Essays on dynamic non-cooperative games based on simulations and experiments

David Lara Arango

Thesis for the Degree of Philosophiae Doctor (PhD)
University of Bergen, Norway
2018

UNIVERSITY OF BERGEN

Essays on dynamic non-cooperative games based on simulations and experiments

David Lara Arango

Thesis for the Degree of Philosophiae Doctor (PhD)
at the University of Bergen

2018

Date of defence: 27.04.2018

Year: 2018

Title: Essays on dynamic non-cooperative games based on simulations and experiments

Name: David Lara Arango

Print: Skipnes Kommunikasjon / University of Bergen

Essays on dynamic non-cooperative games based on simulations and experiments

by

David Lara-Arango

Dissertation for the degree philosophiae doctor (PhD)

System Dynamics Group, Social Science Faculty

Supervised by Prof. Erling Moxnes

University of Bergen

December 2017

Acknowledgements

This thesis has been possible thanks to the System Dynamics group at the University of Bergen, Norway, the Decision Science department at the Universidad Nacional de Colombia, in Medellín, and the Institute of Management at the Università della Svizzera Italiana in Lugano, Switzerland. I express my heartfelt gratitude to these institutions.

I would like to give special thanks to my supervisor, Prof. Erling Moxnes. His interest in both my work and my development as a researcher has helped me to improve my skills beyond what I can express in this space. Thank you for your lessons, challenges, comments and never ending red ink. Your clarity and sharpness are two things I aspire to have one day.

I have nothing but thankful words to my co-supervisor, Prof. Santiago Arango-Aramburo. Thank you so much for supporting me throughout these years in my academic career. Thank you for your advice and comments, but most of all, thank you for your friendship and for being a role-model to me. I also want to express my gratitude to Prof. Erik Larsen for being interested in my work, providing very valuable insights and inviting me to the Università della Svizzera Italiana. I have learned a great deal of things about being a good researcher from you, thank you very much.

I want to give special thanks to Prof. Birgit Kopainsky. Thank you for your advice, lessons, coaching sessions, and most of all, your friendship. I want to express my gratitude to Prof. Pål Davidsen and Prof. David Wheat. Thank you for your willingness and openness to discuss different topics and give me advice. I also want to thank my dear friends from the "original office crew" and from "the league"; thank you Omar, Andreas, Erika, Aklilu, Eduard, Stian and Santi. My Ph.D and my life in Bergen would not have been the same without you all. I want to give a special mention to Omar; those discussion sessions proved to be most valuable not only for my PhD but also my life in general!

I want to thank "mi tutu" for her enduring love and support during the last years of my Ph.D. Thank you for bringing so much joy into my life, my love. Finally, I want to thank the person to whom I owe all I am and ever will be, my mom. ¡Te adoro mami!

Abstract

This thesis studies ways to improve Non-Cooperative Game theory (NCGT) as a policy making methodology. NCGT allows to categorize a wide range of situations and to provide solutions. However, the effectiveness of such solutions depends on whether cooperative behavior can arise. The common presence of cooperative behavior in real life systems often threaten NCGT solutions reliability. Furthermore, most studies in NCGT literature fail to account for important dynamics of real life problems. Understanding such dynamics is of critical importance for policy analysis. We analyze three well-known non-cooperative games: the Cournot oligopoly, the public goods game, and the dictator game. To link this thesis with the existing literature, we analyze these games in contexts they have been applied before such as commodity markets and climate change conferences (COPs). We use simulations and experiments as means to test the solutions provided by NCGT in each specific case.

We start by using the Cournot oligopoly to study the case of deregulated electricity markets. In a first study, we use simulations to test the effectiveness of two capacity mechanisms (i.e. mechanisms intended to stabilize capacity investments in the market) under different uncertainty levels in the market. Contrary to theoretical predictions, capacity mechanisms present substantial differences in market stability, market welfare and sensitivity to uncertainty. We found the most market oriented mechanism to be the best option overall.

In a second study, we conduct an experiment to tests the findings of the first study. We found unexpected market reactions to one of the mechanisms leading to much worse results than what the literature suggests. We found the most market oriented mechanism to be the best one once more.

In a third study, we use an experiment to test Meadows' dynamic hypothesis for the hog market cycle (Meadows ,1970). Contrary to economic theory and previous laboratory experiments, we found strong evidence of lasting price cycles.

In a fourth study, we carry out a public goods laboratory experiment to compare two procedures used in climate negotiations (COPs), one of which has been deemed as ineffective by the NCGT literature. We found no significant difference between the two procedures in terms of promoting cooperation. We also found significantly higher contributions than theory predicted.

In the fifth article, we study whether the Dark triad framework can be a good predictor of people's decisions in the dictator game, and whether those decisions are consistent with theoretical predictions. Using a laboratory experiment, we do not find evidence to consider the Dark triad a good predictor in this case.

This thesis contributes to a better understanding of the existing limitations in NCGT as methodology for policy analysis. This should call for further efforts to understand these limitations in particular contexts and propose solutions to them. Also, the combined methodology proposed in this thesis should serve as a motivation to improve NCGT theoretical predictions in light of dynamic complexity and realistic decision making.

Table of contents

Introduction

1. Overview

Non-cooperative game theory (NCGT) provides a platform to study different situations with no enforceable cooperation (Nash 1950, 1951; Smith, 1982; Owen, 1995; Gibbons, 1992; Rasmusen, 2007; Bolton and Dewatripoint, 2005). Such situations often lead to problematic interactions between players with various decisions affecting and influencing each other (Aumann and Schelling, 2005). Examples of this can be found in several social disciplines such as management, political science, international relations, social psychology, law, among others (e.g. Nutter, 1964; Rapoport and Chammah, 1965; Fudenberg and Maskin, 1986; Rosenthal, 1981; McKelvey and Palfrey, 1992; Ledyard, 1997; Aumann, 2003). In spite of their applicability to a wide variety of problems, NCGT present an intriguing predicament when used to propose real life problem solutions in non-cooperative situations. If a real-life problem solution is not consistent with NCGT predictions, such a solution is likely to fail. If a solution is consistent with NGCT predictions, such a solution is also likely to fail, given that its success is contingent on the absence of cooperative behavior[1]. Therefore, both theoretically inconsistent and consistent solutions are likely to fail in real life. This issue challenges NGCT models' capacity to design public policies, as McCain (2009, p4) states:

"*Non-cooperative behavior is common enough so that a social arrangement that is unstable in the face of non-cooperative behavior will probably fail. However, solutions based on non-cooperative game theory may be unstable in the face of cooperative or collusive behavior, and cooperative behavior is common enough that such solutions will themselves often fail. Thus, non-cooperative game theory is far less effective as a prescriptive tool for public policy*".

[1] Consider a market as an example. If there is no cooperation, NCGT anticipates could anticipate a Nash equilibrium solution. If there is cooperation, monopoly-like behavior is likely to arise, which in turn steers the market away from the Nash equilibrium i.e. the NCGT proposed solution.

One possible solution for this issue, is to improve the non-cooperative models by including the most important elements of cooperative models, since it seems that separating cooperation and non-cooperation is not possible in real life (Maskin, 2004). A formal definition of the non-cooperative and cooperative branches of Game theory is offered by Chatain (2014, p1): “Cooperative game theory focuses on how much players can appropriate given the value each coalition of players can create, while non-cooperative game theory focuses on which moves players should rationally make”. In principle, the fundamental definitions of the two do not suggest any intrinsic challenge to create models based on both branches of Game theory. However, there are several limitations and challenges to do so in practice. First, it is not clear how competitive behavior can explain the mechanisms that lead to coalition formation, organization and competition between coalitions (Aumann and Dreze, 1974; Carraro, 2003; McCain, 2009). Moreover, much of the literature on both non-cooperative and cooperative models is based on highly simplified assumptions, which on the one hand allow researchers to gain insights about different solutions, but on the other hand, limits these games applicability to real life problems, as well as the possibilities to integrate both branches (Osborne and Rubinstein, 1994; Rasmusen, 2007; Schelling, 2010).

Another possible solution is to improve the decision-making theories that support non-cooperative games. Several studies have been conducted to accomplish this (e.g. Rapoport and Chammah ,1965; Poundstone, 1992; Cooper et al.,1990; Van Huyck et. al., 1990;). Improving decision-making theories is especially important when one acknowledges that many real-life problems are embedded in complex dynamic environments. Complex dynamic environments can compromise the accuracy of theoretical predictions, since such environments are known to facilitate unexpected (and undesired) outcomes, such as instabilities in markets (Arango and Moxnes, 2012), disturbances in supply chains (Sterman, 1989), and suboptimal resource management policies (Moxnes, 2004). Therefore, having a clear understanding of the dynamics players face in real life and how such dynamics influence players’ decisions is essential to design effective policies (McCain, 2009). In addition, an improved knowledge of these issues is of course interesting to producers, investors and financial agents.

Despite the importance of dynamic components in real non-cooperative situations, most modern textbooks in game theory scarcely address how such components can impact expected outcomes (e.g. Osborne and Rubinstein, 1994; McCain, 2009; Dinar et al., 2008; Ott, 2006; Carmichael, 2004; Aumann and Hart, 2002). Journal articles concerning these types of games explore feedback driven dynamic behavior such as strategy evolution (Nax and Pradelski, 2015) and directional learning adjustments (Nax, 2015). Besides few exceptions (e.g. Van Long, 2010; Ding et al., 2014a; Ding et al., 2014b), non-cooperative studies generally do not take important aspects of complexity into account, such as accumulation and time delays, which are important in many real-life systems (Arango and Moxnes, 2012; Pierson and Sterman, 2013; Moxnes, 2012). Moreover, in absence of dynamic elements, many non-cooperative models may prescribe policies that can lead to unexpected outcomes in reality, given that these elements may alter the assumptions such games are based on (Aumann, 1973; Stuart, 2001). This thesis aims to contribute in correcting this issue. By using simulations and experiments, this thesis studies different situations that are conceived as one of three well-known games; namely, the Cornout oligopoly, the public goods game, and the dictator game. By doing so, this thesis aims to first, point out ways to improve theory in these three games, and second, propose policies that are not only supported by the principles of Non-cooperative game theory but also by simulations and experiments in which human decision-making biases and realistic dynamics are accounted for.

2. Background

Each of the studies in this thesis is based on one of three well-known games in the scientific literature, namely Cournot Oligopoly, public goods, and dictator game. This section makes a brief literature review of these games and discusses difficulties of using such games as a basis for policy formulation.

2.1. Cournot oligopoly

Ever since Cournot (1838) proposed his first oligopoly model in his work “Recherches sur les Principes Mathématiques de la Théorie des Richesses”, Cournot’s oligopoly model has been widely used to study market competition (See e.g. Von Mouche and Quartieri, 2016 for a comprehensive review). With almost two centuries of history, one could expect that its relevance today would be diminished. However, the literature concerning market competition seems to suggest otherwise (Von Mouche and Quartieri, 2016). This, of course, does not mean that problems in the original formulation by Cournot have not been found, in fact, the model's solution has been challenged on more than one occasion (e.g. Palander, 1939; Theocharis, 1960; Puu, 2006). Nevertheless, the model's main contributions about market convergence towards an equilibrium still remain (Cánovas et al., 2008).

Simulation studies based on the Cournot oligopoly have been widely used to study different issues. Recent journal articles using simulations, use Cournot's formulation to study contingent workforce effects (Matsumoto et al., 2015), strategic minerals markets (Hecking and Panke, 2015), isoelastic demand markets (Fanti et al, 2013), multi-product oligopolies (Wu and Ma, 2014), among many other issues. Experimental studies on oligopoly competition have been the subject of many publications for many decades (Smith, 1962; Fouraker et al., 1961; Hoggatt, 1959). Modern articles about these studies focus on a wide variety of issues, such as quantity and price competition, exogenous timing, price dispersion, tacit collusions among others (See Potters and Suetens, 2013 for a review of modern oligopoly experimental studies).

Regarding its main predictions, the Cournot Oligopoly suggests that players will converge towards the Nash equilibrium, even with large number of players (Cournot, 1838). While the convergence towards the Nash equilibrium is generally accepted as a benchmark of rational behavior, the stability of such equilibrium has been shown to be weak, even if one assumes firms to be identical (Agiza, 1998; Ahmed and Agiza, 1998; Puu, 2006). However, a number of experimental studies have shown that such equilibrium can be reached as players gain more experience in the game (e.g. Milgrom and Roberts, 1991). In fact, the typical strategy chosen by players can either lead to the game's convergence (see e.g. the market-inertia based reply process in Kandori et al., 1993) or divergence (see e.g the best reply process in Theocharis, 1960). This overall strategy has been shown to be highly information-sensitive (Huck et al., 1999). Therefore, one must be careful when using the Cournot Oligopoly to study a specific market. Failing to give an important piece of information or giving the wrong information will directly affect the competition level which in turn will affect the game's applicability to a given market (Huck et al., 2004). This thesis uses simulation and experiments based on the Cournot Oligopoly to assess whether players' decisions can generate specific market behaviors based on a given set of conditions and a given set of available information that resemble what is found in specific real markets.

2.2. Public goods game

Common resources, public goods and public bad have, for a very long time been regarded as a great challenge for scientists and policy makers alike. Problems associated with these goods seem to be especially difficult to solve, even when the cause of the problem is well known to the actors involved (Ostrom, 1990). The management and adequate procurement of common resources have been study extensively, ever since Hardin (1968) published his article “The tragedy of the commons”. Along with Hardin’s influential article, many authors consider Samuelson’s paper “The pure theory of public expenditure” (Samuelson, 1954) to be the foundation of modern theory of public goods (Pickhard, 2006). In this context, the public good game offers the possibility to study to what extent agents contribute to a public good given a set of conditions. The game’s main theoretical prediction is the convergence towards free riding i.e. the Nash equilibrium (Andreoni, 1995).

Modern articles using simulation on the public goods game are mostly focused on punishment/reward mechanisms to sustain cooperation. Some of the issues in this regard include threshold-driven cooperation (Mikkelsen and Bach, 2016), the effect of adaptive reputation (Chen et al, 2016), and collective punishment (Gao et al, 2015). Similarly, modern experimental studies deal with institutional frameworks, participation mechanisms and group structure effects on sustained cooperation. Topics currently discussed include: institutional deterrence (Kingsley and Brown, 2016), voluntary participations in public goods (Hong and Lim, 2016), institutional reciprocity (Ozono et al, 2016), group size inefficacies (Diederich et al, 2016) among others.

Predictions based on the public goods can present a number of weaknesses. Both theoretical and experimental studies suggest that players may voluntarily contribute to public goods to a greater extent than the Nash equilibrium predicts (Bergstrom et al. 1986; Bernheim, 1986; Ledyard, 1995). While such contributions do not typically reach a socially efficient level, they do pose important questions about what the triggering factors for such contributions are. Previous research has shown that player contributions decrease as the uncertainty about the public good payoff increases (Burger and Kolstad, 2008). This finding poses great challenges for real life situations that are commonly characterized as public good games such as the climate change agreements. In this sense, having certainty about the consequences/benefits of contributing to climate emission abatement is a crucial point to ensure a high level of collaboration among nations. This becomes even more challenging when one considers that countries have asymmetric consequences/benefits derived from climate emissions abatement. The fourth paper of this thesis uses simulations and experiments to study how climate change agreements could be improved, in light of asymmetry and certainty about different nations' payoff structures.

2.3. The dictator game

The distribution of income has been one of the most popular concerns in the experimental games literature (Engel, 2010). In fact, this issue has been of interest for many researchers, especially since Daniel Kahneman proposed the Ultimatum game in 1986 (Kahneman et al. 1986). One of the main questions since then is, what makes people behave differently than what theory predicts when there are no apparent rational reasons to do so? The dictator game was first developed to answer this specific question, by taking out any "fear of punishment" effect from the ultimatum game. The dictator game has been a popular game among experimentalist, thus providing a vast body of findings to different research questions. For this reason, the dictator game has been highly praised in the literature, both for its usefulness and its simplicity (Forsythe et al.,1994). In fact, this game has allowed researchers to challenge the traditional profit-maximizing decision making that has been traditionally believed to be the norm in economics (Kahneman et al., 1986).

Modern simulation studies that focus on the dictator game are primarily concerned with two issues. The first issue is how well can individual beliefs and neural-cognitive models explain dictators' behavior (e.g. Beullens et al., 2012). The second issue is the role of institutional punishment in societies, see for instance Gyorgy, (2008) who uses prescriptive agents (agents played by the computer) along with human subjects to explore this issue. Recent articles shows an interest in the effects of social values and psychological traits on the dictators' behavior. See for instance studies of generosity as a result of self-worth (Przepiorka and Liebe, 2016), social value orientation (Wei et al., 2016), reputation and cooperation (Wu et al., 2016) and social contingency (Rutledge et al., 2016).

The standard prediction of the dictator game is the convergence to the Nash equilibrium, that is, the dictator will give zero percent of his endowment to his counterpart. However, many studies that use the dictator game have shown that subjects give significantly more to their counterparts than theory predicts, in fact, subjects give around 28% on average according to Engel (2010). These findings have created additional questions such as: do people give away part of their endowment out of altruism? if not, is it because of strategic reasons derived from a fear of being in the other's position later on? There is no definite answer to these questions yet, but different studies point out that it is not only altruism that explains players giving away part of their endowment (Bolton and Ockenfels 1998; Fehr and Schmidt 1999; Charness and Rabin 2002; Bradsley, 2008). Furthermore, players' generosity has been linked to several context-dependent factors such as age of the dictator, perceived deservedness of the recipient, anonymity in the game, among others (See Engel (2010) for a meta study of such factors).

3. Papers

This section presents the central *problems* and explains the *hypotheses* to be tested in each of the papers presented in this thesis. All the null hypotheses are based on rational expectations (Muth, 1961). Thus, these hypotheses state convergence to the Nash equilibrium. Random and statistically insignificant deviations from the Nash equilibrium do not constitute a reason to reject this hypothesis; only systematic and significant deviations do. Most of the alternative hypotheses are based on Bounded rationality (Simon, 1979), which implies the use of heuristics (Tversky and Kahneman, 1987). Heuristics arise when task complexity outweighs the subjects' cognitive capabilities (Kleinmuntz, 1993). While such heuristics can lead to satisfying results in simple problems, their effectiveness tend to diminish as the complexity of the problem increases (Sterman 1989; Diehl and Sterman, 1995; Moxnes, 2004; Arango and Moxnes, 2012).

3.1 Paper 1: Uncertainty and the long-term adequacy of supply: Simulations of capacity mechanisms in electricity markets

By David Lara-Arango, Santiago Arango-Aramburo, and Erik R. Larsen

This paper studies two effects of capacity policies on a the welfare generation of deregulated electricity markets and on price stability. The method is simulations with different levels of uncertainty. The two policy mechanisms are Procurement for long-term strategic reserves contracting and Centralized auctioning for capacity contracts (Finon and Pignon, 2008). If rational expectations were assumed, neither of the two mechanisms would have a significant effect, since the market would be capable of converging to the Nash equilibrium by itself. However, it is clear that real deregulated electricity markets do not show such convergence (Olsina et al., 2006; Arango and Larsen, 2011; Olaya et al., 2015). Hence, this paper assumes bounded rationality for investment decisions. Specifically, the model assumes adaptive expectations (Nerlove, 1958) that can lead to price cycles (Meadows, 1970), as has been observed in real electricity markets (Arango and Larsen, 2011). In addition, the paper hypothesizes that uncertainty plays a significant role when determining the effectiveness of a given capacity mechanism, i.e. uncertainty can make a given mechanism fail even if such a mechanism is the theoretically most sound.

3.2 Paper 2: Towards a long-term economic welfare in deregulated electricity markets: Testing capacity mechanisms in an experimental setting.

By David Lara-Arango, Santiago Arango-Aramburo, and Erik R. Larsen

This paper is closely related to the previous one as it uses experiments to study the same two policy mechanisms; Procurement for long-term strategic reserves contracting and Centralized auctioning for capacity contracts (Finon and Pignon, 2008). The null hypothesis is based on rational expectations and thus, the experiments are expected to converge towards the Cournot-Nash equilibrium. The alternative hypothesis is based on simulations using adaptive expectations heuristics (Nerlove, 1958). Hypothesized heuristics are built using data from previous experimental studies (Arango and Moxnes, 2012; Lara-Arango, 2014; Alcaraz, 2010). In correspondence with such simulations, Procurement for long-term strategic reserves contracting is hypothesized to present the best results for the market both in terms of economic expected value and economic stability.

3.3 Paper 3: Testing meadows' hog cycle theory by laboratory experiment

By David Lara-Arango and Erling Moxnes

As the title indicates, this paper tests Meadows (1970) theory about cycles in the hog market in the US through a Cournot experiment. The hog or pork cycle is a well-known example of price fluctuation in commodity markets. Meadows' model presents features that differ from other commodities, for instance, the possibility to hold inventory and the fact that capacity expansions (livestock) lead to an immediate reduction in production (slaughtering). The null hypothesis states that players behave with Perfect rationality and thus, the market price will be stable and converge towards the Cournot-Nash equilibrium. The alternative hypothesis is based on bounded rationality and states that players will behave according to Meadows' heuristics formulation. Meadows' heuristics is based on adaptive expectations (Nerlove, 1958).

3.4 Paper 4: Making climate conferences more effective?

By Erling Moxnes and David Lara-Arango.

The academic community showed marked skepticism towards pledges in the COP 21 (Inman, 2009, Cooper, 2010; Cramton et al., 2015; Gollier and Tirole, 2015) while pledges received considerable support from the public at large (Solutions, 2015). This study aims to contribute to the COPs literature by directly comparing a pledges-based procedure such as the one used in the COP 21 with a negotiation procedure such as the one use in the Kyoto protocol. The method is laboratory experiments. If players behave with perfect rationality, both procedures will converge to the Nash equilibrium and there will be no difference between them. If players behave with Bounded rationality, neither the social optimum nor the Nash equilibrium will be achieved and differences between procedures can be expected. Specifically, the negotiations procedure (resembling the Kyoto protocol) is expected to yield higher contributions given the absence of punishment mechanisms in the pledges-based procedure (Stiglitz, 2015).

3.5 Paper 5: Socially aversive personalities and income distribution: Can the dark triad predict behavior in the dictator and gangster games?

By David Lara-Arango.

This paper studies how well the Dark triad of human personality (Paulhus and Williams, 2002) can predict income distribution decisions in controlled environments. The Dark triad of human personality has been a widely-researched topic in behavioral psychology (Furnham et al, 2013). It comprises three elements; Narcissism, Machiavellianism and Psychopathic traits. The combination of these three elements (at subclinical levels) is believed to be a powerful predictor of aversive behavior towards others (Jones and Paulhus, 2011b). In fact, previous studies suggest that people with the highest scores tend to be more aggressive when seeking their goals and are more likely to disregard others' well-being in the process (Jonason and Krause, 2013). This tendency is also consistent with higher selfishness and entitlement (Campbell et al, 2004). Hence, a positive relationship between the scores in the questionnaire and selfish behavior on both frames of the dictator game could be expected. The null hypothesis is that the Dark triad scores and the dictator games outcomes will be completely unrelated. Conversely, the alternative hypothesis for in this paper states that players with the highest scores in the test will also be the ones who give less or take more in their respective games.

4. Methodology and main findings

This section introduces the *methodology* that was employed in each of the papers, and presents their main *findings*. The methodology employed throughout the thesis consists of four phases

- Game theory phase: A theoretical model is proposed to address a particular problem.
- Hypotheses phase: Hypotheses are proposed and formalized.
- Experimental phase: Hypotheses are tested by economic experiments or computer simulations.

- Comparative phase: Results from the Experimental phase are contrasted with the Hypothesized outcomes.

4.1 Paper 1: Uncertainty and the Long-Term Adequacy of Supply: Simulations of Capacity Mechanisms in Electricity Markets

By David Lara-Arango, Santiago Arango-Aramburo and Erik R. Larsen

This paper proposes a series of stylized electricity market models to explore the effectiveness of two capacity mechanisms in terms of market welfare and security of supply. The two mechanisms are Procurement for long-term strategic reserves contracting and Centralized auctioning for capacity licenses (Fignon and Pignon, 2008). The first mechanism allows the regulator to influence market capacity either by making contracts with generators or investing trough a state-owned firm (Meunier and Finon, 2006). This regulator will make capacity investments when there is a perceived need for new capacity. The second mechanism allows the regulator to have control over the total market capacity in the form of capacity licenses, which it is ultimately auctioned to the generators (Finon and Pignon, 2008). The generators bid for licenses to build capacity, that is, they compete to expand their capacity at the best possible license price (Vasquez et al., 2003).

The proposed economic model is based on Arango and Moxnes (2012). By assuming generators to be price-takers, these models represent a closed-loop formulation, in which players decide on a new capacity based on their price expectations. In turn, the price results from the generators' decisions (Wogrin et al, 2013). The paper proposes three economic models, a base case using the Arango and Moxnes (2012) formulation, a second case in which the first mechanism is implemented and a third case in which the second mechanism is implemented. Thereafter, four different simulation scenarios are proposed: no stochasticity, low stochasticity, medium stochasticity, and high stochasticity.

Simulation results show that, in absence of uncertainty, Centralized auctions for capacity licenses lead to a higher market stability. These results are consistent with previous works about this mechanism (de Vries and Hakvoort, 2004). On the other hand, Procurement for long-term strategic reserves leads to a higher welfare. However, this mechanism seems to be less sustainable than the previous one, since it may lead to sustained reduced margins for generators. As uncertainty is introduced into the model, both the performance of both mechanisms decrease, to the point that it is no longer clear that neither mechanism is recommendable. Centralized auctions for capacity licenses presents the poorer result in the high uncertainty treatment after being arguably the best option under no uncertainty.

This paper points out the impact of uncertainty when assessing different capacity mechanisms. Failing to recognize its importance, may lead to wrong conclusions about the adequacy of policy mechanisms. In fact, theoretical assessments made in absence of uncertainty may not hold when uncertainty in the market increases. Thus, policies should carefully consider the benefits of a given intervention in light of its robustness to different levels of market uncertainty.

High levels of uncertainty seem to favor generator interests by inducing semi-permanent shortages with high prices. This semi-permanent shortage works in a somewhat similar way as when generators are allowed to mothball capacity (Arango et al., 2013), which could give them excessive market power. This points out the importance of considering welfare generation when assessing capacity mechanisms.

4.2 Paper 2: Towards a long-term economic welfare in deregulated electricity markets: Testing capacity mechanisms in an experimental setting.

By Santiago Arango-Aramburo, David Lara-Arango, and Erik R. Larsen

This paper proposes an experimental design to test the same two capacity mechanisms studied in the previous paper, namely Procurement for long-term strategic reserves and Centralized auctioning for capacity licenses. Experiments have been used to study various problems in electricity markets such as energy efficiency (Ramos et al., 2015), green technologies (Sundt and Rehdanz, 2015), regulatory designs (Rassenti et al., 2003) and security of supply (Brandts et al., 2008; Arango et al., 2013; Islyaev and Date, 2015). The present paper aims to contribute to the literature on security of supply, by using economic experiments to assess the potential of the aforementioned capacity mechanisms in terms of both market stability and welfare.

Three treatments are considered. The first treatment is Arango and Moxnes' (2012) last treatment, and represents a deregulated electricity market without interventions. The second treatment introduces a regulatory firm that invests in the market when a capacity shortage is perceived. The third treatment introduces a centralized auctioning system through which players can bid for licenses to build new capacity. The data for the first treatment was taken from Arango and Moxnes (2012). This treatment was not conducted again because its main purpose is to serve as a benchmark for the other two, and the same subject pool was used. In addition, the same format was followed in the subsequent treatments, which makes the three treatments directly comparable. All treatments were carried out under the standard protocol for economic experiments (Friedman and Cassar, 2004).

Experimental results suggest that Procurement for long-term strategic reserves does not represent an improvement for the market in neither welfare nor stability terms. Moreover, players seem to bear substantial and sustained losses when this mechanism is implemented, which compromises its sustainability. In addition to this, a higher price volatility was found in treatment 2 than in treatment 1. Therefore, this mechanism may lead to a worse outcome than if the market is left alone with no intervention. In fact, our analyses suggest that this mechanism may lead to excessive competition, which as theory suggest, can compromise the market's ability to reach an equilibrium (Huck et al., 1997). In contrast, Centralized auctioning for capacity licenses presents an improvement in experimental market performances as market welfare is modestly improved while market stability is substantially improved. Furthermore, players' profits are close to the normal profit, which implies a sustainable market setting for generators.

This paper shows that some capacity mechanisms can be detrimental for both security of supply and welfare in a deregulated electricity market. Unexpected reactions from the market actors to an interventionist mechanism can lead to unforeseen and undesirable results. This paper's findings are consistent with previous studies that argue in favor of market oriented mechanisms (Meunier and Finon, 2006). Although market oriented mechanisms may not be able to counteract high uncertainty, as Paper 1 suggests, they can still be plausible to implement, given the intrinsic risk aversion of both consumers and producers. Moreover, the cost of blackouts and shortages are typically considered more severe than the cost of the mechanism.

4.3 Paper 3: Testing Meadows' hog cycle theory by laboratory experiments

By David Lara-Arango and Erling Moxnes

This paper uses a Cournot oligopoly experiment to test Meadows (1970) dynamic hypothesis. Meadows model aims to explain the causes of cycles in the US hog industry during the 50's and 60's. At the time, instabilities in this industry had been commonly associated with instabilities in the corn supply (a critical food source for hogs). After the US government implemented the Agricultural Adjustment Act of 1938, through which corn supply was stabilized, the hog market cycles did not dismiss. In fact, oscillations two decades after the Act was implemented were even larger than before (Dean and Heady, 1958). Meadows' model starts by postulating that it is not the corn price that determines profitability for pig farmers, it is rather the ratio of hog price and corn price. When this ratio increases, it is more profitable to sell hogs, when this ratio decreases, it is less profitable to sell hogs (Meadows, 1970).

Meadows model argues for an endogenously generated cyclical behavior. He argues that by stabilizing corn availability and price, the Agricultural Adjustment Act of 1938 actually enabled farmers to freely expand or contract their hog stock faster than before, which ultimately led to even greater oscillations than in previous years (Breimyer, 1959). In order to test his hypothesis, he proposes a dynamic model expanding the Cobweb theorem. Meadows model differs from the original Cobweb formulation in two important aspects. First, the price perceived by farmers is distinct and different from the commodity retail price and second, any change in capacity (breeding stock) will immediately have an effect on the slaughtering rate. Hog farming is modeled using an estimated heuristic, according to which there is a positive relationship between hog-corn price ratio and the desired breeding stock.

This paper uses Meadows' model as a base to develop a Cournot oligopoly experiment. The findings of this experiment suggest that players do not behave with perfect rationality, as none of the markets shows convergence towards the Nash equilibrium. Furthermore, the study finds strong indications of cyclical behavior in most of the experimental markets with some of them exhibiting a period length that resembles the one proposed by Meadows. Regressions over players' decisions indicate that the hog price strongly influences players' decisions as Meadows predicted.

This paper findings indicate that market policies should aim to stabilize prices endogenously. This means considering the likelihood of an increase or a decrease in investment given the current price. In this sense, market policies should create contingencies around players' decisions.

4.4 Paper 4: Making climate conferences more effective?

By Erling Moxnes and David Lara-Arango.

This paper compares two Climate conference (COP) procedures, one that resembles the Kyoto protocol and one that resembles the Paris agreements in 2015. In order to compare these two procedures, the paper proposes a novel game design based on previous public good games (Andreoni, 1995). Unlike the traditional game, player payoffs are determined by the agreement reached in the last round only rather than by all rounds. In addition, asymmetries among player payoff functions are introduced to capture the effect of asymmetries among countries. In the same way as previous threshold public good games (Brick and Visser, 2015; Tavoni et al., 2011), players are also informed about what the social optimum is and what their expected average contribution is. This social optimum is an interior solution in the proposed game, which means that the social optimum is less than the sum of players' endowments. This feature accounts for the fact that the climate social optimum does not require that all of a countries' budget has to go to emission abatement.

The two procedures are tested in two experimental treatments, namely Negotiations (NG) and Individual quantity pledges (IP). In the NG treatment, players are expected to reach an agreement by stating their individual investments. If at least one player disagrees with the contribution scheme, the there will be no agreement. If so, the negotiations will carry on until a last round is reached. If no agreement is achieved in the last round, all players will gain the Nash equilibrium payoff (i.e. zero contributions by all of them). In the IP treatment, players are free to state their own investment in the public good without the need to reach an agreement. All pledges become common knowledge after each round. Instead of being asked to agree or disagree as in NG, players in IP are asked to revise their own pledges until the game stops at a point in time that is unknown to them. A within-subject design was used.

This paper presents two main findings. First, the proposed game leads to significantly larger contributions than what is typical in public good (bad) games. Moreover, contributions relative to social optimum do not differ much from what was achieved in the Paris COP 21 conference, were pledges were around half of what is needed to reach a stated goal of 2°C warming. In spite of the climate change problem often being framed as a public bad game, this paper results suggest that this issue becomes more of a public good situation when players are asked to contribute towards an agreement, which generates higher cooperation, given that public good frames are known to generate higher cooperation than public bad frames (Andreoni, 1995; Sonnemans et al., 1998). In addition, giving players a focal point by publicly announcing the social optimum also elicits higher cooperation than if the social optimum is not known for certain (Barrett and Dannenberg, 2012). Second, the study finds no significant difference between the two studied procedures, negotiations (NG) and individual quantity pledges (IP). However, NG was found to be unable to reach an agreement in two occasions, which implies that the risk of zero contribution in NG is higher than in IP. This finding is consistent with the failed COPs after Kyoto (Depledge, 2000). These two findings imply that pledges may not necessarily perform as poorly as theory would suggest (Cramton et al., 2015; Gollier and Tirole, 2015; Stiglitz, 2015). The experimental results coincide with the Paris COP 21 in the sense that players were unable to reach the stated goal.

This paper suggests that improved designs are needed for future COPs, in order to increase the chances of reaching a desired goal. One option to improve COPs could be to enhance face to face communication, as it has shown significant benefits when it comes to increase cooperation (Ostrom, 1990; Hackett et al., 1994; Rege and Telle, 2004). While face to face communication is present among negotiators, politicians and country leaders are far away from the venues of the COPs and thus, this communication benefit is reduced. Taking advantage of communication benefits across different stakeholders in different countries seems important to foster higher cooperation.

4.5 Paper 5: Socially aversive personalities and income distribution: Can the Dark Triad predict behavior in the dictator and gangster games?

By David Lara-Arango.

This paper studies to what extend the Dark triad (Paulhus and Williams, 2002) can predict players' decisions in the dictator and gangster games. The triad is composed of three aspects that can be measured separately, namely Narcissism (Morey et al., 2012), Machiavellianism (Jones and Paulhus, 2009) and Psychopathy (Hare and Neumann, 2008). However, previous studies have shown that these three traits have a higher predictive power when they are considered jointly, as a constellation of traits rather than isolated parts (Paulhus and Williams, 2002). Hence, this study uses the Dirty dozen questionnaire (Maples et al, 2014) to assess players aggregated Dark triad scores.

This study uses the dictator and gangster game (inverted dictator game). Master students in Economics and System Dynamics were recruited. Since a within-subject design was used for this experiment, players were given a dirty dozen questionnaire to fill in once they were finished with both games. The questionnaires were answered in separate work stations to ensure privacy. They were also received and processed anonymously. By using the same code for questionnaires and games, it was possible to test for a relationship between the two.

This paper's findings suggest no significant relationship between players' Dark triad scores and their decisions in the experiment. Only marginally significant relationships between Dark triad components and Economics students' behavior in the dictator game were found. Since both games used in this paper are anonymous one-shot games, it was not possible for the players to build reputation. Not allowing players to build reputation obscures the distinction between Psychopathy and Machiavellianism (Kessler et al., 2010; Jones and Paulhus, 2011a). The paper findings suggest that personal attributes such as having a callous personality has little to no relation to income distribution decisions. Further research is needed to account for situational circumstances, such that personal attributes can be better put in context and thus, maybe have a higher predictive power for player's decisions.

5. Conclusions and future research

Non-cooperative game theory (NCGT) is a powerful method that allows decision makers and researchers to understand, conceptualize, and solve problems. Its theoretical nature allows researchers to find principles that can explain a wide range of situations and their respective expectable outcomes. This theoretical nature however, also entails challenges when one intends to use NCGT to formulate policies for real life problems. Hence, there is a clear need for a bridge between a powerful theoretical tool such as Game theory and real-life policy making. This thesis proposes simulations and experiments as suitable methodologies that can help to build such a bridge by improving our understanding of how people make decisions.

Non-cooperative situations are often regarded as difficult to solve and in need of better understanding of the problem they entail. The present thesis shows that dynamic non-cooperative games can present unexpected behaviors that are endogenously generated and are often not predicted by theory. These endogenous dynamics are often the result of relationships between decision makers' actions and system features such as delays, non-linearities and feedback loops. Hence, successful policies need to be built on a solid understanding of these relationships. Both simulations and experiments can be useful in this respect. By providing a structure-based causal framework, simulation methodologies such as System dynamics can offer a context in which Game theory solutions can be tested, and refined to suit a more complex reality. On the other hand, experiments allow for a deeper understanding of players' decision rules that can be later used to improve theory. A combined use of Game theory, simulations, and experiments allows to have more reliable theories as basis for policy making.

Regarding this works' limitations, the papers comprised in this thesis leave a number of questions open that call for future research.

5.1. Cournot oligopoly applications to study commodity markets

Regarding electricity markets, this thesis suggests that market oriented mechanisms have a higher chance of improving market performance than interventionist mechanisms, which is consistent with previous studies (de Vries and Hakvoort, 2004; Meunier and Finon, 2006; Finon and Pignon, 2008). Since only two capacity mechanisms (one interventionist and one market oriented) were considered, future research is needed to further validate (or refute) these findings with other mechanisms and other market conditions. Future research is also needed to test the implementation of these and other capacity mechanisms in different energy generation matrixes, particularly in the context of the current energy transition many countries are undergoing. In this respect, market stochasticity also needs to be addressed with different approaches that are not only limited to production uncertainty e.g. uncertainty in capacity construction projects, institutional uncertainty, changing demand patterns, etc.

Regarding endogenously generated instabilities, the fourth paper of this thesis supports Meadows (1970) theory of endogenously generated cycles. Rather than being the result of exogenous phenomena, these cycles result from players' decision making, which can be explained in terms of adaptive expectations (Nerlove, 1958). These findings point out a need to formulate policies that can endogenously mitigate the effects of such strategies e.g. implementing financial mechanisms that discourage investments during price booms and promotes investments during price busts. Further research is needed to test this postulate and to test the effectiveness of specific stabilization policies in Meadows model, as well as in other types of commodity markets. Producer education is another possibility that should be studied.

5.2. Public good game applications to study COPs

When comparing pledges against commonly agreed-upon quotas, it is important to note that the former eliminates the problem of assigning individual quotas. Assigning quotas has been identified as one of the main impediments in COPs before Paris, to successfully address the climate problem (Depledge, 2000; Gollier and Tirole, 2015; Stiglitz, 2015). Experimental evidence in this thesis suggests that players tend to behave accordingly to their payoff function. Thus, if countries behave as players in such experiments, pledges will be better in representing countries payoff functions than assigned quotas. Further research is needed to test this hypothesis by exploring to what extent pledges in the COP 21 reflected countries' payoff functions.

Regarding positive influences on players' contributions, this thesis suggests that leading nations should set an example to other nations by showing willingness to contribute. Leading nations could encourage other nations to contribute by acting as active leaders promoting cooperation (Moxnes and Van der Heijden, 2003). In this respect, further research is needed to propose ways in which the benefits of face-to-face communication (Ostrom, 1990; Hackett et al., 1994; Rege and Telle, 2004) can be transmitted across different stakeholders in the negotiation, such that cooperation is increased and this leadership effect is effectively used. For negative influences, the presented experimental evidence shows that high standard deviations in contributions reduces future contributions. Thus, policies aiming to enforce and sustain cooperation should consider reasonable and graduated punishments (Ostrom, 1990) for free riders in order discourage this behavior, while preventing future retaliations by punished free riders (Grechenig, 2010). Future research is needed to investigate different punishment mechanisms and their corresponding effectiveness on preventing free riding and fostering cooperation.

5.3. Dictator game applications to study income distribution

When it comes to variables that can predict behavior on income distribution problems, this thesis suggests that the Dark triad of human personality (Paulhus and Williams; 2002) does not effectively predict how people decide on these issues. This suggests that, in a similar way as other predictors such as IQ, the Dark triad prediction power can be highly context-dependent (Murray, 1998; Kamphaus, 2005; Neisser, et al., 1996). As such, the Dark triad can constitute a form of the fundamental attribution error (Ross, 1977). However, two components of this triad seem to present a low, yet interesting chance of being relevant when a situation requires giving money to a counterpart. Machiavellianism and Psychopathy are two personality traits that are strongly linked to how likely a person is to disregard others well-being (Kessler et al., 2010; Hare and Neumann, 2008). Therefore, these two traits are correlated with selfish behavior, which in principle could give an indication of how likely a person is to "give something away". Future research is needed to propose experimental designs that would allow to further test the Dark triad as a behavioral indicator for income distribution decisons.

