Essays on Political Economy

Anastasia Papadopoulou

Department of Economics, Finance and Accounting,

School of Business,

University of Leicester

A thesis submitted for the degree of

Doctor of Philosophy at the University of Leicester.

Submitted: March 2021

Final version: October 2021

To my mother Chrysoula, my father Chrysostomos, and my brother Alkinoos
for being always there.

Essays on Political Economy

by

Anastasia Papadopoulou

Abstract

This thesis studies theoretically and experimentally voting decisions and their welfare implications in three chapters.

Chapter 2 discusses the consequences of heterogeneous beliefs on voting behaviour and the welfare implications driven by them. We assume an asymmetric change in voters' beliefs about the ability of politicians to govern after an implicit negative economic shock. Voters lose faith in the incumbent politician and trust an outsider challenger without prior office experience. Using an electoral competition model, we provide a connection between voters' beliefs and the post-shock equilibrium policies. Moreover, our welfare analysis shows that competition may under-provide public goods compared to a utilitarian social planner, depending on voters' belief distribution regarding the challenger politician.

Chapter 3 explores the role of identity in voters' decision to retain corrupt politicians. We build up a model of electoral accountability with pure moral hazard and bring it to the lab. Politicians must decide whether to invest in a public project with uncertain returns or keep the funds for themselves. Voters observe the outcome of the project but not the action of the politician. We run two treatments; a control and a treatment where subjects are assigned an identity. Our main result is that, upon observing a failed project, voters approve politicians of their same identity group more often than in the control, and compared to politicians of a different group.

Chapter 4 introduces a general median voter model of linear taxation with voters having different valuations of public expenditure, according to their income. Moreover, we assume non-quasi-linear preferences and we allow for distortive taxation. We discuss the welfare implications of these different assumptions separately and we see that competition might under-provide public goods, which comes in contrast with the well-known result of the literature, where competition always leads to over-provision, assuming a rightly-skewed income distribution.

Acknowledgement

Ithaka gave you the marvelous journey.
Without her you wouldn't have set out.
She has nothing left to give you now.
And if you find her poor, Ithaka won't have fooled you.
Wise as you will have become, so full of experience,
you'll have understood by then what these Ithakas mean.

C. P. Cavafy, part of "The City" from C.P. Cavafy: Collected Poems. (Princeton University Press, 1975)

First of all, a huge heartfelt thank you to my supervisors, Subir and Piercarlo, for all their crafting over the last years. Thank you for stirring continuously and subtly the ship towards Ithaka and making me a researcher. I will always remember our discussions in every office over the years and your continuous support. I will be forever grateful to your supervision and I will cherish our collaboration (and hopefully our friendship) in the years to come. Subir an extra thank you for being our seawall as a Ph.D. director; I am sure many things would have been far worse had you not been there.

A very big thank you to Maria and Santi, not only for our collaboration but also for being the first to help me start building my ship. I am immensely grateful for all your generosity, time, support, and encouragement.

Thirdly, I owe a lot to the Department of Economics. The environment I faced when I first arrived in Leicester was of a faculty that cares a lot about their Ph.D. students, offering their advice and support at any moment. Special thanks to the core of that community: David, Andre, Fabrizio, Curro, Asako, James, Dimitris, ... A separate thank you to Aris for his support, and especially for all his help when I was running my experiment. I would also like to express my gratitude to the University of Leicester, especially for the library facilities and the GTA scholarship, without which I would not have been able to pursue my Ph.D. studies. Also, Sam thank you very much for being there for us.

There are no words to describe how thankful I am to the Ph.D. center community of Astley Clarke, which I have terribly missed the last couple of years. Many thanks to all my cohort and especially to Melisa, Taya, Sara, and Barbie, but also to Pasqua, Stelios, Valeria, Emma, Sebastian, Nadia, Mehman, and Michalis. Vaggelis I cannot even describe how important your presence has been for the existence of this community; a very big thank you for your friendship, your support, and all our discussions.

Arpita thank you so very much for all the support throughout this journey. You know how much I value you in my life and I hope that we will manage to keep laughing with(out) reason for many more years.

Evelyn, there is not even enough space to list the things that I am thankful for. I will only say here a big thank you for being the first to believe that I can reach Ithaka. My thanks also to the Department of Economics at the University of Athens for everything I learned during my undergraduate and postgraduate studies. I would like to thank also Alexandros, Dimitris, Spyros, Myrto, Andreas, and Makis for their friendship and the unforgettable years of my master's studies. Fivos thank you very much for your friendship and support all these years; without you insisting I might not have ended up here today. Nana and Giorgos I will never forget the year we spent together; thank you for your true care and support. Thank you also to Lydia, Olga, Mariarina, Vasiliki, and Victoria for all the years of friendship and laughs.

I would like to close this section with the alpha and the omega of this journey. Infinite thanks for all the support (and patience) throughout the years to my parents, Chrysostomos and Chrysoula, and my brother Alkinoos. Without you, I doubt I would have been able to achieve this, and most importantly be the person I am today. Thank you for everything! A special thank you to Litsa for all her unconditional love all these years. Last but not least, I would like to thank Giannis, Tasia, and Iordanis. They might not be around us any more, but had not met them, or had not heard about them, I would not have been the same person.

Declaration

An early version of chapter 2 has been presented at the Genova Summer School in Political Economy (University of Genova, 2019). Chapter 3 is a joint work with Dr. Maria Cubel and Dr. Santiago Sánchez-Pagés and it has been presented as a poster in the 13th Annual NYU-CESS Conference (New York University, 2020).

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Politics play a major role in peoples' lives. The civil right to elect representatives is considered one of the most important ones and electoral decisions about representatives affect the way a democratic society operates. Therefore, studying peoples' voting decisions and their welfare implications is crucial to understand the electoral process in play. This thesis studies how voting behaviour is affected by the presence of heterogeneous beliefs or social identity considerations and if the electoral result is always an efficient outcome. More specifically, it answers the following main questions: What is the electoral outcome, if voters form heterogeneous beliefs about the efficiency of politicians to provide public goods and services after an (implicit) adverse economic shock? What are the welfare consequences of such an electoral competition? Do voters always punish (possibly) corrupt politicians? What is the effect of group identity on their voting behaviour? And what are the welfare implications of median voter models, once we depart from the standard assumptions of the relevant literature. In the next paragraphs we summarize the work of each chapter separately.

Chapter 2 of this thesis analyses the effect of a negative economic shock on voters' behaviour. Our motivation stems from observing the severe and seemingly lasting effects of the last financial crisis in the political scene of most western countries. Several politicians lost the office and many politicians without office experience received a lot of support after the beginning of the crisis. We use a median voter model of electoral competition with linear taxation, where our source of heterogeneity is voters' income and their different valuations (conditional on their income) about the ability of politicians (an incumbent and a challenger) to govern after an implicit adverse economic shock. We assume that people after the shock

trust an outsider politician more than the incumbent, and that different income voters form the same beliefs about the incumbent, but different beliefs about the unknown challenger. We show that the challenger politician will always win irrespective of the taxation platform that the incumbent proposes. Moreover, we show that, under standard assumptions of the income distribution, such a post-shock political competition can result in under-provision of public goods relative to the utilitarian social optimum. More specifically, if the middle class is less optimistic about the challenger politician after the shock compared to the average belief of the society, then competition under-provides public goods and services compared to the utilitarian social optimum. Therefore, the welfare result depends on voters' beliefs about the challenger and contradicts the standard welfare implications of median voter models, which predict that competition always over-provides public goods under the same assumptions about the income distribution.

Chapter 3 of this thesis discusses the role of social identity in voters' decisions to hold politicians accountable. Corruption is an important issue in many societies with the average corruption perception index being 55 out of 100 (Transparency International, 2019), pointing to the fact that voters seem not to punish corrupt politicians. In order to understand the role of social identity in voters' reluctance to vote out (possibly) corrupt politicians, we use the controlled environment of a lab experiment. We build up a model of electoral accountability, where rather than assuming that politicians differ in their level of competence, we assume that the level of public good provided is stochastic. Politicians should decide whether to invest funds in a public project with stochastic returns or keep the funds for themselves. Voters see if the project is successfully, but not if the politician invested the funds, and they decide to approve or not the politician. We run two treatments; a control and a treatment where participants are assigned in the beginning of the experiment an identity using the minimal group paradigm. Our main result shows that, after observing a failed project, voters approve politicians significantly more often when they belong to the same identity group compared to the different identity group, and the control treatment. Moreover, we observe that subjects

acting as politicians are more honest than expected by the equilibrium prediction, losing around 28.3% of the expected earnings of a dishonest politician. Also, we see that voters on average thought that the same identity politicians are more honest.

Chapter 4 of this thesis discusses the welfare implications of median voter models. These models have been used extensively in the literature, but their welfare implications are quite rigid. If a society has a rightly-skewed distribution of income, then the result of the electoral competition will always be over-spending in public goods and services compared to the first best of a utilitarian social planner. However, we saw in the second chapter of this thesis that this might not be always the case. We generalize this idea in an electoral competition model of linear taxation and public good provision with voters having different valuations of public consumption, assuming non quasi-linear preferences between private and public consumption and distortionary taxation. We discuss the welfare consequences of each assumption separately, and we see that the welfare implications of the model depend mainly on the convexity or concavity of the marginal effects of taxation across the population. More specifically, we see that if people perceive the benefits of public consumption differently then competition might under-provide public goods and services. The intuition is that the median voter receives less marginal benefit than the average marginal benefit of the society, and therefore her preferred taxation is less than the first best. A similar argument applies if we remove the quasi-linearity of the preferences between private and public consumption. Then the welfare result depends on the shape of the marginal cost of taxation, and hence the marginal utility of private consumption with respect to income. Finally, we also discuss the effects of distortive taxation in this framework. In this case, the welfare implications depend on the shape of the marginal income across the population.

