Essays in the Political Economy of Organisations: Power, Leadership and Coordination

Lina María Restrepo-Plaza

PhD in Economics

University of East Anglia

School of Economics

July 2017

Abstract

This study provides with behavioural insights on three components of the political economy of organisations namely power, leadership and coordination. In the second chapter, we implement a lab experiment to study how the experience of power affects the value given to powerful positions. We contribute to the literature on the value of decisions rights by replicating in the lab a real-life-like power asymmetry. Moreover, we manipulate the status and the legitimacy of the powerful positions, and elicit the value of power from the powerful and the powerless reference points. We find that those who have experienced and exercised power are willing to let go substantial pecuniary compensations to remain in charge.

The third chapter experimentally compares teams versus individual in a coordination task. Since in our setting teams do not necessarily outperform individuals, our results partially contrast the state of the art on this arena. We contribute to the literature on organisational behaviour by pointing out the importance of coordinating institutions such as voting rules, communication, joint incentives, etc., to promote teams' coordination.

Finally, we empirically study the role of formal leaders in real-life organisations: football teams. We contribute to the literature in organisational behaviour in sports by implementing an adaptation of the red card into the sport production function to estimate the relative effect of a captain dismissal in both Northern and Southern European leagues. We find context specific results, as we find that the captain dismissal only affects the match outcomes in Southern leagues when the away captain is sent-off. This result might open a window to incorporate leadership and cross cultural studies in the literature of the *home team advantage*.

Content List

Chapter 1. General Introduction 8

Chapter 2. The Behavioural Toxicity of Power 11

1. Introduction 12

2. Related Literature 15

3. Experimental Design and Procedures 17

4. Results 21

4.1 Ruling on rates 21
4.2 No Power 24
4.3 Minimal Power 25
4.4 Disentangling the components of the bids for power 29
4.5 Summary 33
4.6 Underbidding bias 34
4.7 The Effect of Power on Individual Productivity 37

5. Conclusions 39

Chapter 3. The Cost of Dissent 42

1. Introduction 43

2. Related Literature 45

3. Experimental Design 47

4. Experimental Results 52

4.1. Individual effort 52
4.2. Wasted Effort 57
4.3. Conditional Contribution Adjustment 58
4.4. Efficiency 60

5. Conclusions 61

Chapter 4. Leadership and Team Performance: A Comparison between Northern and Southern European Soccer 64

1. Introduction 65

2. Related Literature 66

3. Northern and Southern Leagues Background 70

4. Methodology 74

5. Results 79

5.1. Descriptive statistics 79
5.2. The Relative Effect of a Red Card 83

Chapter 5. General Concluding Remarks 89

References 94

Appendix 101

A.1.1. Complementary figures and tables chapter 1 102

A.1.2. Risk aversion analysis 106

A.1.3 Experimental instructions chapter 1 107

A.2.1. Experimental instructions chapter 2 119

A.3.1. Description of the Squawka performance score chapter 3 121
A.3.2. Complementary tables and figures chapter 3 122
A.3.3. Complementary bivariate analysis chapter 3 124
A.3.4. Complementary captain absence analysis chapter 3 126
A.3.5. Captain attributes chapter 3 127

Table List

Table 1.1	Experimental design	19
Table 1.2	Treatments summary Table	21
Table 1.3	WTP, WTA and Gap: BL-High vs NP	26
Table 1.4	WTP, WTA and Gap 2x2 Conditions	28
Table 1.5	Participation of risk aversion in the actual value of power by treatment and role	31
Table 1.6	Participation of the material benefits in the actual value of power	32
Table 1.7	Participation of noise (confusion-demand-framing)	33
Table 1.8	Corrected WTP	36
Table 1.9	The WTP and WTA for the A position and the Gap by treatment (control)	102
Table 1.10	Participation of the components in the actual value of power by treatment (controls) and role	103
Table 1.11	Individual ruling categories by categorisation method	104
Table 2.1	Descriptive statistics	52
Table 2.2	Contribution and Provision Estimations	53
Table 2.3	Contribution Estimation by Nash Scenario	56
Table 2.4	Wasted Effort and adjustment (I-WLM versus T-WLM)	57
Table 2.5	Conditional Contribution	59
Table 2.6	Conditional Contribution Adjustment	60
Table 2.7	Earnings Estimations	61
Table 3.1	Southern and Northern League's Characteristics	73
Table 3.2	Variables' Definitions	77
Table 3.3	Descriptive Statistics of home and away teams	79
Table 3.4	Home vs. Away Red Cards	80
Table 3.5	Captains and Non-Captain Characteristics	81
Table 3.6	Red Card Events and Time Played by Captains	82
Table 3.7	Linear Regressions for Goal Differences, Standard errors in parentheses	86
Table 3.8	Linear Regressions for Goal Differences by Leagues	122
Table 3.9	Means of Captain Characteristics by League	123
Table 3.10	T-Tests results for Captain Attributes' Differences	123
Table 3.11	Bivariate Regressions by Region	125
Table 3.12	Linear Regressions by Region including Injuries and Substitutions	126
Table 3.13	Dismissed Captain Attributes	127

Figure List

Figure 1.1	Experimental stages	18
Figure 1.2	Average rate per quintiles	22
Figure 1.3	Ruling categories by treatment (2x2 conditions)	23
Figure 1.4	WTA and WTP in the No power treatment	25
Figure 1.5	WTA and WTP BL High vs NP	26
Figure 1.6	The value of *minimal* power: WTP and WTA in the 2x2 conditions	27
Figure 1.7	The Components of the WTA and WTP values	33
Figure 1.8	WTP, corrected WTP and WTA	37
Figure 1.9	The effect of power on individuals' productivity	39
Figure 1.10	Rates Distribution across all treatments	102
Figure 1.11	Average assigned rates by productivity quintile per treatment	103
Figure 1.12	Ruling category by treatment (controls)	103
Figure 1.13	Real and Corrected Bids Distribution	105
Figure 2.1	Contribution and provision by treatment	52
Figure 2.2	Contribution versus provision by group and treatment	54
Figure 2.3	Contribution by decision makers over time	55
Figure 2.4	Distributions of individual contributions over time	55
Figure 3.1	Southern and Northern Leagues Geographical Distribution	70
Figure 3.2	Histogram of the Red Card per Match by Region	83
Figure 3.3	Pitch Areas, took from Squawka.com	121

Acknowledgements

To Juanfer…

Chapter 1. General Introduction

The purpose of this thesis is to provide some behavioural insights regarding the political economy of organisations. I focus my study on three main research questions: “what are the determinants of the value of power?”, explored in Chapter 2; “what are the dynamics that differentiate teams and individuals performing coordination tasks?”, presented in Chapter 3. Finally, Chapter 4 looks at the role of formal leaders on team performance.

Two different methodological approaches are adopted. In the second and third chapters the conclusions are supported by the results of computer-based lab experiments conducted recruiting college students. In chapter two, we generate a power asymmetry among participants to analyse how the experience of power affects individual behaviour, and in the third chapter we recreate a day-to-day coordination challenge to compare individuals and teams’ performance in the absence of further coordinating institutions. In contrast, for the fourth chapter we use data provided by real interactions of individuals that are performing real life tasks. Specifically, we use a database of football matches from six major European football leagues, and their captains’ main attributes to carry an “applied” analysis, in the traditional economics sense.

The second chapter studies the intrinsic value of power. We analyse the value of power in a real effort task experiment in which a minority of individuals (participants A) have decision rights on the rate at which most individuals (participants B) earn, and have no control on their own fixed rate. Participants were matched in groups of three in which one individual A, decided the piece rate for two individuals B. Even when power has no extrinsic economic value, it has both procedural (from the asymmetric distribution of rights) and consequential (the rates at which most participants earn) content. After experiencing power, participants A may sell it to participants B, before being randomly re-matched and play the same real effort task again. We systematically manipulate the attributes of power (legitimacy and entitlement), and the status of the decision rights associated with it. Our findings suggest that the willingness to pay of those who have not experience of power (participants B) can be fully rationalized by subjects’ risk attitudes (a preference for a fixed rate), the material benefits of the A-role and the demand effect or individual priors regarding bidding or being labelled as A. Participants A, with an experience of power, are willing to suffer substantial losses to retain decision rights.

The third chapter studies the challenges of teams’ coordination in the absence of communication, consensus and other institutions that favour team learning. We compare the effectiveness of teams and individuals in a setting in which individual incentives may crowd-out choices that would enhance social welfare. In our main treatment, groups of four teams jointly produce in a quasi-continuous version of the minimum effort game. As the overall output is determined by the minimum team contribution, the strategic interaction between teams follow the logic of a coordination game. However, as individual decisions within teams are costly and not necessarily unanimous, the game becomes a social dilemma within each team, in which individuals have an incentive to free ride on the efforts of others. The game combines some interesting properties of coordination and public goods games in a realistic setting in which the individual's effort is always costly. We compare this main treatment with two different baselines in which individuals rather than teams jointly produce outcomes with either a weakest link (as in the interaction between teams in our main treatment) or a linear technology (as in the interaction within teams). Our findings suggest that in this more complex and natural setting, individuals outperform teams. Teams suffer significant efficiency losses, not because they coordinate worse than individuals, but because they coordinate in an inefficient equilibrium.

Finally, in the fourth and final chapter we explore the extent to which leaders facilitate cooperation and coordination within real world setting, in this case a sporting environment. We study the effect of the *regular captain* on football team performance by using an innovative approach that looks at captain versus non-captain dismissals. We implement an adaptation of the red card dismissal into the *sport production function* methodology to isolate the effect of the team captain on team performance. This analysis differs conceptually with respect to the team coach leadership in that captains are horizontally positioned regarding other on-the-field players, but they are still different in terms of status. It also differs to the studies already carried out by sport psychologist, based on survey data, in that we are not looking at individual perceptions but actual team outcomes. We analyse data from six major European football leagues, three Northern leagues and three Southern leagues between 2012 and 2015. We find that while, relative to non-captains, home captains dismissals do not affect the team outcomes, home teams of Southern leagues appear to be able to exploit the dismissal of the away captains. This result does not hold for Northern leagues, suggesting that Northern teams appear to exhibit certain dynamics that allow them to overcome the absence of their regular leader. Further research considering the criteria for captain appointments within each team, and individual team players' performance would complement the findings presented here.

In the remainder of this document, chapters two, three and four are presented, comprising the introduction, related literature, methodological approach (or experimental design), and empirical results. Conclusions will also be provided by the end of each section. Finally, some general concluding remarks will be presented in the fifth chapter.

Chapter 2. The Behavioural Toxicity of Power[1]

[1] Co-authored with: Enrique Fatas, School of Economics, University of East Anglia, e.fatas@uea.ac.uk

1. Introduction

Power is not a means; it is an end […] The object of power is power George Orwell, *1984*

Most social relationships are asymmetric in nature. Within Firms, teams, universities, and families, some individuals have the resources needed to request others to follow their commands and make decisions they would not make otherwise. In the political science literature, this is called power (Dahl, 1959; Emerson, 1962), and social psychologists have long documented how power is intrinsically valued by most individuals and its effects on behaviour (see Sturm and Antonakis, 2015 for an interesting survey on this topic).

In this branch of the literature, power asymmetries are generated by priming, recalling or creating imaginary environments in which participants experience a power imbalance. In these environments power is desired because it brings along autonomy from and influence over others (Lamers et al, 2016)[2], and because it creates the conditions to obtain higher material benefits than those received by powerless participants (Bendahan et al 2015). Power enhances individuals' creativity (Galinsky et al 2008, and Guinote, 2007), as well as cognitive skills (e.g. speediness to act), attention (e.g. information selectivity), and flexibility (e.g. on information processing). When power perception becomes strong, and individuals see their power well above the power of others, they are more optimistic, confident, and risk seeking (Anderson and Galinsky, 2006), and more prone to initiate competitive interactions (Magee et al 2007). However, powered individuals more frequently make corrupt and self-interested actions (Bendahan et al 2015).

In Economics, the study of power asymmetries, and their effects on economically relevant decisions, is quite recent, focusing on the instrumental and consequential value of decision rights in a principal-agent framework (Fehr et al, 2013; Bartling et al, 2014; Bobadilla-Suarez et al, 2016; Charness et al, 2012) and the subjective value of being in control (Charness and Gneezy, 2009; Sloof and Siemens, 2014; Owens et al, 2014, Coats and Rankin, 2015).

---

[2] Lamers et al (2016) conducted nine experimental studies in which participants were induced to imagine or to recall a highly (lowly) influential or a highly (lowly) autonomous experience. They found that, while influencing others have a secondary role, the desire for power is mainly driven by the need for autonomy.

In this study, we bridge the psychological and economic literature studying how the experience of power changes individual behaviour in a real effort task laboratory experiment. In our setting, individuals are asymmetrically endowed with *minimal power*, defined as the ability to affect others' outcomes (by choosing the rate at which they earn) without being able to affect their own or to extract any surplus (as they earn at a fixed rate). As participants repeatedly interact with each other, the first part of the experiment is designed to generate an asymmetric experience of power. In a second stage, power can be exchanged, and individuals may sell their powerful positions, or buy them, and we measure how the asymmetric experience of power changes its value.

The contributions of this study are multiple. First, we manipulate the institutions through which power is allocated and exercised to identify the effect of these institutions on the behaviour of the controlling and subordinated parties. We run a 2X2 between-subjects experimental design. In the first dimension, following Ball and Eckel (1996) and Ball et al (2001), powerful positions are assigned to either top or bottom performers (as per their individual score in a 1st block, in which they face the same real effort task): *High and Low treatments*, respectively. In the other dimension of our factorial design, we manipulate the information given to participants regarding how power is allocated: while in the *Baseline (BL) treatments* subjects receive no information about the selection procedures, in the *Legitimacy (LT) treatments* subjects are aware of the mechanism. [3]

Second, we carefully control for any effect of power on individual and group performance. As groups are always formed by one participant with high power (participants A) and two individuals without power (participants B), we use individual performance in 1st block to classify participants in three categories: *high, medium,* and *low*, as per their tercile. Whereas in the *BL-High* and *LT–High* conditions individuals in the 1st tercile become participants A, in the *BL–Low* and *LT–Low*, those in the 3rd tercile become participants A, and the rest become B. As we know the individual productivity of A and B participants in every condition, we can measure the impact of power on their productivity.

[3] We also controlled *for high-performance-enforcement* and *high-performance-centrality* (as a proxy for charisma/visibility). The findings are aligned with the LT results and are presented in the Appendix 1.1.

Third, we decompound the Willingness to Pay (WTP) and Willingness to Accept (WTA) for power and calculate the proportion associated to any residual material benefits of power (an *extrinsic value of power* linked with higher rates being assigned to specific productivity levels), any insurance value of being a participant A (linked with *risk aversion* and uncertain rates) and any confusion, demand effects or priors regarding the A-label itself, generated by our elicitation mechanism (with an additional *No Power* treatment (see Zizzo, 2010; Bardsley, 2005; Bardsley, 2008, Eckel et al, forthcoming).

Our results strongly suggest that power generates substantial and significant losses for those holding it in block 2. Participants with an experience of power in that block (A) set very high selling prices, relative to the buying prices offered by participants with no experience of power (B). Differences are significant and large in every experimental condition (*BL-High*, *BL-Low*, *LT-High* and *LT-Low*). The only (and informative) exception is the *No power* condition, in which participants A and B face an identical environment to the one faced by participants in the other four treatments, the same real effort task, and a minority of top performers (as in *BL-High* and *LT-High*) are selected into A roles using the same framing and labels; however, without power asymmetries, selling and buying prices of the A-position are indistinguishable in this control treatment.

By setting extraordinarily high selling prices, participants A forego large monetary earnings. In that sense, this study provides evidence supporting the *toxicity* of power on individual behaviour only for those who have experienced it first. In sharp contrast, participants B buying prices can be fully rationalised as a combination of risk aversion, material gains, and demand effects. While the literature discussed at the beginning of this introduction has documented the value of power and authority (e.g. Bartling et al 2014), this is the first time such effect has been studied. Moreover, we do not find evidence that a clear majority of individuals intrinsically value power, as those without experience of power react rationally to uncertainty, minor gains and noise (e.g. demand effects). The intrinsic value of power sharply increases for those who enjoyed it and exercised it.

Interestingly, and contrary to Fehr et al (2013) and Bartling et al (2014), we do not find evidence of a positive impact of power on productivity. Additionally, our design allows for a classification of participants A by their ruling types. Participants A may use their power to equalise participants B's payoffs (assigning higher rates to the less productive participant B), reward (if the more productive participant B gets a higher rate) or passive (if she keeps the status quo and assigns the same piece rates to both participants B). We study if the way power is handled generates significant differences in its value and find that this taxonomy does not account for differences in the value assigned to power by As (and do not affect individuals' productivity).

The remainder of the chapter is organised as follows. The next section presents the related literature. The third section illustrates the experiment design and procedures. Results are provided in the fourth section. Finally, the main conclusions are available in the last section of the chapter.

2. Related Literature

Previous research linking behaviour and power has focused on how individuals value control and autonomy in agency settings in which principals are endowed with the autonomy to make payoff relevant decisions on behalf of both players. Individuals are willing to forgo large monetary benefits to remain autonomous and retain their decision rights (Fehr et al, 2013; Owens et al, 2014; Coats and Rankin, 2015; Neri and Rommeswinkel, 2014; Bobadilla et al, 2016). Fehr et al (2013, p 2) summarises well the main result: "*The allocation of decision rights has non-pecuniary consequences that inhibit the delegation of authority. In our experiment, the fact that the principals are willing to sacrifice some of their earnings to keep authority suggests a preference for the decision righ*t". This result is even stronger in other studies in which individuals are willing to pay to have the (illusory) control on results that are randomly determined (Charness and Gneezy, 2010; Sloof and von Simmens, 2014).

Holding the agents' autonomy comes at a cost. Fehr et al (2013) studies a delegation setting in which principals may give up their decision rights without a binding effort contract. The authors find that delegation happens rarely, and this has a significant and positive effect on effort: while principals retaining control choose effort levels above the optimum, the opposite happens to agents with no decision rights. These results are consistent with similar effects found in Charness et al (2012) and Falk and Kosfeld (2006). In Charness et al (2012)'s gift-exchange game individuals giving up the control upon efforts and wages provide higher efforts for the principal's benefit, and delegation is a Pareto-improvement strategy. Falk and Kosfeld (2006) conducted a two-stage principal-agent game in which the principal is entitled to restrict the agent's choice set to protect her own payoffs. They found that more controlling principals induce less performer agents as the latter perceived this action as a sign of distrust and autonomy inhibition.

Bartling et al (2014) is the experimental study closest to our research question. Their setting differs to the one used by Fehr et al (2013) in that principals may request a binding minimum effort level to agents to delegate their decision rights, being effort levels of whoever is in control directly related to the project's probability of success. They study the intrinsic value of decision rights comparing the monetary values of two lotteries: a controlling and a delegation lottery. Both lotteries are the gambles that would take place if the principal remains in control, or if the decision is made by the agent, respectively. The difference between the value of both lotteries is the intrinsic value of retaining the decision rights:

Value of control lottery + Intrinsic value of decision rights = Value of the delegation lottery

The value of the delegation lottery is significantly higher than the control lottery, and the intrinsic value of decision rights is positive and substantial for a large majority of principals, and it increases in settings in which incentives are not aligned and there exists a conflict of interest between principals and agents. Interestingly, there is no explicit control of individuals' risk attitudes and the ex-post differences between the value of both lotteries (as higher effort levels increase the probability of winning the lottery).

In our setting, we measure the intrinsic value of retaining decision rights over someone else's payoffs controlling for individual risk attitudes, confusion and any residual value of power. As individuals earn by participating in a real effort task, the analysis of any effort enhancing effect of power on productivity is more direct. The details of our design are explained in the next section.

3. Experimental Design and Procedures

Figure 1.1 presents our between-subjects factorial design. In each condition, different participants make decisions in an identical sequence of four blocks. In block 1, participants are asked to add 2-digit numbers for three minutes, being paid for their individual productivity in this *real effort task* (RET), at a rate of 5 experimental currency units (ECU) per correct answer, while mistakes do not count. Block 1 serves two purposes. First, it generates a clean measure of productivity differences across individuals. Second, it allows us to rank them into productivity terciles: *high, medium and low*.

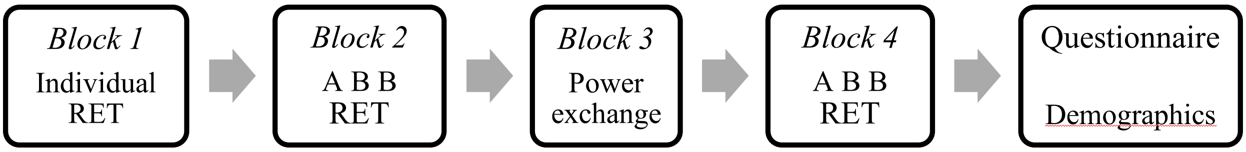
In the block 2, one third of the participants are selected into, and labelled as, participants A[4], while the rest are labelled as B. The labels chosen (A versus B) and the explanations given to participants in the experiment introduce this role allocation as a selection process (using a diagram in which participants A are above participants B). Other than that, the experimental instructions are unloaded, and use neutral terms.

Subjects are randomly allocated to groups of 3, with one participant A and two participants B, and they all face the same RET for 10 rounds of 1 minute. While participants A earn at the same, constant piece rate of 5 ECU per correct answer, the rate at which participants B earn is chosen by participants A satisfying a simple *budget* constraint: in every round, group and condition, the average rate at which participants B earn must be also 5 ECU. In other words, *power* (being A) gives no explicit advantage to participants A, other than the exercise of power itself.

Participants A make the additional decision on rates at the end of each round, so each subject knows the rate at which she earns when the round starts. After receiving information about the individual productivity of participants B (their scores), and their own score, participants A assign rates for the next round. Participants B do not observe the productivity of any other participant, and only receive information about their round payoffs.

[4] From the 1st or the 3rd tertile

Figure 1.1: Experimental stages



In the block 3 both A and B-subjects receive an additional endowment of 150 ECU[5] that is directly added to their earnings. In this block, we directly ask participants A for their willingness to accept (WTA) for selling their role in the last block of the experiment, and participants B are invited to bid their willingness to pay (WTP) to get the A-position. While the mechanism used for participants A is a standard BDM mechanism (as in Becker, DeGroot and Marschak, 1964), participants B set their WTP in an ascending price auction, being common knowledge that the top half WTP will be randomly matched with one WTA. The role is exchanged whenever the WTP of participants B is at least as high as the WTA of the participant A they are randomly matched with. The ascending price auction has an interesting property: as only half of the WTP are matched with the WTA, the number of participants A and B is kept constant for the last block of the experiment (one third and two thirds, respectively). [6]

At the beginning of block 3, the existence of a last block of ten rounds (block 4) is publicly announced. If the random matching between one WTA and one WTP finishes in a role exchange, the entire WTP value is debited from the B-participant experimental account and transferred to the A-Player's account. Otherwise, no exchange happens and participants keep their roles in block 4. By making participants B to pay their WTP in full, we generate a strong test for any endowment effect linked with power experience. [7] Rational participants A with no intrinsic interest in the unprofitable exercise of power (e.g. choosing rates) should rationally post a low WTA price, maximizing their chances of selling their role. Block 4 is identical to block 2, participants are randomly re-matched in brand new groups (following a pure stranger matching protocol), and they

[5] The 150 ECU given to all participants controls for any differences in earnings generated in previous blocks.

[6] As we explain later in this section, the session size is always 18, with 6 participants A and 12 participants B per session.

[7] The endowment effect (see Thaler, 1980) is a direct legacy of loss aversion as in prospect theory (Kahneman and Tversky, 1979). The discrepancy between WTP WTA happens when individuals overweighting losses over gains by using a reference point, and losing what is theirs is less desirable than obtaining an equivalent gain. We specifically control for any endowment effect with the No Power treatment, as we discuss in the following section.

face 10 rounds with the same RET and rules. The experiment finished with a questionnaire.

Table 1.1 below presents the 2x2 factorial design used in the experiment (plus one additional control treatment, described below). We classify participants in three groups per their individual productivity in the 1st block: *high* (1st tercile), *middle* (2nd tercile) and *low* ability individuals (3rd tercile). *High* ability individuals become participants A in the *High* treatments, and *low* ability participants get the role in the *Low* conditions. While in the *Baseline* conditions (*BL-High* and *BL-Low*), and in the *No Power* treatment, subjects were not aware of the mechanism used to assign roles, in both *Legitimacy* treatments (*LT-High* and *LT-Low*) participants were aware of the selection criterion. [8]

Table 1.1: Experimental design[9]

	High	**Low**
No Power	NP	--
Baseline	BL-High	BL-Low
Legitimacy	LT-High	LT-Low

The experimental design targets a set of research questions. First, we want to study the effect of productivity differences on the valuation of power. Do high and low ability individuals value power in different ways? We are agnostic on the interaction between these two variables.

If participants A always chose to reward the most productive B individuals with higher rates, power could come at a cost for high ability individuals if they expect to earn at a rate higher than 5 if becoming B. We should observe then low WTA values in the *High* condition (high ability individuals anticipate higher rates if selling power) and low WTP in the *Low* conditions (high ability individuals are not interested in power, as they earn

[8] As additional controls, we also implemented two treatments in which we manipulate participants A's *centrality* and their ability to *enforce* rates beyond the budget constraint. In the *centrality* treatment, we increased participants A' visibility by reporting their productivity to the other group members at the end of each round (see Carpenter et al, 2012, and Fatas et al, 2010 for other experiments on centrality). Even when participants in these sessions were not aware of how roles were assigned, as in the BL treatments described in the main text, this information could allow participants B to learn about any productivity differences (e.g. participants A being more productive in the High condition). In the *enforcement* treatment, the sum of both rates at which participants earned could be lower than 10 ECU. As we found no significant differences with the BL conditions, we report the results of these two additional treatments in the appendix A.1.2. Results, and procedures, area available upon request.
[9] Instructions are available in the Appendix A.1.3.

at high rates as B). If rates do not follow productivity, but aim to equalize earnings by assigning higher rates to less productive participants, power could be potentially attractive to high ability participants if they earn at a rate lower than 5 as B. Therefore, we should observe high WTA (WTP) in the *High* (*Low*) conditions among high ability participants.

Second, we are also interested in learning if the *legitimacy* of power has any effect on the way the A positions are valued by individuals. In our experiment, we proxy legitimacy by the information participants get about how roles A and B are assigned. As only in the LT treatments subjects are aware of the criterion used to assign roles, we can compare *BL-High* versus *LT-High* and *BL-Low* versus *LT-Low*, keeping all other experimental features constant, to see any effect of this information. By revealing in the *LT-High* treatment that the best performers become participants A, we expect both WTA and WTP to be higher than when it is common knowledge than the low ability individuals get the job of assigning rates (as in the *LT-Low* condition). This effect should not be observed in the BL-High versus BL-Low comparison, as subjects are not aware of the selection criterion, and cannot value more (less) the A role because it was originally assigned to the best (worst) performers in block 1.

Participants in the experiment could very well get confused in block 3. They could also bid for role A just because they positively react to the implicit demand of bidding (see Zizzo et al 2010). Individuals could bring strong priors and beliefs to the lab about the meaning of getting an A rather than a B label, or about the convenience and prestige of leaving the lab having been selected for a role. Therefore, we control for the potential *noise* created by the specifics of this setting, and run a control treatment: *No Power* (*NP*).

In the NP treatment participants do the initial RET in block 1, and the top third performers become participants A, as in any *High* treatments, and the selection criterion remains unknown, as in the *BL* condition. We still use the same diagram presenting a pyramid with participants A over participants B, and the same wording describing the assignment of roles as a selection process (a minority of subjects selected as A, a majority becoming B). The differences with every other treatment is that in blocks 2 and 4 all participants (A and B) always earn at a rate of 5 ECU per correct answer, and always receive information about their own score and earnings. As in any other block 3, they may still sell and buy the A-position for block 4. If the difference between WTA and WTP were generated by confusion, demand effect, priors on labels, or any other spurious endowment effect, the gap in the valuation of role A should still be positive in the NP treatment.

We conducted 14 computer-based sessions (2 sessions per treatment) programmed in Z-Tree (Fischbacher, 2007) at the CBESS lab at the University of East Anglia. Sessions were symmetric in that we kept the session size constant: 18 participants, of which 6 were allocated the role A and 12 role B. The sessions lasted 80 minutes and participants earned £11.4 on average (US$18), including a £2 participation fee. We recruited 180 participants, 36 for each treatment.[10] Since individuals were grouped up in the block 2, we had 12 independent observations for participants A and 12 independent observations of participants B per treatment (see Table 1. 2). Payments during the experiment were made in Experimental Currency Units (ECU) at an exchange rate of £1 = 150 ECU (Experimental Currency Units).

Table 1.2 Treatments summary Table

	No power		*Baseline*		*Legitimacy*		*Total*	
	A	B	A	B	A	B	*A*	*B*
High	12	24	12	24	12	24	*36*	*72*
Low	-	-	12	24	12	24	*24*	*48*
# Ss	*36*		*72*		*72*		*180*	

4. Results

We start the results section by addressing a simple question: how is power used by participants A in block 2, before studying bids in block 3 of the experiment in subsections 4.2 (for the No Power condition, our natural benchmark) and 4.3 (for the 2 x 2 factorial design). In subsection 4.4 we will separate the different elements of power's valuation. We will finish this part with the analysis of power on productivity by comparing performance in Blocks 2 and 4, controlling for participants' roles.

4.1 Ruling on rates

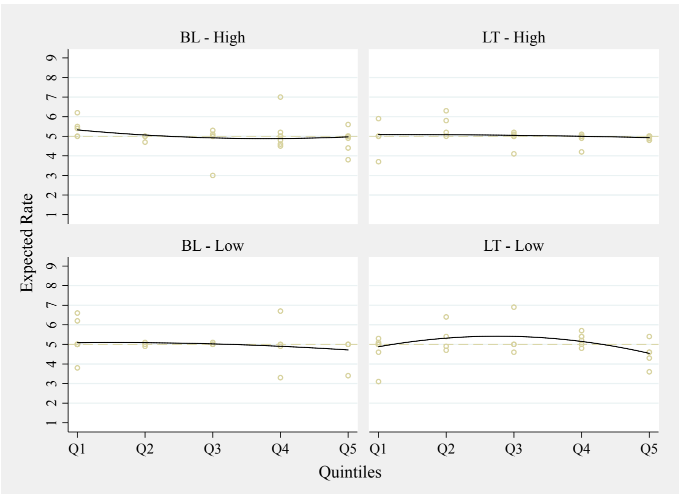
In block 2, participants A receive information about the productivity of participants B in their group in that round, and then decide the rates at which they will earn in the

[10] 252 participants in total, including the 72 participants in the Centrality and Enforcement treatments described previously.

following round. Figure 1.2 below plots the average rates at which participants B were paid, per quintiles of total earnings (being Q5 the highest earners' quintile), and for each experimental condition, including a non-linear fit. By inspection, not a single participant earned at a rate lower than 3 (higher than 7) in any treatment. [11] Participants A in both Baseline treatments are on average moderately more egalitarian than in Legitimacy treatments. Participants getting the low rates (from 1 to 3) are more productive than those getting the high rates (from 7 to 9): +2.8 correct answers per round in BL-High and +3.6 in BL-Low. There is no difference in scores in the two Legitimacy conditions: -0.5 for LT-High and +0.7 for LT-Low.

As the non-linear fit lines document, the connection between average rate and total earnings is weak and insignificant, mainly because participants A do not regularly use extreme rate values to reward or punish participants: on average, 22% of the times. Extreme rates (9-1, 8-2 and 7-3) are used slightly less frequently by top performers (17.5% of times in BL-High and 16.7% in LT-High) than by bottom performers (23.3% in BL-Low and 30% in LT-Low).

