Robust policy schemes for differential R&D games with asymmetric information*

Anton Bondarev†

Abstract

I consider a general differential r&d game with finite set of firms which may generate a multiplicity of strategically-driven outcomes like prevention of entry or strategic delay. It is assumed that while interacting firms are fully aware of their potentials, the government is uncertain over the true state of the industry and thus may be unable to predict such strategic behavior. The choice function over the set of potential outcomes is defined and robust welfare optimal and individually optimal regimes are compared. The belief of the planner over the market shapes the set of potential policy schemes. The ordering of such schemes with respect to uncertainty, costs and welfare-improving potential is established. At last the optimal level of robustness for a given policy is found.

Keywords: axiom of choice; strategic interaction; robust policy sets; welfare ordering; optimal robustness

JEL classification: C02, C61, O31, O38

*This research is part of the activities of SCCER CREST (Swiss Competence Center for Energy Research), which is financially supported by the Swiss Commission for Technology and Innovation (CTI) under contract KTI.2014.0114.

†Department of Business and Economics, University of Basel, Peter Merian-Weg 6, 4002 Basel, Switzerland, e-mail: anton.bondarev@unibas.ch

1 Introduction

It is now widely acknowledged that in many cases the government regulation cannot achieve the first-best outcome due to the lack of information on the market developments. This is even more true for the cases when the market under consideration is dynamic. In this case the robust control theory is needed to define welfare-improving policies. Explicit treatment of the dynamic nature is particularly important in cases of R&D and environmental policy issues, as in Ben-Youssef and Zaccour (2014).

Numerous studies have shown that almost unavoidable market failures can lead to a technology lock-in; typical examples are lock-ins caused by market power that is due to patents for new technologies (see, e.g., Krysiak (2011)) or externalities caused by network effects in technology adoption (see Arrow (1962)), Arthur (1989)), Unruh (2000), or Unruh (2002)). In such cases, it is not sufficient to only set a price for environmental damages to ensure that the best clean technologies are developed; more specific incentives are necessary, (e.g. Krysiak (2011)).

The size and duration of such specific interventions will typically depend strongly on different cases of market failures. For example, the development of a new promising technology might only be delayed or it could be prevented completely, rendering different interventions necessary (see e. g. Bondarev and Krysiak (2017), Ludkovski and Sircar (2016)).

Robust control has been used in a number of such applications in environmental and energy economics. Studies on climate negotiations use robust optimization, as in Babonneau et al. (2013) or Ben-Tal et al. (2009). The robust control approach has been used to investigate government interventions in environmental problems, in particular related to the precautionary principle, as, for example, in Athanassoglou and Xepapadeas (2012) or Vardas and Xepapadeas (2010). Other applications are found, for example, in asset management, see Vardas and Xepapadeas (2015).

These studies are based on the minmax approach, where a planner tries to minimize certain threat and the realization of this threat is given by uncertain variables chosen by a malevolent nature to maximize damages. One particularly interesting formal analysis of the robustness approach is given by Todorov (2009), where a Kullback-Leibler entropy measure is used.

Somewhat different approach based on model uncertainty is used in Gonzalez (2018). There the robust decision-making is based on Markovian chains and the planner is not certain over the true dynamics of the system (time-varying model mistrust).

Current paper contributes to both directions in this literature by developing a generalized robustness approach, where the robustness criteria naturally follows from the choice function over the set of outcomes. Namely the min-max approach is used to define the social welfare under uncertainty, but the underlying dynamic game is allowed to have arbitrary many outcomes other than the standard decentralized equilibrium. This feature may be looked upon as a version of the model mistrust, but the multiplicity of market outcomes are defined by the optimizing logic of firms.

The closest paper to the present approach is Brock et al. (2014), where the notion of hot spots is introduced to mark the cases where uncertainty may break down the regulation or lead to instability of the underlying system. In the current contribution the general idea of such hot spots is further elaborated and it is demonstrated that under mild conditions there generically exists a sequence of such uncertainty thresholds. These thresholds separate domains of efficiency of policy schemes with increasing robustness levels.

The main contribution of the paper is the development of a general theory of robust policy which may be applied to a variety of cases. This theory differs from previous approaches in minimal assumptions being made (only countability of outcomes and axiom of choice are needed) and provides versatile tools for construction of (suitably defined) optimal robust policy sets. In particular, the main theorem of algebra is used to defined the potential set of outcomes of the dynamic game both in decentralized and socially optimal cases. Thresholds separating those regimes turn to be roots of associated algebraic polynomials and the selector over the set of outcomes is then defined via a choice function over sets.

These abstract results are of interest since they allow the construction of robust policies in a fairly general setting, requiring only the countability of potential outcomes and the axiom of choice (see e. g. Herrlich (2006)). At the same time in applications with finitely many regimes the sequence of increasingly robust policies may be easily constructed.

The rest of the paper is organized as following. Section 2 describes the general setup, Section 3 contains all main results of the paper and Section 5 concludes.

2 The model

The general setup consists of dynamically interacting market players (firms) and the government authority. To specialize, I study the underlying market as an r&d differential game in the spirit of (Ben-Youssef and Zaccour, 2014), (Bondarev, 2014) neglecting any production side dynamics, which can be easily incorporated as in (Krysiak, 2011).

Any alternative setup may be used, but for illustration purposes I develop the framework in terms of the R&D game and R&D policy. In particular assume that in the production sector, there is perfect competition and final producers are price takers. In the R&D sector, the firms get a patent for their developments and are thus monopolistic suppliers of their technology. Some of the firms have an initial advantage (their technologies being somewhat more developed initially) and thus might act strategically to forestall the use and development of the new technologies.

In the analysis which follows I abstract from further market imperfections such as environmental externalities, assuming it is already taken care about by proper remuneration schemes in case technologies at hand are dirty and clean ones or both are green. By doing so I apply this study to the case of general innovations setting with green technologies being a specific (but rather important) example of those.

2.1 The general r&d game

In the r&d sector, there is a finite $N \subset \mathbb{N}$ number of firms. Each firm $j \in N$ can invest in r&d and set prices for its own technology. Owing to the patent, each firm is the sole supplier of its technology, thus the market is monopolistically competitive one.

Denote with q_j the state of technology j (subject to investments), p_j associated price, g_j investments of firm j into technology's development.

Without loss of generality further denote $\mathcal{O}$ the set of all possible outcomes of this game and by $\mathcal{F}$ the set of feasible outcomes (under given initial conditions and parameters vectors). We note that these sets have finitely many elements as long as the number of players N is finite. Examples of those elements are issues of strategic behaviour of one or more firms on the market, leading to temporary/permanent prevention of entry of new firms. Denote further $m \in \mathcal{O}$ an arbitrary outcome of the r&d game and by $q_j^m(t)$ the state of technology j in regime m at time t.

In the following I assume the following general properties of the game:

Assumption 1.

- *Free entry condition holds with no sunk costs of entry*[1]*.*
- *Costs c_j associated with development of technology j depend only on firm's j own investments (but can be heterogeneous across firms).*
- *Prices p_j are continuous functions of the state vector of technologies of market participants, $\vec{q}$.*
- *There exists (unique) equilibrium state vector $\vec{q}^m$ of the r&d game, to which technologies converge in the long run in regime m*[2]*.*

The evolution of technology j may depend on own firm's investments as well as on investments of other firms and on the current state of own technology and technologies of other firms:

$$\forall j \in N : \; \dot{q}_j(t) = f_j(q_j(t), q_{-j}(t), g_j(t), g_{-j}(t)), \tag{1}$$

[1] The absence of sunk costs is crucial for the results, since if these are present, Nash equilibrium logic would prevent any strategic behavior of incumbent firms: indeed, they can always price optimally with the threat of lowering price upon new entry, thus preventing the entry of new technology due to not negligible sunk costs

[2] of course this vector may include infinite elements if some technologies do not have a steady state, but I restrict the analysis to those cases where parameters' space allow only for finite values of this vector.

The objective of every firm is to maximize its discounted stream of profits (value) for a given discount rate r choosing optimally price schedule and investments:

$$J_j = \max_{p_j, g_j} \int_0^\infty \mathrm{e}^{-rt} \left\{ \pi_j(p_j(t), q_j(t), p_{-j}(t), q_{-j}(t)) - c_j(g_j(t)) \right\} dt. \tag{2}$$

where here and throughout the paper index j denotes player's j quantitative and index $-j$ quantities associated with all other players except j. profit of firm j, π_j may be a function of prices of all firms as well as of technologies.

The dynamics of all technologies (1) forms a controlled N-dimensional ODE system with state-space and control space restrictions:

$$\forall j \in \mathbb{N}: \; q_j \geq 0; \; g_j \geq 0; \; p_j \geq 0 \tag{3}$$

So that both the state of every technology and associated investments are non-negative[3].

Given some final producers' demand system

$$\forall j \in N: \; Q_j^D = F_j(p_j, p_{-j}, q_j, q_{-j}) \tag{4}$$

for N firms present at the market, we get N reaction functions for prices of technologies j and as a result an N-dimensional system for price schedules as functions of technology states of all the firms, $\forall j \in N: \; p_j = w_j(q_j, q_{-j})$.

We thus reduce the problem of N firms given by (1), (2) to the differential game over technologies states q_j with controls g_j. To keep the constructive nature of the exposition I limit myself to the open-loop solution concept, since closed loop one does not always exist. As long as the controlled system (1) permits for the solution vector $\vec{q}^{*}$[4] denote

$$\Pi_j^*(\vec{q}(0)) = \max_{g_j} \int_0^\infty \mathrm{e}^{-rt} \left\{ \pi_j(\vec{p}(\vec{q}^*), \vec{q}^*(t),) - c_j(g_j(\vec{q}^*)) \right\} dt \tag{5}$$

value function of the firm j as a function of initial states of all technologies $j \in N$ with superscript $*$ denoting the simultaneous development regime (solution of N-players differential game).

[3] this is not essential, but follows standard economic intuition

[4] e. g. f_j are Lipschitz and bounded, then via Picard-Lindelöf theorem for each admissible $\vec{g}$ the system (1) admits a solution.

2.2 Government

The government has the objective of maximizing the net social benefit from all the technologies. This net social benefit consists of a marginal benefit β_j attached to each unit of production with technology j minus locational costs, minus the costs of developing the technologies. For simplicity, assume that the social planner uses the same discount rate r as the r&d firms[5].

