Engineering Institutional Change

Sayantan Ghosal*, Anna Malova†, Sharun W. Mukand‡

Abstract

In this paper, we consider a problem of practical mechanism design: how does an international policy-maker create a policy with the aim of exploiting a national government's domestic constraints and catalyse institutional change (e.g. entering an international agreement, trade organisation) within an accelerated time frame? With limited transferability of payoffs and mobility frictions, we show that institutional change can be engineered even when political leaders are opposed to the adoption of the new institution. The approaches are: (a) take advantage of dynamic inconsistency: identify/construct a "coalition " of agents who either currently or in the future will benefit from the change by small shifts in policy that have the effect of changing the 'facts on the ground' (b) leverage network externalities by targeting 'pivotal' set of countries.

JEL classification: D72; H11; O19; F55

Keywords: Institutions; Appropriate institutions; Voting Behaviour; Political Economy; Economic development; Networks

1 Introduction

"Laws and institutions must go hand in hand with the progress of the human mind... with the change of circumstances, institutions must advance also, and keep pace with the times."(Jefferson 1816)

Institutions tend to persist and do not "keep pace with the times".

Persistence of inefficient institutions is a central issue in discussions of contemporary political economy, especially in the context of entering/exiting international agreements, for instance trade protectionism (the unwillingness to dismantle existing tariff barriers or the incentives to impose new tariff barriers) or climate change (the

*University of Glasgow, email: Sayantan.Ghosal@glasgow.ac.uk
†University of Glasgow, email: a.malova.1@research.gla.ac.uk
‡University of Warwick, email: S.Mukand@warwick.ac.uk

inability of national governments to commit to early adoption of low carbon technologies). Persistent inefficient institutions are also a central issue in the political economy of development (Acemoglu and Robinson 2013). Institutions are central for countries' growth and development. History matters for much of the development because of institutional persistence (North 1991).

How do we transform/replace inefficient institutions? One view suggests that structural conditions are all that matter: "Men make their own history, but they do not make it as they please; they do not make it under self-selected circumstances, but under circumstances existing already, given and transmitted from the past" (Marx 1852). Change is driven by long-term evolution of technology, control of resources or threat of revolution (Acemoglu and Robinson 2013).

From this perspective, institutional change is far too gradual and subject to the vagaries of history.

In this paper, in contrast, we adopt the viewpoint of practical mechanism design. How can inefficient institutions be replaced when national governments have no political interest in doing so, i.e. national governments have a stake in perpetuating the status quo? From Shapley and Shubik (1969) we know that when negative externalities are present, core can be empty. Hence adoption of efficient institutions need not always hold, even allowing for unlimited transferability of payoffs. Consequently, throughout this paper, we assume limited or no transferability of payoffs between agents within countries, between time periods, across national boundaries.

Under such conditions, how does an international policy-maker who internalizes relevant global externalities devise a policy with the aim of exploiting national government's domestic constraints and catalyse institutional change in an accelerated time frame?

We model a continuum of voters who are split into two unevenly sized groups. At each time period a country can be in one of the two political states (corresponding to different institutions). The majority group of voters in the beginning of the game prefers socially inefficient state since this is when they enjoy the highest social rent; the minority prefers an efficient alternative. Hence, the status quo option is the one where inefficient institutions prevail. Political state can change at each of the three periods of the game provided that voters choose an alternative political state over the status quo. At the beginning of each period, an incumbent politician makes a policy proposal. She can be challenged by other politicians, but there is a cost to that. As long as she proposes the programme which satisfies the majority, she will be re-elected; otherwise she will be challenged and will lose the office with probability one. In our model, due to frictions that impede intergroup mobility, a representative voter in the majority group may not vote for the status quo. Depending on the magnitude of the mobility friction, we show that, in equilibrium, either the efficient institution is adopted early or it is never adopted. The latter corresponds to inefficient institutions persistence.

In settings where inefficient institutions persist, we examine three mechanisms

that ensure the adoption of efficient institutions. All three involve engineering the majority in favour of the socially preferred alternative while taking as given the level of the domestic mobility friction.

First, when domestic mobility frictions are not too large, an outside agent (an international policymaker who internalizes global externalities) could ensure that a carefully chosen subset of the group moves with probability one. This approach is in essence similar to creating a new political elite which is interested in abandoning the status quo and secures early adoption of efficient institutions. It can be applied to the formation of international treaties such Kyoto Protocol, Montreal Protocol, Paris Agreement: identifying a specific industry (or firms within an industry) and targeting it with subsidies to ensure a switch to a low-carbon technology. Among these – UK's "plug-in grant" and Norway's incentives scheme for electric vehicles, to name some.

When domestic mobility frictions are stronger, an outside agent could choose a fraction of voters whose volunatury but costly movement between the groups precedes the shift in the political state, effectively creating a committment device for the future change in institutions. This way, there exists a group of voters who will hold an incumbent's feet to the fire if she proposes an inefficient policy: whether she loses the power, or she wins the election, either way, efficient institutions will be implemented. This would ensure that in the next period farsighted voters, anticipating the benefits of choosing the alternative, would initiate an institutional change. This approach establishes delayed adoption of efficient institutions.

The third approach explores the opportunities of network leveraging. In an interconnected global economy, every country is connected to other countries through a network, be it for trade, international aid or any other purposes. Thus, country's payoff from being in a certain political state may depend on the number of other countries in the same state. For instance, nations who wish to explore the potential of wind energy can share research and develpoment costs with other nations who are also laying eyes on wind as a substitute for conventional sources of energy. As a prerequisite, however, we must allow for technological spillovers, which might not be an easy task (Bayer and Urpelainen 2013). An outside agent can manipulate the payoffs of the pivotal country to ensure that inefficient status quo is no longer a subgame perfect equilibrium. Consequently, voters in this country will prefer an alternative political state, which will lower the payoffs from adhering to status quo for the rest of the countries and induce them to abandon it in favour of the alternative. Examples can include subsidies for German car manufacturers to produce electric vehicles.

Our results suggest that, exploiting only a limited amount of resources, we can ensure a switch to socially efficient institutions not just in one country, but in a network of them.

2 Related Literature

Several main issues delineate the scope of the problem studied in the current paper. Institutions, as defined by North (1991), can be a constraint for economic and political problems faced by society; institutions can be inefficient; institutions tend to persist, thereby defining and limiting the options of future trajectories for society's development.

Douglass North (1991, p.97) defined institutions as "the humanly devised constraints that structure political, economic and social interaction". Therefore, institutions determine the set of tools available to a society to effectively resolve its issues. Some institutions are correlated with better economic performance, such as better enforcement of property rights of a large fraction of the citizens and comparative equality before the law (North and Weingast 1989; Acemoglu et al. 2001). Extractive institutions are generally associated with poor economic performance and noticeably lower levels of economic development, although they still can emerge in equilibrium (Binger and Hoffman 1989; Acemoglu et al. 2002; Acemoglu and Robinson 2013). Less developed countries with institutions of poor quality may not benefit from trade, provided that trade and institutions interact bi-directionally (Segura-Cayuela 2006). This raises the question of institutional efficiency. How does one measure it, which institutions are more efficient, what prevents institutional evolution?

Discussion about institutional efficiency and how it relates to national economic performance dates back to Adam Smith (1963). Since then, a bewildering array of ideas has been introduced to public scrutiny about what affects economic performance. Among the candidates for the determinants of growth are geography and factor endowments (Diamond 1997), trade (Frankel and Romer 1999; Dollar and Kraay 2003), human capital (Glaeser et al. 2004), national leaders (Jones and Olken 2005), and, certainly, institutions (Acemoglu et al. 2001; Góes 2016). The preponderance of evidence suggests that the latter may be the largest contributor to economic success (Rosenberg et al. 1986; Hall and Jones 1999; Rodrik et al. 2004; Acemoglu et al. 2014), but institutions have to be efficient.

Traditionally, adequate enforcement of property rights for a large fraction of the society and equality before the law are mentioned when discussing economically efficient institutions (North and Weingast 1989; Knack and Keefer 1995; Hall and Jones 1999; Acemoglu et al. 2001). In the present research, we abstract from these ideas and slightly abuse the term 'institutions' by limiting it to underlying determinants of the payoff structure. Therefore, the notion of efficiency in this context is confined to purely social welfare account. As emphasized in Acemoglu et al. (2002), equilibrium institutions may be extractive and not possess conditions for economic growth. In our paper, we do not discuss how effective our newly established institutions are for long-run growth, but we are certain that they maximize social welfare at present.