The difference between Machiavellianism and Psychopathy is mainly based on the strategic use of self-reputation by the former and the risk-taking, non-empathic behavior of the latter. The experimental design proposed in this thesis fails to capture the effect of players' reputation, which implies that the border between the two traits becomes unclear and it is not possible to clearly separate the effects of these two traits on decision making (Jones and Paulhus, 2011a). While this does not affect the results presented in this thesis, future research is needed to explore specific relationships between the individual elements of the Dark triad and human behavior on specific situations.

Regarding differences across players' backgrounds, the studies in this thesis show that System dynamics master students and Economics master students were equally insensitive to changes in framing. However, only Economics students presented significant relationships between their Dark triad scores and their income distribution decisions. Future research is needed to explore relationships between specific backgrounds and income distribution decisions.

6. References

Agiza, H.N. (1998). Explicit stability zones for Cournot games with 3 and 4 competitors. Chaos, Solitons & Fractals 9, 1955-1966.

Ahmed, E., Agiza, H.N. (1998). Dynamics of a Cournot game with n competitors. Chaos, Solitons & Fractals 9, 1513-1517.

Alcaraz, S. (2010). Free competition vs. auctioning of licenses for electricity production. M.Phil. thesis. System Dynamics Group, Department of Geography, University of Bergen.

Andreoni, J. (1995) Warm-Glow versus Cold-Prickle: The Effects of Positive and Negative Framing on Cooperation in Experiments. The Quarterly Journal of Economics 110:1-21.

Arango, S., Castañeda, J.A., Larsen, E.R. (2013). Mothballing in power markets: an experimental study. From regulation to competition. Energy economics 36, 125-134.

Arango, S., Moxnes, E. (2012). Commodity cycles, a function of market complexity? Extending the cobweb experiment. Journal of economic behavior and organization 84 (1), 321-334.

Arango, S., Larsen, E.R. (2011). Cycles in deregulated electricity markets: empirical evidence from two decades. Energy policy 39 (5), 2457-2466.

Aumann, R. Hart, S. (2002). Handbook of Game Theory with Economic Applications. North Holland ed.

Aumann, Robert J. (1973). Disadvantageous Monopolies. Journal of Economic Theory, 6 (1), 1–11.

Aumann, R. J., J. H. Dreze. (1974). Cooperative Games with Coalition Structure. International Journal of Game Theory, 3, 217–37.

Aumann, R., Schelling, T. (2005). Robert Aumanns and Thomas Schellings contributions to game theory: Analyses of conflict and cooperation. Advanced information on the Bank of Sweden Prize in Economic Science in Memory of Alfred Nobel, 10.

Barberis, N.C. (2013) Thirty years of prospect theory in economics: a review and assessment. J. Econ. Perspect. 27, 173–195.

Bardsley, N. (2008). Dictator game giving: altruism or artefact? Experimental Economics, 11(2), 122–133. doi:10.1007/s10683-007-9172-2.

Barrett S, Dannenberg A (2012) Climate negotiations under scientific uncertainty. P. Natl. Acad. Sci. U.S.A. 109:17372-17376.

Beullens, P., Zaibidi, N. Z., Jones, D. F. (2012). Goal programming to model human decision making in ultimatum games. International Transactions in Operational Research, 19(4), 599-612. doi:10.1111/j.1475-3995.2011.00826.x

Bergstrom, T., Blume, L., and Varian, H. (1986). On the Private Provision of Public Goods. Journal of Public Economics. 29. p.25-49.

Bernheim, B.D. (1986). On the Voluntary and Involuntary Provision of Public Goods. American Economic Review. 76(4).

Bischi, G.I., Naimzada, A. 2000. Global analysis of a dynamic duopoly game with bounded rationality, in: J.A. Filar, V. Gaitsgory, K. Mizukami (Eds.), Advances in Dynamic Games and Applications, vol. 5 of Annals of the International Society of Dynamic Games, Birkhauser Boston, 361–385.

Bolton, G., Ockenfels, A. (1998). Strategy and Equity: An ERC-Analysis of the Güth-van Damme Game. Journal of Mathematical Psychology, 62, 215–226.

Bolton, P., Dewatripoint M. (2005). Contract Theory. Boston, MA: MIT Press.Brandts, J., Pezanis-Christou, P., Schram, A. (2008). Competition with forward contracts: A laboratory analysis motivated by electricity market design. Economic Journal, 118(525), 192-214. Retrieved from <Go to ISI>://WOS:000251859800009

Breimyer, H. F. (1959). Emerging phenomenon: A cycle in hogs. Journal of farm economics 41, 760-768.

Brick K, Visser M (2015) What is fair? An experimental guide to climate negotiations. Europ. Econ. Rev. 74:79-95.

Burger, N, E., Kolstad, C, D. (2008). Voluntary Public Goods Provision, Coalition Formation, and Uncertainty. Working Paper 15543. National Bureau of Economic Research.

Campbell, W. K., Goodie, A. S., Foster, J. D. (2004). Narcissism, confidence, and risk attitude. Journal of Behavioral Decision Making, 17, 297–311.

Cánovas, J.S., Puu, T, Ruiz, M. (2008) The Cournot-Theocharis problem reconsidered. Chaos, Solitons and fractals, 37 (4): 1025-1039

Carmichael, F. (2005). A Guide to Game Theory. Pearson Education Limited

Carraro, Carlo. (2003). The Endogenous Formation of Economic Coalitions, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.

Charness, G. and Rabin, M. (2002). Understanding Social Preferences with Simple Tests. Quarterly Journal of Economics, 117, 817-869.

Chatain, O. (2014). Cooperative and Non-Cooperative Game Theory in Augier, M., Teece, D.J. The Palgrave Encyclopedia of Strategic Management. Springer.

Chen, M. H., Wang, L., Sun, S. W., Wang, J., Xia, C. Y. (2016). Evolution of cooperation in the spatial public goods game with adaptive reputation assortment. Physics Letters A, 380(1-2), 40-47.

Cooper R (2010) The case for charges on greenhouse gas emissions. Post-Kyoto international climate policy: Implementing architectures for agreement:151-178.

Cooper, R, W., Douglas V. D, Robert, F. Thomas, R, W. (1990), Selection Criteria in Coordination Games: Some Experimental Results. American Economic Review, 80 (1),218–33.

Cournot, A.A. (1838). Recherches sur les Principes Mathematiques de la Theorie des Richesses. Bibliothèque des Textes Philosophiques.

Cramton, P., Ockenfels, A., Stoft, S. (2015) An international carbon-price commitment promotes cooperation. Economics of Energy and Environmental Policy 4:51-64.

Clot, S., Grolleau, G., Ibanez, L. (2013). Anti-social behaviour and self-licensing: experimental evidence from a gangster game. Presented at 30. Journées de Micraoeconomie Appliquée, Nice, FRA (2013-06-06 - 2013-06-07). http://prodinra.inra.fr/record/264362.

de Vries, L.J., Hakvoort, R.A. (2004). The question of generation adequacy in liberalized electricity markets. Working paper no. 120.2004, Fondazione Eni Enrico Mattei.

Dean, G.W., Heady, E.O. (1958). Changes in supply response and elasticity for hogs. Journal of farm economics 40, 845-860.

Depledge, J. (2000). Tracing the Origins of the Kyoto Protocol: an Article-by-article Textual History. United Nations, UNFCCC Framework Convention on Climate Change, United Nations, UNFCCC Framework Convention on Climate Change.

Diederich, J., Goeschl, T., Waichman, I. (2016). Group size and the (in)efficiency of pure public good provision. European Economic Review, 85, 272-287.

Diehl, E., Sterman, J.D. (1995). Effects of feedback complexity on dynamic decision-making. Organizational behavior and human decision processes 62(2), 198-215.

Dinar, A., Albiac, J., Sánchez-Soriano, J. (2008). Game Theory and Policy Making in Natural Resources and the Environment. Routledge.

Ding, Z. W., Wang, Q., Jiang, S. M. (2014a). Analysis on the dynamics of a Cournot investment game with bounded rationality. Economic Modelling, 39, 204-212.

Ding, Z. W., Zhu, X. F., Jiang, S. M. (2014b). Dynamical Cournot game with bounded rationality and time delay for marginal profit. Mathematics and Computers in Simulation, 100, 1-12.

Engel, C. 2010. Dictator Games: A Meta Study. Max Planck Institute for Research on Collective Goods.

Fehr, E. and Schmidt, K.M. (1999). A Theory of Fairness, Competition and Cooperation. Quarterly Journal of Economics, 114, p817-868.

Finon, D., Pignon, V. (2008). Electricity and long-term capacity adequacy: the quest for regulatory mechanism compatible with electricity market. Utilities policy, 16, 143 – 158.

Forsythe, R., Joel, L. H, N.E. S., Martin, S. (1994). Fairness in Simple Bargaining Experiments. Games and Economic Behavior 6: 347-369.

Fouraker, L., Shubik, M. and Siegel, S. (1961) Oligopoly bargaining: the quantity adjuster models. Research Bulletin, Pennsylvania State University, Department of Psychology.

Friedman, D., Cassar, A. (2004). Economics lab: an intensive course in experimental economics. London: Routledge.

Fudenberg, D., E. Maskin. (1986). The Folk Theorem in Repeated Games with Discounting and with Incomplete Information. Econometrica, 54, 533–54.

Furnham, A., Richards, S. C., Paulhus, D. L. (2013). The Dark Triad of personality: A 10-year review. Social and Personality Psychology Compass, 7, 199-216.

Gao, L., Wang, Z., Pansini, R., Li, Y. T., Wang, R. W. (2015). Collective punishment is more effective than collective reward for promoting cooperation. Scientific Reports, 5. doi:10.1038/srep17752

Gibbons, R. (1992). Game Theory for Applied Economists. Princeton University Press.

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

Grechenig K, Nicklisch A, Thöni C (2010) Punishment despite reasonable doubt—a public goods experiment with sanctions under uncertainty. Journal of Empirical Legal Studies 7:847-867.

Gyorgy, M. (2008). Praising Punishment. Szociologiai Szemle(3), 109-123. Retrieved from <Go to ISI>://WOS:000262039400005.

Hackett, S., Schlager, E., Walker, J., (1994) The role of communication in resolving commons dilemmas: experimental evidence with heterogeneous appropriators. J. Environ. Econ. Manage. 27:99-126.

Hardin, G. (1968). The tragedy of the commons. Science, 162, 1243-1248.

Hare, R. D., Neumann, C. S. (2008). Psychopathy as a clinical and empirical construct. Annual Review of Psychology, 4, 217246.

Hecking, H., Panke, T. (2015). The global markets for coking coal and iron ore - Complementary goods, integrated mining companies and strategic behavior. Energy Economics, 52, 26-38.

Hoggatt, A. (1959) An experimental business game. Behavioral Science 4: 192–203.

Hong, F. H., Lim, W. (2016). Voluntary participation in public goods provision with Coasian bargaining. Journal of Economic Behavior and Organization, 126, 102-119.

Huck, S., Normann, H-T., Oechssler, J. (1999). Learning in Cournot Oligopoly - An Experiment. The Economic Journal 109 (454) C80-C95

Huck, S., Normann, H.T., Oechssler, J. (2004). Two are few and four are many: number effects in experimental oligopolies. Journal of Economic Behaviour and Organization: 53(4): 435-446.

Inman M (2009) The climate change game. Nature Reports Climate Change:130-133.

Islyaev, S., Date, P. (2015). Electricity futures price models: Calibration and forecasting. European Journal of Operational Research, 247(1), 144-154. doi: 10.1016/j.ejor.2015.05.063.

Jonason, P. K., Krause, L. (2013). The emotional deficits associated with the Dark Triad traits: Cognitive empathy, affective empathy, and alexithymia. Personality and Individual Differences, 55, 532–537.Jones, D. N., Paulhus, D. L. (2009). Machiavellianism. In M. R. Leary, & R. H. Hoyle (Eds.), Handbook of individual differences in social behavior
(pp. 93108). New York: Guilford.

Jones, D. N., Paulhus, D. L. (2011a). Differentiating the Dark Triad within the interpersonal circumplex. In L. M. Horowitz, and S. Strack (Eds.), Handbook of interpersonal psychology: Theory, research, assessment, and therapeutic interventions (pp. 249267). New York: Wiley.

Jones, D. N., Paulhus, D. L. (2011b). The role of impulsivity in the Dark Triad of personality. Personality and Individual Differences, 51, 679–682.

Kamphaus, R. W. (2005). Clinical assessment of child and adolescent intelligence. Springer. ISBN 0-387-26299-7.

Kandori, M., Mailath, G. and Rob, R. (1993). Learning, Mutation, and Long Run Equilibria in Games, Econometrica, 61, 29-56.

Kahneman, D., Knetsch, J. L., Thaler, R. (1986). Fairness as a Constraint on Profit Seeking: Entitlements in the Market. American Economic Review, 76(4), 728-741.

Kessler, S., Bandelli, A., Spector, P., Borman, W., Nelson, C., & Penny, L. (2010). Re-examining Machiavellianism. Journal of Applied Social Psychology, 40, 1868-1896.

Kingsley, D. C., Brown, T. C. (2016). Endogenous and costly institutional deterrence in a public good experiment. Journal of Behavioral and Experimental Economics, 62, 33-41. doi: 10.1016/j.socec.2016.03.005

Lara-Arango, D. 2014. Capacity mechanisms in electricity markets: An experimental study using system dynamics modeling. LAP Lambert Academic Publishing.

Ledyard, J.O. (1995). Public Goods in Eds. John H Kagel and Alvin E. Roth The Handbook of Experimental Economics. Princeton: Princeton University Press.

Ledyard, J. (1997). Public goods: A survey of experimental research. Retrieved from McCain, R.A. (2009). Game Theory and Public Policy. Edward Elgar Publishing Limited.

Maples, J. L., Lamkin, J., Miller, J. D. (2014). A test of two brief measures of the Dark Triad: The Dirty Dozen and Short Dark Triad. Psychological Assessment, 26, 326331.

Maskin, Eric. (2004). Bargaining, Coalitions and Externalities. Plenary lecture, Second World Congress of the Game Theory Society, Marseille.

Matsumoto, A., Merlone, U., Szidarovszky, F. (2015). Oligopolies with contingent workforce and unemployment insurance systems. Communications in Nonlinear Science and Numerical Simulation, 27(1–3), 52-65.

McKelvey, R. D., T. R. Palfrey. (1992). An Experimental Study of the Centipede Game. Econometrica, 60 (4), 803–36.

Meadows, D.L. (1970). Dynamics of Commodity Production Cycles. Wright-Allen Press, Cambridge, Massachusetts.

Meunier, G., Finon, D. (2006). Market power in electricity markets, suboptimal investment and ISO procurement, Presentation to the 26th USAEE/IAEE Conference.

Michael Pickhardt (2006). Fifty Years after Samuelson's "The Pure Theory of Public Expenditure": What are we Left With? Journal of the History of Economic Thought, 28, pp 439-460 doi:10.1017/ S105383720000941X

Mikkelsen, K. B., Bach, L. A. (2016). Threshold Games and Cooperation on Multiplayer Graphs. Plos One, 11(2). doi: 10.1371/journal.pone.0147207.

Morewedge, C, K., Giblin, C, E. (2015). Explanations of the endowment effect: an integrative review. Trends in Cognitive Sciences 19, 6.

Morey, L. C., Hopwood, C. J., Markowitz, J. C., Gunderson, J. G., Grilo, C. M., McGlashan, T. H., et al. (2012). Comparison of alternative models for personality disorders, II: 6-, 8-, and 10-year follow-up. Psychological Medicine, 42, 1705-1713.

Moxnes, E. (2004). Misperceptions of basic dynamics: the case of renewable resource management. System Dynamics review 20(2), 139-162.

Moxnes E, Van der Heijden E (2003) The effect of leadership in a public bad experiment. J. Conflict Resolution 47:773-795.

Moxnes, E. (2012). Individual transferable quotas versus auctioned seasonal quotas. An experimental investigation. Marine Policy, 36, 339-349.

Muth, J.F. (1961). Rational expectations and the theory of price movements. Econometrica 29(3), 315-335.

Murray, C. (1998). Income Inequality and IQ. AEI Press. ISBN 0-8447-7094-9.

Nash, J.F. (1950). Equilibrium points in n-person games. Proc. National Academy of Sciences, 36(1):48-49.

Nash, J.F. (1951). Non-cooperative games. Annals of Math. 54(2):286–295, September.

Nax, H. H. (2015). Equity dynamics in bargaining without information exchange. Journal of Evolutionary Economics, 25(5), 1011-1026. doi:10.1007/s00191-015-0405-9

Nax, H. H., Pradelski, B. S. R. (2015). Evolutionary dynamics and equitable core selection in assignment games. International Journal of Game Theory, 44(4), 903-932. doi:10.1007/s00182-014-0459-1.

Neisser, Ulrich; Boodoo, Gwyneth; Bouchard, Thomas J.; Boykin, A. Wade; Brody, Nathan; Ceci, Stephen J.; Halpern, Diane F.; Loehlin, John C.; Perloff, Robert; Sternberg, Robert J.; Urbina, Susana (1996). "Intelligence: Knowns and unknowns" (PDF). American Psychologist. 51: 77–101. doi:10.1037/0003-066x.51.2.77. ISSN 0003-066X. Retrieved 9 October 2014.

Nerlove, M. (1958). Adaptive expectations and cobweb phenomena. Quarterly Journal of Economics, 72, 227–240.

Nutter, W. (1964). Duopoly, Oligopoly and Emerging Competition. Southern Economic Journal, 30 (April), 342–52.

Olaya, Y., Arango, S., Larsen, E.R. (2015). How capacity mechanisms drive technology choice for power generation. Working paper, University of Lugano.

Olsina, F., Garcés, F., Haubrich, H.J. (2006). Modeling long-term dynamics of electricity markets. Energy policy 34(12), 1411-1433

Osborne, M, J., Rubinstein, Ariel. (1994). A Course in Game Theory. Massachusetts Institute of Technology.

Ostrom, E. (1990). Governing the Commons. Cambridge, Cambridge University Press.
Ott, U. (2006). International Joint Ventures an Interplay of Cooperative and Non-Cooperative Games Under Incomplete Information. Palgrave Macmillan
Owen, G. (1995). Game Theory. 3rd edition, Academic Press.
Ozono, H., Kamijo, Y., Shimizu, K. (2016). Institutionalize Reciprocity to Overcome the Public Goods Provision Problem. Plos One, 11(6).
Palander T. (1939). Konkurrens och marknadsj¨ amvikt vid duopol och oligopol. Ekonomisk Tidskrift, 41:124-145, 222-250.
Paulhus, D. L., Williams, K. M. (2002). The dark triad of personality: Narcissism, Machiavellianism, and psychopathy. Journal of research in personality, 36(6), 556-563.
Pierson, K., Sterman, J. D. (2013). Cyclical dynamics of airline industry earnings. System Dynamics Review, 29(3), 129-156.
Potters, J. and Suetens, S. (2013), Oligopoly experiments in the current millennium. Journal of Economic Surveys, 27: 439–460. doi:10.1111/joes.12025
Poundstone, W. (1992). Prisoner’s Dilemma, New York: Doubleday.
Puu T. (2006). On the stability of Cournot equilibrium when the number of competitors increases. (submitted).
Przepiorka, W., Liebe, U. (2016). Generosity is a sign of trustworthiness-the punishment of selfishness is not. Evolution and Human Behavior, 37(4), 255-262. doi: 10.1016/j.evolhumbehav.2015.12.003
Rapoport, A., Chammah, A. M. (1965). Prisoner's dilemma: A study in conflict and cooperation (Vol. 165): University of Michigan press.
Rasmusen, E. (2007). Games and Information: An Introduction to Game Theory. Malden, MA: Blackwell Publishing.
Rassenti, S.J., Smith, V.L., Wilson, B.J., (2003). Controlling market power and price spikes in electricity networks: demand-side bidding. Proceedings of the National Academy of Sciences of the United States of America 100 (5), 2998–3003.
Rege M, Telle K (2004) The impact of social approval and framing on cooperation in public good situations. J. Public Econ. 88:1625-1644.
Rosenthal, R. (1981), Games of Perfect Information, Predatory Pricing and the Chain Store Paradox. Journal of Economic Theory 25, 92–100.
Ross, L. (1977). The intuitive psychologist and his shortcomings. Advances in Experimental Social Psychology, Vol 10, ed. L. Berkowitz (New York: Academic Press).
Rutledge, R. B., de Berker, A. O., Espenhahn, S., Dayan, P., Dolan, R. J. (2016). The social contingency of momentary subjective well-being. Nature Communications, 7. doi:10.1038/ncomms11825.
Samuelson, P. A. (1954). The pure theory of public expenditure. The review of economics and statistics, 387-389.
Schelling, TC. (2010). Game theory: a practitioner’s approach. Economics and Philosophy 26(1): 27–46.
Simon, H.A. (1979). Rational decision making in business organizations. The American economic review 69(4), 493-513.
Smith, V. (1962). An Experimental Study of Competitive Market Behavior. Journal of Political Economy, 70(2), 111-137.
Smith, J.M. (1982) Evolution and the Theory of Games. Cambridge University Press.
Solutions CfCaE (2015) Outcomes of the U.N. climate change conference in Paris.
Sonnemans, J., Schram, A., Offerman, T. (1998). Public good provision and public bad prevention: The effect of framing. J. Econ. Behav. Organ. 34:143-161.

Sterman, J. (1989). Modeling Managerial Behaviour: Misperceptions of Feedback in a Dynamic Decision Making Experiment. Management Science, 35(3): 321-339.

Stiglitz, J.E. (2015). Overcoming the Copenhagen Failure with Flexible Commitments. Economics of Energy and Environmental Policy, 4 (2).

Stuart, H.W. (2001), Cooperative Games and Business Strategy. K. Chatterjee and William Samuelson (eds), Game Theory and Business Applications, Norwell, MA: Kluwer, pp. 189-211.

Sundt, S., Rehdanz, K. (2015). Consumers' willingness to pay for green electricity: A meta-analysis of the literature. Energy Economics, 51, 1-8.

Tavoni, A., Dannenberg, A., Kallis, G., Löschel, A. (2011). Inequality, communication, and the avoidance of disastrous climate change in a public goods game. P. Natl. Acad. Sci. U.S.A. 108:11825-11829.

Theocharis, R. (1960), On the Stability of the Cournot Solution on the Oligopoly Problem", Review of Economic Studies, 73, 133-134

Tversky, A. Kahneman, D. (1987). Rational choice and the framing of decisions. In: Hogarth, R.M. and Reder, M., Editors, 1987. Rational choice: The contrast between economics and psychology, University of Chicago Press, Chicago.

Van Huyck, J., R. Battalio and R. Beil. (1990). Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure. American Economic Review, 80, 234–48.

Van Long, N. (2010). A Survey of Dynamic Games in Economics. First edition. World Scientific.

Vásquez, C., Battle, C., Rivier, M., Perez-Arriaga, I. (2003). Security of supply in the Dutch electricity market: the role of reliability options. Report IIT-03-0841C, Universidad Pontificia Comillas, Spain

Von Mouche, P., Quartieri, F. (2016). Equilibrium Theory for Cournot Oligopolies and Related Games: Essays in Honour of Koji Okuguchi. Springer.

Wei, Z. Y., Zhao, Z. Y., Zheng, Y. (2016). Moderating Effects of Social Value Orientation on the Effect of Social Influence in Prosocial Decisions. Frontiers in Psychology, 7, 9. doi:10.3389/fpsyg.2016.00952.

Wogrin, S., Centeno, E., Barquin, J., (2013). Generation capacity expansion analysis: Open loop approximation of closed loop equilibria. Power Systems, IEEE Transactions on 28 (3), 3362–3371.

Wu, J. H., Balliet, D., Van Lange, P. A. M. (2016). Reputation, Gossip, and Human Cooperation. Social and Personality Psychology Compass, 10(6), 350-364. doi:10.1111/spc3.12255

Wu, F., Ma, J. H. (2014). The Chaos Dynamic of Multiproduct Cournot Duopoly Game with Managerial Delegation. Discrete Dynamics in Nature and Society. doi:10.1155/2014/206961.

Paper 1

Uncertainty and the Long-Term Adequacy of Supply: Simulations of Capacity Mechanisms in Electricity Markets

This paper is published in Energy Strategy Reviews:

Lara-Arango, D., Arango-Aramburo, S., Larsen, E.R. 2017. Uncertainty and the long-term adequacy of supply: Simulations of capacity mechanisms in electricity markets. Energy Strategy Reviews, 18: 199, 211.

Uncertainty and the long-term adequacy of supply: Simulations of capacity mechanisms in electricity markets

David Lara-Arango
System Dynamics Group
University of Bergen
Norway

Santiago Arango-Aramburo
Decision Science Department
Universidad Nacional de Colombia
Colombia

Erik R. Larsen
Institute of Management
Universita della Svizzera italiana (Lugano)
Switzerland

Abstract

Deregulation in electricity markets has changed the conditions for maintaining long-term adequacy of supply. Particularly in the last decade, security of supply has become a major issue for policymakers due to a number of changes in technology, especially the introduction of renewables, where regulators have introduced capacity mechanisms. In this paper, we focus on the use of two different capacity mechanisms: procurement for long-term strategic reserves contracting, and centralized auctioning for capacity contracts. We investigate the effect of uncertainty on the effectiveness of these two mechanisms in maintaining a stable and sufficient supply of capacity. We use simulation to establish the behavior as the level of uncertainty is increased. Our results suggest that a market's level of uncertainty plays an important role in the effectiveness of these two interventions. The results raise questions about when it is appropriate to introduce either of them.

Keywords: Security of supply, Adequacy of supply, Capacity mechanisms, Strategic reserves, Centralized auctions, Uncertainty.

1. Introduction

Over the last two decades, electricity markets in many countries have gone through several major restructurings, from an initial deregulation to changes in the structure, pricing mechanisms, and regulations when problems or policies required alignment of the markets [1]. The initial focus was on making sure that the deregulation was efficient and effective, which means delivering the promises in terms of new investments, reliability and, in many cases, lower prices. This led regulators and policy makers to focus on the short to medium-term promotion of competition [2, 3] and the prevention of market power [4]. More recently, the discussion has moved on to the long-term security of supply, i.e. the market's ability to deliver enough new investments (and power) at a required time, in order to avoid shortages [5]. This concern has resulted from a number of issues in the last decade, such as the desire to withdraw nuclear capacity in Europe [6], and financial problems for companies in the electricity sector [7], among others.

While there are many elements in security of supply, we shall focus on capacity adequacy, i.e. making sure that there is enough available capacity to deliver electricity at a reasonable market-based price [8]. We start from what might be seen as the result of the investment behavior in deregulated electricity markets: the occurrence of capacity cycles. These capacity cycles are generally seen as a major threat to markets' sustainability and society's welfare [9]. Cycles in generation capacity have been discussed during the last two decades [10, 11, 12, 2, 13]. More recently, there has been empirical evidence that cyclical behavior does occur in deregulated markets [14]. When there is excess capacity, the capacity cycle creates a situation with relatively low prices, benefiting the consumers, while generation companies have low or no profit. This in turn will lead to limited investment in new generation capacity, as the economic return is not sufficient, which will eventually erode the excess capacity and create a shortage, thereby reversing the benefit between the consumer and the generation companies. This situation might compromise the adequacy of capacity as prices could soar and blackouts might occur more frequently. Such cyclical behavior takes a significant amount of time to correct because of the interplay of reluctant investors (due to the previous period of low return) and the long lags in adding new generation capacity, which typically takes from three years for CCGT, and up to a decade for big hydro and nuclear plants.

There is a possibility that security of supply will be further compromised in the future; see, e.g. [16]. One reason for this is the numerous policy initiatives for introducing renewable energy. Some countries, like Germany, have now reached more than 49 percent of installed capacity and 25 percent of production from renewables [15]. While renewable energy has a number of advantages in terms of the environment, less dependence on fossil fuels etc., it has the potential to create an issue for security of supply. As renewables often get first priority in scheduling (as well as a different pricing mechanism), the residual demand for the remaining generation capacity, such as CCGT and coal, is significantly lower, i.e. a fraction of the total demand is "reserved" for renewables. The immediate consequence of this is that generation companies with thermal and nuclear generation capacity produce less and thereby get lower revenue and profit. There are examples in, for example Germany, where CCGT plants produce during only one out of four days. This has led to closure of thermal plants, e.g. in England, where the regulator has expressed concern about the future reserve margin [16]. The reason for this is that, in periods when the renewable generation has a relatively small production, e.g. due to weather conditions, there is a need for the thermal generation plants to make up the missing production. However, because of the low economic return the required thermal capacity might have been decommissioned or mothballed. Even if prices increase, it is unlikely that utility companies are going to invest in new thermal capacity on the basis of market conditions, if they do not believe that they can meet their minimum threshold return on the investment.

We have observed similar behavior in other cases where there has been a large dependence on hydro, particularly in South America, where the Pacific weather system has created situations of excess water in some periods, followed by a shortage in others. This has led to a high volatility in prices and reluctance to invest in thermal capacity to offset the variability in water, due to the relatively long period of excess water during which the thermal capacity would not produce [17].

Regulators and policymakers have become increasingly aware of this issue and have started to be concerned that this might eventually lead to higher probability of blackouts and a general threat to the security of supply [16]. The response has been to make changes in the regulatory frameworks to include options for the regulators to try to solve these issues, particularly in the area of ensuring adequate thermal capacity to maintain a reserve margin large enough to offset the variability in the production of renewables. Different policies have been proposed and implemented in deregulated electricity markets to maintain adequate capacity and prevent the cycles, such as capacity mechanisms [18], mothballing [5], and forward markets [19].

In this paper, we investigate two of the capacity payment mechanisms suggested in the literature. The first is procurement for long-term strategic reserves contracting, and the second is centralized auctioning for capacity contracts. The first one is an interventionist mechanism that introduces a regulator-owned firm into the electricity market. The second is a market-oriented mechanism that consists of the implementation of a centralized auctioning system, where the market participants bid for capacity contracts [18]. By using simulation, we test these two mechanisms under different levels of uncertainty to understand which of them is the most efficient in maintaining a desirable level of generation capacity and in avoiding capacity cycles.

The paper is organized as follows: the next section presents the capacity mechanisms we consider. The third section explains our economic models. The fourth section shows the simulation's results and finally, the fifth section presents the conclusion and discussion of our findings.

2. Capacity mechanisms

We focus on two capacity mechanisms in order to test their economic impacts on investment, in a stylized electricity market. We have selected one interventionist mechanism, procurement for long-term strategic reserves contracting, and one market-oriented mechanism, centralized auctioning for capacity. We investigate whether they both represent an economic improvement for a market base line, and if so, which of the two yields the best results, under different levels of uncertainty. We select these two specific mechanisms because they both have a good theoretical foundation in the literature, and they represent two theoretically opposed ways of solving the issue of maintaining adequate capacity. The first one provides partial market control (influence) for the regulator, while the second is an integrated part of the market dynamics. One might argue that both, in their own ways, are interventions that partly set aside the idea of a market, i.e. interventions that to some degree suspend the market. While this is not necessarily a bad thing, given that regulators have overseen and intervened in the market since the beginning of deregulation, one has to consider that market principles must be preserved. It can be argued that such interventions can be necessary in order to maintain a well-functioning market.

Procurement for long-term strategic reserves contracting allows a governmental institution to use production capacity. This mechanism has typically been implemented in the form of an agreement between the TSO and generators [45, 46]. However, different authors consider that the regulator can also intervene in the market through a state-owned firm [18]. We choose to introduce this mechanism in the form of a regulatory firm that makes investments in capacity for reasons of transparency, but it should be noted that the effects on the market are similar for the two mechanisms. Countries like France, Sweden, and New Zealand have implemented this mechanism. The results in these countries and the academic discussion of them have portrayed high efficiency in capacity adequacy as its main advantage, and a reduced compatibility with market principles as its main disadvantage, as it interferes directly in the market and is not linked to a market-based mechanism [20, 21].

Centralized auctioning for capacity licenses is the second mechanism we investigate, where the government or regulator has control over the total market capacity and holds auctions for licenses to build new capacity when there is a perceived need. In our model, generators do not need to be successful in the auctions for capacity market licenses in order to keep a plant open. However, they do need to obtain licenses in the capacity market to build new plants. The reason for this is that we assume a constant electricity price that is equal to generators' production costs. Therefore, there is no incentive to build a new plant with zero profit-margin; and so, building a new plant is only attractive for the generators when such plant is the result of winning licenses in the capacity market. In other words, our model focuses on the dynamics of the capacity market while leaving the electricity market constant. The generators bid for the right to build capacity, e.g. who will require the lowest subsidy for adding a certain amount of capacity [18]. New England is one area where this mechanism has been implemented [22]. The literature points out capacity-adequacy targeting and market compatibility as its main advantages, and lack of control over physical plants as its main disadvantage [22, 23].

3. Economic model

The analysis of the capacity mechanisms discussed above can be done at different levels of analysis: from a model calibrated to a particular context, such as a country, to a more stylized model, that provides more general insight. We chose the second option, a stylized model for a deregulated electricity market, as it helps to understand the main implications of the two market interventions. In reality, generators adjust capacity year-on-year by closing stations when there is excess capacity and such excess is expected to persist. This adjustment may not be sufficient to eliminate price cycles, but it can dampen them. In fact, previous works on capacity mechanisms have found that the possibility of mothballing capacity can significantly reduce price cycles [4]. Since we consider a stylized market with no interventions as a base case, we decided not to include this feature. Furthermore, mothballing has also been criticized in the literature for enabling generators to raise prices[5].

The model is based on Arango and Moxnes [24]; we extend the model by including the possibility of testing the two capacity mechanisms discussed above.

3.1. Base model

The base model of the electricity market follows the model developed by Arango and Moxnes [24], where investors make investment decisions for capacity in a market-based system. The model represents a stylized electricity market with long capacity lifetimes and investment delays (i.e. capacity construction time). This market setting also resembles other capital-intensive industries, although the lifetime of investment is normally shorter in other industries. Assuming that generators are price takers in our market, this model follows the closed-loop model formulation, in which they choose a capacity to maximize their profits and where the price is the market equilibrium response to the generators' decisions [25]. The model accounts for the main features of electricity markets, namely, the non-storability (i.e. no inventory) and the inelasticity of the demand (demand always matches supply). For simplicity, we assume that only a single generation technology is available. The price of electricity at time t is determined by

$$P_t = Max(A - B * Q_t, 0) \tag{1}$$

where P_t is the electricity price, Q_t is the market production, and A and B are price function parameters. The generators' profit at the time t is given by

$$\pi_t = (P_t - C) * Q_t \tag{2}$$

where C is the marginal cost (SRMC), which includes both operational and capital cost. Since there have been previous works showing similar results for constant and increasing marginal costs [26, 27], we assume the cost (SRMC) to be constant and equal to 1 for the entire market, given that we are only considering one technology. Capacity utilization is assumed to be 100%, which implies that the market's production is equal to the market's capacity. In addition, we assume a capacity construction time of V years and a capacity lifetime of L years, hence, the market's installed capacity and production, in period t, is given by

$$\frac{\partial IC}{\partial t} = CC_t - Dep_t \tag{3}$$

$$\frac{\partial CuC}{\partial t} = X_t - CC_t \tag{4}$$

$$CC = CuC_t/V \tag{5}$$

$$Dep = IC_t/L \tag{6}$$

where IC is the installed capacity, CC is capacity coming online (capacity emerging from the construction queue), Dep is capacity depreciation, CuC is capacity under construction, and X represents the investments. Power stations are, in reality, discrete units; however, we aggregate the capacity, which allows us to use average (continuous) values as a proxy for capacity and the various changes in capacity [49]. We have run a series of simulations with discrete time for our economic models and found no significant difference in the behavior of the price cycles.

In order to model the investments, we use the heuristic proposed by Arango and Moxnes [24]. The heuristic is based on adaptive expectations [28], which can cause cycles in prices and capacity due to the updating of expectations. Previous works have supported Nerlove's adaptive expectations approach [29, 30]. This approach has also been used for capacity construction times to explain endogenous cycles in different industries [31, 24]. We are modeling deregulated electricity markets, where adaptive expectations-based heuristics are generally seen as being an appropriate assumption for investments [14, 47]. The adaptive expectations approach establishes an expected price ExP_t, which is a function of the current price and the current expected price. The current price expectation comes from the price at the immediately preceding time-period. Therefore, ExP_t is given by

$$ExP_t = (1 - \beta_1) * ExP_{t-1} + \beta_1 * P_{t-1} + \beta_0 \tag{7}$$

where β_1 and β_0 are parameters for the calculation of the expected price. As we are using the same heuristic as [24], we adopt the same values they use for the parameters, such that β_1=0.31, and β_0=0.02. Following adaptive expectations, the market investment function has two main parts: partial adjustments for capacity and adjustments for depreciation, similar to the neoclassical investment function [32]. However, as our base case is for an electricity market with cyclical behavior, we assume that the market fails to compensate fully for both capacity and depreciation adjustments. To capture this, we use a linearized version of the investment heuristic proposed by Arango and Moxnes [24] in order to produce cycles in our base-case market. The market's investment function is thus given by:

$$X_t = Max(\partial * Dep_{t+4} + \gamma * Cap_{t+4} + \rho * ExP_t + \delta * P_t + \omega, 0) \tag{8}$$

where X_t is the investment in capacity in period *t*, and ∂, γ, ρ, δ and ω are the market's decision parameters for depreciation, capacity, expected price, actual price and decision adjustment, respectively. Dep_{t+4} and Cap_{t+4} are depreciation and installed capacity four years ahead of period *t*, ExP_t is the expected price in period *t*, and P_t is the market's actual price of electricity in period *t*. With these settings, we actually have a stock management problem [33]. The values of the coefficients are taken from Arango and Moxnes' regressions on experimental results [24].

3.2. Regulator inclusion

To test the procurement for long-term strategic reserves contracting, we introduce a regulatory firm in the base simulation model. The regulatory firm can invest in capacity year by year, with a maximum of *RI* units/year. This capacity limit is set to ensure competition in the market, i.e. preventing the regulator from acquiring market power and thereby maintaining some competition in the market. We do not account for network congestion in our model, which implies that generators are not able to exercise local market power [34], giving all consumers the same level of access to power. An important question regarding this mechanism concerns the use of the reserve capacity: can it generate only in shortage conditions, or can it be used under "normal" conditions as well. In general, the literature suggests that reserve capacity should only be used in periods when the system is under stress [18]. Otherwise, reserve capacity becomes a type of mothballing, through which it is possible to quickly use extra capacity when profits are expected. Hence, in our model, the regulator invests if the capacity four years ahead is going to be less than *CE*, which is the perfect competition equilibrium (where the generators' marginal revenue is equal to their marginal costs). Thus, the decision rule of investment (Z_t) of the regulatory firm is

$$Z_t = if\ Q_{t+4} < CE \begin{cases} IR & if\ CE - Q_{t+4} \geq IR \\ CE - Q_{t+4} & if\ CE - Q_{t+4} < IR \end{cases} \tag{9}$$

where Q_{t+4} is the market's capacity four years ahead, which represents the construction time i.e. the regulator knows the amount of capacity that will enter the market in the next 4 years. With the inclusion of this regulator-controlled firm, we have an additional firm that aims to orient the market toward its perfect competitive equilibrium [24], in order to maximize social welfare.

3.3. Auctioning system

To include centralized auctioning for capacity licenses, we use an auctioning mechanism proposed by Alcaraz [35], where the generators make bids through a bidding curve that determines the quantity of licenses desired at each possible license price. The market's capacity is fixed at *CE* units (perfect competition equilibrium), and the licenses' price is determined through market bidding. Thus, our model focuses on the capacity market dynamics while keeping the electricity market constant. Regarding the licenses' price range, we took the values suggested by Alcaraz [35]. The license price ranges from LP_L to LP_H. A negative price indicates subsidized licenses for the generators, and a positive price indicates that the generators have to pay for the licenses. Therefore, a negative price represents an income and a positive price represents a cost for the generators [35]. The generators have an incentive to make a negative bid for the license, as this will increase their profit. However, a large number of bids will put an upward pressure on the license price, leading to positive values. Conversely, few bids will lead to a lower, i.e. negative, price for the license. The number of licenses is distributed according to the market-bidding curve at the given price. Generators' profits are then given by

$$\pi_{i,t} = (P_t - C) * Q_t - Y_{i,t} * LP_t \tag{10}$$

where $Y_{i,t}$ is the number of capacity licenses given in the market at time *t* and LP_t is the license price at time *t*.