Chapter 2

Over-austerity in a Model of Electoral Competition with Heterogeneous Beliefs

2.1 Introduction

Economic downturns seem to create important changes in a country's political reality. Historically there are many such evidences, like Germany after the Weimar Republic and the Great Depression, Argentina after the 2001 crisis, etc. After a long economically and politically stable period, the global economic crisis of 2008 severely affected the political sphere of the western societies. Most of the incumbent politicians failed to get re-elected and many outsider politicians gained a lot of popularity and support. For example, Donald Trump in the U.S., Nigel Farage in the U.K., Syriza in Greece, the green party in Austria or the communist party in Portugal.[1] It seems that the crisis led people to trust these outsider politicians more, even though they were not trusted to govern before.

Therefore, the following questions arise: What is the effect of an adverse economic shock to the electoral result and under which conditions the resulting policy might lead to over-austerity? In order to answer these questions we use a model of electoral competition with linear taxation that finances public expenditure. We assume two politicians; an incumbent being in office when the shock occurs and an

[1]Part of the political science literature discusses extensively this effect in Europe, like the papers by Torcal (2014) or Hernandez and Kriesi (2016). Also, a recent working paper by Guiso et al. (2017) analyze the demand and supply side of most of these parties and they find evidence that the crisis was the main reason for their presence and increase in their vote share.

outsider without office experience.[2] This assumption depicts the main characteristic of the majority of the politicians or parties mentioned above; their inexperience. More specifically, most of them were absent from the political scene before or they obtained a really small percentage of the voting share.

However, after the crisis people seemed to perceive these politicians in a different way than the incumbents. There might be various reasons why those politicians differ from the mainstream ones for the voters. It might be the case that for some people the different rhetoric plays a role, while for others it is their complete diverse image. What is crucial is that despite the (possibly) different reasons, voters seemed to find these outsider politicians more appealing after the economic crisis, compared to the incumbents. Naturally the magnitude of this effect is not the same for everybody. People may form different opinions about politicians for whom they do not have any prior information of being in office. Some people may argue that an outsider politician will be far better in handling the situation, while others may claim that they cannot be that much different from the incumbents, expressing an opinion closer to the idea of all politicians are alike.

We capture this idea in our model by implicitly modelling the shock as a perceived inability of the incumbent politician to govern after the shock.[3] Most importantly, we assume that voters correlate their beliefs about the incumbent with their beliefs about the challenger. Even though they agree about the inability of the incumbent, they disagree about the expected ability of the unknown outsider. In other words, by losing their faith to the incumbent, voters create differential beliefs about the challenger. We can also think about it as voters having different beliefs about the responsibility of the incumbent politician for this shock. Some people think that the incumbent is not the one to be blamed and therefore any challenger will perform almost the same. In contrary, others believe that this

[2]We deliberately abstract from any specific definition of populism. There are different definitions of what constitutes a populistic candidate or party. Guriev and Papaioannou (2020) provide an extensive and detailed discussion on this.

[3]In other words, voters attribute the negative shock to the incumbent politician. Some economic downfalls can be direct consequences of the incumbent's wrongdoing, while others (such as the subprime mortgage crisis) might be a global phenomenon. However, even in the second case, the crisis management can still be attributed to the incumbent politician.

shock was mainly the incumbent's fault and anybody else can perform better in such condition.

We use a model of electoral competition with linear taxation and we show that the challenger politician wins this competition irrespective of the policy proposed by the incumbent politician. Most importantly, we show that, under standard assumptions of the income distribution, such a post-shock competition can result in under-provision of public goods compared to the utilitarian social optimum. This result depends on the underlying belief distribution on the ability of the challenger to govern after the shock. If in a society after a crisis the middle class believes that the challenger politician is less competent to govern compared to the belief of the rest of the population on average, then the democratic result is an over-reaction to the shock and requires less taxation compared to the first best. However, in the opposite scenario, competition requires higher taxation and public expenditure compared to the social optimal. This result provides important analytical contributions to the literature on electoral competition, which on the contrary, predicts over-provision of public good under the same assumptions about the income distribution.

The rest of the chapter is organised as follows. The next section details the relevant literature on the topic. Section 2.3 describes the model. Section 2.4 proceeds with the analysis of the electoral competition, and Section 2.5 with the welfare analysis. Section 2.6 concludes.

2.2 Literature Review

The classical model of electoral competition traces back to the work of Hotelling (1929), Black (1948), and later Downs (1957). The model assumes that politicians' objective is to win the election, without caring for the proposed or implemented policies per se, i.e. they are just office motivated. At the same time voters care only to elect the policy most beneficial to them. In this decision politician's past performance can serve as a way to foresee future actions. Politicians announce their platforms simultaneously and they commit to them with no intentions of

shirking, being corrupted, etc. Each voter has an ideal policy and it is shown that politicians' proposed platforms converge to the median voter's ideal policy, which is the Condorcet winner (i.e. it defeats any other policy in pairwise comparisons). This model serves as a benchmark for a large part of the field of political economics and has been used to incorporate different extensions.[4] Using the implications of the median voter theorem, Romer (1975), Roberts (1977), and later Meltzer and Richard (1981) provided a framework for the analysis of taxation and redistribution. They use a model of linear taxation to study the implications of the electoral result, i.e. the optimal tax rate of the median voter. Their conclusion is that the higher the income inequality (measured as the distance between the income of the median and the average voter), the greater the redistribution.

One of the extensions of the spatial electoral competition model in the literature has dealt with the concept of valence. Although in our model the difference in voters' beliefs about the politicians is introduced in a different way through a negative shock, a brief review of this strand of the literature is worth presenting. The main way of incorporating the valence idea is to add a component in voters' preferences that differs for each politician. Thus, indicating that politicians vary in some characteristic, which can be thought as their personal charisma. Enelow and Hinich (1982) were of the first to incorporate this component in a one-dimensional model of electoral competition discussing the implications for the result of the competition and the optimal policies. Later, Ansolabehere and Snyder (2000) dealt with the problem in a multidimensional space providing necessary and sufficient conditions for the existence of equilibria, and their characterization. Moreover, Aragones and Palfrey (2002) show that the existence of a valence advantage together with the uncertainty in median voter's ideal point may force parties to choose polarised ideology, even without having any policy preferences. Also, Aragonès and Xefteris (2017) characterized all Nash equilibria in a model where voters have different valuations about the non-policy characteristic of a candidate.

Another phenomenon widely discussed in the political economics literature is

[4] A detailed presentation of the main models in the literature can be found in Persson and Tabellini (2002).

incumbency (dis)advantage. Politicians who are already in office tend to have a (dis)advantage in winning in the next electoral period. In our model the incumbent politician loses the election by being perceived as less competent in the post-shock competition compared to the challenger. This contradicts the main underlying assumption for incumbency effects, i.e., that the incumbent and challenger politicians are of the same quality, so there are other reasons that the incumbent may face this (dis)advantage. The discussion on the incumbency advantage effect started from explanations regarding the sophomore surge or the retirement slump (Erikson (1971)) followed by a seminal paper by Gelman and King (1990). They used a new empirical methodology and provided the first actual evidence of a positive incumbency advantage in the U.S. Lately, the discussion has shifted into identifying the proper empirical methodology in order to disentangle different effects to answer this question, with the regression discontinuity design (RDD) approach to be the most prominent one until today. While the largest part of this literature focuses mainly on incumbency advantage, there is evidence for the exact opposite phenomenon of incumbency disadvantage, especially in developing economies (Uppal (2009)).[5] Theoretical explanations of incumbency effects (mainly incumbency advantage) have been discussed in the literature. For example, Berganza (2000) uses a dynamic multi-period principal-agent model to show that if the economic performance of the incumbent is bad (either due to lack of ability or corruption), he will not be re-elected irrespective of the reason. The effect of signalling on the possibilities of an incumbent to remain in the office is another reason that may create incumbency advantage, as discussed by Caselli et al. (2014). Moreover, Kartik and Weelden (2019) provide an explanation for both incumbency effects mentioned above in a dynamic model of electoral accountability with term limits.

Even though, in our model we abstract from a formal representation of populism, we think that a brief overview of the literature on this topic is relevant to our discussion. In recent years studies on populism or radicalisation have increased. Acemoglu, Egorov, and Sonin (2013) model populist politicians as being harmful

[5]There are studies that find evidence for incumbency disadvantage also in developed countries. For example, Eggers and Spirling (2014) find such evidence with data from British elections.

to the rich, but also not in favour of the poor. They conclude that populist policies are used as a signal by politicians in order to convey to the voters that future policies will be close to median voters' preferences. Moreover, Acemoglu, Egorov, and Sonin (2015) discuss how institutional changes, i.e. changes in the political regime of a country, can occur in a dynamic environment during periods of turmoil. Lately, Altomonte, Gennaro, and Francesco (2019) provide a behavioural explanation of this phenomenon based on findings of social psychology and confirm their theoretical predictions analysing the emergence of the UK Independence Party in the national elections of 2010 and 2015. Moreover, Panunzi, Pavoniz, and Tabellini (2020) assume voters with reference dependent preferences who suffer from loss aversion and show that a negative economic shock can create policy divergence that might result to a risky (populist) candidate winning the election proposing lower taxation. Levy, Razin, and Young (2019) show that populist policies can prevail if we consider a misspecified model of politics. Finally, Karakas and Mitra (2020) study a model of electoral competition between an established and an outsider politician with different abilities to commit to their announcements, where the outsider is described as being more likely to deliver a policy far away from the status-quo. They show that in equilibrium there might be policy divergence and the outsider politician will target the voter group that dominates (skilled vs. unskilled).

2.3 The Model

We consider a model of electoral competition with linear taxation. We assume two competing office-motivated candidates, $j \in \{R, C\}$, with R denoting the ruling politician, and C a challenger politician.[6] There is a continuum of voters with unity mass indexed by i. They differ in their income, y_i, which is a continuous function of the voter's identifier index i and is distributed on the interval $[y_0, y_1]$ according to the $pdf\, f(y_i)$.

[6]The politicians could be thought also as two parties represented each by one politician, as we abstract from modelling any intra-party dynamics.