Social welfare is thus given by:

$$W_m := \int_0^\infty \mathrm{e}^{-rt} \sum_{j=1}^{N} \left\{ \beta_j Q_j^m(t) \left(q_j^m(t) + \Xi_j^m(t) \right) - c_j \left(g_j^m(t) \right) \right\} dt, \tag{6}$$

where $\Xi_j(t)$ denotes the average effect of used locations on output for technology j. In general these are functions of prices and technologies' states, p_j, p_{-j}, q_j, p_{-j}.

The government may use two policy instruments. First, there is the remuneration for green production (measured by β_j). Second, it might subsidize the initially disadvantaged technologies for some time.

These instruments have to correct three market failures, two of which are well known. First, without the remuneration, there would be no green production. Second, as r&d firms have market power, they will set socially suboptimal prices and thus green technologies will be used less than in the social optimum. Finally, the firms developing the initially more advanced technologies might use their market power to set prices that keep the other firms from investing into developing the newer technologies.

In this paper, I will focus on this third problem, because this is new and could be particularly detrimental, as the development of a technology with high potential might be prevented indefinitely. We thus directly consider the point raised in the introduction: Should a government only provide a general incentive for using green technologies (such as a price for GHG emissions) or should it also steer technological change by using technology-specific subsidies?

[5] In general equilibrium models such an assumption would be tedious, but the framework developed here is aimed on industrial level regulation studies and as such does not take the difference in time preferences into account following many IO papers like Dawid et al. (2010). However, this can be easily incorporated into the model.

2.3 One-sided uncertainty

To investigate a setting that is both scientifically interesting and practically relevant, I focus on the case, where the government knows initial technology vector $\vec{q}(0)$ which is invariant across regimes of the game, but does not know the long-run potential of some of the new technologies $i \in I \subset N$. Firms themselves are fully informed over characteristics of both old and new technologies. This is often the case in real industries, since industry players put more efforts into learning their competitors capabilities than the regulating authority. Without loss of generality assume further that at any moment there is exactly one new technology $i \in N$ which is entering the market and the planner does not know its potential, but has some limited information about boundaries of this potential[6]. Since I do not specify the dynamics of technologies, (1) I also restrain from specifying where this uncertainty comes from. It is sufficient for our purposes to assume that at any moment in time government knows the state of the new technology i with some certainty:

$$q_i(t) \in [q_i(t) - \epsilon, q_i(t) + \epsilon],\ \varepsilon \overset{iid}{\sim} \mathcal{N}(0, \sigma_\varepsilon),\ \sigma_\varepsilon = \epsilon^2 \tag{7}$$

and the government is not able to learn the true state over time (otherwise the problem becomes trivial). In a general setting this amounts to the statement that noise vector $\vec{\epsilon}_W$ is non-zero for at least some technologies (i. e. its norm is positive).

As will be showed, an increasing value of ε implies that the government is less able to differentiate between different cases of strategic and non-strategic behavior of the incumbent firms and thus to ascertain that and how long a subsidy is required. We thus look for a robust policy that can cope with several cases at once.

2.4 Standing assumptions

Throughout the rest of the paper in addition to the list of assumptions on the properties of the game, (Assumption 1), three standing assumptions are employed:

[6] indeed this is not a binding assumption, since in the case of simultaneous entry of several firms the problem may be decomposed into the sequence of problems with single entry, see discussion on the sequence of pairwise games of the leader and the follower in Bondarev and Greiner (2017) and in the current paper later in Sec. 3

Assumption 2. *Value functions of all players Π_j^m and social welfare functions W_m have countably many zeros as functions of initial states*[7]*.*

This assumption is necessary to make sense of thresholds separating regimes of the game. It simply means that every value function does have only countably many regions where it is positive or negative with respect to the initial states of technologies separated by (isolated) zeros. In this case we can consider differences in these values across regimes of the game and study those intervals of initial conditions where one or the other outcome is preferred by players and/or the social planner. If Assumption 2 does not hold, some of value functions may have uncountably many zeros which are not isolated implying we could no longer define (dense) regions in the state space where the given regime of the game admits positive value and the approach developed in this paper cannot be applied. Observe that this assumption is not very restrictive, since all value functions considered in economic applications so far enjoy this property. However once we consider stochastic dynamic games, this assumption may fail which is one of the reasons of not studying the fully stochastic game setting here.

Assumption 3. *The (weak version of) axiom of choice holds, i. e. given any selector function it is possible to choose an element of the set with countably many elements.*

This is implicitly assumed in almost all economic papers, but there exist alternative settings providing different mathematical theories (see Herrlich (2006) for example). At the same time if this holds it is guaranteed, that the choice function may be defined for any countable collection of sets.

Assumption 3 plays a crucial role in the analysis that follows. In particular it guarantees that given some predefined rule of choice (choice function) we are able to select exactly one element from every set in a (countable or uncountable) collection of sets. I do not require the original axiom of choice which states that this is possible for any uncountable collection of sets, but only require a countable counterpart, since otherwise

[7]Observe that any analytic function of a real argument enjoys this property, but I do not restrict attention to analytic functions only, since piecewise or absolute value functions are known to be not analytic.

the results even if hold would become too cumbersome to prove. This is the reason of imposing Assumption 2 on value functions of players. Axiom of choice then enables us to select the best outcomes in a sense that a relationship of mutual dominance of (countably many) values of these outcomes of the general game may be established.

In the absence of the axiom of choice many standard mathematical tools would stop working or at least would require different formulations. For example, the differential calculus becomes essentially different (see the book Robert (2011) for example). There are many alternative versions of this axiom with different equivalence relations. It could be possible to reconstruct some of results of the paper employing alternative axioms, but this is the task well beyond a current study.

Assumptions 2 and 3 together is all we need to develop the theory of robust policy schemes for asymmetrically informed agents.

Define next the noise vector for player j as a confidence interval over the state of the vector of technologies:

$$\forall j \in N: \ \vec{q}(t) \in [\vec{q}(t) - \vec{\epsilon}_j, \vec{q}(t) + \vec{\epsilon}_j], \vec{\varepsilon}_j \overset{iid}{\sim} \mathcal{N}_N(0, \Sigma_{\varepsilon_j}) \tag{8}$$

Concerning the information the player is fully informed if its noise vector over the state of all the technologies is zero. In general asymmetry of information is exhausted by the assumption:

Assumption 4. *The subset of fully informed players is non-empty. That is* $\exists j \in N$: $\vec{\epsilon}_j = 0$.

Indeed one could in principle allow for any countable set of less informed players (not only the government). What is important, there must be a reference point, that is, at least one player makes optimal decisions on the basis of full information.

In particular, assume all R&D firms know with certainty technology characteristics of each other, but the government experiences some uncertainty over the potential of some of the technologies (the newer ones) in a form of (7). The government may implement a subsidizing scheme to prevent the strategic behavior of the more developed technology owners, but does so only if this is welfare improving.

3 Analysis of the model

3.1 Characterization of the r&d game outcomes

I first characterize the multiplicity of outcomes of the underlying r&d game under full information on behalf of competing firms. Once the solution for the game exists, it defines the vector of value functions of participating firms as a function of initial states of all the technologies, (5). Those firms, which have initial advantage, $l \in N : q_l(0) > q_j(0)$ may choose strategic behavior to keep competitors out of the market, creating the multiplicity of outcomes. Whether or not such strategic behavior is optimal depends on comparison of values generated by competitive and strategic behavior for every such firm l. Assume for certainty there is a ranking of initial technologies states, such that:

$$q_1(0) \geq q_2(0) \geq .. \geq q_N(0) \tag{9}$$

so that the firm 1 has initial advantage over all other firms and the next firm has advantage over the rest of $N-2$ firms, etc. Then firm 1 decides whether or not to implement strategic pricing and at which level as following: it may set the price at the level p_1^1 such that the profit for all other firms is zero if they enter the market and keep this schedule for some time. If this turns to be not feasible or not profitable, it can set the price at the level p_1^2 as to keep all competitors except the closest one out of the market. Continuation of this argument yields a descending sequence of strategic prices for firm 1, $p_1^1 < p_1^2 < .. < p_1^N$ such that the latest strategic price keeps off the market only the firm N.

If the first firm sets the strategic price p_1^i allowing for entrance of $i < N$ firms, those firms upon entrance may play the best response price or again act strategically. However the leader under full information may predict actions of all the followers and sets the strategic price effectively determining the number of competitors. We thus may reduce the problem to the case of only two firms, since every next competitor just repeats the decision process of preceding firms upon the strategic price setting.

Denote the incumbent firm j and the new entrant (closest competitor) $-j$. I do not specify the number of potential strategic regimes other than assuming them to be finite in number. Denote by Π_j^m the value generated by the outcome m of the underlying r&d

game for the firm j. Observe next that from the above discussion it follows that it is the leader (firm with maximal initial state of technology) which defines the regime of the game. At last note that value function of any firm j is a function of initial states of technologies and parameters only. Denote by $\delta(0) = q_j(0) - q_{-j}(0)$ the initial technological gap between the leader and the closest competitor (potential entrant). Denote further by $\delta_z^j(m)$ the z-th real-valued root yielding to zero the difference in values generated by the game for the firm j across outcomes i, m[8]. The set of such roots is given by real-valued solutions of the equation:

$$\{\delta_z^j(m)\}: \; \Pi_j^i(\delta(0)) - \Pi_j^m(\delta(0)) = 0 \tag{10}$$

This set by Assumption 2 hoes countably many elements.

We then may characterize the comparison of different outcomes of the game in terms of such roots of the algebraic equation (10)[9]:

Proposition 1 (Algebraization of r&d game outcomes)**.**
The outcome $i \in \mathcal{F}$ of the r&d game is individually optimal if for a given initial state $\delta(0)$ there exists the choice function for the leader j:

$$\exists\Theta(\mathcal{F}) : \delta(0) \in \bigcup_{m\in\mathcal{F}} \bigcap_{m\in\mathcal{F}} \{[\delta_z^j(m); \delta_{z+1}^j(m)] | \Pi_j^i(\delta(0)) - \Pi_j^m(\delta(0)) \geq 0\} \implies$$
$$\Theta(\mathcal{F}) = \arg\max_{m\in\mathcal{F}} \Pi_j^m = i. \tag{11}$$

Proof. see Appendix A □

As examples of elements of $\mathcal{O}$ we might think of cases of strategic pricing, preventing the entry of a competitor, temporary delay of such an entry or any other behavior different from the normal equilibrium of the game. Whether or not such behavior is feasible, defines the size of the set $\mathcal{F}$ and the choice function simply selects the outcome (and type of behavior) giving the highest accumulated profit (value) to a given firm. This does not

[8] I assume the value functions of all potential regimes has finitely (or countably many) complex roots, which is the case for any value function over $\delta(0)$ satisfying Ass. 2

[9] Because of this characterization of outcomes through algebraic roots I termed the approach as the *algebraization* i. e. conversion to algebraic form

put any constraints on the pre-commitment (firm may decide to stop strategic pricing at any time or renew it) since every element in $\mathcal{F}$ is the optimal value for a given time-varying strategy of the firm j. For a particular application of such an approach one may consider the multiplicity of outcomes in Bertrand competition studied in Bondarev and Krysiak (2017).

