy (A.10.6) for all $\zeta \in [0, 1]$. If the equilibrium allocation of the simpler problem is to satisfy these constraints then the following must hold:

$$(1-\zeta)p^{\ddagger}\left(\frac{C_{\mathrm{p}}^{\ddagger}}{C_{\mathrm{w}}^{\ddagger}}-1\right)+\delta\zeta p^{\ddagger}\leq\delta p^{\ddagger},$$

which does indeed follow from $C_{\mathrm{p}}^{\ddagger}/C_{\mathrm{w}}^{\ddagger}\leq 1+\delta$. Since the allocation satisfies all the no-rebellion constraints and maximizes the incumbent payoff subject to the constraint (A.10.7), which is weaker than the general constraint (A.10.6), it follows that this allocation is optimal when subject to the full set of no-rebellion constraints. It is thus an equilibrium of the original problem.

Now consider the converse. Take an allocation that is an equilibrium of the original problem with $\{p^{\ddagger},C_{\mathrm{p}}^{\ddagger},C_{\mathrm{w}}^{\ddagger}\}$. With $p'=p^{\ddagger}$ and $C_{\mathrm{p}}'=C_{\mathrm{p}}^{\ddagger}$, this allocation clearly satisfies the single no-rebellion constraint (A.10.7), which is a special case of (A.10.6). Now suppose it is not optimal when subject only to (A.10.7). This means there is an alternative allocation with $\{p,C_{\mathrm{p}},C_{\mathrm{w}}\}$ satisfying (A.10.7) that yields a higher incumbent payoff $C_{\mathrm{p}}>C_{\mathrm{p}}^{\ddagger}$. Observe that for any $\zeta\in[0,1]$:

$$(1-\zeta)p^{\ddagger}\left(\frac{C_{\mathrm{p}}^{\ddagger}}{C_{\mathrm{w}}}-1\right)+\zeta p^{\ddagger}\mathbb{1}[C_{\mathrm{p}}\leq C_{\mathrm{p}}^{\ddagger}]\left(\frac{C_{\mathrm{p}}^{\ddagger}}{C_{\mathrm{p}}}-1+\delta\right)=(1-\zeta)p^{\ddagger}\left(\frac{C_{\mathrm{p}}^{\ddagger}}{C_{\mathrm{w}}}-1\right)\leq(1-\zeta)\delta p\leq\delta p,$$

where the first equality follows from $C_{\mathrm{p}}>C_{\mathrm{p}}^{\ddagger}$, the first inequality follows from the allocation satisfying (A.10.7), and the second inequality follows from $0\leq 1-\zeta\leq 1$. This shows that the alternative allocation satisfies all the no-rebellion constraints (A.10.6) of the original problem and yields a higher incumbent payoff, contradicting the optimality of the allocation with $\{p^{\ddagger},C_{\mathrm{p}}^{\ddagger},C_{\mathrm{w}}^{\ddagger}\}$. This contradiction demonstrates that the allocation must be optimal subject only to the single no-rebellion constraint (A.10.7), and is thus an equilibrium of the simpler problem. This confirms the claim in the proposition.

Equilibrium allocations subject to the binding no-rebellion constraint for workers

With the equilibrium allocation having no capitalists among non-incumbents ($k=0$), the resource constraint in (A.10.4) implies that incumbents' consumption is:

$$C_{\mathrm{p}}=\frac{C-(1-p)C_{\mathrm{w}}}{p}, \qquad \text{[A.10.8]}$$

where the level of aggregate resources C available for consumption is given in (A.10.4) (which depends only on the predetermined capital stock K). There is only one relevant no-rebellion constraint (A.10.7) (where the post-rebellion incumbent group would comprise only workers), and this constraint must be binding in equilibrium. Rearranging the constraint shows that consumption of workers is given by:

$$C_{\mathrm{w}}=\frac{C_{\mathrm{p}}'}{1+\delta\frac{p}{p'}}. \qquad \text{[A.10.9]}$$

Now define $\psi_{\mathrm{p}}\equiv C_{\mathrm{p}}/C$ and $\psi_{\mathrm{w}}\equiv C_{\mathrm{w}}/C$ to be the per-person shares of consumption of incumbents and workers respectively. Dividing both sides of (A.10.8) by C implies:

$$\psi_{\mathrm{p}}=\frac{1-(1-p)\psi_{\mathrm{w}}}{p}. \qquad \text{[A.10.10]}$$