Figure 1.2 Average rate per quintiles



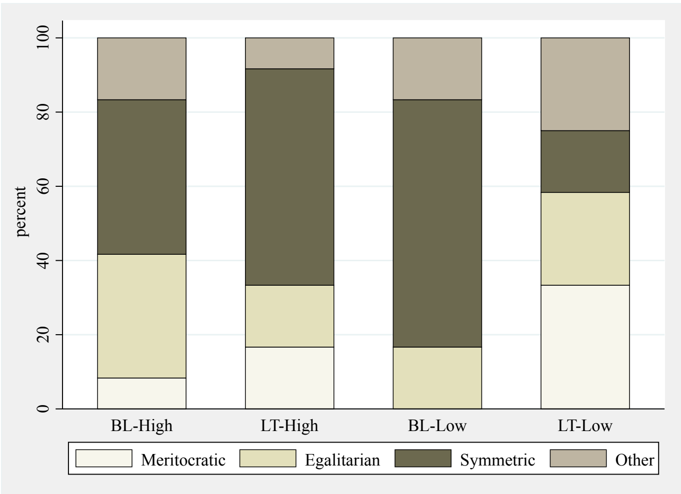
Given the strict rate constraint Participants A face in our experiment (both rates must sum 10 each round) we can easily categorise participants A as *egalitarian, meritocratic* or *symmetric*: if they reward the most productive participant B with a higher rate, we label them as *meritocratic*; if they compensate the less productive participant B with a higher rate, we label them as *egalitarian,* and if they give both participants B the same rate, we label them as *symmetric*.

[11] Figure 1.11 in the Appendix A.1.1 represents the average rates paid for each quintile in the control treatments.

We follow a combination of two simple rules to categorise ruling behaviour across all decisions in Block 2. As the three rules are mutually exclusive, we first count the number of rounds each participant A behaves as *meritocratic*, *egalitarian* and *symmetric* and if she has a unique mode, we assign her to the corresponding type. We also use a simple a *scoring rule* and calculate the squared difference between the rates they assigned and the rates they should have assigned if they behaved as *meritocratic*, *egalitarian* or *symmetric* types. We assign them the type that minimises the squared distance across all rounds in Block 2. For discrepancies between both methods, we create an additional *Others* type. [12]

Our results are illustrated in Figure 1.3. Almost half of all participants A (45%) can be categorised as symmetric, 23% as *egalitarian* and 15% as *meritocratic*. Interestingly, if the best performers in Block 1 are selected as participants A, they tend to be slightly more *symmetric* and less *egalitarian* when they are aware of the selection rule (even though, the distribution of types in BL-High and LT-High are not significantly different). When bottom performers are selected as participants A, they get much more meritocratic when they are aware of the selection method: the distribution of categories in BL-Low is significantly different from the one we observe in LT-Low, as no *meritocratic* types are observed in the former and 38% of subjects A are classified as *meritocratic* in LT-Low. [13]

Figure 1.3 Ruling categories



[12] Given the low number of participants A per treatment we opt for this simple taxonomy strategy. A detailed report of the categorization by both rules is provided in Table 1.11 in the appendix. We did not find differences between the CT and the EN treatments relative to the BL. Figure 3 shows that both methods coincide in more than 80% of cases.
[13] We compare the distribution of types in BL-Low and LT-Low with a Fischer Exact test (p-value< 0.038).

4.2 No Power

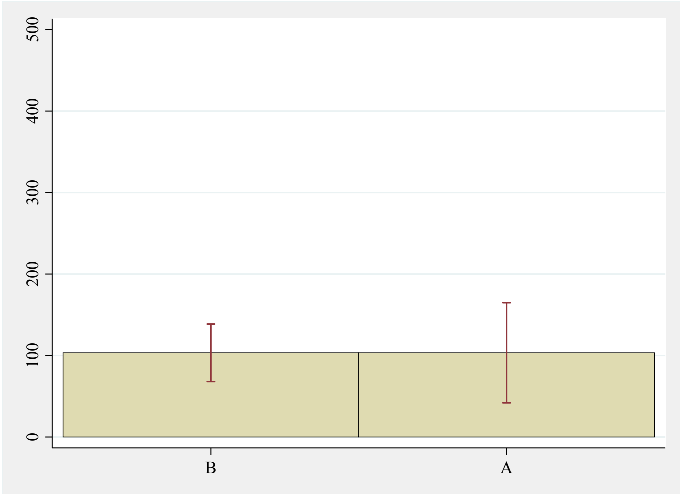
In the NP condition, the A-position does not bring along neither additional material benefits nor extraordinary decision rights. Are individuals still willing to buy and sell an apparently meaningless A-position? We answer this question in this subsection. If participants see the A-position as empty of any valuable content, those holding the A-position in block 2 should be willing to sell it at any positive price, but participants B should still refrain from posting any positive bid. Furthermore, perfectly rational decision makers (aware of similar levels of rationality in the other participants) will set a WTP = WTA = 0.

We find at least three realistic motives for participants to submit positive bids in the *No Power* condition: (i) individual are confused about the power exchange mechanism in block 3 and/or do not understand the instructions (as in Andreoni, 1995), (ii) they are willing to do what is implicitly asked to them in the laboratory by the experimenter (demand effects, as in Zizzo, 2010), and (iii) participants may intrinsically value being selected into the A-role, hold by a minority of them (one third) and labelled with the A letter (linked with the framing of the experimental setting). As this treatment is introduced as a control treatment for all these motivations together, we will not try to disentangle between these three competing explanations, but report the differences between those holding the A-label in block 2 and those holding the B-role. [14]

Figure 1.4 below presents the average WTP and WTA in the NP condition. By pure chance, A and B participants' WTA and WTP are on average the same: 103.3 ECU. Hence, bids are positive and close to the endowment they receive at the beginning of Block 3, consistent with a sizeable demand effect (and/or moderate levels of confusion or framing effects). As the gap between WTP and WTA is zero, we can conclude that in the power-free setting of treatment NP, the A-position per se does not generate any endowment effect (Khanemann and Tversky, 1979; Thaler, 1980; Kahneman et al 1990; Isoni et al, 2009), and that the experience of holding the label A does not alter how participants value it.

[14] All participants passed three very short comprehension quizzes at the beginning of blocks 1, 2 and 3, including multiple choice and open questions. Participants easily passed all quizzes in all treatments, and we did not notice any noticeable impact on their valuation of A-role. Even if their individual performance in the comprehension quiz were informative about the level of confusion, it would be highly speculative to conjecture about the magnitude of the other two motivations (demand and framing effects) from their decisions and their answers in the questionnaire.

Figure 1.4 WTA and WTP in the No power treatment



We summarise Figure 1.2 in our first result:

Result 1: When the A-role has no additional decision rights, no differences are observed between the willingness to sell the role and the willingness to buy. In other words, we do not find an 'endowment effect' in the No Power condition, and being selected as player A or B makes no difference in terms of their willingness to accept/pay in the power exchange stage.

4.3 Minimal Power

Figure 1.4 documents positive but indistinguishable bids for both roles when participants A are not provided with an additional decision right. We start studying the exchange of *minimal power* in block 3 in the simplest comparison. The only difference with respect to the NP treatment is that in the *Baseline-High (BL-High)* condition participants A now make decisions on the rates at which participants B earn. The selection mechanism used to allocate roles (Block 1's best performers become participants A) and the (lack of) information about this selection procedure is identical in both treatments. Thus, any differences across these two treatments should be driven by the existence of *power asymmetries* in *BL-High*.

Figure 1.5 presents the average WTA and WTP across treatments and roles in block 2. Whilst the WTA substantially (and significantly) increases in BL-High with respect to NP, the WTP is not significantly different in both conditions. In other words, while the WTA-WTP gap is zero in the NP treatment, it goes up to 280.0, more than twice the average WTP, in the BL-High treatment.

Figure 1.5 WTA and WTP BL High vs NP

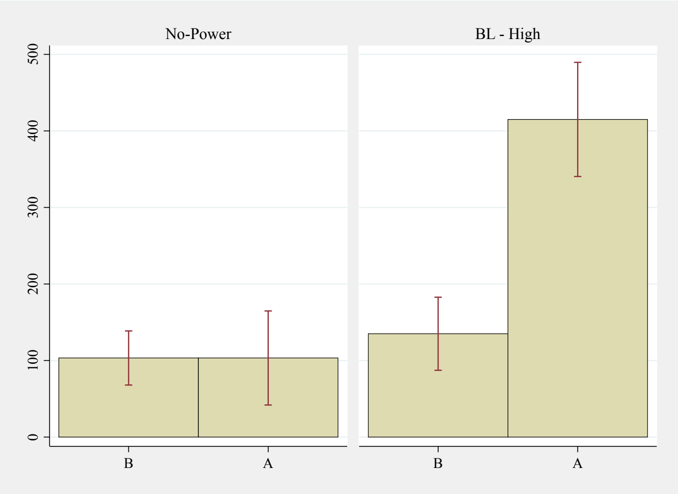


Table 1.3 WTP, WTA and Gap: BL-High vs NP

	A	B	Gap
NP	103.3	103.3	0
	(96.7)	(83.6)	(116.6)
BL-High	415.0	135.0	280.0
	(117.3)	(113.1)	(149.3)

We summarise Figure 1.5 and Table 1.3 in our second result:

Result 2: Keeping the selection procedure and the information constant, the exposure to power generates a vast and significant disparity between the willingness to accept and the willingness to pay for the A-role (that is, for power) in our experiment. While the willingness to pay for power of participants B in the BL-High condition does not differ from the WTP of participants B in the NP treatment, the average WTA of those participants with a previous experience of power is four times larger.

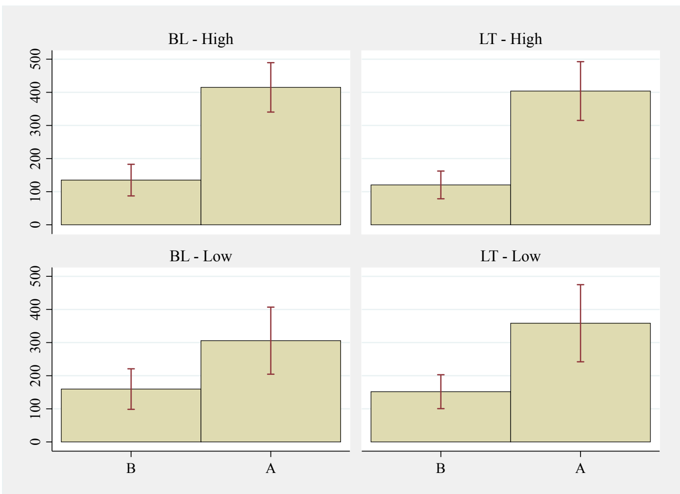
As explained in the previous section, in our 2x2 experiment design we carefully control for the role assignment mechanism (top- versus bottom-performers are selected as participants A in High and Low conditions) and the status of the decision rights by making the selection procedures common information only in the Legitimacy treatments (LT). By comparing BL-High and BL-Low we can tell if the exposure to power generates a different willingness to retain it for high and low skilled participants A. Note that even when in these two treatments of our experiment participants were never informed about the selection protocol, it could be that any exposure to power effect could be mediated by subjects' skills. The rationale for these differences are twofold.

First, being all our participants university students, and being the real effort task loosely related to the tasks they face as students, we cannot discard that high skilled individuals felt more entitled to more powerful positions because they had priors about their superior relative performance. Second, even when the selection procedure was not revealed to subjects, those participants A could easily compare their own individual performance in Block 2 with the performance of participants B in their group, as they were informed of individuals scores at the end of each round. This information could reinforce any positive correlation between individual skills and their sense of being entitled to power (role A), and this *entitlement effect* should reduce the distance between WTA and WTP in the BL-Low condition.

Similarly, by making explicit the selection procedure in the LT treatments, we exogenously manipulate the status of power. While participants are aware that the role A is reserved for the best performers in the LT-High condition, subjects can see that participants with lower relative performance in Block 1 were granted the same role in the LT-Low condition. This *status effect* should lower the WTA and WTP in LT-Low with respect to LT-High. We are quite agnostic on the net effect on the gap between WTA and WTP across both LT treatments.

The WTP, the WTA and the gap values in our four main treatments (BL-High, BL-Low, LT-High and LT-Low) are reported in Figure 1.6 and Table 1.4. Interestingly, we find that the second part of Result 2 is not affected by the *entitlement* and *status* effects defined above. The value given to the A-role by Participants B in the NP, as measured by their WTP for it, is not significantly different to the value given in any treatment in which the A-position comes with an additional decision right. Not even when the WTP values are pooled across *High* and *Low* conditions, or across *Baseline* and *Legitimacy* treatments. Hence, the WTP for power when participants have not been directly exposed to it (they have no previous experience of it) does not significantly change when participants see they are more entitled to it (as in BL-High) or when they do know the A-role is reserved to the best performers (as in LT-High) relative to the WTP observed in the BL-Low and LT-Low conditions.

Figure 1.6 The value of *minimal* power: WTP and WTA in the 2x2 conditions



The first part of Result 2, on the large and significant effect of power exposure on WTA does not change across the different treatments, yet they are significantly higher than the ones observed in NP (p-values of the Wilcoxon-Mann-Whitney test are 0.0001 for BL-High, 0.0044 for BL-Low, 0.0003 for LT-High, and 0.0014 for LT-Low), and significantly larger than the WTP observed in all four treatments. The magnitude of these significant differences make these results economically relevant, as participants in the experiment are posting WTA values in the 50% of their Block 2 earnings' range. The large WTA values (on average 409.4 and 332.0 in the *High* and *Low* conditions, and 360,4 and 381,1 ECU in the *Baseline* and *Legitimacy* treatments) generate substantial losses to those exposed to power in Block 2, as they fail to sell their decision rights in Block 3. We will carefully explore this finding in section 4.4.

The *entitlement* effect on the gap is significant when pooling data from High and Low treatments (p-value < 0.0465, see means and standard deviations at the bottom of Table 1.4). This result is driven by the entitlement effect alone; relative to the BL-High, the BL-Low gap is cut in half when participants A learn about their weak relative performance, free of any legitimacy or status effect (p-value = 0.0567). The gap is not statistically different in LT-High than in LT-Low, even when the former is 37% larger. The entitlement effect never eliminates the large disparity between WTA and WTP, as highly skilled participants B in the Low conditions bid significantly less than their lowly skilled counterparts, even when comparing the most productive tercile. [15]

Table 1.4 WTP, WTA and Gap 2x2 Conditions

		High	**Low**	**H & L**
Baseline	WTA	415.0 (117.3)	305.8 (159.4)	360.4 (147.8)
	WTP	135.0 (113.1)	159.6 (144.9)	147.3 (129.2)
	Gap	*280.0* *(149.3)*	*146.2* *(155.6)*	
Legitimacy	WTA	403.8 (139.6)	358.3 (183.2)	381.05 (161.0)
	WTP	120.4 (99.1)	151.7 (121.2)	136.05 (110.7)
	Gap	283.4 (163.7)	206.7 (191.9)	
BL & LT	WTA	409.4 (126.3)	332.0 (170.1)	370.7 (153.3)
	WTP	127.7 (105.5)	155.6 (132.2)	141.7 (119.8)
	Gap	281.7 (148.2)	176.4 (173.6)	

[15] See Table 1.9 in the Appendix for a similar analysis of Centrality and Enforcement treatments.

It is interesting to note that the *status* effect plays a tiny, marginal role with respect to the effect of being exposed to power in Block 2. The WTA in the BL-High and the LT-High (same selection rule not being or being common knowledge) are almost identical (415 versus 403.8), and the WTA in the BL-Low and the LT-Low are not statistically different (305.8 versus 358.3). Being exposed to a high or a low status power does not change the way power holders are willing to sell it. Similarly, the WTP in the BL-High and the LT-High (BL-Low and LT-Low) are strikingly similar: 135.0 versus 120.4 (159.6 versus 151.7). We cannot discard this could be a consequence of the information A players receive about their superior (inferior) relative performance in the BL-High (BL-Low treatments). However, it is worth noting that WTA are significantly and substantially higher than the WTP within all four conditions.

We summarise these findings in the following result:

Result 3: The exposure to power generates a large and substantial effect in every treatment in our factorial design, as the WTA are always significantly and substantially larger than the WTP. The difference is partially mediated by an entitlement effect, as the gap between them becomes significantly smaller in Low conditions. The status of power generates no sizeable effect, as the WTA and the WTP are never significantly different when comparing LT-High and LT-Low treatments.

4.4 Disentangling the components of the bids for power

Given the magnitude of the difference between WTA and WTP, and its limited sensitivity to the experimental manipulations we implemented in the laboratory, the natural question is to ask ourselves why participants in our experiment are willing to let substantial amounts of money go to retain power, when they decision rights attached to it have minimal economic value. In this section, we explore at the individual level how much of individuals' value of power can be explained by three rational reasons: risk attitudes, skill-dependent economic benefits and a residual linked with confusion, demand and framing effects.

(i) Risk Attitudes

As participants A earn at a certain and constant rate of 5 ECU, rational risk averse players may be willing to spend sizeable amounts of money to buy it, or may be reluctant to sell it, unless they are offered a sufficiently large amount to compensate for the loss of this certainty. In this subsection, we give risk aversion the best chance to explain the bids we observe before we deal with any potential material benefits of earning at a constant rate in the next subsection.

We use Holt and Laury (2002) seminal paper to calculate the WTA and WTP that extremely risk averse subjects would be willing to pay to keep or buy power in Block 3. To calculate the bid value compatible with the utility gains of being a participant A or B, we use the cut-off point that Holt and Laury (2002) used to define highly risk-averse subjects (between 1% and 4% of the population): Constant Rate of Risk Aversion (CRRA), $\theta = 0.97$. Assuming that individuals are able to anticipate the true distribution of rates (see Figure 1.10 in the Appendix 1.1), we can compare the certain outcome of earning at a constant rate with the uncertain payoff of a participant B being paid at an uncertain rate, that we obtain from the actual distribution of rates observed in the experiment with. [16]

We use the weighted probability of being paid at any rate observed in the corresponding treatment to compute the utility of an extremely risk averse individual, and then calculate the WTP (WTA) that would make her switch from the certain (uncertain) scenario and sell (buy) the A-role at the cost (benefit) of certainty. In other words, for each individual, we obtain the maximum bid for power he would make if being extremely risk averse. [17]

Table 1.4 presents the WTA and WTP values that highly risk-averse individuals would have set in the BL-High, BL-Low, LT-High and LT-Low conditions. [18] Even extremely risk averse individuals would only be willing to bid moderately to eliminate uncertainty. To put these numbers into perspective, Table 1.4 shows the percentage of the actual bids that these extreme values represent. An average slightly below 30 ECU represents, approximately, 8% of the WTA and 21% and WTP we observed. The cut-off WTA values are not statistically different from the WTP ones. [19]

[16] The assumption is of course unrealistic. It could be that individuals have unrealistic beliefs about the rates they could be paid. Given the very high CRRA we use, we still believe this is a strong test for the findings presented in the previous section. Moreover, any astronomically high CRRA would only generate higher WTA and WTP, and would never explain the difference. We run alternative assumptions about participants' beliefs in the next section.

[17] See Appendix A.1.2 for details.

[18] We report the same results for the CT and the EN treatments in the Table 1.10 at the Appendix 1.1, with no significant differences.

[19] We do not find significant differences for the riskier condition, namely the EN, in which risk aversion would make completely sense (see Table 1.10 in the Appendix A.1.1)

Table 1.5 Participation of risk aversion in the actual value of power by treatment and role

	Baseline		Legitimacy	
	A	**B**	**A**	**B**
High	39.58	28.00	38.67	26.46
%	10%	21%	10%	22%
Low	17.50	30.04	20.83	32.79
%	6%	20%	6%	22%

We summarise this finding in our result 4:

Result 4: Risk aversion can only explain a moderate proportion of the bids observed in the experiment, even if we assume very high CRRA for all participants: 8% of WTA and 21% of WTP.

(ii) Material Benefits

Even though our design imposes an identical average rate for A and B participants, a constant rate of 5 ECU could still be economically attractive for those participants B earning at a rate lower than 5 ECU in block 2, [20] or for those participants A fearing a similar outcome in block 4. A powerful position would protect participants form this potential payoff loss, giving them another rational motive to bid for it. In this subsection, we study how different individuals could measure the material benefits by looking at the only information participants A had when deciding on rates: their productivity. [21]

We approximate the expected earnings of participants holding the A role in block 4 ($\pi_{i,A}^{e}$) multiplying their individual score in block 2, S_i^{b2} , by the constant rate of 5. The potential earnings for the same individual in the B role ($\pi_{i,B}^{e}$) would be her score in the same block 2 multiplied by the average rate at which similarly productive participants B, in the same treatment, earned (r_B^{avg}). The material benefits of the A role would the difference:

$$\overline{\pi}_i = \pi_{i,A}^{e} - \pi_{i,B}^{e} = S_i^{b2}\left(r_A^{avg} - r_B^{avg}\right) \quad (1)$$

The economic benefits of the role A are reported in Table 1.6. We find that the material benefits are less than 10% the bid values in all conditions. [22] In fact, we find that the economic benefits of the A-position can be negative when participants A are more

[20] 5.2% of participants B earned at a rate lower than 5 in at least 6 rounds in block 2.

[21] In this part of the chapter we exclusively consider individual productivities. A more sophisticated exercise could consider the relative performance of individuals within each group, rather than the absolute level of performance.

[22] Participants A assign rates of 5 ECU 60% of the time. These results are basically the same for the EN and CT treatments as reported in Table 1.11 in the Appendix A.1.1.

generous with participants B that are as productive as they are (as in the BL-Low). In general, we do not have strong evidence to suggest that the WTP and WTA values are largely driven by the material benefits linked to holding the power, as Table 1.6 and Result 5 document:

Table 1.6 Participation of the material benefits in the actual value of power

	BL-A	BL-B	LT-A	LT-B	All A	All B
High	30.2	3.7	17.0	2.1		
%	7.3%	2.8%	4.1%	1.8%	5.70%	2.30%
Low	-23.2	6.2	24.4	2.6		
%	-7.6%	3.9%	6.8%	1.7%	-0.40%	2.80%
	-0.15%	3.35%	5.45%	1.75%	2.65%	2.55%

Result 5: The material benefits linked to the difference between the participant A rate, 5 ECU, and the rates awarded to participants B with specific productivity levels, only explains a very moderate proportion of the bids observed in the experiment (on average, 2.65% of all WTA and 2.55% of all WTP).

(iii) Noise Residual (demand, confusion, plus framing effects)

Since in the NP condition participants A and B certainly earn at the same rate, the WTP and WTA values cannot be explained by risk aversion, material benefits, the additional decision right (power), or its experience. Still, positive bids submitted by participants A and B in the NP treatment could be generated by an experimenter demand effect, as described by Zizzo (2010), confusion regarding the mechanism used in Block 3, or a preference for the role A, as identified by Eckel et al (2016), and mentioned in section 4.2. As this residual could be driven by demographic characteristics of subjects, we estimated different models to control for any differences across the different treatments. Even when we admit the sample in the NP treatment is quite small, we did not find any systematic and significant demographic explaining the size of the demand-confusion-framing effect. Thus, we use the average WTP and WTA values in the NP condition to control for the proportion of *noise* that the specifics of the experiment might be generating on individuals' bids. Table 1.7 below puts into perspective how WTA/WTP observed in the NP treatment count as a fraction of the average bids observed in each treatment:

Table 1.7 Participation of noise (confusion-demand-framing)

	BL-A	BL-B	LT-A	LT-B	All A	All B
High	415.00	135.00	404.80	120.40		
%	24.89%	76.52%	25.57%	85.80%	25.23%	81.16%
Low	305.8	159.60	358.30	151.70		
%	33.78%	64.72%	28.83%	68.09%	31.31%	66.41%
	28.66%	70.13%	27.10%	75.93%	28.27%	73.78%

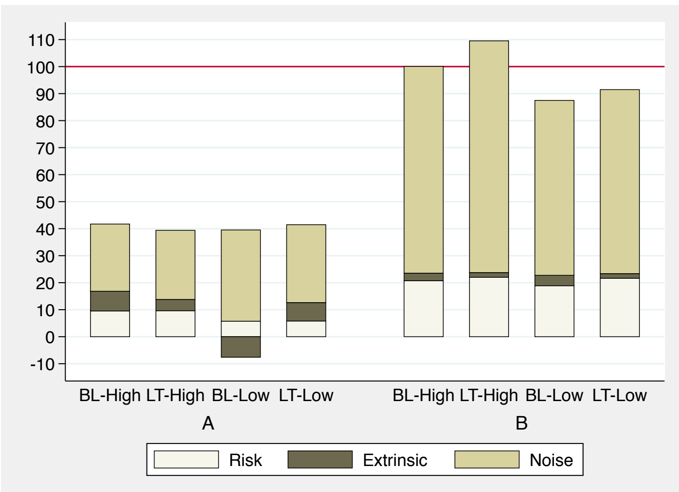
While we cannot exclude a large fraction of the WTP (almost 73% across all four treatments) is consistent with the combination of confusion, demand effects and framing, the percentage goes down to 31% of the WTA, and we summarise this finding in our Result 6:

Result 6: The effect of confusion, demand and framing is quite limited for participants A (28.3% of actual WTA, across all four treatments) and quite substantial for participants B (73.4% of the actual WTP, across all four conditions).

4.5 Summary

Figure 1.7 below pools all three components into one single visual result. For each treatment and role, we can see how the different explanations may account for the bids we obtained in our experiment, in percentages of the actual values. From the relative size of bars, by inspection, we can see that the combination of risk aversion, extrinsic material benefits and noise (confusion, demand and framing effects) may explain most of the WTP submitted by participants B in our experiment (around 95% of the average bid, across all four treatments). Even when we are giving each component its best scenario, the value of power when participants in the experiment have no previous experience of it can be fully rationalised by a combination of effects, none of them related to power per se.

Figure 1.7 The Components of the WTA and WTP values



The big picture is very different for those participants with a direct exposure to power. The left part of figure 1.7 shows that using the same method we applied to participants B, in each experimental condition, we can only explain 40% of the bids submitted by participants A. In other words, more than half of the bids need to come from the only difference between those subjects on the left and those in the right part of Figure 1.7: the experience of power. We summarise this finding in our next result:

Result 7: Not all participants in the experiment were willing to pay significant and substantial amounts of money for power (the additional decision right). While we are able to almost fully rationalise bids submitted by Participants B (without exposure to power) using a combination of standard economic (material benefits, risk attitudes) and behavioural factors (confusion, demand and framing effects), at least 50% of the bids submitted by participants A, with a previous power exposure, cannot be explained by this combination of factors. This gap is consistent with the existence of a vast, intrinsic, substantial and significant value of power in every experimental condition.

4.6 Underbidding bias

The market exchange we created in block 3 of the experiment consisted of a two-side auction in which A and B participants could sell and buy an additional decision right on rates ('power'). The WTA values chosen by participants A in each session were randomly matched with the top 50% WTP values chosen by participants B. Rational participants A have an incentive to truly reveal their WTA particularly when, in case of a successful transfer of power, they receive in full the WTP value they are matched with. If they chose a WTA higher than their true value, they would miss the possibility of receiving an acceptable compensation for the A-position they hold. If they chose a WTA below their optimal bid, they could be transferring the A-position they hold for a price below their true valuation.

We were fully conscious that this mechanism worked very differently for participants B, as there is no incentive compatible mechanism for both sides of a market like the one we implemented in our experiment. The first half of participants B leaving the auction had every incentive for truly representing their individual valuation for power in their bid. By bidding higher, they would pay their bid in full, if matched with a sufficiently low WTA, and the bid would be above their true value. By leaving the auction before the ascending price reached their true valuation of power, they would be missing the opportunity of buying it at a price below their true valuation. Hence, assuming full rationality, this half of participants B are also truly revealing their value for the A position in the bid.

For the other half of participants B, the WTP elicitation method might generates a bias because their bids are always paid in full, if successfully matched. This bias could explain the disparity we observed in the experiment between WTA and WTP values, as participants B could best respond to the expected WTA by posting a bid below their true value of power. We explore how this possibility could change our results interpretation in this section.

We estimate this bias by making different assumptions about the expectations of participants B on the WTA. We use four different distributions of WTA values for both *high* and *low* conditions, and compute the WTP that maximises their expected earnings, given each distribution of WTA. To obtain the probability of buying the power, we use the real distribution of WTA observed in each treatment, a normal distribution, a Poisson distribution (in both cases using the actual means and standard deviations of each treatment), and a uniform distribution. Rational participants B should simply react to these distributions by choosing the WTP_i that maximises their earnings, given their true value (V_i) [23]:

$$\underset{WTP}{\operatorname{argmax}} \sum (V_i - WTP_i) * P(WTP \geq WTA) \quad (2)$$

The idea behind this exercise is simple. For one expected distribution of WTA values, say uniform with one WTA in the ten multiples of 50 going from 50 to 500, a rational bidder getting in the second half of the auction, with a true value of, say, 200 would be better off by bidding below 200. By bidding 200, this participant B would have the chance of buying the A role 4 times out of ten, and misses the opportunity in the other 6 (when matched with WTA of 250, 300, 350, 400, 450 and 500). However, her expected earnings are zero, as she always pays the price in full, and the price exactly matches her true value.

If bidding 150, this same participant B buys the decision right 3 times out of ten, and misses the other 7 (when matched with WTA between 200 and 500). Now the bidder makes the difference between her true value (200) and the price paid (150, her bid) each time she buys, making an expected profit of 15, (0.3 * 50). It is not difficult to see that

[23] Being $P(WTP \geq WTA)$ the probability of buying i.e. the probability for the WTP being at least as high as the WTA.

the optimal bid for this participant B would be 100, as she would still buy 0.20 of the times (when matched with WTA of 50 and 100), making 100 each time (and the expected earnings would go up to 20). By bidding below the optimal WTP (say, 50), she is increasing the profit to 150, but she is also reducing the probability to 0.10, and the expected earnings to 15 (as when bidding 150).

For each assumed distribution of WTA, we can interpret the bids (a WTP of 100 in the example) as the rational WTP submitted by a participant B who is trying to maximise her expected earnings, and estimate their real valuation of power from there.[24] Table 1.8 and figure 1.8 summarise the corrected WTP, estimating the average optimal bid participants B should submit if all were fully rational.

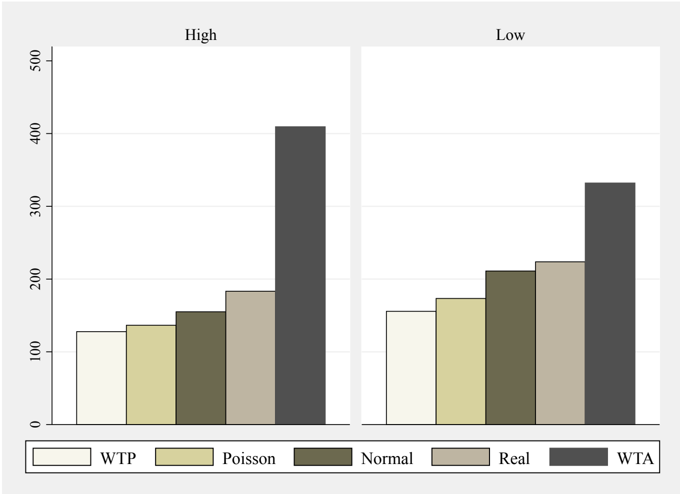
Table 1.8 Corrected WTP

	WTP	Normal	Poisson	Uniform	Actual	WTA
High	127.7	155.0	136.5	226.7	183.2	409.4
	(105.5)	(135.4)	(122.3)	(225.9)	(154.9)	(126.3)
Low	155.6	211.0	173.3	278.5	223.7	332.0
	(132.2)	(196.7)	(167.8)	(285.8)	(207.7)	(170.1)

After controlling for the underbidding bias, the WTA values are still significantly higher than the corrected WTP, for each distribution. Even if we pool data from High and Low conditions, all corrected WTP values are significantly lower than the corresponding WTA. In the *high* conditions the WTA is significantly higher than the WTP with a p-value<0.001 (Wilcoxon-Mann-Whitney test) for all the distributions considered. In the *low* conditions, WTA are always larger than the corrected WTP (p-value<0.001 for the Normal and Poisson distributions, p-value<0.05 for the *actual* distribution, and p-value<0.06 for the Uniform).

[24] This method is sometimes inconclusive (when two WTP generate the same expected earnings for one specific true value) and may generate a second order bias, as a rational bidder with a true value between 200 and 249 and the beliefs described in the example would still bid 100. We assume that all possible true values are equally likely and compute the average. The number of inconclusive cases is quite moderate, given that only 50% of participants B can gain by budding below their true value.