We now move to the uncertain part of the problem. Since government experiences uncertainty over the true potential of technology $-j$, it cannot assign the first-best subsidy as usual[10]. I thus start with definition of social policy in this setup, then define social optimality under uncertainty and work out what is called robust policy schemes preventing strategic behavior.

3.2 Social welfare under uncertainty

Assume from now on initial level of technology j being fixed as well as other parameters of the model except for $q_{-j}(0), Q_{-j}$.

First define the social optimality measure under robustness ϵ for given q_{-j}[11]:

Definition 1 (Social optimality under uncertainty).
The outcome of the r&d game $s \in \mathcal{O}$ is (weakly) social welfare improving over the outcome m with robustness level ϵ if

$$\forall \varepsilon \in [-\epsilon, \epsilon] : \min_{\varepsilon \in [-\epsilon, \epsilon]} \{W_s^\varepsilon(q_{-j})\} \geq \min_{\varepsilon \in [-\epsilon, \epsilon]} \{W_m^\varepsilon(q_{-j})\} \tag{12}$$

it is strongly welfare improving if

$$\forall \varepsilon \in [-\epsilon, \epsilon] : \min_{\varepsilon \in [-\epsilon, \epsilon]} \{W_s^\varepsilon(q_{-j}) - W_m^\varepsilon(q_{-j})\} \geq 0 \tag{13}$$

where $W_{s,m}^\varepsilon(q_{-j})$ are given by integral Eq. (6) *with $Q_{-j} = Q_{-j} + \varepsilon$.*

The outcome s is (weakly) socially optimal with robustness level ϵ if

$$\forall \varepsilon \in [-\epsilon, \epsilon], \forall m \in \mathcal{O} : \arg\max_{m \in \mathcal{O}} \min_{\varepsilon \in [-\epsilon, \epsilon]} \{W_m^\varepsilon(q_{-j})\} = s \tag{14}$$

[10]in the first-best case the government plays a Stackelberg differential game being the first mover and setting the subsidy size and duration such as to maximize social welfare

[11]I omit time argument in integral quantities W, Π and understand $q_{-j} = q_{-j}(0), \delta = \delta(0)$ in what follows if this does not lead to confusion.

it is strongly welfare optimal if

$$\forall \varepsilon \in [-\epsilon, \epsilon], \forall m \in \mathcal{O}: \ \arg\max_{m\in\mathcal{O}} \min_{\varepsilon\in[-\epsilon,\epsilon]} \{W_s^\varepsilon(q_{-j}) - W_m^\varepsilon(q_{-j})\} = s \tag{15}$$

The definition of social optimality (14) requires to obtain minimum possible welfare for every regime m over realization of the noise ε and then to take the maximum across regimes. The regime which provides maximal welfare under the most unfavorable circumstances ($\min_\varepsilon$) is socially optimal with certainty (robustness) level ϵ, if its minimum strictly dominates other minima. Such a definition is in line with others used in the literature (e. g. Babonneau et al. (2013), Brock et al. (2014)) assuming the worst-case realization and in line with model mistrust literature (e. .g Gonzalez (2018)) in accounting for all potentially realizable outcomes (set $\mathcal{O}$).

Strong optimality requires that regime s has higher welfare under the most unfavourable circumstances then other regimes have under the most favourable ones, since $\min\{x-y\} = \min\{x\} - \max\{y\}$. Apparently the strong optimality holds if there are no intersections of welfare functionals as functions of $\varepsilon \in [\pm\epsilon]$ and corresponds to the case when uncertainty is *inessential.*

It is straightforward that under $\epsilon \to 0$ the full certainty social welfare difference $\min_\varepsilon\{W_s^\varepsilon(q_{-j})\} - \min_\varepsilon\{W_m^\varepsilon(q_{-j})\} = W_s(q_{-j}) - W_m(q_{-j})$ is recovered.

To establish social optimality it thus suffices to consider the differences in social welfare across different regimes of the r&d game.

$$\forall \in \mathcal{O}: \ D_{s,m}(W) \stackrel{def}{=} W_s(q_{-j}) - W_m(q_{-j}) \tag{16}$$

and their robust counterparts as:

$$D_{s,m}^\epsilon(W) \stackrel{def}{=} \min_{\varepsilon\in[-\epsilon;\epsilon]}\{W_s^\varepsilon(q_{-j})\} - \min_{\varepsilon\in[-\epsilon;\epsilon]}\{W_m^\varepsilon(q_{-j})\} \tag{17}$$

The regime which is robust welfare optimal would yield positive differences with all other regimes m, but we cannot apply max operator over these differences to select the best outcome as in Def. 1.

To obtain such a procedure to select the socially optimal robust outcome I establish the result concerning social welfare. To this end we make use of the algebraization tools applied to the difference in minima, (17).

First observe that robust social welfare values, $W_m^\epsilon(q_{-j}) = \min_{\varepsilon\in[-\epsilon;\epsilon]}\{W_m^\varepsilon(q_{-j})\}$ may be treated as analytic functions in $q_{-j}(0)$ since min operator gives a unique ε value, which is simply

$$\varepsilon_m^\epsilon = \arg\min_{\varepsilon\in[-\epsilon;\epsilon]}\{W_m^\varepsilon(q_{-j})\} \tag{18}$$

by the theorem on the average value of a function. However, these values are in general different for different m, thus $D_{s,m}^\epsilon(W)$ depends on both $\varepsilon_m^\epsilon, \varepsilon_s^\epsilon$ values. Still it is analytic in $q_{-j}(0)$ with coefficients depending on given robustness level ϵ, since $\varepsilon_{m,-m}^\epsilon \in [\pm\epsilon]$.

For each outcome s denote roots of differences in social welfare as functions of $q_{-j}(0)$ by $q_z^m(\epsilon)$ where z is the index of the root such that $q_{z+1}^m(\epsilon) > q_z^m(\epsilon)$, i. e.:

$$\{q_z^m(\epsilon)\}: \; W_s^\epsilon(q_{-j}) - W_m^\epsilon(q_{-j}) = 0. \tag{19}$$

The full certainty case is recovered with $\epsilon \to 0$.

The following Proposition provides criteria for selecting the socially optimal outcome:

Proposition 2 (Social welfare algebraization under uncertainty)**.**
The outcome $s \in \mathcal{F}$ of the r&d game is socially optimal among outcomes $\mathcal{F} \subseteq \mathcal{O}$ under robustness level ϵ if $q_{-j}(0)$ lies in the union of intervals where social welfare is higher under outcome s than under any other $m \in \mathcal{F}$, i. e. there exists the choice function:

$$\begin{aligned}
&\Psi_\epsilon(\mathcal{F}) : q_{-j} \in \bigcup_{m\in\mathcal{F}}\bigcap_{m\in\mathcal{F}}\{[q_z^m(\epsilon); q_{z+1}^m(\epsilon)] | D_{s,m}^\epsilon(W) \geq 0\} \implies \\
&\Psi_\epsilon(\mathcal{F}) = \arg\max_{m\in\mathcal{F}}\min_{\varepsilon\in[-\epsilon;\epsilon]} W_m^\varepsilon(q_{-j}) = s.
\end{aligned} \tag{20}$$

The outcome of the game s is robust welfare maximizing up to the level ϵ_s^W if $\forall\epsilon < \epsilon_s^W$: $\Psi_\epsilon(\mathcal{F}) = s$.

Proof. As long as the worst-case outcome s is better than the worst-case outcome m for given ϵ, it follows that $D_{s,m}^\epsilon(W) \geq 0$. These objects are analytic functions in $q_{-j}(0)$ (see discussion around (18)), depending on robustness ϵ. Hence they have at most countable zeros (roots given by (19)) which are isolated. Thus roots $q_z^m(\varepsilon_{s,m}^\epsilon)$ form a sequence of intervals where outcome s is social welfare improving or not over m. Select those ranges of $q_{-j}(0)$ which yield positive value for this polynomial. Repeat this process for all $m \in \mathcal{F}$. Outcome s is better than any collection of other outcomes from $\mathcal{F}$ as long as all differences $D^{\epsilon}_{s,m}(W)$ are positive (since Def. 1). Ranges of $q_{-j}(0)$ where this condition remains valid are given by the union of all intervals associated with positive difference for all $D^{\epsilon}_{m,s}(W)$. Union of those intervals gives the total range, where outcome s is better than any $m \in \mathcal{F}$ hence (20). The last claim is just an observation that choice function depends on the uncertainty level: once we change ϵ, it could be the case that $q_{-j}(0)$ no longer lies in intervals of positive sign and outcome s is no longer maximizing worst-case welfare (although it still can be optimal in full certainty case). □

This proposition gives the criteria for comparing any regimes in terms of social welfare of the R&D game for fixed ϵ: we need to compute values W^{ϵ}_{m} , and then compare them for the given $q_B(0)$. As long as the difference between $m, -m$ functions W is positive, regime m is robust welfare improving over regime $-m$. Thus computing roots of this difference provides the range of $q_B(0)$, for which this ordering holds. Since ϵ is fixed, these roots are functions of parameters and uncertainty level. Thus for any given error size the ordering or regimes of the underlying r&d game in terms of social welfare can be established. Formally speaking the relationship (20) provides a selector function in the space of functions type $D^{\epsilon}_{m,-m}(W)$: once the condition of positive difference is fulfilled, it checks whether given $q_B(0)$ falls into one of the provided intervals.

Observe also, that the choice of robust optimal regime depends both on the size of the set $\mathcal{F}$ and the noise level ϵ: it could be the case that competitive outcome is better than strategic forestall for any ϵ, but not so if comparing with the monopolistic development.