We use a heuristic based on the bidding dynamics described by Cramton and Stoft [36], to model this second mechanism. The bidding curve generated by the heuristic starts with the market's desired capacity at the maximum subsidized licenses' price (LP_L). Then, the curve decreases at a constant rate until it reaches zero capacity desired, at the maximum non- subsidized price (LP_H), or before. The generators can ask for zero licenses at a price that is less than LP_H, which means that the market-bidding curve will have zero licenses from such a price until LP_H). Since the assumptions and simplifications of our auctioning model do not allow us to use empirical data, we use regressions on two previous experimental studies (where a similar auction model was used) as a basis for estimating the parameters [35, 37]. In such studies, the generators' profits, capacities and numbers of licenses are found to be the determinants of the generators' bids. These levels apply for all periods of the simulation. The proposed heuristic to determine the desired capacity at a price of LP_L is

$$DCap_{-2.1} = Min(Max(z + x * GProf + y * Cap, 0), ML) \tag{11}$$

where $GProf$ is the generators' profit, ML is the maximum number of licenses the generators can bid for, and Cap is the market's capacity. The letters z, x, and y represent parameters estimated on the basis of pilot experimental results reported by Lara [37] and Alcaraz [35]. With these values, we can determine the heuristic for the subsequent point in the curve, and then repeat the process for all other points. The heuristics for the subsequent points of the bidding curve are given by

$$DCap_{Lp} = Max\left(i + l * DL_{Lp-m}, 0\right) \tag{12}$$

where $DCap_{Lp}$ is desired capacity at a license price Lp, DL_{Lp-m} is the desired capacity at the previous curve point (m is the separation between the curve points), and i, l are coefficients estimated for the particular curve point (Lp, $DCap_{Lp}$) using the results of Alcaraz [35] and Lara [37].

Using the model described above, we introduce different levels of uncertainty, allowing us to test the capacity mechanisms' robustness when markets become uncertain and volatile. The inclusion of uncertainty is expected to reduce the optimal investment decisions [38, 39]; however, we are interested in understanding how the effectiveness (both absolute and relative) of the two capacity mechanisms is affected by the presence of uncertainty. The levels of uncertainty represent a number of external factors such as: the availability of installed capacity, fuel prices, water availability, weather conditions, demand uncertainty etc., representing different types of electricity markets [14]. We use the consumer, producer, and economic surplus as performance measures, in order to obtain an assessment of the market's economic results [40]. The consumer surplus for the three cases is calculated as

$$Consumer\ surplus_t = \frac{(A - P_t\ ) * Q_t}{2} \tag{13}$$

where A is the maximum price parameter in the price function, P_t is the market price at time t, and Q_t is the market production at time t. Since the introduction of uncertainty creates a separation between capacity and production, we have to differentiate between production costs and capacity costs. Thus, the cost function is now

$$C = \frac{\propto *\ Cap_t + \ \beta * Q_t}{Q_t} \tag{14}$$

where $\propto$ and β are constants, Cap_t is the market capacity, and Q_t is the market production.

The producer surplus for the base model is calculated as the generators' profits, that is

$$Producer\ surplus_t = (P_t - C) * \ Q_t \tag{15}$$

where C is the individual production cost. The producer surplus for the regulator inclusion case is calculated in the same way, with the difference that the regulator's part of the surplus is subtracted. Since strategic reserves are managed by a state-owned firm in our model, we subtract this firm's profits from the producer surplus. By doing so, we make sure that the producer surplus only accounts for the actual generators' profits.

$$Producer\ surplus_t = (P_t - C) * \ (Q_t - R_t) \tag{16}$$

where R_t is the regulator's production at time t. For the auctioning system case, we calculated the producer surplus as the generators' profits, that is,

$$Producer\ surplus_t = (P_t - C) * Q_t - Y_{i,t} * LP_t \tag{17}$$

where $Y_{i,t}$ is the number of capacity licenses given in the market at time t and LP_t is the license price at time t. The economic surplus is the sum of the consumer and producer surplus

$$Economic\ surplus_t = Consumer\ surplus_t + Producer\ surplus_t \tag{18}$$

For the auctioning system case, we subtract the cost for the government from the original Economic surplus formulation, as follows

$$Econ\ surplus_t = Cons\ surplus_t + Prod\ surplus_t - Cost\ to\ the\ gov_t \tag{19}$$

The cost for the government is set as twice the generators' revenue from licenses, when the price is negative. The reason for this is that first, the government has to pay for the licenses when the price is negative (subsidies). Second, the money the government uses to subsidize the generators could have been available for other governmental needs if there were no need to subsidize the generators.

4. Simulations

We start with a series of simulations with the three markets: namely, no capacity payments, regulator intervention, and capacity auction, all without any uncertainty, to establish a benchmark before uncertainty is introduced. For the simulation, we initiate the model with the values shown in Table A1 in the Appendix.

4.1. Without uncertainty

We are interested in the price behavior as well as the consumer, producer, and economic surplus. The capacity is fixed in the implementation of the auctioning system; therefore, the electricity price remains constant and equal to one. Furthermore, notice that the producer profit in the auctioning model might be affected by subsidies from the government, i.e. the negative license price, which is why we include the government's cost in our economic surplus calculation.

Fig. 1a shows the electricity price for the case of no capacity payments, Fig. 1b for long-term strategic reserves contracting, and finally Fig. 1c shows the license payment price in a system with capacity auctions. All simulations in Fig. 1 are without uncertainty. Table 1 shows the corresponding comparison of the basic statistics for consumer, producer, and economic surplus in each of the three cases.

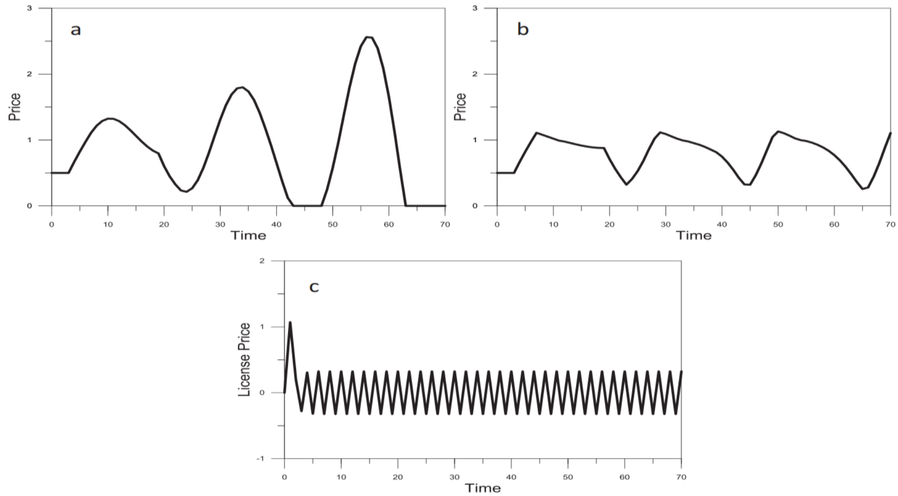


Figure 1. Simulation results for the three markets' prices without uncertainty: a) no market mechanism (electricity price), b) strategic reserve planning, (electricity price), and c) capacity auctions, (license price)

As expected, the base case simulation, i.e. no intervention in the market, shows cyclical price behavior. Furthermore, the oscillations increase in amplitude for each cycle, creating increasing volatility. In the third cycle, price varies between zero and 2.5. While zero might sound unrealistic, we have recently observed negative prices in some markets, i.e. consumers are paid for using electricity [41]. Investment in electricity markets is characterized by lumpiness in investments [42, 43]. We do not include this feature in our stylized model, as it would require detailed assumptions about the investment functions of the individual generators. Note that Figure 1c shows negative prices for capacity. The interpretation of this in our model is that a negative price implies that the generators have been paid to build new capacity, while a positive price implies that the generators have paid to build capacity (see Equation 10). As we stated before, cycles in electricity markets have been identified as a detrimental phenomenon for long-term security of supply [14] since concentrations of investment in high price periods leads to a future overcapacity and low prices, which in turn discourages investment, leading again to high prices and so on [9]. Such a detrimental effect is evidenced in Table 1, where one can see a negative value for the average producer surplus and a high value for the consumer surplus (higher than the optimum and out of the optimum equilibrium). Although the value for the consumer surplus is high, the economic surplus is smaller than the optimum because of the negative producer surplus.

Table 1. Comparison of the three scenarios without uncertainty

		Consumer Surplus	Producer Surplus	Economic Surplus
Base Case	Average	133.59	-12.46	121.13
	Standard Deviation	35.43	37.12	5.09
Reserve Contracting	Average	136.31	-11.86	124.45
	Standard Deviation	13.33	14.05	0.76
Auction System	Average	125.00	-0.06	124.34
	Standard Deviation	0.00	1.07	0.50
Optimum		125.00	0.00	125.00

The second scenario in Fig. 1b, reserve contracting, also generates cycles in the electricity price; however, they are smaller than in the non-intervention scenario. Despite these cycles also expanding, they do it at a much slower rate than what shows in Fig 1a. This scenario generates the largest economic surplus of the three experiments, although the difference is less than 3 percent, as shown in Table 2. The standard deviation is also smaller, as a consequence of the smaller amplitude of the fluctuations. This mechanism, as in the previous case, presents poor results for the producers, with an average producer surplus of -11.86, which raises questions about the system's long-term sustainability, i.e. it is unlikely that investors will continue to invest in an industry where the average result is a significant loss [14].

Fig. 1c represents the capacity auctioning alternative intervention where the capacity is fixed; therefore, the electricity price is stabilized by default and thus there are no electricity price cycles (not shown). However, the instability of the system is represented in the license price, which can be seen in Fig. 1c. In this case, without uncertainty, the producers' profits are determined by the license price, which may represent a reduction in the economic surplus if there are subsidized (negative) license prices. We observe in Fig. 1c that the license price in some periods is positive, i.e. companies pay for being allowed to build capacity, while in other periods they are paid by the regulator for adding capacity. Table 1 shows that the producer surplus is very close to zero, which indicates that the system should be sustainable as companies receive the expected (normal) rent from their investments.

From the above, in the no-uncertainty case, centralized auctioning for capacity is the only sustainable option for electricity markets to ensure capacity adequacy and lower the chances of blackouts.

4.2. Including uncertainty

We now add uncertainty, an important part of real electricity systems as it influences companies' decisions about whether or not to invest. Uncertainty can be generated by many different factors: from technology, weather conditions, fuel cost, regulatory changes, and others. To include uncertainty, we add noise to the model, following the experimental design used by Arango et al. [5]. In their experiment on mothballing, capacity utilization was seen as a decision variable for the subjects, i.e. the subjects could order capacity as well as decide what percentage of their current capacity they wanted to keep online. However, in this case, we are testing for the effect of uncertainty on investment decisions, and we assume that capacity utilization is mainly driven by external factors, i.e. the generators do not make decisions about the fraction of capacity they make available. We assume that they make available for generation the maximum capacity they have, i.e. the capacity utilization decision is imposed on the generators in the form of one or more exogenous phenomena e.g. by a hydrological cycle, or by a maintenance cycle. We are using a stylized electricity market model with one "generic" generation technology. However, we can interpret each of the levels of uncertainty as a different composition of the generation technologies matrix, where the highest levels of uncertainty correspond to a hydro-dominated market and the lower levels correspond to a thermal-intensive market.

To determine how much of the capacity is going to be available for production, we introduced a noise (N) with a normal distribution and adjustable values for the mean and standard deviation. In this way, we have the capacity utilization as an external factor, and we do not have the market constraint on capacity utilization as in Arango et al. [5].

Since we have capacity and production as separate issues, we need to separate capital and operational costs. Therefore, we used the same cost function proposed by Arango et al. [5]

$$Cost = \beta * Q_t + \alpha * IC \tag{20}$$

where Q_t is the market's production, IC is the market's installed capacity and β, α are operational and capital unit costs. We use the same values as Arango et al. [5] for the cost function coefficients, β=0.4, and α=0.6.

We performed a sensitivity analysis for each case of uncertainty: a low uncertainty scenario $N(0.9,0.1)$, a medium uncertainty scenario $N(0.75,0.25)$, and a high uncertainty scenario $N(0.5,0.5)$. When going from one uncertainty-scenario to another, it is important to notice that we are reducing the mean of the normally-distributed noise while increasing the standard deviation (e.g. going from $N(0.9,0.1)$ to $N(0.75,0.25)$). Reducing the mean of the noise and increasing the standard deviation could give rise to two effects: one from the increase in standard deviation, and a second from the reduction of the mean. However, we find that the second effect is negligible; in other words, reducing the mean of the noise while increasing the standard deviation is not significantly different from keeping the mean of the noise constant while increasing its standard deviation. Table 2 shows a comparison between having a constant mean for the noise with an increasing

standard deviation versus a reduction in the mean of the noise mean with an increasing standard deviation. Table 2 shows that the economic surpluses do not change much for the cases of low and medium uncertainty, while in the case of high uncertainty they change significantly. These changes do not alter our analysis when comparing the three cases, i.e. the ranking remains the same regardless of whether we have a constant or reduced mean of the noise. However, the average capacity-utilization shows that keeping the mean of the noise constant is more likely to yield unrealistic production (capacity utilization). As uncertainty increases, one may find periods when the production is higher than what is normally produced (capacity utilization is higher than 100%). Take the case of medium uncertainty as an example where the simulations with a reduced mean $N(0.75,0.25)$, and constant mean $N(1,0.25)$, both have production above 100%. However, having such periods of "overproduction" is more likely in the constant-mean case than in the reducing-mean case (132% vs 111%). Since a reduced mean is less likely to yield extremely high "overproduction" periods and does not affect the relative performance of our three cases, we use this method for our analysis.

Table 2. Comparison between reducing the mean of the noise (mean availability is reduced as the standard deviation of cost is increased) and constant mean of the noise (constant mean for availability while increasing standard deviation of cost).

	(Reducing mean /Constant mean)	Base Case	Reserve Planning	Auction System
Low uncertainty	Economic Surplus	121.17/119.90	121.07/118.84	114.46/113.22
	Mean capacity utilization %	95/103	95/103	95/103
Medium uncertainty	Economic Surplus	112.32/112.96	111.31/112.81	99.77/104.30
	Mean capacity utilization %	111/132	111/132	111/132
High uncertainty	Economic Surplus	90.56/106.89	85.66/106.21	69.22/101.45
	Mean capacity utilization %	137/179	137/179	137/179

We run 40 simulations with different sequences of random numbers for each of the uncertainty scenarios and report the average, and the 10^{th} and 90^{th} percentiles. Beginning with low uncertainty $N(0.9,0.1)$, Figure 2 shows the sensitivity analysis for our three cases. Table 3 summarizes the results.

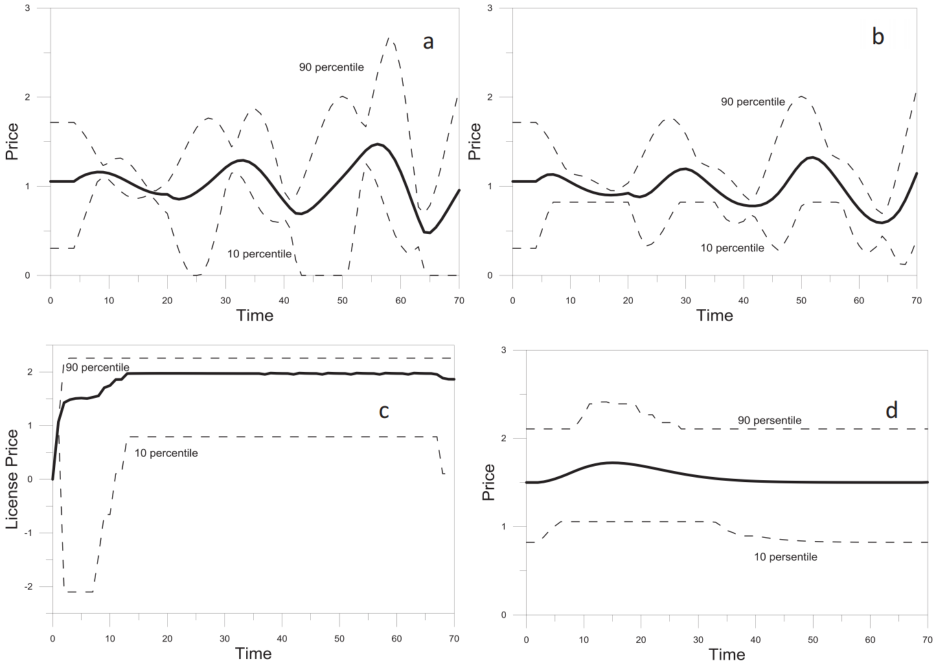


Figure 2. Sensitivity analysis (90% C.I.) for each of the three cases under low uncertainty $N(0.9{,}0.1)$

Fig. 2c shows that there are few cases of negative subsidies (as shown by the 10th percentile) and on average, there are high bids by the companies for the right to build capacity, which is a reflection of the relatively high electricity price seen in Fig. 2d. Overall, the base case presents the highest economic surplus, where before it had the lowest, shown in Table 3; in fact, it has changed little from the case without uncertainty. The two cases with interventions present lost economic value, the auction-based scenario in particular has lost almost 10 percent. The auction-based case has now an important surplus for the producers, which might be problematic in relation to the consumers, as their surplus is significantly lower in this case than it is in the two other cases. There is relatively little volatility in the auction-based system compared to the other two cases, however it is a question of whether the consumers will be willing to pay for this certainty. We can therefore conclude that for low uncertainty there is no strong reason to include any capacity mechanism, or in other words, with low uncertainty, the market would be better off left to itself.

Table 3. Economic results for the inclusion of low uncertainty $N(0.9{,}0.1)$ in the three cases

		Consumer Surplus	Producer Surplus	Economic Surplus
Base Case	Average	124.39	-3.23	121.17
	Standard Deviation	11.38	12.77	3.86
Reserve Planning	Average	126.56	-5.49	121.07
	Standard Deviation	8.97	9.76	3.80
Auction System	Average	98.55	15.91	114.46
	Standard Deviation	3.39	4.02	3.65

We now turn to the case of medium uncertainty, $N(0.75{,}0.25)$. Fig. 3 shows the sensitivity analysis and Table 4, the economic results. The evolution of prices in the three scenarios is similar to that of the prices in the low-uncertainty case shown in Fig. 2; that is, the prices are higher in the no-intervention and capacity auction and lower in the case of strategic reserve contracting. As in the low-uncertainty case, we can see that the no-intervention case provides the best economic value, while the auction system provides the producers with a large surplus, both in absolute and relative terms. We generally observe an increase in both price and volatility in the simulations compared with the previous case. Due to the increase in producer surplus, and the consumer surplus decrease, the consumers are then better off in the case of reserve planning, while the capacity auction is by far the worst case for the consumers. The observed large difference in electricity prices in the case of the auction system is explained by the increased volatility of the production i.e. the electricity price presents broader bands as the production is allowed to have a higher variation in this medium-uncertainty case than in the previous low-uncertainty case. Furthermore, the case of reserve planning is (as it was in the low-uncertainty case) relatively close to the results of the base case. However, it does provide slightly lower electricity prices, visible in Fig. 3b. As in the case of low uncertainty, we find no incentive to pick either of the two intervention mechanisms, as the no-intervention scenario is, on average, more balanced.

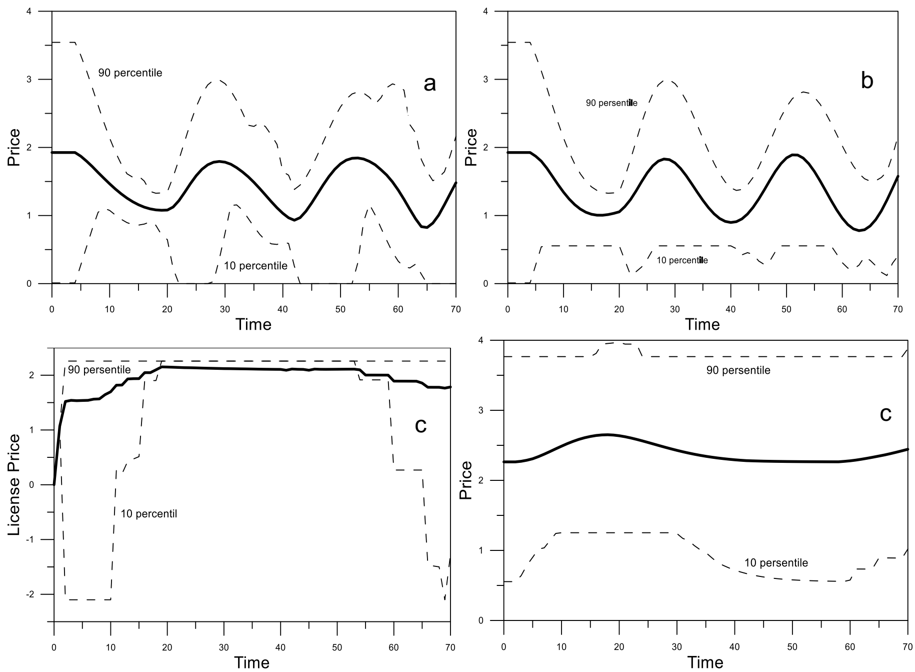


Figure 3. Sensitivity analysis (90% C.I.) for each of the three cases under medium uncertainty $N(0.75,0.25)$

Table 4: Economic results for the inclusion of medium uncertainty $N(0.75,0.25)$ in the three cases.

		Consumer Surplus	Producer Surplus	Economic Surplus
Base Case	Average	104.73	7.59	112.32
	Standard Deviation	14.95	18.10	10.93
Reserve Planning	Average	108.08	3.23	111.31
	Standard Deviation	16.86	18.64	10.29
Auction System	Average	65.29	34.47	99.77
	Standard Deviation	4.64	9.93	10.98

Finally, we test the high-uncertainty scenarios assuming $N(0.5,0.5)$, which makes it difficult for any market agent to foresee a clear sense of the future. Fig. 4 shows the sensitivity analysis in terms of the price evolution over time, and Table 5 presents the economic results. The first observation is that it is difficult to learn much from the simulation; electricity prices seem to almost cover the whole spectrum from close or zero to six, the maximum price allowed in the model. The electricity prices are generally higher in all the scenarios. Table 5 shows that there is even less economic surplus than was previously the case, with the base case still showing the largest economic surplus. In all cases, the producers get a higher surplus than in any of the previous simulations; in the case of the capacity auction system the producers get a higher surplus than the consumer, which has not been observed in the previous cases. On average, the base case presents the most desirable market scenario, although the case of strategic reserve has a slightly smaller producer surplus. Given this, we conclude, based on our simulations, that in an electricity market with a high stochastic component there is no reason to believe that a strategic reserve or an auctioning system will improve the results by much. A broader conclusion might be that given the degree of uncertainty, it is unclear which system will perform best – but at least, the average suggests that not much can be gained by intervening in the system.

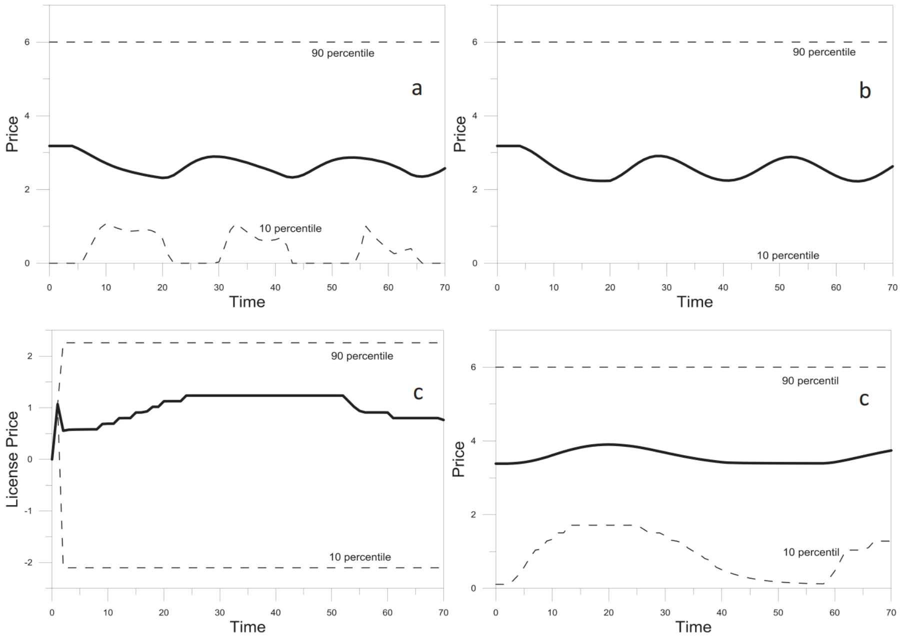


Figure 4. Sensitivity analysis (90% C.I.) for each of the three cases under high uncertainty $N(0.5{,}0.5)$

Table 5: Economic results for the inclusion of high uncertainty $N(0.5,0.5)$ in the three cases.

		Consumer Surplus	Producer Surplus	Economic Surplus
Base Case	Average	55.77	34.80	90.56
	Standard Deviation	8.07	18.55	18.42
Reserve Planning	Average	58.64	27.02	85.66
	Standard Deviation	9.78	12.08	13.72
Auction System	Average	29.70	39.52	69.22
	Standard Deviation	4.28	19.71	20.08

In order to improve understanding of the relationship between performance and the level of uncertainty, we ran additional simulations to establish how producer, consumer, and economic surpluses develop as a function of uncertainty. The results of this analysis are shown in Fig. 5.

All the cases follow a similar trend. The consumer surplus decreases as the uncertainty increases, because the greater the uncertainty, the lower the average production and investment. Following from the lower electricity production, the price starts to increase, which causes a decrease in the consumer surplus. This decreasing consumer surplus is more accentuated in the capacity auctioning system than in the other two cases. The reason for this is that capacity is fixed in the auctioning system, and therefore it is not possible to compensate a lower capacity utilization with an increase in capacity.

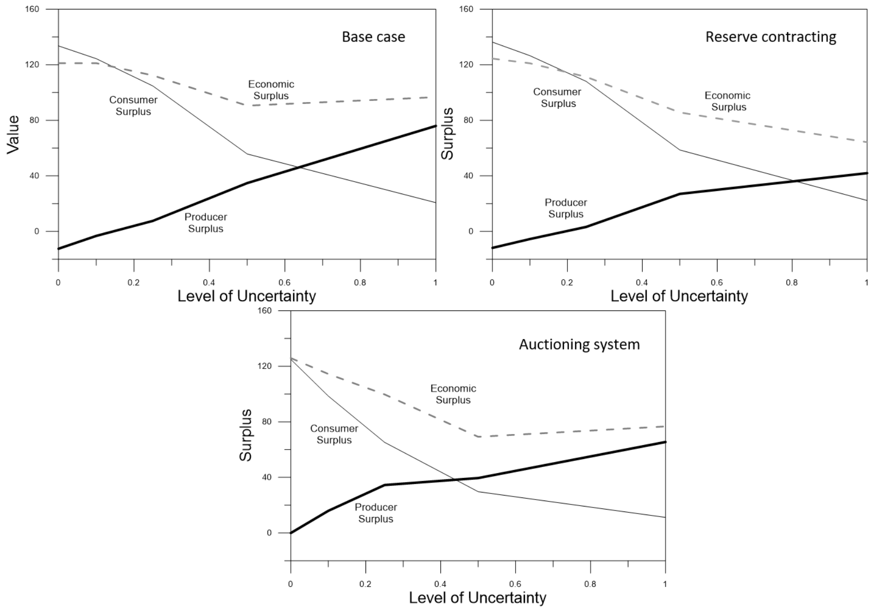


Figure 5. Consumer, producer, and economic surpluses at different levels of uncertainty

The producer surplus starts from relatively low values, negative or zero. Thereafter, it increases as the uncertainty reduces the average production, which in turn increases the price and the profit margins. The lower production is more than offset by the increase in price in all three cases, increasing the overall surplus for producers. The base case scenario shows the steepest increase compared with the other two cases. The first mechanism shows a less steep slope due to the introduction of the regulator, which provides a fraction of the needed capacity, and thus takes market share away from the other generators, which reduces their surplus. The second mechanism restricts the increase of the producer surplus by limiting the generators' profits to the trading of licenses, given that licenses have a much lower margin than electricity (maximum margin of \$5 for electricity vs \$2.1 for licenses). The increase of the producer surplus continues as the uncertainty increases, implying that an increase in uncertainty will be beneficial and should be supported by the producers.

From inspection of the economic surplus we can make two observations. First, the maximum for the economic surplus in the base case and the strategic reserve case is almost constant for low levels of uncertainty, i.e. the graphs are almost flat until a value of uncertainty of around 0.2. On the other hand, in the case of the capacity auction the economic surplus exhibits a steady decline until an uncertainty level of approximately 0.5. Second, in two of the three cases, namely, in the base case and capacity auctions, the economic surplus shows a minimum followed by a small continued increase. In both cases, the minimum is around a level of uncertainty of 0.5. In the case of strategic reserve, we observe a continuous decline in the economic surplus as the uncertainty increases. Thus, we conclude that consumers and producers will always have opposing interests, and the role of the regulator is to balance the interests of these two groups. The producers will always have an interest in trying to increase the uncertainty while the consumers and, in this case the regulator, try to limit the uncertainty. It is equally important for the regulator to ensure that the situation does not end up at the minimum, as that is where the whole system “loses”.

5. Discussion and conclusion

We propose a standard economic model for an electricity market in which we test two capacity mechanisms, procurement for long-term strategic reserves contracting, and centralized auctioning for capacity contracts. We establish a benchmark by means of results obtained from the model without any uncertainty, followed by the introduction of three levels of uncertainty. The performance of the two capacity mechanisms is then compared, each with the other, and in the case of the market without intervention. By doing so, we aim to understand, first, how uncertainty influences performance in these three cases, and second, whether these two capacity mechanisms improve the market performance.

In the case without uncertainty, we find the centralized auctioning for capacity to be the only sustainable option, both in producer surplus and in the market’s stability. The other options give rise to significant negative surplus for the producers, which make them unlikely to be sustainable. These results are consistent with previous work that has shown the advantages of this particular mechanism [20, 21, 18]. When uncertainty is introduced into the model, the results change significantly, as the best results were, on average, from not including any capacity mechanism in the market. Although including a strategic reserve, in some cases improves the results marginally, in reality, it may add complications that are likely to consume more than the marginal benefit when it is implemented. We see that in the last part of the analysis, no-intervention seems to be doing better as the uncertainty increases. The results suggest that there is no incentive to implement either of the two mechanisms we considered when there is significant uncertainty in the market.

Overall, our results show that even though uncertainty leads to worse economic scenarios for society as a whole (i.e. decline in economic surplus), it has on average a positive effect for generators. The reason for this is that a higher uncertainty limits available capacity, which in turn leads to higher prices and higher profit margins. Moreover, such uncertainty also negatively affects the willingness to invest in new capacity, creating a permanent "semi" shortage of capacity and persistent high prices. These results are consistent with [5], who reported higher prices when producers were allowed to mothball capacity.

We return to the old question: what is the best policy for ensuring a sustainable adequacy of capacity in an electricity market? The answer implied in this work is that there is no best model. Our results suggest that the best policy for one market is not necessarily the best for another market. Some production-related uncertainty sources, such as water availability and fuel price, are relevant points to take into account when it comes to deciding on the most sustainable policy. As our results suggest, centralized auctioning for capacity contracts could be a good market policy for stable markets, such as France [14] but it might not be a good market policy for markets with highly stochastic behavior, such as Colombia [44]. In general, we see that even a relatively small degree of uncertainty can make the interventions less effective, making the no-intervention market model a more attractive option. However, having a more controlled market, i.e. a market with a capacity mechanism, may still be appealing due to the risk-averse behavior of both consumers and producers. This behavior might put pressure on politicians to introduce a capacity mechanism, even when it implies a cost for society, especially when we consider the harmful effects of blackouts and shortages, which might outweigh the cost of the regulation. Considerations about uncertainty are more relevant as renewable generation takes a larger share of the market, leaving a decreasing market share for thermal generation. Finally, policy implementation in electricity markets is a difficult task that involves several stakeholders and although the real effectiveness of a given policy is a question than can only be answered in real life, behavioral simulation models and experimental economics can give us helpful insights about market policies' effects.

References

[1] Helm, D. The New Energy Paradigm, Oxford University Press, 2007.

[2] Larsen, E.R., Bunn, D.W. Deregulation in electricity: understanding strategic and regulatory risk. The Journal of the Operational Research Society, 1999, 50(4), System Dynamics for Policy, Strategy and Management Education, 337-344.

[3] Newbery, D.M. Electricity liberalization in Britain and the evolution of market design. In: Sioshansi, F.P., Pfaffenberger, W. (Eds.), Electricity Market Reform 2006: An International Perspective. Elsevier, Oxford, pp. 109–144.

[4] Stoft, S. Power System Economics, 2012. Wiley.

[5] Arango, S., Castañeda, J.A., Larsen, E.R. Mothballing in power markets: an experimental study. From regulation to competition. Energy economics, 2013, 36, 125-134.

[6] Smedley, T. Goodbye nuclear power: Germany's renewable energy revolution. The Guardian, 2013, Friday May 10th (http://www.theguardian.com/sustainable-business/nuclear-power-germany-renewable-energy).

[7] Bloomberg. EON files to close unprofitable Irsching gas power plants [WWW Document]. Bloomberg.com. URL: http://www.bloomberg.com/news/articles/2015-03-30/eon-files-to-close-two-unprofitable-irsching-gas-power-plants (accessed 4.27.15).

[8] EIA. Annual Energy Outlook with projections to 2040, 2014.

[9] IEA. Security of supply in electricity markets: evidence and policy issues. 2002, Paris: Organization for Economic Co-operation and Development.

[10] Bunn, D.W., Larsen, E.R. Sensitivity of reserve margin to factors influencing investment behaviour in the electricity market of England and Wales. Energy Policy, 1992, 20(5), 420-429.

[11] Ford, A. Cycles in competitive electricity markets. Energy policy, 1999, 27, 637–658.

[12] Ford, A. Waiting for the boom: a simulation study of power plant construction in California. Energy Policy, 2001, 29(11), 847-869.

[13] Olsina, F., Garcés, F., Haubrich, H.J. Modeling long-term dynamics of electricity markets. Energy Policy, 2006, 34(12), 1411-1433.

[14] Arango, S., Larsen, E.R. Cycles in deregulated electricity markets: empirical evidence from two decades. Energy policy, 2011, 39 (5), 2457-2466.

[15] Fraunhofer ISE. Electricity production from solar and wind in Germany in 2014. Freiburg, Germany, 20-Oct, 2014.

[16] OFGEM. Electricity Capacity Assessment, 2013. London.

[17] Olaya, Y., Arango, S., Larsen, E.R. How capacity mechanisms drive technology choice for power generation. Working paper 2015, University of Lugano.

[18] Finon, D., Pignon, V. Electricity and long-term capacity adequacy: the quest for regulatory mechanism compatible with electricity market. Utilities Policy, 2008, 16, 143 – 158.

[19] Álvarez, K. Efecto de la introducción de mercados de futuros en la estabilización de ciclos en los precios de commodities mediante experimentos y simulación. MSc. Thesis. Decisions Science Department, Universidad Nacional de Colombia. 2013.

[20] de Vries, L.J., Hakvoort, R.A. The question of generation adequacy in liberalized electricity markets. Working paper no. 120. 2004, Fondazione Eni Enrico Mattei, 2004.

[21] Meunier, G., Finon, D. Market power in electricity markets, 2006, suboptimal investment and ISO procurement, Presentation to the 26th USAEE/IAEE Conference.

[22] Vásquez, C., Battle, C., Rivier, M., Perez-Arriaga, I. Security of supply in the Dutch electricity market: the role of reliability options. Report IIT-03-0841C, 2003, Universidad Pontificia Comillas, Spain.

[23] Bidwell, M. Reliability options. A market-oriented approach to long-term adequacy. The Electricity Journal, 2005, 18 (5), 11-25.

[24] Arango, S., Moxnes, E. Commodity cycles, a function of market complexity? Extending the cobweb experiment. Journal of Economic Behavior and Organization, 2012, 84 (1), 321-334.

[25] Wogrin, S., Centeno, E., Barquin, J. Generation capacity expansion analysis: Open loop approximation of closed loop equilibria. Power Systems, 2013, IEEE Transactions on 28 (3), 3362–3371.

[26] Sutan, A. An experimental investigation of the educative abilities of the agents in negative feedback environments, Ph.D. Dissertation 2005, University Louis Pasteur, Strasbourg.

[27] Huck, S., Normann, H.T., Oechssler, J. Stability of the Cournot process: experimental evidence. International Journal of Game Theory, 2002, 31 (1), 123–136.

[28] Nerlove, M. Adaptive expectations and cobweb phenomena. Quarterly Journal of Economics, 1958, 72, 227–240.

[29] Miller, R.M. Experimental Economics: How we can build better financial markets, 2002. John Wiley and Sons, Hoboken, New Jersey

[30] Plott and Smith. Handbook of Experimental Economics Results, 2008. North Holland.

[31] Meadows, D.L. Dynamics of Commodity Production Cycles. Wright-Allen Press, 1970, Cambridge, Massachusetts.

[32] Jorgenson, D. Econometric studies of investment behavior: a survey. Journal of Economic Literature, 1971, 9 (4), 1111–1147.

[33] Sterman, J.D. Modeling managerial behavior: misperceptions of feedback in a dynamic decision making experiment. Management Science, 1989, 35 (3), 321–339.

[34] Lesieutre, B., Rogers, K., Overbye, T., Borden, A. A sensitivity approach to detection of local market power potential. Power Systems, 2011, IEEE Transactions on 26 (4), 1980–1988.

[35] Alcaraz, S. Free competition vs. auctioning of licenses for electricity production. M.Phil. thesis. System Dynamics Group, Department of Geography, University of Bergen. 2010.

[36] Cramton, P., Stoft, S. The convergence of market designs for adequate generating capacity, with special attention to the CAISO's resource adequacy problem, A White Paper for the Electricity Oversight Board, California, April 2006.

[37] Lara, D. Capacity mechanisms in electricity markets: An experimental study using system dynamics modeling. 2014, LAP Lambert Academic Publishing.

[38] Heck, N., Smith, C., Hittinger, E. A Monte Carlo approach to integrating uncertainty into the levelized cost of electricity. The Electricity Journal, 2016, 29 (3) 21–30

[39] Santos, M, J., Ferreira, P. A methodology to incorporate risk and uncertainty in electricity power planning, 2016, http://dx.doi.org/10.1016/j.energy.2016.03.080.

[40] Baran, P.A., Sweezy, P.M. Some theoretical implications. Monthly Review, 2012, 64 (3), 24.

[41] Mayer, J.N., Kreifels, N., Burger, B. Kohleverstromung zu Zeiten niedriger Börsenstrompreise — Kurzstudie." August, 2013. Frauhofer-Insitut.

[42] Schuler, R. Pricing the use of capital-intensive infrastructure over time and efficient capacity expansion: illustrations for electric transmission investment, 2012.

[43] Nagl, S., Fursch, M., Lindenberger, D. The costs of electricity systems with a high share of fluctuating renewables: A stochastic investment and dispatch optimization model for Europe. The Energy Journal, 2013, 34 (4).

[44] Arango, S. Simulation of alternative regulations in the Colombian electricity market. Socio-Economic Planning Sciences, 2007, 41(4), 305-319.

[45] EC. Proposition de directive du Parlement européen et du Conseil concernant des mesures visant à garantir la sécurité de l'approvisionnement en électricité et les investissements dans les infrastructures. 2003.

[46] EC. Directive 2005/89. /EC of the European Parliament and the Council concerning measures to safeguard security of electricity supply and infrastructure investment. 2005.

[47] Carlson, J.A. The stability of an experimental market with a supply-response lag. Southern Economic Journal, 1967, p. 305-321.

[48] Roques, F.A. Market design for generation adequacy: healing causes rather than symptoms. Utilities Policy, 2008, 16, 171–183.

[49] Sterman, J.D. Business Dynamics. Irwin McGraw-Hill, 2000.

Appendix

Description	**Value**	**Reference**
RI	10	
CE	50	
LP_L	-2.1	Alcaraz (2010)
LP_H	2.4	Alcaraz (2010)
Z	10.09	Lara (2014)
X	1.87	Lara (2014)
Y	1.46	Lara (2014)
A	6	
$\propto$	0.6	Arango et al (2013)
β	0.4	Arango et al (2013)
∂_i	-0.02	Arango and Moxnes (2012)
γ_i	-0.04	Arango and Moxnes (2012)
ρ_i	0.42	Arango and Moxnes (2012)
δ_i	0.16	Arango and Moxnes (2012)
ω_i	0.54	Arango and Moxnes (2012)
A	6.0	Arango and Moxnes (2012)
B	0.1	Arango and Moxnes (2012)
V	4	Arango and Moxnes (2012)
L	16	Arango and Moxnes (2012)
ML	20	Alcaraz (2010)
C	1	Arango and Moxnes (2012)

Table A1. Values used in the simulation experiments.

Paper 2

Towards long-term economic welfare in deregulated electricity markets: Testing capacity mechanisms in an experimental setting.

This paper is published in The Electricity Journal:

Lara-Arango, D., Arango-Aramburo, S., Larsen, E.R. 2017. Towards long-term economic welfare in deregulated electricity markets:Testing capacity mechanisms in an experimental setting. The Electricity Journal, 30: 53,71

Towards long-term economic welfare in deregulated electricity markets: Testing capacity mechanisms in an experimental setting.