Voters' preferences over private consumption, c, and universally provided public goods, g, are described as follows:

$$u_i = c_i + W(g),$$

where

$$c_i = (1 - \tau)\, y_i,$$

with $\tau \in [0, 1]$ being a common tax rate and $W(\cdot)$ being a continuous, increasing, and concave function of the public good.

The government's budget constraint is balanced when the aggregate public expenditure, g, equals the total tax revenue, T:

$$g = T,$$

where

$$T = \tau \int_{y_0}^{y_1} y_i f(y_i) dy_i = \tau \overline{y}.$$

2.4 After shock competition

Our analysis starts at a stage, where the ruling politician is already in the office, when a negative economic shock occurs. After the realization of the shock, voters update their beliefs about the ruler's competence in providing public goods in the next period, and form heterogeneous beliefs about the ability of the unknown challenger. Politicians announce their proposed taxation platforms simultaneously, and after the elections, the winning politician implements his policy.

The following figure summarizes the timing of the model.

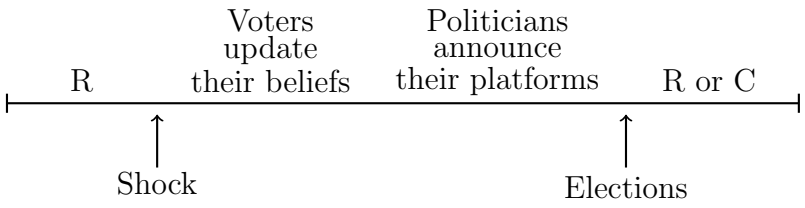


Figure 2.1: Timeline

2.4.1 Preferences

As a result of the different perceptions of voters after the shock, their expectations about the amount of public good provision each politician can deliver in the next period differ.

The expected utility of a voter if she votes for the ruler, denoted as EU_i^r, is the amount of private consumption plus the expected amount of public good provision, where parameter λ captures the probability with which she believes that the ruler will be able to provide the public good in the next period:

$$\mathrm{EU}_i^r(\tau^r, y_i) = (1-\tau^r)\, y_i + E_i[W(g^r)] = (1-\tau^r)\, y_i + \lambda\, W(\tau^r\, \overline{y}). \tag{2.1}$$

The beliefs are homogeneous across the population, as the ruler is the politician already in office and voters having prior information about him, form common beliefs about his competence in dealing with the shock.

On the other hand, the expected utility of a voter, if voting for the challenger, is:

$$\mathrm{EU}_i^c(\tau^c, y_i) = (1-\tau^c)\, y_i + E_i[W(g^c)] = (1-\tau^c)\, y_i + \phi(y_i)\, W(\tau^c\, \overline{y}). \tag{2.2}$$

In this case the probability with which voters believe that the challenger can provide the public good in the next period is a function of voters' income, i.e. $\phi(y_i) \in (0, 1]$.[7] This is a way to reflect peoples' different opinions about the challenger after a negative economic shock. Our assumption is that, because voters do

[7]There might be other individual characteristics (such as education) that can potentially be used as a way of heterogeneity, but we consider income to be an appropriate proxy of these variables for our analysis.

not have any prior information about the challenger, they create different beliefs about his ability to deal with any potential shock. Driving our motivation from real situations, we think that it is a natural way to incorporate heterogeneity in peoples' beliefs. For example, after the economic crisis not all people had the same opinion about the outsider politicians fighting for office, as discussed in the introduction of the chapter. Most importantly, we assume that the challenger is more efficient in the production of public goods after the shock compared to the the ruler, i.e. $\lambda < \phi(y_i)$.[8] So, voters disagree on the relative difference between their efficiency levels but they are unanimous in the absolute comparison between them. This follows from the fact that the incumbent politician is in charge when the negative shock happens, so voters attribute the situation to him. At the same time by losing faith to the incumbent politician, voters trust an outsider challenger to be able to deal better with such a situation. We could think about this assumption also as voters losing their trust to the status-quo and believing that an outsider politician could deal better with this shock environment.

2.4.2 Assumptions

Before proceeding with the analysis, we would like to summarize and present formally the main assumptions of our model:

Assumption 2.1 (Public good technology)**.** $W(\cdot)$ *is continuous, increasing, and strictly concave.*

Assumption 2.2 (Limit of $W'(\cdot)$)**.** $\lim_{x\to 0} W'(x) \to +\infty$.

Assumption 2.3 (Efficiency difference)**.** $\lambda < \phi(y_i) \,\forall\, i, \phi(y_i) \in (0,1] \,\forall\, i, \lambda \in (0,1)$.

Assumption 2.4 (Monotonicity of preferences)**.** $\frac{\partial \tau_i^j}{\partial y_i} < 0 \,\forall\, y_i \in [0,1]$.

The last assumption describes the decreasing monotonicity on voters optimal taxation points. The standard result of a symmetric electoral competition model is the application of the median voter theorem, i.e. both politicians converge in

[8] A short discussion on the effects if this inequality is flipped is presented in the conclusion of the chapter.

the optimal taxation platform of the median voter, as she is pivotal. A condition for the theorem to hold is the single-peakedness of the preferences. In our case there is a second type of heterogeneity regarding the expected public good provision. Assuming voters' preferences are monotonically decreasing with their income allows us to use a similar separation argument. The monotonicity of preferences is related to the super-modularity of the utility function and the single crossing property (SCP) (see Gans and Smart (1996)). More specifically SCP will imply monotonicity, requiring the cross-derivative of the function to be negative.[9]

It is important to discuss the conditions under which the assumption holds for both the preferences regarding the ruler and the challenger. For the preferences of the voters regarding the ruler the monotonicity follows directly from the concavity of $W(\cdot)$. However, for the preferences regarding the challenger the concavity of $W(\cdot)$ is not enough. The following lemma provides the condition needed for those preferences to be monotonically decreasing in income.

Lemma 2.1 (Monotonicity of preferences)**.**

$$\frac{\partial \tau_i}{\partial y_i} < 0 \,\forall\, y_i \in [0,1]$$

$$\textit{iff}$$

$$-\frac{-1+\phi'(y_i)\,W'(\tau_i\,\overline{y})\,\overline{y}}{\phi(y_i)\,W''(\tau_i\,\overline{y})\,\overline{y}^2} < 0. \tag{2.3}$$

Proof. See appendix A. □

Notice that a sufficient condition for the preferences regarding the challenger to be monotonically decreasing is that $\phi(y_i)$ is decreasing, i.e. $\phi'(y_i) \leq 0$. But, as the condition is not necessary, there might be cases where $\phi(y_i)$ is monotonically

[9]SCP can be written as follows (see also Amir (1996) and Amir (2005) for an overview on super-modular games and the connection with SCP):

If $\tau^c > \tau^{c'}$ and $y_i^{'} < y_i$, or if $\tau^c < \tau^{c'}$ and $y_i^{'} > y_i$ then

$$EU_i^c(\tau^c, y_i) \geq EU_i^c(\tau^{c'}, y_i) \Rightarrow EU_i^c(\tau^c, y_i^{'}) \geq EU_i^c(\tau^{c'}, y_i^{'}).$$

It can be shown that the property will require as a sufficient condition the $\phi(y_i)$ function to be decreasing.

increasing or even non-monotonic, but the monotonicity of preferences still holds. This means that in a society where more affluent voters tend to believe that more or less all politicians are alike, while the poorer voters have increased hope in the competence of the challenger compared to the ruler, monotonicity of preferences is guaranteed. But at the same time this condition does not exclude cases, where $\phi'(y_i) > 0$, i.e., the richest citizens are the ones who believe that after the shock a completely unknown politician is far better in handling the situation.

2.4.3 Equilibrium

In this subsection we analyse the equilibrium solution of the electoral competition. We first discuss briefly that under our assumptions the ruler cannot win this competition (lemma 2.2) and then we show the equilibrium platform (proposition 2.1).

Given that the challenger is more competent in providing public goods after the shock in the eyes of all voters (assumption 2.3), it follows that the ruler cannot win the competition. The challenger can always best respond with the same taxation platform and be preferred by all voters. The following lemma formally shows this argument.

Lemma 2.2 (R cannot win)**.** *For any $\tau^r \in [0,1]$ there exists a winning best response of the challenger, τ^c, such that he collects all the votes.*

Proof. See appendix A. □

Our next step is to show the equilibrium platform of the model. Proposition 2.1 below proves that if the challenger chooses a taxation platform equal to the optimal point of the median voter (if she would have voted for him), then the majority of the population votes for him, unconditional of the platform the ruler chooses. We start our proof by assuming that the challenger proposes the optimal taxation platform of the median voter. We first show that in this case the median voter prefers the challenger compared to the ruler for any taxation platform that the ruler proposes. Then, we split the proposed taxation platforms by the ruler

in three cases: equal to the proposed platform of the challenger, lower, and higher than that. We show that in all cases a majority of the population supports the challenger.