Figure 1.8 WTP, corrected WTP and WTA



We believe we are again giving our data the strongest test, as we consider this method to be an upper bound of any underbidding bias for two reasons. First, for each distribution of WTA considered, but the Poisson which does not differ from the real distribution, and each experimental condition, the correction breaks the distribution of bids in two parts, creating arbitrary and artificial jumps in exactly the 50% cut-off point between the lower half and the upper half of WTP (as shown in Figure 1.13 in Appendix A.1.1). Second, the method would generate a massive jump of WTP in the NP condition, making the WTP to be significantly larger than the WTA we observe. [25] We summarise these findings in our next result:

Result 8: Even correcting for the underbidding bias and considering different types of beliefs distributions, the exposure to power in Block 2 generates a substantial and significant increase in the bids in Block 3.

4.7 The Effect of Power on Individual Productivity

We finish the analysis of our results by specifically exploring one result of previous studies. Fehr et al (2013) and Falk and Kosfeld (2006) found that controlling parties tend to overprovide efforts, and subordinating parties do the opposite. This is consistent with other studies that found that the control over someone else's efforts crowds out the intrinsic motivation for individuals to perform on the principal's benefits (Charness et al

[25] Figures with the distributions and the NP results are available in the Appendix.

2012). In this section, we study to what extent being a participant A (controlling party) or being a participant B (controlled party) shapes individuals' performance.

Given that participants in the experiment face an identical real effort task once in Block 1, and ten times in Block 2, we focus in the productivity gains of participants A and B calculating the difference between the average score per minute in blocks 1 and 2. Figure 1.9 below presents the average productivity gains of participants A and B once the roles are assigned in the *NP, High and Low* conditions, pooling the data coming from *BL* and *LT* sessions.[26]

In every treatment, we use individual productivity in Block 1 to assign roles. As explained above, we rank individuals by their individual score and assign roles (A and B) based on whether they are in the top or bottom tercile. In both BL-High and LT-High (BL-Low and LT-Low) treatments, subjects in the top (bottom) tercile in Block 1 become participants A in the second Block. We compare the productivity gains of those participants in the top and bottom tercile across the different conditions, and depending on whether they are a 'controlling' (getting role A) or 'controlled' (B) party in Block 2. [27]

Figure 1.9 below shows the average productivity gains in absolute terms for two groups: the top and bottom tercile of participants (left and right side, respectively). For each group, we plot the productivity gains using three bars: the left one for the *No power* treatment, the one in the middle for High treatments, and the one in the right of each box for Low conditions. The two boxes tell a very different story.

Top performers (left box of Figure 1.9) do better when assigned the A role (as in High conditions) than when they become participants B (as in Low conditions), but the difference is not significant (and the magnitude of the effect is not significantly different from what we observe in the NP control). Bottom performers do significantly react to the role assigned to them, but in a counterintuitive way, as being in a *powerful* role (as in the *Low* treatments) generates no significant productivity gains above the ones observed in the NP control. However, they do significantly better when they become

[26] We pol data to make the visual comparison easier. Results do not change if we present all treatments independently. Figures are available upon request.
[27] We do not include the 2nd tertile in the graph to make sure that A and participants B' productivities are comparable. While participants A in the *high* (*low)* condition and participants B in the *low (high)* condition belong to the 1st (3rd) tertile.

powerless participants B (as in the *High* conditions): their productivity gain almost double the ones observed for participants A in the *Low* treatments (p-value = 0.0048) and almost triple the productivity in the NP condition (p-value = 0.0008).

Figure 1.9 The effect of power on individuals' productivity

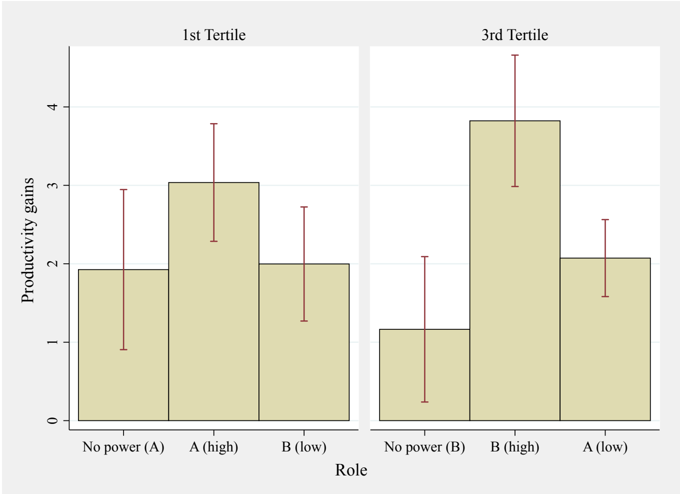


Figure 1.9 does not control for the rates chosen by Participants A, for whom we found some average evidence of egalitarianism (i.e. they rewarded less productive participant with higher rates and more productive participants with lower ones). Even when the egalitarian behaviour of participants A reduces the incentives for participants B to be more productive and does not invalidate figure 1.9, the connection between productivity and roles should be taken with extreme caution in this setting.

We summarise our findings in the last result:

Result 9: We fail to replicate the positive effect of power, control and authority on individual's performance (as in Fehr et al 2013 or Falk and Kosfeld, 2006). Instead, we find that low-skilled participants B (controlled individuals) exhibit significantly higher productivity gains than low-skilled participants A (controlling individuals)

5. Conclusions

The toxicity of power in daily life has brought us to this study. The aim was to understand the determinants of the value of power that are not driven by extrinsic benefits or strategic interactions. We recreate in the lab real-life features of power relationships based on an asymmetric distribution of rights that do not necessarily imply the possibility of getting pecuniary benefits from it. We elicit the value of power and create a "market of power" by using contingent valuation methods from those who have and have not being exposed to powerful positions i.e. the WTP (from participants B) and WTA (from participants B).

We find that even though the WTP for the A-position is positive and non-negligible, over 90% of this value can be rationalised by other attributes, such as the material benefits of the A-role, participants' risk attitudes, and the specifics of the experimental setting.

The WTA value, on the other hand, cannot be explained by those components. On the contrary, participants A who have been in touch with power and have had the taste of exercising it, are willing to forgo the possibility of a profitable role exchange; even by conceding high levels of risk aversion, and controlling for material benefits and the potential *noise* of the setting and individuals' priors. They develop a strong attachment to the experience of power that leads them to set a WTA value that is, on average, 3-times the WTP.

The results obtained in the NP condition suggest the existent gap between the WTA and WTP is not necessarily generated by the power asymmetry. Since, the WTP is the same in the No-power condition and in treatments with power asymmetries, the gap is entirely driven by the WTA value. Therefore, it is the *experience* of power what is leading the whole reaction of our participants when they are facing the possibility of a role exchange. Participants A, who have had the power during 10 rounds in a row, are more reluctant to give up their position and consequently, they ask for a compensation that exceeds 70% of participant Bs' earnings in the block 2. This result prevails even after controlling for risk attitudes, the limited economic benefits of being appointed as A and the potential underbidding bias.

It seems that the WTP-WTA gap here observed is partially mediated by an entitlement effect, as we find that the gap between them becomes significantly smaller in Low conditions. The status of power introduced in the legitimacy treatments generates no sizeable effect, as the WTA and the WTP are never significantly different when comparing LT-High and LT-Low treatments.

We also find that the WTP and WTA variance is not explained by other attributes given to participants A such as the ability to enforce or their visibility at their group members.

Finally, we find that individuals exhibit a passive way of exercising power (45% were classified as *symmetric*), and if using it they would compensate participant B's earnings (23% were classified as *egalitarian)*. power does not boost individual productivity and control does not reduce the subordinating party productivity when the decision right owners cannot control their efforts (as in Bartling et al 2014, and Charness et al 2012); even if they have the right to decide upon their rates.

Chapter 3. The Cost of Dissent[28]

[28] Co-authored with:Rachel T. A. Croson, Department of Economics, Michigan State University, crosonra@msu.edu, and Enrique Fatas, School of Economics, University of East Anglia, e.fatas@uea.ac.uk

1. Introduction

Both coordination problems and teams' interactions are ubiquitous in a variety of important economic contexts. This observation has fuelled the large body of research in economics on coordination games in general (predominantly games with multiple, praetor-ranked equilibria), and on team decision-making. However, it was not until the paper of Francesco Feri et al (2010) that researchers brought these two critical topics together.

There are two main concerns when it comes to boost team performance: coordination and free riding. In this study, we study two polarised social composition functions (SCF) (Harrison and Hirshleifer, 1989) that have been widely studied in the team production literature: the weakest-link (WLM) and the voluntary contribution mechanism (VCM). Our main objective is to compare these SCF with one hybrid mechanism through which we attempt to get closer to how companies or groups of teams actually organise their production process[29].

We study the underlying behaviour that differentiates individuals and teams whilst producing a public good with a WLM technology. To avoid any nominal confusion, hereafter we will refer to a *team* as the set of individuals that are jointly deciding on the team contribution to the public good production. Therefore, the set of contributors to the public good, either teams or individuals, will be appointed as a *group*[30].

The WLM is a SCF through which group production is in the end as good as the minimum contribution made by one of the group members. Each individual receives *r/n* times (r>1 and n the number of individuals) the amount produced by the group. As suggested by Harrison and Hirshleifer (1989), this SCF gives to each group member a kind of veto power over the extent of the collective achievement. Coordination failures may lead groups (or organisations, and society) to a disastrous result. By construction, free riding is impossible in the WLM, as low contributors do not benefit from the decisions made by cooperative individuals.

The WLM addresses coordination problems as it incentivises all individuals to contribute as much as their group mates have contributed. Theoretically, the WLM is a multi-equilibria mechanism on all the strategy space. Hence, this mechanism could generate coordination at high and low levels of production, depending on the pathway followed by the teammates. The cost of coordination comes from the wasted effort exerted by individuals, as any contribution above the minimum is lost.

[29] Social Composition Function (SCF) (Harrison and Hirshleifer, 1989) refers to the mechanism through which individual contributions are alloyed into an available social aggregate of the public good
[30] Mind that this is only a nominal precision. Groups, in this context, are the social figure through which all contributions are computed. Thus, groups are what in the team production literature are known as a *team.*

Group production under the VCM, on the other hand, is the "democratic" result of the average of all contributions made by the group members. As for the WLM, Each subject receives *r/n* times the amount produced by the group. The VCM is a unique-equilibrium game on *zero* as individuals always have incentives to contribute a bit less than their group mates in order to take advantage of the group account. Thus, the VCM jeopardises the incentives for individuals to contribute to the public good, as it is more profitable to *free ride*.

Here a real anecdote of how individual incentives might not be aligned to the team's. The 28th of November 2016 an aircraft flying, from Bolivia faced a fuel emergency and crashed on Colombian ground generating 71 casualties. According to the news, the flight crew failed to adequately manage the airplane's fuel load, and did not use the protocol to report the emergency fuel situation to air traffic control. Apparently, by evading the protocol the captain was simply avoiding further investigations from the aviation authorities.

In this study, we explore the performance of teams in a SCF that combines the VCM structure with the WLM. Our rationale for this combination of two well-known SCF comes from real-life: organisations embed their production within a WLM structure between teams (the ones facing the coordination challenge) in which individuals still face a non-trivial VCM-like dilemma within each team (and team members face a social dilemma challenge).
In our main setting, teams contribute to a group account. The minimum team contribution determines the joint production of the group, so each team has an incentive to contribute as much as other teams have contributed (and free riding is not profitable for teams). We employed the average of the team members' contributions as the decision rule to compute the team contribution to the public good. Hence, individual team members still benefit from the contribution of other teammates (opening a back door for individual free riding).

A close example of these types of organisations is the Continental Airlines' crews. Several teams of an airline work together to make sure that the airplanes depart and land on time, and that the baggage is properly allocated. Each team yields the average effort of each team member, yet the airline production responds to the minimum effort exerted among the ground crew teams.

We experimentally test the effectiveness of teams in this environment using a between-subjects design. Because our environment has elements of both the WLM SCF (between teams) and the VCM SCF (within teams), we use a double benchmark: a repeated team production game played by individuals with a WLM SCF (I-WLM hereafter) and a VCM SCF (I-VCM hereafter). In our main treatment, teams jointly produce outcome using a WLM SCF, while individuals within teams still bear in full the cost of their contributions (T-WLM hereafter).

The recent literature comparing teams versus individuals suggests that results are context dependant, and that groups may be more sophisticated that individuals with an open effect on social welfare (see lessons 1 and 3 in the excellent survey of Charness and Sutter, 2012). Feri et al (2010) impose unanimity as a team decision rule and eliminate free riding within teams (e.g. effort levels are always identical for team members). In their experiment, teams coordinate better and are more productive than isolated individuals. In our experiment, we find the opposite result, in line with the literature suggesting that teams may be more sophisticated and more self-interested than individuals: individual contributions to the group account, and the overall team production, are significantly smaller in T-WLM than in the two individual benchmarks. However, wasted efforts (contributions above the minimum) are not significantly higher in teams. I-WLM exhibits a much larger variance of results, generating three different group categories: those who coordinate in a high-efficiency equilibrium, those who coordinate in a low-efficiency one, and those who try to coordinate, and fail. Groups of teams in the T-WLM condition are only outperformed by successful teams in the I-WLM treatment. In summary, teams do not coordinate worse than individuals in our setting, but they coordinate in a low effort equilibrium.

Our design, however, does not consider other important elements of team interaction in real organisations. There is still room to investigate the effect of communication (without forced consensus), team identification, inter and intra team punishment, peer monitoring, leadership, gradual team integration and neighbourhood selection. All these other elements are left for future research comparing teams and individual performance.

The rest of this chapter goes as follows: in the second section, we present relevant literature comparing teams' and individuals' behaviour; in the third section, we describe our experimental design and procedures; in the fourth section, we present our results; and, finally, in the fifth section we conclude.

2. Related Literature

The literature comparing team and individual decision-making suggests that the relative performance of teams and individuals is context dependant ('lesson three' in the excellent survey written by Charness and Sutter, 2012). The evidence surveyed in this paper cleanly supports the superiority of teams over individuals overcoming cognitive limitations (becoming its 'lesson one'). As an example, Kocher and Sutter (2005) find that groups are more sophisticated than individuals in beauty context games, as even though they both start with similar predictions, groups apparently are better at learning from their results in previous rounds and converge faster to the Nash equilibrium. In other words, teams are more sophisticated than individuals because they make more rational choices in a standard game-theoretic sense.

The welfare implications of cognitive sophistication depend on the strategic environment. In a trust game in which senders and responders can be either teams (of three) or individuals, Kugler et al (2007) allow for communication within teams, imposing unanimity to reach a joint decision about how much to send or to return. Teams get closer to the equilibrium of the game and tend to send less (trust less), yet return roughly the same (equal trustworthiness), than individuals. As in trust games the second mover has a mere redistributing role, teams significantly reduce social welfare.

'Lesson 2' in Charness and Sutter (2012) gives team a welfare enhancing channel, as peer effects may boost the performance of individuals when making decisions in teams. Peer and monitoring effects have been documented in the laboratory and the field. Falk and Ichino (2006) cleanly show that low productivity individuals perform better when paired with other individuals doing the same simple task (stuffing envelopes), even when their individual performance does not affect co-workers' pay. High-productive supermarket cashiers have a similar positive influence in Mas and Moretti (2009). Even when customers can freely move from one slow cashier to a faster line, generating incentives to free-ride on the efforts of others, low-productive workers increase their performance when working with more productive peers (and the productivity of highly skilled workers is not hurt by the presence of low-skill co-workers). The positive spillovers are observed only among cashiers when located close in the supermarket and when they frequently share the shift, suggesting that peer effects can be related to social pressure[31].

Paraphrasing Charness and Sutter, the lesson we learn from the literature review so far is that peer effects may boost team performance by increasing the productivity of low-ability workers, making it attractive for high-ability workers to engage onto group compensation conditions. Using data from a garment factory, in which individuals can choose the team they want to belong to, Hamilton et al (2003) find that peer monitoring reduces free riding, and group piece-rates may generate large and significant productivity gains. Interestingly, the positive effects of peer monitoring may disappear when free riding is impossible. Guryan et al (2009) find no peer effects in randomly generated pairs of players competing in golf tournaments, in which players compete against all other participants in the tournament.

[31] Kato and Shu (2008) and Bandiera et al (2009) find similar results among Chinese textile workers and British fruit pickers, respectively.

Experimental studies investigating differences between teams and individuals in coordination environments are rare. Charness and Jackson (2007) use a Stag-Hunt game and teams of size two to compare teams and individuals decision-making in a coordination task by changing both the incentives and the decision-implementation mechanism (unanimity vs unilaterality). They find that the decision rule critically affects the results: while in the latter scenario (as in *unilaterality*) individuals choose Stag more frequently than individuals, the opposite results hold if both team members must vote Stag (as in *unanimity*) to implement it as their team choice. In other words, while individuals playing against other individuals follow an incentive-based logic, the decisions of teams playing against other teams is more likely to be driven by the voting rule mechanism.

Feri et al (2010) follow van Huyck et al (1990) canonical design and compare teams and individuals' performance. Decision makers (5 individuals or 5 teams of 3 individuals) choose a number (i. e. effort) between 1 and 7 in a coordination game with 7-pareto ranked equilibria. Teams could communicate and had to reach a unanimous decision before submitting a decision. Note that, by design, forced consensus excludes free riding within teams, as all individuals within each team exert the same effort level (i.e. identical payoffs are obtained by all three team members). In this free-riding free environment, teams coordinate more often than individuals (e.g. decision makers choose the same number), and they coordinate better (e.g. on higher numbers). Chaudhuri et al (2015) replicate Feri et al (2010) with teams of two players, and find that their results only hold when the composition of groups is kept constant along the experiment, and peer monitoring is possible. When group composition changes from round to round, individuals outperform teams.

3. Experimental Design

Our study contributes to the literature comparing teams and individuals' behaviour in coordination tasks, allowing free riding situations within teams. Consensus is not imposed and individual payoffs within each team may or may not be identical. Following Chaudhuri et al (2015), the composition of teams does not change. As teams and individuals make anonymous decisions in a controlled laboratory environment, peer monitoring between the participants in our experiment is much weaker than among cashiers (as in Mas and Moretti, 2009), farm workers (as in Bandiera et al, 2009) or employees in the garment plant (as in Hamilton et al, 2003). The feedback provided to the participants in our setting, however, allows for a weak form of peer monitoring, giving teams a fair chance to outperform individuals.

To examine the impact of the *within*-team decision-making process on the *between*-teams' coordination outcomes, we present a new set of experiments focusing on the weakest link game. Contrary to Feri et al (2010), we compare individuals and teams *without* enforcing consensus on effort levels. Instead, we allow individuals within each team to choose an effort level (much like the Continental Airlines' ground crews in our example above). The team effort is determined by the average of the individual efforts of its members. In this sense, our design falls between the two treatments of Charness and Jackson (2007, 2008), as the team members decisions are all used to calculate the team effort.

It is worthy to stress the attention on the boundaries of communication and unanimity on team performance. Teammates are normally allowed to communicate to each other. Yet, this communication does not necessarily happen in the very same moment when actions are taking place (the pilot do not talk to the flight attendant to consult a landing issue whilst landing). Individuals rather make decisions on their own even when those would affect the team performance. Besides, while it is true that teammates might commit with certain action beforehand, spoken commitments are not necessarily binding, and subjects can always deviate to a decision that represents a better deal for them. That is why we created a team-contribution aggregation rule that allows for self-interested decisions; but, at the same time, we imposed a group coordination challenge that should interfere in individuals' incentives to free ride.

We examine three games in which groups of individuals or teams produce a joint outcome with WLM and VCM SCF, and then we compare individual and team treatments. In our individual treatments (I-VCM and I-WLM), four individuals made a group. In our team treatment (T-WLM), four teams of three individuals each made a group[32]. So, *decision makers* were individuals or teams (of three individuals) depending on the condition.

In each group, individuals or teams faced a simple WLM or VCM environment. Each participant received a 50-unit endowment at the beginning of each round, which they could assign to a group or a private account. In both I-WLM and T-WLM treatments, the minimum contribution to the group account made by the decision makers (individual in the I-WLM or team in the T-WLM) determined the group production, and each participant received twice the minimum value. In the I-VCM the average contribution determined the group production, and each participant received twice the average contribution to the group account.

[32]Although we used groups of size four rather than size five, this was held constant across the individual and team treatments. Previous work has suggested that coordination is easier in smaller groups (e.g. Roberto Weber 2006), thus it was possible that our smaller-sized groups could have increased coordination success in both the team and individual treatment. However, we find no difference in effort or coordination among individuals between our study and Feri et al (2010) (and lower efforts and coordination among teams), thus we conclude that this was not a significant difference in the experimental design.

In the T-WLM treatment, the three-team members' allocations to the group account were averaged to decide the team allocation. [33] Contrary to other studies (Kugler et al, 2006; Feri et al, 2010), we did not impose consensus within the team; all three team members independently entered their decisions, and paid the cost of their decision in full. T-WLM follows the Continental Airlines example used in the introduction, as the low performance of one team reduces the entire group's outcome (e.g. if the ground crew delays the plane's departure, then the plane will not land on time, regardless of the effort levels of the flight crew or the arriving ground crew). At the same time, unanimity is not magically imposed within each team (e.g. the ground crew), and individuals still face a non-trivial dilemma (e.g. members of the ground crew may free ride on the contribution of other ground crew members).

The payoff functions of the two individual treatments are the following:

$$\pi_{ind}^{I-VCM} = (50 - c_{ind}) + 2 \cdot c_{ind}^{avg} \quad (1)$$

$$c_{ind}^{avg} = \sum_{i=1}^{i=4} c_i \quad (2)$$

$$\pi_{ind}^{I-WLM} = (50 - c_{ind}) + 2 \cdot c_{ind}^{min} \quad (3)$$

$$c_{ind}^{min} = \min\,(c_1, c_2, c_3, c_4) \quad (4)$$

The payoff function for the T-WLM treatment includes a WLM SCF for the group account (as in the I-WLM):

$$\pi_{ind}^{T-WLM} = (50 - c_{ind}) + 2 \cdot C_{team}^{min} \quad (5)$$

and a VCM SCF within each team (as in the I-VCM):

$$C_{team}^{min} = \min(C_1, C_2, C_3, C_4) \quad (6)^{34}$$

$$C_j = avg(c_{ind}) = \frac{\sum_{i=1}^{i=3} c_i}{3} \quad (7)$$

The I-VCM has a clear theoretical prediction in the stage game. The chosen parameters promote free-riding attitudes, assuming self-interested preferences. Whereas each unit of effort allocated to the private account yields one unit for the individual utility, each unit of effort allocated to the group account (effort) yields a half. In other words, from any symmetric contribution profile, individuals can always unilaterally deviate reducing their contribution to improve their profits: if one subject reduces her individual contribution in one, she increases her individual payoff in half unit. From the group's perspective, full contributions are the efficient strategy as whereas each unit of effort allocated to group account earns a return of two (divided equally among four members), each unit of effort allocated to individual activity earns a return of one.

[33] As we aimed to use the VCM mainly as a theoretical control, and a team VCM production would be just as a 12-people-size VCM, we decided only comparing our results with the traditional VCM played by four individuals.

[34] Uppercases refer to teams contributions, and lowercases represent individual contributions.

The strategic problem of each participant in the I-WLM is the trade-off between the opportunity costs of exerting too little effort and the costs of wasted effort from exerting more than the minimum within her team. Each symmetric contribution profile (i.e. any allocation in which every participant makes the same effort) becomes a pure strategy Nash equilibrium of the stage game. In other words, from any symmetric contribution profile, an individual should not unilaterally deviate at a profit. If she increases her contribution in one unit, she decreases her profit in half (without affecting the group minimum). If she reduces her contribution in one unit, she would decrease everyone's individual payoffs (because the group minimum will be reduced in one unit, and in the group benefits in two units). Equilibria are Pareto-ranked in the sense that the pay-offs in an equilibrium is the same to all subjects and increases linearly in the minimum effort. The pay-off–dominant equilibrium is the one with maximum effort, and the zero-contribution equilibrium is risk dominant (in the sense of Harsanyi and Selten, 1988, see Croson et al 2005 and Fatas et al 2006 for a discussion).

The T-WLM is more interesting to analyse. Teams cannot profit by deviating from any symmetric team contribution profile, but individuals within teams in the T-WLM face a social dilemma, as in the I-VCM. Interestingly, even when the theoretical prediction for T-WLM, assuming self-interested, identical and rational individuals in a common knowledge environment, does not differ from the prediction described above for the I-VCM, the opportunity cost of exerting too little effort is very different than the one in the I-VCM. We describe the differences in some detail.

Participants in the T-WLM and the I-VCM have incentives to reduce their contributions from any symmetric contribution profile, but the incentives are not the same. While in the I-VCM, a unilateral one-unit reduction of the individual contribution always generates a payoff increment of half-unit, in the T-WLM the net effect depends on the relative team position within the group, with three possible scenarios.

a) If the team is contributing above the group provision (that is, $C_j > C_{team}^{min}$the minimum), a one-unit reduction pays a full unit back, as any contribution above the group minimum is lost. A one-unit increase generates a full unit loss for the very same reason.

b) If the individual is in a team contributing the group minimum (that is, $C_j = C_{team}^{min}$), reducing the individual contribution in one-unit only generates an individual benefit of one-third: it increases the private benefits in one unit, and it reduces the public benefits (that is, $2 \cdot C_{team}^{min} = 2 \cdot C_j$) in two thirds of a unit, as it reduces C_j in one third ($C_j = \frac{1}{3} \cdot \sum_{i=1}^{i=3} c_i$). Interestingly, the cost of a one-unit increase is also one third, as it reduces the private benefits in one unit, at a benefit of two thirds.

c) In any symmetric contribution profile in which all teams are contributing the same amount, reducing the individual contribution still generates a moderate loss of one-third,[35] and a one-unit increase generates a full unit loss, as it does not increase the group minimum.

To sum up, incentives to adjust downwards may be larger (if in a team above the minimum) or smaller (if in a team contributing the minimum) than in I-VCM. The cost of unilaterally increasing the contribution may also be larger (when in a team above the minimum, or in a symmetric contribution profile), or smaller than in the I-VCM (if in a team contributing the minimum).

As the strategy maximizing social welfare is identical in all three conditions (full efforts), and the within team decision rule allows participants to compensate the low effort exerted by other team members, T-WLM gives a more realistic and interesting setting to explore the effectiveness of teams versus individuals because it reduces the cost of moving towards the social welfare maximizing solution (and it reduces the incentives to adjust down the individual contribution). [36]

The information provided to participants in our experiment did allow for signalling. After each round, participants learned the group provision (determined by an individual or a team, depending on the treatment), their individual earnings, and the average contribution in their team (only in the T-WLM treatment). As the team size was three, and participants were explicitly reminded of their individual decisions, they could compute in a relatively easy way whether someone else in their team was increasing (or decreasing) their contribution. [37]

The experimental procedures were quite similar to the ones used in other team production experiments. The composition of each team and group was randomly determined at the beginning of the experiment and remained stable across the 20 rounds of the experiment. 208 subjects participated in the computerised experiment. In the I-WLM treatment 36 subjects generated nine independent observations (one per group); in the T-WLM treatment, 120 subjects generated ten independent observations (ten groups of 12 subjects, in four teams of three participants each); in the I-VCM 52 participants generated 13 independent observations.

[35] Unless, of course the contribution is zero, and a decrease in the individual contribution is not possible.
[36] Colman et al (2008) document evidence in favour of team reasoning in two experiments in which individual and team reasoning are primed in different coordination games. As Colmin et al (2008) admit, team reasoning typically assumes that individuals pick out the profile that maximizes collective payoff if this profile is unique; if not unique, the theory is indeterminate. The team reasoning theory goes back to Loomes and Sugden (1992), and the evidence provided by Colman et al (2008) is disputed in Sugden (2008).
[37] We discarded the alternative of providing them with full information about each contribution within their team to keep the complexity of the experiment similar across the three conditions.

We used the zTree software (Urs Fischbacher 2007) to implement the computerized experiment. All experiments were run at the University of Valencia (Spain).[38] No subject participated in more than one session and participants had no prior experience in similar games. The sessions lasted 60 minutes on average, and participants earned an average of €14 (US$15).

4. Experimental Results

4.1. Individual effort

We start by comparing contribution and provision (e.g. group production) in the individual and team treatments (I-WLM vs T-WLM vs I-VCM) in Figure 2.1 and Table 2.1. Figure 2.1 represents the average contribution and provision across all treatments, being standard deviations computed using group data in all cases (and represented in parenthesis in Table 2.1).

Figure 2.1: Contribution and provision by treatment

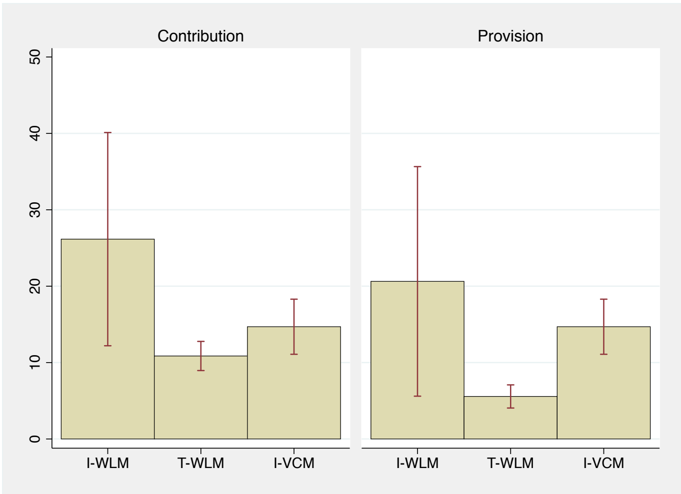


Table 2.1: Descriptive statistics

	Contribution	Provision
I-WLM	26.16	20.63
	(17.36)	(19.55)
T-WLM	10.86	5.57
	(2.54)	(2.12)
I-VCM	14.69	14.69
	(5.79)	(5.97)

Contrary to Feri et al (2010), Figure 2.1 shows that teams do not outperform individuals in this setting. Individuals exert higher efforts than teams in the WLM (26.16 versus 10.86) and in the VCM environment (14.69 vs 10.86). More importantly, provision is significantly lower in the T-WLM than in the other two conditions (20.63 and 14.69 for individuals vs 5.57 for teams); being differences significant at the 1% and 10%, respectively (using to the Wilcoxon-Mann-

[38] Complete instructions (translated from Spanish) are available in the appendix A.2.1.

Whitney test, with only one observation per group). Table 2.2 confirms these findings through different panel data estimations.

Table 2.2: Contribution and Provision Estimations

	Contribution		Provision	
	(1)	(2)	(3)	(4)
	1st round	All rounds	1st round	All rounds
T-WLM	-3.838	-3.830**	-10.60***	-9.127***
	(2.387)	(1.809)	(2.375)	(1.744)
I-WLM	8.987***	11.47*	-9.263**	5.935
	(2.515)	(6.021)	(4.018)	(6.462)
Period	--	-0.564***	--	-0.378***
		(0.0695)		(0.0953)
Constant	21.60***	20.61***	21.60***	18.66***
	(1.929)	(1.880)	(1.985)	(2.064)
I-WLM vs T-WLM	12.82***	15.50***	1.33	15.06**
	(2.14)	(5.86)	(3.73)	(6.29)
Observations	208	4,160	32	640
R-squared	0.112	0.1856	0.319	0.214
Number of subjects	208	208	32	32

Note: the dependent variable in regressions 1 and 2 is the individual contribution, and in regressions 3 and 4 is the public good provision. Standard errors are clustered by group and the regressions control for random effects at individual level. Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. We use the lincom post-estimation test to calculate differences between the I-WLM and the T-WLM coefficients.

The dependent variable in regressions 1 and 2 is the individual contribution to the group account, and in regressions 3 and 4 the group provision. The regressions report one observation per decision-maker (individual or team) and include random effects at the level of the decision maker and clustered standard errors at the group level.