David Lara-Arango
System Dynamics Group
University of Bergen
Norway

Santiago Arango-Aramburo
Decision Science Group
Universidad Nacional de Colombia, Sede Medellín, Medellín, Colombia

Erik R. Larsen
Department of Management
Aarhus University
Denmark

Abstract:

Capacity adequacy has become one of the main concerns in electricity markets over the past decade, when a number of capacity mechanisms have been suggested, studied and implemented. In this paper, we focus on two of them, i.e., procurement of long-term strategic reserve contracting and centralized auctioning for capacity contracts. The first mechanism aims to stabilize price behavior by including a regulator who controls generation. The second mechanism proposes an auction system for capacity licenses to control the total generation capacity. We test and compare the desirability of these two capacity mechanisms with a "free" market using laboratory experiments. Our results suggest that the centralized auctioning for capacity contracts stabilizes laboratory markets and provides economic welfare comparable to a free market. The procurement of long-term strategic reserve contracting, on the other hand, does not seem to improve either of the two aspects in comparison to a non-regulated market.

Keywords: Electricity markets, economic welfare, stabilization, capacity mechanisms, experiments.

1. Introduction

The electricity sector in many countries has gone through significant changes in the form of deregulation and other reforms over the past three decades. Regulators and policy makers have had different concerns about processes and outcomes, prioritizing the prevention of market power and lower prices over issues such as the environment and renewables (Helm, 2007; Sioshansi and Pfaffenberger, 2006). More recently, there has been renewed concern about markets' ability to deliver enough investments and to ensure electricity's security of supply (Elia, 2015), especially since the share of renewables has increased in the national generation portfolio in many countries (e.g., OFGEM, 2013). These concerns shape the ongoing discussion of how to ensure security of supply, i.e., to ensure the uninterrupted availability of electricity at an affordable price (IEA, 2014). Despite three decades of deregulation, security of supply in power markets is not yet properly understood (Joskow, 2008). Security of supply can be seen as the long-term "economic welfare" of the market, as it aims at balancing the needs of generators and consumers (IEA, 2007; Roques, 2008; Arango and Larsen, 2011; Larsen et al. 2016).

To meet producers' and consumers' needs, there is a need for sustainable and reliable markets. To ensure this, a number of solutions have been proposed, many of which fall into the category known as capacity mechanisms. These capacity mechanisms aim to strengthen systems' long-term stability by encouraging a desired investment behavior, i.e., ensuring that enough generation capacity is available and avoiding periods of volatile price oscillations and blackouts (De Vries, 2007). Capacity mechanisms have been analyzed through actual case studies (Cramton and Stoft, 2006; CEER, 2006; Barroso et al., 2006; Cámac et al., 2006; Fignon and Pignon, 2008), simulations and experiments (Liu et al. 2010; Van der Veen, et al. 2012; Arango et al., 2013), while other studies have used theoretical models (Zou, 2009; Fu and Ren, 2011).

In this paper, we focus on two mechanisms, i.e., the procurement of long-term strategic reserve contracting and centralized auctioning for capacity contracts. We study the effects of these two mechanisms on the long-term economic welfare of an electricity market using laboratory experiments. The use of laboratory experiments is a recognized methodology for studying systems that involve decision-making processes (Cárdenas and Carpenter, 2005) and, in the particular case of electricity markets, to test different market settings (Rassenti et al, 2003; Arango et al, 2013).

We design our experiment based on the experiment presented by Arango and Moxnes (2012), using their experiment as our benchmark for the mechanisms' performance. Arango and Moxnes (2012) represent the "no intervention" case because there is no regulatory intervention. Thereafter, we create treatments for each capacity mechanism. The rest of this paper is organized as follows: The next section explains our two selected mechanisms. The third section describes the experimental design and experimental procedure. The fourth section presents our hypotheses about market behavior. The fifth section presents the experimental results, and the sixth and final section discusses our findings and concludes the paper.

2. Capacity mechanisms

We investigate two different capacity mechanisms in this paper, i.e., the procurement of long-term strategic reserve planning and centralized auctioning for capacity contracts. We select these specific mechanisms for two reasons. First, the literature indicates that they both perform well. Second, they are theoretical opposites; one of them seeks to force the market towards a desired state, and the other determines the optimal market state and then uses incentives to reach it.

With the procurement of long-term strategic reserve contracting, a government institution, i.e., a regulator, can intervene in the market with capacity. Capacity can be obtained from either a state-owned or a privately-owned generator (EC, 2003; EC, 2006). Sweden, Belgium, Germany and New Zealand implement variations of this mechanism (Fignon and Pignon, 2008; EC, 2016). Both empirical results and the theoretical literature suggest that the mechanism provides sufficient capacity adequacy in the market. However, this mechanism has been criticized in the literature for its lack of compatibility with market principles; it is seen as a direct market intervention that is not in line with a deregulated market (de Vries and Hakvoort, 2004; Meunier and Finon, 2006).

The second mechanism discussed here is the centralized auctioning of capacity licenses. With this mechanism, the government has control over—i.e., determines—the total market capacity by using auctions for licenses to build the newly required generation capacity to maintain the desired reserve margin. Thus, generators bid for licenses to build new generation capacity (Fignon and Pignon, 2008, Maurer and Barroso, 2011). The UK, New England and Italy are example cases that have implemented this mechanism (Vasquez et al., 2003; EC, 2013; 2016). The literature and empirical results suggest that high capacity adequacy and market compatibility are its main advantages and that the difficulty to ensure that licenses become actual capacity is its main disadvantage (Vasquez et al., 2003; Bidwell, 2005).

3. Experimental design

Experiments have been performed to study different issues in the electricity area. These issues include designs (Rassenti et al., 2003), consumers' willingness to pay for power (Morita and Managi, 2015), green electricity (Sundt and Rehdanz, 2015), energy efficiency in the residential sector (Ramos et al., 2015), and commercial and industrial demand response to time-of-use electricity pricing (Jessoe and Rapson, 2015), among others.

In the area of security of supply, experiments have been conducted to test different capacity mechanisms, such as mothballing (Arango et al., 2013), forward contracts (Brandts et al., 2008) and future price calibration and forecasting (Islyaev and Date, 2015). Our experiment contributes to this literature by evaluating procurement of long-term reserve contracting and centralized auctioning for capacity contracts in terms of security of supply and economic benefits.

Below, we outline the three treatments used in the experiments discussed in this paper. As mentioned above, we use the results of Arango and Moxnes (2012) as the first treatment, representing a "free market". The subsequent two treatments represent capacity mechanisms with the same market design. We assume ideal market and institutional conditions for both mechanisms aiming to have the cleanest possible experimental design for testing purposes. Each of the three treatments consists of six experimental markets, and each market has five players who are encouraged to compete to maximize their profits. The three treatments are described as follows.

3.1. Treatment 1

This treatment is the same as the one used by Arango and Moxnes (2012) in their fourth and more complex treatment. This is a symmetrical Cournot market with five players, linear demand and constant marginal costs, following Huck's standard conditions (Huck et al., 2004). In the experiment, each player represents a firm in the market. The electricity price is determined by a linear inverse demand function with a nonnegative restriction. Each year's profits and electricity price information are made available to all subjects. The capacity utilization is fixed and equal to the current capacity with a time step of one year, a four-year investment delay (i.e., from investment to the time that the new generation capacity comes online, representing the planning and building lag) and a lifetime of capacity of sixteen years. The subjects decide how much they want to invest in new capacity every year. There is a maximum size for each subject (firm) of 20 units to ensure minimal competition in the market. The experiment has 70 rounds, i.e., 70 years. The market price, P_t, in period t is

$$P_t = Max\left(6 - 0.1 * \sum_{i=1}^{5} q_{i,t}\,, 0\right) \tag{1}$$

where $q_{i,t}$ is the nonnegative production of subject i in period t, and 6 and 0.1 are scaling constants. The profit function for subject i in period t is

$$\pi_{i,t} = (p_t - C) * q_{i,t} \tag{2}$$

where C is the marginal cost, which includes operational and capital cost and is constant at 1 (monetary units/capacity units). The production of player i in period t is

$$q_{i,t} = \sum_{i=t-19}^{t-4} X_{i,t} \tag{3}$$

where $X_{i,t}$ is the investment decision of subject i in period j=t-19 to j=t-4. Total capacity or supply in time t (S_t) is

$$S_t = \sum_{i=1}^{5} q_{i,t} \tag{4}$$

3.2. Treatment 2

This treatment represents the regulators' procurement of long-term strategic reserve contracting (Fignon and Pignon, 2008). The underlying logic is the same as in Treatment 1. The only difference is that in addition to the generators, we have a regulator who can invest in new capacity in the case that demand is not met; which could also be understood as state-own firm. Since we assume ideal market conditions, the regulator has perfect information of the market. The regulators' task is to invest when production is expected to be less than demand. Thus, the regulator invests in the market to fill the gap between the capacity required to meet the demand and the actual capacity. The regulator cannot have more than 10 units (20% of the market optimum) to ensure that he does not have the power to dominate the market. In this way, we preserve the characteristics of a competitive market.

The conditional function that determines the regulator actions is

$$Z_t = if\ Q_{t+4} < 50 \begin{cases} 10 & if\ 50 - Q_{t+4} \geq 10 \\ 50 - Q_{t+4} & if\ 50 - Q_{t+4} < 10 \end{cases} \tag{5}$$

where Q_{t+4} is the capacity in four years. The regulator aims for the market optimum (50 units), which is the competitive equilibrium, i.e., the maximum economic welfare.

3.3. Treatment 3

The third treatment represents the case of a centralized auction for capacity contracts (Alcaraz, 2010). The market is as described as in Treatments 1 and 2. In this case, the model calculates the amount of capacity to be auctioned, ensuring that the electricity price is equal to the marginal cost (perfect competition equilibrium), also known as the optimal total investment (OTI; Reichmann, 2007). While in reality both capacity and electricity prices fluctuate, and generators bid for capacity based on anticipated revenues from electricity sales, we assume the electricity price to be equal to the marginal cost and design a subsidies scheme for capacity. This subsidies scheme is designed to reflect different situations in the electricity market without explicitly including it. For instance, high subsidies for capacity reflect a shortage of generation relative to demand, and low or negative subsidies (costs to build capacity) reflect an excess in generation relative to demand. We include this scheme to have more control over our experiment's input and simplify our experimental design for the participants, by providing them one market to deal with instead of two.

Players bid for capacity licenses every year. They make a bid curve in which they write how many licenses they want at each possible license price. The bid curves are aggregated and matched with the OTI. Then, the equilibrium license price LP is found. Based on the LP licenses and the individual bid curves, licenses are given to the individual generators. To ensure that the equilibrium can be found, and to avoid logical errors, the model accepts only monotonically decreasing curves for desired investments. Players bid every year by drawing investment curves in a desired number of licenses vs the license price chart. Negative license prices denote that players require subsidies to make additional investments. With the purpose of having more control, we have assumed ideal institutional conditions, and thus all licenses become actual capacity after their allocation in the auction i.e. players are obliged to build the capacity they have gotten licenses for.

The cost function of this license or subsidy for subject i in period t is

$$B_{i,t} = X_{i,t} * LP_t \tag{6}$$

where $X_{i,t}$ is the licenses assigned to subject i in period t, and LP_t is the license price. The profit function in monetary units for subject i in period t is

$$\pi_{i,t} = (P_t - \alpha) * q_{i,t} - B_{i,t} \tag{7}$$

where α is the operational cost constant equal to 1 monetary unit per unit of capacity, and $B_{i,t}$ is the license cost or cost function of the mandatory right to build capacity. Since the capacity is fixed, the electricity price will remain constant throughout the whole treatment and will be equal to the cost; therefore, the profits can only come from the trading of licenses.

3.4 Economic equilibrium for the three treatments

Each experimental treatment presents a different scenario for competition and market behavior. To be able to benchmark the behavior, we calculate the economic equilibria in the cases of a joint maximization, Cournot-Nash and perfect competition (see equilibria calculations in Appendix A). Table 1 shows the results for the three equilibria in the three different treatments.

Table 1: Players' individual investments (ID), market capacity (Cap), price for joint maximization (JM), Cournot-Nash (CN) and perfect competition (PC) equilibria.

	T1			T2			T3		
	ID	*Cap*	*Price*	*ID*	*Cap*	*Price*	*ID*	*Cap*	*Price*
JM	0.31	25.0	3.50	0.25*	30.0*	3.00	NA	NA	NA
CN	0.52	41.7	1.83	0.42*	43.3*	1.67	NA	NA	NA
PC	0.63	50.0	1.00	0.50-0.63*	50.0*	1.00	0.63	50.0	1.00

*In this treatment, a regulator firm intervenes, so the individual investments of the players do not match with the market production.

As observed in Table 1, the optimal production for maximizing the economic welfare, i.e., the competition equilibrium, is 50 units given the calibration we have presented in the equations above, and this production amount becomes the one we use in the following.

3.5. Experimental procedure

For treatment 1 (T1), we use the results of Arango and Moxnes (2012). We do not repeat these experiments because their main purpose for use is to act as a benchmark for the two other treatments. Furthermore, we follow the same format and use the same population, which makes their results directly comparable with ours.

For treatments 2 and 3, we run six experimental markets for each treatment. We follow the standard protocol for experimental economics (Friedman and Sunder, 1994; Friedman and Cassar, 2004). All participants were fourth- and fifth-year students of management and industrial engineering at the Universidad Nacional de Colombia in Medellín, Colombia. The participants could earn a payoff between COP $10.000 and COP $40.000 (at the time, equaling between US $5 and US $20) in one and a half hours, depending on their performance, which was measured as accumulated profits in the experiment. The payoff exceeded the opportunity cost estimated for the students for the experiment's duration time.

Participants were organized randomly at the workstations used for the experiment so that they were not able to identify their competitors in the market. The instructions were then distributed among participants (see Appendix B for the experiment instructions). After 15 minutes, we gave participants an additional explanation of the market dynamics and posed general questions to make sure they understood their task. We asked participants to make their decisions and write them down on a record sheet we gave them as a backup for the experiment. We started the experiments under the same initial conditions of Arango and Moxnes (2012), that is, a market capacity of 55, which translates into a price of 0.5 and negative profit of -0.5 as the starting point. They were told that the software would not allow them to have a capacity over 20 units to ensure minimum competition. It was also explained how the interface worked and what information they would receive every year. They were informed by the software about their own production capacity each year (discriminated in years of remaining lifetime), yearly profits, accumulated profits, market capacity and market price. We ran the experiments using a computer network with Powersim Constructor 2.51 (see Appendix C for the experiment interface). The software ran year by year automatically once all participants made their decision and kept records of all variables, including participants' decisions. The experiment is available from the authors upon request.

4. Dynamic Hypotheses

We present our hypotheses for Treatments 2 and 3 based on the simulations performed by Lara (2014) for treatments 2 and 3. For Treatment 1, we show the simulations reported by Arango and Moxnes (2012).

4.1. Treatment 1

In this treatment, we use the heuristic and parameters estimated by Arango and Moxnes (2012) in their fourth treatment. This heuristic is based on Nerlove's (1958) adaptive expectations approach, which states that players create price expectations based on previous forecasts they have made and the current market price. The expected price at time t, ExP_t, is given by

$$ExP_t = (1 - \beta_1) * ExP_{t-1} + \beta_1 * P_t + \beta_0 \qquad (8)$$

where β_1 is the weight on past expected prices, and β_0 is a scaling constant. We use the values estimated by Arango and Moxness (2012), i.e., β_1=0.31 and β_0=0.02.

Given the expected price, we can estimate the total market investment (X_t), which is determined by

$$X_t = Max(\partial * Dep_{t+4} + \gamma * Cap_{t+4} + \rho * ExP_t + \delta * P_t + \omega, 0) \quad (9)$$

where x_t is the market's investments in capacity in period t, and $\partial, \gamma, \varepsilon$, μ and ω are market decision parameters for depreciation, capacity, expected price, actual market price and decision adjustment, respectively. Dep_{t+4} and Cap_{t+4}, are market depreciation and installed capacity four years ahead of period t, ExP_t is the market's expected price in period t, and P_t is the market's actual price in period t. The base case simulation, using the estimated parameters from Arango and Moxness (2012), is shown in Figure 1.

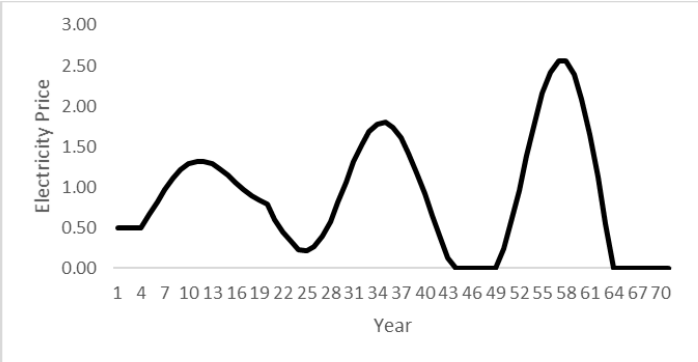


Figure 1: Simulated prices of the base case (Arango and Moxnes, 2012)

As Fig. 1 shows, the parameters give rise to cyclical behavior. We use this as the benchmark of the comparison with the two intervention mechanisms we have selected because this represents a "free market" in the sense that there is no direct regulatory intervention.

4.2. Treatment 2

Based on the same parameters used in Treatment 1, Lara (2014) performs a number of simulations with the model, including regulator's investment in long-term strategic capacity. Regulator capacity is assumed to be constant the entire time period and equal to 10 units, which gives the regulator significant intervention power (20% of the capacity goal) but not enough power to completely control the system's capacity or the price (Meunier and Pignon, 2006). Representative results from these simulations are shown in Figure 2.

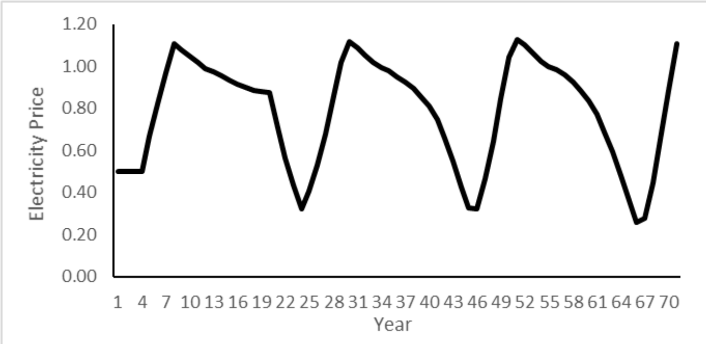


Figure 2: Simulated price with regulator (Lara, 2014)

Figure 2 clearly shows that although the regulator does not eliminate cyclical behavior, the amplitude of the cycles is significantly lower than in Figure 1. Regulators' intervention contributes to reducing the magnitude of the oscillations. This behavior is, in that respect, more preferable than the free market in Treatment 1. We can conclude that it improves the system, although the calculation of welfare later will raise some doubts about this.

4.3. Treatment 3

In this treatment, we examine the license prices for permissions to build new capacity, given that the electricity price is constant. To simulate market bidding for this treatment, we use the heuristic proposed by Lara (2014), in which the bidding curve's initial point is at the lowest possible price for the licenses (-2.1) and is determined by the market's profit and capacity. Thus, this initial point is given by

$$DCap_{-2.1} = Min(Max(z + x * MProf + y * Cap, 0), 20) \tag{10}$$

From this initial point, we can estimate the following points on the bidding curve, which are given by

$$DCap_{Lp} = Max\left(i + l * DL_{Lp-m}, 0\right) \tag{11}$$

where $DCap_{Lp}$ is the desired capacity at a license price Lp, DL_{Lp-m} is the desired capacity at the previous point on the curve (m is the distance between the points on the curve), and i and l are coefficients estimated for each particular point (Lp, $DCap_{Lp}$).

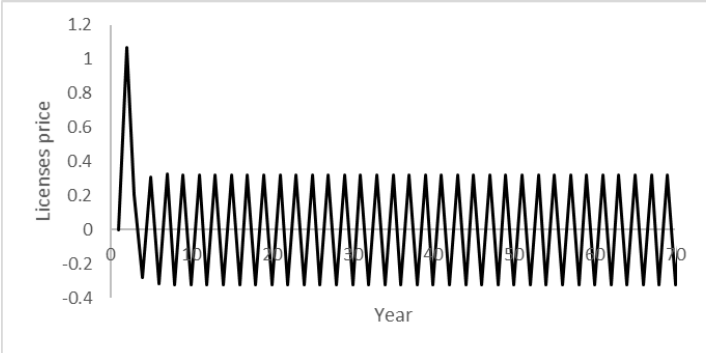


Figure 3: Simulated licenses prices

With the auctioning system, the market capacity is fixed; therefore, the electricity price is stable. However, market instability is present in the license prices, as Figure 3 shows, because the producers' investments and profits are determined by the license prices. If the license price is sufficiently negative, i.e., generators are paid to build capacity, there might be a reduction in the overall economic surplus. This auction intervention fosters decision making similar to what we observed in the previous cases, i.e., the market makes high orders for capacity when the price is attractive and low orders for capacity when the price is not attractive. Thus, this decision-making approach may lead to oscillations in the license prices (Lara, 2014). However, it is worth noting that the uncertainty is less for the generators in this treatment because the future price is known, and the cost or subsidy is known before the investment is made, leading to relatively small corrections in the license price.

4.4 Welfare considerations

To obtain more insight about the advantages and disadvantages of the different treatments, we calculate the economic welfare in the three markets. We calculate the consumer, producer and economic surpluses for every year; thus, we can see how the economic surplus is affected in terms of both expected value and stability. For Treatments 1 and 2, we use the following equations:

$$Consumer\ surplus = \frac{(Maximum\ Price - Market\ Price) * Q_t}{2} \tag{12}$$

$$Producer\ Surplus = (Market\ Price - Producer\ Cost) * Q_t \tag{13}$$

$$Economic\ Surplus = Consumer\ Surplus + Producer\ Surplus \tag{14}$$

For Treatment 2, we take the regulator firm results into account, i.e., the profits or losses reported by this firm are added to (or subtracted from) the economic surplus. For treatment 3, we include the cost for the government. Therefore, the economic surplus for treatment 3 is the sum of the consumers' and producers' surplus minus the cost for government. The cost for the government is calculated as follows:

$$Cost\ for\ gov = \begin{cases} 2 * L_p * X_t & If\ prices\ are\ subsidized \\ 0 & If\ prices\ are\ not\ subsidized \end{cases} \tag{15}$$

where X_t is the total assigned licenses in period t, and LP_t is the license price, as before. When there are subsidized prices, the government incurs a double cost, i.e., the direct cost of the subsidies and the opportunity cost of the public resources going to subsidies when they could be used to meet other public needs. We test economic surplus using the average value and the standard deviation to test for both expected value and stability. Table 2 shows the results for the simulations discussed previously.

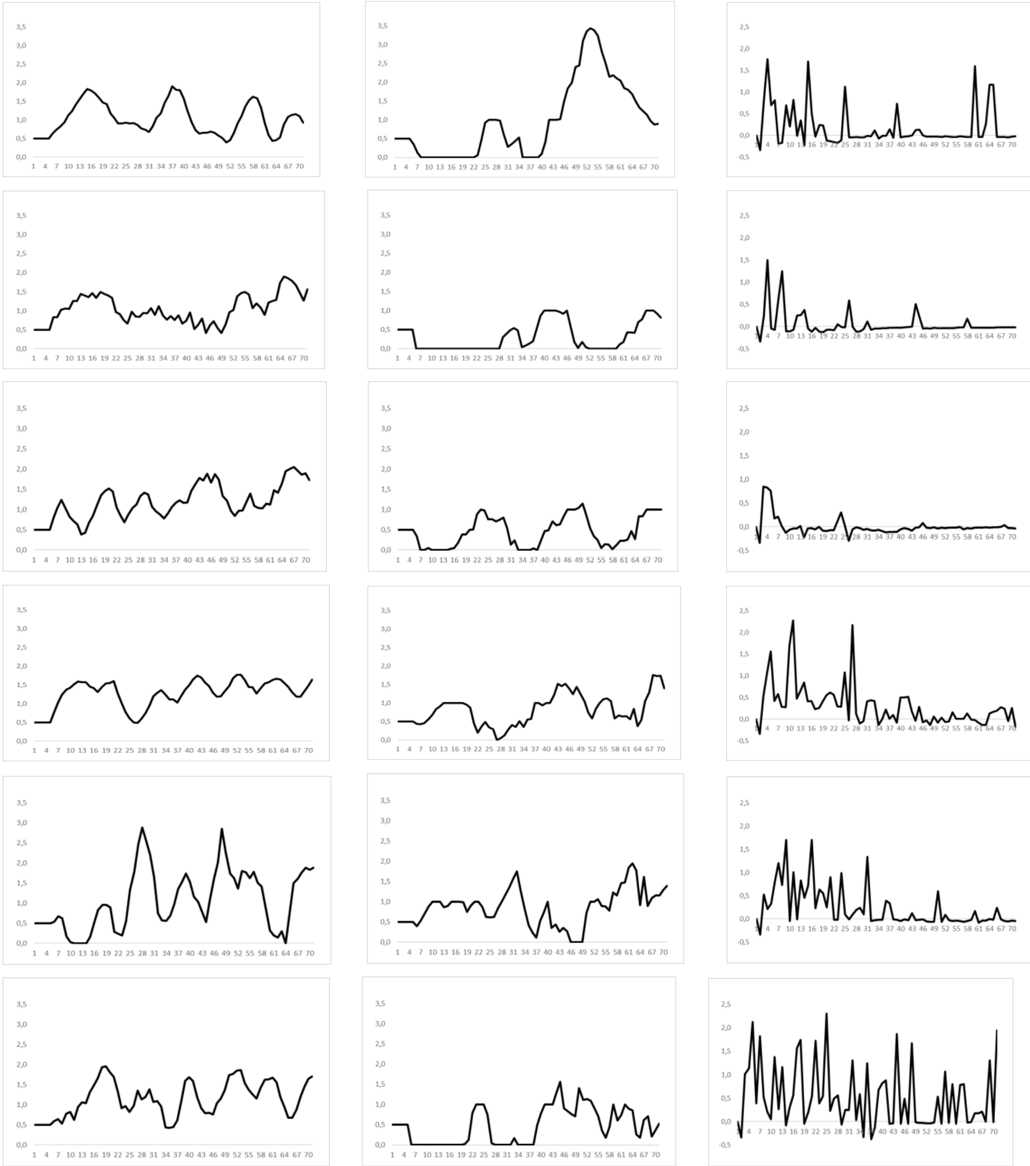
Table 2: Consumer, producer and economic surplus results for the simulations (Lara, 2014)

	Consumer surplus	Producer surplus	Economic surplus
Base case			
Average	133.59	-12.46	121.13
Std. Dev.	35.43	37.12	5.09
Regulator			
Average	136.31	-11.86	124.45
Std. Dev.	13.33	14.05	0.76
Auctioning			
Average	125.00	-0.06	123.98
Std. Dev.	0.00	1.07	0.31
Optimum	125.00	0.00	125.00

As Table 2 shows, Treatment 1 presents a negative value for the average producer surplus and a high value for the consumer surplus (higher than the optimum and out of the optimum equilibrium). Despite the high value for consumer surplus, the economic surplus is less than the optimum due to the negative producer surplus. For Treatment 2, the consumer surplus value exhibits a value greater than the optimum, along with a negative producer surplus. This combination results in an economic surplus that is less than the optimum but greater than the base case. Since licenses are the only factor determining market profits in Treatment 3, market investors may be prompt to make a decision similar to the ones we observed in the other treatments, i.e., investors make high orders for capacity when the price is attractive and low orders for capacity when the price is not attractive. In comparison, Treatment 2 presents the best average value and the second-lowest standard deviation for the economic surplus of all three cases. However, these results raise some questions about the system's sustainability because this treatment shows sustained economic losses (on average) for the producers, which would be unbearable for the producers in the long term (Arango and Larsen, 2011). Therefore, we consider that Treatment 3 presents the best results in terms of sustainability, average value and standard deviation for economic surplus (Lara, 2014). We now show our experimental results to see whether they are consistent with these simulations.

5. Experimental Results

Figure 4 shows the price develoment over the periods for the different treatments. It illustrates our treatments 2 and 3 along with the results of Arango and Moxnes (2012) in T1, which is used as our base case. Visual inspections show cyclical tendencies in T1 with different amplitudes and frequencies. We can also observe that there seems to be an almost constant adjustment of the price in all six markets, as the price tends to regress to its mean. For T2, with the procurement of long-term reserves, we observe a change in the dynamics compared to T1, with a higher anplitude and lower frequencies on average, i.e., more pronounced cycles as well as periods with a stable price in a number of the experimental markets. From T3, we observe a different pattern compared to the first two treatments, given that the electricicty price is constant in this treatment, and we observe the license price, with relatively short cycles intercepted by periods of no change. In most markets, we also observe a significant change over time as the amplitude seems to become smaller and in some cases almost dissapear. T2 presents a higher likelihood for prices equal to zero. This can be explaind by the presence of the regulator, in the sense that having an extra player investing in the market (the regulator) can have the effect of periods with more excess capacity and thus a price that drops to zero.



Electricity price T1 ($/Unit) Electricity price T2 ($/Unit) Licenses price T3 ($/License)

Figure. 4: Observed market prices in different markets for T1 and T2 and license prices in T3

Visual inspection is only the first step in understanding the results. We are interested in how the behavior and particularly the oscillations we have observed affect the economic welfare of the market. This analysis allows us to understand the cost and benefits of the different types of interventions in the market. We calculate the consumer, producer and economic surplus for all markets for each period in the experiment to be able to make these comparisons. We start by looking at the aggregated results across treatments and then discuss the disaggregated behavior. Table 3 summarizes the aggregated results across the three treatments.

In Table 3, T3 shows the largest expected value and the second-lowest standard deviation, which makes this mechanism the best option for the market. Regarding the economic surplus, i.e., the aggregated performance, we observe that there is less than a 5 percent difference between the best case (124.44 in T3) and the worst case (118.71 in T2). Moreover, the no-intervention scenario (T1) is close to the results of T3, with 123.70. However, the disaggregated results show larger differences, where the results of T2 are different from the other two. The consumers do better in both intervention cases compared to the market base case; however, they do better in T2 than in T3, where the consumer surplus increases by almost fifteen percent, from 125.00 to 142.76. The opposite is the case for the producer surplus, which is lower in T2 (-23.75) than in T3 (-0.67). However, T2 has the largest producer surplus with 3.60.

One question that we need to solve is whether the treatments' economic surpluses are significantly different from each other. To assess this, we perform a t-test to compare them. Table 3 shows that all economic surplus values are statistically different from each other in terms of both expected value and standard deviation (t-tests at 5% significance).

Table 3: Consumer, producer and economic surplus aggregated results of the experiments. j = 1, 2 and 3, different from Tj (T-test). All tests at 5% level. N=18 (each market is one observation)

	Consumer surplus	Producer surplus	Economic surplus
Base case (T1)			
Average	120.10	3.60	123.70[2,3]
Std. Dev.	13.16	12.96	0.66[2,3]
Regulator (T2)			
Average	142.76	-23.75	118.71[1,3]
Std. Dev.	30.02	29.21	13.18[1,3]
Auctioning (T3)			
Average	125.00	-0.67	124.44[1,2]
Std. Dev.	0.00	1.49	0.90[1,2]
Optimum	125.00	0.00	125.00

We observe that there are significant differences in the standard deviation in Table 3. T2 has a large standard deviation, which raises questions about the desirability of this option, where the variation is significantly higher than in the other two treatments. This implies that there is more uncertainty in the prices, which is consistent with our visual inspection of Figure 4. Thus, while consumers on average will be better off, there might be periods where the prices will be very high. In many cases, this is not desirable, as large price differences might make the market look unstable and discourage both new investment in generation (slowly increasing the role of the regulator) and general investments in manufacturing if there is uncertainty about the future electricity prices. In T3, there is no variation in the price because it is determined by the regulator and the market-based solution, although there is uncertainty in this treatment, such uncertainty is lower than in T2.

The economic surplus across treatments is, although different, relatively similar; there is only a 5% difference between the best and the worst. On the other hand, one could argue that a four- to five-percent difference at the national level is a significant amount of money. At a first glance, we might argue that the case of procurement of long-term reserves (T2) should be the preferred option, given that it creates the most value for consumers, i.e., it is the case where the consumers are best off. However, this is also the scenario where the producers are worst off, and the consequence might be that producers' activity will not be sustainable in the long term, which will require the regulator to become increasingly involved. Although the auction-based system presents a negative producer surplus, it is very small and should as such not endanger the long-term survival of the generators.

We now turn our attention to the disaggregated results in Table 4, and we look more closely at the differences between markets within a treatment, i.e., we explore the reasons for the different standard variation in Table 3. In Treatment 1, the free market condition, we observe that four out of the six markets have a positive producer surplus and in some cases, such as T1-M4, have a relatively high surplus, which corresponds to relatively high prices and smaller volatility, as observed in Figure 4. We observe little variation (2% difference between best and worst) in the economic surplus across the different markets in T1; therefore, the tradeoff is between who captures the surplus, the producer or the consumer. The largest economic surplus, although the difference is relatively little, is in the first two markets, T1-M1 and T1-M2. These two cases represent the two markets where the producer surplus is closest to zero.

For T2, the procurement of capacity by the regulator, we observe that producers have a negative surplus in all markets, and in all cases, it is significantly higher than the greatest loss in any of the markets in T1. The variation between the markets is also significantly higher, as shown by the standard deviation in Table 3. The reason for the negative surplus can partly be explained by the period of low prices in T2 (see Figure 4). This also explains the markets where there is a relatively high consumer surplus, e.g., T2-M3 and T2-M6. Looking at the consumer surpluses, we can observe that even the smallest, T2-M3, is larger than the largest in T1, T1-M5. The economic surplus is very low only in the case of T2-M2.

T3 shows relatively little variation across the markets compared to the two other treatments. The reason for this is that the electricity price is stabilized with the capacity auctions. The producer surpluses, though negative, are close to zero in all markets.

As an overall impression, both mechanisms seem to induce a loss for regulators to bear in order to improve the economic surplus of the market. However, the procurement of strategic reserve planning puts more stress on the producers' surplus, arguably because it reduces the generators' ability to increase their profits by reducing their investments to raise the price. This stress on the producers' surplus ultimately leads to both reducing and destabilizing the economic surplus.

Table 4 Producer, consumer and economic surplus for the disaggregated markets in the three treatments

\		Producer surplus		Consumer surplus		Economic surplus	
		Average	*Std. Dev.*	*Average*	*Std. Dev.*	*Average*	*Std. Dev.*
T1	M1	-0.38	20.28	124.47	20.72	124.10	0.96
	M2	0.79	18.05	123.51	18.36	124.29	0.84
	M3	7.01	19.63	116.91	20.51	123.91	1.40
	M4	11.63	17.52	112.33	17.86	123.96	0.76
	M5	-2.49	34.85	124.51	35.82	122.02	3.58
	M6	5.07	20.54	118.87	21.16	123.94	1.12
T2	M1	-16.41	39.03	131.54	43.69	117.92	7.81
	M2	-43.94	27.09	147.70	27.66	102.70	25.43
	M3	-30.26	21.53	154.44	20.26	122.52	2.53
	M4	-11.49	19.40	135.77	21.38	123.95	1.15
	M5	-9.23	21.19	131.87	23.17	123.81	1.76
	M6	-31.21	26.44	154.33	24.99	121.35	4.00
T3	M1	-0.55	1.44	124.75	0.38	124.20	1.31
	M2	-0.15	0.86	124.75	0.31	124.61	0.78
	M3	-0.01	0.60	124.71	0.39	124.40	0.85
	M4	-0.94	1.54	124.84	0.35	124.50	0.74
	M5	-0.71	1.35	124.86	0.28	124.45	0.78
	M6	-1.64	2.10	124.83	0.46	124.48	0.76

We focus now on the analysis of potential cycles in the experiment markets, i.e., cyclical tendencies in the electricity prices in T1 and T2 and the license price in T3. We perform an autospectra and autocorrelogram analysis of the three treatments (see Appendix D). From the autospectra, we identify frequency concentrations in all T1 markets. These frequency concentrations imply price cycles from approximately 13 to 40 years in this set of markets. The autocorrelograms show that for the markets in T1, there are significant autocorrelations for at least the first three lags in all groups. With the same analysis for T2, the autospectra show significant frequency concentrations in all groups, which suggests cycle lengths ranging from approximately 6 to 40 years. For T2, the autocorrelograms show significant correlations for at least the first four lags in all groups. For the case of license prices in T3, the autospectra analysis shows significant frequency concentrations in groups 4, 5 and 6. The cycle lengths in these groups range from approximately 5 to 20 years, i.e., significantly shorter in the two previous treatments. The autocorrelograms for T3 show more variation across the groups than was the case in the previous treatments. Here, groups 1 and 6 have no significant autocorrelation, groups 2 and 3 present significant autocorrelations in only one lag, and groups 4 and 5 show significant autocorrelations in 4 and 5 nonconsecutive lags, respectively.

There are different reasons for and implications of the difference between T1 and T2 versus T3. Although the first two treatments present electricity prices (T1 and T2) and the last (T3) presents license prices, we expect the same dynamic evolution because it is driven by similar decisions around investment in capacity. However, we observe significant differences between the outcomes of T1 and T2, which might be useful to discuss in more detail. The bidding mechanism in T3 forces players to compete for a fixed capacity; that is, when one player takes a portion of the fixed capacity, the maximum capacity the other players can obtain is automatically restricted, which is not the case in T1 and T2. To understand how the subjects, make decisions, we assess subjects' decision rules using regressions, with *Investment* as a dependent variable and *capacity*, *depreciation*, *expected price* and *price* as independent variables. Table 5 shows the results of the regressions' average coefficients per market and per treatment. There is a significant difference across treatments in terms of the percentage of significant regressions. The coefficients do not seem to differ much between treatments.

Table 5 Average coefficients of the regressions (different than zero) for T1 (results from Arango and Moxnes (2012)) and T2. a_d is the coefficient for *depreciation*, a_c is the coefficient for *capacity*, a_e is the coefficient for *expected price*, a_p is the coefficient for *price*, and a_0 is the constant term.

Treatment	a_d	a_c	a_e	a_p	a_0	R^2	P-Value< %5
T1	**-0.02**	**-0.04**	**0.42**	**0.16**	**0.54**	**0.36**	**87%**
M1	0.27	0.13	-0.24	0.59	-15.99	0.48	100%
M2	-0.65	-0.63	0.47	-1.00	7.88	0.48	80%
M3	-0.02	0.21	-0.06	-0.71	1.27	0.09	80%
M4	-0.43	0.36	0.56	0.26	15.26	0.45	100%
M5	0.50	0.02	0.93	1.09	-6.11	0.09	80%
M6	0.19	-0.32	0.89	0.72	0.95	0.54	80%
T2	**0.04**	**-0.07**	**0.31**	**0.05**	**1.03**	**0.15**	**58%**
M1	0.03	0.10	0.48	0.05	0.98	0.13	20%
M2	-0.09	-0.13	0.30	-0.75	1.44	0.11	60%
M3	0.18	-0.11	0.28	0.10	1.06	0.22	80%
M4	-0.19	-0.29	0.22	0.04	0.98	0.11	80%
M5	-0.03	0.79	0.53	0.18	0.33	0.17	20%
M6	0.08	-0.18	0.44	0.31	1.37	0.18	40%

Expected price and price are reinforcing variables (positive coefficients), meaning that a price that is greater than zero will always generate more investment, especially when the investment has a positive expected return. This will eventually put pressure on the price and drive it down. Capacity shows a negative overall effect on investments, meaning that the more the capacity increases, the more reluctant the players will be to invest. Depreciation shows a seemingly ambiguous effect. The subjects in T1 have a negative effect on investments, which suggests that players in T1 tend to prefer a lower capacity, i.e., they do not necessary replace all retired capacity. This leads to an increase in the price as the total capacity decreases. However, subjects from T2 seem to have the opposite reaction: They invest more as the depreciation increases to either maintain or increase their current capacity. This behavior argues for a higher degree of competition in T2 than in T1, leading to lower prices.

There is a reduction in the percentages of significance when T1 is compared to T2. This reduction in percentages can be explained by the fact that T2 markets present a higher random component in their decisions. This notion makes sense if we conceive T2 as T1 with an added actor, i.e., the regulator firm. The introduction of this additional firm (with full foresight) seems to make the market less predictable. In terms of the coefficients, there is no distinguishable change from the coefficients in T1 to T2 in general. However, it is interesting to note that the price expectations in T2 are seemingly less widespread than the expectations in T1. Although it is not conclusive, this finding raises the question of whether the introduction of the regulatory firm can somehow lead to more predictable price expectations and thus more predictable (at least marginally) investments. Therefore, further research is required.

Based on the average coefficients portrayed in Table 5, we can run simulations using these and compare them with our original simulations. Since we use Arango and Moxnes' (2012) results for T1, we will make such a comparison only for T2 and T3. Fig. 5 shows our original simulations from Figure 2 along with the new simulations using the average coefficients displayed above for T2. Fig. 5. shows a significant resemblance between our original simulation and the new simulation based on the experiment. We observe a similar frequency with the maximum for the simulation being the same and the minimum being higher compared to the first simulation in Figure 2. The simulation with the new decision rules shows a smoother price transition than our original simulation. This can be attributed to the smaller coefficients for expected price and price, the more negative coefficient for capacity and the positive coefficient for depreciation. The smaller coefficients in expected price and price suggest that players are slightly more conservative when observing increases or decreases in price. The more negative capacity coefficient accounts for a more constrained investment policy, in which a higher capacity implies higher reductions in investments than in the original case. The positive depreciation coefficient implies that depreciated capacity is compensated to some extent rather than indicating a decreasing tendency (investment going down as depreciation goes up), as in the previous case. These three elements suggest a less aggressive investment policy, which accounts for the smoother prices we find in Figure 2.