Proposition 2.1 (Equilibrium)**.** *(τ_m^c, τ^r) is an equilibrium where the challenger would win the competition against any $\tau^r \in [0,1]$.*

Proof. Assume that the challenger chooses τ_m^c such that it maximizes the expected utility of the median voter regarding him, i.e.

$$\tau_m^c \text{ is such that } \operatorname*{argmax}_{\tau} \mathrm{EU}_m^c(\tau, y_m) = (1-\tau)\, y_m + \phi(y_m)\, W(\tau\, \overline{y}). \tag{2.4}$$

For any $\tau^r \in [0,1]$ chosen by the ruler, the median voter prefers the challenger, i.e.

$$\mathrm{EU}_m^r(\tau^r, y_m) < \mathrm{EU}_m^c(\tau_m^c, y_m). \tag{2.5}$$

To see this assume that the ruler chooses τ_m^r such that it maximizes the expected utility of the median voter regarding him, i.e.

$$\tau_m^r \text{ is such that } \operatorname*{argmax}_{\tau} \mathrm{EU}_m^r(\tau, y_m) = (1-\tau)\, y_m + \lambda\, W(\tau\, \overline{y}). \tag{2.6}$$

Then by assumption 2.3 we have that

$$\mathrm{EU}_m^r(\tau_m^r, y_m) \leq \mathrm{EU}_m^c(\tau_m^r, y_m). \tag{2.7}$$

But by the definition of τ_m^c we know that

$$\mathrm{EU}_m^c(\tau, y_m) \leq \mathrm{EU}_m^c(\tau_m^c, y_m) \text{ for any } \tau \in [0,1].$$

Therefore, since $\tau_m^r \neq \tau_m^c$:

$$\mathrm{EU}_m^c(\tau_m^r, y_m) < \mathrm{EU}_m^c(\tau_m^c, y_m). \tag{2.8}$$

Combining inequalities (2.7) and (2.8) we get that:

$$\mathrm{EU}_m^r(\tau_m^r, y_m) < \mathrm{EU}_m^c(\tau_m^c, y_m). \tag{2.9}$$

We know also that by the definition of τ_m^r

$$\mathrm{EU}_m^r(\tau, y_m) \leq \mathrm{EU}_m^r(\tau_m^r, y_m) \text{ for any } \tau \in [0,1]. \tag{2.10}$$

Then inequality (2.5) holds after combining inequalities (2.9) and (2.10).
The ruler has three different options:

1. Choose τ^r such that $\tau^r = \tau_m^c$.
   Note that $\tau_m^c \neq 0$ and by assumption 2.3 we have that

   $$\mathrm{EU}_i^r(\tau_m^c, y_i) < \mathrm{EU}_i^c(\tau_m^c, y_i) \text{ for any } y_i \in [0,1].$$

   So, in this case all voters prefer the challenger.

2. Choose τ^r such that $\tau^r < \tau_m^c$.
   By assumption 2.4 it follows that the optimal taxation platforms considering either politician, τ_i^j, are greater than τ_m^j for any voter with less income than the median income. Therefore:

   $$\mathrm{EU}_i^c(\tau_m^c, y_i) > \mathrm{EU}_i^c(\tau^r, y_i) \text{ for any } y_i < y_m. \tag{2.11}$$

   Also following assumption 2.3 we know that

   $$\mathrm{EU}_i^c(\tau^r, y_i) > \mathrm{EU}_i^r(\tau^r, y_i). \tag{2.12}$$

   Combining inequalities (2.11) and (2.12) we see that in this case any voter with income $y_i < y_m$ votes for the challenger.
   Now, considering the voters with more income than the median, we show that there exists a voter y_i such that $y_i = y_m + \varepsilon$, with $\varepsilon > 0$ and small

enough, for which:

$$\mathrm{EU}_i^c(\tau_m^c, y_i) > \mathrm{EU}_i^r(\tau^r, y_i) \Leftrightarrow$$
$$(1-\tau_m^c)(y_m+\varepsilon) + \phi(y_m+\varepsilon)W(\tau_m^c\,\overline{y}) > (1-\tau^r)(y_m+\varepsilon) + \lambda W(\tau^r\,\overline{y}).$$

To see this, note that:

$$\lim_{\varepsilon\to 0}\{(1-\tau_m^c)(y_m+\varepsilon)+\phi(y_m+\varepsilon)W(\tau_m^c\,\overline{y})-(1-\tau^r)(y_m+\varepsilon)+\lambda W(\tau^r\,\overline{y})\} =$$
$$= \mathrm{EU}_m^c(\tau_m^c, y_m) - \mathrm{EU}_m^r(\tau^r, y_m) > 0.$$

3. Choose τ^r such that $\tau^r > \tau_m^c$.

   The argument is symmetric with the one in the second case.

As a result if the challenger chooses τ_m^c, he gains the vote of the majority of the population for any $\tau^r \in [0,1]$. □

We would like to stress that this equilibrium, τ_m^c, holds for any possible τ^r chosen by the ruler. Therefore even if the ruler was randomizing his chosen platform, τ_m^c would have still been a winning platform by the challenger, even though he would not have known the platform chosen by the ruler.[10] This is also a reason why we focus on this particular equilibrium platform, even though it is common to deal with multiple equilibria in this framework, following the asymmetry of the model.

2.5 Welfare Analysis

In this section we focus our attention on the welfare implications of our model. We start our analysis by presenting and discussing the welfare benchmark that we use. Then, we split our analysis into two subsections. First we discuss the welfare implications of the model for the equilibrium platform that we proved in

[10]This equilibrium platform would have survived also if we were to use certain equilibrium refinement concepts. For example, because it is robust to any possible strategies chosen by the ruler, it would have survived a risk dominance refinement or a trembling hand one.

the previous section, and then we extend our analysis to any possible alternative equilibria. Recall that, as we discussed above, the optimal platform of the median voter might not be the unique equilibrium. However, this equilibrium platform is robust to any mixing strategies by the ruler, and thus is the one we focus on in subsection 2.5.2.

2.5.1 Utilitarian benchmark

There are two points to consider regarding a social welfare analysis in this framework. The first is which type of beliefs the planner should consider and the second is how the planner aggregates heterogeneous preferences.

For the first point, we assume that the social planner considers the problem at the same time as the voters, i.e. after the shock has happened but before the next election. As a result, under the assumption that all voters agree that the challenger is more efficient than the ruler at that point (assumption 2.3), it seems natural to assume that the social planner considers the expected utility regarding the more efficient candidate, i.e. the challenger.

Regarding the second point, there is an ongoing debate on the most suitable way to sum up heterogeneous beliefs, especially in situations, where people with different beliefs may engage in some type of trading. This behaviour is mainly observed in financial markets, and these points are stressed by the financial literature discussing the causes of financial bubbles. In that case because of this belief heterogeneity people may end up in departing from risk-sharing, or in engaging in hedging activities. This might be in principle Pareto optimal but there is no actual gain for the society from this trading.[11] Proposals of ways of how to measure the welfare implications under this framework, include the use of different weights with which the social welfare function incorporates the beliefs, but also an idea of using risk-sharing and not Pareto optimal conditions to measure efficiency.[12]

[11] A famous example is attributed to Kreps (2012) and his discussion of the pillow fight.

[12] Some of the main papers in this literature are the ones by Kim (2012), Brunnermeier, Simsek, and Xiong (2014), Gilboa, Samuelson, and Schmeidler (2014), and Gayer et al. (2015). An overview can be found in Xiong (2013).

We do argue that in our framework voters hold heterogeneous beliefs about the possible competence levels of the two politicians after the shock and they vote sincerely without any engagement in cooperation or trade. So, we believe that the main concern for questioning a utilitarian aggregation approach of heterogeneous beliefs does not really apply to our case. A weighted welfare function will indicate that the planner "respects" more the opinions of some voters compared to others. This seems quite a restrictive assumption, as we see the social planner as an outsider institution that should take into account equally the opinions/beliefs of all citizens without trying to be condescending to the electorate. As a result assigning the same weight to all different voters' beliefs reflects the above idea.

Therefore the socially optimal taxation rate, τ_s^c, is the result of the optimization problem of the aggregate preferences of all voters, i.e.

$$\tau_s^c = \underset{\tau}{\operatorname{argmax}} \int_{y_0}^{y_1} \mathrm{EU}_i^c(\tau^c, y_i) f(y_i)\, dy_i.$$

Substituting equation (2.2) into the integral and applying the FOC we have that τ_s^c solves the following equation:

$$W'(\tau_s^c\, \overline{y}) = \frac{1}{\overline{\phi}}, \tag{2.13}$$

where $\overline{\phi} = \int_{y_0}^{y_1} \phi(y_i) f(y_i)\, dy_i$.

The standard welfare result of median voter models of redistribution is that predictions on efficiency depend on the relative position between the median and the mean income voters. More precisely, if the distribution of income is symmetric across the electorate, then competition is efficient, as the median and the mean voters coincide. In other words the pivotal voter of the electoral competition is the average income voter, whose utility is also the objective of the social planner. However, adopting a rightly-skewed income distribution drives competition to always over-provide public goods compared to the social optimum.[13]

[13]Related discussions can be found amongst others in Stiglitz (2000) or Acemoglu and Robinson (2005).

In the following subsections we show that this post-shock competition may revert the standard welfare implications of median voter models and result in under-provision of public goods, even in the case of a rightly-skewed income distribution.

2.5.2 Prominent equilibrium

In the next two propositions we show the conditions for under- or over-provision for the prominent equilibrium shown in section 2.4. These have a clear interpretation in terms of the shape of the voters' beliefs about the relative efficiency of the ruler and the challenger politicians. We conclude by discussing these interpretations.

We start proposition 2.2 by showing that if voters' belief function regarding the challenger, $\phi(y_i)$, is strictly convex in income, then competition leads to over-austerity even with a symmetric income distribution. By over-austerity we mean that after-shock competition diminishes the amount of public goods below the level a social welfare benchmark (in our case the utilitarian benchmark) would prescribe. This follows from the fact that the belief of the average voter is lower than the average belief of the society that the social planner considers. Then we assume a rightly-skewed income distribution and we show that the result holds, if the distance between the median and the average income is sufficiently small. It is important to mention that this result depends only on the convexity of the belief function, and not on the type of monotonicity of the function.

Proposition 2.2 (Under-provision)**.**

$$\tau_m^c < \tau_s^c$$

if $\phi(y_i)$ is a strictly convex function of income and y_m is sufficiently close to $\overline{y}$.