Table 2.2 confirms that teams do not do better than individuals. Whereas in the first round (Model 1) effort (or contribution) in I-WLM is higher than in I-VCM, T-WLM is not significantly above the I-VCM and it is significantly below the I-WLM by a massive margin of 12.82 endowment units. Across the 20 rounds (Model 2), T-WLM does significantly worse than I-VCM and I-WLM, and the differences between I-WLM and I-VCM are only marginally significant, even when the magnitude of the difference (and the coefficient) goes up; we will come to this below.

Differences observed in provision in the first round (Model 3) come from the WLM SCF complex strategic environment. Provision is significantly and substantially lower in the I-WLM and T-WLM than in the I-VCM in the first round; interestingly, no significant differences are observed between I-WLM and T-WLM in the first round. Model 4 suggests that, once all rounds are included, that teams have a negative effect on the provision levels, but the provision of individuals' treatments are not significantly different to each other.

By inspection, Figure 2.1 and Table 2.1 suggested that the I-WLM exhibited a much higher dispersion than in the other two treatments. Table 2.2 shows that the coefficient capturing differences in contributions between I-WLM and I-VCM go up, as well as its standard deviation. Maybe not surprisingly, this larger dispersion could be linked to the very different occurrence of symmetric contribution profiles: while individuals coordinate in one equilibrium 35% of the times in the I-WLM condition, participants never reached a symmetric contribution profile in the T-WLM (more on this below). We explore the heterogeneity of results within each treatment in Figure 2.2 and Table 2.2 below.

Figure 2.2: Contribution versus provision by group and treatment

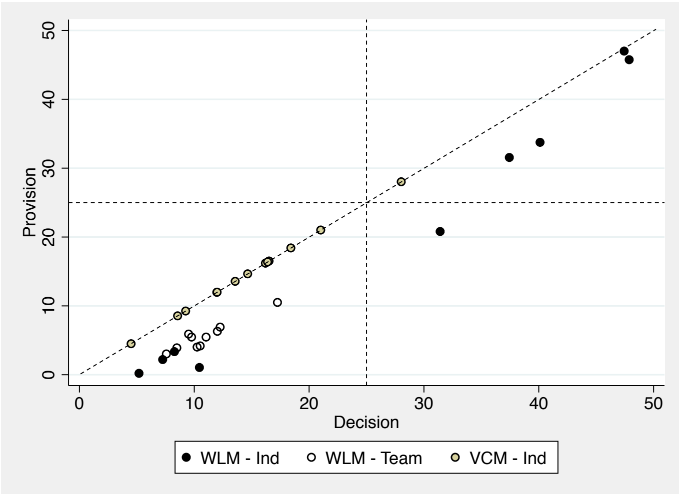


Figure 2.2 illustrates a two-dimension diagram of decisions (group contributions) on the *x* axe, and provisions on the *y* axe. All I-VCM groups are over the 45 degrees line (green circles), as by construction, every contribution to the group account counts and the average contribution and the provision are always identical. Solid black and hollow circles represent the I-WLM and the T-WLM groups, respectively. By inspection, the polarisation in I-WLM groups is apparent. Almost half of them are trapped in the bottom left quarter (with low contributions and low provision levels), while the other half manage to stay in the right and top quarter (high contributions and high provision levels).

Figure 2.2 allows us to identify at least two interesting features of the T-WLM treatment performance. First, the horizontal distance between each point in Figure 2.2 and the 45-degree line represents a rough measure of wasted effort (in other words, how far the average contribution is from the average provision, or how above contributions are from the minimum). Before analysing this pattern more rigorously, Figure 2.2 also reveals that teams in T-WLM do not waste more effort than individuals in I-WLM. Second, the distribution of T-WLM dots overlaps with the low performance groups in I-WLM.

How do teams get trapped in the left-bottom space of low contribution and provision? Figures 2.3 and 2.4 below illustrate the distribution of contributions by decisions makers across all periods. Beyond the massive dispersion of I-WLM (left panel), contributions to the group account in the T-WLM condition are more compact and exhibit a decline similar to the one observed in the I-VCM. Yet, they are slightly further from zero than in the other two treatments.

Figure 2.3: Contribution by decision makers over time

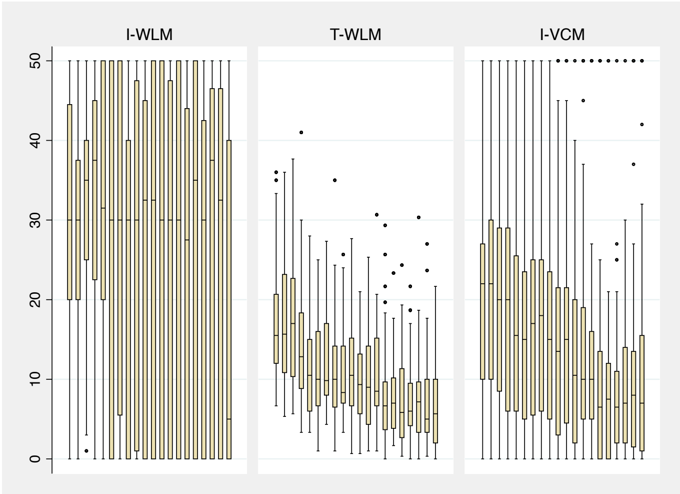


Figure 2.4 below confirms that individual contributions get polarised over time only in the I-WLM, while both I-VCM and T-WLM follow similar patterns. These two figures together show that even when provision in round 1 is not significantly lower in teams than in individuals when the experiment starts, the decline and lack of success in increasing contributions starts immediately after.

Figure 2.4: Distributions of individual contributions over time

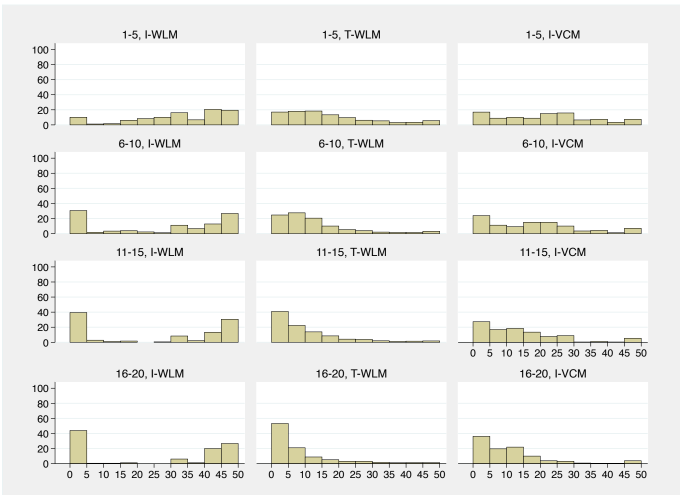


Table 2.3 documents the polarisation in the I-WLM, and how teams in the T-WLM condition are indistinguishable from the low performing teams in the I-WLM treatment:

Table 2.3: Contribution Estimation by Nash Scenario

	(1)
T-WLM	-3.830**
	(1.810)
I-WLM*	
Non-Nash	9.732*
	(5.151)
Zero Nash	1.996
	(4.478)
Non-Zero Nash	21.98***
	(5.141)
Period	-0.556***
	(0.0633)
Constant	20.54***
	(1.861)
Observations	4,160
R-squared	0.2900
Non-Nash versus T-WLM	13.56***
	(4.95)
Zero Nash versus T-WLM	5.83
	(4.28)
Number of subjects	208

Note: the dependent variable in this regression is the individual contribution. The I-WLM contributions were divided per the NE scenario they belong to. Standard errors are clustered by group, and the regressions control for random effects at individual level. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. We use the lincom post-estimation test to calculate differences between the Non-Nash and Zero-Nash I-WLM and the T-WLM coefficients.

In Table 2.3 we disaggregate I-WLM in three categories: not in equilibrium, *zero* equilibrium (all participants in one group contribute zero in that period) and *non-zero* equilibrium (all participants in one group contribute the same positive quantity). Two main results come up from this table: First, that participants in I-WLM contribute significantly more than participants in the T-WLM as long as they are not trapped in the zero-equilibrium; by construction, a symmetric profile of zero contributions is an equilibrium in all three conditions. Second, while I-WLM groups coordinated in *zero* are not significantly different to individuals in the I-VCM and T-WLM conditions, participants in the T-WLM still contribute significantly less than those in the I-VCM, probably suggesting that by seeking coordination, we have creating an underling mechanism pressing down the teams' contributions, other than the individual incentive for free-riding.

In summary, when the experiment starts in round 1, the mechanisms work transparently for which provision is not different in the T-WLM and the I-WLM, and contribution in T-WLM and I-VCM are not significantly different. Polarisation starts immediately after in the I-WLM, and the VCM-like decline drags teams into low levels of contribution and provision. Even out of equilibrium, groups in the I-WLM contribute more than teams in the T-WLM condition. Teams fail to address the challenging environment, and following the graphical in Figure 2.2, we explore now to what extent wasted effort is driven by the incentives to free ride within teams in the T-WLM. We provide further evidence in the next subsection.

4.2. Wasted Effort

We measure *wasted effort* as the absolute deviation of each effort level chosen by individuals from the minimum in her group every period (i.e. provision). Effort is wasted because it comes at an individual cost and does not generate any group surplus. Following Feri et al (2010), we also investigate the *adjustment* dynamics, defined as the absolute distance between the individual (or team) chosen effort in a round and the group minimum in the previous round. Table 2.4 presents the results of three different models, being the dependent variable of models 1 and 2 wasted effort (in round 1 and all rounds, respectively) and the dependent variable of model 3 the adjustment in contributions. The method is identical to the one described above in Table 2.3, and the regressions only include data from the I-WLM and the T-WLM treatments, as no wasted effort is generated in the VCM.

Table 2.4: Wasted Effort and adjustment (I-WLM versus T-WLM)

	Decision Makers		
	Wasted Effort		Adjustment
	(1)	(2)	(3)
	1st round	All rounds	All rounds
T-WLM	-8.045	1.974	1.289
	(6.263)	(1.858)	(1.927)
Period	--	-0.304***	-0.356***
		(0.0668)	(0.0716)
Constant	16.11***	8.458***	9.792***
	(4.543)	(1.520)	(1.605)
Observations	19	380	361
R-squared	0.088	0.0562	0.0582
Number of groups	19	19	19

Note: the dependent variable in regressions 1 and 2 is the difference between the team/individual contribution and the public good provision; also, known as the *wasted effort*. The dependent variable in regression 3 is the difference between the current contribution and the public good provision in the last period namely the *adjustment*. Standard errors are clustered by group, and the regressions control for random effects at individual level. Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 2.4 shows the lack of significant differences between individuals and teams for both wasted effort and adjustment. In other words, teams contribute less than individuals in the WLM SCF, but the magnitude of wasted effort generated and the adjustment process seem to be very similar, confirming the finding in Figure 2.2 i.e. despite the differences in contributions, team are only outperformed by trapped (in a bad equilibrium) individuals.

4.3. Conditional Contribution Adjustment

Croson et al (2005) show that conditional contribution as a behavioural driving force is context specific. While participants react to the average contribution of others in the VCM, they react to the minimum in the WLM, being the average and the minimum contribution imposed by the team production technology. In our T-WLM treatment, the joint production of teams is determined by the minimum contribution made by the four decision-makers (teams), while the contribution of each team is the average contribution of all three team members. The conditional contribution pattern could set the reference point at the group (the minimum contribution made by any team in the group) or the team level (the average contribution in the team).

The characterization of the within group dynamics goes well beyond the purely academic exercise, as it could help us to understand why teams fail to coordinate without forced consensus.

Model 1 in Table 2.5 shows how decisions made by individuals (first two columns) and decision-makers (individuals or teams, depending on the treatment, last column) follow the public good provision in the previous round, using interaction terms (treatments' dummies times lagged provision). The large, positive, and significant coefficients show that conditional cooperation is a strong behavioural force in all three treatments.

As participants in the T-WLM may follow the team or the group performance, the positive coefficient of "T-WLM group-prov [t-1]" could be an expression of conditional contribution towards the group or the team. Model 2 disentangles between these two levels incorporating the (lagged) average contribution within the team. Results clearly indicate that individual actions are mainly driven by the within team dynamics: even though, they do not significantly follow the lagged public good provision made by the group, they are positively affected by their teammates behaviour. Results are confirmed by Model 3 that is a similar regression run with decision makers (again, the coefficient for T-WLM group-prov [t-1] is not significantly different from zero). In other words, by following decisions at the team rather than the group level, individuals are willing to either waste effort (if contributing more than other teams) or to drag down group performance (if contributing less).

Table 2.5: Conditional Contribution

	(1) Individuals	(2) Individuals	(3) Decision Makers
Period	-0.420***	-0.388***	-0.464***
	(0.0508)	(0.0527)	(0.0680)
I-WLM group-prov [t-1]	0.681***	0.706***	0.667***
	(0.0811)	(0.0782)	(0.0892)
T-WLM group-prov [t-1]	0.334***	0.0931	0.206
	(0.0675)	(0.0600)	(0.146)
T-WLM team-cont [t-1]		0.252***	
		(0.0525)	
I-VCM group-prov [t-1]	0.418***	0.455***	0.376***
	(0.0502)	(0.0516)	(0.0564)
Constant	13.58***	12.20***	15.26***
	(1.084)	(1.209)	(1.645)
Observations	3,952	3,952	2,242
Number of subjects	208	208	98

Note: the dependent variable in regressions 1 and 2 is the individual contribution. The dependent variable in regression 3 is the decision maker contribution (individual/team). All the explanatory variables, but the period, are lagged one round. Variables referring to provision are basically the group provision; those referring to *team contribution* consider the decision maker contribution in the previous round, individual or team. Standard errors are clustered by group, and the regressions control for random effects at individual level. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

In Table 2.6 we study positive and negative conditional contribution patterns by analysing the decisions' adjustment – the difference between the current and the previous contributions, $C_t - C_{t-1}$– of those whose contributions deviated from the group provision, or the team contribution (only for the T-WLM). The coefficients should be interpreted as the effect of the explanatory variables on the *contribution adjustment* (contribution change). We find that individuals reduce their contributions, make a negative adjustment, when they are above the group provision in all three treatments (models 1 to 4). However, when contributing below the group, individuals adjust up their effort levels only when making decisions individually (as in I-VCM model 1), and not when they are teamed up (model 4).

The conditional cooperation adjustment is asymmetric in all three experimental conditions (as the different values of the above and below coefficients suggest). While for the I-VCM (model 1) and the I-WLM (model 2) the results are pretty much consistent with previous empirical results on peer effects (Mas and Moreti, 2009; Kato and Shu, 2009, Bandiera et al 2009) i.e. individuals respond by adjusting their decisions towards their peers, the same logic does only partially hold for the teams' condition. The coefficient of upwards adjustments is not significantly different from zero in the T-WLM, neither at teams or group level (models 3 and 4).

From Table 2.5 we know that individuals looked more closely their peers' behaviour rather than the group performance to make decisions on their contributions. However, we find a different pattern for their contribution adjustment. As model 5 in Table 2.6 suggests, when we control for both the team and the group reference points, downwards adjustments strongly follow the group (the provision), and it is only marginally affected by the team (their immediate peers). This probably means that teams eventually learn how their payoffs are actually determined. If any, the team performance would generate a negative effect on the contribution change, reinforcing the tendency for teams get trapped in the bad equilibrium.

Table 2.6: Conditional Contribution Adjustment

	(1) I-VCM	(2) I-WLM	(3) T-WLM	(4) T-WLM	(5) T-WLM
Period	0.0170	-0.351***	-0.0124	-0.0699**	-0.0613**
	(0.0302)	(0.0863)	(0.0231)	(0.0305)	(0.0279)
Distance above group [t-1]	-0.537***	-0.495***		-0.550***	-0.430***
	(0.106)	(0.106)		(0.0346)	(0.0609)
Distance below group [t-1]	0.103**			-0.0611	0.0418
	(0.0439)			(0.0899)	(0.107)
Distance above team [t-1]			-0.815***		-0.219*
			(0.0742)		(0.125)
Distance below team [t-1]			-0.0158		-0.0735*
			(0.0338)		(0.0415)
Constant	1.018*	5.993***	2.360***	3.867***	3.846***
	(0.563)	(1.064)	(0.386)	(0.521)	(0.422)
Observations	988	684	2,280	2,280	2,280
Number of subjects	52	36	120	120	120

Note: the dependent variable in these regressions is the difference between the contribution in the current round and the contribution in the previous one. The explanatory variables represent the absolute distance between the individual contribution and the group provision or the team contribution. The variables are divided in those who contributed above and below the group/team. Standard errors are clustered by group, and the regressions control for random effects at individual level. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

4.4. Efficiency

We measure now earnings in the three treatments as a synthetic indicator of the relationship between institutions, decisions and efficiency. As very different levels of wasted effort are compatible with the same contribution and provision profiles, neither contributions nor provision are enough to characterize efficiency in coordination games. Table 2.7 presents the results of three regressions in which the dependent variable is individual earnings, as defined by the payoff functions used in section 3.

Table 2.7: Earnings Estimations

	(1) 1st round	(2) All rounds	(3) All rounds
Period	--	-0.221***	-0.303***
		(0.0794)	(0.0602)
T-WLM	-17.55***	-14.69***	-14.69***
	(2.441)	(1.725)	(1.725)
I-WLM	-27.71***	0.131	--
	(6.263)	(6.999)	
I-WLM *			
Non-zero Nash	--	--	21.41***
			(5.139)
Zero Nash	--	--	0.136
			(4.663)
Non-Nash	--	--	-7.145
			(5.371)
Constant	71.79***	67.28***	68.15***
	(1.945)	(1.923)	(1.850)
Non-Nash versus T-WLM	--	--	7.549
			(5.15)
Observations	208	4,160	4,160
R-squared	0.281	0.191	0.351
Number of subjects	208	208	208

Note: the dependent variable in these regressions is the individual profits per round. The I-WLM contributions were divided according to the NE scenario they belong to. Standard errors are clustered by group and the regressions control for random effects at individual level. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. We use the lincom post-estimation test to calculate differences between the Non-Nash I-WLM and the T-WLM coefficients.

Our results again depart from the ones obtained by Feri et al (2010) when comparing individuals and teams. Participants in T-WLM earn significantly less than individuals in I-WLM and I-VCM (with the only exception of the first round, in which a Lincom test shows a p-value of 0.108 when comparing earnings in I-WLM and T-WLM). Model 3 reveals that participants in T-WLM teams earn less than individuals in the I-WLM condition only when they reach an equilibrium other than the inefficient zero-equilibrium. While Table 2.3 confirmed that off-equilibrium groups did contribute significantly more than teams in the T-WLM condition, Table 2.7 suggest that they do not earn more, due to the efficiency losses associated with an off-equilibrium position.

5. Conclusions

These results together support the observation that institutions are critical to assure better performed teams. In Feri et al (2010), when teams needed to reach consensus, teams significantly outperformed individuals in effort levels, earnings, and wasted efforts. In our study, when consensus was absent from the process of team decision-making, teams contributed and earned less but did not necessarily miss-coordinate more. Across all rounds, participants in the T-WLM contributed less than in the I-VCM but just as much as I-WLM coordinated in *zero.*

The pivotal and innovative paper of Feri et al (2010) combines two important but previously independent streams of literature: coordination games and team decision-making. They show convincingly that when team members are forced to reach consensus on their effort decisions, teams outperform individuals. However, the reliance on consensus seems an important one. We report results on a similar weakest-link game using both teams and individuals but without requiring consensus on the part of the teams. We find the opposite result as Feri et al (2010): without consensus teams do worse than individuals.

In our setting teams do not necessarily waste more efforts than individuals. They get trapped hanging around the bad equilibrium, probably because of the free riding incentives, their inability to make upward adjustments, or the combination of both. Even though extensive previous research demonstrates that communication helps coordination (e.g. Russell Cooper et al. 1992, Gary Charness 2000, John Duffy and Nick Feltovich 2002, Andreas Blume and Andreas Ortmann), these papers typically examine communication between the coordinating parties (here, the groups of four [or five] teams or individuals) rather than communication within the decision-maker (the team or the individual). In our study, neither teams nor individuals engaged in explicit communication. We replace communication with an aggregation rule i.e. the average of the teammates contributions. While it could be reprehensible that teammates do not get any communication, it is also quite implausible that they would consult other team members in the precise moment where the action is taking place. Any interaction between teammates might come either before or after the decision-making moment.

Another matter is whether the lack of explicit communication affects or not the behaviour in the individual treatments. In strategic settings, a few papers have used a "think-aloud" protocol in order to gain insight into the decision-making process, although they generally do not directly compare decision-making outcomes with and without self-communication (e.g. Uri Gneezy et al. 2007, Ondrej Rydval et al. 2009, Monica Capra 2009). However, two recent papers demonstrate differences in decision-making in games when individual players are required to explicitly communicate with themselves.[39] Bjorn Frank (2009) examines the impact of enforced reflection in different games by asking subjects to write a message to themselves. He finds that self-communication significantly increases the probability of winning in beauty contest games, the amount given in the solidarity game and the proportion of equitable offers in ultimatum games. Scott Rick and Roberto Weber (2010) report evidence on learning in different games, including a coordination (stag-hunt) game. They find that the effect of thinking aloud before

[39] The idea is far from new in cognitive psychology; it is associated with the 'thinking aloud' protocol of K. Anders Ericsson and Herbert Simon (1984). The intuition is that when subjects are forced to 'talk to themselves' about a problem, the solution is easier to find. Michelene Chi et al. (1994) found that participants who communicate aloud with themselves were able to perform better in a variety of individual decision-making settings. For example, Elizabeth Jefferies et al. (2004), show that 'thinking aloud' enhances the capacity of working memory. Alexander Renkl et al. (1998) show that it helps subjects better calculate compound interest, Yair Neumann and Baruch Schwarz (1998) show that it helps subjects to solve analytic reasoning problems.

deciding promotes learning of the strategic nature of games, allowing for a greater transfer of learning across games.

In summary, we believe that the question of teams and individuals in coordination decisions, and in economic and game-theoretic decisions more generally, is still critical and understudied. We hope that the results of this study highlight the importance of studying not only the outcomes of team versus individual decision-making, but also the process by which teams (and individuals) make their decisions. Here we argue that teams outperform individuals per se. They require a proper combination of incentives and institutions to put the coordination challenge on top of the self-interested bias. Further work is required to identify other factors that cause teams to outperform or underperform individuals in a variety of economically relevant settings.

Chapter 4. Leadership and Team Performance: A Comparison between Northern and Southern European Soccer[40]

[40] Co-authored with: Peter Dawson, School of Economics, University of East Anglia, peter.dawson@uea.ac.uk

1. Introduction

"The strength of the group is the strength of the leaders."
– Vince Lombardi

Leadership is, according to Calvert (1992), an institution created to overcome coordination and cooperation problems within groups. Northouse (2001) defines it as a process whereby a group is influenced by an entitled individual in order to achieve a common goal. Even though leaders do not always have the means to enforce or to reward others' demeanour, they have charisma and prestige to encourage individuals make coordinated decisions.

Leaders apparently play a non-negligible role in sport teams' performance. They are appointed or elected by the coach or the team members ((Gould, Hodge, Peterson, and Petlichkoff, 1987; Glenn and Horn; 1993) to guide the team in motivational, task-related, social and external aspects (Loughead et al. 2006; Fransen et al. 2014a; Fransen et al. 2014b). Studies conducted by Cotterill (2013), Dupuis, Bloom and Loughead (2010), Holmes, McNeil, and Adorna (2010) and Apitzsch (2009) have highlighted the relevance of having an effective athlete leadership to encourage team members and to improve the on-field experience.

Besides the players' talent, sport team production is directly affected by the team members' emotions, expectations and their interaction with powerful Figures. Caliendo and Radic (2006) and Mechtel et al (2011) studied the motivational effect of player dismissals. Dawson et al (2000) estimated the effect of the interaction between team members and coaches from the production function perspective. Finally, several sport psychologists have approached the roles of the team leaders (i.e. the captain) and the team members' perceptions (Loughead et al. 2006; Fransen et al. 2014a; Fransen et al. 2014b) to recognise their effects on team performance. Nevertheless, the actual effect of the leaders on the field on team attainments has, to our best of knowledge, not been previously estimated empirically.

This study aims to connect our learning in both sport team production and sport psychology in order to disentangle the indirect effect of the *regular captain* on football team performance. To do so, information of captains' and teams' performance in six European leagues during 2012 – 2015 has been collected. We studied the captain influence in a quasi-natural experiment environment in which the presence or absence of the captain, due to a red card, is assumed as an exogenous event. A combination of econometric methods and the methodological manipulation of a natural experiment allowed us to recognise whether the presence (absence) of the captain enhances (reduces) team performance.

We find that home teams in Southern leagues (Ligue 1, Serie A and La Liga) appear to benefit from dismissed away captains. No such result however is found for Northern leagues (Premier, Bundesliga and Eredivisie). Our results for non-captains suggest that red cards typically reduce team performance (Ridder et al 1994; Bar-Eli et al 2006) and, contrary to Carmichael and Thomas (2005), we find that this negative effect is not necessarily smaller for away teams. Our results are also consistent with Caliendo and Radic (2006) and Mechtel et al (2011), in that controlling for the time to go, red cards do not always have the negative expected effect on the away team performance. When issued to the away team, red cards can increase the team morale and reduce the goal difference.

The remaining of this chapter is structured as follows. The next section describes the relevant literature on the determinants of sport team production and team performance, as well as the role of leadership in sports. A background of Northern and Southern European leagues is also provided in the third section. The fourth section describes the methodology. The fifth section illustrates the econometric approach and the results. Section six concludes.

2. Related Literature

We focus on the literature linked to sport team production function and the effect red card on teams performance, on one hand; and the effect of leadership on teams' performance, on the other.

❖ *Sporting Production Literature*

The seminal contribution of the idea of a sporting production function was introduced by Scully (1974) and followed by several authors that have estimated the production function in different sports such as baseball, basketball, American football, cricket, rugby and football (Zech, 1981; Scott et al., 1979; Atkinson et al., 1988; Schofield, 1988; Carmichael and Thomas, 1995; Carmichael et al. 2000; Carmichael et al 2001; Lee and Berri, 2008).

These studies analyse sport teams as firms or market organisations[41]. As such, they study the marginal and joint effects of a set of the *inputs* (players, talent, abilities, managers, etc.) on the generation of *outputs* (goals, tournaments, points, ranks, etc.). Over time, the production function studies in football have significantly improved in quality after the appearance of Opta Index, which has provided access to match-based data (Carmichael et al 2000; Carmichael et al 2001) such as individual playing talents, and the players' ability to win duels, recover the ball,

[41] Another approach to the production process in sport teams is the *frontier analysis*, which directly studies the production factors' efficiency (Porter and Scully, 1982; Dawson et al 2000). Dawson et al (2000) model and estimate the indirect coaching effect on team performance by including ex-ante measures of the playing talent. By implementing time-variant and time-invariant procedures, the authors found that coaching efficiency is not strongly correlated with team outcomes, but it has more to do with the playing talent. They conclude suggesting that coaching evaluation should be mediated by the outcomes obtained relative to the available playing talent.

attack, defend and finally to score[42]. The interactions of both playing talents and player's actions have been also included in the analysis of the determinants of team performance.

Carmichael et al (2000), for instance, estimated the production function for English Premier League using match-performance as the unit of observation. Their estimation lies on the assumption that the production function is linearly explained by the production factors, which are the play and team characteristics. They utilise the goals difference as outcome variable and found that while players' effectiveness to attack positively affects the goals difference, indiscipline and pressure on the field (red cards and send-offs decisions) have a negative effect.

Carmichael et al (2001) studied the efficiency and inefficiency sources of match plays for the 20 Premier League clubs in the 1997 to 1998 season by computing residual-based models. The authors estimated models for the cumulated points by the end of the league, the goals conceded, the goals scored and ball possession. The residuals of those estimations were ranked and utilised as proxies for the relative efficiency of the 20 Premiership clubs by calculating the difference between the estimated and the true values for each club. The computation of the residuals showed that the ball possession might increase team efficiency as it reduces the chances of the opposition to score and increase the probabilities to attack. However, whereas possession is the basic determinant of goals, it does not determine team output. It is the players' skills and their team-working relationships what determines teams' efficiency and their ability to convert opportunities into goals.

Findings regarding the impact of red cards have been controversial. Ridder et al (1994) analyse matches of the Eredivisie League (Dutch league) from 1989 to 1992 by estimating and comparing the results of both a time-homogeneous Poisson model and an OLS regression. Bar-Eli et al (2006) conducted an analysis of the effect of players' dismissals on team performance in German football clubs between 1963 and 2003. Non-parametric statistics as well as Multinomial logistic regressions were used to estimate regression weights and the odds ratios of red-card related variables and the match outcomes. These two studies found that red cards always have a negative effect on the punished team performance[43]. The later study also suggests that, relative to the away team, home teams might have a natural advantage as their probability of receiving a red card is lower, and therefore are less likely to get trapped in the behavioural downward spiral of crisis.

Carmichael and Thomas (2005), Caliendo and Radic (2006), Merchtel et al (2011) found that the impact of a dismissal is not necessarily the same for home and away teams. Carmichael and

[42] We also rely on betting data as they provide a good proxy for team quality.
[43] the effect of match score, the team to which the red card was issued, and the time interval the card was issued on immediate (i.e. first goal after the issue of the card) and final match outcomes

Thomas (2005) estimated a match-based production function by using cross-sectional data for 380 matches of 20 English Premiership clubs. The authors estimate linear regressions of the goals scored and shots at goal for both home and away teams. While red cards are always detrimental for English teams' performance, away teams are better at accommodating their strategy after a player dismissal. Thus, the effect of a red card is stronger for home than away teams.

Caliendo and Radic (2006) used World Cup data between 1930 and 2002, following Ridder et al (1994). Their econometric strategy consisted of a conditional maximum likelihood estimation in which they obtained the impact of a red card on the punished team by looking at the difference between goals scored before and after a red card. Red cards negatively affect team performance, but if issued after the first half of the match, teams do not necessarily do worse.

Merchtel et al (2011) analyse the German Bundesliga from 1999 to 2009 by implementing both, a goals-based (goal difference) and a result-based (win, draw and lose categories) approach. They found that although red cards issued against home teams dramatically reduces their chances of winning, late dismissals of away teams' players (issued after the 70th minute) might have a positive motivational effect on the away team.

Finally, Lago-Peñas et al (2016) carried out a descriptive analysis of 75 matches of the top-five professional football leagues in the 2015/2016 season[44]. They calculated a set of relative indexes regarding the team performance (ball possession, shots, total passes, long passes, short passes, successful passes, touches and defence) before and after a red card was issued. They found that while the disadvantaged team reduces performance in all dimensions, the advantaged team improves in all areas except for defence.

[44] Premier League (n=9), French Ligue 1 (n=16), Spanish La Liga (n=19), Italian Serie A (n=21) and German Bundesliga (n=10)

❖ *Team Leader Effect*

The studies identified above look at team performance and player dismissal in general. Analyses on the team leader have generally been conducted in psychology studies by implementing survey methods. Sport psychologists have expanded these analyses to show that professional team members are not behaviourally immune leadership on the field.

According to Northouse (2001) leadership is "a process whereby an individual influences a group of individuals to achieve a common goal". Coaches have positively valued this motivational effect, and this is the reason why formal leader (i.e., captains) elections or appointments are carried on (Gould, Hodge, Peterson, and Petlichkoff, 1987; Glenn and Horn; 1993).

Loughead et al (2006) studied how rugby, football, field lacrosse, volleyball, field hockey, ice hockey and basketball team members play a leadership role. Roles were categorised in task (tactics, individual responsibilities, goal focus), social (harmony, communication, support) and external (promotion, needs representation). Team members were asked to list the team leaders fitting within the three categories and they also participated in a two-waves-survey study. The authors found that both captains and regular teammates are sources of leadership, and formal leaders (e.g., captains) were more likely to be identified as team leaders by their teammates.

Fransen et al. (2014a) have pointed out that *motivational* leadership should be acknowledged as a distinct role besides the other three already mentioned by Loughead et al (2006) (task, social and external). They conducted an online survey asking both coaches and players to allocate within each of the four leadership categories the one team player that best fit in each role. 1,258 coaches and 3,193 players within nine different team sports in Flanders (Belgium) participated in the survey (i.e., basketball, volleyball, soccer, handball, netball, hockey, rugby, water polo, and ice hockey). In contrast to Loughead et al (2006), the authors found that 44% of the participants did not perceive their captain as the principal leader on any of the four roles, and they explained this result by claiming that instead of being experts just at one role, leaders should comply with a combination of characteristics that allow them to exercise their leadership.