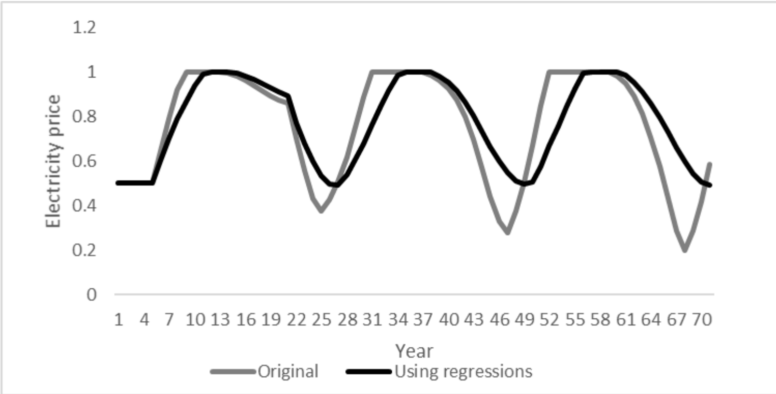


Figure 5: Original simulation vs new simulation using the average estimated coefficients from Table 4 for T2

To perform the same comparison between the original and new investment functions for T3, we run a number of regressions of T3 data following our original investment function (eq. 9) formulation, that is, using capacity and accumulated profits as independent variables and investment as the dependent variable. Table 6 shows the average coefficients per market and the average for the treatment.

Table 6: Average coefficients of the regressions (different than zero) for T3. MProf is the coefficient for *profits*, a_c is the coefficient for *capacity,* and a_0 is the constant term.

Treatment	MProf	a_c	a_0	R^2	P-Value< %5
T3	**1.80**	**1.63**	**-12.19**	**0,82**	**90%**
M1	0.73	2.29	-15.52	0.76	80%
M2	2.41	2.42	-6.53	0.80	100%
M3	1.73	1.14	-38.21	0.82	100%
M4	2.17	0.78	-3.92	0.83	80%
M5	1.46	1.54	-2.67	0.85	100%
M6	1.83	1.67	-6.31	0.86	80%

It is interesting to note that 90% of the regressions are statistically significant and that the R^2 value of 0.82 is relatively high. The coefficients are also close to the simulation coefficients, as in T2. Market-to-market comparisons show a rather narrow distribution of coefficients for profits and capacity. Similar to observations in Table 5 for T2, it also raises the question of whether the introduction of capacity mechanisms (auctions, in this case) may lead to more predictable investments. Such investments may not be the optimal ones, but they might be more predictable; nevertheless, this question is beyond this paper scope, and it may be addressed by future research. Based on the coefficients displayed in Table 6, we perform a comparison between the original investment function and a new function with these coefficients. As Fig. 6 shows, there is a significant resemblance between the original and the new investment functions.

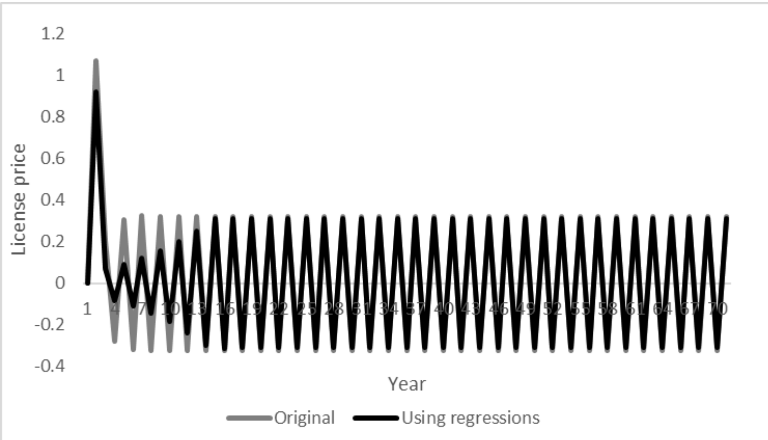


Figure 6: Original simulation vs new simulation using the average estimated coefficients from Table 6

There is a significant resemblance between the original and the new investment functions, as shown in Fig. 6. These results show consistency between our original simulation and the simulations that result from using the experiment-inferred decision rules. Such consistency suggests that our results for T2 and T3 are in line with the results that were used to build the original simulation parameters i.e., our experiment validates the robustness of the experiments on which the original simulations are based. We now turn to the last section, which discusses our findings and concludes the paper.

6. Conclusion and Discussion

Capacity mechanisms aim to improve the long-term security of supply or specifically one aspect of security of supply, i.e., capacity adequacy. The regulator intervenes in the system by activating a mechanism to increase capacity when it is considered necessary (Fignon and Pignon, 2008). In this paper, we perform an experimental study to test two different capacity mechanisms, i.e., the procurement of long-term strategic reserve contracting and capacity auctions.

We investigate whether either of the two interventions in the market significantly improves the capacity adequacy of the market and thereby not only provides insurance against blackouts but also provides affordable prices for consumers, which is another important dimension of security of supply. We do this in a number of different ways. We look at the economic welfare of both producers and consumers as well as overall economic performance, and we look at the autospectra and autocorrelograms to establish the presence of cycles. We estimate the general investment function in the market with a simple equation. We also look at the stability of the market, as cycles imply that there will be high uncertainty regarding investments, i.e., firms will generally be cautious in investing if there is great uncertainty regarding the future return (Arango and Larsen, 2011).

Our experimental results suggest that the procurement of long-term strategic reserve planning neither improves the market's expected benefits nor benefits stability in the market. Moreover, the experimental results suggest that generators exhibit substantial losses on average across the markets and in each single market. Furthermore, looking at the behavior of the price over time, the potential problems for the generators become even more significant if such losses are sustained. As observed in Figure 4 and confirmed by the autocorrelograms, there are large and long-term fluctuations in the price, including a number of cases with a very low price over long periods. We observe larger amplitudes of the fluctuations compared to the base case, T1, which increases uncertainty for investments. This is also confirmed by the regressions (see Table 5), where it is shown that this type of intervention is unlikely to bring more stability to the market and, in most cases, might be worse than leaving the market "alone".

The use of centralized auctioning for capacity contracts improves the experimental markets' performance. The overall economic welfare as well as the consumer and producer welfare is higher than the second treatment, T2, and only slightly worse for the producers compared to the first treatment, T1. In addition, we find that this mechanism presents a more sustainable scenario, given that the expected value and standard deviation for the producers are closer to the normal economic profit than in the first mechanism's case. License price is characterized by random patterns with no autocorrelation, which indicates no evidence of cyclical behavior. Furthermore, since the regulator sets the electricity price, the uncertainty that the generators experience is much less than in the previous cases. This is likely to lead to more timely investments. However, as in the other intervention cases, this is far from a "free" market and might to a large extent depend on how appropriately the regulator manages the price, ensuring a return for generators and an affordable price for consumers.

To summarize, our experimental results suggest that centralized auctioning for capacity contracts can have positive results in electricity markets by improving their economic welfare in terms of both stability and expected economic surplus. Conversely, our results also suggest that the procurement of long-term strategic reserve planning can actually have detrimental effects in the market by reducing the expected economic surplus and increasing the surplus' instability. These results are consistent with the fact that market-oriented mechanisms tend to be more preferable than interventionist mechanisms (Meunier and Finon, 2006).

The selection and implementation of capacity mechanisms in real life are difficult tasks because we cannot know for sure what the real outcome is going to be (Larsen and Arango, 2013). However, simulation models and laboratory experiments provide insights about what the effects of different capacity mechanisms might be in real markets; therefore, they can offer a better basis to select and implement such mechanisms.

References:

Alcaraz, S. 2010. Free Competition vs. Auctioning of Licenses for Electricity Production. M.Phil Thesis. System Dynamics Group, Department of Geography, University of Bergen.

Arango, S., Castañeda, J.A., Larsen, E.R. 2013. Mothballing in power markets: an experimental study. From regulation to competition. *Energy economics* 36, 125-134.

Arango, S., Moxnes, E. 2012. «Commodity cycles, a function of market complexity? Extending the cobweb experiment». *Journal of Economic Behavior & Organization* 84 (1) (septiembre): 321-334. doi:10.1016/j.jebo.2012.04.002.

Arango, S., Larsen, E. 2011. Cycles in deregulated electricity markets: Empirical evidence from two decades. *Energy Policy* 39 (5) (mayo): 2457-2466. doi:10.1016/j.enpol.2011.02.010.

Barroso, L.A., Bezerra, B., Rosenblatt, J., Guimar~aes, A., Pereira, M.V., 2006. Auctions of Contracts and Energy Call Options to Ensure Supply Adequacy in the Second Stage of the Brazilian Power Sector Reform. IEEE PES General Meeting 2006, Montreal, Canada.

Bidwell, M. 2005. Reliability options. A market-oriented approach to long-term adequacy. The electricity journal 18 (5), 11e25.

Brandts, J., Pezanis-Christou, P., Schram, A. 2008. Competition with forward contracts: A laboratory analysis motivated by electricity market design. Economic Journal, 118(525), 192-214. Retrieved from <Go to ISI>://WOS:000251859800009

Cámac, D., Ormeño, V., Espinoza, L., 18-22 June 2006. Assuring the efficient development of electricity generation in Peru. Power Engineering Society General Meeting, 2006. IEEE, p. 5.

Cárdenas, J.C., Carpenter, J.P., 2005. Experiments and Economic Development: Lessons from field labs in the developing world. middlebury college economics discussion paper no. 05-05

CEER (Council of European Energy Regulators), June 2006. Survey of Capacity Support Mechanisms in the Energy Community. Ref: C06-SEE-15-06, 6.

Cramton, P., Stoft, S., 2006. The Convergence of Market Designs for Adequate Generating Capacity, with Special Attention to the CAISO's Resource Adequacy Problem, White Paper for the Electricity Oversight Board, March.

De Vries, L.J., 2007. Generation adequacy: Helping the market do its job. *Utilities Policy* 15 (1), 20e35

De vries, l.J., Hakvoort, R.A. 2004. The question of generation adequacy in liberalized electricity markets. Working papers no. 120.2004, fondazione eni enrico mattei.

EC, 2003. Proposition de directive du Parlement européen et du Conseil concernant des mesures visant à garantir la sécurité de l'approvisionnement en électricité et les investissements dans les infrastructures {SEC(2003) 1368}/* COM/2003/0740 final - COD 2003/0301 */

EC, 2006. Directive 2005/89./EC of the European Parliament and the Council concerning measures to safeguard security of electricity supply and infrastructure investment, Brussels, 18 January 2006

EC, 2013. Capacity Mechanisms in Individual Markets within the IEM. DG ENER - DIRECTORATE B

EC, 2016. Report from the Commission: Interim Report of the Sector Inquiry on Capacity Mechanisms

Elia, 2015. Etude relative à la sécurité d'approvosionnement pour la Belgique. Besoin de réserve stratégique pour l'hiver 2016-17.

Finon, D., Pignon, V., 2008. Electricity and long-term capacity adequacy: The quest for regulatory mechanism compatible with electricity market. *Utilities Policy*, 16, 143 – 158.

Fu, S., Ren, Y. 2011. A motivating mitigation mechanism for generators on condition of coordinated regulation of emission and electricity market. International Journal of Electrical Power & Energy Systems 33 (5) (junio): 1151–1160. doi:10.1016/j.ijepes.2011.01.008.

Friedman, D., Cassar, A. 2004. Economics lab: an intensive course in experimental economics. London: Routledge.

Friedman, D., Sunder, S. 1994. Experimental methods: a primer for economists. Cambridge: Cambridge university press.

Helm, D. (ed.) 2007. *The new energy paradigm*. Oxford University Press. Oxford, UK.

Huck, S., Normann, H., Oechssler, J. 2004. Two Are Few and Four Are Many: Number Effects in Experimental Oligopolies. Journal of Economic Behavior & Organization, 53(4): 435-446

IEA, 2007. Tackling Investment Challenges in Power Generation in IEA Countries. OECD/IEA, Paris, France, 210 pp.

IEA, 2014a. IEA - Energy security. URL https://www.iea.org/topics/energysecurity/

Islyaev, S., Date, P. 2015. Electricity futures price models: Calibration and forecasting. European Journal of Operational Research, 247(1), 144-154. doi:10.1016/j.ejor.2015.05.063

Jessoe, K., Rapson, D. 2015. Commercial and industrial demand response under mandatory time-of-use electricity pricing. Journal of Industrial Economics, 63(3), 397-421. Retrieved from <Go to ISI>://WOS:000363900800001

Joskow, P., 2008. Lessons Learned from Electricity Market Liberalization. The Energy Journal, Special Issue on The Future of Electricity, Papers in Honor of David Newbery, pp. 9–42.

Lara-Arango, D. 2014. Capacity mechanisms in electricity markets: An experimental study using system dynamics modeling. Lambert Academic Publishing.

Larsen, E.R., Arango, S. 2013. Thinking about the future: System Dynamics and the process of deregulation. Energy policy modelling in the 21 century.

Larsen, E.R., Osorio, S., van Ackere, A. (2016) Security of supply in the electricity sector: a framework for evaluation. Working Paper, University of Lugano.

Liu, Z., Zhang, X., Lieu, J. 2010. Design of the incentive mechanism in electricity auction market based on the signaling game theory. Energy Economics 35 (35), 1813–1819.

Morita, T., Managi, S. 2015. Consumer willingness to pay for electricity after the great east Japan earthquake. Economic Analysis and Policy 48, 82-105.

Meunier, G., Finon, D. 2006, Market Power in Electricity Markets, Suboptimal Investment and ISO Procurement, Presentation à la 26 USAEE/IAEE Conference.

Nerlove, M., 1958. Adaptive expectations and cobweb phenomena. Quarterly Journal of Economics 72, 227–240.

OFGEM, 2013. Electricity Capacity Assessment 2013 (Report to the Secretary of State No. 105/13). London.

Ramos, A., Gago, A., Labandeira, X., Linares, P. 2015. The role of information for energy efficiency in the residential sector. Energy Economics, 52, S17-S29. doi:10.1016/j.eneco.2015.08.022

Rassenti, S.J., Smith, V.L., Wilson, B.J., 2003. Controlling market power and price spikes in electricity networks: demand-side bidding. Proceedings of the National Academy of Sciences of the United States of America 100 (5), 2998–3003.

Reichmann, T. 2007, *Controlling*. Spinger-Verlag, Berlin

Roques, F.A., 2008. Market design for generation adequacy: healing causes rather than symptoms. Utilities Policy 16 (3), 171–183.

Sioshansi, F.P., Pfaffenberger, W (Eds.). 2006. Electricity market reform, Amsterdam, Elsevier.

Sundt, S., Rehdanz, K. 2015. Consumers' willingness to pay for green electricity: A meta-analysis of the literature. Energy Economics, 51, 1-8. doi:10.1016/j.eneco.2015.06.005

Van der, V, Reinier A.C., Abbasy, A, Hakvoort, R,A. 2012. «Agent-based analysis of the impact of the imbalance pricing mechanism on market behavior in electricity balancing markets». Energy Economics 34 (4) (julio): 874–881. doi:10.1016/j.eneco.2012.04.001.

Vasquez, C., Battle, C., Rivier, M., Perez-Arriaga, I., 2003, Security of supply in the Dutch electricity market: the role of reliability options. Report IIT-03-0841C, (prepared for the Office for Energy Regulation of the Netherlands-DTE), Comillas University, IIT.

Maurer, L. T. A., Barroso, L. A. 2011. Electricity Auctions: An Overview of Efficient Practices. A World Bank study.

Zou, X. 2009. «Double-sided auction mechanism design in electricity based on maximizing social welfare». Energy Policy 37 (11) (noviembre): 4231–4239. doi:10.1016/j.enpol.2009.05.019.

Appendix A: Equilibria derivation

Treatment 1:

Joint Maximization:

Consider the main equations in this economic model, i.e., the price equation and the players' profit equation.

$$P_t = A - B * Q_t \qquad (1a)$$

$$\pi_{i,j} = \left(P_t - C_{i,j}\right) * q_{i,j} \qquad (2a)$$

Since the sum of the players' individual capacities is total market production Q and we assume that all n players are identical, we have

$$Q_t = n * q_{i,j} \qquad (3a)$$

By replacing (3) in (1), we have

$$P_t = A - B * n * q_{i,j}$$

(4a)

Then, we replace (4) in (2), and we have

$$\pi_{i,j} = A * q_{i,j} - B * n * q_{i,j}^2 - C_{i,j} * q_{i,j} \qquad (5a)$$

We now derive (5) with respect to $q_{i,j}$ and equal it to zero to find the maximum value

$$\frac{\partial \pi_{i,j}}{\partial q_{i,j}} = A - 2 * B * n * q_{i,j} - C_{i,j} = 0 \qquad (6a)$$

We now find the value of $q_{i,j}$ given that A=6, B=0.1, n=5 and $C_{i,j}$=1:

$$q_{i,j} = \frac{A - C_{i,j}}{2 * B * n} \qquad (7a)$$

$$q_{i,j} = \frac{6 - 1}{2 * 0.1 * 5} = 5$$

Since we now have the value of $q_{i,j}$, we can find the values for Q_t and P by replacing the value of $q_{i,j}$ in (3) to obtain the value of Q_t and then replacing this value in (1) to find P:

$$Q_t = 5 * 5 = 25$$

$$P_t = 6 - 0.1 * 25 = 3.5$$

Cournot-Nash:

In this equilibrium, we follow the same process we followed in the joint maximization. However, we now assume that each player maximizes his or her own benefits, while the other players' benefits remain constant. This implies that the marginal change in the total market capacity is equal to the change of the player making the profit maximization. In mathematical terms, the previous statements can be described as

$$\frac{\partial Q_t}{\partial q_{i,t}} = 1 \qquad (8a)$$

The starting point is the same as the joint maximization equilibrium with the price and profit equations. As in the previous case, we have the following equation:

$$\pi_{i,j} = A * q_{i,j} - B * n * q_{i,j}^2 - C_{i,j} * q_{i,j} \qquad (9a)$$

We also know that

$$Q_t = n * q_{i,j} \qquad (10a)$$

We can rewrite (2) as

$$\pi_{i,j} = A * q_{i,j} - B * q_{i,j} * Q_t - C_{i,j} * q_{i,j} \qquad (11a)$$

Since we have the assumption stated in (1), the derivation is

$$\frac{\partial \pi_{i,j}}{\partial q_{i,t}} = A - \left(B * q_{i,j} * \frac{\partial Q_t}{\partial q_{i,j}} + Q_t * B\right) - C_{i,j} = 0 \qquad (12a)$$

$$\frac{\partial \pi_{i,j}}{\partial q_{i,t}} = A - \left(B * q_{i,j} + Q_t * B\right) - C_{i,j} = 0 \qquad (13a)$$

By replacing (3) in (6), we have

$$\frac{\partial \pi_{i,j}}{\partial q_{i,t}} = A - \left(B * q_{i,j} + n * q_{i,j} * B\right) - C_{i,j} = 0 \qquad (14a)$$

We then find the value for $q_{i,j}$, and then, we find the value for Q_t and

$$q_{i,j} = \frac{A - C_{i,j}}{B*(n+1)} \qquad (15a)$$

$$q_{i,j} = \frac{6-1}{0.1*(5+1)} = 8.33$$

$$Q_t = 8.33 * 5 = 41.66$$

$$P_t = 6 - 0.1 * 41.66 = 1.83$$

Perfect Competition:

This equilibrium is achieved when the price is equal to the marginal cost, that is, when the profits are theoretically zero. This theoretical zero does not mean that the firms actually receive zero profits; rather, it means that the firms receive the normal economic profit (no extra gains).

Margin cost = 1

Price = Margin cost

$$P_t = A - B * Q_t = Margin\ cost$$

$$Q_t = \frac{A - Margin\ cost}{B}$$

$$Q_t = \frac{6-1}{0.1} = 50$$

Now that we know the value for Q_t, which makes $P_t = 1$, we proceed to find the values for P_t and $q_{i,j}$:

$$q_{i,j} = \frac{50}{5} = 10$$

Treatment 2:

Joint Maximization:

In this treatment, we base our derivations in the basic scheme of Treatment 1, but we now have to consider the regulatory firm. Therefore, we start with the same two equations of the previous treatment and the regulator contribution (RC = regulator contribution) to the total production capacity of the market:

$$P_t = A - B * Q_t$$

(16a)

$$\pi_{i,j} = \left(P_t - C_{i,j}\right) * q_{i,j} \quad (17a)$$

$$Q_t = n * q_{i,j} + RC \quad (18a)$$

Since we know that the RC value depends on $n * q_{i,j}$, we can form 3 intervals for this value:

$$If\ n * q_{i,j} \leq 40, \qquad RC = 10$$

$$If\ 40 > n * q_{i,j} < 50, \qquad RC = 50 - n * q_{i,j}$$

$$If\ n * q_{i,j} \geq 50, \qquad RC = 0$$

By examining these 3 intervals, we can conclude that the third interval does not have the joint maximization equilibrium since the players' profits are 0 or negative in it. We can also conclude that since the price ranges from 2 to 0 in the second interval, this one does not have the equilibrium value since the maximum unitary profit in it is 1. Therefore, we know that the equilibrium lies in the first interval, where RC=10. Thus, we take RC as a constant equal to 10.

We replace (3) in (1):

$$P_t = A - B * n * q_{i,j} - B * RC \quad (19a)$$

We replace (4) in (2):

$$\pi_{i,j} = A * q_{i,j} - B * n * {q_{i,j}}^2 - RC * q_{i,j} * B - C_{i,j} * q_{i,j} \qquad (20a)$$

Now, we derive with respect to $q_{i,j}$ an equal to zero to find the maximum:

$$\frac{\partial \pi_{i,j}}{\partial q_{i,j}} = A - 2 * B * n * q_{i,j} - RC * B - C_{i,j} = 0 \qquad (21a)$$

$$q_{i,j} = \frac{A - RC * B - C_{i,j}}{2 * B * n} \qquad (22a)$$

Now, we find the values of$q_{i,j}$, Q_t and P_t:

$$q_{i,j} = \frac{6 - 10 * 0.1 - 1}{2 * 0.1 * 5} = 4$$

$$Q_t = 5 * 4 + 10 = 30$$

$$P_t = 6 - 0.1 * 30 = 3$$

Cournot-Nash:

For this equilibrium, we assume that Q_t changes as a result of a maximization process made by one player, while the rest of the variables remain constant.

$$\frac{\partial Q_t}{\partial q_{i,t}} = 1 \qquad (23a)$$

This is the same as saying that for this equilibrium, the derivation with respect to $q_{i,t}$ is the same as the one obtained with the derivation with respect to Q_t. Therefore, we have

$$\frac{\partial P_t}{\partial q_{i,t}} = \frac{\partial P_t}{\partial Q_t} = -B \qquad (24a)$$

With the consideration exposed in (2), we proceed to derive the profit equation:

$$\frac{\partial \pi_{i,j}}{\partial q_{i,t}} = P_t * \frac{\partial q_{i,t}}{\partial q_{i,t}} + \frac{\partial P_t}{\partial q_{i,t}} * q_{i,t} - C_{i,j} \qquad (25a)$$

$$\frac{\partial \pi_{i,j}}{\partial q_{i,t}} = P_t - B * q_{i,t} - C_{i,j} = 0 \qquad (26a)$$

We consider price equation (5) and production capacity equation (6):

$$P_t = A - B * Q_t \qquad (27a)$$

$Q_t \ = n * q_{i,j} \ + RC$ (28a)

We replace (6) in (5) to replace them in (4)

$$\frac{\partial \pi_{i,j}}{\partial q_{i,t}} = A - B * n * q_{i,j} - RC * B - B * q_{i,j} - C_{i,j} = 0 \qquad (29a)$$

$$q_{i,j} = \frac{A - RC * B - C_{i,j}}{B(n+1)} \qquad (30a)$$

We now find the value for $q_{i,j}$, Q_t and P_t (we also assume RC=10 for this equilibrium for the margin intervals we explain in the joint maximization)

$$q_{i,j} = \frac{6 - 10 * 0.1 - 1}{0.1 * (5 + 1)} = 6.66$$

$$Q_t \ = 5 * 6.66 + 10 = 43.33$$

$$P_t = 6 - 0.1 * 43.33 = 1.67$$

Perfect Competition:

This equilibrium is achieved when the price is equal to the marginal cost, that is, when the profits are theoretically zero. This theoretical zero does not mean that the firms actually receive zero profits; rather, it means that the firms receive the normal economic profit (no extra gains).

Marginal cost = 1

Price = Marginal cost

$$P_t = A - B * Q_t = Marginal\ cost$$

$$Q_t = \frac{A - Marginal\ cost}{B}$$

$$Q_t = \frac{6 - 1}{0.1} = 50$$

Since Q_t = 50, $q_{i,j}$ and RC can make infinite combinations that sum 50, we have the following ranges, $q_{i,j}$ and RC, that could add up to 50 units in combination:

$q_{i,j} = \ [8, 10]$

$$RC = [10, 0]$$

Treatment 3:

Since the market production capacity is fixed in this treatment, there is only one capacity and price scenario that is equivalent to the perfect competition equilibrium; that is

$$Q_t = 50$$

$$P_t = 1$$

In this treatment, the players bid in auctions; in other words, the players compete for a share of the total market production capacity (which is fixed). Therefore, $q_{i,j}$ behaves differently for every individual according to the auction results but always keeping the perfect competition equilibrium in the market ($Q_t = 50$ and $P_t = 1$).

To summarize the equilibria:

	T1		T2		T3	
	Cap	Price	Cap	Price	Cap	Price
JM	25.0	3.50	30.0*	3.00	NA	NA
CN	41.7	1.83	43.3*	1.67	NA	NA
PC	50.0	1.00	50.0*	1.00	50.0	1.00

Market capacity (Cap) and price for joint maximization (JM), Cournot-Nash (CN) and perfect competition (PC) equilibria.

Appendix B: Experiment instructions (translated from Spanish)

T2

INSTRUCTIONS

WARNING: Do not touch the computer until you are instructed to do so!

Welcome. In this game, you will play the role of an electricity producer. Every year, you decide how much new production capacity (power plants) you want. Your goal is to maximize the benefits for all periods of the experiment.

You are one of five electricity producers in a market and do not know who your competitors are and how they operate individually.

Power plants have a lifespan of 16 years and a construction period of 4 years, from the time you order new capacity until this new capacity begins to produce electricity.

Annual profits are taken for production multiplied by the difference between the price and unit costs. Unit costs are constant and equal to 1 unit of experimental money.

Production benefits = (Price - Unit costs)

The production capacity of each player cannot be negative and must be below 20 units. Think of this upper limit as a government regulation to maintain a minimum market competition.

Each year, the capacity and electricity production are given by investments in previous years. The capacity utilization is always 100% for all players, i.e., it is assumed that all plants are always working at 100% capacity.

All electricity produced by the market is consumed every year. The larger the total production, the lower the price at which electricity is sold. The exact relationship is given by a demand curve and is expressed as

Price = 6 to 0.1 • Total production ≥ 0

The price cannot be negative, and there is no economic growth influencing demand and electricity prices. The relationship between total output (x axis) and electricity prices (Y axis) is shown in the following figure.

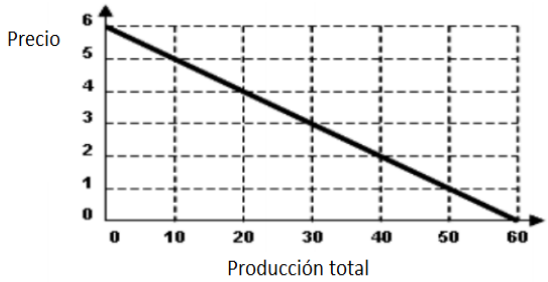


Figure 1. Relationship between total production and the price.

Every year, you should make investment decisions in new capacity (you can choose 0 units). After 4 years, new investments are added to the existing capacity. This capacity lasts for 16 years and is automatically reduced as the old plants are discarded.

In addition to the other participants, you will interact with a market regulator, operated by the computer. This regulator has a maximum capacity of 10 units and will invest if it detects that the total market capacity will be less than 50 units.

At the beginning of the experiment, the former managers have invested a constant amount of 0.6875 units / year. 5 companies are equal players with the same unit costs and the same initial capacity. The market starts with an initial capacity of 11 units for a total capacity of 55 market units. Consequently, the price is equal to 0.5 (experimental money) and the initial benefits are equal to -0.5.

TASKS OF THE EXPERIMENT

BE CAREFUL NOT TO PUSH "accept the decisions" unless you really mean it.

After pressing "accept the decisions", this decision for a particular year can no longer be changed.

1. Look at the information available for the company and the market.
2. Make your investment decision and type your decision in the box. Please note, you have to make an active decision every year. If you do not write anything in the box, the decision will be the same as the previous year because the previous investment decision is not automatically deleted. The program does not allow you to choose negative investments or investments that exceed the maximum capacity.
3. Type your investment into the assigned paper sheet.
4. Click "accept the decisions."
5. Wait until all market participants have made their decisions for the current year, and start at point 1 again, when information for decision-making for next year becomes available.

The game will continue until it stops at some unknown future year.

NOTE: This experiment requires that you do not share any information (verbal, written, gestures, etc.) with other participants. Please respect these rules because they are important for the scientific value of the experiment.

If it is required, you can direct questions to the personnel in charge of the experiment.

Thank you for participating in this game, and do your best!

T3

INSTRUCTIONS

CAUTION: Do not touch the computer until instructed!

In this game, you will play the role of an electricity producer. Every year, you decide how much you want to order in new capacity (power plants). Orders are made through bidding at an auction for licenses to build new capacity.

Your objective is to maximize their accumulated benefits. You are one of five electricity producers in a market and do not know who your competitors are or how they perform individually. Power plants have a useful life of 16 years, and they take 4 years to be built; that is, if you decide to build a plant today, you have will it within four years, and it will last 16 years from that time. Annual earnings are given by the production multiplied by the difference between the price and unit costs, minus the number of new licenses purchased multiplied by their price.

Benefits = Prod * (Price-Cost unit) - New licenses * License Pricing

Unit costs are constant and equal to 1 (experimental money / unit). The new licenses reduce benefits if the auction generates a positive price for licenses. If the auction generates a negative price for the licenses, the benefits increase. Think of the latter as if the government were subsidizing licenses.

The capacity of each player cannot be negative and must be below 20 units. Think of this upper limit as a government regulation to maintain a minimum market competition. Each year, the ability to produce electricity is given by licenses issued in prior years. The capacity utilization is always 100% for all players and all the electricity produced by the five producers consumed each year.

The government determines the total number of new licenses each year. The amount is configured such that each year the total number of new licenses is equal to the total capacity that has become obsolete. This means that the total capacity in the market remains constant over time. Since we assume no growth in demand over time, the price of electricity will also be constant and equal to 1 (experimental money / unit). This means that annual earnings depend only on the number and price of new licenses.

Each year, new licenses are auctioned. This is done through a graphical bidding curve where you can specify the number of new licenses you would like to receive for different prices. You can manipulate the graph by moving the cursor on the graph by clicking and dragging. The license prices range from negative values (subsidies) to positive values. The graph bid must be defined in such a way that you order fewer licenses as the price increases, and the bid for a price of 2.4 must be zero. If these requirements are not met, the program will ask you to edit the graph.

Each year, the government takes all the graphs of supply and assigns new licenses so that the total number of new licenses is equal to the amount of capacity that becomes obsolete. All players end up paying (or receiving) the same price (or subsidies) for new licenses. Each player receives the number of new licenses he or she has specified by the resulting equilibrium price. New licenses automatically lead to orders for new capacity. After four years, the new licenses are added to the existing capacity. The capacity lasts for sixteen years and is automatically reduced when old plants are discarded.

The previous business managers have received new licenses and invested a constant amount of 0.625 units per year, for a long time. 5 companies are the same, with the same unit costs and the same initial capacity. The market starts with an initial individual capacity of 10 units for each company and a market total capacity of 50 units.

BE CAREFUL NOT TO PUSH "accept the decisions" unless you really mean it. After pressing "accept the decisions", your decision for that particular year can no longer be changed.

1. Look at the information available to the company and the market.

2. Draw your offer with the graphics cursor. The program does not allow you to ask for more licenses when the price increases (logical error) or to request a number higher than zero when the price is 2.4 (maximum price). Note that you have to make an active decision every year. If you do not make any changes to the chart offer, your decision will be taken as equal to the previous year, as the previous offer is not automatically deleted.

3. Click "accept the decisions."

4. Wait until all market participants have made their decisions for the current year and receive information from the bids next year. Enter your assigned number in the paper reporting licenses.

5. Start at point 1 again. The game will continue until it stops itself at some unknown future year.

NOTE

According to the purpose of the experiment, it is required that you not share any information (verbal, written, gestures, etc.). Please respect these rules because they are important for the scientific value of the experiment. You can ask clarifying questions.

Thank you for participating in this game, and give the best to you!

Appendix C: Experiments' interfaces.

T2

General information of the year

Player's capacity (Unit)	11,0
Other players' capacity (Unit)	40,0
Regulator capacity (unit)	0,0
Total capacity (unit)	55,0
Price ($/Unit)	0,5
Production unitary cost ($/Unit)	1,0
Profit Margin ($/Unit)	-0,5

Player's capacity (Unit)

Under construction 1 year	0,7
Under construction 2 year	0,7
Under construction 3 year	0,7
Total capacity under construction	2,1
1 year old capacity	0,7
2 years old capacity	0,7
3 years old capacity	0,7
4 years old capacity	0,7
5 years old capacity	0,7
6 years old capacity	0,7
7 years old capacity	0,7
8 years old capacity	0,7
9 years old capacity	0,7
10 years old capacity	0,7
11 years old capacity	0,7
12 years old capacity	0,7
13 years old capacity	0,7
14 years old capacity	0,7
15 years old capacity	0,7
16 years old capacity	0,7
Total online capacity	11,0

Decisions

New capacity (Unit)	0,0
Expected price ($)	0,0

Player's performance

Profit this year ($)	-5,5
Acumulated profits ($)	-5,5

Year 1

T3

Total market licences 3,13

Year 1

General information

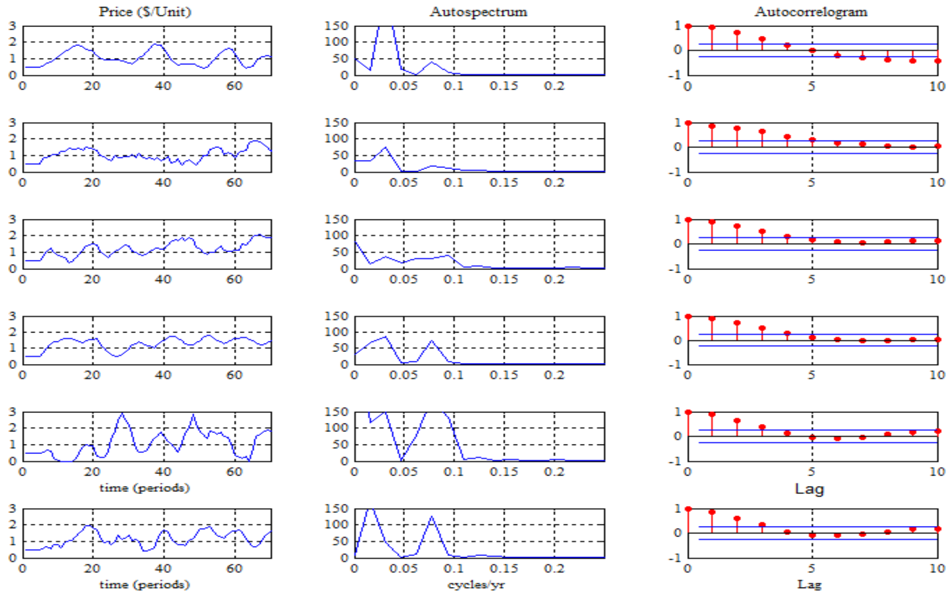
Player's production (unit)	10,00
Other players' production (unit)	40,00
Total production (unit)	50,00
Electricity price ($/unit)	1,00
Accumulated profits ($)	1,31

Price and quantities of the year

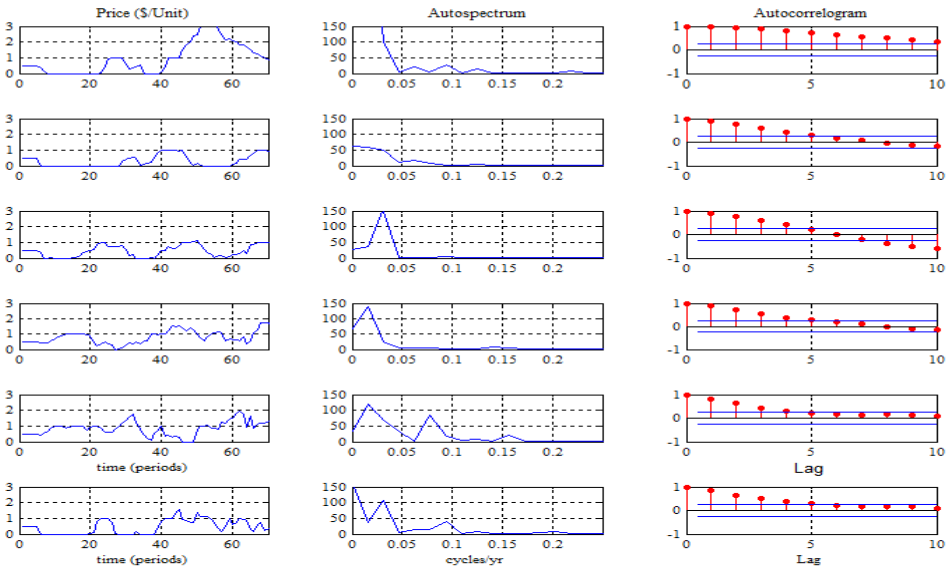
License's price	-2,10
Number of granted licenses	0,625

Appendix D: Autospectra and autocorrelogram for all three treatments

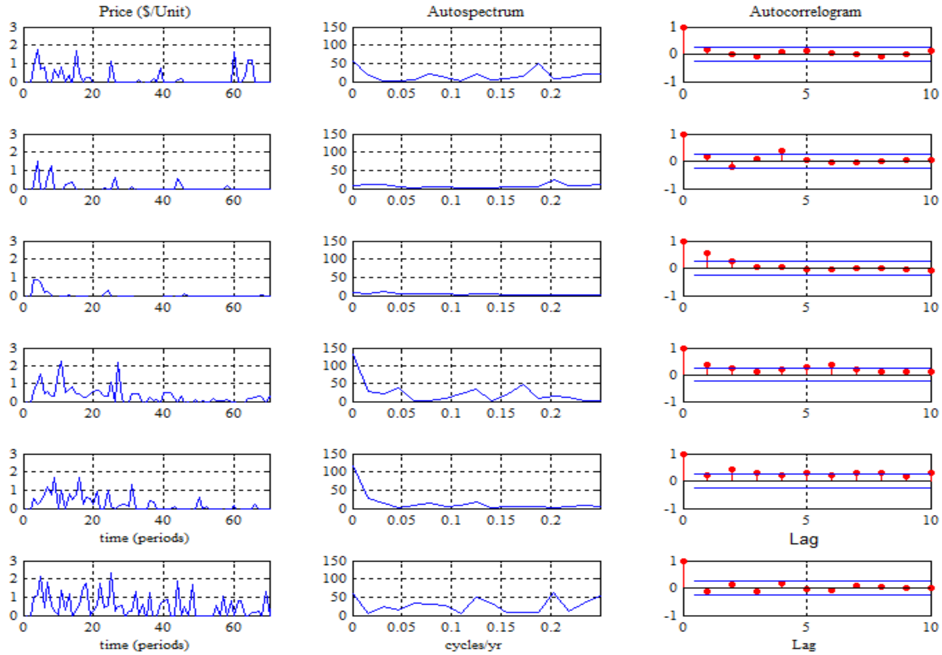
T1: Electricity price in groups 1 to 6 (from top to bottom)



T2: Electricity price in groups 1 to 6 (from top to bottom)



T3: License prices in groups 1 to 6 (from top to bottom)



Paper 3

Testing Meadows' hog cycle theory by laboratory experiment

This paper was presented as a parallel presentation in the International Conference of System Dynamics 2017 in Cambridge, USA.

Testing meadows' hog cycle theory by laboratory experiment

David Lara-Arango
Erling Moxnes
System Dynamics Group - University of Bergen, Norway

Abstract

Commodity prices are known to fluctuate. Cyclical tendencies are described in terms of period lengths, amplitudes, and regularity. Such fluctuations cause problems for producers, consumers, labor, and national economies. Here we take a closer look at the well-known hog cycle, which has been observed in hog markets since major markets were established. The Cobweb theory suggests that the internal working of the market causes these fluctuations. However, this theory has been rejected in economic literature because cycles can be shown to disappear when assuming a simple behavioral decision rules among producers. Furthermore, laboratory experiments have failed to replicate lasting cycles. This has opened up for theories that explain price fluctuations as caused by random, external shocks. However, there exists a more advanced dynamic model of the hog market, which produces lasting cycle, the Meadows model. Here we test this theory by a laboratory experiment. Without random shocks, the experiment produces price cycles similar to the Meadows theory and to historical observations. Meadows' claim that livestock adjustments are driven by current price-cost ratios is not rejected.