Due to the vagueness of our use of the term "institutions", one can locally think of the process of institutional change as being synonymous to a political/economic reform, and hereafter, we will use terms reform, institution and policy interchangeably.

The question we set out to tackle is best illustrated by Fernandez and Rodrik (1991), although they describe it in terms of reform rather than institution. In their work, ex post beneficial reform does not carry the day ex ante due to uncertainty at the individual level about future gains and losses. This is the culprit that prevents the shift in our model as well, but rather than focusing on modelling such a political conundrum, we concentrate on the practical ways to promote an efficiency-enhancing reform in a democratic society with free and fair, simple majority elections. We thus introduce an international benevolent agent who, unlike in other recent similar models (Galiani et al. 2019), does not serve as a provider of funds to compensate the members of the coalition that bears all the costs, but eliminates the individual-specific uncertainty about the ex-post indentities of winners and losers. She does so by revealing ex-ante who will be drawn from the losing to the winning group, so that these voters would be willing to abandon the status quo and create the majority in favour of the alternative.

Inefficient policies are infamous for their ability to persist for prolonged periods of time, even when a more efficient alternative is just around the corner. This may be due to already mentioned individual-specific uncertainty (Fernandez and Rodrik 1991), political failure to recognize cost of adjustment to the new policy as sunk which causes even more extensive support in the future (Coate and Morris 1999), or a holdup problem (Espín-Sánchez 2017). Inefficient water allocation in the cities of Lorca and Mula that are today parts of Spain was stealing in welfare from the people for more than 700 years (Espín-Sánchez 2017). Land policy in the New World may have prolonged higher extent of inequality (Sokoloff and Engerman 2000). Dealing with feeble institutions can be a dubious enterprise, especially if politicians are dynamically and/or time-inconsistent (Harstad 2016). Conventional wisdom, in this case, holds that inefficient institutions and political instruments are used strategically to the benefit of a ruling politician with little or no account for the inconsistency (Alesina and Tabellini 1990; Battaglini and Harstad 2016). In certain conditions, thus, autocratic regimes may do better, but not in others as they may, for instance, default on debt altogether where a democratic leader would pay it back for the possibility of returning to the office in the future (Amador 2003). Here, we do not assume dynamic inconsistency in voters or in the incumbent politician, but it nonetheless arises as a structural attribute of the model, suggesting that aggregate political inconsistency does not stem exlusively from agents' inconsistency. This is in line with the overwhelming majority of the literature that features the inability to commit to decisions of yesterday and uses it to explain inefficient institutional persistence. In section 5, we propose a solution for this deadlock, which we dub a political equivalent of behavioural economics' commitment device (Thaler and Benartzi 2004).

The problem of the commitment of the government to its promises is central to our analysis, along with exploiting network externalities when countries are interconnected. For the mechanism designed in section 6, necessary assumptions are voters' farsightedness (Dutta et al. 2005) and spillovers to ensure that technology can spread

between countries with no or minimal barriers[1]. There is no myopic adjustment as there are no shocks to the model, which is crucial to ensure the common knowledge in all periods and, consequently, to establish the equilibrium path to efficient institutions across the network of countries.

The paper is motivated by a seeming inability of national governments to promote emissions reduction in such a scale as to limit the average temperature increase to 2°C. The novelty of the approach used in the Paris Agreement, while supported by some (Harstad 2019), was heavily criticized for the lack of adequacy and ambition (Gollier and Tirole 2015). Proponents of carbon pricing believe it is the first-best solution to the game of climate change, however, for the complexity of the task, it does not seem attainable, and if so, we need some simple mechanisms that satisfy political constraints and at the same time deliver noticeable positive results in an accelerated timeframe. How do we model policy in such a way as to avoid a runaway climate change in the near future, when both politicians and voters implicitly prefer this option, but the world is locked up in the inefficient status quo? Our paper suggests three ways that are discussed in details in Sections 4.1, 5 and 6. Section **??** describes the basic model, section 4 presents the solution. Section 7 provides some applications for the approaches developed in the paper, while Section 8 generalizes and concludes.

3 Benchmark: a simple political game of institutional adoption

3.1 Voter preferences

There are three time periods, $t = 0, 1, 2$. A mass one of voters is divided into two groups, G_1 (initial mass μ_0) and G_2 (mass $1 - \mu_0$), with $\mu_0 > \frac{1}{2}$. Voters in G_1 lead a high-carbon lifestyle, and voters in G_2 are environmentally friendly (driving electric/hybrid cars or cycling, eating a vegetarian/vegan diet, travelling with a train, recycling or using alternative materials). At every t one of the two political states is possible: status quo, a, or alternative, b. Hence, a political state at time period t is $p_t \in \{a, b\}$. Each state favours one of the lifestyles and penalises the other. Status quo, which is defined as the political state at time period $t = 0$ with the majority of voters, is a state of business-as-usual and is suited for those with little or no consideration for the environment. The alternative, in turn, penalises such people and promotes green lifestyle. This is reflected in the payoff structure in Table 1 and per-period payoffs are defined as follows.

1. If $i \in G_1$, α if a is the current political state, 0 if b is the current political state (consider 0 to be a normalisation without much interpretational load);

2. if $i \in G_2$, $\alpha - \delta$ if a is the current political state where δ is the (implicit) cost of

[1] Bayer and Urpelainen (2013) show that technology transfer is not always possible and certain conditions on the technology and on the host country must be met for the transfer to happen successfully.

	G_1	G_2
a	α	$\alpha - \delta$
b	0	β

Table 1: Per-period payoffs to the voters.

being an environmentalist under the status quo, β if b is the current political state.

We further assume that group membership is endogenous and evolves according to the following rule:

1. if $p_t = p_{t-1}$, $\mu_{t+1} = \mu_t$ (note that μ_1 is undefined under this rule, hence set $p_{-1} = a$ to emphasize that the status quo by the beginning of period 1 is a);

2. for some externally defined and fixed $\varepsilon \geq 0$, if $p_t \neq p_{t-1}$ where $p_{-1} = a$,

   (a) $\mu_{t+1} = \min\{\mu_t + \varepsilon, 1\}$ if $p_t = a$ where with probability one a voter in G_1 remains in G_1, and with probability $\max\{1 - \frac{\varepsilon}{1-\mu_t}, 0\}$ (respectively, $\min\{\frac{\varepsilon}{1-\mu_t}, 1\}$) a voter in G_2 remains in G_2 (respectively, moves to G_1), and

   (b) $\mu_{t+1} = \max\{\mu_t - \varepsilon, 0\}$ if $p_t = b$, where with probability one a voter in G_2 remains in G_2 and with probability $\max\{1 - \frac{\varepsilon}{\mu_t}, 0\}$ (respectively, $\min\{\frac{\varepsilon}{\mu_t}, 1\}$) a voter in G_1 remains in G_1 (respectively, moves to G_2). Here, ε is regarded as a structural feature of the economy which cannot be altered. It is the stiffness of the country towards change, its embeddedness in the current order. More speculation on the role of ε and the nature of the group membership is provided below.

'Moving' between the groups is a poor, but convenient term to express the idea that voters can make changes in their lifestyles and become more or less environmentally friendly. The change is, however, costly, and it costs c for a single person to move between the groups. The cost can be reflected in the price of buying an electric car and inconvenience of charging it, deteriorating social connections due to changed lifestyle and eating habits, the time spent learning the basic principles of recycling and travelling to the recycling centre every so often. In the simplest version of the model, the cost is kept constant.

At the beginning of the game, the groups are in equilibrium and no one prefers to change their allegiance provided that $\alpha - c < \alpha - \delta$, or, equivalently, $c > \delta$. Moreover, the state where everyone leads a low-carbon lifestyle is a socially superior state at $t = 0$ for any μ_0 compared to the status quo, thus we assume $(2\beta - c)\mu_0 + 3(1-\mu_0)\beta = \beta(3-\mu_0) - \mu_0 c > 3\alpha\mu_0 + 3(\alpha-\delta)(1-\mu_0) = 3(\alpha - (1-\mu_0)\delta)$. However, on a per-period basis, it is not individually optimal to change one's lifestyle when the alternative is a prevailing political state as $\beta > \alpha > \beta - c$.