Proof. τ_m^c is a winning platform for the challenger and recall that it is the optimal taxation point of the median voter regarding him, i.e.

$$\tau_m^c \text{ is such that } \underset{\tau}{\operatorname{argmax}}\, \mathrm{EU}_m^c(\tau, y_m) = (1-\tau)\, y_m + \phi(y_m)\, W(\tau\, \overline{y}).$$

Considering F.O.C. we have that:

$$\frac{\partial \mathrm{EU}_m^c}{\partial \tau^c} = 0 \Rightarrow -y_m + \phi(y_m)\, W'(\tau^c\, \overline{y})\, \overline{y} = 0 \Rightarrow W'(\tau_m^c\, \overline{y}) = \frac{y_m}{\overline{y}} \frac{1}{\phi(y_m)}. \tag{2.14}$$

Recall also that τ_s^c is the optimal choice of the social planner and it is defined by the following equation:

$$W'(\tau_s^c\, \overline{y}) = \frac{1}{\overline{\phi}}. \tag{2.15}$$

By concavity of W (assumption 1) we have that

$$\tau_m^c < \tau_s^c \Leftrightarrow W'(\tau_m^c\, \overline{y}) > W'(\tau_s^c\, \overline{y}). \tag{2.16}$$

Substituting equations (2.14) and (2.15), inequality (2.16) becomes

$$\frac{y_m}{\overline{y}} \frac{1}{\phi(y_m)} > \frac{1}{\overline{\phi}} \Rightarrow \overline{\phi} > \frac{\overline{y}}{y_m} \phi(y_m). \tag{2.17}$$

Assuming a symmetric (e.g. uniform) income distribution, inequality (2.17) becomes

$$\overline{\phi} > \phi(\overline{y}). \tag{2.18}$$

In this case, the above relationship follows from Jensen's inequality as ϕ is strictly convex.

Relaxing the assumption of a symmetric distribution, we focus here on the case where $y_m < \overline{y}$, as a rightly-skewed income distribution is empirically relevant. In this case a simple continuity argument of the ϕ function assures that inequality (2.17) continues to hold when the difference between the median and the mean income is sufficiently small. More precisely, let us denote this difference by δ, i.e. $\delta = \overline{y} - y_m$, with $\delta > 0$. Continuity of ϕ assures that:

$$\lim_{\delta \to 0} \{ \frac{\overline{y}}{y_m} \phi(y_m) \} = \phi(\overline{y}),$$

that is, more specifically relevant for our purposes, function $\phi(y_m)$ smoothly approaches $\phi(\overline{y})$ as delta approaches zero from positive values. It must then exist a

critical value for the difference between the mean and the median income. Let us denote this value by $\overline{\delta}$, such that inequality (2.17) holds for any positive δ lower than $\overline{\delta}$. Reformulating condition (2.17) as an equality, $\overline{\delta}$ can be implicitly defined as:

$$\overline{\delta} = \overline{y}(1 - \frac{\phi(\overline{y})}{\overline{\phi}}(1 + \Delta(\overline{\delta}))), \tag{2.19}$$

where $\Delta(\overline{\delta}) = \frac{\phi(y_m) - \phi(\overline{y})}{\phi(\overline{y})}$.

□

In the following proposition we discuss the condition for an over-provision result. Following from the result above, in this case, competition over-provides public goods, even with a symmetric income distribution, as long as the belief function is strictly concave. We show that for rightly-skewed income distribution, if the belief function is also decreasing, then the result still holds without any additional assumption. If the ϕ function is increasing, then competition can still over-provide public goods, if the function is not very steep in the region between the median and the average voter.

Proposition 2.3 (Over-provision)**.**

$$\tau_m^c > \tau_s^c$$

if $\phi(y_i)$ is a strictly concave function of income.

Proof. Recall, that by concavity of W (assumption 1) we have that

$$\tau_m^c > \tau_s^c \Leftrightarrow W'(\tau_m^c\, \overline{y}) < W'(\tau_s^c\, \overline{y}). \tag{2.20}$$

Substituting equations (2.14) and (2.15), inequality (2.20) becomes

$$\overline{\phi} < \frac{\overline{y}}{y_m}\phi(y_m). \tag{2.21}$$

Assuming a symmetric income distribution, inequality (2.21) becomes

$$\overline{\phi} < \phi(\overline{y}). \tag{2.22}$$

In this case, the above relationship follows from Jensen's inequality since ϕ is strictly concave.

Relaxing the assumption of a symmetric distribution, we focus here on the case where $y_m < \overline{y}$, as a right-skewed income distribution is empirically relevant. In this case if ϕ is a decreasing function of income, then inequality (2.21) holds, as

$$\overline{\phi} < \phi(\overline{y}) < \frac{\overline{y}}{y_m}\phi(y_m).$$

However, if ϕ is increasing, then (2.21) holds, if

$$\frac{\partial\left(\frac{\overline{y}}{\overline{y}-\varepsilon}\phi(\overline{y}-\varepsilon)\right)}{\partial\varepsilon} > 0$$

with $\varepsilon > 0$. Notice that in this case we restrict ϕ to be increasing up to the point that the monotonicity of preferences (assumption 2.4) is not reversed.

□

We see that after the shock distorts the beliefs of individuals the welfare implications of the model depend on the way different parts of the population perceive the two politicians. In the absence of the shock all voters would perceive both politicians equally, and thus combined with the quasi-linearity of the utility function the welfare result would depend on the relative effect of taxation and redistribution. However, the existence of heterogeneous beliefs create one more channel that affects the welfare implications in our case.

If ϕ is strictly convex in income then there is an effect that works in the opposite direction than the one of the quasi-linearity. It is like competition overreacts to the shock by under-providing public goods compared to the first best. More specifically, in order for the challenger to win the competition with a platform that creates under-provision of public goods, it should be the case that the average

pessimism of the population for the ability of the ruler (compared to the challenger) is lower than the pessimism of the median voter. In other words, other parts of the population feel much more confident about the ability of the challenger to provide public goods after the shock than the middle class. In a way the middle class that has voted for the ruler in the previous period has an inertia that prevents them of becoming too pessimistic about him. If ϕ is also increasing in income then this increased optimism about the competence of the challenger comes from the richest part of the population, while the poorer people do not believe that the two politicians are very different. On the other hand, if the function is decreasing, the poorer voters are the ones who develop higher faith on challenger's ability. If ϕ is strictly concave then the effect of concavity works towards the same direction with the quasi-linearity effect. The reason is that, in this case, it is the middle class that loses its faith to the ruler and perceives the unknown challenger as much more competent to govern after the shock. Therefore, the median voter prefers higher public good provision from the challenger compared to the social optimal that reflects the average beliefs of the population.

So, the welfare implications depend strongly on the type of the shock and how it affects the beliefs of the population. If a shock forces the middle class to become over-optimistic about the challenger, then competition will inefficiently spend more in public expenditure. Whereas if the middle class, after a shock, does not blame the ruler much more than the challenger relative to the rest of the society, then the equilibrium result will lead to over-austerity.

2.5.3 Alternative equilibria

In this subsection we continue our analysis by showing that the over-provision result is robust even if there exist multiple equilibria. In this case, under some more stringent conditions, the taxation platform that is the upper bound of the equilibrium set is still lower than the social optimal tax rate. The intuition of our result is that if the two politicians are not very different in the eyes of the median voter and if the challenger proposes a taxation platform that is higher than the

optimal point of the median voter (regarding him), then the ruler could have won the competition by proposing a platform equal to the optimal point of the median voter, τ_m^r. Then for sufficiently close values of the two probabilities, the upper bound of the equilibrium set will still be below the social optimal.

We start our analysis by showing that if the ruler proposes a lower taxation platform than the challenger, then the supporting vote of the median voter is a necessary and sufficient condition for any politician to win the election. To show this, we show that the difference between voters' expected utility from the challenger and the ruler is monotonic in income. Therefore, if the median voter votes for the challenger then the majority of the population votes for the challenger and vice versa. For the case, where the ruler proposes a lower taxation platform than the challenger, this difference in the expected utilities is monotonically decreasing in income if voters' beliefs are also decreasing in income.

Lemma 2.3. *For any $\tau^r < \tau^c$, $EU_i^c(\tau^c) > EU_i^r(\tau^r)$ for the majority of the population iff $EU_m^c(\tau^c) > EU_m^r(\tau^r)$.*

Proof.

- If $\mathrm{EU}_m^c(\tau^c) > \mathrm{EU}_m^r(\tau^r)$ then $\mathrm{EU}_i^c(\tau^c) > \mathrm{EU}_i^r(\tau^r)$ for the majority.

  Consider the derivative of the difference between the two expected utilities w.r.t. income:

  $$\frac{\partial(\mathrm{EU}_i^c(\tau^c) - \mathrm{EU}_i^r(\tau^r))}{\partial y_i} = \tau^r - \tau^c + \phi'(y_i)W(\tau^c\overline{y})$$

  If this difference is monotonically increasing in y_i then it follows that if the median prefers the challenger then all voters with $y_i > y_m$ prefer also the challenger. Alternatively, if the difference is monotonically decreasing then all voters with $y_i < y_m$ prefer the challenger. In both cases there exists by continuity a mass of voters to the left, or the right of the median voter respectively who also prefer the challenger. For $\tau^r < \tau^c$ it is sufficient that $\phi(y_i)$ is decreasing, so for the difference to be also monotonically decreasing.

- If $\mathrm{EU}_i^c(\tau^c) > \mathrm{EU}_i^r(\tau^r)$ for the majority then $\mathrm{EU}_m^c(\tau^c) > \mathrm{EU}_m^r(\tau^r)$.

  By the monotonicity of the difference of the expected utilities, the set of

voters voting for the challenger cannot be disjoint. As a result, if the majority of voters prefer the challenger then the median voter votes also for the challenger.

□

We continue the analysis by showing that, if the challenger proposes a platform that is higher than the prominent equilibrium rate, then the ruler will propose a platform less than the proposed platform of the challenger. The intuition of this result is that if the ruler and the challenger are of almost equal ability in the eyes of the median voter, then the ruler by proposing the optimal taxation platform of the median voter could win the competition.

Lemma 2.4. *For any $\phi(y_m)$ sufficiently close to λ the ruler's best response for any $\tau^c > \tau^c_m$ is such that $\tau^r = \tau^r_m$.*

Proof. The ruler has three options:

1. $\tau^r = \tau^c$. By assumption 2.3, i.e. $\phi(y_i) > \lambda$, all voters prefer the challenger.