We also need the robust criteria of individual optimality. It is done in the same way as for social welfare, albeit for profit functionals of the players. Denote by $\Pi^{m}_{j}(\varepsilon)$ total profit of player j in regime m under uncertainty level ε. Observe that this is valid only for social planner, since players do not experience uncertainty at all.

Definition 2 (Robust outcome of the r&d game)**.**
The outcome i of the r&d game is believed to be (weakly) individually optimal across (feasible) outcomes $\mathcal{F} \subseteq \mathcal{O}$ for the leader j with robustness ϵ, if

$$\arg\max_{m \in \mathcal{F}} \min_{\varepsilon \in [\epsilon;\epsilon]} \Pi^{m}_{j}(\varepsilon) = i \tag{21}$$

It coincides with actual realization (the belief is robust)

$$\arg\max_{m\in\mathcal{F}} \Pi_j^m = i = \arg\max_{m\in M} \min_{\varepsilon\in[\epsilon;\epsilon]} \Pi_j^m(\varepsilon) \tag{22}$$

if it is strongly individually optimal:

$$\arg\max_{m\in\mathcal{F}} \min_{\varepsilon\in[\epsilon;\epsilon]} \{\Pi_j^i(\varepsilon) - \Pi_j^m(\varepsilon)\} = i \tag{23}$$

Comment: By the belief in this definition the belief of the social planner is understood (since this is the only one not fully informed actor in this setup). This belief is measured by the confidence level ϵ the planner puts into its knowledge of the state of the new technology.

This definition provides the criteria for a planner, how to define which regime of the game to expect in the absence of regulation. Still, as the second part points out, the believed regime is not always the actual one, so there is a room for mistake which is that higher, the higher is ϵ.

Denote with $\delta_z^m(\epsilon)$ the z-th root (zero) of the equation

$$\{\delta_z^m(\epsilon)\} : \min_{\varepsilon\in[-\epsilon;\epsilon]} \{\Pi_j^i(\varepsilon)\} - \min_{\varepsilon\in[-\epsilon;\epsilon]} \{\Pi_j^m(\varepsilon)\} = 0 \tag{24}$$

which is the robust counterpart of (10).

Corollary 1 (Algebraization of robust outcomes of the r&d game)**.**
The outcome $i \in \mathcal{F}$ of the r&d game is expected to realize among outcomes $M \subseteq \mathcal{O}$ with robustness level ϵ if $\delta(0)$ lies in the union of intervals where worst-case profit is higher under outcome s than under any other $m \in M$ for player j (denoted as the leader), i. e. there exists the choice function:

$$\Theta_\epsilon(\mathcal{F}) : \delta(0) \in \bigcup_{m\in\mathcal{F}} \bigcap_{m\in\mathcal{F}} [\delta_z^m(\epsilon); \delta_{z+1}^m(\epsilon)] | \min_{\varepsilon\in[-\epsilon;\epsilon]} \{\Pi_j^i(\varepsilon)\} - \min_{\varepsilon\in[-\epsilon;\epsilon]} \{\Pi_j^m(\varepsilon)\} \geq 0 \implies$$

$$\Theta_\epsilon(\mathcal{F}) = \arg\max_{m\in M} \min_{\varepsilon\in[-\epsilon;\epsilon]} \Pi_j^m(\varepsilon) = i. \tag{25}$$

The outcome i is the robust realization of the r&d game with certainty level ϵ_i^O if $\forall \epsilon < \epsilon_i^O : \Theta_\epsilon(\mathcal{F}) = i$.

Proof. Amounts to application of results of Prop. 2 to the value functions of the underlying r&d game: Given confidence level ϵ compute the value function of player j under the worst case scenario (minimal profit). Since by Assumption 2 these are analytic functions, they have countably many zeros. Select those intervals with positive difference in these worst-case values and consider their intersections for all outcomes m. Once $\delta(0)$ (which is known to the planner) falls within a union of these intersections, outcome i is believed (in the sense of Def. 2) to be chosen by the players. At last, this selection depends on the chosen level of certainty, ϵ_i^O, since for higher noise level some of intervals in (25) may cease to exist or change signs. □

3.3 Robust subsidies under uncertainty

The social planner may implement the policy scheme consisting of the subsidy and its duration to one of the players to prevent strategic behavior. Under full certainty the first-best subsidy might be implemented, but under the uncertain potentials of technologies this is not the case. The implementation of a subsidy follows multiple steps:

1. Social welfare is computed for all possible regimes of the game, and the best one in the sense of Prop. 2 is selected;
2. The expected regime of the game is defined via Cor. 1;
3. Subsidy is assigned to one of the players in such a way, as to incentivize players to switch to the desired regime;
4. Social welfare for the resulting regime (subsidized) is computed and checked against the otherwise realised non-perturbed regime and profit incentives of players in the resulting regime.

I thus require the subsidy to be robust and social-welfare improving, but not necessarily optimal. Under uncertainty the planner does not know profit incentives of players, but only expected ones subject to the error ε. Hence the definition of socially desirable robust subsidy:

Definition 3 (Robust welfare-optimal policy scheme)**.**
A policy scheme is the pair $\Sigma_k : \{\sigma_k, t^k\}$ *which defines size* (σ_k) *and duration* (t^k)*of the subsidy assigned to player* $-j$ *for certainty.*

For each ϵ *the robust welfare-improving policy scheme* $\Sigma_k^\epsilon(i,s)$ *switching the game from* i *to* s *is characterized by following:*

1. *Regime* i *is expected to realize without the subsidy in the sense of Def. 2 but regime* s *is socially optimal in the sense of Def. 1*

2. *The policy scheme* Σ_k^ϵ *is (weakly) social welfare-improving under uncertainty level* ϵ*:*

$$\min_{\varepsilon\in[-\epsilon;\epsilon]}\{W_s^\varepsilon(q_{-j},\Sigma_k^\epsilon)\} - \min_{\varepsilon\in[-\epsilon;\epsilon]}\{W_i^\varepsilon(q_{-j})\} \geq 0 \tag{26}$$

   It is strongly welfare improving if

$$\min_{\varepsilon\in[-\epsilon;\epsilon]}\{W_s^\varepsilon(q_{-j},\Sigma_k^\epsilon) - W_i^\varepsilon(q_{-j})\} \geq 0 \tag{27}$$

3. *The policy scheme is (weakly) robust under uncertainty level* ϵ*, if this policy allows for the prevention of switching back from the subsidized regime in all cases considered in Proposition 1:*

$$\forall m \in \mathcal{F}: \min_{\varepsilon\in[-\epsilon;\epsilon]}\{\Pi_j^s(\Sigma_k^\epsilon)\} - \min_{\varepsilon\in[-\epsilon;\epsilon]}\{\Pi_j^m(\Sigma_k^\epsilon)\} \geq 0 \tag{28}$$

   It is strongly robust if

$$\forall m \in \mathcal{F}: \min_{\varepsilon\in[-\epsilon;\epsilon]}\{\Pi_j^s(\Sigma_k^\epsilon) - \Pi_j^m(\Sigma_k^\epsilon)\} \geq 0 \tag{29}$$

The policy scheme Σ_*^ϵ *is (weakly) optimal among all (weakly) robust welfare-improving policy schemes switching from* i *to* s*,* $\Sigma_k^\epsilon \in \Sigma(i,s)$ *if*

$$\arg\max_{\Sigma_k^\epsilon\in\Sigma(i,s)} \min_{\varepsilon\in[-\epsilon;\epsilon]}\{W_s^\varepsilon(q_{-j},\Sigma_k^\epsilon)\} = \Sigma_*^\epsilon \tag{30}$$

It is strongly optimal if

$$\arg\max_{\Sigma_k^\epsilon\in\Sigma(i,s)} \min_{\varepsilon\in[-\epsilon;\epsilon]}\{W_s^\varepsilon(q_{-j},\Sigma_*^\epsilon) - W_s^\varepsilon(q_{-j},\Sigma_k^\epsilon)\} = \Sigma_*^\epsilon \tag{31}$$

where by the argument Σ_k^ϵ in social welfare I understand the social welfare obtained under policy scheme Σ_k^ϵ and by $\Pi_j^s(\Sigma_k^\epsilon)$ the value for r&d firms obtained under the given policy scheme in regime s[12].

Robustness is thus understood in this paper as the ability of a policy to perform a given task (prevention of strategic behavior) albeit crucial information is missing while preserving the social welfare at least not lower than in the worst-case scenario under alternative regimes. This concept of robustness is close to the usual min-max approach, since we use maximal confidence intervals for uncertainty and compare worst-case scenarios in terms of welfare. At the same time the suggested notion allows for immediate application due to the the algebraization approach and can be used in the setting with many alternative regimes of the model.

Since there are multiple regimes of the game possible, the set of robust and welfare improving policy schemes will be different depending both on which regime is the target of subsidy (where the planner wants the game to switch to) and on the actual realization (which regime realizes in the absence of the planner), but policy schemes themselves are defined independently of regimes of the game.

Thus to find an optimally robust subsidy in terms of Def. 3 social planner has to define both the socially desirable outcome with the help of Prop. 2 and the regime of the game which would actually realize in a non-distorted case. Prop. 2 and Cor. 1 provide tools necessary to define the starting position of a subsidy: where the system would go in unperturbed case and where the planner wants it to go as well as the criteria for optimality and robustness of it:

Corollary 2.
There is a need to subsidize regime s only if $s = \Psi_\epsilon(\mathcal{F}) \neq \Theta_\epsilon(\mathcal{F}) = i$ at $\epsilon < \epsilon_s^W$. Policy scheme Σ_k^ϵ is optimal and robust in switching from i to s at the level ϵ_s^W in the sense of Def. 3 only if

$$\Psi_\epsilon(\mathcal{F}) = \Psi_\epsilon(\mathcal{F}_{\Sigma_k^\epsilon}) = \Theta_\epsilon(\mathcal{F}_{\Sigma_k^\epsilon}) = s \tag{32}$$

where $\mathcal{F}_{\Sigma_k^\epsilon}$ denotes the set of feasible outcomes under the policy scheme Σ_k^ϵ.

[12]this is necessary since introduction of the subsidy changes both $\delta(t), q_{A,B}(t)$ dynamics and thus values for the planner and firms differ from those in regime m without subsidies

In the present setup it is not the case that simultaneous development of both technologies is always socially desirable. We thus need a general criteria to select the appropriate policy scheme among suitable (welfare-improving and robust) ones.