Fransen et al. (2014b) used the same study described above to understand the effect of leadership on group/team identity. They examined the quality of the four athlete leader categories (i.e., the task, motivational, social, and external leader) and explored the impact of athlete leaders' quality on collective efficacy (team members' shared belief in their conjoint ability to achieve an outcome) and team outcome confidence (group's shared belief of their ability to outperform the opposition and to obtain a goal). The authors utilised a Structural Equation Modelling (SEM) with AMOS to connect the leadership quality and group identity with the both kinds of team identity variables (efficacy perception and confidence). They concluded that both formal and informal leaders play an important role in team performance as they can optimize teams' collective efficacy and team outcome confidence.

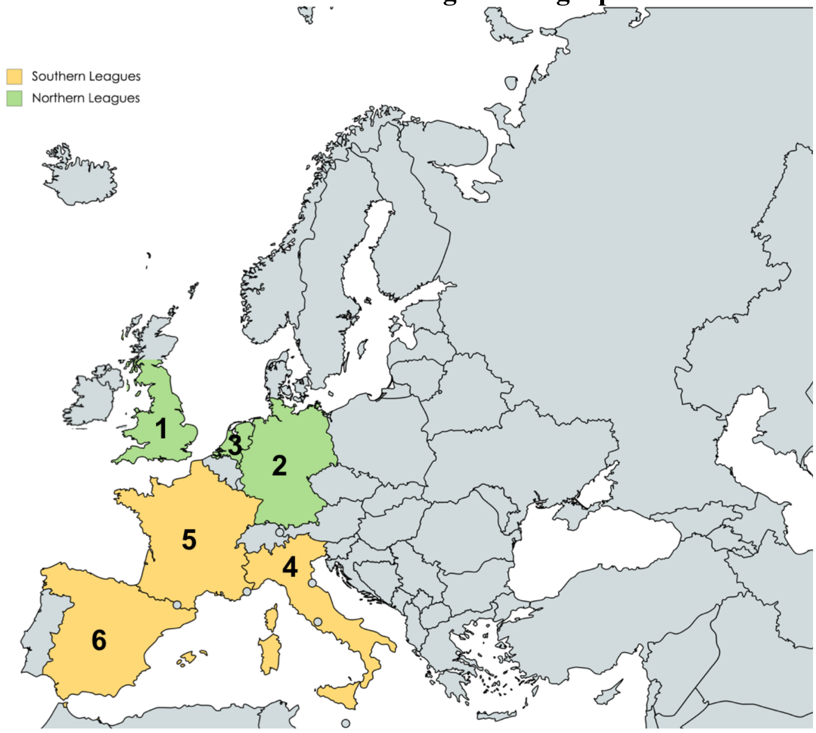
We are aware that formal and informal leaders normally coexist within teams. However, here we study the effect of formal leadership on team performance i.e. the *regular (first) captain*. We are interested in disentangling and estimating the effect of the regular captain on match outcomes. Our main assumption is that besides their own performance, captains play a behavioural role that affects the performance of team partners (Gould, Hodge, Peterson, and Petlichkoff, 1987; Glenn and Horn; 1993). Moreover, even though captains are not entitled to make tactic decisions and they have no voice to recruit or include players in a match, their role is still prominent with respect to the other team members. As a matter of fact, there is always one captain on the field who has been elected or appointed by the coach or the team.

Empirical studies looking at the impact of the captain on team performance on the field are rare. The majority of these studies are survey-based analyses focused on the effect of leadership on team members' perceptions. To complement these, this chapter focuses on the effect of formal leadership on football teams' performance. We use an adaptation of the *sport production function* methodology in an attempt to disentangle the effect of the regular team captain on team performance. Our analysis differs conceptually with respect the team coach leadership in that captains are horizontally positioned regarding other on-the-field players, but are vertically different in terms of status. We depart from studies already carried out by sport psychologist in that we use actual team outcomes rather than individual perceptions.

3. Northern and Southern Leagues Background

We acknowledge that leadership is an attribute that can vary across cultures and regions. Therefore, we included a geographical dimension into ours. We consider six European football leagues, including the top five leagues in the UEFA ranking for club competition, and the Netherlands (currently ranked 13 in the same list). The geographical location of those leagues is illustrated in Figure 3.1.

Figure 3.1 Southern and Northern Leagues Geographical Distribution



i. Northern Leagues (in green): The Premier League (1. England and Whales), Bundesliga (2. Germany) and Eredivisie (3. The Netherlands).
ii. Southern Leagues (in yellow): Serie A (4. Italy), Ligue 1 (5. France) and La Liga (6. Spain).

In the Northern leagues, there is one main club that accounts for approximately half of the trophies since the League foundation (Manchester United, Ajax and Bayern Munich). In Southern leagues, there are typically more than one star clubs. However, by looking at the distribution of the trophies, some heterogeneity in terms of the teams' quality is observed: the Ligue 1 awards are pretty much distributed between 19 clubs; 79% of the Serie A awards have been earned by 3 teams (Juventus, Inter de Milan and Milan); and 64% of the La Liga tournaments were earned by two teams (the Real Madrid and Barcelona) (See Table 3.1 for more details).

European Leagues operate within a meritocracy ("open" leagues) compared to the American sports which operate closed leagues. Relegations are automatic for the three-least-ranked clubs playing in the Premier League, but the promotion is only automatic for the top two clubs in the Football League Championship. Promotion and relegation (P/R) is automatic for two clubs in the Bundesliga (a the 2. Bundesliga) and one club in the Eredivisie (and the Eerste Divisie). In Southern Leagues P/R is automatic for three teams, except for La Liga in which the promotion is automatic just for two teams. In all leagues, non-automatic P/R is determined after a series of play-offs with top clubs competing in the second division.

Football clubs/leagues operate in very different economic and social environment. Northern leagues receive 15% more revenues from broadcasting than the Southern leagues, and the relative attendance to football matches is over 45% higher in the north. Thus, we can state that Northern leagues face larger levels of social "pressure" (See Table 3.1 for more details)[45].

Financially, Northern leagues receive 24% more revenues and pay 12% more to their players than Southern leagues. Moreover, both revenue and wages seem to be more unequal within the Northern leagues (headed by the English Premier League) than within the Southern ones (see Table 3.1).

Northern tournaments have increased in popularity relative to Southern contests for two main reasons: wide broadcasting and football quality. The Premier League is one of the most popular football events in the world, not only because it is *competitive*, but also because it is very well broadcasted around the world. The Bundesliga is one of the most-supported leagues and perceived as a *teamwork* spectacle in which players are well paced and are always ready to exercise more physical pressure on the opposition; or as they call it, '*gegenpressing'*. The Dutch football clubs have developed a style that mixes up the triangular-ball flow to elevate the rate of ball *possession*, with the German *'gegenpressing'* to keep a fluent and permanent attack on the opposition. In general, Northern teams' styles consist of attacking, keeping possession and overwhelming the other team with numbers on the offence. It is not necessarily a stylish strategy, but it is efficient and effective.

Northern and Southern leagues share similar players' formation on the field, but they exhibit different philosophies in teams style. The 4-4-2 formation is the most popular across clubs. Other systems such as the 3-5-2 and the 4-3-3 tend to be more popular in Northern leagues (especially in German and Dutch leagues) as they engage in a more aggressive, effective, fast and organised style.
Southern leagues appear to be more confident with the ball, good dribblers, and innovative. However, there are not many style elements in common between them. Firstly, The Italian clubs are perceived very *cautious* with respect to attacking strategies; thus, the Serie A is one of the lower scoring leagues in Europe. It also presents a declining pattern that has been attributed to the lack of competition after becoming a 20-teams contest (as it used to be 16-teams tournament) and the doping and referee corruption scandals. Secondly, la Liga is one of the main exponents of the *Tiki-taka* game. It is possession oriented and technically impressive.

[45] The effect of the crowd on referees' behaviour has been studied in sports psychology. Nevil et al (2002) found that relative to matches in silence, referees viewing the challenges with background crowd noise were more uncertain in their decision making and tend to issue significantly fewer fouls (15.5%) against the home team. Downward and Jones (2007) studied the effect of the crowd size and found a non-linear and negative correlation between the crowd size and the probability of a yellow card being awarded against the home team. Thus, we cannot rule out that social exposure is also affecting teams' performance.

Nevertheless, the financial inequality between the Spanish clubs have induced to a quality imbalance. Thus, there are only two teams that concentrate the best players and awards i.e. the Real Madrid and the FC Barcelona; situation that has probably weaken the Spanish football audience. Finally, the French league, Ligue 1, lacks a general attribute that describes the league style. It seems, however, that French football clubs are lately emphasising more on physical strength rather than on technical quality.

Table 3.1 Southern and Northern League's Characteristics

Leagues	Premier	Bundesliga	Eredivisie	Serie A	Ligue 1	La Liga
Foundation	1992	1962	1956	1929	1932	1927
Promotion and relegation	3	2 + 1	1 + 2	3	3	3
Clubs	20	18	18	20	20	20
Matches per team	38	34	34	38	38	38
Broadcasting revenue[46]	£1.97b	£0.69b	£0.07b	£0.94b	£0.51b	£0.86b
League revenues	£3.77b	£2.06b	£0.15b	£1.54b	£1.20b	£1.80b
Average player wage a year[47]	£2.3m	£1.5m	£0.2m	£1.3m	£1m	£1.2m
Average attendance[48]	36,461	43,300	19,387	22,162	20,896	28,568
Attendance per 100 thousand inhabitants	68.8	52.7	114.0	36.5	31.3	61.5
Seasons	23	53	59	86	83	88
Competing teams	47	54	53	66	76	60

Notes: The Premier League was created by the Football League First Division to break away from the Football League, which was founded in 1888. The system of promotion and relegation of the Premier League exists between the Premier League and the Football League Championship. The three lowest placed teams in Premier League are relegated to the Football League Championship, and the top two teams from the Football League Championship promoted to Premier League, with an additional club promoted after a series of play-offs involving the third, fourth, fifth and sixth placed clubs.
The Bundesliga was initially called the Oberliga DFV and it was created in 1958. For the promotion and relegation, the bottom two finishers in the Bundesliga are automatically relegated to the 2. Bundesliga, with the top two in the 2. Bundesliga taking their places. The third-bottom club in the Bundesliga will play a two-legged tie with the third-place team from the 2. Bundesliga, with the winner taking up the final place in the following season's Bundesliga
In the Eredivisie, the club at the bottom is automatically relegated to the second level of the Dutch league system, the Eerste Divisie (First Division). At the same time, the champion of the Eerste Divisie will be automatically promoted to the Eredivisie. The next two clubs from the bottom of the Eredivisie go to separate promotion/relegation play-offs. The play-offs are played in two groups. Each group has one Eredivisie club and three high-placed clubs from the Eerste Divisie.

[46] According to the Deloitte UK annual footbal finance report (2016), other sources or revenues are the Matchday tickets, sponsorship and comercial, and other commercial activities. Figures are presented in billions.
[47] The average wage was obtained from https://fourfourthreefootball.wordpress.com/2014/11/14/average-salaries-from-major-world-football-leagues-revealed/. Figures are reported in millions.
[48] This information corresponds to the average attendance in the 2015 season according to the European Football Statistics Web Site.

The relegation in the Italian, French and Spanish leagues work similarly i.e. The three lowest-placed teams are relegated to Serie B, Ligue 2 and the Segunda División, respectively. The promotion works a bit different. Whilst the French league directly promotes the tow three teams, the Spanish and the Italian leagues directly promote the top two and create a play-offs tournament to decide the third promotion.

4. Methodology

We estimate the effect of the *regular captain* on team outcomes of six European Football leagues. We understand that the marginal and collective contribution of players to team performance is not trivial. However, football matches can be exploited as a quasi-natural-field experiment that provides with sufficient information to disentangle the effect of any particular player, including the captain, on team performance.

In an absolute controlled environment, we would evaluate the effect of the absence of the captain relative to the effect of the absence of an equally productive player. Yet, this information requires a sophistication that is beyond the available dataset. Instead, we evaluate the effect on team outcomes of a red card issued to the captain relative to the effect of a red card issued to another player.

Even though one could argue that red cards are the result of a series of endogenous interactions, they are statistically clean events that exogenously change the match conditions, and there are not reasons to make ex-ante (before the match) beliefs regarding players' dismissal. To make this study comparable with other references estimating the red card effect, and to more closely respond to our research question, we estimate the effect of the regular captain dismissal on team relative to other players' dismissals. [49]

We collected data from specialised sources that are in turn based on the Opta database (i.e., Squawka.com and football-data.co.uk). We merged information relative to the match results, red cards, captain absences and the match odds for three seasons (2012-2013, 2013-2014 and 2014-2015) of the six top European leagues that accounts for 144 teams, 6,328 matches i.e. 12,656 individual teams' performances. These leagues are: the Premier League (England), La Liga (Spain), Serie A (Italy), Eredivisie (Netherlands), the Bundesliga (Germany) and Ligue 1 (France).

In our dataset, observations are matches in which two teams A and B played against each other. To identify the red card effect, we control for time-invariant differences between matches that

[49] Injuries and substitutions of the captain, before and during the match, did not report significant effects on team performance. Results are available in Table 3.12 in the Appendix A.3.4

might be constant within seasons and leagues. That is why we implement fixed effect estimations in which the home-team is used as the panel variable. [50]

Prior studies that have analysed the impact of red cards have not assessed the *role* of the player being dismissed (apart from the playing position). We start with a simple model to compare the red card effect among European leagues divided by regions i.e. north and south. We control for team quality, season and the home team.

The following equation describes the functional relationship of the goal difference and the covariates:

$$GD_l = \propto_1 + \propto_2 Hwin_l + \propto_3 HRC_l + \propto_4 ACC_l + \propto_5 HRCT_l + \propto_6 ARCT_l + \propto_7 Season + \varepsilon_l \quad (1)$$

Each match l is considered an observation. GD_l represents the goal difference in match l. Although we focus the analysis on the goal difference as the main dependent variable, similar results are reported in the appendix A.3.3 in which regressions for the number of goals in a bivariate ordered Probit are presented. $Hwin_l$ is a continuous variable that collects the average of the home team winning probability extracted from the odds provided by Ladbrokes, William Hill and Bet Brain. It is used as a proxy for the expected team performance. HRC_l and ARC_l are dummy variables representing the red card event for both home and away teams, respectively. $HRCT_l$ and $ARCT_l$ are continuous variables representing the time to go after the red cards for home and/or away teams, if any, were issued. ε_l is the error term (the estimation results are displayed in Table 3.7).

The second step consists of breaking down the red card effect into captain and non-captain. The regressions keep the same format as in (1). Nevertheless, this time we include new variables relative to the captain performance and his role in the team, which will allow us to estimate a cleaner effect on the goal difference. Equation 2 describes the functional relationship between the dependent and the independent variables:

$$GD_l = \propto_1 + \propto_2 Hwin_l + \propto_3 HCRC_l + \propto_4 ACRC_l + \propto_5 HORC_l + \propto_6 AORC_l + \propto_7 HCRCT_l + \propto_8 ACRCT_l + \propto_9 HORCT_l + \propto_{10} AORCT_l + \propto_{11} HCPM_l + \propto_{12} ACPM_l + \propto_{13} HCPP + \propto_{14} ACPP + \propto_{15} HOPP + \propto_{16} AOPP + \propto_{17} Season + \varepsilon_l \quad (2)$$

$HCRC_l$ and $ACRC_l$ are dummy variables that take the value of *one* if the home or the away captain received a red card, 0 otherwise, respectively. $HORC_l$ and $AORC_l$ follow exactly the same logic, but now it takes the value of *one* if a non-captain player received the red card.

[50] Housman tests comparing random and fixed effects were conducted. The test suggests that the fixed effect estimation suits best for the data and the estimation purposes. We also estimate a bivariate probit model as robustness check. The results are in general consistent and are available in Table 3.11 in the Appendix A.3.3.

$HCRCT_l$, $ACRCT_l$, $HORCT_l$ and $AORCT_l$ represent the time to go after the red cards were issued. The former two are related to the cards issued against the captains, and the latter two are related to non-captain players. $HCPM_l$ and $ACPM_l$ are continuous variables of the time played by the home and away captains. Finally, $HCPP_l$, $ACPP_l$, $HOPP_l$, $AOPP_l$ are vectors of three dummy variables for the home and away captain (as for the former two) and non-captain (as for the latter two) playing position, being goalkeeper the baseline (Table 3.7 for more details).

The third and last step consists of including individuals' attributes that might differentiate Southern and Northern captain selection such as seniority (age) or nationality. [51] We estimated the following model:

$$GD_t = \propto_1 + \propto_2 Hwin_l + \propto_3 HCRC_l + \propto_4 ACRC_l + \propto_5 HORC_{\ l} + \propto_6 AORC_l + \propto_7 HCRCT_l + \propto_8 ACRCT_l + \propto_9 HORCT_l + \propto_{10} AORCT_l + \propto_{11} HCPM_l + \propto_{12} ACPM_l + \propto_{13} HCPP + \propto_{14} ACPP + \propto_{15} HOPP + \propto_{16} AOPP + \propto_{17} HLOCAL + \propto_{18} ALOCAL + \propto_{19} HAGE + \propto_{20} AAGE + \propto_{21} HCRC_l * HLOCAL + \propto_{22} HCRC_l * HAGE + \propto_{23} ACRC_l * ALOCAL + \propto_{24} ACRC_l * AAGE + \propto_{25} Season + \varepsilon_l \qquad (3)$$

$HLOCAL_l$ and $ALOCAL_l$ are dummy variables for home and away teams that take the value of 1 if the nationality of the captain coincides with the league country. $HAGE_l$ and $AAGE_l$ are continuous variables for the home and away captain age.

Table 3.2 presents the variables from the equations with a short description for each of them. In the third and fourth columns, we provide the code used in the estimations and the prior (expected) effect, respectively.

[51] Other variables such us charisma, market value and moral solvency would be interested to include in a follow up of this study

Table 3.2 Variables' Definitions

Variable Name	Definition	Code	Expected
Goal Difference	It is the difference between the home and the away teams' score. It is the main dependent variable.	GD	---
Home team winning probability	Calculated by using the average of the odds reported by Ladbrokes, William Hill and Bet Brain: (h_odd_av-1)/h_odd_av	h_win_prob	+
Home team red card	Dummy variable that takes de value of 1 if, or a given match, a red card is issued to the home team and 0 otherwise.	home_red	-
Away team red card	Dummy variable that takes de value of 1 if, for a given match, a red card is issued to the away team and 0 otherwise.	away_red	+
Remaining time home red card	Time to go after the home team has received a red cart.	home_red_time	+
Remaining time away red card	Time to go after the away team has received a red cart.	away_red_time	-
Home captain red card	This variable takes the value of one if the red card was issued to the home team captain.	home_capt_red	-
Away captain red card	This variable takes the value of one if the red card was issued to the away team captain.	away_capt_red	+
Home non-captain red card	This variable takes the value of one if the red card was issued to a non-captain of the home team.	home_nocapt_red	_
Away non-captain red card	This variable takes the value of one if the red card was issued to a non-captain of the away team.	away_nocapt_red	+
Time to go home captain red card	Time to go after the home team captain has received a red cart.	home_capt_red_time	+
Time to go away captain red card	Time to go after the away team captain has received a red cart.	away_capt_red_time	_
Time to go home non-captain red card	Time to go after the home team non-captain has received a red cart.	home_nocapt_red_time	+
Time to go away non-captain red card	Time to go after the away team non-captain has received a red cart.	away_nocapt_red_time	-
Home captain played time	Minutes played by the regular captain of the home team.	home_capt_time	+
Away captain played time	Minutes played by the regular captain of the away team.	away_capt_time	-
Home captain local	Dummy variable that takes the value of one if the home captain nationality corresponds to the league nationality.	home_local_capt	+/-

Home red-carded captain local	Interaction term of home_local_capt & home_capt_red	home_capt_local_red	+/-
Away captain local	Dummy variable that takes the value of one if the away captain nationality corresponds to the league nationality.	away_local_capt	+/-
Away red-carded captain local	Interaction term of away_local_capt & away_capt_red	away_capt_red_local	+/-
Home captain age	Home captain age at the beginning of a given season	home_capt_age	+/-
Home red-carded captain age	Interaction term of home_capt_age & home_capt_red	home_capt_red_age	+/-
Away captain age	Away captain age at the beginning of a given season	away_capt_age	+/-
Away red-carded captain age	Interaction term of away_capt_age & away_capt_red	away_capt_red_age	+/-
Home captain defender	Dummy variable that takes the value of 1 if the home captain plays a defender position and 0 otherwise	home_capt_defender	+/-
Home captain forward	Dummy variable that takes the value of 1 if the home captain plays a forward position and 0 otherwise	home_capt_forward	+/-
Home captain midfield	Dummy variable that takes the value of 1 if the home captain plays a midfield position and 0 otherwise	home_capt_midfield	+/-
Away captain defender	Dummy variable that takes the value of 1 if the away captain plays a defender position and 0 otherwise	away_capt_defender	+/-
Away captain forward	Dummy variable that takes the value of 1 if the away captain plays a forward position and 0 otherwise	away_capt_forward	+/-
Away captain midfield	Dummy variable that takes the value of 1 if the away captain plays a midfield position and 0 otherwise	away_capt_midfield	+/-

5. Results

5.1. Descriptive statistics

There are 88 teams that participated in all three seasons' championships; 30 were present in two seasons, and 26 appeared in a single season. As presented in Table 3.3, home teams score significantly more goals than away teams (1.45 vs. 1.13 goals), are less likely to receive a red card (9.8% vs 13.1%) and their expected probability of winning is significantly higher (45.8% vs. 29.1%). These general statistics support the *home advantage* hypothesis that has been broadly studied in sport economics (Boudreaux et al 2015; Lago-Peñas et al 2016). [52] Additionally, the proportion of red cards issued to captains, and their played time with and without expulsion, is not significantly different between home and away teams.

Table 3.3 Descriptive Statistics of home and away teams

	Home	Away	t-test (p-value)
Goals	1.451802 (0.0133254)	1.130057 (0.0127121)	0.0000
Winning probability	0.4581226 (0. 4982826)	0.2910872 (0 .4542995)	0.0000
Red cards	0.0979772 0.3148639	0.1314791 0.3627619	0.0000
Captain red cards	0.0063211 (0.07926)	0.0075853 (0.0867697)	0.3922
Non-captain red cards	0 .0916561 (0.3040321)	0.1238938 (0.3539984)	0.0000
Captain played time (minutes)	57.72685 (42.45859)	56.59246 (40.99365)	0.1360
Captain red card time (minutes)	64.925 (20.96932)	61.5 (25.23254)	0.4959
Non-captain red card time (minutes)	65.54159 (22.98855)	65.42319 (22.35102)	0.9283

Note: we represent means and standard deviation in parenthesis.

As Table 3.4 indicates, 80% of matches did not involve red cards. Complementary, 1,367 red cards were issued across the three seasons and the six leagues, 88 of those were delivered to a captain (6.4%). In general, away teams account for 57% of the dismissals and the distribution of red cards issued to home and away teams is significantly different according to the Fischer Exact Test (p-value = 0.000). This result goes in line with Thomas et al (2006) and Glamser (1990) who found that the *home team crowd* might affect officials' judgement in favour to the home team, which in turn leads to aggressive acts from the away team.

[52] Lago-peñas et al (2016) found that home teams scored first 57.8% of the matches; and those who scored first ended the games scoring an average 1.88 goals more than their opponents

Table 3.4 Home vs. Away Red Cards[53]

		Away red cards				
		0	**1**	**2**	**3**	*Total*
Home red cards	**0**	5,083	624	34	1	*5,742*
	1	439	102	9	2	*552*
	2	25	8	0	1	*34*
	Total	*5,547*	*734*	*43*	*4*	*6,328*

Table 3.5 displays a comparison of the captains' individual characteristics between Northern and Southern leagues. The average team captain in Southern leagues is significantly older than in Northern leagues (31.5 vs. 29.0 years old). The captain playing position is also significantly different between regions (Fischer exact test = 0.000): 37.2% (53.6%) of the Southern (Northern) captains played at the defender position; 45.6% (36.1%) are midfielders, 8.3% (6.0%) are goalkeepers and 8.9% (4.2%) are forwards. Southern leagues exhibit a significantly larger proportion of captains that were originally born in the league country, 82%, than Northern leagues, 61%.

In general, we find significant differences between the attributes of Northern and Southern team captains, for the playing position, the age and the proportion of locals that were appointed as leaders. The age and the playing position differences can be attributed to the playing style that seems to be more physical in Northern leagues. The local proportion is consistent to the distribution of national and international players in the teams, as according to The CIES Football Observatory in Switzerland (2014), Northern leagues exhibit a larger proportion of international players than Southern leagues. This proportion was for the Premier League 42.5% (the highest in Europe), followed by the Bundesliga with 35.6%. The Dutch League ranked 9th with 15.7%. In Southern leagues the figures are: 30.1% for The Serie A, 22.3% for Ligue 1 and 20.7% for La Liga.

Captains played full time games around 57% of the time and did not play at all in 28.4% of the matches (3,575 times, either for pre-match substitutions[54], 20% of the cases, or injury reasons, 8.4%). They were substituted 14.3% of the time (1,807 times), received a red card 7% (88 times) and were injured during the match in 3% of the games (39 times). Regarding their performance, the Squawka index suggests that captains and the average player do not differ between regions. Nevertheless, captains outperform the average player, and this captain

[53] A detailed table of all red cards is provided in Table 3.13 at Appendix A.3.5
[54] Strategic substitutions are those that do not respond to an injury but are related to tactical reasons.

premium is significant for both Southern and Northern leagues (p-value<0.01), yet not different between regions (p-value = 0.3050) (Table 3.5). [55]

Table 3.5 Captains and Non-Captain Characteristics

	Southern	Northern	p-values
Captain Performance per minute	0.1984982	.2223963	0.3164
	(0.2073548)	(0.219926)	
Non-captain performance per minute	0.1675356	0.1701271	0.8329
	(0.1103398)	(0.1095284)	
Captain premium	0.0300107 ***	0.0522692 ***	0.3050
	(0.1836322)	(0.2055409)	
Captain age	31.47778	29.0241	0.0000
	(3.506507)	(3.465766)	
Captain BMI	23.015	22.6988	0.2884
	(2.008376)	(3.39811)	
Captain local	82.22%	61.44%	0.0000
Defender	37.22%	53.61%	0.0000
Midfielder	45.56%	36.14%	
Forward	8.89%	4.22%	
Goalkeeper	8.33%	6.02%	

Note: The *captain premium* is the difference between the captain and the average player performance. We tested the difference between these two with the Wilcoxon Match Pair Signed test and we find that the captain performance is significantly higher than the average player at p-value<0.01, for which we placed three stars *** in the premium variables. We compared Southern and Northern leagues by using the t-test for continues variables, and the Fischer Exact test for proportions (local captains and playing position).

Home and away captains do not exhibit differences in the time played within Southern and Northern leagues. Northern and Southern captain time does not have major differences except for the case of the non-red-carded-away captains. In the absence of red cards, away captains tend to play longer in Northern teams relative to Southern (Table 3.6). Red cards, on the other hand, tend to be delivered to captains roughly at the same time between regions (approximately 1.15 red cards per season, after the minute 60), and team types (home or away).

[55] Squawka index details are provided in the appendix A.3.1. The average player index is obtained by dividing over ten players, the difference between team performance and the captain performance.

Table 3.6 Red Card Events and Time Played by Captains

	Southern		Northern	
	Home	Away	Home	Away
Captain red card events	25	29	15	19
Non-captain red card events	348	441	169	249
No red card time	77.27969	77.94384[a]	83.06151	81.56925[a]
	(15.44847)	(15.42137)	(9.270077)	(10.67747)
Red card time	62.25	66.13636	61.29412	65
	(27.29079)	(22.88126)	(20.80494)	(17.70795)

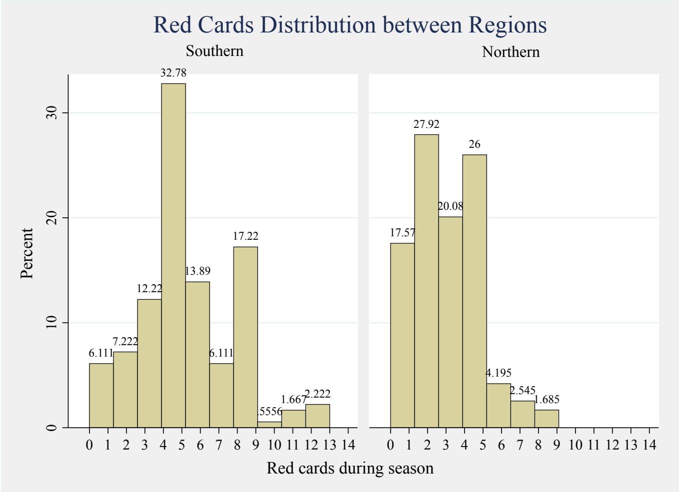
Note: We used the Fischer exact test to test differences in the red card proportions between Northern and Southern leagues, whilst teams were playing home or away. Nonetheless, we find no significant differences between regions. We tested the difference in the number of minutes played by the captains in Southern and Northern leagues by using the t-test. We placed a superscript at the values that reported significant differences. We only found significant differences at 1% between Southern and Northern leagues (a-superscript) in Non-red carded away teams (a-p-value = 0.000).

On average, a football team receives between four and five red cards each season. As Figure 3.2 indicates, the distribution of red cards is significantly different between Northern and Southern teams (Kolmogorov-Smirnov test p-value = 0.000). Specifically, Southern captains received significantly more red cards across the three seasons (54) compared to Northern captains (34) (t-test p-value = 0.0046). [56]

The difference in the red cards' distribution probably responds to differences in the football style and/or to the leagues' social exposure. As stated above, Southern leagues are more ball-possession oriented and often exhibit defensive kind of games. This strategy, combined with cultural preconceptions, might lead players to engage in more aggressive actions. Moreover, as Northern leagues are more socially exposed in terms of broadcasting and match attendance, it is possible that they feel higher levels of pressure to behave appropriately on the field.

[56] Leagues are all significantly different to each other, except for the Premier League and Bundesliga which t-test p-value is 0.2940

Figure 3.2 Histogram of the Red Card per Match by Region



Note: The histogram illustrates the percentage of red card events during one season. On the x axis we have the average number of red cards received by one team of Southern/Northern leagues across the three seasons; and on the y axe we have the percentage of occurrence. The probability of receiving less than four red cards during a specific season is 2.6 times larger for Northern leagues. Complementary, the probability of receiving more than five red cards is four times larger in Southern leagues than in Northern leagues.

5.2. The Relative Effect of a Red Card

The first three columns of Table 3.7 represent the estimation results of Equation (1). We present the coefficients for the full model in the first column, and the Southern and Northern leagues in the second and third columns, correspondingly. Robust standard errors are in parentheses. As expected, we find that the home team quality positively (h_win_prob) affects the goal difference in favour to the home team and this coefficient is significant in all regressions. Red cards issued to the away (away_red) team increase the goal difference in around 0.9 goals, and expulsions against the home team (home_red) reduces it in around one goal.

We find that the time to go after a red card time (home_red_time) has no effects when it is issued to the home team, but it increases the goal difference when it is given to the away team (away_red_time) in Northern leagues, meaning that the longer the remaining time after a red card, home teams are more able to take advantage of the away team weakness. We break down the red card effect into captain (hom_capt_red and away_capt_red) and non-captain (hom_nocapt_red and away_nocapt_red) dismissals in the other models. We include several explanatory variables related to the captain participation to capture his relevance for teams' attainments.