Keywords: Commodity markets, price cycles, learning equilibrium.

1. Introduction

Commodity cycles and their effects for producers, consumers and national economies have been a recurring topic of research (Deaton, 1999; Akiyama et al 2003). Typically, economic literature explains price fluctuations as a result of external shocks such as weather events, economic and political instabilities (Deaton, 1999). While random external shocks certainly affect commodity prices, they do not provide an explanation for lasting cyclicality (Deaton and Laroque, 2003). The Cobweb model provides such an explanation. However, this model has a weak standing in the economic literature, and it is frequently omitted in economic textbooks. Nerlove (1958) showed that the cycle disappears if suppliers make use of adaptive price expectations. Moreover, laboratory experiments have failed to produce lasting cycles (Carlson, 1967; Miller, 2002; Plott and Smith, 2008). Since many real commodity markets tend to be cyclical, it seems pertinent to ask whether the highly-simplified Cobweb model fails to capture essential characteristics of commodity markets. In this respect, a study by Arango and Moxnes (2012) shows that when complexity and realism in the Cobweb theorem is increased, laboratory experiments can produce cycles.

We focus on the hog market, which provides one of the most recognized examples of commodity cycles. For this market, Meadows (1970) developed a more complete dynamic model than the Cobweb model. Until now, this model has not been tested in the laboratory. This is what we do in this paper. Unlike previous experiments to test the Cobweb model or Cournot markets (Carlson, 1967; Sonnemans et al 2004; Arango and Moxnes, 2012; Arango et al, 2013), Meadows has a more detailed description of such market. He captures the livestock of hogs, stocks of growing pigs, and inventories of pork. Livestock is important because an increase in livestock requires a reduction in slaughtering, thus creating a short-term backward bending supply curve. Pork inventory is important since it carries mismatches between supply and demand into the future. Both livestock and inventories complicate the model and make it more difficult to form rational expectations. Based on data of the U.S hog market between 1956 and 1966, Meadows hypothesizes that breeding decisions (or decisions to increase or decrease the livestock), are simply based on observations of price-cost ratios. As a result, his model produces price cycles similar to those observed.

The experiment has N=9 markets, each with five players. We observe similar price cycles to those predicted by Meadows' model and as observed in US hog markets. We cannot reject Meadows' simple decision rule. First, we present Meadows' model, next our hypotheses and the experimental design. Results and conclusions follow.

2. Meadows' Model

Meadows' hog cycle model seeks to describe and explain the dynamics of hog markets in the US in the 1950's and 1960's. Hog price cycles had for long been associated with random variations in the price of corn, the main forage for hogs (Meadows, 1970). As a consequence of this belief, the US government enacted the Agricultural Adjustment Act of 1938 in order to stabilize corn supplies through the use of corn inventories. However, contrary to expectations, amplitudes increased as a result of the policy (Dean and Heady, 1958). This paradoxical result motivated Meadows to develop a market model to explain price cycles when there are no random external shocks.

Meadows' model is a continuous-time simulation model. To adapt the model to our experimental design we make several inconsequential modifications, see the comparison of the original and the modified model behaviors in Figure 1. The major reason for these modifications is to simplify the documentation of the model and the introduction to the players. Where Meadows use piecewise linear functions to describe nonlinearities, we use linear functions. Units are changed from the Imperial system to the Metric system, which is more familiar to our players. We explain other modifications as we go through the model equations. Table 1 at the end of this section shows values and units of constants and initial values.

At the center of the model is the inventory of pork,

$$I = \int (S_P - C_P)dt + I(0) \tag{1}$$

which grows or declines with discrepancies between pork supply S_P and pork consumption C_P. $I(0)$ denotes the initial amount of pork in inventory. Pork supply measured in terms of weight

$$S_P = (S_M + S_L) * W * Y \tag{2}$$

is given by the number of slaughtered mature hogs S_M plus the number of slaughtered livestock sows S_L. W is the average live weight of slaughtered hogs, and Y is the average hog-dressing yield.

Pork consumption

$$C_P = (1 - e^{-I/I_L})c_P P_{US} \tag{3a}$$

is the product of the total US population P_{US} (in the 1960s) and per capita consumption of pork c_P. Inventory never goes to zero in Meadows' model simulations. Since we do not know what the players will do, we limit consumption if inventory approaches or falls below a low limit I_L. [Unfortunately, in the experiment reported in this paper, the following erroneous formulation was used

$$C_P = \begin{cases} c_P * P_{US} & , \ c_P * P_{US} \leq I \\ I & , \ c_P * P_{US} > I \end{cases} \tag{3b}$$

where consumption is limited to what is in the inventory at any point in time. Since the inventory is low compared to the monthly throughput, this formulation leads to permanent rationing of consumption. In turn, there is no longer need for high prices to limit consumption, and consequently, there is less need for hog production.]

Per capita consumption

$$c_P = a_c P_R + b_c \tag{4}$$

depends on the retail price of pork P_R and parameters a_c and b_c. Here we have linearized the nonlinear function used by Meadows. The original function has an upper and lower limit for consumption. Simulated price variations tend to be in the nearly linear portion of the relationship. A regression on the original function produced values for a_c and b_c ($R^2 = 0.97$).

The retail price

$$P_R = P_W/Y + M \tag{5}$$

is given by the wholesale price of hogs P_W divided by the average hog dressing yield *Y* to get the wholesale price of pork. To this expression is added the sum of margins *M* for all middlemen in the process before the pork is presented in retail stores. The wholesale hog price

$$P_W = \max(a_P^1 I_R + b_P^1,\ a_P^2 I_R + b_P^2,\ a_P^3 I_R + b_P^3; 0) \tag{6}$$

is a declining and stepwise linear function of the relative inventory coverage I_R. While Meadows' nonlinear relationship was close to linear (a linear regression gives $R^2 = 0.98$), we use a convex relationship with two kinks. This choice implies that it takes a higher inventory than in the original model for the price to go as low as zero. Similar negative relationships between inventory and price have been documented for commodities such as copper (Klein and Marquez, 1989), oil (Fattouh, 2009) and agricultural products (UN, 2011). The relationship reflects risks of inventories approaching capacity limits and risks of stockouts. The relationship also presumes that there is an underlying understanding in the market of what the price should be for some learned equilibrium inventory level.

The relative inventory coverage

$$I_R = I_C/I_{DC} \tag{7}$$

is given by the inventory coverage I_C divided by the desired inventory coverage I_{DC}, which is assumed to stay constant over time. The inventory coverage

$$I_C = {}^{I}/_{C_E} \tag{8}$$

is given by the inventory I divided by the expected consumption rate C_E, which is set constant and equal to the expected average consumption over the long run.

Then we turn to the supply side. Slaughtering of mature hogs is given by

$$S_M = M_S/T_F \tag{9}$$

where M_S is the stock of mature hogs and T_F is the optimal length of the feeding period for mature hogs. The stock of mature hogs

$$M_S = \int (M_R - S_M - A_L)dt + M_S(0) \tag{10}$$

increases with the maturation rate M_R, it decreases with slaughtering S_M, and it decreases with adjustments to increase the stock of livestock A_L. The maturation rate is given by a third-order delay of the breeding rate B_R. In the last step, the maturation rate M_R follows from the oldest stock of piglets P_S^3

$$M_R = P_S^3/(T_M/3)$$

$$P_S^3 = \int (M_R^2 - M_R)dt + P_S^3(0)$$

$$M_R^2 = P_S^2/(T_M/3)$$

$$P_S^2 = \int (M_R^1 - M_R^2)dt + P_S^2(0) \tag{11}$$

$$M_R^1 = P_S^1/(T_M/3)$$

$$P_S^1 = \int (B_R - M_R^1)dt + P_S^1(0)$$

where P_S^1, P_S^2, and P_S^3 represent the stocks of piglets in three age classes each with a lifetime of $T_M/3$, where T_M is the total time it takes for piglets to mature. M_R^1, M_R^2, and

M_R represent maturations out of the respective age classes. By using three short delays, the pulse response will be more narrowly distributed than for one long delay and more widely distributed than for a discrete lag. The resulting distribution reflects that some piglets grow faster and reach the slaughter age earlier than others.

The breeding rate

$$B_R = L_S p_L l_M \tag{12}$$

is given by the livestock L_S, the number of piglets per litter p_L, and the number of litters per month per livestock l_M. The livestock

$$L_S = \int (A_L - S_L)dt + L_S(0) \tag{13}$$

varies with positive or negative adjustments of the livestock A_L, and decreases with the slaughtering of livestock sows

$$S_L = L_S\, f_S / T_L \tag{14}$$

where f_S denotes the fraction of sows in the livestock and T_L is the average productive life of sows. Hence, it is only sows that leave the livestock as slaughtered hogs. Boars go back to the stock of mature hogs if the livestock is adjusted downwards. The adjustment of the livestock

$$A_L = (L_D - L_S)/T_A \tag{15}$$

is given by the difference between the desired livestock L_D and the actual livestock L_S. The adjustment time T_A reflects that it takes some time to reallocate the necessary resources to handling the hogs (man-hours, sheds, feed stores, etc.). The expression implies that in the medium-term, farmers can adjust their capacity fairly rapidly, with little need for adjustments in facilities. This is consistent with empirical evidence (Estabrook, 2015).

Different from the continuous decisions (IL_D) in the original model, the experiment requires that decisions be made at distinct points in time and are held constant between these points. Desired livestock is updated according to

$$L_D = \int (IL_D - L_D)\delta(t, 3{,}0)dt + \ L_D(0) \tag{16}$$

where Dirac pulses $\delta(t, 3{,}0)$ ensure that every third month, starting at time zero, the desired livestock is updated from the current desired livestock L_D to the continuously updated indicated desired livestock. We linearize Meadows' function for indicated desired livestock to

$$IL_D = \ b_D + a_D E_{RPC} \tag{17a}$$

This is a close approximation to Meadows' function ($R^2 = 0.98$). [Due to the erroneous formulation in Equation 3b, we will simulate with a similar equation

$$IL_D = \ b_D^* + a_D E_{RPC} \tag{17b}$$

where the intercept b_D^* is lowered to compensate for the reduced consumption rate.]

The expected ratio of price to cost

$$E_{RPC} = \int \{(R_{PC} - E_{RPC})/T_R\}dt + \ E_{RPC}(0) \tag{18}$$

is given by a continuous version of the standard adaptive expectation formation. T_R denotes the time needed to form expectations and to make decisions about adjusting the livestock. The ratio of price to cost

$$R_{PC} = P_W / C_m \tag{19}$$

is given by the ratio between the wholesale price P_W and the operating cost C_m.

Meadows compares hog prices to a constant cost of corn. In order to account for increasing marginal costs (manure, hygiene, crowding etc.) associated with increasing livestock, we modify and allow for increasing marginal costs

$$C_m = \begin{cases} C_{m0} & , \; L_S \leq L_E \\ C_{m0} + \left(a_m \frac{L_S - L_E}{L_E}\right)^2 & , \; L_S > L_E \end{cases} \tag{20}$$

The marginal cost[2] is constant and equal to C_{m0} for livestock in the range from zero to the perfect competition equilibrium livestock L_E. Then marginal costs increase for larger livestock. The main reason for introducing increasing marginal costs is to ensure that the game is stationary.

2 C_{m0} is estimated based on data found in the Animal Nutrition Handbook (Chiba 2014).

Table 1: Parameter values and initial conditions.

Parameter	Value	Initial condition	Value
W	109 kg	$I(0)$*	200x10^6 Hog
Y	0.58 dimensionless	$M_S(0)$*	13x10^6 Hog
a_c	-0.05 dimensionless	$P_S^3(0)$	9.76x10^6 Hog
b_c	7.91 USD/hog	$P_S^2(0)$	9.76x10^6 Hog
M	31.55 USD/hog	$P_S^1(0)$	9.76x10^6 Hog
a_P^1	-42.1 USD/hog	$L_S(0)$*	8.2x10^6 Hog
b_P^1	40 USD/hog	$E_{RPC}(0)$	23.07 USD/Hog
a_P^2	-15.55 USD/hog	P_{US}	200x10^6 People
b_P^2	31.8 USD/hog		
a_P^3	-2.18 USD/hog		
b_P^3	7.04 USD/hog		
I_{DC}	0.36 Months		
C_E	9.9x10^8Pounds of pork		
T_F	2 Months		
T_M	10 Months		
p_L	7 Hog/litter		
l_M	0.17 Litter/month		
F_S	0.6 dimensionless		
T_L	36 Months		
T_A	5 Months		
a_D	3.2x10^6 Hog		
b_D	5.3x10^6 Hog		
b_D^*	1.06x10^6 Hog		
T_R	6 Months		
a_m	2 USD/hog		
C_{m0}	11.84 USD/ hog^2		
L_E	38.4x10^6 Hog		
I_L	1.74x10^6 Pounds of pork		

*Denotes that stocks are initialized out of equilibrium in Meadows' model.

Figure 1 shows the behaviors of the original and the modified Meadows models simulated with the initial values in Table 1. If it was not for the erroneous rationing of consumption, the two model versions would produce cycles with nearly identical amplitudes and with hardly any damping. Hence, the cyclical phenomenon is not sensitive to "correct" modifications we made. [With rationing in place, the modified version produces dampened cycles. This is the phenomenon we try to replicate in the experiment]. Initial conditions are different due to the linearization of the price function. The original function produces an initial price of $20.96, whereas the linear function produces an initial price of $23.07.

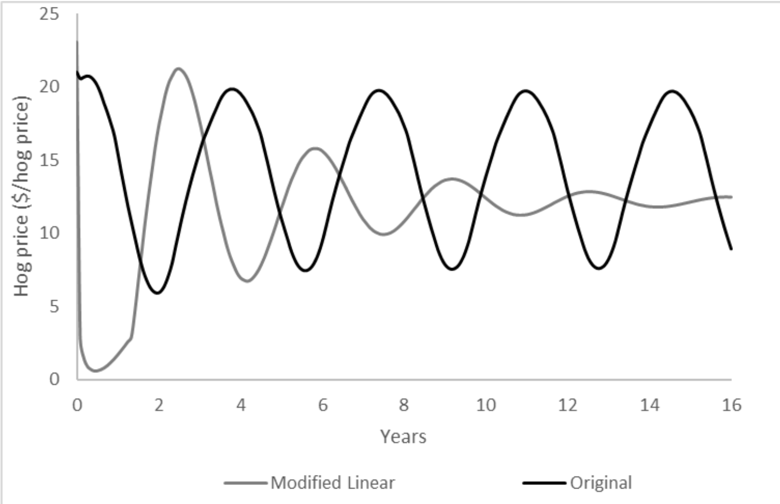


Figure 1: Behaviors of original and modified model with the same initial conditions.

3. Hypotheses

Meadows' model is based on much prior information about the hog market from statistical sources and conversations with scholars and hog farmers. The model produces lasting cycles similar to those observed in the US hog market. The present experiment enables an additional test of Meadows' hypothesis in an environment with no ongoing external disturbances. The experiment produces detailed data to examine decision rules, decision delays, and prices relative to known equilibria.

The experiment is interesting because it represents a novel approach to see if cycles can appear in experimental markets. Except for one treatment in Arango and Moxnes (2012), previous experiments have failed to produce lasting cycles (Sonnemans et al, 2004; Huck et al, 2002). There are two main reasons why Meadows' model is more likely to produce cycles than Cournot markets and Cobweb markets. First, Meadows introduced an inventory, which adds an extra state variable (stock) in the feedback loop on the supply side of the market. Second, decisions to increase livestock and meat production have a counteracting short-term effect since hogs are added to the livestock instead of being slaughtered. Hence a first reaction to a price-induced desire to increase production is reduced supply and an upward pressure on price.

Specifically we test Meadows' assumption that adjustments in livestock are determined by recent ratios between the hog price and the marginal production cost (originally the cost of corn). This corresponds to Nerlove's (1958) hypothesis about adaptive price expectations and a linear relationship between price and investment. Meadows' hypothesis differs from Nerlove's hypothesis that supply follows investments after a simple time lag. As can be seen from the preceding section, Meadows' model gives a far more detailed and realistic description of the supply side of the hog market. Observations of persistent price fluctuations in hog markets and previous observations of cycles in experimental markets (Arango and Moxnes, 2012) motivate the hypothesis.

The experiment also allows for a study of how quickly the experimental markets approach equilibria around which they may cycle. We hypothesize a gradual approach where adjustments are faster the further away from equilibrium the market is. We compare the observed (learned) equilibria to three theoretical equilibria. Equilibrium prices (in 1968 prices) are as follows: joint maximization ($21.82/hog), Nash equilibrium ($14.64/hog), perfect competition ($11.84/hog), see Appendix A for derivations.

4. Experimental design

The design is a Cournot market with increasing marginal costs, under Huck's standard conditions[3]. The experiment consists of 9 markets each with 6 players. The modified

[3] Standard conditions (Huck et al., 2004, p.106): a. Interaction takes place in fixed groups; b. Interaction is repeated over a fixed number of periods; c. Products are perfect substitutes; d. Costs are symmetric; e. There is no communication between players; f. Participants have complete information about their own payoff functions; g. Participants receive feedback about aggregated supply, the resulting price, and their own individual profits; h. The experimental instructions use an economic frame (instructions use economic terms such as "market", "price", "consumption", etc)

version of Meadows' model is used to capture the structure of the experimental market. To get a reasonable scale, each player is assumed to make up 1/10,000 of the entire US hog market. Initial conditions and parameters that reflect scale are adjusted accordingly. The total size of the market is given by the sum over the six players. Appendixes B and C show the player interface and the instructions. The experiment is programmed in Livecode 7.0.1.

Player *i*'s payoff in NOK

$$PF_i = AP/1586 + 40 \geq 40 \tag{21}$$

is a linear function of the accumulated profits *AP* above a minimum level of NOK 40 ($5).

The interface and instructions give players information about the current hog price, and the hog prices one and two months ago. The players get information about own production and livestock, marginal costs, and profits. Players do not get information about consumption and inventory, however, they could estimate inventory levels from a graph showing the exact relationship between inventory and hog price. Players have perfect information about the demand curve. [Obviously they did not get information about the erroneous effect of the first order control on consumption.]

To reduce the burden of analysis, players have access to a profit calculator. Using their own price expectations, they can find their own and other players' optimal production and livestock. Players are also informed that an increase in the livestock will not give an immediate effect on slaughtering. They get to know that it takes 5 months from a decision to increase the livestock is made until new sows start to give birth, and that it takes 12 months from births to slaughtering.

A few changes were made to simplify the task. Players get to see demand as a function of the hog price rather than the proportional pork price; the pork price is not reported.

.

Another simplification is that in the instructions the livestock is treated as if it consists of only sows. This is of no practical concern since the program and the profit calculator keep track of the balance between sows and boars. Finally, players get information that slaughtering and sales take place automatically, and that they are not involved in setting prices, prices are determined by the inventory coverage only. While this is an approximation to what happens in real markets, players get precise information that this is how the experimental market works.

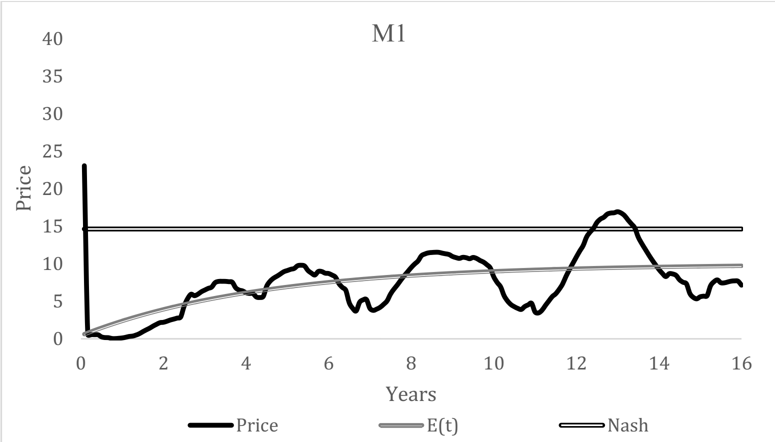
The perception and decision delay of 6 months (Equation 18) is not included in the laboratory experiment. Players use their own procedures for expectation formation, we observe and compare to the estimate in Meadows' model.

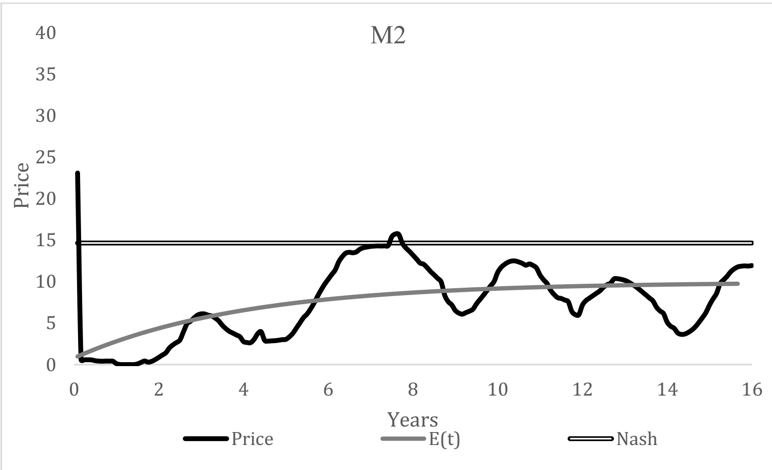
Experimental procedure:

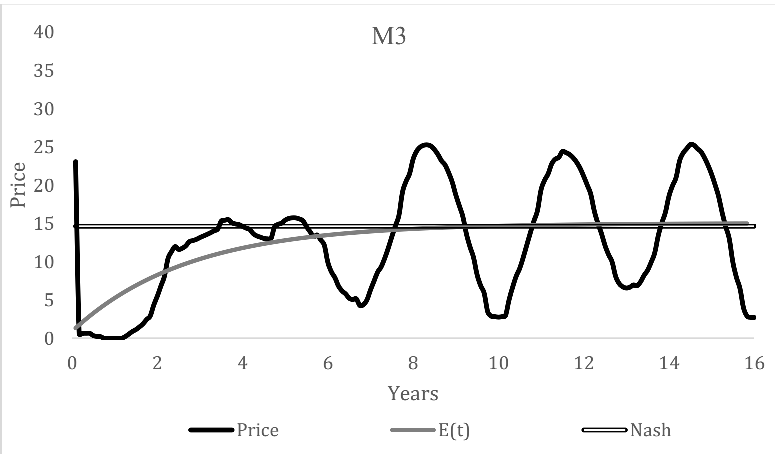
We recruited 54 students in total, 22 master students from Economics and 32 master students with varied backgrounds at the University of Bergen, Norway. Participants were randomly assigned to different computers, and did not know whom they were competing with. The participants received written instructions. They were allowed to ask technical questions and to play the three first rounds of the experiment to familiarize themselves with the game. After the try-out rounds, the experiment was reset and the students were reminded that their payoffs depended on their own performance and that the game would last for 64 periods (192 months or 16 years).

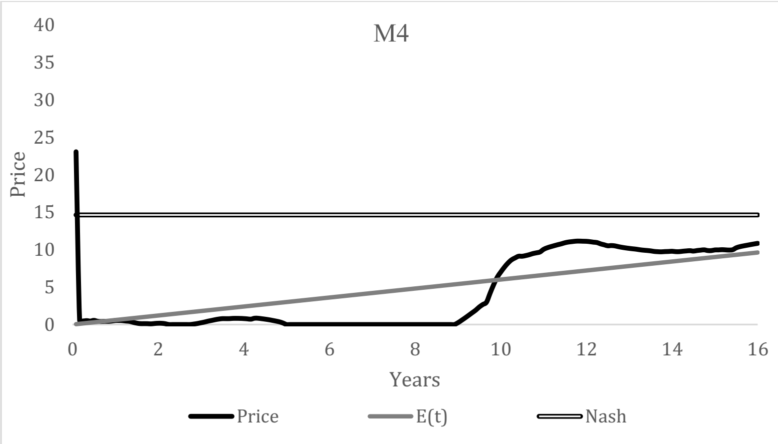
5. Results

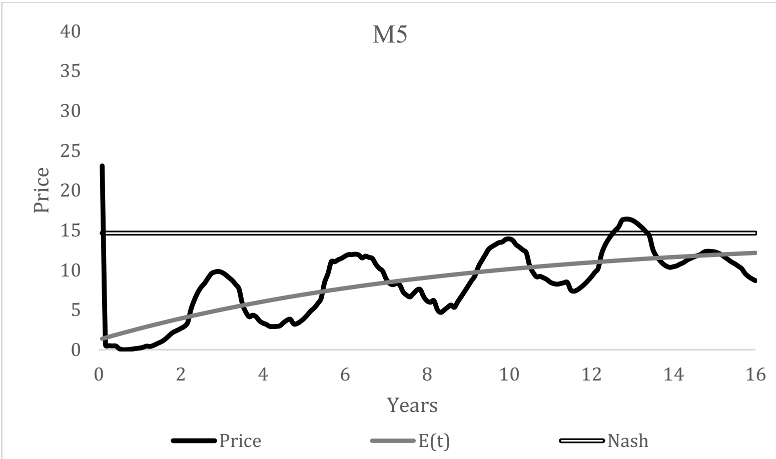
Figure 2 shows price developments for all nine markets. With the exception of M4, all markets show clear signs of cycles with expected period lengths. To estimate period lengths we consider the time intervals with the largest number of consecutive and well-defined peaks and divide by the number of peaks minus one. Table 2 shows the period lengths. The average period length is close to the period length of the modified Meadows model.

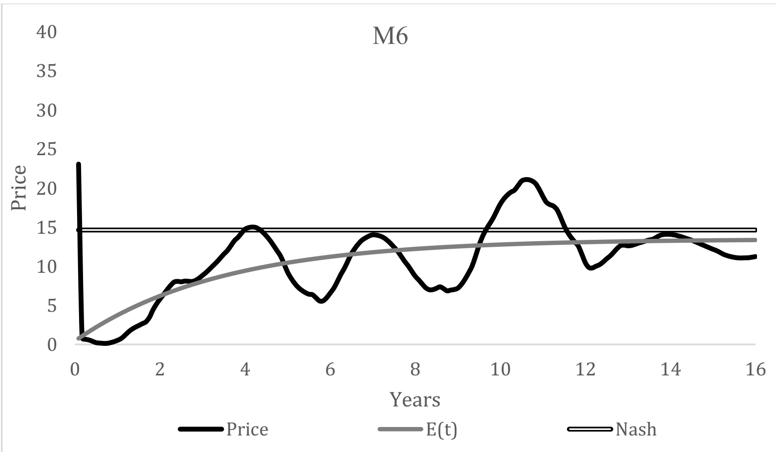

M1
40
35
30
25
20
15
10
5
0
Price
0
2
4
6
8
10
12
14
16
Years
Price
E(t)
Nash

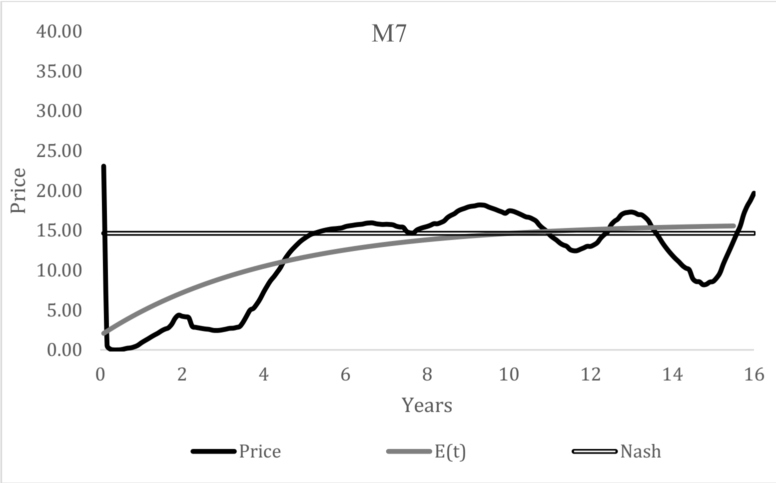

M2
40
35
30
25
20
15
10
5
0
Price
0
2
4
6
8
10
12
14
16
Years
Price
E(t)
Nash

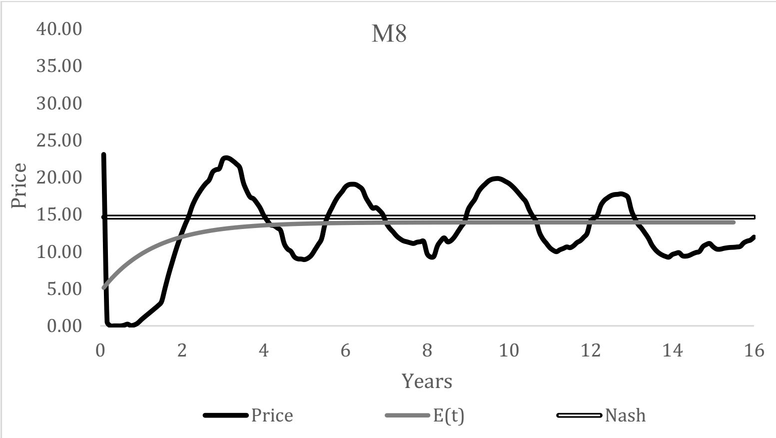

M3
40
35
30
25
20
15
10
5
0
Price
0
2
4
6
8
10
12
14
16
Years
Price
E(t)
Nash


M4
40
35
30
25
20
15
10
5
0
Price
0
2
4
6
8
10
12
14
16
Years
Price
E(t)
Nash


M5
40
35
30
25
20
15
10
5
0
Price
0
2
4
6
8
10
12
14
16
Years
Price
E(t)
Nash


M6
40
35
30
25
20
15
10
5
0
Price
0
2
4
6
8
10
12
14
16
Years
Price
E(t)
Nash





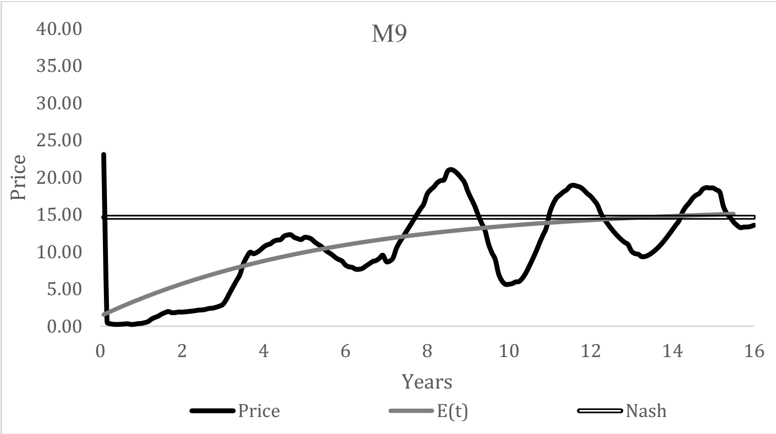


Figure 2: Observed market prices and learned equilibrium E(t) for all nine markets M1 to M9.

Table 2: Estimated period lengths for all markets and time intervals with consecutive peaks.

Market	Time interval for estimation	Period length
M1	3.3-12.8	3.2
M2	2.8-15.6	3.2
M3	5.0-14.4	3.1
M4	n.a.	n.a.
M5	2.8-14.8	3.0
M6	3.9-14.0	2.5
M7	1.9-15.8	3.5
M8	2.9-12.6	3.2
M9	4.4-14.8	3.4
Average		3.1
Meadows modified (Fig.1)	3.2-14.7	4

Next, we observe from the price curves that there is an upward trend from the early low prices that follow from the initial disequilibrium. It takes time and experience to establish equilibria. We assume that learning takes place through a feedback process where learning is rapid when prices are far away from desired and slower as equilibrium is approached. The result of such a process is captured by the following equation

$$E_t = (E_\infty - E_0)\left(1 - e^{-t/T}\right) + E_0 \geq 0 \tag{22}$$

where E_∞ denotes the ultimate equilibrium value, E_0 the starting value, and T the time constant.

Table 3: Ultimately learned equilibrium price (E_∞), initially perceived equilibrium price (E_0) and adjustment time for learning (T) for all markets.

Market	E_0	E_∞	T
M1	0.5	10	4.4
M2	0	9.9	4.0
M3	0	15	2.7
M4	0	2490024	124742
M5	1.3	14.2	8.7
M6	0.4	13.5	3.4
M7	0	15.9	4.2
M8	0	13.9	1.3
M9	0	16.1	5.7
Average	0.3	13.6	4.3

Whenever a regression gives a negative value of E_0, the regression is repeated with no intercept. All parameters are significantly different from zero at the 5%-level.

Table 3 shows the regression results. Most markets start out with a learned equilibrium price close to zero and far below the ultimate equilibrium price. On average the markets tend towards a learned equilibrium price of 13.6 (M4 not included), which is between the competitive equilibrium (11.8) and the Nash equilibrium (14.9). The average adjustment time T is 4.3 years.

Table 4: Average prices as percentages of the Nash equilibrium (NE) and the Perfect competition equilibrium (PCE) for various experiments.

	Present paper E_∞/Aver.	Chamberlin (1948)	Harstad et al (1998)[1]	Huck et al (2000)[2]	Arango & Moxnes (2012)[3]
Percentage of NE	91/64	91	33	99	64
Percent of PCE	115/81	94	110	n.a	117

[1] First treatment with players as individual price setters

[2] Cournot market treatment (basic)

[3] Last treatment (T4)

Table 4 shows that the experiment produces equilibrium or average prices, which fall within the range of earlier experiments of the Cournot type. All experiments tend to end up with prices below the Nash equilibrium and close to or above the Perfect Competitive equilibrium. The average price of our experiment is particularly low because of the initial disequilibrium condition and the long time needed to learn equilibrium.

Next, we test Meadows' decision rule for desired livestock. Since his rule represents market decisions, we test a model for aggregate desired livestock as a function of the average ratio of price to costs R_{PC}. Since we have no observations of the expected price-cost-ratio E_{PCR} we reverse the sequence of Equations 17 and 18. This can be done since the decision rule is linear. We use the following linear regression model

$$L_{D,t} = \alpha_0 + \alpha_1 L_{D,t-1} + \alpha_2 \bar{R}_{PC,t} + \sum_{j=1,j\neq 6}^{9} d_j \, D_{j,t} + \varepsilon_t \qquad (23)$$

Expectation formation is captured by a first order autoregressive model (Koyck lag). The average ratio of price to cost over all players in a market is denoted $\bar{R}_{PC,t}$. We pool the data for all nine markets, and for that reason we introduce dummies $D_{j,t}$ for all markets except M6. This is motivated by the different equilibrium values E_∞ reported in Table 3. Finally, note that due to the fact that decisions are made only once every third month, time *t* moves in steps of three months. This gives 63 data points for each market when the first data point is omitted due to lacking information about $L_{D,t-1}$, altogether 567 data points for the pooled data.

The coefficient α_1 represents the weight on the previous value for the livestock. The time to form expectations can be calculate as

$$T_R = -1/\ln(\alpha_1) \qquad (24)$$

The decision rule for indicated livestock in Equation 17b follows from the regression in the following way

$$IL_{D,t} = (\alpha_0 + \tfrac{1}{9}\sum_{j=1}^{9} d_j \, D_{j,t} + \alpha_2 \bar{R}_{PC,t})/(1-\alpha_1) = \hat{b}_D + \hat{a}_D \bar{R}_{PC,t} \qquad (25)$$

where $d_6 = 0$.

Table 5 shows estimated parameters for different sets of data. Of greatest interest is the slope coefficient $\hat{a}_D$. Whether M4 is included or not is of little importance for the slope. When the early years are excluded such that prices cycle around a relatively constant mean value (learned equilibrium), slopes are steeper and are quite close to the slope in the down-scaled Meadows' model.

Table 5: Results of regressions on pooled observations and parameters in the modified Meadows model.

	$\bar{d}$	α_0	α_1	α_2	T_R[6]	$\hat{b}_D$	$\hat{a}_D$
Experiment[1]	9.5	76**	0.73**	41**	9.5	317	152
Experiment[2]	6.5	106**	0.63**	59**	6.5	304	159
Experiment[3]	7.6	38	0.68**	94**	7.8	143	294
Experiment[4]	3.8	54*	0.45**	167**	3.8	105	304
Meadows[5]					6.0	534	329

Significance levels: ** $p<0.0001$, * $p<0.001$

[1] All data, N=567

[2] All data except M4

[3] Last 10.5 years

[4] Last 10.5 years except M4

[5] Parameters are estimated from the scaled down close to linear function used in Meadows' model.

[6] Time delay is reported in months (not 3-month steps)

M4 has a strong effect on the estimate for the expectation delay. We also see that the delay time is reduced when the early years are excluded from the data. In this case and when M4 is excluded, the estimate of 3.8 months is 2.2 months shorter than the estimate in Meadows' model. When the effect of making decisions only every three months is taken into account, the experiment produces a delay time that seems consistent with Meadows' model. When the individual markets are simulated with estimated parameters dampened cycles result. With a minimum of internally or externally generated randomness, cycles persist with stable characteristics.

6. Conclusions

We ran nine experimental markets to test Meadows' hog cycle hypothesis. In spite of an erroneous formulation of consumption in the experiment, it showed cyclical behavior of the same type as produces by Meadows' model and as observed in hog markets. The behavior was similar to that produced by Meadows' model when corrupted by the same error. We cannot reject Meadows' bounded rationality assumption that decisions to change the size of the livestock were influenced by adaptive price expectations. A linear regression produces a slope coefficient similar that that estimated by Meadows. Adjustments towards an implicit equilibrium indicate learning over time, which is consistent with previous studies that have shown significant learning effects in repeated Cournot experiments (Milgrom and Roberts, 1991). The learned equilibrium price relative to the Nash equilibrium ends up similar to previous laboratory experiments.

The results suggest that lasting market instabilities can indeed be generated endogenously, with no exogenous influences. Hence, Meadows' model should be appropriate for testing stabilization policies. For instance, there may be a potential for policies aiming to influence hog farmer decision-making. Further experiments could be used to test this proposition and also to test effects of forward markets.

References

Akiyama, T., Baffes, J., Larson, D., Varangis, P. 2001. Commodity market reforms: lessons of two decades. World Bank regional and Sectoral Studies, Washington, D.C.

Arango, S., Moxnes, E. 2012. Commodity cycles, a function of market complexity?. Extending the cobweb experiment. Journal of economic behavior and organization 84 (1), 321-334.

Arango, S., Castañeda, J.A., Larsen, E.R. 2013. Mothballing in power markets: an experimental study. From regulation to competition. Energy economics 36, 125-134.

Carlson, J.A., 1967. The stability of an experimental market with a supply-response lag. Southern Economic Journal 23 (3), 305–321

Chamberlin, E.H. 1948. An experimental imperfect market. The journal of political economy, 16 (2) 95-108.

Chiba, L.I. 2014. Pig nutrition and feeding. Animal Nutrition Handbook. Auburn University - College of Agriculture.

Dean, G.W., Heady, E.O. 1958. Changes in supply response and elasticity for hogs. Journal of farm economics 40, 845-860.

Deaton, A., 1999. Commodity Prices and Growth in Africa. The Journal of Economic Perspectives,13(3): 23-40.

Deaton, A., Laroque, G. 2003. A model of commodity prices after Sir Arthur Lewis. Journal of Development Economics, 71: 289– 310.

Estabrook, B. 2015. Pig Tales: An Omnivore's Quest for Sustainable Meat. First edition. Norton and Company, Inc.

Fattouh, B. 2009. Basis Variation and the Role of Inventories: Evidence from the Crude Oil Market. Oxford Institute for Energy Studies. WPM 38.

Forrester, J. W. (1961). Industrial dynamics. Productivity Pr; Student Edition.

Forrester, J. W. (1968). Principles of systems. Cambridge, EUA : Wright-Allen.

Franses, P,H., van Oest, R. 2004. On the econometrics of the Koyck model. Econometric Institute Report 2004-07.

Frisch, R. 1933. Propagation problems and impulse problems in dynamic economics. Economic Essays In Honour of Gustav Cassel.

Harstad, R., Martin, S., Normann, H.T. 1998. Experimental tests of consciously parallel behaviour in oligopoly. In L.Philips, ed., Applied Industrial Organization, Cambridge: Cambridge University Press.

Huck, S., Normann, H.T., Oechssler, J. 2000. Does information about competitors' actions increase or decrease competition in experimental oligopoly markets?, International Journal of Industrial Organization, 18:39–57.

Huck, S., Normann, H.-T., Oechssler, J. 2002. Stability of the Cournot Process: Experimental Evidence. International Journal of Game Theory, 31(1): 123-136.

Huck, S., Normann, H.-T., Oechssler, J., 2004. Two are few and four are many: number effects in experimental oligopolies. Journal of Economic Behavior and Organization 53 (4), 435–446

Klein, L, R. Marquez, J. 1989. Economics in Theory and Practice: An Eclectic Approach.Kluwer Academic Publishers.