Given the state variable K, the optimality of domestic free exchange and international free trade already determines the value of C (equation 2.9). It follows that the optimality condition of equilibrium only requires that the remaining aspects of the allocation maximize the incumbent share ψ_{p} subject to the no-rebellion constraint (A.10.9) (the resource constraint has already been accounted for in A.10.10). Using independence of irrelevant history, since any allocation established after a subsequent rebellion would face the same value of the only fundamental state variable K, and since domestic free exchange and international free trade will remain necessary for optimality of those allocations, aggregate resources available for consumption will remain the same after a rebellion ($C'=C=C^{\ddagger}$). Dividing the binding no-rebellion constraint (A.10.10) by this common total amount of resources implies:

$$\psi_{\mathrm{w}} = \frac{\psi'_{\mathrm{p}}}{1 + \delta\frac{p}{p'}}. \qquad \text{[A.10.11]}$$

This can be substituted into the expression for ψ_{p} in (A.10.10) to obtain:

$$\psi_{\mathrm{p}} = \frac{1 - (1-p)\frac{\psi'_{\mathrm{p}}}{1+\delta\frac{p}{p'}}}{p}. \qquad \text{[A.10.12]}$$

Optimality of the equilibrium allocation requires that power sharing p maximizes ψ_{p} above (subject to $0 \leq p \leq 1/2$), taking p' and ψ'_{p} as given, with the independence of irrelevant history conditions $p = p'$ and $\psi_{\mathrm{p}} = \psi'_{\mathrm{p}}$ holding in equilibrium because fundamental state variables (and the level of C) are the same following any further rebellion. Finally, it remains to be confirmed that incumbents who would lose power following a rebellion are willing to join the loyal faction, that is, $C_{\mathrm{p}} > C'_{\mathrm{w}}$, or equivalently, $\psi_{\mathrm{p}} > \psi'_{\mathrm{w}}$.

The first derivative of the expression for ψ_{p} in (A.10.12) with respect to p is:

$$\frac{\partial \psi_{\mathrm{p}}}{\partial p} = \frac{1}{p}\left(\frac{\psi'_{\mathrm{p}}}{1+\delta\frac{p}{p'}}\left(1 + \frac{\frac{\delta(1-p)}{p'}}{1+\delta\frac{p}{p'}}\right) - \frac{1}{p}\left(1 - \frac{(1-p)\psi'_{\mathrm{p}}}{1+\delta\frac{p}{p'}}\right)\right), \qquad \text{[A.10.13]}$$

and the first-order condition is:

$$\frac{1 - (1-p)\frac{\psi'_{\mathrm{p}}}{1+\delta\frac{p}{p'}}}{p} = \left(1 + \frac{\frac{\delta(1-p)}{p'}}{1+\delta\frac{p}{p'}}\right)\frac{\psi'_{\mathrm{p}}}{1+\delta\frac{p}{p'}}. \qquad \text{[A.10.14]}$$

Using (A.10.13), the second derivative of ψ_{p} with respect to p evaluated at a point where the first-order condition (A.10.14) holds is:

$$\left.\frac{\partial^2 \psi_{\mathrm{p}}}{\partial p^2}\right|_{\frac{\partial \psi_{\mathrm{p}}}{\partial p}=0} = -\frac{2\delta}{pp'}\left(1 + \frac{\frac{\delta(1-p)}{p'}}{1+\delta\frac{p}{p'}}\right)\frac{\psi'_{\mathrm{p}}}{\left(1+\delta\frac{p}{p'}\right)^2}.$$

This is unambiguously negative, so it follows that ψ_{p} is a quasi-concave function of p. The first-order condition (A.10.14) is necessary and sufficient for the global maximum (assuming $0 \leq p \leq 1/2$ is satisfied).

Given the equilibrium conditions $p = p' = p^{\ddagger}$ and $\psi_{\mathrm{p}} = \psi_{\mathrm{p}} = \psi^{\ddagger}_{\mathrm{p}}$, the objective function in equation (A.10.12) and the first-order condition in equation (A.10.14) imply:

$$\psi^{\ddagger}_{\mathrm{p}} = \frac{1 - (1-p^{\ddagger})\frac{\psi^{\ddagger}_{\mathrm{p}}}{1+\delta}}{p^{\ddagger}}, \quad \text{and} \qquad \text{[A.10.15a]}$$

$$\frac{1 - (1-p^{\ddagger})\frac{\psi'_{\mathrm{p}}}{1+\delta}}{p^{\ddagger}} = \left(1 + \frac{\frac{\delta(1-p^{\ddagger})}{p^{\ddagger}}}{1+\delta}\right)\frac{\psi^{\ddagger}_{\mathrm{p}}}{1+\delta}. \qquad \text{[A.10.15b]}$$