Moreover, so far we defined robust social welfare, profit incentives and criteria of choice for policy schemes for a fixed level on uncertainty ϵ. It might happen that at some level ϵ^* one of the choice functions changes, i. e. predicts different outcome of the game as socially optimal and/or profit maximizing. Denote with $q_z^k(\epsilon)$ the $z-th$ root (zero) of the equation

$$\{q_z^k(\epsilon)\}: \min_{\varepsilon\in[-\epsilon;\epsilon]}\{W_s^{\Sigma_x}\} - \min_{\varepsilon\in[-\epsilon;\epsilon]}\{W_s^{\Sigma_k}\} = 0 \tag{33}$$

comparing the social welfare difference in regime s achieved after implementing policy schemes x and k respectively at robustness level ϵ.

We thus arrive to robust policy thresholds:

Proposition 3 (Selection and ordering of robust policy schemes)**.**

1. *If there exists* $\epsilon_k^R(i,s) < \epsilon_s^W$ *such that* $\Theta_{\epsilon>\epsilon_k^R(i,s)}(\mathcal{F}_{\Sigma_k^\epsilon}) \neq \Theta_{\epsilon<\epsilon_k^R(i,s)}(\mathcal{F}_{\Sigma_k^\epsilon}) = s$*, the policy scheme* Σ_k^ϵ *is robust in switching* i *to* s *up to level of uncertainty* $\epsilon_k^R(i,s)$*, otherwise it is globally robust for the pair* i,s*.*

2. *If there exists* $\epsilon_k^S(i,s) < \epsilon_s^W$ *such that* $\Psi_{\epsilon>\epsilon_k^R(i,s)}(\mathcal{F}_{\Sigma_k^\epsilon}) \neq \Psi_{\epsilon<\epsilon_k^R(i,s)}(\mathcal{F}_{\Sigma_k^\epsilon}) = s$*, the policy scheme* Σ_k^ϵ *is welfare improving for* s *up to level of uncertainty* $\epsilon_k^S(i,s)$*, otherwise it is globally improving for the pair* i,s*.*

3. *Policy scheme* Σ_k^ϵ *is admissible for the pair* $\{i,s\}$ *only for* $\epsilon \leq \epsilon_k^*(i,s) = \min\{\epsilon_k^R(i,s), \epsilon_k^S(i,s)\}$*.*

4. *At any given* $\epsilon \leq \epsilon_s^W$*, if the set* $\Sigma_\epsilon(i,s) \stackrel{def}{=} \{\Sigma_k^\epsilon\}, k \in K$ *of admissible policy schemes switching the game from* i *to* s *is non-empty, then there exists the choice function*

$$\exists\Lambda(\Sigma_\epsilon(i,s)):\ q_{-j}(0) \in \bigcup_{\Sigma_k^\epsilon\in\Sigma_\epsilon(i,s)} \bigcap_{\Sigma_k^\epsilon\in\Sigma_\epsilon(i,s)} \{[q_z^k(\epsilon); q_{z+1}^k(\epsilon)] | \min_{\varepsilon\in[-\epsilon;\epsilon]}\{W_s^{\Sigma_x}\} - \min_{\varepsilon\in[-\epsilon;\epsilon]}\{W_s^{\Sigma_k}\} > 0\} \Longrightarrow$$

$$\Lambda(\Sigma_\epsilon(i,s)) = \arg\max_{\Sigma_k^\epsilon\in\Sigma_\epsilon(i,s)} \min_{\varepsilon\in[-\epsilon;\epsilon]}\{W_s^{\Sigma_k^\epsilon}\} = \Sigma_x^\epsilon \tag{34}$$

selecting the best policy scheme Σ_x^ϵ *among those welfare-improving and robust ones at the level* ϵ*.*

5. *Denote* $\epsilon_1^*(i,s) = \min_{k\in K}\{\epsilon_k^*(i,s)\}$*, the uncertainty threshold of policy scheme* $\Sigma_x^\epsilon = \Lambda(\Sigma_{\epsilon<\epsilon_1^*(i,s)}(i,s))$*. There exists an increasing sequence* $\Omega^*(i,s) = \{\epsilon_1^*(i,s), .., \epsilon_k^*(i,s), .., \epsilon_K^*(i,s)\}$ *of uncertainty thresholds for all policy schemes in* $\Sigma_\epsilon(i,s)$ *such that the choice function* $\Lambda(\Sigma_\epsilon(i,s))$ *changes its value at each of them.*

6. *This forms a sequence of robust policy schemes increasing in uncertainty tolerance level* $\Sigma^*(i,s) = \{\Sigma_1^{\epsilon_1^*}, .., \Sigma_K^{\epsilon_K^*}\}$ *where each next element is more robust and welfare optimal under given robustness level.*

Proof. see Appendix B □

The (34) is another choice function, which selects the policy scheme among those social welfare improving and robust under ϵ. It selects the one which yields the highest welfare under regime s in worst case once policy scheme is applied.

The sequence of robust policy schemes is formed by increasing ϵ: once it crosses the threshold $\epsilon_k^*(i,s)$ the choice function changes and selects another scheme. It is important to note that these thresholds differ not only across schemes (which are independent of the regime) but also across the switches $\{i,s\}$: a given policy scheme may be more robust and/or welfare improving in one switching than in the other. These uncertainty thresholds may be found through application of choice function Λ to different values of ϵ. There is exactly the same number of thresholds as of welfare-improving robust policy schemes for switch $\{i,s\}$. Naturally for full certainty case the set of uncertainty thresholds is a singleton with $\epsilon_1^* = \infty$ and the policy scheme with the highest welfare is selected.

Corollary 3 (Optimal robustness level)**.**
Assume $\epsilon < \epsilon_s^W$*. The level of robustness* $\epsilon^{**}(i,s) = \epsilon_k^*(i,s)$ *is optimal for the switch* $\{i,s\}$*, if both*

$$\begin{aligned}
&\min_{\epsilon_k^*}\{W_s^{\Sigma_k^{\epsilon_k^*}}\} - \min_{\epsilon_{k+1}^*}\{W_s^{\Sigma_{k+1}^{\epsilon_{k+1}^*}}\} + ||\min_{\epsilon_{k+1}^*}\{W_s^{\Sigma_k^{\epsilon_k^*}}\} - \min_{\epsilon_{k+1}^*}\{W_{m_k}^{\Sigma_k^{\epsilon_k^*}}\}|| \leq 0,\\
&\min_{\epsilon_{k-1}^*}\{W_s^{\Sigma_{k-1}^{\epsilon_{k-1}^*}}\} - \min_{\epsilon_k^*}\{W_s^{\Sigma_k^{\epsilon_k^*}}\} + ||\min_{\epsilon_k^*}\{W_s^{\Sigma_{k-1}^{\epsilon_{k-1}^*}}\} - \min_{\epsilon_k^*}\{W_{m_{k-1}}^{\Sigma_{k-1}^{\epsilon_{k-1}^*}}\}|| \geq 0,\\
&m_k = \Theta_{\epsilon_{k+1}^*}(M_{\Sigma_k^{\epsilon_k^*}})
\end{aligned} \tag{35}$$

It is unique for any $\{i,s\}$*.*

Proof. Both lines in (35) are sums of welfare loss and gain from increasing the robustness level by one threshold. The first term is the difference in worst-case welfare in regime s under two successive schemes from $\Sigma^*(i,s)$ under associated robustness thresholds. This is always non-positive, since at every threshold the maximum welfare is selected. The second term is the potential error from applying the preceding (less robust) scheme under higher robustness level. There are two kinds of potential errors: either this scheme is no longer worst-case welfare improving, or it is not robust. Both cases are described by the difference under modulo operation and differ only in sign. Taking absolute value gives positive value of error avoidance. Once the sum of those two terms is not increasing, there is no further gain in increasing robustness level for the planner. Since the sequences of thresholds and policies are increasing, this choice is always possible and unique one. □

Corollary 3 partially endogenizes the robustness concept and relates back to the well-known Arrow-Fisher quasi-option value (see e. g. Mensink and Requate (2005) for discussion). One may interpret the changing level of ϵ as learning new information (uncertainty decreases) and (35) as a stopping rule taking into account the value of the new information gain and loss in efficiency. The advantage oft his last formulation is that again it does not relate on any specific structure of the underlying model except the existence of the value at hand and countability of the set of outcomes.

4 Application example

In this section I present a specific model based on the fully deterministic game of Bondarev and Krysiak (2017) to illustrate the main findings of the paper.

4.1 Specification of the model

Consider the 2-firms game $\{A, B\}$ whereas firm A has an initial advantage and firm B is trying to enter the market with the new technology. Both firms are fully informed on both initial states, $q_A(0), q_B(0)$ and efficiencies of investments. The social planner knows with certainty only the initial state of both technologies as well as the evolution of A, but is uncertain over the true evolution of B in time (e. g. because it does not know the efficiency of B's development). The production sector is modeled in the same fashion as in Krysiak (2011): a continuum of producers consider choosing technology A or B given locational costs x, current prices and states of both technologies.

The profit that each producing firm can obtain by using technology A or B is given by

$$\pi_A^{Prod} = z\,(q_A(t) - x) - p_A(t), \tag{36}$$

$$\pi_B^{Prod} = z\,(q_B(t) + x) - p_B(t). \tag{37}$$

We assume $x \in [-\bar{x}_A, \bar{x}_B] \subset \mathbb{R}_{>0}$. Thus, depending on the choice of $\bar{x}_A, \bar{x}_B$, locations could be on average better suited for technology A or for technology B.

To calculate the demand for each technology, we take into account that each firm buys one unit of equipment and each locations hosts a single firm. As long as $z\,q_j > p_j$ holds for $j = A, B$, all locations are used and thus the demand for technology j is determined by the distance between $\bar{x}_j$ and the location where a firm is indifferent between both technologies. This implies the following demand functions

$$Q_A^{Prod} = \bar{x}_A - \frac{1}{2}\left(\frac{p_A - p_B}{z} - q_A + q_B\right), \tag{38}$$

$$Q_B^{Prod} = \bar{x}_B - \frac{1}{2}\left(\frac{p_B - p_A}{z} - q_B + q_A\right). \tag{39}$$

We assume that $\bar{x}_j \leq q_j/2$ for $j = A, B$, which implies that in the equilibrium derived later, the conditions $z\,q_j > p_j$ for $j = A, B$ will always hold. Thus these demand functions characterize the case where both technologies are available.