Meadows, D. L. 1970. Dynamics of commodity production cycles. Cambridge, Mass., Wright- Allen Press.

Milgrom, R. J., Roberts. 1991. Adaptive and Sophisticated Learning in Normal Form Games. Games and Economic Behavior, 3, 82-100.

Miller, R.M., 2002. Experimental Economics: How we can Build Better Financial Markets. John Wiley and Sons, Hoboken, New Jersey

Nerlove, M. 1958. Adaptive expectations and cobweb phenomena. Quarterly Journal of Economics, 72, 227–240.

Plott, C.R., Smith, V.L. (Eds.), 2008. Handbook of Experimental Economics Results. North-Holland, Amsterdam.

Rodrigues, C. 2013. The Phenomenological Value of a "Vagabond" Approach to Environmental Education. INTED2013 Proceedings: 929-934.

Sonnemans, J., Hommes, C., Tuinstra, J., Velden, H. 2004. The Instability of a Heterogeneous Cobweb Economy: A Strategy Experiment on Expectation formation. Journal of Economic Behaviour and Organization, 54 (4): 453-481.

Smith, V.L., Suchanek, G.L., Williams, A.W. 1988. Bubbles, crashes, and endogenous expectations in experimental spot asset markets. Econometrica, Vol. 56, No. 5, 1119-1151.

United Nations conference on trade and development and Arbeiter kammer W .2011. Study prepared by the secretariat of the United Nations Conference on Trade and Development, New York and Geneva.

Appendix A: Equilibria derivation

BS is Breeding Stock (Variable)

LPH is Litters per hog (Constant and equal to 0.17)

PSL is Piglets save by litter (Constant and equal to 7)

WSF is Weaning survival factor (Constant and equal to 0.7)

LW is the hogs live weight (Constant and equal to 240 pounds or 108.86 in Kg)

DY is the hogs dressing yield (Constant and equal to 0.58)

$Hog\ Cost$ is hogs individual cost (Constant and equal to 11.84)

EC is the market expected consumption (Constant and equal to 1000 pounds or 453.59 in Kg)

Prerequisites

By assuming the market is in equilibrium, we can express the following relationships in the supply side:

$$Hogs\ slaughtered = \frac{MatureStock}{2} + \frac{BreedingStock}{36}\ (1)$$

$$Maturation\ rate = \frac{MatureStock}{2}\ \ (2)$$

$$Maturation\ rate = BS * LPH * PSL * WSF\ (3)$$

By replacing 3 in 2, we have:

$$Mature\ Stock = 2 * BS * LPH * PSL * WSF\ (4)$$

$$Hogs\ slaughtered = BS * LPH * PSL * WSF + \frac{BS}{36}\ (5)$$

Since the market is assumed to be in equilibrium, we can also establish the following relationships on the demand side:

$$Inventory = Pork\ production\ (6)$$

$$Pork\ production = \ Hogs\ slaughtered * LW * DY\ (7)$$

By replacing 5 into 7, we have

$$Pork\ production = \left(BS * LPH * PSL * WSF + \frac{BS}{36}\right) * LW * DY\ (7)$$

$$Inventory = \left(BS * LPH * PSL * WSF + \frac{BS}{36} \right) * LW * DY \ \ (8)$$

Joint Maximization equilibrium

Farmers' profits are determined by the difference between the hog price and the hog cost, multiplied by the number of hogs that were slaughtered in the farms.

$$Profits = Hogs\ slaughtered * (Hog\ price - Hog\ Cost)\ (9)$$

$$Hog\ price = \frac{-15{,}54 * \left(BS * LPH * PSL * WSF + \frac{BS}{36} \right) * LW * DY}{EC * DC} + 31{,}81\ (10)$$

By replacing 10, 5 and including the value for $Hog\ Cost$ in 9, we have

$$Profits = \left(BS * LPH * PSL * WSF + \frac{BS}{36} \right) * \frac{-15{,}54 * \left(BS * LPH * PSL * WSF + \frac{BS}{36} \right) * LW * DY}{EC * DC} + 31{,}81 - 11{,}84 \ \ (11)$$

By deriving 11 with respect to BS to find the maximum value for the expression, we have

$$\frac{dprofits}{dBS} = \left(BS * LPH * PSL * WSF + \frac{BS}{36} \right) * \frac{-15{,}54 * \left(LPH * PSL * WSF + \frac{1}{36} \right) * LW * DY}{EC * DC} + \left(LPH * PSL * WSF + \frac{1}{36} \right) * \left(\frac{-15{,}54 * \left(BS * LPH * PSL * WSF + \frac{BS}{36} \right) * LW * DY}{EC * DC} + 31{,}81 - 11{,}84 \right) = 0\ (12)$$

By replacing all the constants values in equation 12, we can know the value for BS that corresponds to Joint maximization equilibrium

$$BS = 1{,}9305\ (Millions\ of\ hogs)$$

Nash Equilibrium

For this equilibrium we assume that, by definition, every market actor maximizes his own profit assuming that the other players are going to do the same (Best response). Therefore, the total market variation is the result of individual maximization. In mathematical terms this is represented by

$$\frac{dBS}{dbs} = \frac{dbs}{dbs} = 1$$

Where bs is the Breeding stock of one player. Since we are considering 6 in the market we have that

$$BS = 6 * bs \ (13)$$

By replacing 13 in 11, we have

$$Profits = \left(6 * bs * LPH * PSL * WSF + \frac{6 * bs}{36} \right) * \frac{-15{,}54 * \left(6 * bs * LPH * PSL * WSF + \frac{6 * bs}{36} \right) * LW * DY}{EC * DC} + 31{,}81 \ - Hog\ cost\ (14)$$

By rearranging 14, we have

$$Profits = 36bs^2 * \left(LPH * PSL * WSF + \frac{1}{36}\right)^2 * \frac{(-15{,}54 * LW * DY)}{EC * DC} + (31{,}81 - Hog\ cost) * 6 * bs * \left(LPH * PSL * WSF + \frac{1}{36}\right)\ (15)$$

Where

$$36bs^2 = 6 * bs * 6 * bs = 6 * bs * BS \quad (16)$$

By deriving equations 15 and considering the best response requisite for Nash equilibrium, we have

$$\frac{dProfits}{dbs} = 42 * bs * \left(LPH * PSL * WSF + \frac{1}{36}\right)^2 * \frac{(-15{,}54 * LW * DY)}{EC * DC} + (31{,}81 - Hog\ cost) * 6 * \left(LPH * PSL * WSF + \frac{1}{36}\right) = 0\ (17)$$

By finding the value for bs in equation 17, we have that

$$bs = 0{,}5510$$
$$BS = 6 * bs = 3{,}3096\ (Millions\ of\ hogs)$$

Perfect Competition Equilibrium

For the players to not earn above the normal profit, the following condition must be satisfied

$$Hog\ price - Hog\ cost = 0 \ \ (18)$$

Therefore, by replacing 10 in 18 we have

$$\frac{-15{,}54 * \left(BS * LPH * PSL * WSF + \frac{BS}{36}\right) * LW * DY}{EC * DC} - Hog\ cost = 0\ (19)$$

By finding the value for BS in 19, we have that

$$BS = 3{,}8610\ (Millions\ of\ hogs)$$

Equilibria summary

	BS (Millions of hogs)	Number (in millions) of hogs in the farms	Profits
Joint Maximization	1.93	5.15	16.59
Nash equilibrium	3.31	8.82	8.12
Perfect Competition	3.86	10.29	0.00

Appendix B: Instructions [4]

Welcome! In this experiment you will play the role of a pig farmer. Every third month you will make a decision that influences how many pigs you will have ready for slaughtering at a later point in time. Your farm is one of six identical farms that supply hogs to the slaughtering houses. Your goal as a manager is to maximize your farm's accumulated profits over a 16-year period. Your payoff depends on the accumulated profits and can range from NOK 40 to NOK 500.

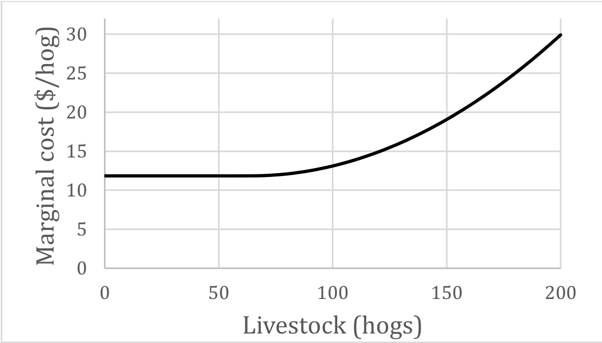
To help you manage, take a look at the computer screen and note the following information. On the left-hand side you find information about your own farm. The first item is the number of livestock, which is the number of sows (female pigs) that can give birth to piglets (offspring). On average, each litter has 5.8 piglets (number of siblings each time a sow gives birth). Each sow gives birth every 10th month. The sows' productive life is 3 years, after which sows are sold to a slaughtering house.

The next item is the number of piglets up to the age of 10 months. Below that you see the number of mature pigs between 10 and 12 months old. When these pigs reach 12 months of age, they are either sold to a slaughterhouse or female pigs may become livestock. To simplify, only pigs that survive birth and breeding are counted for.

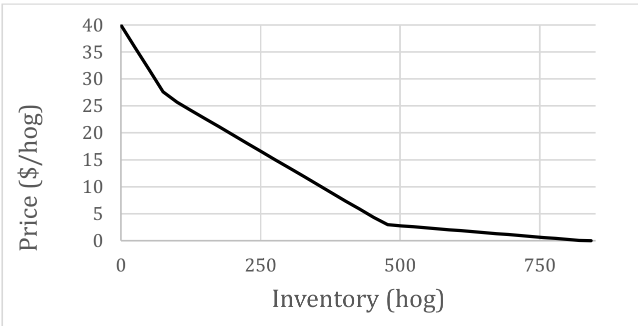
Then you get information about the number of pigs and livestock that are sold in the last three-month period. Selling is automatic and happens exactly when pigs reach the slaughter age. The marginal cost per pig increases with the number of pigs on your farm, the number you see is for the last three-month period. The per unit cost increases because your farm has limited room for pigs, and limited capacity for feeding and cleaning. The below graph shows how the marginal costs vary with the number of sows in the livestock.

---

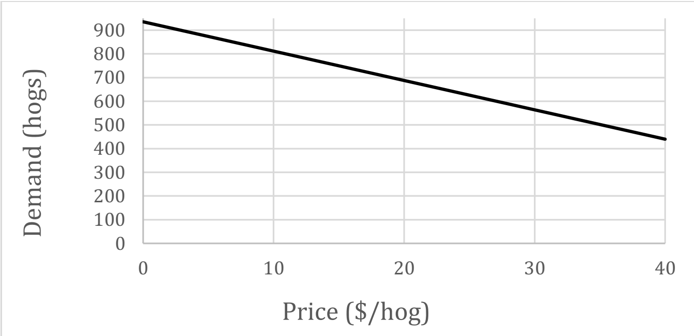
[4] We changed the term "hog" for "pig" and the term "breeding stock" for "livestock" to make the instructions easier to understand, given that most of our participants were not native English speakers.



The rectangle with market information shows market prices per pig for the last three months. The price varies with the number of slaughtered pigs that the slaughter houses have in their inventories. When inventories are nearly full, prices are low. This stimulates consumption of pork (pig meat) and help reduce inventories. When inventories are low, prices are high and reduce the demand for pork. The below graph shows the exact relationship between inventory and price in this market.



There is an immediate effect of price on consumption. The figure below shows the assumed relationship between hog price and demand (a linear demand curve).



Price minus marginal unit cost per pig gives the unit profit per pig sold. Total profits for a sale is given by the average unit profit times the number of pigs sold. The last piece of farm information shows the total profits earned over the last three-month period.

Your decision is to set the desired number of livestock. Once you set the desired livestock, it will take on average five months before the livestock reaches the desired size and the sows begin to produce piglets. It also takes time to reduce the livestock because pregnant sows will not be slaughtered before they have given birth. You can set a desired livestock from 0 to 200 pigs.

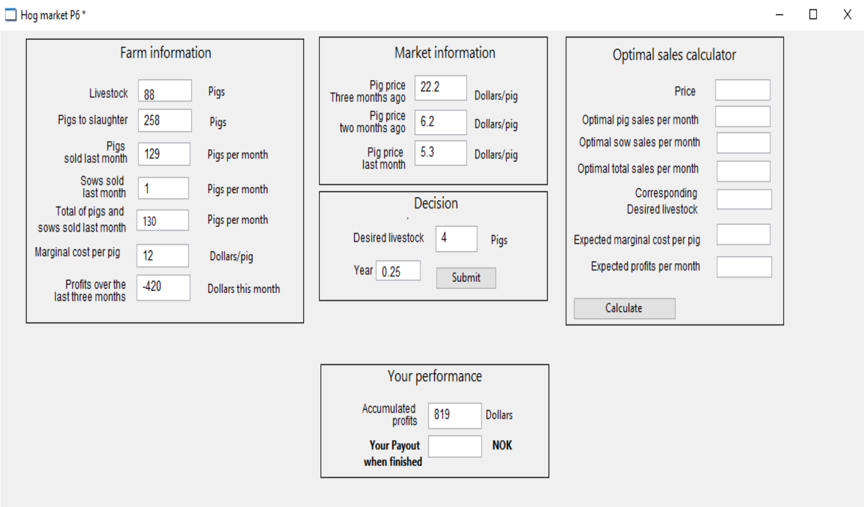
Below the rectangle for decisions you see the accumulated profits for all years. It is the accumulated profits in the last year that determines your payoff. Time is denoted in years such that three months show up as 0.25 year.

On the right-hand side you see a tool that can help you make decisions about the size of the livestock. You enter an assumption about the future price and the tool calculates the profit maximizing sales from your farm. This calculation takes account of the fact that marginal costs per pig rises with the number of pigs on your farm. As a further help, the tool also calculates the needed size of the livestock to reach the optimal sales numbers. Once you have entered a new assumption about the future price, click on the button “Calculate” to see the new recommendations.

Note that the recommendations you receive reflect your own assumptions about what the future price will be, which in turn depends on how many pigs you and your competitors sell to the slaughter houses.

Please use the answers sheet (columns for time period and desired livestock) to record your desired livestock every time.

Appendix C: Experiment interface [5]



5 We changed the term "hog" for "pig" and the word "breeding stock" for "livestock" to make the experiment more understandable, since most of our participants were not native English speakers.

Paper 4

Making climate conferences more effective?

Making climate conferences more effective?

Erling Moxnes
David Lara-Arango
System Dynamics Group, University of Bergen, Norway.

Abstract

Climate conferences (COPs) over the years have not led to sufficient emission reductions to reach generally agreed upon upper limits for climate change. According to mainstream game theory and public good and bad experiments, this should not come as a surprise. Just like individuals, nations are thought to be selfish. Here these results are challenged by a novel design of a public bad experiment. Interestingly, the new design leads to relative emission reductions similar to those obtained in Paris and much larger than predicted by previous games. This calls for some optimism. The results also point to possible improvements in COP procedures and activities. For instance, contrary to most earlier studies we find that individual pledges (IP), as used in the recent COP 21 in Paris, lead to just as strong agreements as negotiations (NG), as used in earlier COPs.

Keywords: Climate conferences, COP 21, Kyoto protocol, free riding, emission quotas, voluntary pledges.

1. Introduction

Experience thus far suggests that negotiations (NG) do not work well for the climate problem. The Kyoto protocol was a weak agreement among a limited number of nations. In follow-up conferences after Kyoto, NGs failed to establish new and stronger protocols. Major obstacles have been a bias towards quantity commitments (Cramton et al., 2015; Stiglitz, 2015) and consequent disagreements on quotas for different parties and on principles for setting such quotas (Depledge, 2000). Poor results are discouraging since NGs with conditional cooperation (I will, if you will) are typically seen as necessary to ensure socially optimal outcomes in commons problems (Cramton et al., 2015; Hauser et al., 2014).

Individual quantity pledges (IP) have been received with marked skepticism (Cooper, 2010; Cramton et al., 2015; Gollier and Tirole, 2015; Stiglitz, 2015; Weitzman, 2015). The main reason for such skepticism is the significant incentive for countries to free ride. Support for this position comes from game theory and public bad experiments showing that unilateral altruism is not sufficient to solve commons problems, even though contributions typically exceed what is predicted by the free rider hypothesis (Inman, 2009; Marwell and Ames, 1981). In spite of this skepticism, before the Paris COP started, "more than 180 countries producing more than 90 percent of global emissions had submitted intended nationally determined contributions (INDCs), a much broader response than many had anticipated" (Solutions, 2015). However, contributions were still much lower than needed to prevent average global temperature from exceeding 2ºC.

Game theoretical arguments supported by public bad experiments are at the core of discussions about COP procedures. In light of the relatively successful use of pledges in the Paris COP, it is important to take another look at the underlying theory. With this question in mind, we designed and used a novel laboratory experiment representing COPs of the NG and IP type. We introduced four major changes from the standard public good and public bad games.

1. Payoffs only depend on the agreement reached in the final round of the game, earlier rounds represent negotiations or announcements of pledges.

2. The social optimum is announced as a common goal, similar to the goal of 2°C. Individual players do not have to conceive of and figure out the social equilibrium on their own.
3. The social optimum is an interior solution reflecting the fact that all investments of the world are not needed to solve the climate problem.
4. Players have different payoff functions reflecting that climate and abatement costs differ between nations.

The results show much larger relative contributions than comparable public bad and public good games. Contrary to what has been claimed, there is no significant difference between average contributions for NG and IP. Towards the end, we discuss explanations for these findings and point to factors that could lead to even larger contributions towards limiting climate change.

2. A novel game

The game builds on the standard public bad game (Andreoni, 1995). Equation 1 shows the payoff function for player *i*

$$\pi_i = \propto_i A_i + (5.66 - 3.64 B_i/60) B_i - 1.16 \sum_{j=1}^{5} A_j \qquad (1)$$

where A_i is player *i*'s investment in project *A*. The last term sums up investments in project *A* for all five subjects. This sum determines the public bad, which reduces the payoff for all players. Hence, project *A* reflects economic activities that lead to emissions of GHGs and to climate change. B_i denotes investments in project *B*, which has no public bad effect and thus represents emission reductions.

To adapt this game to COP settings, payoffs are determined by the agreement reached in the final round rather than by the sum of results obtained over all rounds. To capture asymmetries among countries, players have different payoff functions (different $\propto_i$ coefficients) and these differences are private information. Still, all players get information about the social optimum, as if the goal for emission reductions is announced by an agency with knowledge of all payoff functions (e.g. representing the goal of limiting global warming to 2.0 ºC). To mimic that the social optimum does not require zero emissions and that all of society's investments are not needed for abatement, the social optimum is an interior solution with lower total investments than the total player endowments. Finally, treatments NG and IP differ in that to get an agreement with NG in any year, all players must accept the set of individual proposals. With IP, individual pledges in a non-specified final round become the agreement.

The experiment is similar to previous threshold public good experiments of climate change negotiations (Brick and Visser, 2015; Tavoni et al., 2011) except for having no climate change framing with country names, no predefined strategies to choose from, and no face-to-face communication between players. The NG treatment is also similar to a bargaining game (Özyurt, 2015) except for having identical roles (all players are investors rather than buyers and sellers) and no face-to-face communication among players.

All players have the same endowment of NOK 60. According to Equation 1, per unit return on investments in project B decreases with increasing investments in B. When $B_i = 0$, the per unit return is NOK 5.66; when $B_i = 60$, the per unit return is NOK 2.02. This assumption reflects increasing marginal costs of abatement.

The per unit return in project A varies between players and reflects differences across nations with respect to costs of climate change as well as of abatement (Gollier and Tirole, 2015). Variation is limited and is meant to reflect only real differences that are not known to everyone. Differences that are known and acknowledged as acceptable reasons for different contributions are not captured under the assumption that such differences would be more easily adjusted for in agreements. The $\propto_i$ coefficient is 6.39 for player 1, 6.52 for player 2, 6.79 for player 3, 7.07 for player 4 and 7.20 for player 5. These parameter values imply a social optimum where the sum of player investments in project B equals 191.4. The optimal average investment is 38.28 per player, well below the endowment of 60. Figure 1 shows how payoffs vary for the median player as a function of investments in project B.

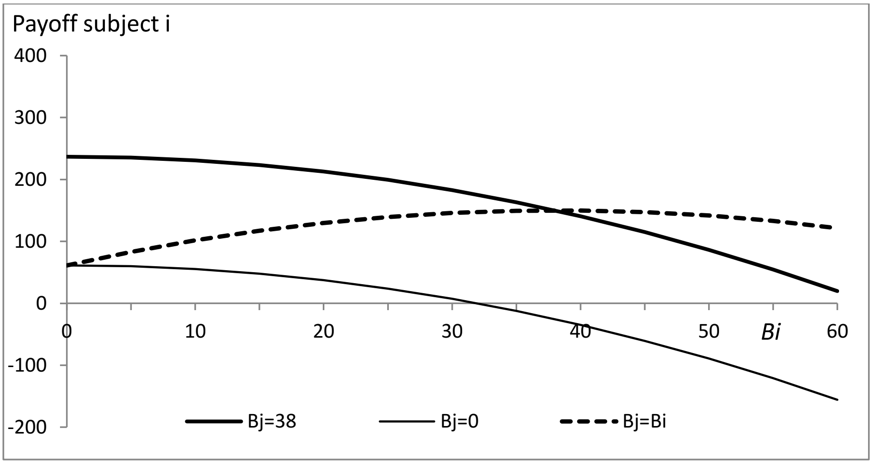


Figure 1: Payoff as a function of investments in project *B* (abatement), for subject *i* with the median payoff function. The dashed line shows payoffs if all players invest the same amount in project *B*. Payoffs are maximized when the player with the median value of $\propto_i$ invests NOK 38 in project *B* (social optimum). Solid lines show that individual payoffs are maximized for zero investments in project B (Nash equilibrium) no matter what the average investment is of the other four players B_j (38 or 0). Hence, there are incentives to free ride as well as to cooperate.

The experiment's two treatments capture two different procedures to reach an agreement, Kyoto negotiations (NG) and Paris individual pledges (IP). The Kyoto procedure is mimicked in the experiment by having each player propose an investment in project *B* for herself. When all players have made a proposal, all five proposals become common knowledge and each player can choose to say yes or no to an agreement based on the current proposals. If all say yes, negotiations end and the individual payoffs are determined by that agreement. If at least one player says no, the negotiations move on to a similar next round. In the 9th and 10th rounds the subjects get warnings that they are respectively in the second to last and in the last round. If they fail to reach an agreement in the last round, the result is zero investments in project *B* by all players, i.e. the Nash equilibrium. There is no verbal communication.

Different from the Kyoto protocol, players in NG cannot leave the negotiations. However, in practice, by proposing zero investments in project *B*, one player can shift the burden of finding an agreement to the remaining four. In the final around, a subject that invests zero in project *B* should have no incentive to vote no to any agreement that the other four come up with. Hence, even if free riders cannot leave negotiations, they cannot prevent agreements among those that choose to contribute.

In the IP procedure, nations are free to state their own national goals for emission reductions. In practice, nations made pledges over a several month-long period preceding the conference in Paris. In the experimental treatment, subjects are asked to announce pledges regarding own investments in project B. After each round, all pledges become common knowledge. Different from the NG treatment, players do not vote yes or no to the current pledges. Rather, they are invited to revise their own pledges in a new round, and so on. At the very beginning they are told that there will be between six and eleven rounds and that the experiment will stop without any last round warning. This should reduce the temptation to try to fool the others by proposing large investments in project B in all rounds except for the last one.

There were 5 subjects in each group, and the groups were formed randomly. With some exceptions, players met different subjects in the first and the second treatment. In no case, subjects could find out who they were playing against and to what extent group memberships had changed. Each player had the same payoff function in both treatments with one and the same □-value. Group membership stayed the same over all rounds within each treatment. Subjects participated twice, once with the NG treatment and once with the IP treatment. There were 95 subjects forming 19 groups playing the NG treatment and 85 subjects and 17 groups playing the IP treatment. Twenty groups played NG first and 16 groups played IP first.

Players read the written instructions (Appendix A), they were given an introduction to the experiment interface (Appendix B), and were encouraged to test out a payoff calculator before the experiment started. After each round, players could see all individual investment proposals or pledges, the total group investment, and their own payoffs. Players were privately paid at the end of the experiment. Average subject payoff was NOK 213 (USD 25) for two treatments.

The first half of the experiments were run about a month before the Paris COP started. The second half was run one and a half year later. The neutral investment wording in the experiment should prevent associations to climate conferences. Participants were recruited among Master level students, 50 Norwegian students studying economics and 40 studying system dynamics; the latter students had varied backgrounds at the Bachelor level and came from different countries.

3. Results

Summary statistics in Table 1 show average last round group contributions of respectively 38 and 44 percent of the social optimum for NG and IP treatments. Players capture 70 and 72 percent of the payoff for the social optimum in respectively NG and IP.

Table 1: Summary statistics. Contributions and payoffs refer to last round contributions.

	NG	IP
Number of groups	19	17
Groups with final agreement	15	17
Contribution in Nash equilibrium (NOK)	0	0
Contribution in social optimum (NOK)	38	38
Average contributions (NOK)	14.6	16.8
- as percentage of social optimum	38%	44%
- as percentage of endowment	24%	28%
Optimal payoff per treatment (NOK)	150	150
- average payoff relative to optimal	70%	72%

In each of four groups in the NG treatment, one of the players voted no in the last round leading to zero contributions (Nash equilibrium). The four players, who voted no, were the ones with the highest proposals in their respective groups in the last round. Three of those who voted no gained in payoffs, the fourth lost, however, much less than the others in the same group. If the four players had voted yes, average last round contributions in NG would have increased from 38% to 47%. The frequency of last round individual proposals and pledges equal to zero were respectively 12% and 19% in NG and IP.

Figure 2 shows the number of players agreeing to the set of proposals in each round. There is a sharp increase in the willingness to agree in the last round. This resembles the tendency in bargains for agreements to be reached as time is about to run out (Ståhl, 1972).

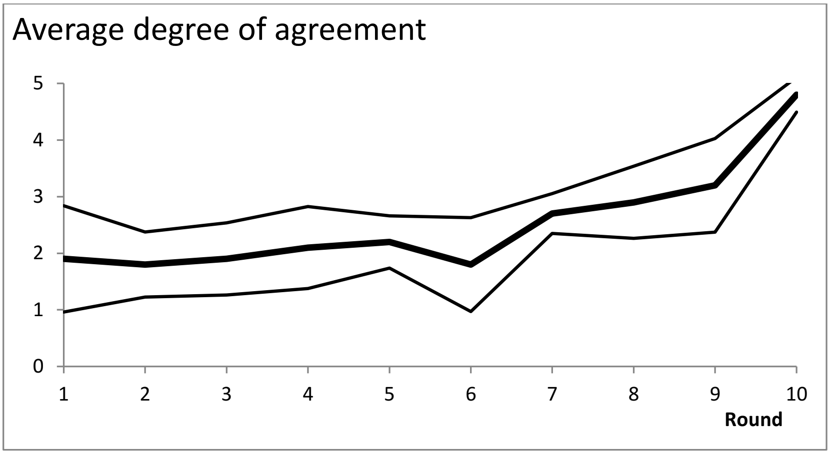


Figure 2: Number of players agreeing on set of proposals, average and 95% confidence interval for average. Data are only available for the experiments carried out in 2017, N=10.

Table 2 shows the results of multivariate regressions were the contributions in the last round is the left-hand side variable, and the right-hand side variables are dummies for treatment NG, first treatment, players being economics students, and player numbers indicating the effects of different cost parameters $\propto_i$.

Table 2: Coefficients and p-values for linear regressions of investments in project *B* in final agreements, N=180. The degrees of freedom have been cut in half to account for the fact that each student participated in both treatments.

	NG	First treatm.	Economics	Player number	Constant
Coefficient	-2.16	0.71	- 1.43	-0.39	18.4
p-value	0.30	0.73	0.50	0.59	$3*10^{-8}$

The constant is highly significant, indicating that the overall average contribution is much larger than the Nash equilibrium of zero. There is no significant difference between the two treatments. A within subject comparison also fails to show a significant difference between the treatments (p=0.12). There is no significant effect of in the order of the treatments. The tendency is in the direction found in public good and bad games where contributions tend to be higher in the first than in the ensuing treatments (Moxnes and Van der Heijden, 2003). There is no significant difference between contributions by economics students and students with mixed backgrounds. The tendency is in the same direction as found in many other studies (Frank et al., 1993).

A clustering of errors analysis (N=29) shows that players starting with NG and continuing with IP tend to increase their contributions in the second treatment (intra-class correlation of 0.025, p=0.48). Conversely, players starting with IP tend to make similar decisions in the following NG treatment (intra-class correlation of 0.622, p=0.009).

There is no significant effect of player number, and hence of payoff function parameter $\propto_i$ on contributions. The social optimum predicts that player 1 should contribute NOK 40.3 while player 5 should contribute only NOK 35.7. Hence if players behaved optimally, we should have obtained a regression coefficient of $(35.7 - 40.3)/(5 - 1) = -1.15$. While the obtained regression coefficient is considerably higher, -0.39, we note that this coefficient is not significantly different from -1.15 (p=0.3). Hence we can neither reject that players behave optimally nor that costs have no effect. Interestingly, if we consider the IP treatment only, this tendency is strengthened. The coefficient for player number is -0.82 (p=0.47), and hence it is less different from -1.15 (p=0.77). The NG treatment yields a coefficient for player number of 0.01 (p=0.99). The groups that did not reach an agreement cause most of the difference from IP.

Figure 2 shows how the average contributions tend to decrease over rounds in both treatments. In NG, the tendency is that contributions drop when it is announced that there are only two rounds left. In IP, contributions start to drop after period five, when the game may stop in any year. In NG, all agreements were made in the last period and the dot shows the average contribution for all groups whether they agreed or not.

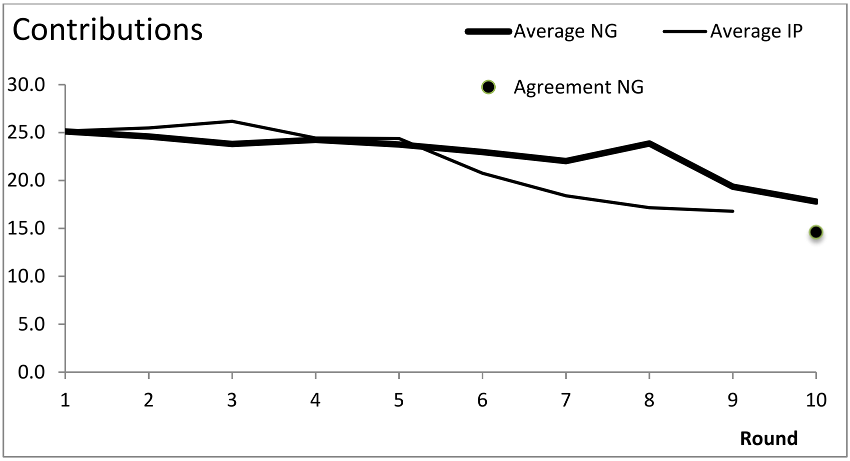


Figure 3: Average proposals for NG and pledges for IP over rounds.

Given literature on leadership in public bad games (Moxnes and Van der Heijden, 2003), it is interesting to explore how individual decisions in any round are influenced by what the other four players contributed in the preceding round. Pooling data for both treatments (N=1440, not including round 10 for NG), we tested a simple regression model for individual *i*'s investment in project *B* in round *k*.

$$B_k^i = 17.9 + 0.50B_{k-1}^{others} - 0.39S_{k-1} - 0.81k + e_k^i \qquad (2)$$

All parameters are established with very low p-values (all lower than 1.3 10^{-7}). First note that the constant is close to the one reported in Table 2. The second to last term picks up the downward trend in contributions over rounds *k*. Individual investments are positively correlated with the average investment for the other four players in the preceding round, B_{k-1}^{others}. Individual investments are negatively correlated with the standard deviation of investments in project *B* over all five players in the preceding round, S_{k-1}. A large standard deviation signals an unfair distribution and pulls individual contributions in the direction of free riding. The effect is considerable. By use of simulation[6] we find that if the standard deviation were reduced from its average of 11.0 to the standard deviation for the social optimum, 2.95, the last round average contribution B_9^i would increase from 17.5 to 23.7, an increase of 35%.

6 The simulation of Equation 2 starts out with the observed first period average for $B_1^{others} = 25.1$ and assumes that $B_{k-1}^{others} = B_{k-1}^i$ in ensuing rounds.

The experimental design with its NG- and IP-treatments differs from more standard public good and bad games, except for mostly modest differences with respect to group sizes and number of rounds. For a sample of studies that are likely to be quite representative, Table 3 shows relative average contributions in the range from Nash equilibrium (0%) to social optimum (100%). Considering averages over rounds, NG has 26 percentage-points higher contributions than the average for public good games and 47 percentage-points higher contributions than the average for public bad experiments. Corresponding numbers for IP are 23 and 44 percentage points. Considering last round agreements, NG contributions exceed the average for public good games by 13 percentage-points and the average for public bad games by 40 percentage-points. Corresponding numbers for IP are 19 and 45 percentage-points.

Table 3: Group size and number of rounds together with average over rounds and last year relative contributions (contribution minus Nash contribution as a percentage of social optimum minus Nash contribution). Numbers are derived from tables and graphs in the quoted publications.

	Andre-oni (1995)	Rege & Telle (2004)	Moxnes & Heijden (2003)[1]	Willinger & Ziegelme-yer (1999)	Dannen-berg (2015)[2]	Chen and Komorita (1994)[3]	NG[4]	IP[4]
Group size	5	10	5	4	4	5	5	5
Rounds	10	1	10	15	10	10	10	9
Public good								
Over rounds	34	34		41	35	31		
Last round	21			39	16			
Public bad								
Over rounds	16.2		18	9.7			61	58
Last round	1.0		3.4	-8.2			38	44

[1] First treatment and no leader condition.

[2] First treatment with no leader and no pledge condition (Ex-base)

[3] Nonbinding pledge condition

[4] Averages taken over 9 rounds for proposed contributions in NG and pledges in IP. Last round includes no agreements in NG.

4. Discussion

Two main findings come out of this study. First, the novel game leads to larger relative contributions than what is typical in public bad and public good games. This is an important finding that gives cause for some optimism. Contributions relative to the social optimum are not so different from what was achieved in the Paris COP where total contributions to emission reductions were about half of what seems needed to reach a stated goal of limiting warming to 2ºC.

Second, the experiment finds no significant difference between average contributions for the two procedures, negotiations (NG) and individual pledges (IP). However, the fact that NG may fail to produce a final agreement implies that the risk of zero contribution is larger for NG than for IP. This is consistent with unsuccessful COPs after Kyoto.

The two main results are more in line with experience than with standard theory. Hence, our novel experiment hints at how theory or models for investigation may be improved. First, traditional public bad games do not present a goal of reaching an explicit social optimum. The announcement of the goal creates a more noticeable and certain focal point (Schelling, 1960) than what players are likely to arrive at using otherwise available information and their own reasoning. Importantly, players are also likely to trust that the other players take notice of the same focal point. Previously, uncertainty about focal points in terms of thresholds has been found to reduce contributions (Barrett and Dannenberg, 2012).

Second, when focus shifts towards reaching an agreement, the public bad problem becomes more of a public good problem. An agreement is in itself a public good. It is know from before that public good framings lead to higher contributions than public bad framings (Andreoni, 1995; Sonnemans et al., 1998). However, this cannot explain why our experiment also achieved somewhat better results than previous public good experiments.

Third, for games to predict outcomes of COPs better, the social optimum should not require contributions to equal to the players' entire endowments. Other experiments show that excess cash can stimulate to larger investments than what is optimal (Caginalp et al., 2001). Similarly, when all of society's resources are not needed to limit global warming, it is easier to move away from the Nash equilibrium and closer to the social optimum. The less costly Montreal treaty to limit emissions of CFC gases illustrate this point.

Neither in the experiment nor in the Paris COP did pledges reach the social optimum. There is also uncertainty whether real emission reductions will match pledges in coming years, in spite of commitments to report emissions and to submit new pledges every five years. Future COPs must do better. Detailed experimental results point to potentials for further improvements in procedures.

Our experiment did not allow for face-to-face communication. Earlier studies (Hackett et al., 1994; Rege and Telle, 2004) show that face-to-face communication leads to considerable increases in contributions. Hackett et al. (1994) introduced face-to-face communication in an asymmetric public bad experiment with different player endowments (private information). Communication increased the average payoff by 33.5 percentage points (ibid, Table VII). Communication is also an important factor for successful management of local common-pool resources (Ostrom, 1990). Hence, if we had allowed for verbal face-to-face communication, our experiment would most probably have given even better results. In real COPs there is face-to-face communication. However, different from public bad experiments and local common-pool resources management, COPs operate with layers of people involved from COP negotiators to home country politicians and voters. While communicating negotiators seem to be more prone to cooperation than most people (Inman, 2009), home country voters and politicians may not be sufficiently involved in the communication to establish focal points and to build mutual trust. Hence, it is a challenge to establish communication between voters and politicians from different countries.

Communication also allows for leading nations to collaborate to establish thresholds by announcing large emission reductions conditional on other nations reciprocating. Thresholds can create equilibrium situations where self-interest lead to increasing contributions over time (Cadsby and Maynes, 1999). Similar to announcing a goal in terms of a social optimum, there may be need for information and awareness about constellations that could produce thresholds and new equilibria.

Still regarding communication, a benefit of IP in real COPs is that there is no need to agree on principles for how emission reductions should be allocated. This has been an obstacle in COPs preceding Paris (Depledge, 2000; Gollier and Tirole, 2015; Stiglitz, 2015) most likely caused by excessive self-interest (Brick and Visser, 2015). Relying on voluntary pledges is likely to reduce the conflict level considerably. In this connection it is interesting that we cannot reject the possibility that player contributions to some extent reflect private information about payoff functions. Hence, a procedure with pledges (IP) could lead to greater efficiency than the proportional reductions often observed in negotiated allocations among exploiters of common property resources (Hackett et al., 1994). The same is the case for negotiations (NG), provided final agreements rely on individual proposals and not on a principle of proportional reductions. Further research could shed light on how well Paris pledges did reflect public (and possibly revealed private) information about individual country "payoff functions".

The experiment shows that what each player contributes in one round is positively influenced by what others contributed in the previous round. Previous public bad experiments have found that when a leader sets a good example that has a positive effect on followers (Moxnes and Van der Heijden, 2003). Similar effects are also seen in field experiments (Frey and Meier, 2004). Moreover, leading nations not only influence the ambitions of other countries, they also have to invent and test out new abatement policies and to develop emission-reducing technologies. In turn, successful policies and technologies diffuse worldwide over time (Rogers, 1995). Information strategies are needed to get as much as possible out of all these leadership effects.

In follow-up COPs after Paris, actual emission reductions are likely to become more important as signs of leadership than pledges. The literature on diffusion shows that people are inspired more by practical experiences than by theory, where "theory" could be seen to include uncertain promises and pledges. Uncertainty typically slows down diffusion (Rogers, 1995). Already today, those who argue for emission reductions point to good practical examples in other countries. Therefore, it is essential for leading as well as other nations to deliver what they have pledged. For this reason it is important that nations do not underestimate the effort and time needed to reduce emissions of GHGs.

Underestimation of the length of delays could slow down implementation of policies (Sterman, 1989). So can failure to take account of delays when making decisions (Brehmer, 1989). A laboratory experiment where the goal was to meet the Kyoto requirements in a cost effective way, found delayed implementations of policies and consequent excessive costs (Moxnes and Assuad, 2012). These findings point to the need for information policies.

Our experiment shows that large standard deviations of contributions in one round have negative effects on contributions in the next round. Fehr and Gächter (2000) find that when free-riders can be punished directly, and not only indirectly by lowering contributions towards the common good, average contributions increase over rounds and start to approach the social optimum. Regarding the climate problem, groups of co-operators could punish free-riding nations by using tariffs to reduce imports of products with large carbon footprints. The potential for retaliation by punished free-riders must be considered (Fehr, 2000; Grechenig, 2010). To avoid unnecessary conflict, punishments should be reasonable and graduated (Ostrom, 1990). Some likely free-riders may be automatically punished. Provided global emissions of GHGs are reduced, reduced demand for fossil energy will force free-riding fossil energy exporting countries to reduce production and emissions.

Finally, there is one more deviation between the experiment and reality that needs attention. In the experiment, the social optimum is clearly defined. In reality, there is uncertainty and confusion about what the socially optimal global emission reductions should be. This uncertainty invites politicians and electorates to opt for wait-and-see strategies (Guy et al., 2013; Moxnes and Saysel, 2009; Sterman, 2011). This strategy plays down the importance of facts and analysis and is likely to lower pledges and actual emission reductions. Thus, there is need for media attention to the problem with wait-and-see strategies. In particular, the general population needs to understand that the current emission rate of GHGs is about twice as large as the removal rate and that it takes many years to reduce emissions towards or below the removal rate. When climate change has reached an unacceptable level, there is little one can do to prevent that it gets even worse.