Substituting (A.10.15a) into (A.10.15b) and cancelling the non-zero term $\psi^{\ddagger}_{\mathrm{p}}$ from both sides leads to:

$$\frac{1}{1+\delta}\left(1 + \frac{\frac{\delta(1-p^{\ddagger})}{p^{\ddagger}}}{1+\delta}\right) = 1,$$

and solving this equation yields an expression for equilibrium power sharing $p^{\ddagger}$:

$$p^{\ddagger} = \frac{1}{2+\delta}.$$

This confirms the expression for $p^{\ddagger}$ given in (A.7.14). Observe that since $\delta > 0$, $0 < p^{\ddagger} < 1/2$, so the constraint $0 \leq p \leq 1/2$ is satisfied. Substituting the expression for $p^{\ddagger}$ back into equation (A.10.15a) leads to a formula for $\psi_{\mathrm{p}}^{\ddagger}$:

$$\psi_{\mathrm{p}}^{\ddagger} = \frac{2+\delta}{2}.$$

In equilibrium, the binding no-rebellion constraint (A.10.11) implies $\psi_{\mathrm{w}}^{\ddagger} = \psi_{\mathrm{p}}^{\ddagger}/(1+\delta)$, and hence

$$\psi_{\mathrm{w}}^{\ddagger} = \frac{2+\delta}{2(1+\delta)}.$$

Note that since $\delta > 0$, it follows immediately that $\psi_{\mathrm{p}} > \psi_{\mathrm{w}}'$ in equilibrium, confirming the earlier conjecture. The expressions for $C_{\mathrm{p}}^{\ddagger}$ and $C_{\mathrm{w}}^{\ddagger}$ in (A.7.14) are verified by noting the definitions of ψ_{p} and ψ_{w}. This completes the proof.

A.11 Proof of Lemma 2

Conditions for an allocation to be efficient

Consider an allocation $\mathscr{A} = \{\mathcal{P}, \mathcal{D}, c_{\mathrm{pE}}, c_{\mathrm{pI}}, c_{\mathrm{kE}}, c_{\mathrm{kI}}, c_{\mathrm{wE}}, c_{\mathrm{wI}}, x_{\mathrm{E}}, x_{\mathrm{I}}\}$ established at the pre-investment stage of Figure 4. If the incentive constraint (4.2) is satisfied then individuals who belong to the set $\mathcal{D}$ will take investment opportunities if they receive one, implying the fraction s of opportunities taken is given by (4.3). Each individual in $\mathcal{D}$ randomly receives an investment opportunity with probability $\nu = \mu/(1-p)$. It follows that an allocation $\mathscr{A}$ implies expected utilities (using 4.1) for individuals at the pre-investment stage that depend only on whether an individual belongs to the sets $\mathcal{P}$, $\mathcal{D}$, or $\mathcal{N} = [0,1]\backslash(\mathcal{P}\cup\mathcal{D})$:

$$U_{\mathrm{p}} = \log C_{\mathrm{p}}, \quad U_{\mathrm{d}} = (1-\nu)\log C_{\mathrm{w}} + \nu\left(\log C_{\mathrm{k}} - \log(1+\theta)\right), \quad \text{and} \;\; U_{\mathrm{n}} = \log C_{\mathrm{w}}. \qquad \text{[A.11.1]}$$

An allocation $\mathscr{A}$ is therefore Pareto efficient if there is no alternative allocation satisfying the resource constraints (2.5), (2.6), (2.10), and the incentive compatibility constraint (4.2) (so that s is determined by 4.3) that leads to a higher value of either U_{p}, U_{d}, or U_{n} while not lowering any of the other expected utilities.

All equilibrium allocations feature a binding incentive constraint (4.2) according to Proposition 7, so attention is restricted to the requirements for efficiency for such allocations. Note that a binding incentive constraint (4.2) implies the expected utilities in (A.11.1) will satisfy $U_{\mathrm{d}} = U_{\mathrm{n}}$. The conditions for efficiency can therefore be derived by maximizing the welfare function:

$$\mathscr{W} = p\Omega U_{\mathrm{p}} + dU_{\mathrm{d}} + nU_{\mathrm{n}}, \qquad \text{[A.11.2]}$$