2. $\tau^r > \tau^c$. The median voter votes for the challenger, as τ^c is closer to her optimal point compared to τ^r. All voters with $y_i > y_m$ prefer less taxation than the median voter from each candidate. Consequently if we assume a challenger choosing τ^c and one choosing τ^r, then the voters will prefer the one choosing τ^c, i.e.,

$$EU^c_i(\tau^c) > EU^c_i(\tau^r). \tag{2.23}$$

We also know that the challenger is more competent than the ruler for all voters (assumption 2.3), i.e., voters prefer the challenger offering τ^r than the ruler:

$$EU^c_i(\tau^r) > EU^r_i(\tau^r). \tag{2.24}$$

Combining both inequalities (2.23) and (2.24) we have that:

$$EU^c_i(\tau^c) > EU^r_i(\tau^r). \tag{2.25}$$

Now consider the voters with $y_i < y_m$. There exists a mass of voters close to the median one, who still vote for the challenger. To see this, consider the following continuity argument:

$$\lim_{\varepsilon \to 0}\{(1-\tau^c)(y_m-\varepsilon)+\phi(y_m-\varepsilon)W(\tau^c\,\overline{y})-(1-\tau^r)(y_m-\varepsilon)+\lambda W(\tau^r\,\overline{y})\} = \\ = \mathrm{EU}^c_m(\tau^c)-\mathrm{EU}^r_m(\tau^r)>0$$

So, the challenger wins the majority of voters.

3. $\tau^r < \tau^c$. In this case by lemma 1 voters' preferences are monotonic in income, i.e. the ruler wins the election iff the median voter votes for him:

$$EU^r_m(\tau^r) \geq EU^c_m(\tau^c). \tag{2.26}$$

Therefore, the ruler chooses the optimal platform of the median voter, τ^r_m. In this case for values of λ sufficiently close to $\phi(y_m)$ the median voter will prefer the ruler compared to the challenger. To see this, consider a platform chosen by the challenger such that $\tau^c = \tau^c_m + \zeta$ with $\zeta > 0$. Then inequality (2.26) becomes

$$EU^r_m(\tau^r_m) \geq EU^c_m(\tau^c_m+\zeta). \tag{2.27}$$

Now assume that $\lambda = \phi(y_m) - \varepsilon$ with $\varepsilon > 0$. Then for small values of ε inequality (2.27) becomes:

$$\lim_{\varepsilon \to 0}\{EU^r_m(\tau^r_m)-EU^c_m(\tau^c_m+\zeta)\} = EU^c_m(\tau^c_m)-EU^c_m(\tau^c_m+\zeta)>0, \tag{2.28}$$

since τ^c_m is such that $\operatorname{argmax}_\tau \mathrm{EU}^c_m(\tau)$.

□

Proposition 2.4 below concludes our analysis. We show that if multiple equilibria exist, then the upper bound of the equilibrium set can still be lower than the social optimum if the median perceives the challenger after the shock as being almost as competent as the ruler. The intuition of our result is that if the belief

of the median voter for the ruler, λ, approaches her belief about the challenger, $\phi(y_m)$, then the the equilibrium set will be shrinking. As a result, the upper bound of the set will be quite close to the prominent equilibrium tax rate and thus less than the social optimum.

Proposition 2.4. *For any $\phi(y_m)$ sufficiently close to λ the upper bound of the equilibrium set is lower than the social optimum.*

Proof. We know from lemma 2 that for any $\tau^c > \tau^c_m$, the ruler best responds by choosing $\tau^r = \tau^r_m < \tau^c$. From lemma 1 we know that in this case the challenger wins the competition iff $\mathrm{EU}^c_m(\tau^c) > \mathrm{EU}^r_m(\tau^r_m)$. Let us denote by $\hat{\tau}^c$ the upper bound of the equilibrium set and assume that $\lambda = \phi(y_m) - \varepsilon$ with $\varepsilon \geq 0$. We distinguish two cases:

- $\varepsilon = 0$. The median voter believes that both politicians are of equal ability. As a result, the equilibrium platform will be $\hat{\tau^c} = \tau^c_m = \tau^r_m$ and from proposition 2.2 we know that $\tau^c_m < \tau^c_s$.

- $\varepsilon > 0$. As τ^c_m is the optimal tax rate of the median voter regarding the challenger, the following inequality holds

$$\mathrm{EU}^c_m(\tau^c_m) > \mathrm{EU}^c_m(\hat{\tau^c}). \tag{2.29}$$

  Also $\hat{\tau^c}$ is a winning platform for the challenger as it belongs to the equilibrium set, i.e.,

$$\mathrm{EU}^c_m(\hat{\tau^c}) > \mathrm{EU}^r_m(\tau^r_m). \tag{2.30}$$

  By combining inequalities (2.29) and (2.30) we have that

$$\mathrm{EU}^c_m(\tau^c_m) > \mathrm{EU}^c_m(\hat{\tau^c}) > \mathrm{EU}^r_m(\tau^r_m). \tag{2.31}$$

  From Proposition 2.2 we know that $\tau^c_m < \tau^c_s$. We also know that for $\varepsilon > 0$, $\tau^r_m < \tau^c_m$ by assumption 2.3. But, by continuity, as ε approaches 0, τ^r_m approaches τ^c_m and the equilibrium set shrinks with $\hat{\tau^c}$ approaching also τ^c_m.

To see this notice that:

$$\lim_{\varepsilon \to 0}\{(1-\tau_m)\, y_m + (\phi(y_m) - \varepsilon)\, W(\tau_m\, \overline{y})\} = (1-\tau_m)\, y_m + \phi(y_m)\, W(\tau_m\, \overline{y}).$$

Therefore for small values of ε, $\hat{\tau^c} < \tau_s^c$, i.e., the upper bound of the equilibrium set is less than the optimal platform of the social planner. The continuity argument refers here to a class of ϕ functions (which must be continuous, monotone, convex, and greater than λ), and a specific position for all of them (distance from λ of their value at y_m). A vanishing distance of $\phi(y_m)$ from λ at the limit just requires that the function becomes arbitrarily flat at $y_i \in [0, y_m]$ or $y_i \in [y_m, 1]$.

□

2.6 Conclusion

We presented a model of asymmetric electoral competition with heterogeneous beliefs about the ability of two politicians (an incumbent and a challenger) to provide public goods, as a result of an implicit economic shock in voters' perceptions. We showed that the challenger politician wins the election securing the majority of votes by proposing a taxation platform that coincides with the optimal point of the median voter, and this platform is robust to any possible mixing by the ruler. Moreover, we discussed the welfare implications of our model, showing that the welfare result can contradict the standard over-provision result of the median voter models. In other words, the equilibrium platform can be less than the utilitarian social optimum, i.e., competition can lead to over-austerity, even with a rightly-skewed income distribution, and the result is robust in the potential existence of multiple equilibria. These welfare implications rely on the convexity or concavity of the belief function of the voters regarding the ability of the challenger to provide public goods after the shock. We argued that if the median voter does not trust the challenger politician more than the population on average, the challenger wins the election by providing less public good than the social optimum. More precisely,

if after the shock the middle class believes that the ruler and the challenger are more or less equally efficient to govern, but other parts of the population become quite pessimistic about the ruler, then this will result to over-austerity. Though, if the median voter becomes really optimistic about the challenger compared to the average belief of the society, then the winning competition platform leads to over-provision of public goods.

An interesting point that stems from our discussion is that the possibility of over-austerity exists because all voters agree that the challenger is more competent to govern after the shock compared to the ruler. However, imagine a standard economic decline during normal times. In most cases we do not observe severe political changes, as small in magnitude downturns do not tend to trigger people to lose their faith to the incumbent politician. In such cases the ruler remains in office providing an amount of public good that is either efficient or higher than the social optimum, conditional on the shape of the income distribution. On the other hand, if we think about the case of an economic crisis that may make people to trust an outsider politician more than the incumbent, then our analysis shows that the challenger can overturn the incumbent due to voters' beliefs. Most importantly, when voters believe that he is better than the ruler, he chooses a taxation platform that can be lower than the one chosen by the social planner, which can potentially depict the austerity imposed in some countries during the last economic crisis. To conclude, we believe that our framework can incorporate interesting real life implications. Our planned future work includes the analysis of the dynamic nature of the shocks. This can lead to interesting results about the severity and the timing needed for a crisis to overturn the ruler politician and the different effects that different types of crises can have to the beliefs of the population.

Chapter 3

Identity and Corruption: A Laboratory Experiment

3.1 Introduction

One of the primary goals of elections is to hold politicians accountable for their past actions. Politicians should be less likely to engage in corrupt behaviour if they know they will be electorally punished for any wrong-doing (Ferejohn (1986)). However, corruption is still rife in democracies, whose average Corruption Perception Index is a measly 55 in a scale from 0 (highly corrupt) to 100 (very clean) (Transparency International (2019)). This is worrisome given the well-documented deleterious effect of corruption on economic growth, social welfare, and redistribution (e.g. Mauro (1995), Méon and Sekkat (2005), and Gründler and Potrafke (2019)).

A reason for the prevalence of corruption seems to be the frequent reluctance of citizens to vote out corrupt politicians.[1] This reduces the incentives of incumbents to behave honestly. Rundquist, Strom, and Peters (1977) argue that voters may not vote against corrupt politicians because they engage in an implicit trade-off between corruption and policies. Fernández-Vázquez, Barberá, and Rivero (2016) provide empirical evidence in favour of this exchange argument between voters and politicians using data from Spanish local elections. They also find evidence that the trading can be quite explicit, in the sense of voters expecting to receive direct side benefits from corruption.