As defined by the structure of payoffs, voters in G_1 (G_2) belong to the advantaged (disadvantaged) group when the society is in a and voters in G_2 (G_1) belong to the advantaged (disadvantaged) group when the society is in b. For simplicity of exposition, we do not assume that voters discount payoffs.[2]

3.2 Motivation: ε (to be rewritten)

The above process of evolution of group membership can be interpreted as a reduced form representation of a situation where there are economic frictions (exogenous to the model) in the mobility across groups. In the simple model, mobility is triggered whenever there is a shift in the political status quo but not otherwise. Once the political status quo has shifted, we assume that all voters who belong to the disadvantaged group by the shift in the political status quo simultaneously decide whether or not to queue up to move to the other group. Of all those voters who queue, a fraction ε is chosen at random to move to the other group. For simplicity of exposition, suppose there is no cost to queuing: then it is a weakly dominant action for all voters belonging to the payoff disadvantaged group to queue and each voter has a probability ε of moving across groups whenever there is a shift in the political status quo. We also assume that voters always play a weakly dominant strategy. By construction, at a subgame perfect equilibrium, in our set-up, voters are time-consistent: they cannot pretend to deviate in the future to obtain higher payoff today (as in Dewatripont and Roland (1992)).

Hence, in our model $\frac{\varepsilon}{\mu_0}$ is a measure of the underlying mobility friction so that $\varepsilon = 0$ is a situation with maximum friction and $\varepsilon = \mu_0$ is a situation with no mobility frictions.

Described here group mobility can be best understood in the context of a trade reform discussed in Fernandez and Rodrik (1991) which, if implemented, would raise basic income of workers in one sector, lower income of workers in the other sector, and draw individuals from the second sector to the first. The authors describe a small open economy with only two sectors in which workers can move between the sectors at a cost not known in advance. Trade reform, if passed, will destabilize existing equilibrium and cause some workers to relocate, but it is not possible to know beforehand who will move and enjoy higher payoff. This example, however, does not motivate the exact magnitude of the movement and its rigidity towards alteration. We can nonetheless attempt to modify it to accommodate those features. For instance, the highest possible mobility can be defined by the education system: the differences between the way two groups of workers are taught are so vast, that only ε of them are capable of changing their career. In addition, relocation between the sectors can require physical movement, which puts pressure on the facilities and infrastructure in the areas with the highest migration levels that can only accommodate certain number of people without provoking unrest among the local population. However, even with these adjustments, the features of mobitility frictions in this pa-

[2]We have solved the model with discounting and the results are quantitatively the same with slightly different boundary values for ε.

per are better thought of with regard to climate change mitigation: introduction of new technologies, and natural resources and assets that are becoming stranded.

When an innovative technology becomes available to consumers, it is likely that the supply will not be enough to satisfy all those wishing to get access to it. Not only the production capacity may not meet demand, new technologies are also oftentimes prohibitively expensive, which prevents otherwise perfectly solvent customers from contributing to the technology's expansion. This translates into congestion in the transition process when not all those who want the technology can have it instantly. Since the introduction of solar photovoltaic panels, their cost stayed a major deterrent to its expansion, and even recently its market value is still far below its cost (Borenstein 2008). At the same time, from the research of the Yale program on climate change communication, it is evident that about half of Americans would pay more for renewable energy (Leiserowitz et al. 2018). However, an average American is only willing to add \$15 per month on top of her electricity bill to have a 'clean' energy provider, which is well below what is needed to promote solar PV panels, and thus without subsidies, the desire to install solar panels will not translate into action – there is congestion on the consumers' side.

In a related vein, stranded resources are becoming an evermore prominent phenomenon with the rise of more vocal environmental concerns. As of 2011, to meet the Paris Agreement target of (less than) 2°C global warming with a 50% chance, 82% of coal has to stay underground, 33% of oil and 49% of natural gas, i.e. become stranded resources (McGlade and Ekins 2015). Apart from implications for relationships between developed and developing countries (Bos and Gupta 2019), this means that with sufficiently fast fossil fuels phase-out, the infrastructure will rapidly become obsolete, which in turn implies colossal losses for shareholders and destabilization of the world economy. This scenario is nevertheless unprobable exactly because capital cannot be thrown away. Thus, whatever changes the economy might undergo, the speed of the transition is contingent on the readiness of the economy and all its agents to declare a significant share of assets stranded and widely introduce substitutes for the lost capital. The proportion of people who are able to rip the fruits of the transition in each period is therefore fixed and determined by the rate of fossil-related assets extirpation. Therefore, our ε is a counterpart of a rate of transition in these examples and is defined by the structure of the economy and its innovative production capacity.

3.3 The Model

A policy p^t at time period t is simply a proposal to implement either of the two political states. In this manner, $p^t = p_t$. Since the voters are farshighted, the programme is limited to one time period and does not suggest anything further in time due to incredibility of any such proposal.

There is an incumbent politician at the beginning of $t = 0$. The incumbent proposes a policy p^0.

After observing the policy chosen by the incumbent, at $t = 0$, an entrant (drawn

from a set of possible candidates) can choose to challenge the incumbent by paying a cost $h > 0$. If there is more than one candidate who chooses to challenge the incumbent, the candidate who actually faces the incumbent is chosen at random with each potential challenge having equal probability of being chosen. Only the chosen challenger pays the cost.

If there is no challenger, the incumbent gets elected with probability one. If there is a challenger, she proposes an alternative policy q^0.

Voting takes place and each voter votes for their preferred policy. The politician with the majority of votes wins; if there is a tie, either candidate has a probability $\frac{1}{2}$ of being elected. The elected politician implements her policy and obtains a per-period payoff that is a fraction $e > 0$ of the social payoff at the implemented political state. We assume that $h > \hat{h} \equiv \frac{e}{2}\left((2\beta - c)\mu_0 + 3\beta(1 - \mu_0)\right)$ so that in no subgame an entrant will choose to challenge the incumbent when both make identical policy proposals. In other words, as long as the incumbent makes an optimal proposal, she will not be challenged. This grants unlimited, unconditional first-mover advantage to the incumbent thereby eliminating political competition and putting a spotlight on the question of social improvement regardless of the political turmoil.

The elected politician from the preceding period becomes the incumbent in period $t = 1$, at which point the political process described above is iterated.

The timeline within a period can be described as follows and is sketched in Figure 1.

1. The share of voters in each group, μ_t, is defined.
2. An incumbent proposes a programme p^t.
3. A challenger (if any) makes an alternative policy proposal q^t and pays the cost c (if several, then, first, they propose their policies, then, one of them is randomly picked, finally, she pays the cost and gets to compete the incumbent).
4. Voters choose a new incumbent by the simple majority rule.
5. The winning politician implements the proposed policy, voters receive payoffs in accordance with the (new) state, politician receives her share of the social payoff.
6. (if there is a change in the status quo) Voters from the disadvantaged group queue to move to the advantaged group, ε of them gets picked. This completes the process and defines μ_{t+1} for the next period, at which point the game repeats from the top.

We solve for the subgame perfect equilibria of the three-period sequential political game of institutional adoption.

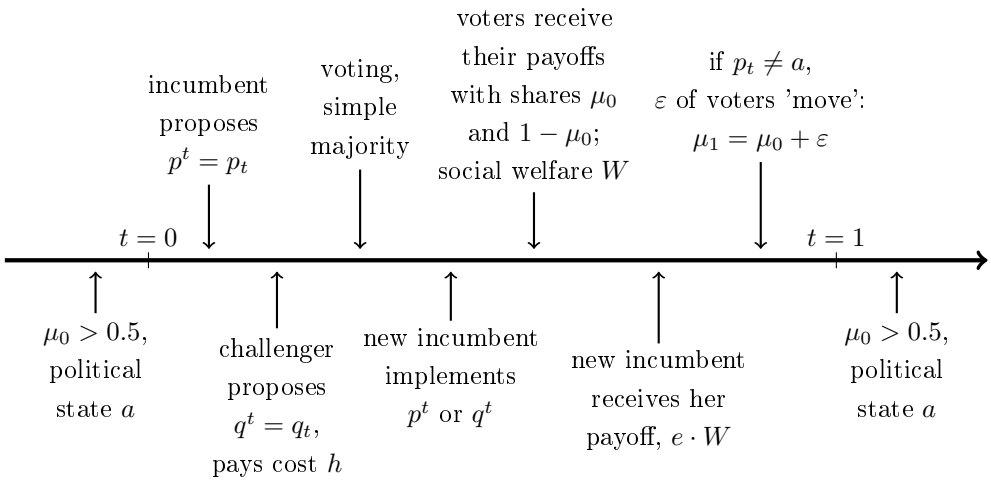

Figure 1: Timeline.