If only technology A is available (which will be the case in some settings), demand for this technology is determined by the distance between $\bar{x}_A$ and the location where a firm receives a profit of zero when using technology A. In this case, demand for technology A is given by

$$Q_{A,-B}^{Prod} = \bar{x}_A - \left(\frac{p_A}{z} - q_A\right). \tag{40}$$

R&D firms sell their technologies and invest into their further development with objective functionals:

$$J_j = \max_{p_j, g_j} \int_0^\infty \mathrm{e}^{-rt} \left\{p_j(t)\,Q_j(t) - \frac{1}{2}\,g_j^2(t)\right\} dt. \tag{41}$$

and given evolution of each technology as:

$$\dot{q}_j(t) = \gamma_j g_j(t) \; - q_j(t). \tag{42}$$

where γ_j is the efficiency of investments into technology for firm j.

We then define $\delta(t)$ as before to be a technological gap across firms and governed by the equation:

$$\dot{\delta}(t) = \dot{q}_A(t) - \dot{q}_B(t) = \gamma_A g_A(t) \; - \gamma_B g_B(t) \; - \delta(t). \tag{43}$$

This structure forms a standard linear-quadratic differential game of R&D so that explicit solution may be easily obtained. However due to the assumption of zero entry costs, the incumbent firm A has a variety of options of strategic behavior by setting its price at a level preventing the entrance of the firm B.

The social planner's objective is the same as (6) with location effect defined as follows. In case that both technologies are available, these costs are given by[13]

$$\Xi_A(t) = \frac{\bar{x}_A^2}{2} - \frac{\left(p_A(t) - p_B(t) - z\,(q_A(t) - q_B(t))\right)^2}{8\,z^2}, \tag{44}$$

$$\Xi_B(t) = \frac{\bar{x}_B^2}{2} - \frac{\left(p_A(t) - p_B(t) - z\,(q_A(t) - q_B(t))\right)^2}{8\,z^2}. \tag{45}$$

In case only technology A is used, we get

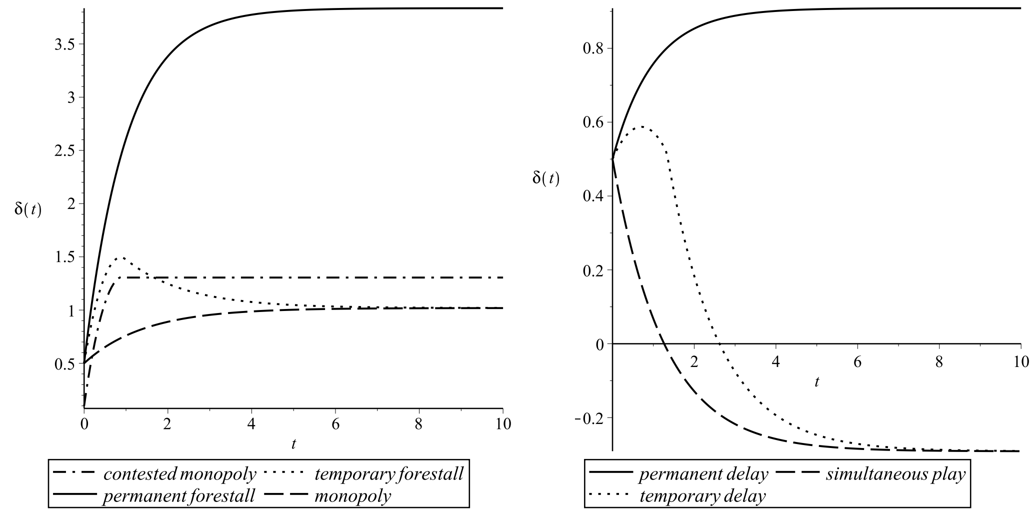
$$\Xi(t) = \frac{\bar{x}_A^2}{2} - \frac{1}{2\,z^2}\,\left(p_A(t) - z\,q_A(t)\right)^2. \tag{46}$$

4.2 Characterization of outcomes

In this setting it has been demonstrated that depending on parameters specifications there are the following possible outcomes:

1. Simultaneous development of both technologies, denoted by $*$,
2. Uncontested monopoly of the firm A whereas firm B never tries to enter the market, denoted M,
3. Permanent strategic pricing of firm A preventing the entrance of B, denoted S,

[13]This follows directly from Eqs. (38)–(39).



(a) Contested monopoly (b) Delay

Figure 1: Piecewise regimes of the model

4. Temporary strategic pricing of A leading to:

   (a) Eventual entrance of firm B (strategic delay), denoted d,

   (b) Eventual switch to a monopoly pricing by A, denote P,

   (c) Eventual switch to a monopoly with a constraint of keeping constant technology level due to ongoing threat of B's entry, denoted C for contested monopoly.

All these outcomes are results of solving the associated optimal control problems and based on the comparison of values achieved by different outcomes for the firm A, it chooses the msot preferred one according to the general Proposition 1. In particular the set $\mathcal{O}$ is finite and has 6 elements: $\mathcal{O} = \{*, M, S, P, C, d\}$ from which the one with the highest value is selected. Detailed solutions for all cases may be found in Bondarev and Krysiak (2017). Illustration of typical dynamics in different regimes is given by the Figure 1.

For illustration purposes I focus only on the case when the simultaneous development of both technologies is socially optimal. However even in this simple setup this is not necessarily the case: depending on market potentials of both technologies, $\bar{x}_A, \bar{x}_B$ it might be socially optimal to select a monopoly of firm A or even the contested monopoly case, such that only technology A is present on the market, but it is developed to a level higher then firm A would find individually optimal.

4.3 Full certainty policy schemes

We first derive a set of (second-best) policy schemes characterized by a subsidy size (paid to a firm B) and its duration, so that $\Sigma_k = \{\sigma_k, t^k\}$ switching the game from any given regime to the simultaneous play, restricting $\Sigma_k(i, s)$ to $\Sigma_k i, *$ only. Since all alternative outcomes in $\mathcal{O}$ are consider the argument $(i, *)$ is omitted in schemes characterization onwards. There are 5 different policy schemes in total which may be defined for this setup:

1. Maximal state-based scheme $\Sigma_{max} = \{\sigma_{max}, t^{max}\}$ defined as the subsidy level making the strategic pricing for A infeasible::

$$\sigma_{max} = z\delta(t) - 2z\bar{x}_B,\ t^{max}:\ \delta_{\sigma_{max}}(t^{max}) = 2\bar{x}_B \tag{47}$$

   where $\delta_{\sigma^{max}}$ denotes the solution to (43) under subsidy size σ_{max};

2. Maximal state-independent scheme $\Sigma_+ = \{\sigma_+, t^+\}$ defined as a constant counterpart of Σ_{max}:

$$\sigma_+ = z\delta_0 - 2z\bar{x}_B\ t^+ = \infty; \tag{48}$$

3. Minimal scheme $\Sigma_{min} = \{\sigma_{min}, t^{min}\}$, defined as a subsidy making the permanent strategic pricing of firm A unprofitable:

$$\sigma_{min}(t):\ \pi_A^S(t)|_{t<t^{min}} = \pi_A^*(t)|_{t<t^{min}}\ t^{min}:\ \Pi_A^S(\delta_{\sigma_{min}}) = \Pi_A^*(\delta_{\sigma_{min}}) \tag{49}$$

   i. e. size of the subsidy equalizes profit streams across strategic and simultaneous regimes for A and duration equalizes the accumulated value after reaching $\delta_{\sigma_{min}}$ gap level.

4. Associated constant scheme $\Sigma_- = \{\sigma_-, t^-\}$ is:

$$\sigma_- = z(\delta_0 - 4\bar{x}_B - 2\bar{x}_A)\ t^- = \infty, \tag{50}$$

5. Sufficient scheme $\Sigma_{suff} = \{\sigma_{suff}, t^{suff}\}$ is defined int the same profitability terms but taking into account the fact that firm A may switch from strategic pricing to monopoly after some time earning higher profits:

$$\sigma_{suff} : \ \pi_A^M(t)|_{t<t^{suff}} = \pi_A^*(t)|_{t<t^{suff}} \ t^{suff} : \ \Pi_A^M(\delta_{\sigma_{suff}}) = \Pi_A^*(\delta_{\sigma_{suff}}) \tag{51}$$

6. Associated constant counterpart denoted $\Sigma_o = \{\sigma_o, t^o\}$.

Precise expressions for these policy schemes may be found in the referred paper Bondarev and Krysiak (2017). It suffices to note here that every constant subsidy is higher then the state-based counterpart and

$$t^{min} < t^{suff} < t^{max} \ \sigma_{min} < \sigma_{suff} < \sigma_{max} \tag{52}$$

so that every next scheme in this sequence is costlier and lasts longer.

So far we have seen that employing a simple linear-quadratic game already allows for a multiplicity of outcomes and policy regulation tools, once we restrain from the first-best policy (which is given by the solution of the associated Stackleberg game with the government being the leader). This last is unique, but rather challenging to obtain and has limited use once we assume information asymmetry.

4.4 Robust subsidies

We now are ready to use the general results of the paper.

Under any level of ϵ the robust counterparts of full certainty subsidies are defined as:

$$\begin{aligned} \sigma_k^\epsilon &= f_k(\max_{\varepsilon\in[-\epsilon,\epsilon]} \delta(t,\varepsilon)); \\ t_\epsilon^k &= t^k(\max_{\varepsilon\in[-\epsilon,\epsilon]} \delta^{\sigma_k}) \end{aligned} \tag{53}$$

with $k \in K_* = \{min, suff, max, +, -, o\}$ defined in above. This is the case since subsidies are assigned only during the time when player A has an advantage ($\delta(t) > 0$), and since planner does not know with certainty true value of $\delta(t)$, the subsidy has to cover the maximal possible error in duration and size.

It follows that

$$\forall k \in K_* : \sigma_k^\epsilon \geq \sigma_k, \; t_\epsilon^k \geq t^k \tag{54}$$

and robust subsidies sizes and durations are increasing in the robustness level ϵ. Then as long as the full certainty subsidies are ordered as in (52), the same is true for robust counterparts for any ϵ.