An alternative explanation is in-group loyalty. Voters might be willing to turn a blind eye when the corrupt official belongs to their same ethnic group or share

[1] Evidence exists for Europe (Bågenholm (2013)), Japan (Reed et al. (1996)), and the United States (Welch and Hibbing (1997)). For an overview, see Golden (2010).

their ideology. Anduiza, Gallego, and Muñoz (2013) show that this partisanship effect is a very important factor in voters' tolerance towards corruption. Using a survey experiment ran in Spain, these authors find that voters tend to consider corruption scandals to be less serious, when the politician involved belongs to the political party they favour. Another paper providing evidence on this partisan bias is by Eggers (2014), which documents a trade-off between voters' punishment of corrupt politicians' and their willingness to see their supported political party in office. Using data from the British expenses scandal, the author shows that electoral punishment of politicians involved in the scandal was weaker in areas where the ideological distance between the two runner-up candidates was larger. This phenomenon helps to explain why candidates in more ideologically polarized constituencies were more likely to be involved in the scandal in the first place.

The role of identity in individual behaviour has been highlighted in the economic literature only recently. Akerlof and Kranton (2000) model the effect of identity ("a person's sense of self") in economic outcomes by incorporating it in agents' utility. The experimental evidence on the significant role of identity in individual decision is growing rapidly. For example, Chen and Li (2009) provided evidence of the role of identity in social preferences. Klor and Shayo (2010) highlighted that group identity affects agents' preferences for redistribution biasing them in favour of their own group. Cornaglia, Drouvelis, and Masella (2019) showed that individuals display more competitiveness towards in-group members.

The main goal of this chapter is to study whether social identity affects the tolerance of voters towards corruption. We first build up a two-period model of electoral accountability with pure moral hazard. Politicians have the same level of competence but their actions when in office are unobservable to voters. The incumbent must decide whether to invest funds in a public project with stochastic returns or to keep the funds for themselves. Voters only observe whether the project is successful or not. Upon observing a failed project, they do not know whether this was the result of bad luck or corruption. Then, voters must decide whether to approve the incumbent or not.

We take this model to the lab to investigate the role of identity in voters' decision to retain a possibly corrupted politician. We run two treatments using a between-subject design. In the control sessions, participants played in pairs under strangers re-matching for 36 independent rounds, half in the role of citizens and half in the role of representatives. In the identity treatment subjects were assigned into a group before participating in the main game. Because we employed the minimal group paradigm (Tajfel et al. (1971), Tajfel et al. (1979)), these identities were constructed to be orthogonal to personal characteristics which may in turn affect choices. Participants in this treatment played the same game as in the control treatment with the only difference being that they were informed about the identity of their representative when playing as citizens.

Our first result is that, after observing that the project is unsuccessful, citizens approve a representative around 7% less often when they do not belong to the same group compared to when they belong. Our second result shows that citizens belonging to the same identity group as the representative approve them about 11% more often compared to the control treatment conditional on the public project being unsuccessful. Our third result is that representatives behave more honestly compared to the theoretical prediction. By being honest they earn around 28.3% less than the expected earnings of a dishonest representative.

The observed differences in approval rates may be due to two reasons. One is pure social preferences. As Chen and Li (2009) showed, people tend to be more generous towards in-group members in the dictator game. Approving a politician grants them a payoff at no direct cost to the voter. It is thus to be expected that in-group favoritism will operate through social preferences. An alternative channel is differential beliefs about honesty. If citizens expect same-identity politicians to be more honest, they might tend to reciprocate that honesty by approving them more often. To explore these mechanisms we elicited beliefs from participants about the honesty of representatives upon having observed a failed project. We observe that subjects believe that representatives of their same identity group are significantly more honest compared to representatives of the other group.

This chapter is closely related to Landa and Duell (2015) and Solaz, De Vries, and Geus (2019). Landa and Duell (2015) study the effect of identity in politicians' effort choices and voters' behaviour in a setting with hidden competence and observable effort. In a lab experiment they find that voters show in-group favouritism by ignoring both politicians' effort and competence in in-group matches. Meanwhile, politicians exert some effort when they belong to the same group as voters, regardless of their competence and even if they expect to be re-elected. On the other hand, Solaz, De Vries, and Geus (2019) study the effect of identity in corruption empirically and experimentally. In their empirical study, they take advantage of the publicity of a major corruption scandal in Spain to show how voters behaviour change after knowing that the incumbent party was corrupt. They find that voters punish the corrupt party, but partisanship eliminates this punishment. The authors also ran a lab experiment inducing artificial identities to see if this in-group effect replicates in the lab. They show that in-group loyalty persists and that voters are more likely to select politicians of their same group even when it is known for certain that they were engaged in corruption.

However, our setting differs in several ways and offers new insights on the role of identity in voting decisions. First of all, we concentrate our analysis on how politicians are disciplined rather than selected. We incorporate only moral hazard incentives by making politicians' actions unobservable and all of them equally competent. In other words, we depart from the selection problem and we concentrate on the issue of accountability. This allows us to derive conclusions regarding the effect of identity on tolerance for corruptive behaviour. Moreover, by allowing for stochastic outcomes, we can elicit voters' beliefs regarding the honesty of politicians. This allows us to disentangle whether any in-group loyalty effect operates through social preferences and/or through differential beliefs about honesty. Also, we avoid any possible reciprocity effects and we are able to capture the pure effect of identity by allowing only voters to know the identity of politicians, and not vice versa. In addition, by removing any pivotality considerations and repeated interaction effects, we avoid any possible confound in our analysis.[2]

[2]For example, Solaz, De Vries, and Geus (2019) have groups of two candidates and three

The remainder of the chapter is as follows. Section 3.2 contains the theoretical model we use to derive benchmark predictions for our experiment, while section 3.3 describes the experimental design. Section 3.4 presents the results of the experiment and section 3.5 concludes. Appendix B includes summary information of the sessions, further robustness checks of the results, a descriptive analysis of the answers to the personal characteristics questionnaire we administered to subjects, and the instructions of the experiment.

3.2 The Model

3.2.1 Benchmark

This model adapts the two-period model of electoral accountability in Persson and Tabellini (2002), where public good provision has now uncertain returns. There is one voter (she) with exogenous wealth y per period and one incumbent politician (he). There is an exogenous tax rate $\overline{\tau}$, so the tax revenue in each period is $\overline{\tau}y$. The incumbent must decide whether to pocket these revenues or to use them to fund a public project. Formally, the incumbent politician chooses in each period $t = 1, 2$ the rent he extracts $r_t = \{0, \overline{\tau}y\}$.

Given this choice, the level of public good provided at period t is given by $g_t(r_t, \Theta) = \Theta(\overline{\tau}y - r_t)$, where $\Theta = \{0, \theta\}$ is a random variable which takes each value with probability $\frac{1}{2}$. The voter does not observe the rents extracted by the incumbent at the first period, r_1, but she observes the outcome of the project. If $g_1 = \theta\overline{\tau}y$, the voter knows with certainty that the incumbent funded the project and the project succeeded, i.e. $r_1 = 0$ and $\Theta = \theta$. But if the voter observes $g_1 = 0$, she does not know whether the project failed due to randomness, i.e. $\Theta = 0$, or because the incumbent kept the funds for himself, i.e. $r_1 = \overline{\tau}y$.

The voter's expected payoff at period t is thus

$$v_t = y(1 - \overline{\tau}) + \frac{g_t(r_t, \theta)}{2}, \tag{3.1}$$

voters with the same composition through blocks of rounds whereas we have pairs of one representative and one voter re-matched in every round.

where the first term is her private consumption and the second term is the expected outcome of the project. The payoff for the incumbent politician in period $t = 1$ is $B_1 + r_1$, where B_1 are the rents from being in office in the first period. The timing of the model is summarized in the following figure:

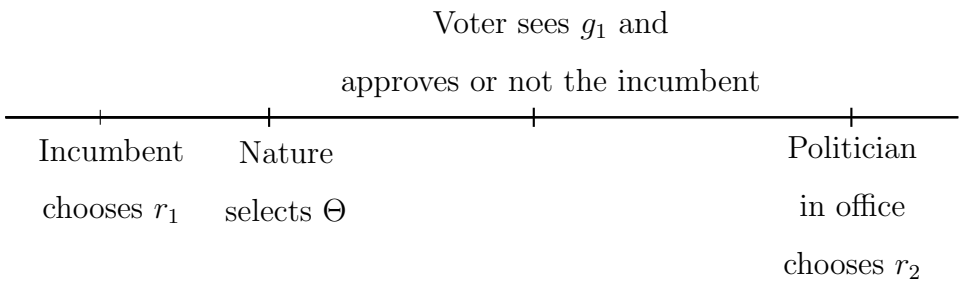


Figure 3.1: Timeline

After observing g_1, the voter must decide whether to approve the incumbent or not. If she approves him, then the incumbent has to choose the amount of rents to extract in the second period, r_2. Since the politician has no incentive to refrain from extracting rents at that point, he chooses $r_2 = \overline{\tau}y$. As a result, his payoff in the second period is $B_2 + \overline{\tau}y$, where B_2 is the rent from being in office in the second period. The voter receives $v_2 = y(1 - \overline{\tau})$. If the voter does not approve the incumbent, he steps down and he does not receive any payoff in the second period. The voter's payoff is $v_2 = y(1 - \overline{\tau})$ in that case too. This is equivalent to what the voter would receive if the incumbent were replaced by another politician who in turn would extract full rents.

Under these assumptions, the voter is indifferent between reelecting the incumbent or not. A multiplicity of equilibria emerge since the voter is indifferent between all her approval strategies. However, she is not indifferent among all equilibria. Because all approval rules are sequentially rational, punishments and rewards are credible and the voter can use them to discipline the incumbent. To be more specific, the voter's preferred outcome is that the incumbent chooses $r_1 = 0$. She can incentivize the incumbent to do this by using outcome-contingent approval strategies as follows.

We denote by a_H and a_L the probabilities with which the voter approves the

politician when the project is successful, i.e. $g_1 = \theta\overline{\tau}y$, and when the project fails, i.e. $g_1 = 0$, respectively. Given this strategy profile, the expected payoff for the incumbent politician at period 1 is

$$u_1 = B_1 + p\overline{\tau}y + \frac{1-p}{2}a_H(B_2 + \overline{\tau}y) + \frac{1+p}{2}a_L(B_2 + \overline{\tau}y), \tag{3.2}$$

where p is the probability with which the incumbent embezzles the funds. The second term in the above expression is the rents extracted in the first period, the third term is the expected value of re-election when the incumbent is honest (i.e., he sends the money) and the project succeeds, and the last term is the expected value of re-election when $g_1 = 0$, either because the politician kept the funds or because he was honest but the project failed.