4 The Benchmark Solution

In this case, $\mu_t \leq \mu_0$ for each $t = 1, 2$ and group membership evolves over time. We solve the model backwards.

At $t = 2$, for any $p_1 \in \{a, b\}$,

1. if $\mu_2 > \frac{1}{2}$, the incumbent politician will make a policy proposal $p^2 = a$, no challenger will arise and the incumbent will be re-elected with probability one and implement a;

2. if $\mu_2 < \frac{1}{2}$, the incumbent politician will make a policy proposal $p^2 = b$; no challenger will arise and the incumbent will be re-elected with probability one and implement b.

At $t = 1$, for any $p_0 \in \{a, b\}$, anticipating the outcome of the subgame at $t = 2$, there are several possible cases to look at.

1. If $\mu_1 > \frac{1}{2}$ and $\mu_1 - \varepsilon > \frac{1}{2}$, then the incumbent politician will propose $p^1 = a$; no challenger will arise and the incumbent will be re-elected with probability one.

2. If $\mu_1 > \frac{1}{2}$ and $\mu_1 - \varepsilon < \frac{1}{2}$, then voters in G_1 anticipate that voting for a proposal $p^1 = b$ will result in the expected payoff

$$0 + \frac{\varepsilon}{\mu_1}(\beta - c) + \left(1 - \frac{\varepsilon}{\mu_1}\right) \cdot 0 = \frac{\varepsilon}{\mu_1}(\beta - c),$$

while voting for proposal $p^1 = a$ will result in the expected payoff of 2α. Notice that the politician should only be concerned with the payoff to the winning group, because she cannot maximize social payoff directly by imposing a Pareto

superior alternative on the society. Hence, the expected payoff is calculated only with respect to the majority group of voters (G_1 in this case, since $\mu_1 > \frac{1}{2}$). Therefore, the incumbent will make the policy proposal $p^1 = b$ if and only if

$$\frac{\varepsilon}{\mu_1}(\beta - c) \geq 2\alpha \iff \varepsilon \geq 2\mu_1 \frac{\alpha}{\beta - c};$$

otherwise, the incumbent will make the policy proposal $p^1 = a$; no challenger will arise and the incumbent will be re-elected with probability one.
Notice that by assumption, $\alpha > \beta - c$, and thus $\frac{\alpha}{\beta - c} > 1$. As long as $\mu_1 > \frac{1}{2}$, ε needed to vote for b at $t = 1$ is greater than 1. This condition is never satisfied, hence under no circumstances b is adopted at $t = 1$.

3. If $\mu_1 < \frac{1}{2}$ and $\mu_1 + \varepsilon < \frac{1}{2}$, then the incumbent politician will be indifferent between making the policy proposal $p^1 = b$; no challenger will arise and the incumbent will be re-elected with probability one.

4. If $\mu_1 < \frac{1}{2}$ and $\mu_1 + \varepsilon > \frac{1}{2}$, then voters in G_2 anticipate that voting for a proposal $p^1 = a$ will result in the expected payoff

$$(\alpha - \delta) + \frac{\varepsilon}{1 - \mu_1}(\alpha - c) + \left(1 - \frac{\varepsilon}{1 - \mu_1}\right)(\alpha - \delta) = 2(\alpha - \delta) + \frac{\varepsilon}{1 - \mu_1}(\delta - c),$$

while voting for proposal where $p^1 = b$ will result in the expected payoff of 2β. Therefore, the incumbent will make the policy proposal $p^1 = a$ if and only if

$$2(\alpha - \delta) + \frac{\varepsilon}{1 - \mu_1}(\delta - c) \geq 2\beta;$$

otherwise, she will make the policy proposal $p^1 = b$, no challenger will arise and the incumbent will be re-elected with probability one. However, note that $\delta < c$, hence $\delta - c < 0$, so the coefficient on ε is negative, and therefore the expression takes the form

$$\varepsilon \leq 2(1 - \mu_1)\frac{\beta - \alpha + \delta}{\delta - c},$$

which is in itself a negative number, and since $\varepsilon \in (0, 1)$, this case cannot arise in equilibrium.

At $t = 0$, given that the status quo is a and $\mu_0 > \frac{1}{2}$, anticipating the outcome of the subgame at $t = 1$, there are several possible scenarios:

1. If $\mu_0 > \frac{1}{2}$ and $\mu_0 - \varepsilon > \frac{1}{2}$, then the incumbent politician will propose $p^0 = a$, no challenger will arise and the incumbent will be re-elected with probability one. Note that we obtain this conclusion even though each incumbent politician, conditional on remaining in power, prefers the alternative b to a, but any attempt to alter the status quo implies that the incumbent politician loses power with probability one.

2. If $\mu_0 > \frac{1}{2}$ and $\mu_0 - \varepsilon < \frac{1}{2}$, then voters in G_1 anticipate that expected payoff from voting for a proposal $p^0 = a$ will be 3α, while voting for proposal $p^0 = b$ will be $0 + \frac{\varepsilon}{\mu_0}(\beta - c) + (1 - \frac{\varepsilon}{\mu_0}) \cdot 0 + \frac{\varepsilon}{\mu_0}\beta + (1 - \frac{\varepsilon}{\mu_0}) \cdot 0 = \frac{\varepsilon}{\mu_0}(2\beta - c)$. Therefore, the incumbent will propose b if and only if

$$\frac{\varepsilon}{\mu_0}(2\beta - c) \geq 3\alpha \iff \varepsilon \geq 3\mu_0 \frac{\alpha}{2\beta - c},$$

otherwise she will make the policy proposal $p^0 = a$, no challenger will arise and the incumbent will be re-elected with probability one.

It is of some interest to note that the switch in the political state from a to b at $t = 0$ is monotonic in ε, and there is a tipping point for the value of ε, which marks early adoption of efficient institutions on one side and inefficient persistence on the other side.

We summarize the above discussion as the following result:

Proposition 1:

(a) (Inefficient Institutional Persistence) Along the equilibrium path of play, if $\varepsilon < \max\left\{\mu_0 - \frac{1}{2}, 3\mu_0 \frac{\alpha}{2\beta - c}\right\}$, $\{p_t = a : t = 0, 1, 2\}$; *at each* t, *the incumbent proposes* $p^t = a, \forall t$, *and gets re-elected. There is inefficient institutional persistence.*

(b) (Early Adoption of Efficient Institution) Along the equilibrium path of play, if $\varepsilon \geq \max\left\{\mu_0 - \frac{1}{2}, 3\mu_0 \frac{\alpha}{2\beta - c}\right\}$, $\{p_t = b : t = 0, 1, 2\}$; *at each* t, *the incumbent proposes* $p^t = b, \forall t$, *and gets re-elected. Socially efficient institution is adopted.*

Proposition 1 calculates the bound on the mobility friction, above which there is inefficient institutional persistence and below which there is early adoption of the efficient institution. The calculated bound is intuitive: the closer μ_0 is to $\frac{1}{2}$, the lower is the bound; the lower is α (the instantaneous loss in payoff to a voter in G_1 when the political state shifts from a to b), the lower is the bound; the higher is $2\beta - c$ (the future gain in payoff to a member of G_1 who successfully becomes a member of G_2 following a shift in the political status quo from a to b), the lower is the bound.

4.1 Engineering a Majority in Favour of b

For this and subsequent sections, we will assume that there is an outside agent (e.g. an international agency/policymaker such as the UN which internalizes all global externalities) who is able to change the set of actions available to the incumbent while taking domestic mobility frictions as given, meaning the outside agent cannot intervene by increasing the value of ε. In addition, we will also assume that the outside agent cannot directly manipulate the values of the underlying payoff parameters (i.e. $\alpha's$ and $\beta's$).