The coefficients corresponding to Equation (2) for the full, Southern and Northern estimations are presented in columns 4, 5 and 6 of the Table 3.7. When it comes to non-captains, our results are consistent with Ridder et al (1994) and Bar-Eli et al (2006). We find that for all cases, red cards issued against home (away) teams reduces (increases) the goal difference of the home team, and these coefficients are not significantly different between regions. [57] For home and away teams, the red card time (home_nocapt_red_time and away_nocapt_red_time) has a small but significant effect on goal differences. Specifically, the earlier a non-captain of the home (away) team gets penalised, the lower (higher) the home advantage in each region and in both regions pooled. Also, we partly support the relative red card time effect in away teams' performance found by Caliendo and Radic (2006) and Mechtel et al (2011) namely that later red cards on the away teams might have a boost on their morale which eventually has a positive effect on team performance.

The playing position for home team captains is insignificant for the full (model 4) and regional specifications (models 5 and 6). Nonetheless, forward away captains (away_capt_forward) may significantly increase the goal difference in Northern leagues, and midfield away captains (away_capt_midfield) may have the same (marginally) effect in Southern leagues[58]. This finding is probably a reflection of the game style of Northern leagues. We pointed out above that Northern leagues exhibit an attacking and organised style. When in Northern matches the leadership of the away team is appointed in the attacking role, it probably weakens the defence area and rises the opportunities for increasing the goal difference.

As Table 3.7 suggests, the effect of home captain dismissals (home_capt_red) on goal differences is similar across regions. However, Southern teams seem to be better at exploiting the lack of leadership of the away teams. More specifically, away captain dismissals (away_capt_red) increase in almost two the goal difference in favour to the home team. This effect is 36% greater, yet not significantly different, than the non-captain's. Interestingly, the red card time accounts for goal differences only for away captains' dismissal at Southern leagues (away_capt_red_time), probably meaning that the motivational effect already suggested by Mechtel et al (2011) is reinforced when the captain is the one being sent off. [59]
Southern home teams take advantage of away team the leader disadvantage, and La Liga seems to be driving a large part of this effect (as presented in Appendix A.3.2). Few questions appear to be relevant to understand this phenomenon. Why away-captain dismissals seem strengthen Southern home teams and are insignificant for Northern teams? Why Southern teams are better

---

[57] Even though, the results here presented control for the home team fix effects, we also estimated OLS regressions of the same model with the sole purpose of conducting the Suest test for equality between coefficients of independent regressions.
[58] We controlled for the playing positions of non-captain dismissals in models 4-9 and appeared insignificant.
[59] We also estimated the model in Equation No 2 for each league separately. We found that the coefficients obtained in the Southern leagues for the away captains are mainly driven by the results of *La Liga,* which is the Spanish professional football league. The results are presented in Table 3.8 in the appendix A.3.2. Additionally, two tables with descriptive statistics and p-values per league are provided in Tables 3.9 and 3.10, respectively in appendix A.3.2.

at utilising the lack of leadership within their *home advantage* than Northern teams*?* To answer these questions, we estimate an additional regression including the captain attributes and their interaction with the red card event. Specifically, we include captain characteristics that, according to the descriptive analysis, accounted for differences between Northern and Southern leagues such as nationality (home_capt_red_local and away_capt_red_local) and age (home_capt_red_age and away_capt_red_age) (Equation 3).

The results of Equation (3) are presented in the last three columns of Table 3.7. The regressions suggest that the positive effect of the away captain dismissal on home score advantage is not only significantly higher than other player dismissal effect, but it is also a robust finding as it continue to hold even after controlling for captains' attributes. However, captain characteristics such as the age and nationality seem not having a substantial role at explaining the goal difference[60].

The lack of effect on Northern leagues could be also influenced by their football style. As they are acknowledged for being more organised, team-synchronised and physical, the lack of leadership seems not to be an issue for ensuring results. Probably meaning that anyone in the field is capable of assuming the job of the dismissed leader (or the team is so well coordinated that there is no room for salient leadership). [61]

[60] We estimated models 7 – 9 including the interaction of captain red cards and their playing position and we did not get different results to the ones here reported. Instead, few variables were dropped due to the number of red carded captains (88).
[61] We estimate three additional regressions that include captains' injuries and substitutions. These types of absences appear to have no impact on the goal difference (see Table 3.12 in appendix A.3.4).

Table 3.7 Linear Regressions for Goal Differences, Standard errors in parentheses

	Full GD (1)	Southern GD (2)	Northern GD (3)	Full GD (4)	Southern GD (5)	Northern GD (6)	Full GD (7)	Southern GD (8)	Northern GD (9)
h_win_prob	3.756***	4.059***	3.443***	3.705***	3.963***	3.382***	3.697***	3.962***	3.357***
	(0.151)	(0.202)	(0.229)	(0.152)	(0.210)	(0.231)	(0.152)	(0.210)	(0.231)
home_red	-1.074***	-1.157***	-0.850*						
	(0.249)	(0.289)	(0.470)						
away_red	0.897***	0.880***	0.924**						
	(0.234)	(0.276)	(0.422)						
home_red_time	-0.00180	-0.00371	0.00241						
	(0.00256)	(0.00301)	(0.00472)						
away_red_time	0.00523**	0.00362	0.00864**						
	(0.00222)	(0.00261)	(0.00406)						
home_capt_red				-0.690	-0.783	-0.696	1.255	-1.506	5.939
				(0.931)	(1.109)	(1.705)	(2.201)	(3.110)	(5.676)
away_capt_red				1.495**	1.974***	0.551	1.319	6.251**	-6.578
				(0.609)	(0.708)	(1.156)	(2.124)	(3.024)	(4.758)
home_nocapt_red				-1.670***	-1.413***	-2.444***	-1.659***	-1.430***	-2.230***
				(0.337)	(0.374)	(0.726)	(0.338)	(0.374)	(0.738)
away_nocapt_red				1.480***	1.264***	1.819***	1.479***	1.264***	1.821***
				(0.328)	(0.395)	(0.575)	(0.328)	(0.395)	(0.575)
home_capt_red_time				0.00226	0.00171	0.000625	-0.00126	0.00533	0.00179
				(0.0131)	(0.0152)	(0.0251)	(0.0136)	(0.0174)	(0.0267)
away_capt_red_time				0.0140	0.0204**	0.000624	0.0118	0.0249**	0.0187
				(0.00904)	(0.0104)	(0.0176)	(0.00916)	(0.0113)	(0.0221)
home_nocapt_red_time				-0.0107***	-0.0111***	-0.0109*	-0.0105***	-0.0108***	-0.00982*
				(0.00318)	(0.00382)	(0.00569)	(0.00318)	(0.00382)	(0.00572)
away_nocapt_red_time				0.0156***	0.0131***	0.0199***	0.0156***	0.0130***	0.0199***
				(0.00277)	(0.00339)	(0.00474)	(0.00278)	(0.00339)	(0.00474)
home_capt_time	0.000841	0.00125*	0.000323	0.000879*	0.00135**	0.000348	0.000881*	0.00130**	0.000373
	(0.000527)	(0.000646)	(0.000894)	(0.000528)	(0.000649)	(0.000894)	(0.000528)	(0.000649)	(0.000893)
away_capt_time	-0.000333	-0.000784	0.000443	-0.000170	-0.000686	0.000692	-0.000184	-0.000685	0.000700
	(0.000507)	(0.000638)	(0.000831)	(0.000508)	(0.000638)	(0.000838)	(0.000508)	(0.000638)	(0.000837)
home_capt_defender				-0.260	-0.108	-0.234	-0.251	-0.0905	-0.227
				(0.206)	(0.260)	(0.377)	(0.206)	(0.260)	(0.377)
home_capt_forward				-0.230	-0.339	0.0142	-0.228	-0.329	-0.0138
				(0.223)	(0.254)	(0.451)	(0.223)	(0.254)	(0.450)

home_capt_midfield				-0.262	-0.0712	-0.274	-0.255	-0.0528	-0.279
				(0.214)	(0.280)	(0.374)	(0.214)	(0.280)	(0.373)
away_capt_defender				0.140*	0.134	0.0895	0.142*	0.139	0.0872
				(0.0829)	(0.103)	(0.146)	(0.0829)	(0.103)	(0.146)
away_capt_forward				0.290***	0.174	0.500**	0.292***	0.177	0.518**
				(0.111)	(0.132)	(0.203)	(0.111)	(0.132)	(0.203)
away_capt_midfield				0.150*	0.181*	0.0495	0.148*	0.180*	0.0429
				(0.0829)	(0.102)	(0.148)	(0.0829)	(0.102)	(0.147)
home_capt_local							0.0976	0.127	0.0947
							(0.0943)	(0.117)	(0.173)
away_cap_local							0.0285	-0.0172	0.0395
							(0.0499)	(0.0722)	(0.0717)
home_capt_age							-0.00886	-0.0274*	0.0400
							(0.0127)	(0.0152)	(0.0255)
away_capt_age							-0.00200	-0.000412	-0.00504
							(0.00654)	(0.00825)	(0.0108)
home_capt_red_local							0.578	-0.242	1.403
							(0.524)	(0.684)	(0.938)
away_capt_red_local							0.972*	1.074	0.620
							(0.508)	(0.672)	(0.867)
home_capt_red_age							-0.0835	0.0349	-0.257
							(0.0725)	(0.107)	(0.164)
away_capt_red_age							-0.0197	-0.152*	0.276
							(0.0645)	(0.0818)	(0.187)
Constant	-1.762***	-1.552***	-2.371***	-3.254**	-3.736**	-2.149	-2.758*	-4.503**	-3.957
	(0.312)	(0.364)	(0.583)	(1.483)	(1.718)	(2.868)	(1.528)	(1.930)	(3.238)
Observations	5,977	3,379	2,598	5,977	3,379	2,598	5,977	3,379	2,598
R-squared	0.126	0.142	0.113	0.133	0.149	0.124	0.134	0.151	0.127
Number of id_home	141	75	66	141	75	66	141	75	66
Season	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Home team FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Non-captain position	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Test hcrc = horc	No signif.	No signif.	No signif.	No signif.	No signif.	No signif.	No signif.	***	No signif.
Test acrc = aorc	No signif.	No signif.	No signif.	No signif.	No signif.	No signif.	No signif.	***	**

Note: The dependent variable of these models is the goal difference between home and away teams. We estimated fixed effect regressions to isolate time-invariant characteristics across matches. We also control for the season and the dismissed non-captain playing positions. We keep the home team as the panel variable. *** p<0.01, ** p<0.05, * p<0.1

6. CONCLUSIONS

We investigate the captain effect on team performance based on the relative difference of a captain send-off with respect to a regular player. To the best of our knowledge, this is the first study that focuses the analysis on dismissals issued to formal leaders.

We find significant differences between geographical regions. Our results suggest that the effect of a captain dismissal on team performance is context dependant. Whilst in Northern leagues captains seem to play no significant roles, in Southern leagues they can affect the goal difference. However, this result requires another layer of specification in that only red cards issued to away captains of a Southern team positively affects the goal difference. Yet, red cards to home captains have no significant effects. Overall, the effect of an away captain dismissal probably suggests that the lack of leadership inside the away team could be an unexplored component in the literature of the *home-team advantage.*

The lack of the captain effect on Northern leagues can be interpreted as a result of their football style. As they are acknowledged for their teamwork, it is possible that the captaincy is only a formal figure that does not play a substantial role in terms of team coordination. Thus, captain dismissals do not threaten team performance in a different way than a non-captain's.

Our results for non-captains are consistent to Ridder et al (1994) and Bar-Eli et al (2006), but contrast to Carmichael and Thomas (2005) as we find that red cards always deteriorate team performance and this effect is not necessarily smaller for away teams. Our results are also consistent to the work of Caliendo and Radic (2006) and Mechtel et al (2011) in that we found that by controlling for the time to go, red cards on the away team may generate a positive impact on the team morale that could lead to goal difference reductions.

Our main drawback is that our database does not allow for comparisons with dismissals of players that are equally productive than the captain. However, we are able to disentangle the effect of a captain dismissal relative to a non-captain dismissal. New studies investigating how captains are appointed and the leadership effect in other sports, might contribute to the study of leadership for team performance.

We have no formal information accounting for differences in the way captains are appointed in Southern and Northern leagues. Selection rules could partly explain why Southern leagues are better at exploiting the lack of leadership of the away teams, or why away teams in those leagues are more sensitive to captain dismissals. We suggest this as the starting point for future work in this area.

Chapter 5. General Concluding Remarks

Three essays broadly relating to the political economy of organisations were presented in this thesis. We utilised organisations re-created in the lab as well as football teams to understand the behavioural implications of the interaction between social constructs, such us power, leadership and peer effects, with traditional economic concepts such us value, individual incentives and team production.

The lack of control over the elements intervening on real-life interactions, as well as the implausibility of creating a systematised database to disentangle the intrinsic value of power, led us to approach power asymmetries in the lab in the first chapter. We differentiate the value of decision rights (Fehr et al 2014, Neri and Rommenswinkel, 2016) from the value of power as we remove any strategic incentive for holding the decision rights. Rather, we create an environment in which participants experience power by making decisions about others, but not about themselves.

We conducted a 2x2 experimental design in which we manipulated the mechanism to assign individuals into powerful positions (high or low skilled individuals), and the awareness regarding how the assignation is done (legitimacy). We also conducted two further treatments manipulating the visibility of powerful individuals (centrality) and their ability to significantly reduce the group members' piece rates (enforcement). A control with no power asymmetries was also implemented.

In the absence of economic incentives, we would have expected individuals to give up power for any positive value. However, we find that the experience of power make individuals letting go substantial economic earnings in order to keep their powerful position. The value of power does not depend upon the legitimacy, centrality or enforcement ability of the powerful, but seems to rely exclusively on the experience of power.

This finding may change the way we approach political interactions, and political changes within organisations (teams, firms, countries, groups of countries, etc.). By acknowledging that the experience of power might change individual preferences by transforming power in an "intrinsically valuable" experience, organisations should create the political institutions that would help them to fairly share and distribute power, and to prevent power abuse and power attachment.

Further research on the value of power could also manipulate the amount of power that one individual has, and his/her hierarchical position within an organisation. Additionally, second order beliefs[62] in further papers could provide us with information regarding the underlying social norms that make individuals demand a significant compensation for giving up the power, and that at the same time it prevents the powerless to paying for it.

Another topic that calls the attention whilst managing the production within organisations, is how to create the proper incentives among individuals to maximise an organisation's outcomes. While some activities are better carried out individually, there are other jobs that generate better outcomes when performed by teams. Comparisons between individual and team performance at coordination tasks have been studied both in the lab (Feri et al 2010 and Chaudhuri et al 2015) and in the field (Hamilton et al 2003). Whilst Feri et al (2010) and Hamilton et al (2003) documented that teams are more productive than individuals, Chaudhuri et al (2015) carried out a modification of Feri's et al matching protocol, and found otherwise. We implemented a different protocol to Chauduri and Feri, yet we also aimed to understand to what extent teams may or may not exhibit certain heuristics that would lead them to outperform individuals in a weakest link public good game.

In contrast to Feri et al (2010), and consistent to Chaudhuri et al (2015), we find that teams do not outperform individuals. Although, both teams and individuals experience similar coordination levels, the majority of teams get trapped in low levels of contributions to avoid high wasted effort. We find that in the absence of coordinating institutions (such as voting, unanimity, communication, etc.) between team members, teams might not reach high contribution levels. Moreover, since participants playing in teams do not consider the history of the total public good produced by the group of teams whilst deciding on their contribution, but only make downward adjustments with respect their teammates' contribution, the low contribution cycle is reinforced after each round.

The second chapter findings point out the challenge for organisations to create an appropriate political environment to aggregate the team members' efforts, and to avoid free-riding incentives. Additionally, our results support the conditional cooperation and reciprocity evidence (Croson et al 2005). Specifically, our results suggest that even though the peer effect literature suggest that peers might generate social pressure in favour of team production (Hamilton et al 2003; Mas and Moretti, 2009; Bandiera et al 2009), when individual incentives are not necessarily aligned to the teams', it might be important to avoid negative messages as they can evolve in a downside peer effect that could harm team performance.

[62] Through first order beliefs we might elicit positive and normative individual's beliefs regarding the others' behaviour. By using second order beliefs we would evoke positive and normative individual's beliefs on how she thinks the others belief the average behaviour would/should be. Second order beliefs allow researchers to understand to what extent individuals are trying to conform with a social norm.

A follow up of this study could compare how different institutions affect individual and team coordination. For instance, the inclusion of leaders, voting rules, communication, punishment, among other possible manipulations, may affect teams and individuals differently. Further work on this topic might provide us with a better characterisation of the relative performance of teams and individuals that are facing coordination challenges.

Finally, we studied the role of leadership on team performance in the third chapter. Like power in the first chapter, a minority of individuals become leaders in organisations. Nonetheless, unlike power, we study leaders as formal figures that do not have the ability to reward or punish other individuals, but have the status to behaviourally influence the team performance. Although, this topic could have also been addressed experimentally, there is public data available to explore it, which give us the opportunity to provide the literature with some insights regarding real-life interactions.

We constructed a database for 144 real organisations namely teams belonging to the six major European leagues. Geographically, three of the leagues operate within Southern Europe, and three within Northern Europe: La Liga, Serie A and Ligue 1 belong to the former; The Premier League, The Bundesliga and The Eredivisie belong to the latter group. The information of team performance was obtained from the available data of football matches for each of the seasons between 2012 and 2015 (6,328 matches).

We assess the effect of captain red card relative to a non-captain dismissal. Although players' dismissal can be understood as an endogenous event, resultant of several technical and emotional interactions on-the-field, the impact of a captain dismissal can be studied as an exogenous event in the sense that we do not know prior to the match that there will be a red card offence. Thus, within each match, we characterised the red card events among captain and non-captain dismissals, and this information was merged with the regular captain characteristics. We acknowledge that the captain is not necessarily the only leader on-the-field, but since they are formally appointed and their status is different, we expect them to have a behavioural effect on their teammates' performance.

We used a *Sport Production Function* framework (e.g. Ridder et al 1994; Dawson et al 2000; Bar-Eli et al 2006, Carmichael and Thomas 2005; Caliendo and Radic 2006; Mechtel et al 2011) to disentangle the captain impact on team performance. Even though the red card effect has been studied by using similar frameworks, from the best of our knowledge, this is the first time that it is used to assess whether the impact of red cards differ between leaders and non-leaders.

We find that the effect of leadership on team performance is context dependent. A captain dismissal is only significant for Southern leagues (mainly driven by La Liga). Moreover, it is only significant when the away captain is the recipient of the red card. Thus, it seems that the combination of the *home advantage* with cultural and technical heuristics make Southern teams playing in away from home round more vulnerable to the lack of leadership. Yet, individual characteristics appear not being affecting such vulnerability.

These findings bring several insights regarding the role of formal leadership in organisations. In general, when teams are technically and tactically trained, as in football teams, formal leaders tend to have a secondary effect on team performance. Additionally, it is possible that well established interactions among teammates, favours the emergence of other leaders that have not necessarily been formally appointed. Leaders can be, however, relevant when the teams are facing a disadvantaged position. In this study, the disadvantage relies on the fact of playing as a guest team, but we can extend this situation to non-sport related teams starting an innovation project, realising a new product, or moving into a new market.

The team (playing) style also might determine how indispensable the leader is. Well-paced teams are less leader dependent (as the Northern leagues); yet, teams that innovate on-the go (as the Southern leagues), might require a front-runner to elevate the teammates morale when the environment turns against them.

We concede that further research on this topic, analysing how captains are appointed, and the effect of regular captains in other sports will enrich the discussion, and generate conclusions that can be extended to different types of organisations.

References

Anderson, C., & Galinsky, A. D. (2006). Power, optimism, and risk-taking. European journal of social psychology, 36(4), 511-536.

Apitzsch, E. (2009). A case study of a collapsing handball team. In S. Jern & J. Näslund (Eds.), Dynamics within and outside the lab (pp. 35–52). Linköping: LiU-Tryck.

Audas, R., Dobson, S., & Goddard, J. (2002). The impact of managerial change on team performance in professional sports. Journal of Economics and Business, 54(6), 633-650.

Ball, S B., & Eckel C. (1996). "Buying Status: Experimental Evidence on Status in Negotiation." Psychology and Marketing, 13(4): 381-405.

Ball, S B., Eckel, C. Grossman, P., and Zame W. (2001). "Status in Markets." Quarterly Journal of Economics, 116(1): 161-88.

Bar-Eli, M., Tenenbaum, G., & Geister, S. (2006). Consequences of players' dismissal in professional soccer: A crisis-related analysis of group-size effects. Journal of Sports Sciences, 24(10), 1083-1094.

Bardsley, N. (2005). Experimental economics and the artificiality of alteration. *Journal of Economic Methodology*, *12*(2), 239-251.

Bardsley, N. (2008). Dictator game giving: altruism or artefact? *Experimental Economics*, *11*(2), 122-133.

Bartling, B., Fehr, E., & Herz, H. (2014). The intrinsic value of decision rights. *University of Zurich, Department of Economics Working Paper*, (120).

Bendahan, S., Zehnder, C., Pralong, F. P., & Antonakis, J. (2015). Leader corruption depends on power and testosterone. The Leadership Quarterly, 26(2), 101-122.

Besson, R., Poli R., & Ravenel, L (2014). *Demographic Study 2014 in Europe.* Edition CIES.

Blume, A., & Ortmann A. (2007). "The Effects of Costless Pre-Play Communication: Experimental Evidence from Games with Pareto-Ranked Equilibria." *Journal of Economic Theory*, 132: 274-290.

Bobadilla-Suarez, S., Sunstein, C. R., & Sharot, T. (2016). Are Choosers Losers? The Propensity to Under-Delegate in the Face of Potential Gains and Losses. *The Propensity to Under-Delegate in the Face of Potential Gains and Losses (February 15, 2016)*.

Boudreaux, C. J., Sanders, S. D., & Walia, B. (2015). A natural experiment to determine the crowd effect upon home court advantage. *Journal of Sports Economics*, 1527002515595842.

Brookshire, David S., & Coursey, Don L. (1987). "Measuring the Value of a Public Good: An Empirical Comparison of Elicitation

Becker, G. M., DeGroot M., & Marschak J. (1964). “Measuring Utility by a Single- Response Sequential Method.” *Behavioural Science*, 9(3): 226–32.Procedures." *A.E.R.* 77 pp: 554-66.

Caliendo, M. &Radic, D. (2006) Ten do it better, do they? an empirical analysis of an old football myth. *Discussion Paper 2158*. Institute for the Study of Labour, Bonn.

Calvert, R. (1992). Leadership and its basis in problems of social coordination. International Political Science Review, 13(1), 7-24.

Capra, M. (2009). “Understanding decision process in P-beauty contest game: A protocol analysis.” Unpublished.

Carmichael, F., & Thomas, D. (1995). Production and efficiency in team sports: An investigation of rugby league football. *Applied Economics*, *27*, 859-869.

Carmichael, F., Thomas, D., & Ward, R. (2000). Team performance: The case of English Premiership Football. *Managerial and Decision Economics*, *21*, 31-45.

Carmichael, F., Thomas, D., & Ward, R. (2001). Production and efficiency in association football. *Journal of sports Economics*, *2*(3), 228-243.

Carmichael, F., & Thomas, D. (2005). Home-field effect and team performance evidence from English premiership football. *Journal of Sports Economics*, *6*(3), 264-281.

Carpenter, J., Kariv, S., & Schotter, A. (2012). Network architecture, cooperation and punishment in public good experiments. *Review of Economic Design*, *16*(2-3), pp 93-118.

Charness, G. (2000). “Self-Serving Cheap Talk: A Test of Aumann’s Conjecture.” *Games and Economic Behaviour,* 33: 177-194.

Charness, G., Cobo-Reyes, R., Jiménez, N., Lacomba, J. A., & Lagos, F. (2012). The hidden advantage of delegation: Pareto improvements in a gift exchange game. *The American Economic Review*, *102*(5), 2358-2379.

Charness, G., & Gneezy, U. (2010). Portfolio choice and risk attitudes: An experiment. *Economic Inquiry*, *48*(1), 133-146.

Charness, G., & Jackson, M. (2007). “Group Play in Games and the Role of Consent in Network Formation.” *Journal of Economic Theory*, 136: 417-445.

Charness, G., & Jackson M. (2008). “The Role of Responsibility in Strategic Risk-Taking.” *Journal of Economic Behaviour and Organization*, 69(3): 241-247.

Charness, G, & Sutter, M. (2012). "Groups make better self-interested decisions." *The Journal of Economic Perspectives* 26.3: 157-176.

Chaudhuri, A., Paichayontvijit, T., & So,. T. (2015). "Team versus individual behaviour in the minimum effort coordination game." *Journal of Economic Psychology* 47: 85-102.

Chi, M., T.H., De Leeuw, N., Chiu, M., & Lavancher, C. (1994). “Eliciting self-explanations improves understanding.” *Cognitive Science*, 18: 439–477.

Coats, J. C., & Rankin, F. W. (2015). The Delegation of Decision Rights: An Experimental Investigation. *Available at SSRN 2443919*.

Cooper, R., V. DeJong, D., Forsythe, R., & Ross, T. (1992). "Communication in Coordination Games." *Quarterly Journal of Economics,* 107: 739-771.

Cotterill, S. (2013). Team psychology in sports: Theory and practice. Abingdon: Routledge.

Coursey, Don L.; Hovis, John L.; & Schulze, William D. (1987)."The Disparity between Willingness to Accept and Willingness to Pay Measures of Value." Q.J.E. 102 pp 679-90.

Dahl, R. A. (1957). The concept of power. *Behavioural science*, *2*(3), 201-215.

Downward, P., & Jones, M. (2007). Effects of crowd size on referee decisions: Analysis of the FA Cup. *Journal of sports sciences*, *25*(14), 1541-1545.

Dawson, P. M. (2012). Experience, social pressure and performance: the case of soccer officials. *Applied Economics Letters*, *19*(9), 883-886.

Dawson, P., Dobson, S., & Gerrard, B. (2000). Stochastic frontier and the temporal structure of managerial efficiency in English soccer. *Journal of Sports Economics*, *1*, 341-362.

Dawson, P., & Dobson, S. (2010). The influence of social pressure and nationality on individual decisions: Evidence from the behaviour of referees. *Journal of Economic Psychology*, *31*(2), 181-191.

de Dios Tena, J., & Forrest, D. (2007). Within-season dismissal of football coaches: Statistical analysis of causes and consequences. European Journal of Operational Research, 181(1), 362-373.

Deloitte (2016). *Annual Review of Football Finance 2016.* Available at: https://www2.deloitte.com/uk/en/pages/sports-business-group/articles/annual-review-of-football-finance.html. Consulted on the 6th January 2017.

Duffy, J., & Feltovich, N., (2002). "Do Actions Speak Louder than Words? An Experimental Comparison of Observation and Cheap Talk." *Games and Economic Behaviour,* 39: 1-27.

Dupuis, M., Bloom, G. A., & Loughead, T. M. (2006). Team captains' perceptions of athlete leadership. Journal of Sport Behaviour, 29(1), 60–78.

Emerson, R. M. (1962). Power-dependence relations. *American sociological review*, 31-41.

Falk, A., & Kosfeld, M. (2006). The hidden costs of control. *The American economic review*, *96*(5), 1611-1630.

Ericsson, K. Simon, A., & Simon, H. (1984). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.

Falk, A., & Ichino, A. (2006). "Clean evidence on peer effects." *Journal of Labour Economics* 24.1: 39-57.

Fatas, E., Meléndez-Jiménez, M. A., & Solaz, H. (2010). An experimental analysis of team production in networks. *Experimental Economics*, *13*(4), pp 399-411.

Fehr, E., Herz, H., & Wilkening, T. (2013). The lure of authority: Motivation and incentive effects of power. The American Economic Review, 103(4), 1325-1359.

Feri, F., Irlenbusch, B.., & Sutter, M. (2010). Efficiency gains from team-based coordination–Large-scale experimental evidence. *American Economic Review, 100: 1892-1912*

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental economics*, *10*(2), 171-178.

Flores, R., Forrest, D., & Tena, J. D. D. (2012). Decision taking under pressure: Evidence on football manager dismissals in Argentina and their consequences. European Journal of Operational Research, 222(3), 653-662.

Frank, B. (2009). "The effects of enforced reflection in three simple experiments." Unpublished.

Fransen, K., Vanbeselaere, N., De Cuyper, B., Vande Broek, G., & Boen, F. (2014a). The myth of the team captain as principal leader: extending the athlete leadership classification within sport teams. *Journal of sports sciences*, *32*(14), 1389-1397.

Fransen, K., Coffee, P., Vanbeselaere, N., Slater, M., De Cuyper, B., & Boen, F. (2014b). The impact of athlete leaders on team members' team outcome confidence: A test of mediation by team identification and collective efficacy. *The Sport Psychologist*, *28*(4), 347-360.

Galinsky, A. D., Magee, J. C., Gruenfeld, D. H., Whitson, J. A., & Liljenquist, K. A. (2008). Power reduces the press of the situation: implications for creativity, conformity, and dissonance. Journal of personality and social psychology, 95(6), 1450.

Glamser, F. D. (1990). Contest location, player misconduct, and race: A case from English soccer. *Journal of Sport Behaviour*, *13*(1), 41.

Glenn, S. D., & Horn, T. S. (1993). Psychological and personal predictors of leadership behaviour in female soccer athletes. *Journal of Applied Sport Psychology*, *5*(1), 17-34.

Gneezy, U., Rustichini, A., & Vostroknutov, A. (2007). "I'll cross that bridge when I come to it: Backward induction as a cognitive process." Unpublished.

Gould, D., Hodge, K., Peterson, K., & Petlichkoff, L. (1987). Psychological foundations of coaching: Similarities and differences among intercollegiate wrestling coaches. *The Sport Psychologist*, *1*(4), 293-308.

Guinote, A. (2007). Behaviour variability and the situated focus theory of power. European review of social psychology, 18(1), 256-295.

Guryan, J., Kroft, K.., & Notowidigdo, M. J. (2009). Peer effects in the workplace: Evidence from random groupings in professional golf tournaments. *American Economic Journal: Applied Economics*, *1*(4), 34-68.

Hamilton, B. H., Nickerson, J. A., & Owan, H. (2003). Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation. *Journal of political Economy*, *111*(3), 465-497.

Hanemann, W. Michael. (1991) "Willingness to Pay and Willingness to Accept: How Much Can They Differ?" *American Economic Review*, June 1991, 81(3), pp. 635-47.

Harrison, G. W., & Hirshleifer, J. (1989). An experimental evaluation of weakest link/best shot models of public goods. *Journal of Political Economy*, *97*(1), 201-225.

Harsanyi, J., & Selten, R. (1988). *A General Theory of Equilibrium Selection in Games*. Cambridge, MA: MIT Press.

Holmes, R. M., McNeil, M., & Adorna, P. (2010). Student athletes' perceptions of formal and informal team leaders. Journal of Sport Behaviour, 33(4), 442–465.

Isoni, A., Loomes G., & Sugden R. (2009). "The Willingness to Pay–Willingness to Accept Gap, the "Endowment Effect", Subject Misconceptions, and Experimental Procedures for Eliciting Valuations: A Reassessment". Discussion paper 09-14. Centre for Behavioural and Experimental Social Science, CBESS. University of East Anglia.

Jefferies, E., Ralph, M. A. L., & Baddeley, A. D. (2004). Automatic and controlled processing in sentence recall: The role of long-term and working memory. *Journal of Memory and Language*, *51*(4), 623-643.

Kahneman, D., Knetsch J., & Thaler, R. (1990). "Experimental Tests of the Endowment Effect and the Coase Theorem." *Journal of Political Economy* 98(6) pp 1325–1348.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the econometric society*, 263-291.

Kato, T., & Shu, P. (2008). Performance spillovers and social network in the workplace: Evidence from rural and urban weavers in a Chinese textile firm.

Kocher, M., & Sutter, M. (2005). "The decision maker matters: Individual versus group behaviour in experimental beauty-contest games." *The Economic Journal* 115.500: 200-223.

Kugler, T., Bornstein, G., Kocher, M. G., & Sutter, M. (2007). Trust between individuals and groups: Groups are less trusting than individuals but just as trustworthy. *Journal of Economic psychology*, *28*(6), 646-657.

Lago-Peñas, C., Gómez-Ruano, M. Á., Owen, A. L., & Sampaio, J. (2016). The effects of a player dismissal on competitive technical match performance. *International Journal of Performance Analysis in Sport*, *16*(3), 792-800.