References

Andreoni J (1995) Warm-Glow versus Cold-Prickle: The Effects of Positive and Negative Framing on Cooperation in Experiments. The Quarterly Journal of Economics 110:1-21.

Barrett S, Dannenberg A (2012) Climate negotiations under scientific uncertainty. P. Natl. Acad. Sci. U.S.A. 109:17372-17376.

Brick K, Visser M (2015) What is fair? An experimental guide to climate negotiations. Europ. Econ. Rev. 74:79-95.

Cadsby CB, Maynes E (1999) Voluntary provision of threshold public goods with continuous contributions: experimental evidence. J. Public Econ. 71:53-73.

Caginalp G, Porter D, Smith V (2001) Financial bubbles: Excess cash, momentum, and incomplete information. The Journal of Psychology and Financial Markets 2:80-99.

Chen X-P, Komorita SS (1994) The effects of communication and commitment in a public goods social dilemma. Organ. Behav. Hum. Decis. Process. 60:367-386.

Cooper R (2010) The case for charges on greenhouse gas emissions. Post-Kyoto international climate policy: Implementing architectures for agreement:151-178.

Cramton P, Ockenfels A, Stoft S (2015) An international carbon-price commitment promotes cooperation. Economics of Energy & Environmental Policy 4:51-64.

Dannenberg A (2015) Leading by example versus leading by words in voluntary contribution experiments. Soc. Choice Welfare 44:71-85.

Depledge J (2000) Tracing the Origins of the Kyoto Protocol: an Article-by-article Textual History. United Nations, UNFCCC Framework Convention on Climate Change, United Nations, UNFCCC Framework Convention on Climate Change.

Fehr E, Gächter S (2000) Cooperation and Punishment in Public Goods Experiments. Am. Econ. Rev. 90:980-994.

Frank RH, Gilovich T, Regan DT (1993) Does studying economics inhibit cooperation? The Journal of Economic Perspectives 7:159-171.

Frey BS, Meier S (2004) Social comparisons and pro-social behavior: Testing" conditional cooperation" in a field experiment. Am. Econ. Rev. 94:1717-1722.

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

Guy S, Kashima Y, Walker I, O'Neill S (2013) Comparing the atmosphere to a bathtub: effectiveness of analogy for reasoning about accumulation. Clim. Change 121:579-594.

Hackett S, Schlager E, Walker J (1994) The role of communication in resolving commons dilemmas: experimental evidence with heterogeneous appropriators. J. Environ. Econ. Manage. 27:99-126.

Hauser OP, Rand DG, Peysakhovich A, Nowak MA (2014) Cooperating with the future. Nature 511:220-223.

Inman M (2009) The climate change game. Nature Reports Climate Change:130-133.

Marwell G, Ames RE (1981) Economists free ride, does anyone else?: Experiments on the provision of public goods, IV. J. Public Econ. 15:295-310.

Moxnes E, Assuad S (2012) GHG taxes and tradable quotas, experimental evidence of misperceptions and biases. Environmental Economics 3:44-56.

Moxnes E, Saysel AK (2009) Misperceptions of global climate change: information policies. Clim. Change 93:15-37.

Moxnes E, Van der Heijden E (2003) The effect of leadership in a public bad experiment. J. Conflict Resolution 47:773-795.

Ostrom E (1990) Governing the Commons. Cambridge University Press, Cambridge.

Rege M, Telle K (2004) The impact of social approval and framing on cooperation in public good situations. J. Public Econ. 88:1625-1644.

Rogers EM (1995) Diffusion of innovations. The Free Press, New York.

Schelling TC (1960) The strategy of conflict. Harvard University Press, Cambridge.

Solutions CfCaE (2015) Outcomes of the U.N. climate change conference in Paris.

Sonnemans J, Schram A, Offerman T (1998) Public good provision and public bad prevention: The effect of framing. J. Econ. Behav. Organ. 34:143-161.

Sterman JD (2011) Communicating climate change risks in a skeptical world. Clim. Change 108:811-826.

Stiglitz JE (2015) Overcoming the Copenhagen failure with flexible commitments. Economics of Energy & Environmental Policy 4:29-36.

Tavoni A, Dannenberg A, Kallis G, Löschel A (2011) Inequality, communication, and the avoidance of disastrous climate change in a public goods game. P. Natl. Acad. Sci. U.S.A. 108:11825-11829.

Weitzman M (2015) Internalizing the climate externality: Can a uniform price commitment help? Economics of Energy & Environmental Policy 4:37-50.

Willinger M, Ziegelmeyer A (1999) Framing and cooperation in public good games: an experiment with an interior solution. Econ. Letters 65:323-328.

Özyurt S (2015) Bargaining, reputation and competition. J. Econ. Behav. Organ. 119:1-17.

Appendix A: Instructions

Overall instructions:

Two experiments.
You will participate in two experiments each in a group with four other players. The instructions on this page are the same for both experiments. You do not know whom you are participating with, and you will interact with different groups in the first and second experiment. Your decisions will determine how much you will be paid when the experiments are finished.

Your task in the experiment is to decide how much to invest in projects A and B. The total amount you can invest is 60 NOK. What you do not invest in project B will automatically be invested in project A; A = 60 - B. The following formula shows how much *you* will earn in NOK in each experiment.

$$Payoff = \alpha A + (5.7 - 3.6B/60)B - 1.2 * (Sum\ of\ all\ investments\ in\ A)$$

[The letter □ was replaced by the respective numbers 6.39, 6.52, 6.79, 7.07, and 7.20 for players 1 to 5.]

Investing all 60 NOK in project A and nothing in B gives the highest direct contribution to your payoff. However, your investments in A, and the four other players' investments in A, lead to costs for everyone. Hence, if all players invest nothing in B, that will be the worst outcome for everyone. Investing in project B has no negative effect for everyone. However, the direct return on investment is lower for project B than for A. Also, the more you invest in project B, the lower the return on each additional NOK invested in B. To be precise, the per unit payoff decreases from 5.5 to 3.5 as your investment in B increases from 0 to 60 NOK.

To help you calculate your payoff for different investments in B, you can use the payoff calculator that you see on the screen. You enter an investment in B for yourself, and you enter assumptions about investments in *B* for the four others players. The calculator gives slightly different payoffs than the formula because numbers in the formula are rounded.

Note that *all players in the group have different returns on investments in project A*; payoff formulas are different. This means that all players do not necessarily want to invest the same amount in project *B*. You do not know the returns on project *A* for the other players, and they do not know your return.

Using information that is not available to you or any of the four other players, it has been found that total payoffs are maximized for a total investment in *B* of 190 NOK, which is an average investment of 38 NOK per players. This maximum requires that investments in project *B* vary among players.

First/Second experiment [if NG-treatment]

In this experiment, you will enter a series of negotiations with the other four players. This works as follows: all players announce how much they are willing to reduce their investments in project *A* by proposing investments in project *B*, thus establishing a contract proposal. When all players have made a proposal, the proposals become common knowledge, and all players are asked to say *yes* or *no* to the current contract. If everyone says *yes*, negotiations stop and you will earn the contract payoff. If at least one player says *no*, the negotiation moves to the next round. Then you will be invited to make a new proposal, and the proposal process starts again and continues until you get an agreement or you reach the maximum number of rounds (10 rounds). If you do not reach an agreement by the tenth round, you get a payoff that corresponds to 0 investments in project *B* by all players.
During the experiment, only do what you are instructed to do. Do not talk to other players during the experiments, do not start any other program on your computer, and do not close down the experiment program. In case something should go wrong, for each round, write down the investment in project B that you propose in the table below. When finished, fold the paper so that nobody can see your proposals. Thanks for following these instructions the best you can!

Round number	Your proposed investments in project B
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	

Obtained payoff in this experiment: ___________________ NOK

Second/First experiment [if IP-treatment]

In this experiment, you will make a series of pledges (voluntary contributions) to the group. This works as follows: all players announce how much they are willing to reduce their investments in project *A* by proposing investments in project B, thus establishing a contract proposal. When all players have made a pledge, the pledges become common knowledge. Based on the current contract proposal, you are free to revise your pledge to the group by announcing another investment in project B. Revisions will go on for at least 6 rounds and possibly up to 11 rounds. None of the players know exactly when revisions of the contract will stop. Hence, after round 6 any pledge you make could be the binding one. You get a payoff that corresponds to the investments in project B by all players in the final round.

During the experiment, only do what you are instructed to do. Do not talk to other players during the experiments, do not start any other program on your computer, and do not close down the experiment program. In case something should go wrong, for each round, write down the investment in project B that you pledge in the table below. When finished, fold the paper so that nobody can see your pledges. Thanks for following these instructions the best you can!

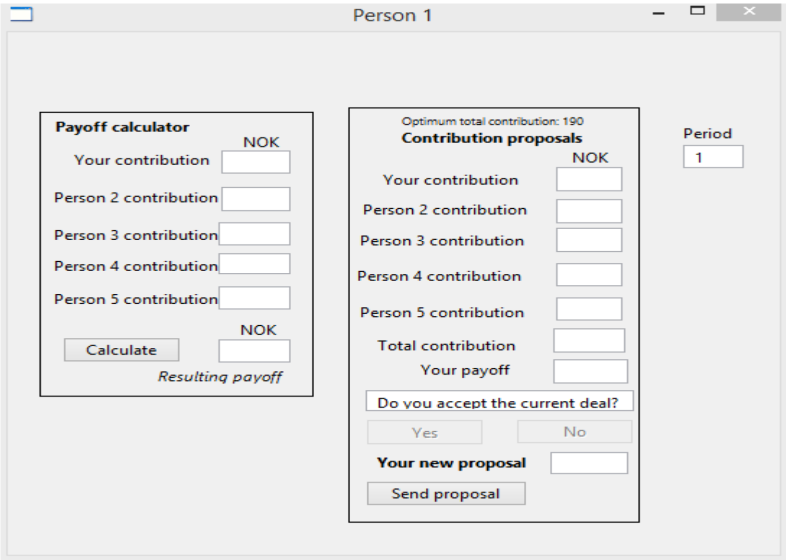
Round number	Your pledged investments in project B
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	

Obtained payoff in this experiment: ____________________NOK

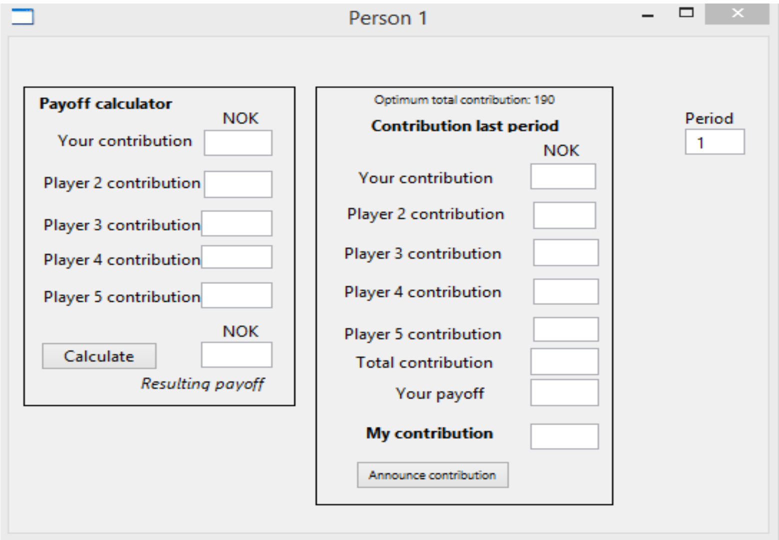
Appendix B:

The experiment was programmed in Livecode, with a program downloaded to each and every computer linking to a common server and database.

Interface NG treatment



Interface IP treatment



Paper 5

Socially aversive personalities and income distribution: Can the Dark Triad predict behavior in the dictator and gangster games?

Socially aversive personalities and income distribution: Can the dark triad predict behavior in the dictator and gangster games?

David Lara Arango
System Dynamics Group, University of Bergen, Norway.

Abstract

Finding reliable indicators for how humans make decisions is a subject of upmost importance. The Dark triad presents itself as a promising indicator of behavior, as it has been extensively suggested in the social psychology literature. People who score high in the triad components, namely Narcissism, Machiavellianism and Psychopathy, have been characterized as being impulsive, selfish and having a generalized lack of empathy. Similar to other indicators such as IQ, these personality traits have been able to predict behavioral feats in various contexts. However, it is unclear whether the Dark Triad can predict people's decisions in a non-cooperative setting, such as the dictator and gangster games. Moreover, it is unclear whether these personality features are sufficient to explain individual behavior in controlled environments, or whether they can be considered as a form of the fundamental attribution error. This paper proposes an experimental design to investigate this issue. The findings suggest that the Dark triad has no significant predictive power in the dictator and gangster games, thus supporting the hypothesis of the triad being a form of the fundamental attribution error.

Keywords: Dark Triad, fundamental attribution error, dictator game, gangster game, Income distribution.

1. Introduction

Finding reliable indicators for human decision making processes has been a problem many scholars have struggled with. In this sense, the Dark triad presents itself as yet another indicator for such decision making processes. Ever since Paulhus and Williams (2002) pointed out that Narcissism, Machiavellianism and Psychopathic traits were a set of variables that could explain socially aversive behavior (at a sub-clinical level), the literature on what thereafter was known as the Dark Triad has been growing quickly (Furnham et al, 2013; Jones and Paulhus, 2011a). Despite the fact that Narcissism, Machiavellianism and Psychopathic traits already had vast separate literatures when Paulhus and Williams (2002) published their work, the Dark Triad generated a substantial response in the academic community (Furnham et al, 2013). The reason behind such a response was that the three components of the Dark triad share the same behavioral core, namely a tendency towards callous manipulation, which at a sub-clinical level can make it difficult to differentiate one of the Triad's component from the other two. Therefore, one of Paulhus and Williams (2002) main contributions was to present a method to better triangulate callous personalities by jointly using Narcissism, Machiavellianism and Psychopathic traits, rather than having three isolated components.

Having the Dark triad as a method to define callous personalities and socially aversive behavior has allowed researchers to find interesting insights about different human behavior in occupational, educational and interpersonal settings (Wiggins and Pincus, 1989; Furnham and Crump, 2005). In occupational settings, the Dark triad has been used to study cases of bad leadership, career success and manipulation of coworkers (Paulhus and Buckels, 2011). In educational settings, it has been used to predict cheating and plagiarism among students (Nathanson et al., 2006; Williams et al., 2010). In interpersonal settings, it has been used to study different issues such as prejudices formation, social dominance orientation, cynicism, among others (Arvan, 2012; Rauthmann, 2012).

As a predictor, the Dark triad presents yet another alternative to anticipate human behavior and performance, much like other predictors such as the IQ, which has been linked to job performance (Schmidt et al, 1998), academic performance (Frey et al.,

2004) and income (Murray, 1998). In this respect, recent studies suggest that the Dark triad can predict counterproductive leadership skills (Furnham, 2010), callous interpersonal behavior (Rauthmann and Kolar, 2012), antisocial behavior (Baughman et al., 2012) and academic entitlement (Turnipseed and Cohen, 2015). Given this predictive power, a natural question to ask could be, can the Dark triad also predict other types of entitlement, such as income entitlement? If so, then the Dark triad can be an interesting set of predictive variables to study how people decide in non-cooperative games that involve income distribution. The link between such a set of variables and behavior in non-cooperative income distribution games has not been established. This paper explores this issue by conducting an experiment consisting of the dictator and the gangster game (inverted dictator game) along with the Dirty dozen questionnaire (Maples et al, 2014) to capture participants' Dark triad scores.

The rest of the paper is organized as follows: the next section provides a brief definition of the three components of the Dark triad, the third section describes our experimental design, the fourth section explains our hypotheses, the fifth section presents our results and finally, the last section discusses our findings and concludes.

2. The Dark Triad components

As was previously mentioned, the three components of the Dark triad have been the subject of a vast number of separated studies. This section succinctly describes each of them.

Narcissism:

Narcissism is often defined as compensatory self-promotion, characterized by grandiosity and attention seeking behavior (Morey et al., 2012). In other words, Narcissism can be understood as an exaggerated sense of self-importance (Miller & Campbell, 2011)

Machiavellianism:

This psychological trait has been typically defined as having a lack of morality, a cynical world view and a tendency to manipulate others (Fehr et al, 1992). A more recent definition of Machiavellianism also accounts for the ability to build reputation, form alliances and plan ahead; critical factors to clearly distinguish this trait from Psychopathy, another component of the Dark triad (Jones and Paulhus, 2009).

Psychopathy:

The classical definition of psychopathy is a self-control deficit combined with callousness (Cleckley, 1941). Unlike Machiavellian individuals, psychopaths have a tendency to act recklessly. In other words, psychopathy refers to the extent to which one can disregard other's well-being when making impulsive decisions (Hare and Neumann, 2008).

3. Experimental design

This paper uses the dictator and gangster games. The gangster game is included in order to explore whether the gangster game setup could be better predicted by the dark triad when compared to the dictator game. Distinctions between the likely behaviors associated with the two games have been pointed out in the literature. Previous research has shown a significant change in the dictator's behavior if he is given the possibility to take money instead (or besides) the possibility to give money. Some findings suggest that if one gives a wider options range to the dictator, the amount of money he will give to the recipient will be significantly smaller than the amount he would give if his only option was to give (Cappelen et al, 2012). However, one interesting question still remains; what happens when the number of options is not changed but the framing is changed, i.e. what happens when people can either give or take money. This question is addressed by comparing individuals' behavior in the traditional dictator game with the behavior obtained in a gangster game. This inverted dictator game has also been called the gangster game (Eichenberger and Oberholzer-Gee ,1998). In the gangster game, an endowment is given to a player, and a second player is asked how much of such endowment he wants to take from the first player. The dictator and the gangster games are theoretically the same, given that a fully rational player will choose to give zero and to take everything, which yields the same profit under the assumption that the endowments is both games are the same.

Previous studies have shown differences between participants' behavior in the dictator and gangster games. One example is Eichenberger and Oberholzer-Gee (1998). In this study, the authors study how fairness considerations affect the decision making of both dictators and gangsters. They found that gangsters ended up leaving less than the dictators gave. Another example can be found in Bardsley (2008). The aim of this study is to see what the effects of having asymmetric endowments in both games are. Bardsley's results suggests that the proportion of subjects willing to give in the dictator game was higher than the proportion of people willing to leave money in the gangster game. According to these two studies, people in the gangster game seem to have the tendency to be more selfish than people in the dictator game.

Experimental procedure:

We developed four computer interfaces for the dictator game and the gangster game (Appendix A). One for the dictator in the dictator game, one for the recipient in the dictator game, one for the gangster in the gangster game and one for the victim in the gangster game. In the instructions (Appendix B) the dictators were told they were given an endowment of 60 NOK (6.9 USD approximately), of which they had to decide how much to share with a second player, who had 0 NOK. They were also told that the second player knew how much their endowment was. The recipients were told they had 0 NOK and that there was another player who had received an endowment of 60 NOK and was entitled to share a fraction of such money with them. In a separate sheet of instructions, the gangsters were told that a second player was given an endowment of 60 NOK and they had to decide how much of those 60 NOK they wanted to take for themselves. The victims were told they had received an endowment of 60 NOK and that there was another player who had received 0 NOK and could decide how much of the 60 NOK he wanted for himself.

We ran two experimental sessions with 15 groups of two people in one and 13 groups of two people in another one. The subjects were System Dynamics and Economics master students from the University of Bergen, Norway. All subjects were recruited from classes. In each session, we ran both the dictator and the gangster game, that is, we used a within-subject experimental design. Since we used a within-subject variation design, 14 out of the 28 groups in did the dictator game first and then the gangster game while the remaining 14 groups started with the gangster game and then continued with the dictator game. Participants were not allowed to take part in more than one session. Subjects were told that they would be paid privately the amount of money they had after the game was over.

Upon arrival, participants were randomly seated behind computers in cubicles. The experiment was designed such that couples (dictator-recipient, gangster-victim) were randomly assigned and each participant could not identify her counterpart. The within-subject design ensured that the ones who were dictators (gangsters) in the first round were gangsters (dictators) in the second round. The instructions were distributed among the participants and they were told they had 10 minutes to read the instructions and ask questions about them. All information was common knowledge.

Participants were playing on linked computers, once the experiment started, the dictators had to type the amount of money they wanted to share with the recipient (any amount from 0 to 60 NOK in integer numbers, was allowed), and the gangster had to type how much of their victims' endowment they wanted for themselves (any amount from 0 to 60 in integer numbers, was allowed). Participants were asked to press their respective interface buttons (see the interfaces in Appendix B). First, the dictators and gangsters typed their decisions and pressed their interfaces' buttons, thus sending their decisions to their respective recipients and victims. Recipients and victims were then asked to press the button "Open" to see what their counterparts had given (or left) to them. After having filled in the Dirty dozen questionnaire (Appendix C) proposed by Maples et al (2014) they were privately paid the amount of money they had when the experiment ended.

4. Hypothesis

This study's null hypothesis states that the Dark triad has no significant predictive power in control environments, given that it may not be character (or personality) that determines participants actions in the dictator and gangster games, but rather participants' circumstances. Previous studies have argued that character is not a good predictor of people's actions. In fact, the majority of people are more influenced by their circumstances (or rather the way they perceive their circumstances) than they are by their characters (Harman, 1999). Failing to account for people's circumstances and assuming that their actions are only a direct reflection of their character is often referred to as the fundamental attribution error in the scientific literature (e.g. Ross, 1977; Flanagan, 1991; Nisbett and Ross, 1991). While both games present particular circumstances for participants, they are both fairly abstracted from real life situations. In this sense, it is mostly participants' character that would define how they decide in either game. Being a character assessment system, one could expect the Dark triad to be a weak predictor of people's behavior.

The alternative hypothesis states that subjects with the highest score in the Dark Triad test will be the ones giving the least in the dictator game and leaving the least in the gangster game. This hypothesis is based on previous findings of the Dark triad, suggesting that subjects that score high on the Dark triad test are more aggressive when seeking their own interest given their lack of empathy for others (Jonason and Krause, 2013), In addition, subjects who score high in the Dark triad have also shown higher impulsivity (Jonason and Tost, 2010) and a tendency towards risk-taking behavior (Adams et al, 2014), which, combined, constitute an ego-satisfying tendency that is consistent with higher entitlement and higher selfishness (Campbell et al, 2004; Jones and Paulhus, 2011b).

5. Results

Since we used a within-subject experimental design with two different sequencing orders for the two games, we performed a dummy regression analysis test to determine whether the sequencing of the games has any effect in how participants decided in one game versus the other. Table 1 shows the dummy variables regression analysis. As Table 1 shows, neither the type of game (dictator/gangster), nor the sequencing of the games had a significant effect in the average amount given by the dictators nor the amount left by the gangsters. Since there is no such effect, we can group all the dictators' decisions and all the gangsters' decisions regardless of the sequence in which such decisions were made for any given person.

Table 1. Dummy variables regression testing the game type and sequencing simultaneously.

	Coeff.value	Std. error	Std.coeff	t	Sig
Constant	17.553	3.872		4.534	0.000
Game	1.037	4.525	0.032	0.229	0.820
Sequencing	2.005	4.528	0.062	0.443	0.660

Regarding differences between the dictator and gangster games, a t-test comparison shows no significant difference between the average amount given by the dictator and the average amount left by the gangsters (dictators average=19.6, gangsters average= 18.5, P-value=0.83). A series of regressions were conducted in order to explore a link between participants Dark triad scores and their decisions, using the questionnaire score for Machiavellianism, Narcissism and Psychopathy as explanatory variables for the amounts given by the dictators and the amount left by the gangsters. Table 2 shows the resulting R-squares and the P-values for each of the three components of the Dark triad. As Table 2 shows, only Machiavellianism seems to play a marginally significant role (with a 90% confidence level) in determining how much a dictator will give to his counterpart.

Table 2. Average scores, scores standard deviations and regression results for each of the Dark triad's elements as explanatory variables of the amount dictators give and gangsters leave to their counterparts.

			Dictator game			Gangster game		
	Av.score*	Std.dev**	R^2	P-Value	Slope	R^2	P-Value	Slope
Narcissism	3.76	1.68	0.00	0.88	-0.30	0.00	0.76	-0.58
Machiavellianism	2.95	1.63	0.09	0.07	-3.10	0.02	0.50	1.15
Psychopathy	4.29	1.92	0.03	0.21	-2.18	0.01	0.56	0.99

* Scores for each of the triad's component range from 1 to 9. Questions 1 to 4 relate to Machiavellianism, 5 to 8 are related to Psychopaty and 9 to 12 are related to Narcisism

** Sample standard deviation

To further explore the relationship between the Dark triad components and subjects' behavior, we divided the group in accordance to their recent educational background. Tables 3 and 4 shows the results for the Economics and System Dynamics students. As these two tables show, there are only two marginally significant relationships in the Economics students case for the dictator game (Machiavellianism and Psychopathy are significant at a 90% confidence level). There seems to be no significant relationship between the Dark triad components and the dictator game for the System dynamics students nor for the gangster game in neither of the two groups.

Table 3. Regressions results for the System Dynamics master students

	Dictator game			Gangster game		
	R-Squared	P-Value	Slope	R-Squared	P-Value	Slope
Narcissism	0.01	0.71	-0.88	0.04	0.47	-1.67
Machiavellianism	0.04	0.51	-1.28	0.01	0.71	0.74
Psychopathy	0.00	0.93	-0.16	0.01	0.78	0.52

Table 4. Regressions results for the Economics master students

	Dictator game			Gangster game		
	R-Squared	P-Value	Slope	R-Squared	P-Value	Slope
Narcissism	0.04	0.52	-3.37	0.03	0.56	2.97
Machiavellianism	0.19	0.08	-8.08	0.04	0.50	3.16
Psychopathy	0.22	0.06	-6.64	0.03	0.57	2.10

6. Conclusions and Discussion

We developed and ran an experiment to study whether it is possible to establish a connection between the Dark triad and people decision making on income distribution in an experimental setting. Our results suggest that there is not a strong link between the Dark triad components of personality and how people decide on income distribution problems. However, it is interesting to note that we found smaller P-values for Machiavellianism and Psychopathy in the dictator game than in the gangster game. In this sense, the Dark triad does seem to be a good indicator for participant behavior in some experiments, while it has been a good indicator for behavior in other settings

(Arvan, 2012; Rauthmann, 2012) and not in others. This property resembles the research on the predictive power of the IQ for different types of performances and accomplishments. In a similar way as the IQ, which has been found to be both a powerful predictor (Schmidt et al, 1998; Frey et al., 2004; Murray, 1998) and a non-reliable predictor of high performance (Kamphaus, 2005; Neisser, et al., 1996) across different settings, the Dark triad prediction power may be highly context-dependent and thus hard to reproduce in some experiments. In this sense, The Dark triad could indeed be considered a form of the fundamental attribution error. Therefore, our results allow us to reject our alternative hypothesis and to not reject our null hypothesis.

The reasons why Machiavellianism and Psychopathy may be important predictive variables when it comes to making decisions about income distribution can be found in the literature. Machiavellians are characterized by being careful strategist who manage their own reputation and their relationship with others in such a way that allows them to get the maximum benefit for themselves (Kessler et al., 2010). Psychopaths are characterized by a marked lack of empathy, which leads them to disregard others wellbeing in relation with their own (Hare and Neumann, 2008). An interesting point comes from the fact that we are using a one-shot game in our experiment such that reputation building is not possible. By not allowing players to build reputation, the boundary between Machiavellianism and Psychopathy is difficult to define (Jones and Paulhus, 2011a). Future research is needed to explore specific links between the Dark triad and people's behavior in controlled environments.

Contrary to previous studies, this article finds no differences between the dictator and gangster games. One possible explanation for this is that previous studies have added circumstances. Such is the case of Eichenberger and Oberholzer-Gee (1998) who introduced fairness considerations and Bardsley (2008), who introduced asymmetric endowments and scale up transfers in both games. It is also interesting to compare this article results with previous studies in the public good and public bad literature such as Andreoni's (1995) work on the motivations behind people's contribution to a common good or a common bad. Andreoni argues that the good feeling of contributing to a common cause (the warm glow) is often stronger than the guilt one may feel when one contributes to a common bad (cold prickle). While Andreoni's (1995) paper and the present article are based on different games, one may argue that the warm glow and the cold prickle would be present in the dictator and gangster games, that is, the warm glow will drive the dictators to give money to their recipients, and the cold prickle will keep the gangsters from taking money away from their victims. Consequently, the dictators should give more than what gangsters leave to their counterparties if Andreoni's explanation is correct. Our results do not provide evidence to support such explanation. Further research is needed to explore alternative explanations for this phenomenon.

References

Adams, H. M., Luévano, V. X., Jonason, P. K. (2014). Risky business: Willingness to be caught in an extra-pair relationship, relationship experience, and the Dark Triad. Personality and Individual Differences, 66, 204–207.

Andreoni, J. (1995).Warm-Glow versus Cold-Prickle: The Effects of Positive and Negative Framing on Cooperation in Experiments. The Quarterly Journal of Economics 110:1-21.

Arvan, M. (2012). A lot more bad news for conservatives, and a little bit of bad news for liberals? Moral judgments and the Dark Triad personality traits: A follow-up study. Neuroethics. doi: 10.1007/s12152-012-9155-7.

Baughman, H. M., Dearing, S., Giammarco, E., Vernon, P. A. (2012). Relationships between bullying behaviours and the Dark Triad: A study with adults. Personality and Individual Differences, 52, 571–575.

Bardsley, N. (2008). Dictator game giving: altruism or artefact? Experimental Economics, 11(2), 122–133. doi:10.1007/s10683-007-9172-2

Campbell, W. K., Goodie, A. S., Foster, J. D. (2004). Narcissism, confidence, and risk attitude. Journal of Behavioral Decision Making, 17, 297–311.

Cappelen, A, W., Ulrik H. Nielsen, U, H., Sørensen, E, Ø., Tungodden, B., Tyran, J,R. (2012). Give and Take in Dictator Games. Norwegian School of Economics.

Cleckley, H. (1941). The mask of sanity. Oxford, England: Mosby.

Eichenberger, R., Oberholzer-gee, F. (1998). Rational moralists: The role of fairness in democratic economic politics. Public Choice, 94, 191–210.

Fehr, B. A., Samsom, D., Paulhus, D. L. (1992). The construct of Machiavellianism: Twenty years later. In C. D. Spielberger, & J. N. Butcher (Eds.), Advances in personality assessment (Vol.9, pp. 77116). Hillsdale: Erlbaum.

Flanagan, O. (1991). Varieties of Moral Personality. Cambridge, Mass.: Harvard University Press.

Frey, M.C. Detterman, D.K. (2004). Scholastic Assessment or g? The Relationship Between the Scholastic Assessment Test and General Cognitive Ability. Psychological Science. 15 (6): 373–8.

Furnham, A. (2010). The Elephant in the Boardroom: The Causes of Leadership Derailment. Basingstoke: Palgrave MacMillan.

Furnham, A., Crump, J. (2005). Personality traits, types and disorders. European Journal of Personality, 19, 167–184.

Furnham, A., Richards, S. C., Paulhus, D. L. (2013). The Dark Triad of personality: A 10-year review. Social and Personality Psychology Compass, 7, 199-216.

Hare, R. D., Neumann, C. S. (2008). Psychopathy as a clinical and empirical construct. Annual Review of Psychology, 4, 217246.

Harman, G. (1999). Moral Philosophy Meets Social Psychology: Virtue Ethics and the Fundamental Attribution Error. Proceedings of the Aristotelian Society, New Series, Vol. 99 (1999), pp. 315-331.

Jonason, P. K., Tost, J. (2010). I just cannot control myself: The Dark Triad and selfcontrol. Personality and Individual Differences, 49, 611–615

Jonason, P. K., Krause, L. (2013). The emotional deficits associated with the Dark Triad traits: Cognitive empathy, affective empathy, and alexithymia. Personality and Individual Differences, 55, 532–537.

Jones, D. N., Paulhus, D. L. (2009). Machiavellianism. In M. R. Leary, & R. H. Hoyle (Eds.), Handbook of individual differences in social behavior
(pp. 93108). New York: Guilford.

Jones, D. N., Paulhus, D. L. (2011a). Differentiating the Dark Triad within the interpersonal circumplex. In L. M. Horowitz, and S. Strack (Eds.), Handbook of interpersonal psychology: Theory, research, assessment, and therapeutic interventions (pp. 249267). New York: Wiley.

Jones, D. N., Paulhus, D. L. (2011b). The role of impulsivity in the Dark Triad of personality. Personality and Individual Differences, 51, 679–682

Kamphaus, R. W. (2005). Clinical assessment of child and adolescent intelligence. Springer. ISBN 0-387-26299-7.

Kessler, S., Bandelli, A., Spector, P., Borman, W., Nelson, C., & Penny, L. (2010). Re-examining Machiavellianism. Journal of Applied Social Psychology, 40, 1868-1896.

Maples, J. L., Lamkin, J., Miller, J. D. (2014). A test of two brief measures of the Dark Triad: The Dirty Dozen and Short Dark Triad. Psychological Assessment, 26, 326331.

Morey, L. C., Hopwood, C. J., Markowitz, J. C., Gunderson, J. G., Grilo, C. M., McGlashan, T. H., et al. (2012). Comparison of alternative models for personality disorders, II: 6-, 8-, and 10-year follow-up. Psychological Medicine, 42, 1705-1713.

Murray, C. (1998). Income Inequality and IQ. AEI Press. ISBN 0-8447-7094-9

Nathanson, C., Paulhus, D. L., Williams, K. M. (2006). Predictors of a behavioral measure of scholastic cheating: Personality and competence but not demographics. Contemporary Educational Psychology, 31, 97–122.

Neisser, Ulrich; Boodoo, Gwyneth; Bouchard, Thomas J.; Boykin, A. Wade; Brody, Nathan; Ceci, Stephen J.; Halpern, Diane F.; Loehlin, John C.; Perloff, Robert; Sternberg, Robert J.; Urbina, Susana (1996). "Intelligence: Knowns and unknowns" (PDF).

American Psychologist. 51: 77–101. doi:10.1037/0003-066x.51.2.77. ISSN 0003-066X. Retrieved 9 October 2014.

Nisbett, R. E. and Ross, L. (1980). Human Inference: Strategies and Shortcomings of Social Judgment. Englewood Cliffs, NJ: Prentice-Hall.

Paulhus, D. L., Williams, K. M. (2002). The dark triad of personality: Narcissism, Machiavellianism, and psychopathy. Journal of research in personality, 36(6), 556-563.

Paulhus, D. L, Buckels, E. E. (2011). The Dark Tetrad of Personality: Relevance to Terrorist Groups. Presented to the Defense Research and Development Canada (DRDC) agency, Toronto, Canada.

Rauthmann, J. F. (2012). The Dark Triad and interpersonal perception: Similarities and differences in the social consequences of narcissism, Machiavellianism, and psychopathy. Social Psychological and Personality Science, 3, 487– 496.

Rauthmann, J. F., Kolar, G. P. (2012). How ‘‘dark’’ are the Dark Triad traits’’ Examining the perceived darkness of narcissism, Machiavellianism, and psychopathy. Personality and Individual Differences, 53, 884–889.

Ross, L. (1977). The intuitive psychologist and his shortcomings. Advances
in Experimental Social Psychology, Vol 10, ed. L. Berkowitz (New York: Academic Press).

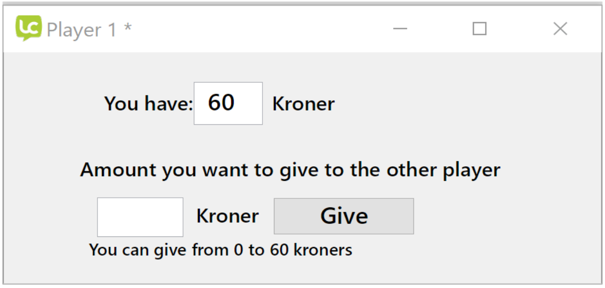
Schmidt, F.L. Hunter, J,E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. Psychological Bulletin. 124 (2): 262–74.

Wiggins, J. S., Pincus, A. L. (1989). Conceptions of personality disorders and dimensions of personality. Psychological Assessment: A Journal of Consulting and Clinical Psychology, 1, 305–316

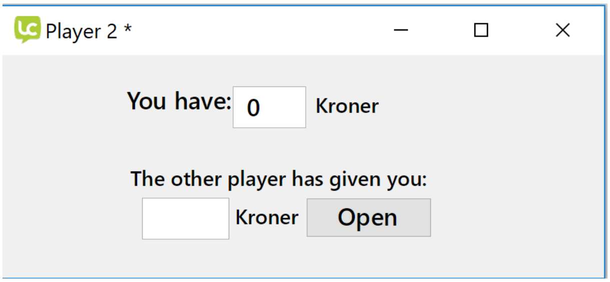
Williams, K. M., Nathanson, C., & Paulhus, D. L. (2010). Identifying and profiling scholastic cheaters: Their personality, cognitive ability, and motivation. Journal of Experimental Psychology: Applied, 16, 293–307.

Appendix A: Interfaces

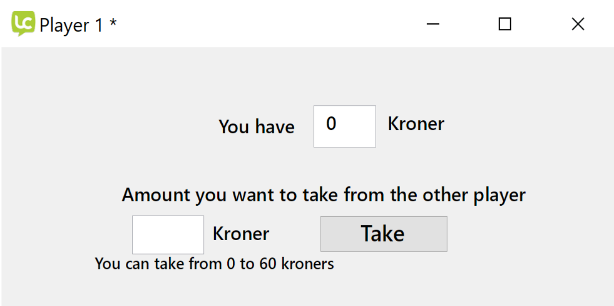
Dictator in the dictator game



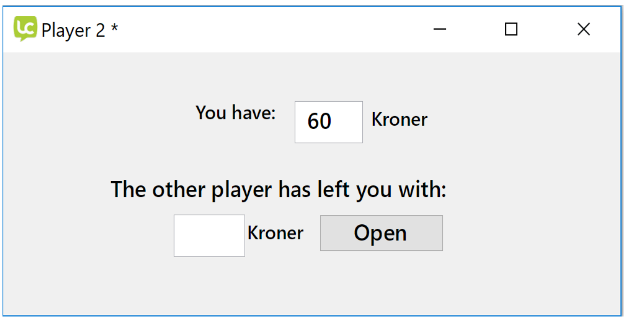
Recipient in the dictator game



Gangster in the gangster game



Victim in the gangster game



Appendix B: Instructions

Instructions for the dictator in the dictator game:

You are Player 1 in a group of 2 players. You start with 60 NOK, and Player 2 is given 0 NOK. You have to make a decision about giving money to Player 2. You are free to give from 0 to 60 NOK. Player 2 knows that you have started with 60 NOK. You cannot identify Player 2, and Player 2 cannot identify you.
Write the amount you will give to Player 2 in the blank space, and click on the Give button. After the game you will be paid the remaining amount of money, and Player 2 will be paid the amount you gave him or her.

Instructions for the recipient in the dictator game:

You are Player 2 in a group of 2 players. You start with 0 NOK, and Player 1 is given 60 NOK. Player 1 can give some of this money to you. Player 1 is free to give from 0 to 60 NOK. Player 1 knows that you have started with 0 NOK. You cannot identify Player 1, and Player 1 cannot identify you.
To be ready to receive information about how much Player 1 decides to give to you, click on the Open button. This is the amount you will be paid, and Player 1 will be paid what remains of the 60 NOK.

Instructions for the gangster in the gangster game:

You are Player 1 in a group of 2 players. You start with 0 NOK, and Player 2 is given 60 NOK. You have to make a decision about taking money from Player 2. You are free to take from 0 to 60 NOK. Player 2 knows that you have started with 0 NOK. You cannot identify Player 2, and Player 2 cannot identify you.
Write the amount you want to take from Player 2 in the blank space, and click on the Take button. After the game, you will be paid the amount of money that you took, and Player 2 will be paid what remains.

Instructions for the victim in the gangster game:

You are Player 2 in a group of 2 players. You start with 60 NOK, and Player 1 is given 0 NOK. Player 1 can take money from you. Player 1 is free to take from 0 to 60 NOK. Player 1 knows that you have started with 60 NOK. You cannot identify Player 1, and Player 1 cannot identify you.
To be ready to receive information about how much Player 1 has decided to leave for you, click on the Open button. This is the amount you will be paid, and Player 1 will be paid the amount that was taken from you.

Appendix C: Dirty dozen questionnaire (Maples et al, 2014)

This is a standard questionnaire that has been used in previous studies all over the world. Do not write your name anywhere on this sheet of paper.

Please rate your agreement with the items below using numbers ranging from Strongly disagree (1) to Strongly agree (9).

1. I tend to manipulate others to get my way.

2. I have used deceit or lied to get my way.

3. I have used flattery to get my way.

4. I tend to exploit others towards my own end.

5. I tend to lack remorse.

6. I tend to be unconcerned with the morality of my actions.

7. I tend to be callous or insensitive.

8. I tend to be cynical.

9. I tend to want others to admire me.

10. I tend to want others to pay attention to me.

11. I tend to seek prestige or status.

12. I tend to expect special favors from others.

Graphic design: Communication Division, UIB / Print: Skipnes Kommunikasjon AS

uib.no

ISBN: 978-82-308-3780-1