In what follows, for simplicity, we will also assume that we are in the setting where model parameters are characterized as in Proposition 2(a) i.e. $\varepsilon <$

$\max\left\{\mu_0 - \frac{1}{2}, 3\mu_0\frac{\alpha}{2\beta - c}\right\}$. Consider the problem of engineering institutional change in such a setting.

With these two constraints in mind, note that it follows that there is no way in which such an outside agent can intervene to engineer a majority in favour of b when $\varepsilon < 3\mu_0\frac{\alpha}{2\beta - c} \leq \mu_0 - \frac{1}{2}$. We will turn to this case in section 6. For the remainder of this section, we focus on the case $\mu_0 - \frac{1}{2} < \varepsilon < 3\mu_0\frac{\alpha}{2\beta - c}$.

We know from the computations characterizing the subgame perfect equilibrium strategy profile at $t = 1$, that if there is a way to engineer a switch to b at $t = 0$, the political equilibrium will ensure that at $t = 1$, b continues to persist. The reason for inefficient institutional persistence in the above scenario is that although voters in G_1 realise that the switch to b at $t = 0$ will ensure that b prevails thereafter, the probability of any individual voter in G_1 of being able to actually move to G_2 is sufficiently small to ensure that voting for a payoff-dominates switching to b. Therefore, the incumbent prefers to adhere to the status quo a, thus ensuring that a prevails in all subsequent periods as well, resulting in institutional persistence.

So how can an outside agent engineer a majority in favour of b?

The intervention that might work in this setting is to ensure that all voters belonging to a specific subset of G_1, say G_1' of mass ε, are guaranteed a switch to G_2 at $t = 1$ if they vote for b at $t = 0$. Straightforwardly, once such a subset is identified, computations symmetric to those underlying Proposition 1 will ensure that if offered a choice between a and b, all voters in $G_1' \cup G_2$ will vote for b at $t = 0$: as the incumbent's payoff from b is higher, at each t, the incumbent proposes $p^t = b$, no challenger emerges, the incumbent always gets re-elected and socially efficient institution is adopted.

In effect, by discriminating between agents in G_1 and favouring voters belonging to a carefully chosen subset G_1' over other agents in G_1, a majority in favour of b is constructed. Possible ways in which G_1' can be formulated include:

1. The marker identifying the subset G_1' could be location, ethnicity or some other marker of identity. In the context of climate change, it could be identifying communities and neighbourhoods that are more predisposed to low-carbon lifestyle and are characterized by some features that could facilitate the transition, and promoting green solutions in such communities.

2. Alternatively, if agents have positive but small heterogeneous queuing costs (for consistency with our existing computations, we will need to assume that the queuing costs are small enough so that it is risk-dominant for all voters in the disadvantaged group to queue up whenever there is a shift in the political status quo), G_1' could be constructed on the grounds of economic efficiency, i.e. the subset of agents in G_1 whose moving costs across the two groups is lowest. This way, marginal voters can be targeted and, rather than subsidizing the whole electric vehicle market, those who are not willing but capable of purchasing an electric car should be presented with an opportunity to rethink their choice.

3. Another possibility is to create a first mover advantage in agreeing to move out of G_1 by creating a fund that compensates or at least subsidizes the first ε voters in the queue. For instance, facilitation of the transition to greener lifestyle via subsidies or other means. Electric vehicles can be promoted using tax rebates, access to high occupancy lanes, special electricity tariffs and free charging points. Decrease in the meat and dairy consumption can be encouraged by providing and possibly subsidizing plant-based alternatives and disseminating the knowledge about the benefits of a plant-based diet. Norway's comprehensive incentives scheme for electric vehicles can serve as a good example. Numerous incentives, including exemption from non-recurring vehicle fees and purchase taxes, exemption from the annual road tax, all public parking fees, and toll payments, as well as permission to use bus lanes, are in force until the Parliament's goal to reach 50,000 electric vehicles is achieved. This ensures the first-mover advantage for 50,000 first buyers, thus making it attractive for them to switch from petrol to zero emission cars.

5 Delayed Adoption as a Commitment Device

For many years one of the predominant matters that occupied the minds of political scientists and economists was, and still is, the problem of commitment of the government to its promises. As Daron Acemoglu and his co-authors write, one of potential reasons that lead to the rise of inefficient institutions is commitment issues (Acemoglu et al. 2005). For instance, transition from dictatorship to democracy requires commitment on the citizens behalf to keep current dictator in power, which they cannot promise, and transition from democracy to dictatorship requires commitment of the government, which it cannot promise either. Since power is a contract between the government and the voters, while there is no third power to enforce this contract, any promises that are not self-enforcing or do not have any other means of being enforced, are not credible.

Similarly, central issue of our analysis here is the problem of commitment of the government to its promises. As in the bulk of the political economy literature, the incumbent in our research cannot promise not to misuse the power in the future, when the society supports new institutions. The majority of voters do not benefit from the transition unless there are wealth transfers between the groups, and all of the promises on behalf of the incumbent about any transfers taking place in the future cannot be credible. Therefore, when current institutions cannot turn to more efficient institutions in a matter of one period, we need a commitment device. Every institution creates a group of people who have a vested stake in perpetuating it, and precisely this leads to a political conundrum: these people will not benefit from a political change and thus do not support the shift and do not buy into promises about any future transfers. In this section we resolve this impasse.

We continue to assume that the outside agent cannot directly manipulate the values of the underlying payoff parameters (i.e. $\alpha's$ and $\beta's$). However, now we will relax the assumption that at the initial time period $t = 0$, the outside agent can

propose that an ε fraction of agents be given an option to move between the groups even without a change in the political state in the same time period. In other words, a carefully chosen fraction ε of the voters from G_1 are offered to move at $t = 0$ preceding the change in the political state at a later time period. This would imply that unless there is a guaranteed shift in the political state in the future, there is no apparent gain for them to do so. On the other hand, the shift cannot be guaranteed since there is no reason for them to trust that an incumbent will fulfil the promise given by the outside agent when period $t = 1$ comes. However, them moving from G_1 to G_2 represents exactly this guarantee, because either the incumbent will propose an alternative policy and win the next election, or she will abandon the promise and propose the status quo, consequently be challenged and lose the office. Therefore, the group of early adopters who forfeit some of their benefits in the present in hope of recovering them in the future when the status quo changes, act as a commitment device to ensure that, come tomorrow, an incumbent has no other option but to propose the alternative as a policy, or to adhere to an inefficient status quo and lose the power.

The fundamental point here is that the shift from the status quo to the alternative *cannot and does not* take place in the same period, thus some early adopters are needed to hold the politician responsible should she fail to perform the promised shift in the political state. In this case, the commitment device is the subset of electorate that is guaranteed some benefits when the status quo is abandoned in favour of the alternative, but in the current period there is immediate cost for them of doing so, and the punishment for the politician who fails to deliver the promise is the loss of the office. This strongly resonates with the problem of dynamic inconsistency, only that in our case all agents are dynamically consistent and the problem of commitment to future hard choices arises from internal structure of the model. The solution we use in this section brings us back to an acclaimed empirical study by Thaler and Benartzi (2004) and has a form of committing in advance to a shift from a socially-inferior status quo by performing an action in the current period, such as offering ε share of people to move between the groups at a cost with no apparent gain in order to create an incentive for the incumbent to ensure institutional change in the next period. Hence, delayed adoption as described in the current section is a political equivalent to the commitment device in behavioural economics.

In contrast with the case studied in the previous section, now a subset of agents in G_1 are offered a choice to be early adopters to act as a commitment device to engineer, in the future, a majority of voters in favour of adoption of efficient institution and thus ensure a political state switch.

We know from the computations characterizing the subgame perfect equilibrium strategy profile at $t = 1$, that if a fraction ε of voters from G_1 adopt b at $t = 0$, the status quo a continues to prevail in that period; however, if $\mu_0 - \varepsilon > \frac{1}{2} > \mu_0 - 2\varepsilon$, so long as ε is sufficiently large, political equilibrium will ensure that at $t = 1$, there will be a switch to b. In this way, the outside agent can engineer a majority in favour

of b via strategy that involves delayed adoption.