Lago-Peñas, C., Gómez-Ruano, M., Megías-Navarro, D., & Pollard, R. (2016). Home advantage in football: Examining the effect of scoring first on match outcome in the five major European leagues. *International Journal of Performance Analysis in Sport*, *16*(2), 411-421.

Lee, Y. H., & Berri, D. (2008). a re-examination of production functions and efficiency estimates for the national basketball association. *Scottish Journal of Political Economy*, *55*(1), 51-66.

Loughead, T. M., Hardy, J., & Eys, M. A. (2006). The nature of athlete leadership. *Journal of Sport Behaviour*, *29*(2), 142.

Magee, J. C., Galinsky, A. D., & Gruenfeld, D. H. (2007). Power, propensity to negotiate, and moving first in competitive interactions. Personality and Social Psychology Bulletin, 33(2), 200-212.

Mas, A., & Moretti, E. (2009). "Peers at work." *The American Economic Review* 99.1: 112-145.

Mechtel, M., Bäker, A., Brändle, T., & Vetter, K. (2010). Red cards: Not such bad news for penalized away teams. *Journal of Sports Economics*, 1527002510388478.

Neri, C., & Rommeswinkel, H. (2014). Freedom, Power and Interference: An Experiment on Decision Rights. *Available at SSRN 2485107*.

Neumann, Y., & Schwarz, B. (1998). "Is self-explanation while solving problems helpful? The case of analogical problem-solving." *British Journal of Educational Psychology*, 68: 15–24.

Nevill, A. M., Balmer, N. J., & Williams, A. M. (2002). The influence of crowd noise and experience upon refereeing decisions in football. *Psychology of Sport and Exercise*, *3*(4), 261-272.

Northouse, P. G.(2001). Leadership Theory and Practice. *Organization Studies-Berlin-European Group for Organizational Studies*, *20*(2).

Owens, D., Grossman, Z., & Fackler, R. (2014). The control premium: A preference for payoff autonomy. *American Economic Journal: Microeconomics*, *6*(4), 138-161.

Porter, P. K., & Scully, G. W. (1982). Measuring managerial efficiency: the case of baseball. *Southern Economic Journal*, 642-650.

Randall, A., & Stoll, J. R. (1980) "Consumer's Surplus in Commodity Space." *American Economic Review*, 71(3), pp. 449-57.

Renkl, A,. Stark, R., Gruber, H., & Mandl, H. (1998). "Learning from worked-out examples: The effects of example vocabulary and elicited self-explanations." *Contemporary Educational Psychology*, 23: 90–108.

Rick, S., & Weber, R. A. (2010). Meaningful learning and transfer of learning in games played repeatedly without feedback. *Games and Economic Behaviour*, *68*(2), 716-730.

Ridder, G., Cramer, J. S., & Hopstaken, P. (1994). Down to ten: Estimating the effect of a red card in soccer. *Journal of the American Statistical Association*, *89*(427), 1124-1127.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology*, *25*(1), 54-67.

Rydval, O., Ortmann, A., & Ostatnicky, M. (2009). Three very simple games and what it takes to solve them. *Journal of Economic Behaviour & Organization*, *72*(1), 589-601.

Scully, G. W. (1974). Pay and performance in Major League Baseball. *American Economic Review*, *64*, 915-930.

Shrogren, J. F., Shin, S. Y., Hayes, D. J., & Kliebenstein, J. B. (1994). Resolving differences in willingness-to-pay and willingness to accept. *American Economic Review*, 84(1), pp 255–270.

Sloof, R., & Von Siemens, F. (2014). Illusion of Control and the Pursuit of Authority.

Sturm, R. E., & Antonakis, J. (2015). Interpersonal power: A review, critique, and research agenda. Journal of Management, 41(1), 136-163.

Thaler, R. (1980). "Toward a Positive Theory of Consumer Choice." *Journal of Economic Behaviour and Organization*, 1(1) pp 39–60.

Thomas, S., Reeves, C., & Smith, A. (2006). English soccer teams' aggressive behaviour when playing away from home. *Perceptual and motor skills*, *102*(2), 317-320.

Van Huyck, J. B., Battalio, R. C., & Beil, R. O. (1990). Tacit coordination games, strategic uncertainty, and coordination failure. *The American Economic Review*, *80*(1), 234-248.

Weber, R. (2006). "Managing growth to achieve efficient coordination in large groups." *American Economic Review*, 96(1): 114-126.

Zech, C. E. (1981). An empirical estimation of a production function: the case of major league baseball. *The American Economist*, *25*(2), 19-23.

Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics*, *13*(1), 75-98.

Websites

The European Football Statistics. "Match Attendance". Available at: http://www.european-football-statistics.co.uk/attn.htm. Consulted on the 6th of January 2017

http://www.sponsorreport.nl/sport/totale-sponsorinkomsten-betaald-voetbal--20372-miljoen/

https://fourfourthreefootball.wordpress.com/2014/11/14/average-salaries-from-major-world-football-leagues-revealed/

Foros about Football styles:

https://www.quora.com/What-is-the-difference-between-Spanish-British-German-Italian-and-French-styles-of-club-football#

https://www.quora.com/What-is-the-difference-between-the-Dutch-and-the-German-style-of-soccer

http://www.soccer-training-guide.com/soccer-formations.html#.WHOzabaLTeQ

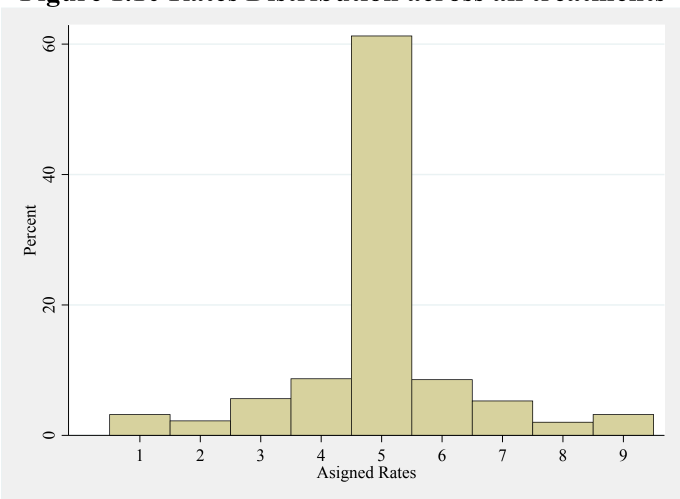
Appendix

A.1.1. Complementary figures and tables chapter 1

Table 1.9 The WTP and WTA for the A position and the Gap by treatment (control)[63]

		Mean	(SD)
NP	**A**	103.3	(96.7)
	B	103.3	(83.6)
	Gap	0	(116.6)
BL - High	**A**	415	(117.3)
	B	135	(113.1)
	Gap	280	(149.3)
EN - High	**A**	565.5	(532.6)
	B	130.8	(112.4)
	Gap	431.7	(484.2)
CT - High	**A**	422.9	(182)
	B	157.9	(117.2)
	Gap	265	(266.2)

Figure 1.10 Rates Distribution across all treatments



[63] *Only significant differences are reported: p-value WMW test between treatments within roles*
A – NP vs. EN= 0.0013; A – NP vs. CT = 0.0006; B – NP vs. CT= 0.0827;
p-value WMW test between roles within treatments, A vs. B
EN = 0.0022; CT = 0.0024
p-value WMW test of the gap between treatments
NP vs. EN = 0.0016

Figure 1.11 Average assigned rates by productivity quintile per treatment

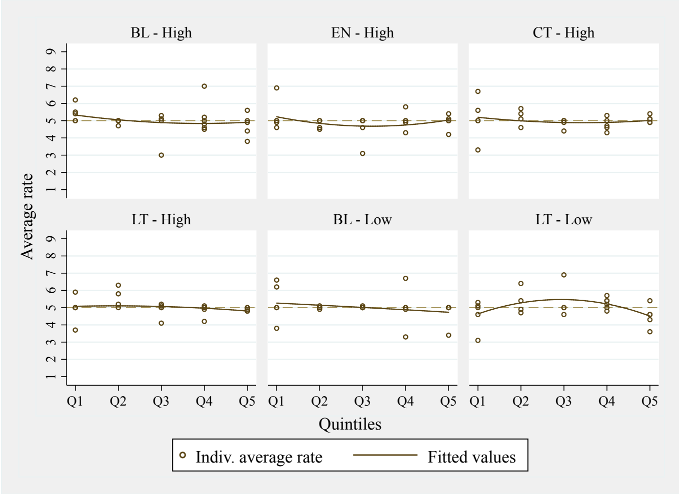


Table 1.10 Participation of the components in the actual value of power by treatment (controls) and role

	BL		EN		CT	
	A	**B**	**A**	**B**	**A**	**B**
Noise (mean)	82.50	104.78	82.50	103.06	82.22	103.06
% actual	19.9	77.6	14.7	78.8	19.4	157.9
Extrinsic (mean)	30.2	3.7	7.9	8.8	-6.0	1.0
% actual	7.3	2.8	1.4	6.7	-1.4	0.65
Risk (mean)	39.58	28.00	33.42	21.25	36.58	25.29
% actual	10	21	6	16	9	16

Figure 1.12 Ruling category by treatment (controls)

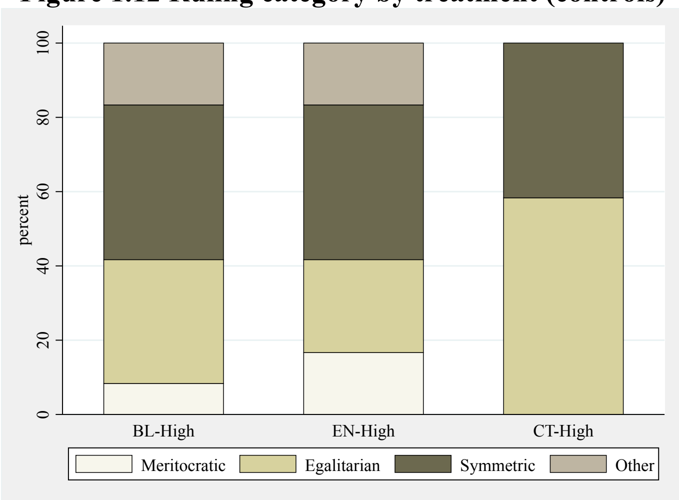
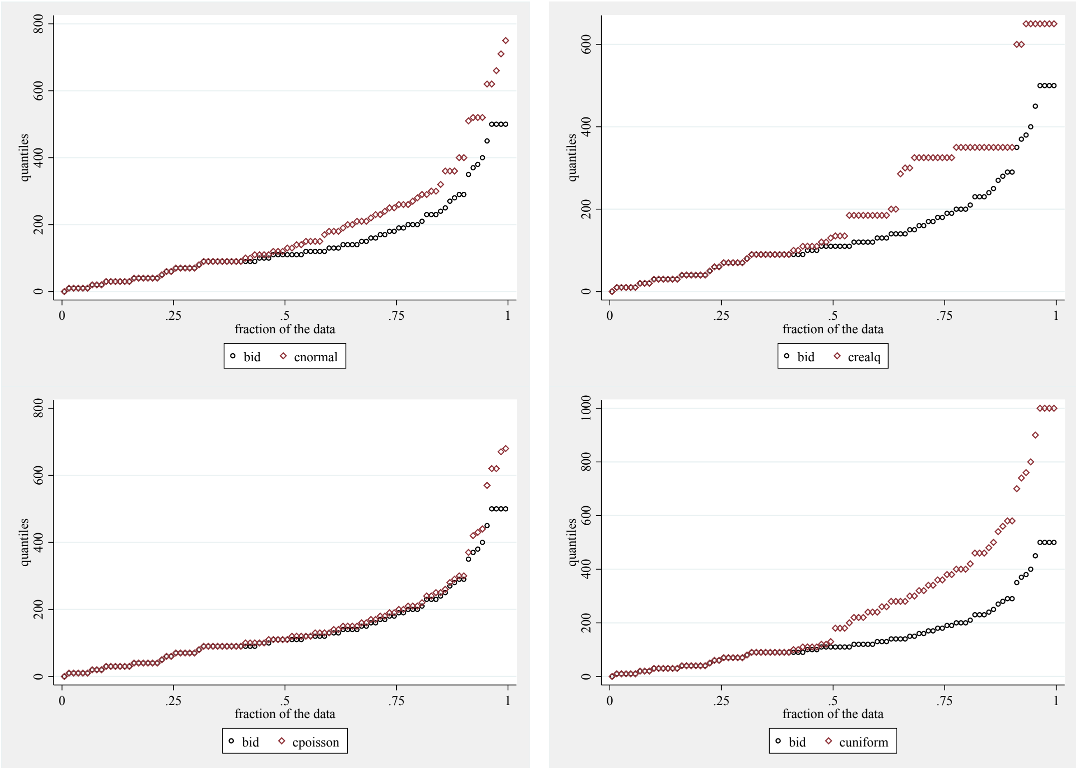


Table 1.11 Individual ruling categories by categorisation method

Treatment	indiv	Scoring rule	frequency rule	Final	Gender	WTA
BL - High	116	Egalitarian	Egalitarian	Egalitarian	Female	450
BL - High	117	Symmetric	Symmetric	Symmetric	Male	500
BL - High	151	Symmetric	Symmetric	Symmetric	Male	300
BL - High	152	Symmetric	Symmetric	Symmetric	Male	350
BL - High	153	Symmetric	Symmetric	Symmetric	Male	450
BL - High	1112	Egalitarian	Symmetric	Other	Male	490
BL - High	1113	Egalitarian	Egalitarian	Egalitarian	Male	600
BL - High	1116	Symmetric	Symmetric	Symmetric	Female	150
BL - High	1118	Egalitarian	Symmetric	Other	Male	500
BL - High	1510	Egalitarian	Egalitarian	Egalitarian	Male	450
BL - High	1511	Egalitarian	Egalitarian	Egalitarian	Male	340
BL - High	1513	Meritocratic	Meritocratic	Meritocratic	Male	400
LT - High	443	Symmetric	Symmetric	Symmetric	Female	500
LT - High	445	Symmetric	Symmetric	Symmetric	Female	501
LT - High	449	Symmetric	Symmetric	Symmetric	Male	600
LT - High	481	Egalitarian	Symmetric	Other	Male	300
LT - High	484	Meritocratic	Meritocratic	Meritocratic	Male	500
LT - High	4411	Symmetric	Symmetric	Symmetric	Male	375
LT - High	4412	Symmetric	Symmetric	Symmetric	Male	300
LT - High	4414	Symmetric	Symmetric	Symmetric	Female	350
LT - High	4810	Egalitarian	Egalitarian	Egalitarian	Female	450
LT - High	4813	Egalitarian	Egalitarian	Egalitarian	Female	500
LT - High	4814	Meritocratic	Meritocratic	Meritocratic	Male	70
LT - High	4816	Symmetric	Symmetric	Symmetric	Male	400
BL - Low	592	Symmetric	Symmetric	Symmetric	Male	500
BL - Low	594	Symmetric	Symmetric	Symmetric	Female	300
BL - Low	597	Symmetric	Symmetric	Symmetric	Female	300
BL - Low	598	Symmetric	Meritocratic	Other	Male	50
BL - Low	5102	Symmetric	Symmetric	Symmetric	Male	299
BL - Low	5105	Symmetric	Symmetric	Symmetric	Male	500
BL - Low	5107	Symmetric	Symmetric	Symmetric	Female	250
BL - Low	5911	Symmetric	Symmetric	Symmetric	Female	120
BL - Low	5918	Egalitarian	Symmetric	Other	Female	450
BL - Low	51012	Egalitarian	Egalitarian	Egalitarian	Male	500
BL - Low	51016	Egalitarian	Egalitarian	Egalitarian	Male	100
BL - Low	51018	Symmetric	Symmetric	Symmetric	Male	300
LT - Low	6111	Meritocratic	Meritocratic	Meritocratic	Male	100
LT - Low	6112	Meritocratic	Meritocratic	Meritocratic	Male	500
LT - Low	6114	Symmetric	Symmetric	Symmetric	Female	300
LT - Low	6119	Meritocratic	Meritocratic	Meritocratic	Female	60
LT - Low	6131	Egalitarian	Egalitarian	Egalitarian	Male	500
LT - Low	6138	Symmetric	Meritocratic	Other	Female	350
LT - Low	61111	Meritocratic	Meritocratic	Meritocratic	Male	450
LT - Low	61118	Egalitarian	Egalitarian	Egalitarian	Male	650
LT - Low	61313	Egalitarian	Egalitarian	Egalitarian	Male	150
LT - Low	61315	Egalitarian	Symmetric	Other	Male	450

LT - Low	61316	Symmetric	Egalitarian	Other	Female	290
LT - Low	61317	Symmetric	Symmetric	Symmetric	Female	500

Figure 1.13 Real and Corrected Bids Distribution



Note: The Two-sample Kolmogorov-Smirnov test for equality of distribution functions suggest that while the Normal and the Poisson corrected bids' distributions are not significantly different to the actual bids' distribution, the corrected distributions using the WTA values and the Uniform density function are both significantly different to the actual bids at 1%.

A.1.2. Risk aversion analysis

We find the bids making the utility level of being B above the utility of being A *i.e.* *u(B)* > *u(A)*. For our purpose, we use the following utility function suggested in Holt and Laury (2002):

$$u(x) = \frac{x^{1-\theta}}{1-\theta}, \text{for } x > 0 \quad (3)$$

θ is the constant relative risk aversion parameter that we assumed equals to 0.97[64]. The equation suggests that for θ =0, subjects value each unit of money by itself and do not experience neither benefits or loses associated to the certainty through which that money is obtained. If θ <0, individuals are categorised as risk loving, which means, the higher the risk, the more utility is reported by the same unit of money. Finally, if θ >0, subjects belong to the risk averse category. They experience disutility when the same amount of money is obtained under higher levels of uncertainty.

Replacing the earnings of A and B participants in block 2 into the utility function, we obtain the following equations:

$$u(A) = \frac{(y_i^{Block\ 2} * r_A)^{1-\theta}}{1-\theta} \quad (4)$$

$$u(B) = \sum_{j=1}^{9} \frac{\left[(p_j * (y_i^{Block\ 2} * j)) + bid\right]^{1-\theta}}{1-\theta}, \forall\, j = 1, \ldots, 9 \quad (5)$$

y_i: total productivity (the score) of the individual *i* in block 2, assumed constant.

j: piece rate at which a B-subject might get paid at. *j* varies from 1 to 9 as participants A are not allowed to pay zero.

p_j: probability *p* of being paid at a rate of *j*. These values were extracted from the actual distribution of rates assigned by participants A along block 2 during the experiment.

θ: Constant rate of risk aversion, assumed as 0.97.

r_A: rate paid to those in the A-position i.e. $r_A = 5$.

bid: WTP or WTA values.

[64] A CRRA of 0.97 is extremely high across similar experiments. As we did not control for individual levels of risk aversion in the experiment, we are just giving our findings the strongest test. Other experiments find much lower CRRA average values: Cox and Oaxaca (1996), Goeree et al (1999), Chen and Plott (1998), and Campo et al (2000) get similar values in different auctions (0.67, 0.52, 0.48, and 0.56, respectively).

A.1.3 Experimental instructions chapter 1

Here we will present all experimental instructions. We provide a full package of the NP and BL-High (which looks identical to BL-Low) experimental instructions to create an overview of what participants received. We only deliver blocks 1 and 3 instructions in the BL-High treatment as these blocks are identical for all conditions. Since block 4 is a reminder of the differences between participants A and B that have been already given to participants in block 2, we only report the instructions of block 2.

- **No Power (NP)**

Welcome to our experiment. This exercise is part of a research project in which we are trying to understand how people make decisions. For attending today, you will receive £2. Additionally, you will be paid according to the decisions you make in the experiment. During the experiment, you will see your earnings as Experimental Currency Units (ECU). Your earnings will be saved in your experimental account and will be anonymously paid to you at the end of the experiment at the exchange rate of 150 ECU = £1.

This experiment is divided in four blocks. You will get a new set of instructions at the beginning of each block. Please follow the instructions carefully. You will participate in a quiz before each block to make sure you have understood the instructions.

After you read the information onscreen or complete a task as instructed, press the “OK” or “Continue” button to move forward.

Please make sure you do not have anything on the desk. Food and drinks cannot be consumed in the lab. Electronic devices such as mobiles, tablets, laptops, iPods, etc. cannot be used whilst in the lab. Please switch them off and leave them under the desk.

You are not allowed to talk to each other during the experiment. This includes laughing aloud and shouting. Please try to remain silent and do not attempt to communicate with other people. If you have any questions, please raise your hand until the experimenter comes to your cubicle.

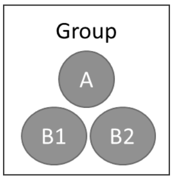
1st Block

In the 1st block of this experiment you will add 2-digit numbers for three minutes. You will earn 5 ECU per correct answer. Mistakes will not be taken into account. At the end of the block you will be informed of your score and earnings. Your earnings in this block will be paid to you at the end of the experiment.

Please press the "OK" button to start the quiz about the instructions and the 1st block.

If you have any questions whilst completing the quiz, please raise your hand and remain in silence until the experimenter comes to your cubicle.

2nd Block



In the 2nd block you will be randomly assigned with two others to form a group of three. Each group will consist of **one Participant A and two Participants B.** Participants A and B will be making exactly the same task, will be paid at exactly the same rate and you will be informed about your role (A or B) at the beginning of this block. The composition of the groups and your role will not change throughout this block. Groups are independent in that your earnings will be determined by your individual performance

As in the 1st block, you and the other participants will add 2 digit numbers for one minute and will be paid at a rate of 5 ECU for correct answer. However, in this 2nd block you will do this task 10 times. We will call each repetition of the task a round.

At the end of the block, you will receive information about your total earnings for that block. Cumulative earnings will be paid at the end of the experiment.

Please press the "OK" button to start the quiz about the instructions. Once you answer the questions you will be able to start the 2nd block.

If you have any questions whilst completing the quiz, please raise your hand and remain in silence until the experimenter comes to your cubicle.

3rd Block

All participants will receive **150 ECU** at the beginning of this 3rd block. Your role will be the same as in the 2nd block. However, all the groups will be dissolved and participants A and B will be making different decisions.

If you are a participant B, you may **buy** the right to be a participant A in the 4thBlock. If you are a participant A, you may **sell** this right to one of the participants B in the experiment. Participants A and B will simultaneously choose the prices at which they are willing to sell/buy the role of participant A in the 4th block.

The 4th block will be the last one, as the experiment will finish afterwards. The 4th block is exactly the same as the second block, so both participants A and B will be making exactly the same task, and both participants A and B will be paid at exactly the same rate. The tasks that participants have to perform are explained as following:

Participants B…

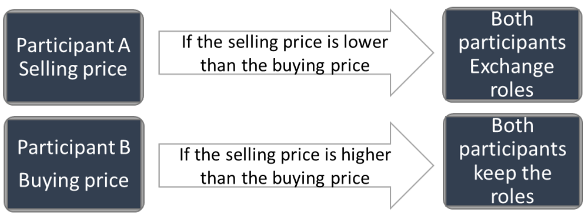
Each of the 12 participants B will choose their buying price to get the role of participant A by participating in a simple auction. The computer will start with a posted price of 10 ECU that will be increased by 10 ECU as long as the auction continues. The auction will stop when all the participants have left or when the maximum price of 500 ECU is reached.

Participants B have to decide whether they are willing to pay the posted price. If they are not, they just have to click the "Leave" button to leave the auction. Each of the 6-highest bids/buying prices will in turn be randomly compared to a different participant A's selling price.

Participants A…

Whilst participants B are choosing the buying price, each of the 6 participants A will be asked for the price at which they would be willing to sell the role as participant A. Afterwards, each of these selling prices will be compared to one of the 6-highest bids resulting from the auction.

If the participant A's selling price is lower than the buying price assigned to them, they will become a participant B in the 4th block. The participant B who made that offer will become participant A. If the selling price is higher than the auction's buying price, the roles will not change for the next block. Please note that both participants A and B will be making exactly the same task, and both participants A and B will be paid at exactly the same rate in the 4th block.



If you are a participant B, your earnings in the 3rd block will be the 150 ECU you received at the beginning of the block, **minus** the buying price you have chosen (if you are a participant B whose offer covers the selling price). If your price did not get to be offered to one participant A, or if you decide not to buy the role of participant A (e.g. by posting a bid of zero) your earnings from this block will still be 150 ECU.

If you are a participant A, your earnings in the 3rd block will be the 150 ECU **plus** the buying price offered to you, if your selling price is covered by a participant B. Alternatively, if you are a participant A whose selling price is not covered by a participant B's offer your earnings from this block will still be 150 ECU.

Please press the "OK" button to start the quiz about the instructions and the 3rd block.
If you have any questions whilst completing the quiz, please raise your hand and remain in silence until the experimenter comes to your cubicle.

4th Block

Tasks in the 4th block are identical to the 2nd block. However, the group members will be completely different. You will not receive information about past decisions or the performance of your group members. Whether you will perform as participant A or B throughout this block is dependent on the result of the previous block.

As in the 2nd block, participants A and B will add 2 digit numbers for one minute and will be paid at a rate of 5 ECU for correct answer. You will face this task 10 times.

At the end of the block, you will receive information about your total earnings for that block. Cumulative earnings will be paid at the end of the experiment.

- **Baseline High (BL-High)**

Welcome to our experiment. This exercise is part of a research project in which we are trying to understand how people make decisions. For attending today you will receive £2. Additionally, you will be paid according to the decisions you make in the experiment. During the experiment you will see your earnings as Experimental Currency Units (ECU). Your earnings will be saved in your experimental account and will be anonymously paid to you at the end of the experiment at the exchange rate of 150 ECU = £1.

This experiment is divided in four blocks. You will get a new set of instructions at the beginning of each block. Please follow the instructions carefully. You will participate in a quiz before each block to make sure you have understood the instructions.

After you read the information onscreen or complete a task as instructed, press the "OK" or "Continue" button to move forward.

Please make sure you do not have anything on the desk. Food and drinks cannot be consumed in the lab. Electronic devices such as mobiles, tablets, laptops, iPods, etc. cannot be used whilst in the lab. Please switch them off and leave them under the desk.

You are not allowed to talk to each other during the experiment. This includes laughing aloud and shouting. Please try to remain silent and do not attempt to communicate with other people. If you have any questions, please raise your hand until the experimenter comes to your cubicle.

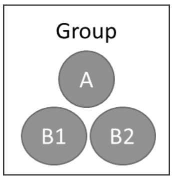
1st Block

In the 1st block of this experiment you will add 2-digit numbers for three minutes. You will earn 5 ECU per correct answer. Mistakes will not be taken into account. At the end of the block you will be informed of your score and earnings. Your earnings in this block will be paid to you at the end of the experiment.

Please press the "OK" button to start the quiz about the instructions and the 1st block.

If you have any questions whilst completing the quiz, please raise your hand and remain in silence until the experimenter comes to your cubicle.

2nd Block



In the 2nd block you will be randomly assigned with two others to form a group of three. Each group will consist of **one Participant A and two Participants B.** Participants A and B will be making different decisions and you will be informed about your role (A or B) at the beginning of this block. The composition of the groups and your role will not change throughout this block. Groups are independent in that your earnings will be determined by the decisions made within your group.

As in the 1st block, you will add 2 digit numbers for one minute. However, in this 2nd block you will do this task 10 times. We will call each repetition of the task a round. We now explain the sequence of decisions in each round and the differences between participants A and B:

Participants A…

- Earn **5 ECU** per correct answer in this block;
- At the beginning of the round they decide the rate which participants B will earn for that round. The sum of both rates has to be **equal to 10 ECU** i.e. if one participant B is awarded 3 ECU for a correct answer, the other participant B must be awarded 7;
- Have to do the same task of adding 2-digit numbers;
- Receive information about their own score/earnings and the scores of both participants B in their group at the end of each round.

Participants B…

- Earn at a rate decided by the Participant A in their group;
- Are informed of their rate at the beginning of each round;
- Have to do the task of adding 2-digit numbers;
- Receive information about their own score/earnings, but will not get information about the score of participant A or the other participant B.

At the end of the block, you will receive information about your total earnings for that block. Cumulative earnings will be paid at the end of the experiment.

Please press the "OK" button to start the quiz about the instructions. Once you answer the questions you will be able to start the 2nd block.

If you have any questions whilst completing the quiz, please raise your hand and remain in silence until the experimenter comes to your cubicle.

3rd Block

All participants will receive **150 ECU** at the beginning of this 3rd block. Now the 18 participants will be enrolled in an auction in order to **buy** the right to be a participant A in the 4thBlock (the last block).

The computer will start with a posted price of 10 ECU that will be increased by 10 ECU as long as the auction continues. The auction will stop when there are 6 participants left or when the maximum price of 500 ECU is reached. When the posted price exceeds your willingness to pay for the right of being participant A, you just have to click the "Leave" button to leave the auction. The 6 remaining participants when the auction stops will perform as participants A in the 4th block.

Your earnings in the 3rd block will be the 150 ECU you received at the beginning of the block, **minus** the last posted price when the auction stops (if you are one of the last 6 remaining participants). If you leave the auction before it stops, your earnings from this block will be 150 ECU.

Please press the "OK" button to start the quiz about the instructions and the 3rd block.

If you have any questions whilst completing the quiz, please raise your hand and remain in silence until the experimenter comes to your cubicle.

4th Block

Tasks in the 4th block are identical to the 2nd block. However, the group members will be completely different. You will not receive information about past decisions or the performance of your group members. Whether you will perform as participant A or B throughout this block is dependent on the result of the previous block.

As in the 2nd block, you will add 2 digit numbers for one minute and you will face this task 10 times. We now explain the sequence of decisions in each round and the differences between participants A and B:

Participants A…

- Earn **5 ECU** per correct answer in this block;
- At the beginning of the round they decide the rate which participants B will earn for that round. The sum of both rates has to be **equal to 10 ECU** i.e. if one participant B is awarded 3 ECU for a correct answer, the other participant B must be awarded 7;
- Have to do the same task of adding 2-digit numbers;

- Receive information about their own score/earnings and the scores of both participants B in their group at the end of each round.

Participants B…

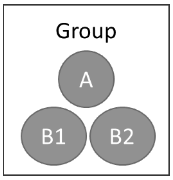
- Earn at a rate decided by the Participant A in their group;
- Are informed of their rate at the beginning of each round;
- Have to do the task of adding 2-digit numbers;
- Receive information about their own score/earnings, but will not get information about the score of participant A or the other participant B.

At the end of the block, you will receive information about your total earnings for that block. Cumulative earnings will be paid at the end of the experiment.

Please press the "OK" button to start 4th block.

- **Legitimacy High (LT-High)**

2nd Block



In the 2nd block you will be randomly assigned with two others to form a group of three. Each group will consist of **one Participant A and two Participants B.** Participants A are those who had the 6-highest scores in the previous block. The other 12 participants will be labelled as participants B. Participants A and B will be making different decisions and you will be informed about your role (A or B) at the beginning of this block. The composition of the groups and your role will not change throughout this block. Groups are independent in that your earnings will be determined by the decisions made within your group.

As in the 1st block, you will add 2 digit numbers for one minute. However, in this 2nd block you will do this task 10 times. We will call each repetition of the task a round. We now explain the sequence of decisions in each round and the differences between participants A and B:

Participants A…

- Earn **5 ECU** per correct answer in this block;
- At the beginning of the round they decide the rate which participants B will earn for that round. The sum of both rates has to be **equal to 10 ECU** i.e. if one participant B is awarded 3 ECU for a correct answer, the other participant B must be awarded 7;
- Have to do the same task of adding 2-digit numbers;
- Receive information about their own score/earnings and the scores of both participants B in their group at the end of each round.

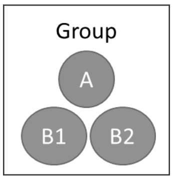
Participants B…

- Earn at a rate decided by the Participant A in their group;
- Are informed of their rate at the beginning of each round;
- Have to do the task of adding 2-digit numbers;
- Receive information about their own score/earnings, but will not get information about the score of participant A or the other participant B.

At the end of the block, you will receive information about your total earnings for that block. Cumulative earnings will be paid at the end of the experiment.