where individuals in $\mathcal{D}$ and $\mathcal{N}$ receive the same weight per person (normalized to one) in the welfare function, and Ω denotes the relative weight per person of individuals in $\mathcal{P}$ (with $\Omega > 0$). The variables p, d, and $n = 1-p-d$ denote the measures of the sets $\mathcal{P}$, $\mathcal{D}$, and $\mathcal{N}$. When the incentive constraint (4.2) is satisfied, $s = d/(1-p)$, hence $d = (1-p)s$ and $n = (1-p)(1-s)$. Using (A.11.1) and $\nu = \mu/(1-p)$, the welfare function (A.11.2) becomes:

$$\begin{aligned}\mathscr{W} &= p\Omega\log C_{\mathrm{p}} + (1-p)s\left(\left(1-\frac{\mu}{1-p}\right)\log C_{\mathrm{w}} + \frac{\mu}{1-p}\left(\log C_{\mathrm{k}} - \log(1+\theta)\right)\right) + (1-p)(1-s)\log C_{\mathrm{w}} \\ &= p\Omega\log C_{\mathrm{p}} + (1-p)\log C_{\mathrm{w}} + \mu s\left(\log C_{\mathrm{k}} - \log(1+\theta) - \log C_{\mathrm{w}}\right). \qquad \text{[A.11.3]}\end{aligned}$$

Given the value of s, the measure of capitalists is μs and the measure of workers is $1-p-\mu s$. The resource constraints (2.6) can therefore be written as:

$$pc_{\mathrm{pE}} + \mu s c_{\mathrm{kE}} + (1-p-\mu s)c_{\mathrm{wE}} = \mathrm{q} - x_{\mathrm{E}}, \quad \text{and} \;\; pc_{\mathrm{pI}} + \mu s c_{\mathrm{kI}} + (1-p-\mu s)c_{\mathrm{wI}} = K - x_{\mathrm{I}}. \qquad \text{[A.11.4]}$$

According to (4.3), the incentive constraint (4.2) is only required to hold if $s > 0$, so this constraint can be equivalently represented by:

$$\mu s\left((1-\alpha)\log c_{\mathrm{kE}} + \alpha\log c_{\mathrm{kI}} - (1-\alpha)\log c_{\mathrm{wE}} - \alpha\log c_{\mathrm{wI}} - \log(1+\theta)\right) \geq 0, \qquad \text{[A.11.5]}$$

which uses the formula for the consumption aggregator in (2.1).

The welfare function (A.11.3) can be maximized subject to the constraints (2.5), (2.10), (A.11.4), and

(A.11.5) by setting up the Lagrangian function:

$$\begin{aligned}\mathscr{L} &= p\Omega\left((1-\alpha)\log c_{\mathrm{pE}} + \alpha\log c_{\mathrm{pI}} - (1-\alpha)\log(1-\alpha) - \alpha\log\alpha\right)\\ &+ (1-p)\left((1-\alpha)\log c_{\mathrm{wE}} + \alpha\log c_{\mathrm{wI}} - (1-\alpha)\log(1-\alpha) - \alpha\log\alpha\right)\\ &+ \mu s\left((1-\alpha)\log c_{\mathrm{kE}} + \alpha\log c_{\mathrm{kI}} - (1-\alpha)\log c_{\mathrm{wE}} - \alpha\log c_{\mathrm{wI}} - \log(1+\theta)\right)\\ &+ \Lambda_{\mathrm{K}}(\mu s - K) + \Lambda_{\mathrm{T}}(x_{\mathrm{E}} + \pi^* x_{\mathrm{I}}) + \Lambda_{\mathrm{E}}(\mathrm{q} - x_{\mathrm{E}} - pc_{\mathrm{pE}} - \mu s c_{\mathrm{kE}} - (1-p-\mu s)c_{\mathrm{wE}})\\ &+ \Lambda_{\mathrm{I}}(K - x_{\mathrm{I}} - pc_{\mathrm{pI}} - \mu s c_{\mathrm{kI}} - (1-p-\mu s)c_{\mathrm{wI}})\\ &\qquad + \mu s\Lambda_{\mathrm{D}}((1-\alpha)\log c_{\mathrm{kE}} + \alpha\log c_{\mathrm{kI}} - (1-\alpha)\log c_{\mathrm{wE}} - \alpha\log c_{\mathrm{wI}} - \log(1+\theta)),\end{aligned} \quad [\text{A.11.6}]$$