In effect, by discriminating between agents in G_1 and affirming voters belonging to subset G_1'' of measure ε that if they move at $t = 0$ before other agents in G_1, the majority in favour of b at $t = 1$ can be constructed and their originally incurred costs will be recovered with the shift in the political state. Possible ways in which G_1'' can be constructed include are the same as in the previous section.

Formally, the game is solved as follows.

At $t = 2$, the solution is the same as in Section 4.

At $t = 1$, the solution is the same as in Section 4.

At $t = 0$, given that the status quo is a and $\mu_0 > \frac{1}{2}$, anticipating the outcome of the subgame at $t = 1$, there are several possible cases to look at:

1. If $\mu_0 > \frac{1}{2}$ and $\mu_0 - \varepsilon > \mu_0 - 2\varepsilon > \frac{1}{2}$, then the incumbent will be indifferent between making the policy proposal $p^0 = a$, no challenger will arise and she will be re-elected with probability one.

2. If $\mu_0 > \frac{1}{2}$ and $\mu_0 - \varepsilon > \frac{1}{2} > \mu_0 - 2\varepsilon$, if an ε fraction of voters in G_1 move in period $t = 0$, each voter who moves anticipates that following her switch from a to b at $t = 0$, given that $\mu_0 - 2\varepsilon > \frac{1}{2}$, the switch a to b will be made with probability one at $t = 1$; although each such voter incurs an instantaneous payoff loss, because the political state does not change at $t = 0$, as the political state shifts at $t = 1$, they gain over time (in terms of intertemporal payoffs) as long as $\frac{\mu_0}{2} - \frac{1}{4} < \varepsilon < \mu_0 - \frac{1}{2}$.

We summarize the above discussion as the following result:

Proposition 2 (Delayed Adoption of Efficient Institution):

Along the equilibrium path of play, if $\mu_0 > \frac{1}{2}$ and $\mu_0 - \varepsilon > \frac{1}{2} > \mu_0 - 2\varepsilon$, then $\{p_0 = a, p_1 = p_2 = b : t = 0, 1, 2\}$; *at* $t = 0$ *, the incumbent proposes* $p^0 = a$ *and an* ε fraction *of voters in* G_1 *are offered to move at* $t = 0$*, while at* $t \geq 1$*, the incumbent proposes* $p^t = b$ *and gets re-elected. Socially efficient institution is adopted with a delay of one period.*

6 Leveraging in a network. The section contains old notations.

Consider what happens when $\varepsilon \leq 2\mu_0 \frac{\alpha}{2\beta - c} \leq \mu_0 - \frac{1}{2}$. Suppose, as before, the outside agent cannot directly manipulate ε. So how can an institutional change be achieved in this case?

So far we have looked at the problem of institutional change form the perspective of an individual country. Now, we extend the formal model to a setting where there are $k = 1, ..., n$ countries and voters within each of these countries have to make a choice between the status quo a and the alternative b. The key modelling change is

that the payoffs within each country at any of the two political states will depend on the network structure linking these countries.

First we need to introduce some new notation. As before, we will assume that there is a mass one of voters in each country i, divided into two groups, G_1^k (initial mass μ_0^k) and G_2^k (mass $1-\mu_0^k$), $\mu_0^k > \frac{1}{2}$. The political state at time period t within country k is $p_t^k \in \{a, b\}$. The political state at t is $\mathbf{p}_t = \left(p_t^k : k = 1, ..., n\right)$. The per-period payoffs to voters in country k is:

1. if $i \in G_1^k$, $\alpha^k(\mathbf{p}_t)$ if a is the current political state, 0^k if b is the current political state;

2. if $i \in G_2^k$, $(\alpha - \delta)^k$ if a is the current political state, β^k if b is the current political state.

Notice that in specifying the per-period payoffs, only the payoff to voters in group G_1^k with the status quo as the political state is assumed to depend on the political state of other countries in the network. This specification has been adopted for simplicity of exposition: nothing essential depends on it.

Let $\mathbf{p}_t^a = \left(p_t^k = a : k = 1, ..., n\right)$.

We assume that for each country $k = 1, 2, ..., n$, the following inequalities hold: $\beta^k > \alpha^k(\mathbf{p}_t^a) > \beta^k - c$, and

$$\mu_0^k \alpha_1^k(\mathbf{p}_t^a) + \left(1 - \mu_0^k\right)\beta_1^k < \mu_0^k \alpha_2^k + \left(1 - \mu_0^k\right)\beta_2^k$$

so that the alternative is the socially efficient political state for the entire network of countries at $t = 0$, although within each country any member of the advantaged group is strictly worse off under the status quo if the network as a whole shifts to the alternative.

As before, we do not assume that voters discount payoffs.

Other than an obvious change in notation, there is no change in the way we specify how group membership evolves, i.e. we assume that the mobility across groups within a country is solely a function of the prevailing political state within that country.

Throughout this subsection, to keep matters simple, we will assume that

$$0 < \varepsilon^k < \frac{3}{2}\mu_0^k \frac{\left(\alpha_1^k(\mathbf{p}_t^a) - \beta_1\right)}{(\beta_2 - \beta_1)} \leq \mu_0^k - \frac{1}{2} \text{ for each } k = 1, ..., n,$$

so that a straightforward extension of Proposition 2 ensures that as long as each country within the network takes as given political states of all other countries, it is a political equilibrium for each country in the network to stick with the status quo in all periods $t = 0, 1, 2$. Note that we obtain this conclusion even though each incumbent politician within each country in the network, conditional on remaining in power, prefers the alternative b to a: the problem is that any attempt to alter the status quo implies that the incumbent politician looses power with certainty.

Suppose we limit the ability of the outside agent to manipulate payoffs from the status quo directly by assuming that the outside agent cannot affect payoffs for all countries except at most one, i.e. the outside agent has a limited quantity of resources to play with. How can the outside agent engineer institutional change by potentially exploiting the cross-country network structure?

To illustrate how such a mechanism might work, we need some notation and a simplifying assumptions. For any non-empty subset of countries C, let

$$\mathbf{p}_t^C = \begin{cases} p_t^k = a \text{ if } k \in C \\ p_t^k = b \text{ if } k \notin C. \end{cases}$$

For each country k, assume that there exists a non-empty subset of countries $C_k = \{k^{'} \geq k\}$ such that $\alpha_1^k \left(\mathbf{p}_t^{C_k}\right) < \beta_1^k$. Note that under this assumption, country 1 is a pivotal member of the network and political shift in it triggers a shift in all other countries, one by one. The outside agent then has to subsidize voters in G_1^1 so that

1. either $\varepsilon^1 > \mu_0^1 - \frac{1}{2}$, or
2. $\alpha_1^1 \left(\mathbf{p}_t^a\right) < s + \beta_1^1$ (whichever is cheaper), ensuring that, assuming all other countries vote for a, it is a dominant strategy for voters in country 1 to vote for b at $t = 0$.

But then it follows that it is a dominant strategy for voter's in country 2 to vote for b as well, and so on for each country in the network, till all countries in the network vote for b thus ensuring a switch from the status quo a, socially inefficient political state, to the socially efficient political state b globally.

What kind of institutions can be transformed? What are characteristics of institutions (and countries) that can change on their own accord and what countries need external intervention? These questions represent a topic for the whole other paper. Here, we will only give some examples of the applications of this mechanism.

In a network of countries using oil as a primary source of energy for cars, all benefit from petrol cars as long as others do the same. Assume that electric cars benefit society much more than petrol cars because they use renewable sources of energy and do not pollute the air with emissions, but the payoff from using petrol cars is higher because of the network externalities. Now, if we were to target one of the pivotal countries in such a network with a subsidy on electric vehicles, then the benefits of other agents in the network would fall. For instance, if a network consists of countries from the EU and the UK, subsidising German car manufacturers would lead to decline in the supply of German petrol cars, which would cause the benefits of the rest of the countries to drop and lead to them slowly switching to electric cars as well.

7 Applications

In this section we discuss some examples that we feel are a good illustration for the points we make in this paper. For convenience, it is separated into three subsections

with each corresponding to one mechanism described above.