We thus may establish the ordering of policy schemes switching the game into the simultaneous development regime, using results of the paper:

Corollary 4. *Assume the game has a structure* (41)-(43), (6)*. As long as* $\Psi_\epsilon(\mathcal{F}) = (*)$ *(i. e. simultaneous development is socially optimal) and* (52)[14]*, holds for* $\epsilon < \epsilon_*^W$*:*

1. *If* $\Theta_\epsilon(\mathcal{F}) = (d)$*, subsidies* $\{\Sigma_{min}^\epsilon, \Sigma_-^\epsilon, \Sigma_{suff}^\epsilon, \Sigma_{max}^\epsilon\}$ *are welfare improving and robust for* $\epsilon < \min\{\epsilon_{min}^R, \epsilon_{min}^S\}$*.*

   a *As long as* $\epsilon < \min\{\epsilon_{min}^R, \epsilon_{min}^S\}$*, the policy scheme* Σ_{min}^ϵ *is welfare optimal and robust;*

   b *As soon as* $\min\{\epsilon_{suff}^R, \epsilon_{suff}^S\} > \epsilon > \min\{\epsilon_{min}^R, \epsilon_{min}^S\}$ *the policy scheme* Σ_{suff}^ϵ *is welfare optimal and robust*

   c *As soon as* $\min\{\epsilon_{max}^R, \epsilon_{max}^S\} > \epsilon > \min\{\epsilon_{suff}^R, \epsilon_{suff}^S\}$ *the policy scheme* Σ_{max}^ϵ *is welfare optimal and robust*

   d *At last if* $\min\{\epsilon_*^W, \epsilon_d^O\} > \epsilon > \min\{\epsilon_{max}^R, \epsilon_{max}^S\}$ *only the policy scheme* Σ_-^ϵ *is welfare optimal and robust.*

2. *If* $\Theta_\epsilon(\mathcal{F}) = m \in \{S, P, C, M\}$*, subsidies* $\{\Sigma_{suff}^\epsilon, \Sigma_o^\epsilon, \Sigma_{max}^\epsilon\}$ *are welfare improving and robust for* $\epsilon < \min\{\epsilon_{suff}^R, \epsilon_{suff}^S\}$*.*

   a *As long as* $\epsilon < \min\{\epsilon_{suff}^R, \epsilon_{suff}^S\}$*, the policy scheme* Σ_{suff}^ϵ *is welfare optimal and robust;*

   b *As soon as* $\min\{\epsilon_{max}^R, \epsilon_{max}^S\} > \epsilon > \min\{\epsilon_{suff}^R, \epsilon_{suff}^S\}$ *the policy scheme* Σ_{max}^ϵ *is welfare optimal and robust*

[14] In the deterministic model there are cases when this ordering may fail, so here I impose this ordering as an assumption

c At last if $\min\{\epsilon_*^W, \epsilon_m^O\} > \epsilon > \min\{\epsilon_{max}^R, \epsilon_{max}^S\}$ *only the policy scheme* Σ_o^ϵ *is welfare optimal and robust.*

3. If $\Theta_\epsilon(\mathcal{F}) = ()$*, only the policy scheme* Σ_+^ϵ *is robust. It is welfare improving if* $\max\{\epsilon_{m\in\mathcal{F}}^O\} < \epsilon < \epsilon_*^W$ *and still* $\epsilon < \epsilon_+^S$*.*

Proof. See Appendix C □

This last example of application of general framework developed in the paper points out to the important distinction in what could be termed inessential, limited and unlimited uncertainty.

Denote by δ_*^{-j} the value of technology gap allowing for (unperturbed) simultanous development of both technologies, by δ_d^j the level leading to strategic delay and by δ_S^j the level leading to the (permanent) strategic pricing.

Fig. 2 illustrates the relation between increasing noise and ability of the government to distinguish different cases.

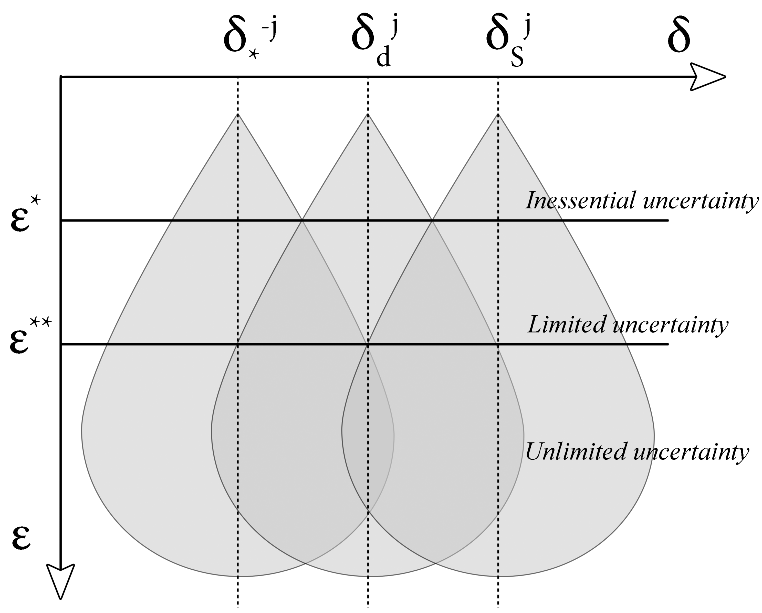


Figure 2: Structure of uncertainty in the model

The last Corollary tells us that there exists ϵ^* such that for uncertainty below this threshold the planner is still able to distinguish all these cases and all policy schemes retain there ordering and efficiency albeit with increased costs and duration (due to robustness requirement). Next there is at least one other threshold ϵ^{**} such that if there is limited but non-zero information on behalf of the state of the system, confidence intervals over values of roots of some of value functions equations overlap, dismissing some policy schemes as not robust. At last, if $e > \epsilon^{**}$ the planner cannot distinguish between different cases and only the most robust (and most expensive) type of regulation may still do the job.

The question of which level of confidence to choose to balance off potential costs of error and additional regulation costs is then answered by Corollary 3 which can be applied to the provided example int he same straightforward way as the Prop. 3.

5 Conclusions

This paper develops a novel approach to robust policies, which combine the usual min-max approach (e.g. Brock et al. (2014)) and model uncertainty (e. g. Gonzalez (2018)). The underlying dynamic game is assumed to allow for the existence of value functions for fully informed players while the government is uncertain over the rue state of the industry (asymmetric information).

Starting with an arbitrary dynamic game with finite number of states measuring the technologies' development, I formulate the ordering of individually and socially preferred outcomes based on the abstract notion of the choice function. The advantage of this abstract approach is that a choice function for a finite collection of sets exists always once axiom of choice is assumed. Moreover it allows for immediate computation of different policy thresholds for a wide class of value functions. In particular if underlying value functions are of the polynomial type power n, there always exist not more than n different thresholds separating different regimes for each player.

Based solely on this notion it is possible to find the ordering of robust policy schemes, derive the criterion for the selection of the most optimal one and formulate the concept of the optimal robustness level relative to the uncertainty.

First the algebraic characterization of full certainty outcomes of the differential game is obtained, then the same is done for the uncertain problem of the planner. This last takes into account the belief over the outcome of the game which is not necessarily coinciding with the true outcome (model mistrust). Next the robust policy schemes are defined as subsidies having size and duration (allowing for time-limited interventions) allowing to switch the state of the game to the desired one given the level of noise.

It comes with no surprise that costlier policies are required to cope with noisier signals for the planner to switch the state from some i to some j within feasible outcomes. What is interesting that even under rather mild assumptions we are able to prove the existence of a sequence of increasingly robust policies for different uncertainty thresholds. For any given level of uncertainty there exists a (suitably defined) optimal choice of the policy scheme for the planner in terms of welfare and it is a unique one.

At last once we think of this uncertainty as being subject to choice (i. e. more resources devoted to study of the market participants capabilities) it turns out that the optimal level of robustness exists in the sense that increasing robustness further will not increase the gain in expected welfare more than will be lost by using costlier (and more reliable) policy schemes. This result is generalization of the well-known Arrow-Fisher-Henry-Hanemann (AFHH) quasi-option value which relates the learning and optimal stopping in investments under uncertainty. Indeed, one may interpret the (increasing) uncertainty level in the last corollary as the result of (dynamic) learning and the optimal threshold as the stage where the 'investment' occurs.

The approach proposed here can be applied to any dynamic game with finite number of players and asymmetric uncertainty. In particular it may be applied to questions of subsidizing green technologies, which are frequently characterized by the uncertain potential and relative disadvantage in comparison with existing older technologies.

At last it has to be noted that the machinery developed in this paper can be easily adapted for any dynamic game not necessarily the differential one. Indeed it suffices for the game to have a finite well-defined value for this approach to be applicable. Such further generalization as well as refinement of information structures are considered as immediate extensions of the approach.

From a more application-oriented perspective, our paper casts some new light on policies that aim to support green technological change. These policies are widely used and are often criticized by economists, as they eliminate competition among technological options. Our results show that there are cases where it is indeed reasonable to temporarily reduce the effects of competition via technology-specific subsidies. Most interestingly, a less informed government should subsidize new technologies more and longer, as long as it can still ascertain that developing the technology is socially desirable. This is illustrated with the help of a particular example in the end of the paper.

References

Arrow, Kenneth J. (1962) The economic implications of learning by doing. *The Review of Economic Studies* 29 (3): 155–173.

Arthur, W. Brian. (1989) Competing technologies, increasing returns, and lock-in by historical events. *The Economic Journal* 99 (394): 116–131.

D'Aspremont, C. and Jacquemin, A. (1988). Cooperative and noncooperative r & d in duopoly with spillovers. *The American Economic Review*, 78(5): 1133–1137.

Athanassoglou, S. and Xepapadeas, A. (2012) Pollution control with uncertain stock dynamics: When, and how, to be precautious. *Journal of Environmental Economics and Management* 63 (3): 304–320.

Babonneau, F., Haurie, A., and Vielle, M. (2013). A robust meta-game for climate negotiations. *Computational Management Science*, 10(4): 299–329.

Ben-Tal, A., El Ghaoui, L. and Nemirovski, A. (2009) Robust Optimization. Princeton University Press, Princeton.

Ben-Youssef, S. and Zaccour, G. (2014). Absorptive Capacity, R&D Spillovers, Emissions Taxes and R&D Subsidies. *Strategic Behavior and the Environment*, 4(1): 41–58.

Bondarev, A. (2014). Endogenous specialization of heterogeneous innovative activities of firms under the technological spillovers. *Journal of Economic Dynamics and Control*, 38(0): 235 – 249.