where the consumption basket equation (2.1) has been substituted into (A.11.3). The Lagrangian multipliers on the constraints (2.10), (2.5), (A.11.4), and (A.11.5) are denoted by Λ_{K}, Λ_{T}, Λ_{E}, Λ_{I}, and Λ_{D} respectively. The first-order conditions with respect to the consumption allocation $\{c_{\mathrm{pE}}, c_{\mathrm{pI}}, c_{\mathrm{kE}}, c_{\mathrm{kI}}, c_{\mathrm{wE}}, c_{\mathrm{wI}}\}$ are:

$$p\Omega\frac{(1-\alpha)}{c_{\mathrm{pE}}} = p\Lambda_{\mathrm{E}}, \quad p\Omega\frac{\alpha}{c_{\mathrm{pI}}} = p\Lambda_{\mathrm{I}}, \quad \mu s(1+\Lambda_{\mathrm{D}})\frac{(1-\alpha)}{c_{\mathrm{kE}}} = \mu s\Lambda_{\mathrm{E}}, \quad \mu s(1+\Lambda_{\mathrm{D}})\frac{\alpha}{c_{\mathrm{kI}}} = \mu s\Lambda_{\mathrm{I}},$$
$$\left(\frac{(1-p-\mu s)-\mu s\Lambda_{\mathrm{D}}}{1-p-\mu s}\right)\frac{(1-\alpha)}{c_{\mathrm{wE}}} = \Lambda_{\mathrm{E}}, \quad \text{and} \quad \left(\frac{(1-p-\mu s)-\mu s\Lambda_{\mathrm{D}}}{1-p-\mu s}\right)\frac{\alpha}{c_{\mathrm{wI}}} = \Lambda_{\mathrm{I}}, \quad [\text{A.11.7}]$$

noting that $1-p-\mu s > 0$. The first-order conditions with respect to the capital stock K and net exports $\{x_{\mathrm{E}}, x_{\mathrm{I}}\}$ are:

$$\Lambda_{\mathrm{K}} = \Lambda_{\mathrm{I}}, \quad \Lambda_{\mathrm{E}} = \Lambda_{\mathrm{T}}, \quad \text{and} \quad \Lambda_{\mathrm{I}} = \pi^*\Lambda_{\mathrm{T}}. \quad [\text{A.11.8}]$$

The derivative of the Lagrangian (A.11.6) with respect to s is:

$$\begin{aligned}\frac{\partial\mathscr{L}}{\partial s} &= \mu\left(\Lambda_{\mathrm{K}} - (c_{\mathrm{kE}} - c_{\mathrm{wE}})\Lambda_{\mathrm{E}} - (c_{\mathrm{kI}} - c_{\mathrm{wI}})\Lambda_{\mathrm{I}}\right)\\ &\qquad + \mu(1+\Lambda_{\mathrm{D}})((1-\alpha)\log c_{\mathrm{kE}} + \alpha\log c_{\mathrm{kI}} - (1-\alpha)\log c_{\mathrm{wE}} - \alpha\log c_{\mathrm{wI}} - \log(1+\theta)),\end{aligned} \quad [\text{A.11.9}]$$

and the first-order condition requires this to be non-positive if $s = 0$, zero if $0 < s < 1$, and non-negative if $s = 1$. Maximizing the welfare function also requires:

$$\Lambda_{\mathrm{D}} \geq 0, \quad \text{and} \quad \mu s\Lambda_{\mathrm{D}}\left((1-\alpha)\log c_{\mathrm{kE}} + \alpha\log c_{\mathrm{kI}} - (1-\alpha)\log c_{\mathrm{wE}} - \alpha\log c_{\mathrm{wI}} - \log(1+\theta)\right) \geq 0, \quad [\text{A.11.10}]$$

because the constraint (A.11.5) is an inequality.

Suppose that the constraint (A.11.5) were slack. It follows that $s > 0$, and by using (A.11.10), $\Lambda_{\mathrm{D}} = 0$. Substituting these into (A.11.7) implies:

$$c_{\mathrm{kE}} = \frac{1-\alpha}{\Lambda_{\mathrm{E}}}, \quad c_{\mathrm{kI}} = \frac{\alpha}{\Lambda_{\mathrm{I}}}, \quad c_{\mathrm{wE}} = \frac{1-\alpha}{\Lambda_{\mathrm{E}}}, \quad \text{and} \quad c_{\mathrm{wI}} = \frac{\alpha}{\Lambda_{\mathrm{I}}},$$

however, substituting the expressions for consumption into (A.11.5) shows that the constraint would be violated. Therefore, the constraint (A.11.5) must be binding:

$$\mu s\left((1-\alpha)\log c_{\mathrm{kE}} + \alpha\log c_{\mathrm{kI}} - (1-\alpha)\log c_{\mathrm{wE}} - \alpha\log c_{\mathrm{wI}} - \log(1+\theta)\right) = 0. \quad [\text{A.11.11}]$$