7.1 Early Adoption

The mechanism we are examinating here is nothing more than a political economy equivalent to the first-mover advantage[3]. As per the model setup, when the society is very close to the transition point but cannot pass it due to individual-specific uncertainty and political constraints, what can be done is a group of people can become early adopters in the following sense. This group (of the size of ε) will be the ones who move after the transition has taken place and consequently the ones to be guaranteed to enjoy a new technology and receive higher utility compared to those who have not adopted it (the 'old' majority), $\beta_2 > \beta_1$, and themselves had the transition not happened, $\beta_2 > \alpha_1$. In other words, all that is left for a politician to do is to *create* the first-mover advantage to engineer the change.

Some examples of early adoption of more efficient institutions have already been mentioned in the paper and include the UK government plug-in grant and the Norway's incentives for zero emission cars scheme. The former is a government's discount for those who wish to buy a new low-emission vehicle, where the buyer does not have to apply for the grant (up to £3,500) because it is already included in the vehicle price. The latter is more of a first-mover advantage because it is only in effect until the country hits a target of 50,000 zero-emission cars. Those who decide to buy one will enjoy exemption from import taxes and 25% VAT, avoid road tolls and tax, pay half price on ferries, get free municipal parking in cities and can usually use bus lanes. This, according to our model, represents a cituation of induced shift from petrol and diesel cars to zero-emission cars, where the government acts as an outside agent and provides a certain group of citizens with incentives to prefer a more environmentally-friendly option. There are some costs of moving between the groups, as opposed to the assumption of zero moving cost in the model, since individuals would need to buy a new car, so to adjust this case to our model, we can think of the subset of people who were already thinking about buying a new car, and they just have to choose now whether it would be a petrol or an electric vehicle.

In fact, in Norway the incentives worked so well that now there is a 12-months waiting list for the country's most popular electric car — Volkswagen e-Golf. German manufacturer's production capacity defines ε, as demand at this stage exceeds supply. Thus, we have a situation with presumably zero queuing costs (as it does not cost anything to be in the waiting list), and it is still risk-dominant for potential buyers to queue up to move to the group with numerous perks on the road.

Very generally speaking (maybe even too generally), any government's grant or subsidy which is aimed at shifting the status quo by creating the first-mover advantage can be potentially regarded as an attempt to engineer majority in favour of a desired alternative. Instances may include government subsidies in the US for

[3]Usually, first-mover advantage is discussed with respect to the production side of the economy (Gal-Or 1985, Lieberman and Montgomery 1988, Kerin et al. 1992). Another common context is the first-mover advantage in sequential public goods games (Varian 1994).

purchases of Tesla's cars or the California Solar Initiative to promote solar panels installations (Hughes and Podolefsky 2015).

7.2 Delayed Adoption

We find the mechanism illustrated in this subsection the most interesting yet based on a simple idea of the first-mover *disadvantage* of consuming technologies that have not yet been widely adopted.[4] Such a situation is most relevant for global warming mitigation: environmentally-superior behaviour is harder to follow, hence not many people do, and the first among those who do are at a disadvantage due to lack of sustainable consumer products alternatives. As such, they undergo a period of inconvenience (for their own moral reasons) because they correctly expect that in the future, more people will follow their example and the shift in the public conscience will occur. These people are the ones to hold politicians accountable for their lack of actions through protests, strikes and voting.

We construct our supporting example on the case of electric cars. Let us consider the market for electric cars in the UK. The UK government has pledged under the Kyoto protocol to fight climate change by moving the country towards cleaner energy and less reliance on the fossil fuels. As a part of this pledge, the government is subsidising the purchases of certain low-emission vehicles, which includes zero-emission cars running exclusively on electricity. This means that the government has to take care of the charging points since such vehicles cannot be 'refueled' at the traditional petrol stations. Installation of charging points is a costly activity, and while the amount of electric cars in the country is relatively low, the stations would not be ubiquitous. This means that the first owners of such cars buy these vehicles at the expense of their own comfort and convenience, because charging them, once out of battery, will require searching for the nearest charging point which may be at a considerable distance to a person's home, and waiting for the car to fully charge, which varies from 30 minutes to more than 10 hours. The increase in the number of cars will ensure the increase in the number of charging points, but people who were the first to swap their petrol vehicles for the zero-emission ones would have to wait to reap the full benefits of this environmentally-friendly alternative. Naturally, the government would not have to specifically target a certain group of car users; the shift will be secured by the self-selection process of those who care enough about the environment and can afford buying an electric vehicle (and not settling for a bike instead).

7.3 Network Externalities

Before the Montreal protocol was signed, the US unilaterally banned the use of chlorofluorocarbons (CFCs) in aerosols in the 1970s. By then, science was not solid on the link between CFCs and ozone layer depletion. Some other countries supported the US and introduced restrictive measures in attempts to curb CFC emissions. The

[4] We are not the ones to coin the term, but we are the first to place it in the context of climate change innovation. It is traditionally employed in the context of firms competition (Gal-Or 1987).

measures, as noted by Scott Barrett, were costly for them to implement (2003, p.223), not the least because they were the first to act and did so unilaterally. In 1987, however, 23 countries signed the protocol which would prove to be almost unique in its preformance and outcome. As of today, it is universally ratified and still successfully in effect with some incidental regional non-compliance. The US in this example served as a starting point for the shift in the rest of the countries. It went through a period of unilateral ban which inadvertently damaged its economy, but because of the size of it, it also eventually resulted in other countries following in the footsteps and lifting the ban or restrictions as well. This event perfectly illustrates how network externalities can be exploited and a positive change can spread to all nations starting with a pivotal among them.

In the same manner, smallpox eradication started with a single country proposal – USSR – to the World Health Assembly to merge global efforts. Eradication dividend from such efforts takes time to mature and implies increased cost in the short and medium term due to heightened vaccination demand and the need for surveillance. According to already mentioned Scott Barrett, India's eradication dividend was much higher than that of the US or even to all industrialised countries (2007, p.51). Yet, India would not become an initiator of such a global programme, but it would agree to be the follower. Hence, whenever there is an obvious benefit to the countries that cannot afford to be the first movers, and there is a (smoewhat smaller) benefit to the more developed world, the latter would inevitably have to contribute more and become initiators of the change. Even if the programme is formally supported by the UN or WHO (World Health Organisation), the funds for these organisations are primarily formed out of contributions of richer nations. A cul de sac is expected if the former cannot, and the latter will not. This is where an interference is needed, and the country who's benefit to cost ratio is the largest should be nudged to create an initial push.

Finally, let us come back to the main application of this paper – climate change. Kyoto Protocol, which was meant to be an equiavalent of the Monteral Protocol, was not an equivalent success story (Barrett 2003). It is suggested that the fundamental reason for this is free riding, and since climate change is a global public good that depends on aggregate efforts of all nations, it is not impervious to this issue (Barrett 2007). We propose that free riding in this case is also intertemporal, where no one wants to meet the prospect of being the first to tackle climate change unilaterally, but almost everyone would agree to be the second. Such multidimensional public good represents the third part of our model and would have to be provided strarting from more developed nations who would bear the most cost. Fortunately, there are externalities in this process, ranging from R&D to the market for end-use products, and these externalities can be leveraged to engineer a global transition to green economy once the pivotal countries can be identified and nudged in the right direction. *Unfortunately,* the pivotal countries do not seem interested in leading the change[5],

[5]We mean th US withdrawal from the Paris Agreement.

and while the transition is still theoretically possible, the prospect of it coming from the most powerful nations looks bleak.

8 Concluding Remarks

The paper is motivated by a seeming inability of national governments to promote emissions reduction in such a scale as to limit the average temperature increase to 2°C. The novelty of the approach used in the Paris Agreement, while supported by some (Harstad 2019), was heavily criticized for the lack of adequacy and ambition (Gollier and Tirole 2015). Proponents of carbon pricing believe it is the first-best solution to the game of climate change, however, for the complexity of the task, it does not seem attainable, and if so, we need some simple mechanisms that satisfy political constraints and at the same time deliver noticeable positive results in an accelerated timeframe. How can change be engineered as to avoid a runaway climate change?

In this paper, in the presence of limited transferability of payoffs and mobility frictions, we explore three mechanisms to engineer institutional change. First, when voters are split amost equally between the alternatives, eliminate uncertainty for individual voters, by identifying a subgroup who will benefit from the reform with certainty. Second, for the countries where the majority in favour of the status quo is higher, construct a "coalition" of voters who will benefit from the change in the future and are willing to sacrifice current payoffs by moving early and in doing so, change the facts on the ground so that delayed adoption of the efficient institution occurs. Third, leverage network externalities by targeting 'pivotal' set of countries to ensure institutional change in all members of this network.