Bondarev, A. and Greiner, A. (2017) Catching-up and falling behind: Effects of learning in an R&D differential game with spillovers. *Journal of Economic Dynamics and Control,*In press online version, doi: 10.1016/j.jedc.2017.10.001.

Bondarev, A. and Krysiak, F. (2017) Temporary and permanent technology lock-ins in the quality-differentiated Bertrand competition. *WWZ Working paper 2017/09.*

Brock, W., Xepapadeas, A., and Yannacopoulos, A. (2014). Robust Control and Hot Spots in Spatiotemporal Economic Systems. *Dynamic Games and Applications*, 4(3): 257–289.

Dawid, H., Greiner, A., and Zou, B. (2010). Optimal foreign investment dynamics in the presence of technological spillovers. *Journal of Economic Dynamics and Control*, 34(3): 296 – 313.

Gonzalez, F.(2018) Pollution Control with Time-Varying Model Mistrust of the Stock Dynamics. *Computational Economics*, 51(3): 541–569.

Herrlich, H. (2006). Axiom of Choice. *Lecture Notes in Mathematics No. 1876.* Springer-Verlag Berlin Heidelberg, Berlin.

Hinloopen, J., Smrkolj, G., and Wagener, F. (2013). From mind to market: A global, dynamic analysis of R&D. *Journal of Economic Dynamics and Control*, 37(12): 2729–2754.

Krysiak, F. (2011) Environmental regulation, technological diversity, and the dynamics of technological change. *Journal of Economic Dynamics and Control* 35 (4): 528–544.

Ludkovski, M. and Sircar, R. (2016) Technology ladders and R&D in dynamic Cournot markets. *Journal of Economic Dynamics and Control* 60(C): 127–151.

Mensink, P. and Requate, T. (2005) The Dixit-Pindyck and the Arrow-Fisher-Hanemann-Henry option values are not equivalent: a note on Fisher (2000) *Resource and Energy Economics*, 27(1): 83–88.

Robert, A. (2011) Nonstandard analysis. *Dover Books on Mathematics*. Dover Publications, Mineola, New York.

Todorov, E. (2009). Efficient computation of optimal actions. *Proceedings of the national academy of sciences*, 106(28): 11478–11483.

Unruh, G. (2000) Understanding carbon lock-in. *Energy Policy* 28 (12): 817–830.

Unruh, G. (2002) Escaping carbon lock-in. *Energy Policy* 30 (4): 317–325.

Vardas, G. and Xepapadeas, A. (2010). Model Uncertainty, Ambiguity and the Precautionary Principle: Implications for Biodiversity Management. *Environmental & Resource Economics*, 45(3): 379–404.

Vardas, G. and Xepapadeas, A. (2015). Uncertainty aversion, robust control and asset holdings. *Quantitative Finance*, 15(3): 477–491.

A Proof of Proposition 1

Proof. The proof simply follows from the definition of value functions Π_j^m and the fact that the choice function and the selector exist for any finite (and countable) collection of arbitrary sets satisfying the axiom of choice.

In particular, Π_j^m are functions of initial state of the game $\delta(0)$ only and thus their differences too. Each of the intervals $[\delta_z^j(m); \delta_{z+1}^j(m)]$ denotes the range of initial value $\delta(0)$ for which the difference in values of the game for the leader j across a given outcome i and some other outcome m has constant sign (positive or negative), since $\delta_z^j(m); \delta_{z+1}^j(m)$ denote the z-th and $z+1$th roots (zeros) of this difference for j as an algebraic expression. By Assumption 2 there are at most countably many such roots for each value function and hence at most countably many roots for differences of these values. Consider next only those such intervals that imply $\Pi_j^i(\delta(0)) - \Pi_j^m(\delta(0)) \geq 0$. If $\delta(0)$ belongs to one of such intervals it implies that outcome i is preferred to an outcome m.

We next compare i to all other m outcomes and consider those intervals where profit difference is positive. The intersection of all such intervals would give a range of $\delta(0)$ where outcome i is preferred to all other feasible outcomes. At last, we need to consider a union of all these intersections since there might be multiple of those.

The choice function $\Theta(\mathcal{F})$ selects the *outcome* i among all feasible ones if and only if the given initial state $\delta(0)$ belongs to the set of values where the value for j under outcome i is greater than under any other outcome, as denoted by the *argmax* term. At this point we need axiom of choice which guarantees that it is always possible to select an element within a set given some criterion (the selector function) whereas the selector is given by the above procedure on intervals between zeros of value functions.

Example: consider as an illustration the example with only 3 possible outcomes, a, b, c and assume for simplicity all values are given as polynomials of 3d degree in $\delta(0)$. In this case for an outcome a to be optimal it is needed that values differences $\Pi^a - \Pi^b, \Pi^a - \Pi^c$ remain positive at the given $\delta(0)$ level. So consider intervals between roots of those polynomials (4 intervals for the case of real roots) for each of alternatives. Assume $\delta(0) \in \{[\delta_1(b), \delta_2(b)|\Pi^a - \Pi^b \geq 0\}$ meaning that a is preferred to b once $\delta(0)$ is in this range. Then find the same interval for c alternative. If the intersection of those two ranges is non-empty and the union of such intersections contains $\delta(0)$, outcome a is optimal in comparison to the other two, that is, the selector function gives a as an *argmax* of value function w.r.t. the set of outcomes (and not $\delta(0)$ value itself). □

B Proof of Proposition 3

Proof. 1. The condition $\Theta_{\epsilon > \epsilon_k^R(i,s)}(\mathcal{F}_{\Sigma_k^\epsilon}) \neq \Theta_{\epsilon < \epsilon_k^R(i,s)}(\mathcal{F}_{\Sigma_k^\epsilon}) = s$ means that at the noise level $\epsilon_k^R(i,s)$ the outcome of the game believed to be realized is changing, so this is a threshold distinguishing the level under which subsidy from i to s is robust from not being robust. Thus by definition the given policy scheme is robust up to this level of noise. If such a change in believed outcome does not exist up to the maximal noise level ϵ_s^W associated with the outcome s (and chosen by the planner), the given scheme will always have an effect in switching the outcome from i to s.

2. The same claim follows for welfare functions: once under a given noise level $\epsilon_k^S(i, s)$ the preferred outcome changes (the choice function Ψ of the planner changes its value), this defines a threshold up to which the given policy is improving the (expected) welfare. Again, if such a change in choice function is not observed for any noise level up to the maximal one, given subsidy scheme is always welfare improving.

3. Combining two previous arguments we observe that the given policy scheme will be implementable (that is, both robust and welfare-improving) only for noise level being under both above defined thresholds (if they exist).

4. Now we form a set of all admissible policy schemes as defined by the previous point for a given noise level and two outcomes i, s. If this set is non-empty, then there is at least one policy scheme for a noise level up to ϵ_s^W which is both robust and welfare improving in switching the regime from i to s. Then there is a problem of selecting the best of such admissible schemes (since they are all only welfare-improving but not optimal). We then may define the welfare associated with each of such policy schemes and formulate a choice criteria in the same way as in previous propositions of the paper: the policy scheme, which yields the best worst-case welfare under given noise level ϵ is selected as the most appropriate one, see (34) which defines the selection criteria int he set of these admissible policy schemes.

5. Next we fix the noise level associated with the selected by Λ policy scheme as the minimal one. Once we allow this value to vary it is straightforward that the choice function (34) will change its value (i. e. will select a different scheme) at some level of noise. This level is the next threshold, where the scheme Σ_x^ϵ ceases either being robust or welfare improving and the set of admissible schemes shrinks. The fact that the set will shrink under increasing ϵ is implied by the *min* operator over ϵ in (34) formulation. Thus there is an increasing sequence of ϵ^* such that the previously admissible scheme stops to be such and the Λ selector will yield a different result.

6. By selecting at each such a threshold a new policy scheme the sequence Σ^* is formed. Since we have countably many schemes, the increasing sequence can be constructed out of it. It follows that each next element is better then previous ones under given noise level, since otherwise those previously selected elements will remain optimal.

□

C Proof of Corollary 4

Proof. 1. Once the belief of the social planner is such that only delay may realize, this outcome may be switched to the simultaneous one in a variety of ways. The set $\Sigma_\epsilon(d,*)$ defined as in Proposition 3 excludes schemes Σ_+, Σ_o since even in the extreme case of no information except the expected realization the scheme Σ_- is robust and by (52) it is less costly then the other two constant schemes. Hence $\Sigma_\epsilon(d,*) = \{\{\Sigma^\epsilon_{min}, \Sigma^\epsilon_-, \Sigma^\epsilon_{suff}, \Sigma^\epsilon_{max}\}\}$. Then we start with robustness level $\epsilon < \min\{\epsilon^R_{min}, \epsilon^S_{min}\}$, for which, by Proposition 3 all the set is admissible. Then the most welfare improving scheme is selected which is by (52) the scheme Σ^ϵ_{min}, giving case a. Once we increase the confidence interval to the level $\min\{\epsilon^R_{suff}, \epsilon^S_{suff}\}$, this scheme stops being robust and the set $\Sigma_\epsilon(d,*)$ looses one element. Thus the next scheme is selected, which is Σ^ϵ_{suff}. The process continues until the maximal level of noise is reached, under which only the constant subsidy remains robust.

2. Once the belief is different and permits multiple other outcomes except the delay, minimal subsidy cannot be robust since it does not prevent firm A from strategic pricing and the same is true for its constant counterpart. A the same time since (52) the scheme Σ_+ is clearly dominated by the Σ_o and thus is also excluded from the admissible set. We get $\Sigma_\epsilon(m,*) = \{\Sigma^\epsilon_{suff}, \Sigma^\epsilon_o, \Sigma^\epsilon_{max}\}$. We next again apply Prop. 3 increasing the uncertainty tolerance level ϵ and selecting at each of the threshold the next optimal scheme after the shrinkage of the admissible set.

3. At last, if there is not enough information for the planner to belief in any particular outcome of the game, only the constant maximal scheme is robust and can prevent strategic pricing. this would be the case if the robustness level is above the one

enabling to select the belief in any particular outcome, $\max\{\epsilon^O_{m\in\mathcal{F}}\}$ but still under the threshold where the planner believes in the social optimality of simultaneous development, ϵ^W_*. The policy scheme is welfare improving (and thus optimal since the set of admissible subsidies is a singleton) if additionally it is robust, i. e. the noise level does not exceed its robustness level.

□