With $s = 0$, note that c_{kE} and c_{kI} are irrelevant (any values satisfy A.11.7). When $s > 0$, equation (A.11.7) implies the welfare-maximizing consumption allocation must satisfy:

$$c_{\mathrm{kE}} = \frac{(1-\alpha)(1+\Lambda_{\mathrm{D}})}{\Lambda_{\mathrm{E}}}, \quad c_{\mathrm{kI}} = \frac{\alpha(1+\Lambda_{\mathrm{D}})}{\Lambda_{\mathrm{I}}}, \quad c_{\mathrm{wE}} = \frac{(1-\alpha)((1-p-\mu s)-\mu s\Lambda_{\mathrm{D}})}{(1-p-\mu s)\Lambda_{\mathrm{E}}},$$
$$\text{and} \quad c_{\mathrm{wI}} = \frac{\alpha((1-p-\mu s)-\mu s\Lambda_{\mathrm{D}})}{(1-p-\mu s)\Lambda_{\mathrm{I}}}, \quad [\text{A.11.12}]$$

noting that this is also welfare-maximizing for $s = 0$ (and hence the equation holds for any s). Suppose equation (A.11.11) holds for both $s = 0$ and $s > 0$, that is:

$$(1-\alpha)\log c_{\mathrm{kE}} + \alpha\log c_{\mathrm{kI}} - (1-\alpha)\log c_{\mathrm{wE}} - \alpha\log c_{\mathrm{wI}} - \log(1+\theta) = 0. \quad [\text{A.11.13}]$$

Using the consumption basket formula (2.1), this implies:

$$C_k = (1+\theta)C_w, \qquad [A.11.14]$$

so (4.2) is binding. Substituting from (A.11.12) into the above shows that the Lagrangian multiplier Λ_D satisfies the equation:

$$1 + \Lambda_D = (1+\theta)\left(\frac{(1-p-\mu s) - \mu s\Lambda_D}{1-p-\mu s}\right),$$

which can be rearranged to deduce:

$$\Lambda_D = \frac{\theta(1-p-\mu s)}{1-p+\mu\theta s}. \qquad [A.11.15]$$

Since $\theta > 0$, $p \leq 1/2$, and $\mu \leq 1/2$ (implied by 4.5), the value of Λ_D is strictly positive, satisfying both conditions in (A.11.10) (given A.11.13). Note that the expression for Λ_D in (A.11.15) implies:

$$1 + \Lambda_D = \frac{(1+\theta)(1-p)}{1-p+\mu\theta s}, \quad \text{and} \quad \frac{(1-p-\mu s) - \mu s\Lambda_D}{1-p-\mu s} = \frac{1-p}{1-p+\mu\theta s}.$$

These are both strictly positive, which in combination with (A.11.7) and (A.11.8) confirms that all of the Lagrangian multipliers Λ_K, Λ_T, Λ_E, and Λ_I are strictly positive. Substituting into (A.11.7) shows that the welfare-maximizing consumption allocation must satisfy:

$$c_{pE} = \frac{(1-\alpha)\Omega}{\Lambda_E}, \quad c_{pI} = \frac{\alpha\Omega}{\Lambda_I}, \quad c_{kE} = \frac{(1-\alpha)(1+\theta)(1-p)}{(1-p+\mu\theta s)\Lambda_E}, \quad c_{kI} = \frac{\alpha(1+\theta)(1-p)}{(1-p+\mu\theta s)\Lambda_I},$$
$$c_{wE} = \frac{(1-\alpha)(1-p)}{(1-p+\mu\theta s)\Lambda_E}, \quad \text{and} \quad c_{wI} = \frac{\alpha(1-p)}{(1-p+\mu\theta s)\Lambda_I}. \qquad [A.11.16]$$

Solutions for the Lagrangian multipliers can be obtained by substituting these expressions into the resource constraints (A.11.4):

$$\Lambda_E = \Lambda_T = \frac{(1-\alpha)(p\Omega + 1 - p)}{\mathrm{q} - x_E}, \quad \text{and} \quad \Lambda_I = \Lambda_K = \frac{\alpha(p\Omega + 1 - p)}{K - x_I}. \qquad [A.11.17]$$

Finally, note that no first-order condition with respect to power sharing p needs to be considered because given the consumption allocation (A.11.16), a change in p cannot make an individual better off without implying that some other individual is worse off. Power sharing p is thus irrelevant for efficiency when investment s is only constrained by incentive compatibility (4.2).