Several simplifying assumptions were made in the model. To name some, we have assumed that there are only three time periods, while infinite horizon could have made more sense. We also presumed that the nature of social welfare gains from institutional change is common knowledge as is the nature of the underlying network externalities across countries. Preferences of the median voter are dynamically consistent, there is no discouting and no inherent structural defect, yet the model is still characterized by aggregate time inconsistency, which leads us to believe that observed political inconsistency in real life is not necessarily a result of individual inconsistency. We contend that this is an interesting insight which demands further elaboration. In the future, we plan to extend the model to allow for above-mentioned extensions as well as to explore the causes of political time inconsistency.

References

Acemoglu, D., F. A. Gallego, and J. A. Robinson (2014). Institutions, Human Capital, and Development. *Annual Review of Economics 6*, 875–912.

Acemoglu, D., S. Johnson, and J. A. Robinson (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review 91*(5), 1369–1401.

Acemoglu, D., S. Johnson, and J. A. Robinson (2002). Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution. *The Quarterly Journal of Economics 117*(4), 1231–1294.

Acemoglu, D., S. Johnson, and J. A. Robinson (2005). Institutions as a Fundamental Cause of Long-run Growth. *Handbook of Economic Growth*, 385–472.

Acemoglu, D. and J. A. Robinson (2013). *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. Broadway Business.

Alesina, A. and G. Tabellini (1990). A Positive Theory of Fiscal Deficits and Government Debt. *The Review of Economic Studies 57*(3), 403–414.

Amador, M. (2003). A Political Economy Model of Sovereign Debt Repayment. *Manuscript, Stanford University*.

Barrett, S. (2003). *Environment and Statecraft: The Strategy of Environmental Treaty-Making*. Oxford University Press.

Barrett, S. (2007). *Why Cooperate? The Incentive to Supply Global Public Goods*. Oxford University Press. Oxford Scholarship Online, 2011. doi: 10.1093/acprof:oso/9780199211890.001.0001.

Battaglini, M. and B. Harstad (2016). The Political Economy of Weak Treaties. *Journal of Political Economy*.

Bayer, P. and J. Urpelainen (2013). External sources of clean technology: Evidence from the clean development mechanism. *The Review of International Organizations 8*(1), 81–109.

Binger, B. R. and E. Hoffman (1989). Institutional Persistence and Change: The Question of Efficiency. *Journal of Institutional and Theoretical Economics (JITE)/Zeitschrift für die gesamte Staatswissenschaft*, 67–84.

Borenstein, S. (2008). The Market Value and Cost of Solar Photovoltaic Electricity Production. Center for the Study of Energy Markets.

Bos, K. and J. Gupta (2019). Stranded assets and stranded resources: Implications for climate change mitigation and global sustainable development. *Energy Research & Social Science 56*, 101215.

Coate, S. and S. Morris (1999). Policy Persistence. *American Economic Review 89*(5), 1327–1336.

Dewatripont, M. and G. Roland (1992). Economic Reform and Dynamic Political Constraints. *The Review of Economic Studies 59*(4), 703–730.

Diamond, J. M. (1997). *Guns, Germs and Steel: a Short History of Everybody for the Last 13,000 Years*. London: Jonathan Cape.

Dollar, D. and A. Kraay (2003). Institutions, Trade, and Growth. *Journal of Monetary Economics 50*(1), 133–162.

Dutta, B., S. Ghosal, and D. Ray (2005). Farsighted Network Formation. *Journal of Economic Theory 122*(2), 143–164.

Espín-Sánchez, J.-A. (2017). Institutional Inertia: Persistent Inefficient Institutions in Spain. *The Journal of Economic History 77*(3), 692–723.

Fernandez, R. and D. Rodrik (1991). Resistance to Reform: Status Quo Bias in the Presence of Individual-Specific Uncertainty. *The American Economic Review*, 1146–1155.

Frankel, J. A. and D. H. Romer (1999). Does Trade Cause Growth? *American Economic Review 89*(3), 379–399.

Gal-Or, E. (1985). First mover and second mover advantages. *International Economic Review 26*(3), 649–653.

Gal-Or, E. (1987). First Mover Disadvantages with Private Information. *The Review of Economic Studies 54*(2), 279–292.

Galiani, S., I. Torre, and G. Torrens (2019). International organizations and the political economy of reforms. *Journal of International Economics 121*.

Glaeser, E. L., R. La Porta, F. Lopez-de Silanes, and A. Shleifer (2004). Do Institutions Cause Growth? *Journal of Economic Growth 9*(3), 271–303.

Góes, C. (2016). Institutions and growth: A gmm/iv panel var approach. *Economics Letters 138*, 85–91.

Gollier, C. and J. Tirole (2015). Negotiating effective institutions against climate change. *Economics of Energy & Environmental Policy 4*(2), 5–28.

Hall, R. E. and C. I. Jones (1999). Why do Some Countries Produce So Much More Output Per Worker than Others? *The Quarterly Journal of Economics 114*(1), 83–116.

Harstad, B. (2016). Technology and Time Inconsistency.

Harstad, B. (2019). Pledge-and-Review Bargaining. *CESifo Working Paper*.

Hughes, J. E. and M. Podolefsky (2015). Getting Green with Solar Subsidies: Evidence from the California Solar Initiative. *Journal of the Association of Environmental and Resource Economists 2*(2), 235–275.

Jefferson, T. (1816). Thomas Jefferson to Samuel Kercheval.

Jones, B. F. and B. A. Olken (2005). Do Leaders Matter? National Leadership and Growth since World War II. *The Quarterly Journal of Economics 120*(3), 835–864.

Kerin, R. A., P. R. Varadarajan, and R. A. Peterson (1992). First-Mover Advantage: A Synthesis, Conceptual Framework, and Research Propositions. *Journal of marketing 56*(4), 33–52.

Knack, S. and P. Keefer (1995). Institutions and economic performance: Cross-country tests using alternative institutional measures. *Economics and Politics 7*(3), 207–227.

Leiserowitz, A., E. Maibach, S. Rosenthal, J. Kotcher, A. Gustafson, P. Bergquist, M. Ballew, and M. Goldberg (2018). *Energy in the American Mind, December 2018*. Yale University and George Mason University. New Haven, CT: Yale Program on Climate Change Communication.

Lieberman, M. B. and D. B. Montgomery (1988). First-mover advantages. *Strategic management journal 9*(S1), 41–58.

Marx, K. (1852). The Eighteenth Brumaire of Louis Bonaparte.

McGlade, C. and P. Ekins (2015). The geographical distribution of fossil fuels unused when limiting global warming to 2 c. *Nature 517*, 187–190.

North, D. C. (1991). Institutions. *Journal of Economic Perspectives 5*(1), 97–112.

North, D. C. and B. R. Weingast (1989). Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-century England. *The Journal of Economic History 49*(4), 803–832.

Rodrik, D., A. Subramanian, and F. Trebbi (2004). Institutions Rule: the Primacy of Institutions over Geography and Integration in Economic Development. *Journal of Economic Growth 9*(2), 131–165.

Rosenberg, N., L. E. Birdzell, and G. W. Mitchell (1986). *How the West Grew Rich*. Popular Prakashan Mumbai.

Segura-Cayuela, R. (2006). Inefficient Policies, Inefficient Institutions and Trade. working paper.

Shapley, L. S. and M. Shubik (1969). On the Core of an Economic System with Externalities. *The American Economic Review 59*(4), 678–684.

Smith, A. and D. Stewart (1963). *An Inquiry into the Nature and Causes of the Wealth of Nations*, Volume 1. Wiley Online Library.

Sokoloff, K. L. and S. L. Engerman (2000). Institutions, Factor Endowments, and Paths of Development in the New World. *Journal of Economic Perspectives 14* (3), 217–232.

Thaler, R. H. and S. Benartzi (2004). Save More Tomorrow: Using Begavioral Economics to Increase Employee Saving. *Journal of Political Economy 112*(S1), S164–S187.

Varian, H. R. (1994). Sequential contributions to public goods. *Journal of Public Economics 53*(2), 165–186.