- **Legitimacy Low (LT-Low)**

2nd Block



In the 2nd block you will be randomly assigned with two others to form a group of three. Each group will consist of **one Participant A and two Participants B.** Participants A are those who had the 6-lowest scores in the previous block. The other 12 participants will be labelled as participants B. Participants A and B will be making different decisions and you will be informed about your role (A or B) at the beginning of this block. The composition of the groups and your role will not change throughout this block. Groups are independent in that your earnings will be determined by the decisions made within your group.

As in the 1st block, you will add 2 digit numbers for one minute. However, in this 2nd block you will do this task 10 times. We will call each repetition of the task a round. We now explain the sequence of decisions in each round and the differences between participants A and B:

Participants A…

- Earn **5 ECU** per correct answer in this block;
- At the beginning of the round they decide the rate which participants B will earn for that round. The sum of both rates has to be **equal to 10 ECU** i.e. if one participant B is awarded 3 ECU for a correct answer, the other participant B must be awarded 7;
- Have to do the same task of adding 2-digit numbers;
- Receive information about their own score/earnings and the scores of both participants B in their group at the end of each round.

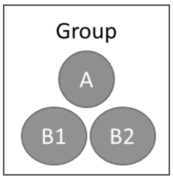
Participants B…

- Earn at a rate decided by the Participant A in their group;
- Are informed of their rate at the beginning of each round;
- Have to do the task of adding 2-digit numbers;
- Receive information about their own score/earnings, but will not get information about the score of participant A or the other participant B.

At the end of the block, you will receive information about your total earnings for that block. Cumulative earnings will be paid at the end of the experiment.

- **Centrality High (CT-High)**

2nd Block



In the 2nd block you will be randomly assigned with two others to form a group of three. Each group will consist of **one Participant A and two Participants B.** Participants A and B will be making different decisions and you will be informed about your role (A or B) at the beginning of this block. The composition of the groups and your role will not change throughout this block. Groups are independent in that your earnings will be determined by the decisions made within your group.

As in the 1st block, you will add 2 digit numbers for one minute. However, in this 2nd block you will do this task 10 times. We will call each repetition of the task a round. We now explain the sequence of decisions in each round and the differences between participants A and B:

Participants A…

- Earn **5 ECU** per correct answer in this block;
- At the beginning of the round they decide the rate which participants B will earn for that round. The sum of both rates has to be **equal to 10 ECU** i.e. if one participant B is awarded 3 ECU for a correct answer, the other participant B must be awarded 7;
- Have to do the same task of adding 2-digit numbers;
- Receive information about their own score/earnings and the scores of both participants B in their group at the end of each round.

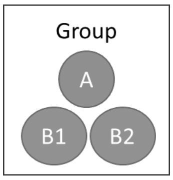
Participants B…

- Earn at a rate decided by Participant A in their group;
- Are informed of their rate at the beginning of each round;
- Have to do the task of adding 2-digit numbers;
- Receive information about their own score/earnings and participant A's score, but will not get information about the score of the other participant B.

At the end of the block, you will receive information about your total earnings for that block. Cumulative earnings will be paid at the end of the experiment.

- **Enforcement High (EN-High)**

2nd Block



In the 2nd block you will be randomly assigned with two others to form a group of three. Each group will consist of **one Participant A and two Participants B.** Participants A and B will be making different decisions and you will be informed about your role (A or B) at the beginning of this block. The composition of the groups and your role will not change throughout this block. Groups are independent in that your earnings will be determined by the decisions made within your group.

As in the 1st block, you will add 2 digit numbers for one minute. However, in this 2nd block you will do this task 10 times. We will call each repetition of the task a round. We now explain the sequence of decisions in each round and the differences between participants A and B:

Participants A…`

- Earn **5 ECU** per correct answer in this block;
- At the beginning of the round they decide the rate which participants B will earn for that round. The sum of both rates has to be **up to 10 ECU** i.e. if one participant B is awarded 3 ECU for a correct answer, the other participant B must be awarded at most 7 ECU;
- Have to do the same task of adding 2-digit numbers;
- Receive information about their own score/earnings and the scores of both participants B in their group at the end of each round.

Participants B…

- Earn at a rate decided by the Participant A in their group;
- Are informed of their rate at the beginning of each round;
- Have to do the task of adding 2-digit numbers;
- Receive information about their own score/earnings, but will not get information about the score of participant A or the other participant B.

At the end of the block, you will receive information about your total earnings for that block. Cumulative earnings will be paid at the end of the experiment.

A.2.1. Experimental instructions chapter 2

- **Weakest Link for Individuals (I-WLM) – Translated from Spanish[65]**

This experiment consists of 20 rounds. Just for being here you have already earned 300 ECU. At the beginning of the experiment you will be assigned to an independent group of four participants, so whatever happens in one group will not affect all the others. The composition of your group will not change during the experiment. You will never know the identity of the other participants in your group.

Every round, every participant gets an endowment of 50 ECU and needs to make a simple allocation decision. You may allocate part or all of this endowment to a collective account for your group. Every group member chooses an allocation to the collective account independently. The difference between your individual decision and your endowment will be automatically assigned to a private account.

Your earnings in the experiment come from both the private account and the collective account. Your earnings coming from the collective account depend on the minimum allocation to the collective account in your group (that is, the lowest allocation to this account made by a member of your group). This minimum allocation is multiplied by 8 and equally shared by every single participant in the group; that is, your earnings from the collective account is always twice the minimum individual allocation to the collective account.

Your earnings coming from the private account match your allocation to this account. They do not depend on the decisions of the other participants.

In summary, your earnings in any round are determined in the following way:

Earnings = Earnings from the private account + Earnings from the collective account

(50 ECU – my allocation to the collective account) + (2 x minimum allocation to the collective account in my group)

After each round, you will get information about your decision and the minimum allocation to the collective account made by the individuals in your group. In addition, you will get information about your earnings in the round, distinguishing between the earnings coming from the private and the collective account. On your screen you will also see a table with the same information corresponding to previous rounds.

At the end of the experiment, your ECU will be exchanged into real Euros at the exchange rate of 100ECU=€1.

[65] The I-VCM experimental instructions follow a similar dynamic to the I-WLM only changing the payoffs function: Earnings = (50 ECU – my allocation to the collective account) + (2 x average allocation to the collective account in my group)

- **Weakest Link for Teams (T-WLM) – Translated from Spanish**

This experiment consists of 20 rounds. Just for being here you have already earned 300 ECU. At the beginning of the experiment you will be assigned to an independent cohort of 12 participants, so whatever happens in one cohort will not affect all the others. Every cohort consists of 4 groups of 3 participants. The composition of your group and your cohort will not change during the experiment. You will never know the identity of the other participants in your group or cohort.

Every round, every participant gets an endowment of 50 ECU and needs to make a simple allocation decision. You may allocate part or all of this endowment to a collective account for your cohort. Every group member chooses an allocation to the collective account independently and the average of these three allocations will be your group allocation to the collective account. The difference between your individual decision (not the group decision) and your endowment will be automatically assigned to a private account.

Your earnings in the experiment come from both the private account and the collective account. Your earnings coming from the collective account depend on the minimum allocation to the collective account in your cohort (that is, the lowest allocation to this account made by a group in your cohort). This minimum allocation is multiplied by 8 and divided by four to get the benefits of every participant; that is, your earnings from the collective account is always twice the minimum group allocation to the collective account.

Your earnings coming from the private account match your allocation to this account. They do not depend on the decisions of the other participants.

In summary, your earnings in any round are determined in the following way:

Earnings = Earnings from the private account + Earnings from the collective account

(50 ECU – my allocation to the collective account) + (2 x minimum allocation to the collective account in my cohort)

After each round, you will get information about the allocation of your group (the average of your three individual decisions) and the minimum allocation to the collective account made by the groups in your cohort. In addition, you will get information about your earnings in the round, distinguishing between the earnings coming from the private and the collective account. On your screen you will also see a table with the same information corresponding to previous rounds.

At the end of the experiment, your ECU will be exchanged into real Euros at the exchange rate of 100ECU=€1.

A.3.1. Description of the Squawka performance score chapter 3

Squawka generates the performance score by processing the Opta database and feeds that are delivered to Squawka live in-game. The score is calculated by an advanced algorithm that computes over 500,000,000 data points from all on-ball and preceding actions (attack, defence and possession), pitch co-ordinates (13 as presented in Figure 3. No xx), action outcomes (success/failed) and the playing position of the player (goalkeeper, Defender, Midfielder, Forward)[66].

Although there is not a particular limit on the number of points that a player can score, on average, players that play over 90 minutes can get between 10 and 20 points. Scores greater than 50 points are a signal of good performance, and negative scores are obtained by poor performers.

Figure 3. 3 Pitch Areas, took from Squawka.com

The team performance index is the addition of all players' performance during the match.

[66] More information at http://www.squawka.com/what-is-the-squawka-player-performance-rating

A.3.2. Complementary tables and figures chapter 3

Table 3.8 Linear Regressions for Goal Differences by Leagues[67]

	La Liga GD OLS	Serie A GD OLS	Ligue 1 GD OLS	Premier GD OLS	Bundesliga GD OLS	Eredivisie GD OLS
h_win_prob	4.161***	3.666***	3.850***	2.784***	4.386***	3.209***
	(0.342)	(0.381)	(0.409)	(0.321)	(0.439)	(0.529)
hcrc	0.925	-2.126	-0.0333	-0.407	3.512	-0.448
	(3.500)	(2.377)	(1.377)	(2.022)	(6.638)	(6.295)
acrc	2.950***	1.084	2.780	-0.274	1.445	3.215
	(1.112)	(1.028)	(2.353)	(1.569)	(1.916)	(5.781)
horc	-1.540***	-1.380***	-1.069**	-1.157**	-1.973***	-2.670***
	(0.452)	(0.430)	(0.456)	(0.586)	(0.625)	(0.907)
aorc	1.663***	0.926**	1.380***	1.989***	1.821***	2.038***
	(0.422)	(0.417)	(0.389)	(0.482)	(0.583)	(0.683)
hcrc_rt	0.0187	-0.0163	0.0163	0.0200	0.0497	-0.0144
	(0.0433)	(0.0313)	(0.0207)	(0.0338)	(0.0826)	(0.0926)
acrc_rt	0.0443***	0.00127	0.0273	-0.0116	0.0194	0.0442
	(0.0169)	(0.0160)	(0.0305)	(0.0224)	(0.0352)	(0.0808)
horc_rt	-0.0137**	-0.0129**	-0.00432	-0.00301	-0.0119	-0.0328**
	(0.00635)	(0.00595)	(0.00687)	(0.00909)	(0.00872)	(0.0132)
aorc_rt	0.0178***	0.00701	0.0138**	0.0212***	0.0186**	0.0174*
	(0.00597)	(0.00575)	(0.00589)	(0.00714)	(0.00814)	(0.0102)
hcpm	0.000616	0.000766	0.00233*	0.000268	-0.000673	0.00355
	(0.00125)	(0.00101)	(0.00122)	(0.00131)	(0.00151)	(0.00230)
acpm	-0.00106	-0.00107	0.000483	0.00192	0.000871	-0.00279
	(0.00115)	(0.00112)	(0.00114)	(0.00120)	(0.00142)	(0.00212)
Constant	-7.630*	0.271	-6.392*	-2.967	-7.936	-3.203
	(4.270)	(3.212)	(3.433)	(3.865)	(8.168)	(11.20)
Observations	1,140	1,139	1,100	1,102	918	578
R-squared	0.186	0.144	0.131	0.111	0.154	0.149
Number of id_home	25	25	25	24	22	20
Captain position	Yes	Yes	Yes	Yes	Yes	Yes
Season	Yes	Yes	Yes	Yes	Yes	Yes
Test hcrc=horc	No signif.	No signif.	No signif.	No signif.	No signif.	No signif.
Test acrc=aorc	No signif.	No signif.	No signif.	No signif.	No signif.	No signif.

[67] Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 3.9 Means of Captain Characteristics by League

Captain Variables	La Liga	Serie A	Ligue 1	Premier	Bundesliga	Eredivisie
capt_age	32.33784	32.52632	30.11594	29.98592	29.11538	27.2202
capt_bmi	23.23243	23.07895	22.97246	22.8	23.49231	23.02
capt_perf_min	0.4965018	0.4402214	0.2818033	0.4087701	0.3096398	0.2374245
capt_time	45.06306	53.32435	56.58204	58.47974	59.22756	70.97607
other_perf_min	0.1519945	0.1905294	0.1429038	0.1830523	0.1418205	0.1504291
premium_min	0.3445073	0.249692	0.1388995	0.2257178	0.1678193	0.0869954
Capt_rc	0.25	0.2763158	0.3009259	0.2072072	0.198718	0.2146465

Table 3.10 T-Tests results for Captain Attributes' Differences

Captain Variables	Southern vs Northern leagues	La Liga vs. other Southern leagues	La Liga vs. all leagues
capt_age	***	**	***
capt_bmi	*	**	**
capt_perf_min	No signif.	No signif.	*
capt_time	***	***	***
other_perf_min	No signif.	No signif.	No signif.
premium_min	No signif.	No signif.	**
Capt_rc	**	No signif.	No signif.

*** p<0.01, ** p<0.05, * p<0.1

A.3.3. Complementary bivariate analysis chapter 3

We estimate a bivariate ordered Probit (bioprobit) as in Dawson and Dobson (2010) and Dawson (2012). In a nutshell, the bioprobit model contains the following structure:

$$h\_goal_l = \propto_1 + \propto_2 Hwin_l + \propto_3 HCRC_l + \propto_4 ACRC_l + \propto_5 HORC_l + \propto_6 AORC_l + \propto_7 HCRC\_RT_l + \propto_8 ACRC\_RT_l + \propto_9 HORC\_RT_l + \propto_{10} AORC\_RT_l + \propto_{11} HCPM_l + \propto_{12} ACPM_l + \propto_{13} HCPP + \propto_{14} ACPP + Season + \varepsilon_{h,l} \quad (4)$$

$$a\_goal_l = \propto_1 + \propto_2 Hwin_l + \propto_3 HCRC_l + \propto_4 ACRC_l + \propto_5 HORC_l + \propto_6 AORC_l + \propto_7 HCRC\_RT_l + \propto_8 ACRC\_RT_l + \propto_9 HORC\_RT_l + \propto_{10} AORC\_RT_l + \propto_{11} HCPM_l + \propto_{12} ACPM_l + \propto_{13} HCPP + \propto_{14} ACPP + Season + \varepsilon_{a,l} \quad (5)$$

Where Corr $(\varepsilon_{h,l}, \varepsilon_{a,l})$ = rho.

Table 3.11 delivers the result of a bioprobit estimation in which the dependent variable is the number of goals scored by home and away teams (h_goal and a_goal) coded in four categories: 0, 1, 2 and +3 (for three or more goals). The interpretation of the coefficients follows the same rules than an ordered Probit as if they were independently estimated for home and away teams. Yet, bioprobit estimations deliver more effective results for our analysis as it takes into account the error correlation between home and away results.

The results are consistent to the ones presented in Table 3.7. Across regions, red cards issued to a non-captain have a negative effect on the punished team and a positive effect on the opposition, and the magnitude of this effect does not differ between home and away teams within the same region.

A captain dismissal has, however, an impact only in Southern leagues. A home (away) captain sent-off decreases the probability for the home (away) team to score an extra goal. Moreover, in Model 3 we find that the impact of the home captain red card on the probability for the home team to get an extra goal, is significantly different the effect of a red card on the away captain. We do not find a similar situation in Model 4, though (the home captain dismissal effect is not significantly different to the away captain dismissal effect). This result reinforces our findings in Table 3.7 in which we state that whilst home teams are better at taking advantage of the away captain dismissal, away teams are less prone to use the same situation in their favour. Interestingly, as mentioned above, this finding does not hold for Northern leagues.

Table 3.11 Bivariate Regressions by Region

	Full		Southern		Northern	
	H_goal	A_goal	H_goal	A_goal	H_goal	A_goal
	(1)	(2)	(3)	(4)	(5)	(6)
h_win_prob	2.136***	-1.736***	2.353***	-1.802***	1.849***	-1.698***
	(0.0803)	(0.0799)	(0.112)	(0.114)	(0.119)	(0.118)
hcrc	-1.233**	-0.208	-1.583***	-0.733	-1.155	0.159
	(0.486)	(0.752)	(0.575)	(0.827)	(0.724)	(1.305)
acrc	0.331	-0.993***	0.625	-1.327**	-0.414	-0.508
	(0.388)	(0.347)	(0.480)	(0.517)	(0.771)	(0.322)
horc	-0.665***	0.634***	-0.539***	0.590***	-0.883***	0.756***
	(0.155)	(0.150)	(0.193)	(0.197)	(0.259)	(0.233)
aorc	0.765***	-0.708***	0.760***	-0.541***	0.800***	-0.950***
	(0.133)	(0.135)	(0.164)	(0.180)	(0.231)	(0.207)
hcrc_rt	-0.00806	-0.00697	-0.0104	-0.0115	-0.0112	-0.00633
	(0.00726)	(0.0108)	(0.00916)	(0.0121)	(0.00999)	(0.0184)
acrc_rt	-1.21e-05	-0.0131**	0.00377	-0.0176**	-0.0117	-0.00729
	(0.00596)	(0.00531)	(0.00719)	(0.00778)	(0.0125)	(0.00446)
horc_rt	-0.00662***	0.00344	-0.00567**	0.00307	-0.00866**	0.00390
	(0.00221)	(0.00211)	(0.00270)	(0.00272)	(0.00388)	(0.00337)
aorc_rt	0.00762***	-0.00756***	0.00673***	-0.00648**	0.00942***	-0.00910***
	(0.00186)	(0.00191)	(0.00228)	(0.00252)	(0.00324)	(0.00298)
hcpm	0.000689**	0.000194	0.000897*	-1.79e-05	0.000203	0.000112
	(0.000350)	(0.000322)	(0.000467)	(0.000398)	(0.000547)	(0.000563)
acpm	-4.76e-05	0.000307	-0.000520	0.000119	0.000342	3.25e-06
	(0.000344)	(0.000348)	(0.000457)	(0.000458)	(0.000539)	(0.000550)
Tests:						
capt_rc=non_capt_rc	No signif.	No signif.	*	No signif.	No signif.	No signif.
h_capt_rc (at h_goal) = a_capt_rc (at a_goal)	No signif.		No signif.		No signif.	
h_non-capt_rc (at h_goal) = a_non-capt_rc (at a_goal)	No signif.		No signif.		No signif.	
h_capt_rc (at h_goal) = a_capt_rc (at h_goal)	*		**		No signif.	
h_capt_rc (at a_goal) = a_capt_rc (at a_goal)	No signif.		No signif.		No signif.	

Note: the dependent variable of these regressions is the number of goals scored by the home/away team. We estimate a bivariate ordered Probit in order to control for the error correlation between home and away results. Robust errors in parenthesis, *** p<0.01, ** p<0.05, * p<0.1

A.3.4. Complementary captain absence analysis chapter 3

Table 3.12 illustrates three linear regressions of the goal difference that include the captain absences due to injuries (hci and aci) or substitutions (hcs and acs). We find that injuries and substitutions do not significantly affect the goal difference, neither in Southern or Northern leagues, and their coefficients are not different to one another. As the main results suggested, a captain red card only affects the Southern leagues, and the coefficient significantly differs to the substitutions' and injuries'. Since we have no information regarding injuries and substitutions of non-captains, we cannot provide evidence of the relative effect between players.

Table 3.12 Linear Regressions by Region including Injuries and Substitutions

	(1) Full	(2) Southern	(3) Northern
h_win_prob	3.853***	4.160***	3.495***
	(0.153)	(0.204)	(0.232)
hcrc	-0.768	-0.842	-0.821
	(0.948)	(1.128)	(1.736)
acrc	1.266**	1.774**	-0.0216
	(0.621)	(0.719)	(1.177)
hcrc_rt	-0.000434	-0.000406	-0.00338
	(0.0133)	(0.0154)	(0.0255)
acrc_rt	0.0144	0.0206*	-0.00263
	(0.00918)	(0.0105)	(0.0179)
hci	0.106	-0.140	0.314
	(0.328)	(0.466)	(0.469)
aci	0.499	-0.107	0.843
	(0.429)	(0.688)	(0.561)
hcs	-0.00618	0.0387	-0.0885
	(0.0613)	(0.0763)	(0.102)
acs	0.0778	0.0792	0.0757
	(0.0584)	(0.0701)	(0.103)
hcpm	0.00105**	0.00149**	0.000333
	(0.000535)	(0.000654)	(0.000912)
acpm	-0.000434	-0.000874	0.000318
	(0.000515)	(0.000646)	(0.000847)
Constant	-2.781*	-3.455**	-0.877
	(1.442)	(1.660)	(2.809)
Observations	5,977	3,379	2,598
R-squared	0.102	0.121	0.086
Number of id_home	141	75	66
Season	Yes	Yes	Yes
Home team FE	Yes	Yes	Yes
Test hcrc = hci	No signif.	No signif.	No signif
Test hcrc = hcs	No signif.	No signif.	No signif
Test acrc = aci	No signif.	*	No signif
Test acrc = acs	*	**	No signif

Note: the dependent variable of these regressions is the goal difference between the home/away team. We estimate three linear regressions in order to compare cultural effects. Standard errors in parenthesis, *** $p<0.01$, ** $p<0.05$, * $p<0.1$

A.3.5. Captain attributes chapter 3

Table 3.13 Dismissed Captain Attributes

	Type	League	Season	Date	Captain	Team	Opposition team	Red card time (min)	Playing position
1	home	Bundesliga	2012-2013	04/11/2012	Simon Rolfes	Leverkusen	Fortuna Dusseldorf	65	Midfielder
2	away	Bundesliga	2012-2013	01/12/2012	Serdar Tasci	Stuttgart	Greuther Furth	53	Defender
3	home	Bundesliga	2012-2013	15/12/2012	Mergim Mavraj	Greuther Furth	Augsburg	76	Defender
4	away	Bundesliga	2012-2013	27/01/2013	Clemens Fritz	Werder Bremen	Hamburg	80	Midfielder
5	away	Bundesliga	2013-2014	24/08/2013	Benedikt Howedes	Schalke 04	Hannover	14	Defender
6	away	Bundesliga	2013-2014	07/12/2013	Benedikt Howedes	Schalke 04	M'gladbach	44	Defender
7	away	Bundesliga	2014-2015	07/02/2015	Fabian Lustenberger	Hertha	Mainz	58	Defender
8	home	Bundesliga	2014-2015	28/02/2015	Lars Stindl	Hannover	Stuttgart	89	Forward
9	home	Bundesliga	2014-2015	16/05/2015	Paul Verhaegh	Augsburg	Hannover	90	Defender
10	away	Eridivise	2012-2013	03/11/2012	Kwame Quansah	Heracles	PSV Eindhoven	52	Midfielder
11	away	Eridivise	2012-2013	26/01/2013	Sanharib Malki	Roda	Den Haag	49	Midfielder
12	away	Eridivise	2012-2013	10/02/2013	Nick Viergever	AZ Alkmaar	Feyenoord	67	Defender
13	away	Eridivise	2012-2013	12/05/2013	Mark van Bommel	PSV Eindhoven	Twente	71	Midfielder
14	away	Eridivise	2013-2014	06/10/2013	Sander Duits	Waalwijk	PSV Eindhoven	85	Midfielder
15	home	Eridivise	2013-2014	02/04/2014	Rasmus Bengtsson	Twente	Den Haag	64	Defender
16	away	Eridivise	2014-2015	24/08/2014	Andreas Bjelland	Twente	NAC Breda	55	Defender
17	home	Eridivise	2014-2015	05/02/2015	Niklas Moisander	Ajax	AZ Alkmaar	62	Defender
18	home	Eridivise	2014-2015	15/02/2015	Willem Janssen	Utrecht	Dordrecht	60	Midfielder
19	home	Eridivise	2014-2015	22/02/2015	Marten de Roon	Heerenveen	Groningen	83	Midfielder

20	home	La Liga	2012-2013	26/08/2012	Francisco Puñal	Osasuna	Barcelona	76	Midfielder
21	away	La Liga	2012-2013	16/09/2012	Nunes	Mallorca	Osasuna	35	Defender
22	away	La Liga	2012-2013	29/09/2012	Casto	Betis	Malaga	11	Goalkeeper
23	home	La Liga	2012-2013	09/02/2013	Manuel Pablo	La Coruna	Granada	95	Defender
24	home	La Liga	2012-2013	24/02/2013	Gabi	Ath Madrid	Espanol	49	Midfielder
25	away	La Liga	2013-2014	25/08/2013	Borja Oubiña	Celta	Betis	85	Midfielder
26	away	La Liga	2013-2014	01/09/2013	Duda	Malaga	Sevilla	91	Midfielder
27	away	La Liga	2013-2014	21/10/2013	Bruno Soriano	Villarreal	Ath Bilbao	46	Midfielder
28	away	La Liga	2013-2014	01/12/2013	Francisco Puñal	Osasuna	Valencia	16	Midfielder
29	away	La Liga	2013-2014	21/12/2013	Juanfran García	Levante	Ath Madrid	90	Defender
30	away	La Liga	2013-2014	16/02/2014	Ricardo Costa	Valencia	Sevilla	50	Defender
31	away	La Liga	2013-2014	28/02/2014	Diego Mainz	Granada	Ath Bilbao	73	Defender
32	away	La Liga	2013-2014	21/03/2014	Duda	Malaga	Celta	66	Midfielder
33	home	La Liga	2014-2015	23/08/2014	Duda	Malaga	Ath Bilbao	88	Midfielder
34	home	La Liga	2014-2015	22/11/2014	Gabi	Ath Madrid	Malaga	88	Midfielder
35	away	La Liga	2014-2015	23/05/2015	Daniel Parejo	Valencia	Almeria	89	Midfielder
36	away	Ligue 1	2012-2013	18/08/2012	Mickael Tacalfred	Reims	Bastia	59	Defender
37	home	Ligue 1	2012-2013	25/08/2012	Andre Luiz	Nancy	Toulouse	49	Defender
38	away	Ligue 1	2012-2013	25/08/2012	Jean-Baptiste Pierazzi	Ajaccio	Valenciennes	38	Midfielder
39	away	Ligue 1	2012-2013	08/12/2012	Ahmed Kantari	Brest	Rennes	89	Defender
40	home	Ligue 1	2012-2013	05/05/2013	Thiago Silva	Paris SG	Valenciennes	43	Defender
41	away	Ligue 1	2013-2014	24/08/2013	Cedric Kante	Sochaux	Montpellier	74	Defender
42	away	Ligue 1	2013-2014	06/10/2013	Maxime Gonalons	Lyon	Montpellier	83	Midfielder
43	home	Ligue 1	2013-2014	09/11/2013	Lionel Mathis	Guingamp	Lille	80	Midfielder
44	home	Ligue 1	2013-2014	04/12/2013	Vitorino Hilton	Montpellier	Lorient	89	Defender
45	home	Ligue 1	2013-2014	08/12/2013	Jonathan Zebina	Toulouse	Montpellier	68	Defender
46	home	Ligue 1	2013-2014	08/12/2013	Yannick Cahuzac	Bastia	Lyon	88	Midfielder

47	away	Ligue 1	2013-2014	06/04/2014	Didier Digard	Nice	St Etienne	90	Midfielder
48	home	Ligue 1	2013-2014	04/05/2014	Bruno Ecuele Manga	Lorient	Ajaccio	25	Defender
49	home	Ligue 1	2014-2015	24/09/2014	Yannick Cahuzac	Bastia	Nantes	40	Midfielder
50	home	Ligue 1	2014-2015	17/10/2014	Jerome Lemoigne	Lens	Paris SG	53	Midfielder
51	home	Ligue 1	2014-2015	24/01/2015	Yannick Cahuzac	Bastia	Bordeaux	62	Midfielder
52	away	Ligue 1	2014-2015	21/03/2015	Ludovic Sane	Bordeaux	Toulouse	92	Defender
53	home	Ligue 1	2014-2015	02/05/2015	Vitorino Hilton	Montpellier	Rennes	93	Defender
54	away	Premier	2012-2013	04/11/2012	Fabricio Coloccini	Newcastle	Liverpool	84	Defender
55	home	Premier	2012-2013	18/11/2012	Brede Hangeland	Fulham	Sunderland	30	Defender
56	away	Premier	2012-2013	13/01/2013	Vincent Kompany	Man city	Arsenal	75	Defender
57	away	Premier	2013-2014	31/08/2013	John O'Shea	Sunderland	Crystal Palace	78	Defender
58	away	Premier	2013-2014	07/12/2013	Kevin Nolan	West Ham	Liverpool	82	Midfielder
59	away	Premier	2013-2014	01/01/2014	Kevin Nolan	West Ham	Fulham	44	Midfielder
60	away	Premier	2013-2014	19/01/2014	Nemanja Vidic	Man United	Chelsea	92	Defender
61	away	Premier	2013-2014	15/03/2014	Vincent Kompany	Man city	Hull	10	Defender
62	home	Premier	2013-2014	16/03/2014	Nemanja Vidic	Man United	Liverpool	77	Defender
63	home	Premier	2013-2014	26/04/2014	Ryan Shawcross	Stoke	Tottenham	52	Defender
64	home	Premier	2014-2015	27/09/2014	Wayne Rooney	Man United	West Ham	59	Forward
65	home	Premier	2014-2015	03/11/2014	Mile Jedinak	Crystal Palace	Sunderland	87	Midfielder
66	home	Premier	2014-2015	28/12/2014	Fabian Delph	Aston Villa	Sunderland	49	Midfielder
67	away	Premier	2014-2015	15/03/2015	Fabricio Coloccini	Newcastle	Everton	59	Defender
68	home	Premier	2014-2015	22/03/2015	Steven Gerrard	Liverpool	Man United	46	Midfielder
69	away	Serie A	2012-2013	26/09/2012	Daniele Conti	Cagliari	Milan	66	Midfielder
70	home	Serie A	2012-2013	30/09/2012	Daniele Gastaldello	Sampdoria	Napoli	66	Defender
71	home	Serie A	2012-2013	11/11/2012	Stefano Mauri	Lazio	Roma	85	Midfielder
72	away	Serie A	2012-2013	27/01/2013	Fabrizo Miccoli	Palermo	Cagliari	0	Forward
73	home	Serie A	2012-2013	28/04/2013	Daniele Gastaldello	Sampdoria	Fiorentina	80	Defender

74	away	Serie A	2012-2013	12/05/2013	Francesco Totti	Roma	Milan	92	Forward
75	away	Serie A	2012-2013	19/05/2013	Massimo Ambrosini	Milan	Siena	68	Midfielder
76	away	Serie A	2013-2014	06/10/2013	Francesco Magnanelli	Sassuolo	Parma	88	Midfielder
77	home	Serie A	2013-2014	10/11/2013	Sergio Pellissier	Chievo	Milan	91	Forward
78	away	Serie A	2013-2014	10/11/2013	Riccardo Montolivo	Milan	Chievo	83	Midfielder
79	away	Serie A	2013-2014	25/01/2014	Gianluigi Buffon	Juventus	Lazio	24	Goalkeeper
80	away	Serie A	2013-2014	16/02/2014	Daniele Gastaldello	Sampdoria	Roma	80	Defender
81	home	Serie A	2013-2014	16/03/2014	Daniele Conti	Cagliari	Lazio	77	Midfielder
82	away	Serie A	2013-2014	27/04/2014	Daniele Portanova	Genoa	Atalanta	50	Defender
83	home	Serie A	2013-2014	04/05/2014	Sergio Pellissier	Chievo	Torino	65	Forward
84	away	Serie A	2013-2014	11/05/2014	Alessandro Lucarelli	Parma	Torino	64	Defender
85	away	Serie A	2014-2015	06/01/2015	Daniele Conti	Cagliari	Palermo	26	Midfielder
86	home	Serie A	2014-2015	11/01/2015	Gianpaolo Bellini	Atalanta	Chievo	89	Defender
87	home	Serie A	2014-2015	31/01/2015	Nicolas Burdisso	Genoa	Fiorentina	81	Defender
88	home	Serie A	2014-2015	22/03/2015	Alessandro Lucarelli	Parma	Torino	36	Defender