Equations (A.11.16) and (A.11.17) imply that the consumption allocation must satisfy:

$$\frac{c_{pI}}{c_{pE}} = \frac{c_{kI}}{c_{kE}} = \frac{c_{wI}}{c_{wE}} = \frac{\alpha}{1-\alpha}\frac{\Lambda_E}{\Lambda_I} = \frac{K - x_I}{\mathrm{q} - x_E}. \qquad [A.11.18]$$

These equations must hold for all values of the Pareto weight Ω, so they are necessary conditions for an efficient allocation. Equations (A.11.8) and (A.11.18) also imply:

$$\frac{\alpha}{1-\alpha}\frac{\mathrm{q} - x_E}{K - x_I} = \pi^*, \qquad [A.11.19]$$

which must hold for all Ω, so this is also a necessary condition for an efficient allocation. The allocation (A.11.16) features $c_{kE} = (1+\theta)c_{wE}$ and $c_{kI} = (1+\theta)c_{wI}$, which confirms a binding incentive constraint (A.11.14) is consistent with efficiency. By using these equations in combination with (A.11.8) and (A.11.13), the expression in (A.11.9) for the derivative of the Lagrangian with respect to s becomes:

$$\frac{\partial \mathscr{L}}{\partial s} = \mu\left(\Lambda_I - \theta(\Lambda_E c_{wE} + \Lambda_I c_{wI})\right) = \mu\Lambda_I\left(1 - \theta\left(\frac{\Lambda_E}{\Lambda_I}c_{wE} + c_{wI}\right)\right).$$

Using (A.11.18), the partial derivative is equal to:

$$\frac{\partial \mathscr{L}}{\partial s} = \mu\Lambda_I\left(1 - \theta\left(\frac{1-\alpha}{\alpha} + 1\right)c_{wI}\right) = \frac{\mu\Lambda_I}{\alpha}(\alpha - \theta c_{wI}).$$

Since Λ_I is strictly positive according to (A.11.17), the derivative above is positive if $\alpha > \theta c_{wI}$, zero if

$\alpha = \theta c_{\text{wI}}$, and negative if $\alpha < \theta c_{\text{wI}}$. Note that these conditions are valid for all Ω. The fraction s must lie in the unit interval, therefore the conditions for efficiency are:

$$\alpha \leq \theta c_{\text{wI}} \;\text{ if }\; s = 0, \quad \alpha = \theta c_{\text{wI}} \;\text{ if }\; 0 < s < 1, \quad \text{and} \;\; \alpha \geq \theta c_{\text{wI}} \;\text{ if }\; s = 1. \tag{A.11.20}$$

The consumption allocation in (A.11.16) implies

$$\frac{c_{\text{pE}}}{c_{\text{wE}}} = \frac{c_{\text{pI}}}{c_{\text{wI}}} = \frac{(1 - p + \mu\theta s)\Omega}{1 - p}, \quad \text{and hence} \;\; \frac{C_{\text{p}}}{C_{\text{w}}} = \frac{(1 - p + \mu\theta s)\Omega}{1 - p}.$$

Furthermore, the binding incentive constraint (A.11.14) implies all non-incumbents receive the same utility $U_{\text{d}} = U_{\text{n}}$, consistent with the equal weighting of all non-incumbents in the welfare function (A.11.3). It follows that an allocation where incumbents receive rents β, so that $C_{\text{p}} = (1 + \beta)C_{\text{w}}$, would be a solution of the welfare maximization problem (assuming all the efficiency conditions derived above are satisfied) for a weight Ω on each incumbent given by:

$$\Omega = \frac{(1 - p)(1 + \beta)}{1 - p + \mu\theta s}.$$

This means that incumbents receiving rents is not inefficient as long as equations (A.11.18), (A.11.19), and (A.11.20) are satisfied. Therefore, the requirements for an allocation to be efficient are those given in (A.9.1), confirming the claim in the proposition.

Equilibrium and efficiency

Since $\mu\psi_{\text{p}}(p)/\tilde{\pi}^{\alpha}$ is strictly positive, (4.9) shows that $C_{\text{p}}(s)$ is increasing in s if $\tilde{\pi} > (1 + \chi(p))\theta Y_{\text{w}}$. The conditions in (A.9.1b) demonstrate that a value of s is inefficiently low if $\alpha > \theta c_{\text{wI}}$. Since Proposition 7 states that equilibrium allocations must be consistent with free exchange, (2.4) implies workers' consumption of the investment good is $c_{\text{wI}} = \alpha Y_{\text{w}}/\tilde{\pi}$, where $\tilde{\pi}$ is the market-clearing relative price from (2.7). Hence, $\alpha > \theta c_{\text{wI}}$ is equivalent to:

$$\alpha > \theta \left(\frac{\alpha Y_{\text{w}}}{\tilde{\pi}} \right).$$

This simplifies to $\tilde{\pi} > \theta Y_{\text{w}}$, confirming the claim in the proposition.

Proposition 7 demonstrates that any equilibrium allocation must be consistent with free exchange of goods domestically and free international trade. Since (2.3) holds for all individuals, the ratios $c_{\text{pI}}/c_{\text{pE}}$, $c_{\text{kI}}/c_{\text{kE}}$, and $c_{\text{wI}}/c_{\text{wE}}$ are all identical and equal to $\alpha/((1 - \alpha)\tilde{\pi})$, where $\tilde{\pi}$ is the market-clearing relative price from (2.7). Furthermore, free international trade $\tau = 0$ implies $\pi^* = (\alpha(\text{q} - x_{\text{E}}))/((1 - \alpha)(K - x_{\text{I}}))$ using (2.7) and (2.21). Therefore, all the efficiency conditions in (A.9.1a) must hold for any equilibrium allocation.

The value of s in an equilibrium allocation must maximize the incumbent payoff function $C_{\text{p}}(s)$ (Proposition 7), implying the following first-order conditions for s:

$$C'_{\text{p}}(s) \leq 0 \;\text{ if }\; s = 0, \quad C'_{\text{p}}(s) = 0 \;\text{ if }\; 0 < s < 1, \quad \text{and} \;\; C'_{\text{p}}(s) \geq 0 \;\text{ if }\; s = 1.$$

Suppose the equilibrium allocation features $s = 1$. Using the above and (4.9) it follows that $\tilde{\pi} \geq (1 + \chi(p))\theta Y_{\text{w}}$. Since $\chi(p) > 0$ this implies $\tilde{\pi} > \theta Y_{\text{w}}$, confirming that the equilibrium allocation with $s = 1$ is efficient. Now suppose the equilibrium allocation features $0 < s < 1$, in which case (4.9) shows that $\tilde{\pi} = (1 + \chi(p))\theta Y_{\text{w}}$ is required. As $\chi(p)$ is strictly positive, this implies $\tilde{\pi} > \theta Y_{\text{w}}$, revealing that the equilibrium allocation with $0 < s < 1$ is inefficient. Finally, for an equilibrium allocation featuring $s = 0$, the allocation would be inefficient if $\alpha > \theta Y_{\text{w}}$, which has been shown to be equivalent to $\tilde{\pi} > \theta Y_{\text{w}}$. This completes the proof.

B Data appendix

The empirical analysis in section 5 uses data from the Polity IV Project on 'Executive Constraints' (XCONST, a score between 1 and 7). Four groups of countries are considered: Britain, the 'European core', Latin America, and Asia. Table 1 lists the countries included in each group and first year data are available for each country (and the treatment of the data where countries have unified or changed names). A simple average of the executive constraints score is calculated in each year for each group using data covering the period 1800–1913 (with the compositions of groups depending on which countries have data available).

Table 1: *Countries with Polity IV data used in the empirical analysis (1800–1913)*

Country groups	Data available	Remarks
Britain	1800	
'European core'		
Austria	1800	Austria-Hungary from 1867 onwards
Belgium	1830	
France	1800	
Germany	1800	Prussia from 1800 to 1867
Italy	1815	Average of Modena, Parma, Papal States, Sardinia, Two Sicillies, and Tuscany between 1815 and 1860
Netherlands	1815	
Spain	1800	
Sweden	1800	
Switzerland	1848	
Latin America		
Argentina	1825	
Bolivia	1825	
Brazil	1824	
Chile	1818	
Colombia	1821	Gran Colombia between 1821 and 1830
Ecuador	1830	
Haiti	1820	
Mexico	1822	
Paraguay	1811	
Peru	1821	
Uruguay	1830	
Venezuela	1830	
Asia		
Afghanistan	1800	
China	1800	
Iran	1800	Persia
Thailand	1800	Siam
Turkey	1800	

Source: Polity IV Project, Center for Systemic Peace (http://www.systemicpeace.org/inscrdata.html).