The incumbent chooses p to maximize (3.2). This optimization problem dictates that the incumbent is honest in period 1 if and only if

$$a_H - a_L \geq \frac{2\overline{\tau}y}{B_2 + \overline{\tau}y}. \tag{3.3}$$

For the sake of exposition we assume that $a_H = 1$, which is a sequentially rational choice for the voter following the arguments discussed above. In this case, the best response of the incumbent politician is to fund the project in the first period if and only if

$$a_L \leq \overline{a}_L \equiv 1 - \frac{2\overline{\tau}y}{B_2 + \overline{\tau}y}. \tag{3.4}$$

In other words, the incumbent can be disciplined if the approval probability when the project fails is low enough.

Proposition 3.1. *There exists an equilibrium of the game where the politician is honest if and only if the citizen approves him with probability $a_L \leq \overline{a}_L$ when she observes $g_1 = 0$ and with probability $a_H = 1$ otherwise. In addition, there exists a continuum of equilibria where the politician is dishonest, i.e. $a_H = 1$ and $a_L > \overline{a}_L$.*

Note that for the honest equilibrium to exist, $\overline{a}_L$ must be non-negative, that is, $B_2 \geq \overline{\tau}y$; in other words, the rents from being in office in the second period must

be high enough compared to the personal rents the incumbent can extract in the first period.

3.2.2 Introducing identity

According to Tajfel and Turner (1986) the categorization of people in different groups can create in-group favouritism and out-group prejudice. Therefore, in our context voters may condition their approval strategies on the identity of the politician.

To be more precise, let us consider again the expected payoff at period $t = 1$ of an incumbent politician belonging to group $i = A, B$. Now (3.2) becomes

$$u_{1i} = B_1 + p\overline{\tau}y + \frac{1-p}{2}a_H(B_2 + \overline{\tau}y) + \frac{1+p}{2}(\zeta_i a_L^s + (1-\zeta_i)a_L^d)(B_2 + \overline{\tau}y), \quad (3.5)$$

where ζ_i is the probability with which the politician is matched with a voter of their same identity, and a_L^s and a_L^d are the voter's approval probabilities when she faces a same and a different identity politician respectively. For the sake of exposition, let us assume that voters approve politicians of either group with the same probability when the project succeeds, i.e. $a_H^d = a_H^s = a_H$. None of the results below hinge on this assumption.

Given (3.5), the best response of the incumbent is to be honest if and only if

$$\zeta_i a_L^s + (1-\zeta_i)a_L^d \leq a_H - \frac{2\overline{\tau}y}{B_2 + \overline{\tau}y}. \quad (3.6)$$

Now, assume that voters feel a "warm glow" when they do approve politicians of their same identity. Formally, voters' utility increases by a fixed amount I with $I > 0$ when they approve a politician of their same group regardless of the outcome of the project. In that case, the voter will always approve the politician of their same group and punishments are no longer a credible threat, i.e. $a_L^s = a_H = 1$. So whether the incumbent is honest or not in equilibrium depends now on the proportion of voters in each identity group.

Condition (3.6) boils down to

$$a_L^d \leq \overline{a}_L^d \equiv \frac{1}{1-\zeta_i}\left(\frac{B_2 - \overline{\tau}y}{B_2 + \overline{\tau}y} - \zeta_i\right). \tag{3.7}$$

Proposition 3.2. *There exists an equilibrium of the game where the politician is honest if and only if the citizen approves politicians of a different identity with probability $a_L^d \leq \overline{a}_L^d$ when she observes $g_1 = 0$ and approves with certainty otherwise. In addition, there exists a continuum of equilibria where the politician is dishonest, i.e. $a_H = a_L^s = 1$ and $a_L^d > \overline{a}_L^d$.*

For an equilibrium to exist in which incumbents of both identity groups are honest, $\overline{a}_L^d$ must be non-negative for both groups. Assuming a fully random matching so that ζ_i for $i = A, B$ now denotes the proportion of voters in each group, an honest equilibrium exists if and only if

$$\frac{2\overline{\tau}y}{B_2 + \overline{\tau}y} \leq \zeta_i \leq \frac{B_2 - \overline{\tau}y}{B_2 + \overline{\tau}y} \quad \text{for} \quad i = A, B. \tag{3.8}$$

In words, this means that the proportion of voters in each identity group must not be too high or too small. Otherwise, incumbents in one of the groups would be too likely to meet a voter who will approve them unconditionally, eliminating any incentive to behave honestly. Note also that a necessary condition for an honest equilibrium to exist is again that the payoff from reelection should be high enough relative to the rents the incumbent can extract in the first period, i.e. $B_2 \geq 3\overline{\tau}y$.

3.2.3 From the theory to the lab

When implemented in the lab, the model described above may generate unwanted reciprocity effects. If a voter expects that reelecting the incumbent will make him more likely to be honest in the second period -even though that would run against his narrow self-interest- it would be rational for her to approve him when the project fails. This expectation of reciprocity may be reinforced by the presence of identity: an incumbent who sees himself approved despite a project failure may update up his beliefs about the voter having the same identity as his and increase

his likelihood of being honest in $t = 2$ if in-group favouritism exists.

To rule out these effects, we implement in the lab a one period version of the game described above which produces the same theoretical predictions. In this version, the game ends after the voter's approval decision and the politician receives $B_2' = B_2 + \overline{\tau}y$ if approved. Assuming again that $a_H = 1$, the politician's expected payoff in the case without identity is

$$u_1 = B_1 + p\overline{\tau}y + \frac{1-p}{2}B_2' + \frac{1+p}{2}a_L B_2'. \tag{3.9}$$

Now $\overline{a}_L \equiv 1 - \frac{2\overline{\tau}y}{B_2'}$ so for an honest equilibrium to exist $B_2' \geq 2\overline{\tau}y$.

If we were to introduce identity in-group favouritism, the one period version would require that for an honest equilibrium to exist voters must approve incumbents in their identity group with probability no greater than $\overline{a}_L^d = \frac{1}{1-\zeta_i}(1 - \frac{2\overline{\tau}y}{B_2'} - \zeta_i)$.

Consider the particular case where $B_2' = 2\overline{\tau}y$. This is the one we implement in our experiment. In that case, we can derive the following theoretical predictions:

Prediction 1: In absence of identity, an honest equilibrium exists if and only if $(a_L, a_H) = (0, 1)$.

Prediction 2: With in-group favouritism, an honest equilibrium does not exist, i.e. $\overline{a}_L^d < 0$.

3.3 Experimental Design

Our experiment was pre-registered at "As Predicted".[3] Experimental sessions were ran at the LExEcon lab of the University of Leicester (UK) and at BEADS lab of the University of Birmingham (UK). We ran nine sessions in total; six for the identity treatment and three for the control. The number of participants in each session was 14 or 16. Overall, we recruited 142 undergraduate and postgraduate students from several departments at the two Universities. The experiment was

[3] http://aspredicted.org/blind.php?x=tu6p5e

programmed in zTree (Fischbacher (2007)).[4] The experiment was run in March and November 2019. A session lasted approximately 40 minutes. The show up fee was £4 and the average payment across all sessions was around £14.9.

Subjects were randomly assigned to one of the two treatments, identity and control. The identity treatment had an initial additional stage where identity was induced using the Minimal Group Paradigm (MGP) as detailed below. Subjects played 36 independent rounds, half as a "citizen" and half as a "representative". We chose these words to provide some context and increase the external validity of the experiment. However, we deliberately refrained from using the word "politician" to avoid any strong connotations as much as possible.

We followed the same strangers matching protocol as Woon (2012) with participants being re-matched in every round throughout the session whilst ensuring that all subjects had played in both roles by the twelfth round. At the beginning of each round each pair received 16 tokens, which serve as tax revenue $\overline{\tau}y$. Subjects acting as representative were told they were in charge of administering that amount, i.e to decide whether to "keep" it or "send" it to the citizen. If the representative decided to keep the tokens, the citizen received 0 tokens. If the representative decided to send the tokens, the citizen received either double the amount (32 tokens) or nothing (0 tokens) with equal probability, i.e. $\theta = 2$. After observing the tokens received, but not the choice of the representative, the citizen decided whether to approve the representative or not. If the citizen approved the representative, the latter received 32 tokens, i.e. $B_2' = 2\overline{\tau}y$. If the representative was not approved, the representative received nothing, i.e., $B_2' = 0$. After this decision, the round ended. Our main interest is the comparison of approval decisions between in-group and out-group matches and the control.

In the identity treatment, citizens were informed about the identity of the representative they were matched with, while representatives received no information about the identity of the citizen. This allows us to pin down the effect of identity on citizens' decisions without worrying about possible reciprocity effects created by representatives knowing citizens' identity too. Our choice of letting citizens

[4]See appendix B for session details and instructions.

but not representatives know stems also from reality. It depicts better a scenario where voters know the identities of politicians but they know only imperfectly the ethnicity or ideology composition of their constituency.

Social identity was artificially induced and assigned at the beginning of each session of the identity treatment. We followed the widely-used Minimal Group Paradigm (MGP) introduced by Tajfel et al. (1971) and Tajfel et al. (1979). This minimal categorization is enough to create in-group favouritism and out-group animosity. The methodology has the advantage that artificially induced identities should be, in principle, orthogonal to subjects' characteristics which may in turn correlate with the behaviour under study. In other words it provides more control over the identity-formation process compared to natural identities (Charness and Chen (2020)). However, according to Lane (2016), who conducted a meta-analysis of identity experiments, discrimination is higher when identity is induced artificially rather than naturally. In other words, it might be the case that the MGP paradigm does not produce a lower bound of identity effects. The reasons for this can be various, including the fact that discrimination might be socially acceptable when the groups are not correlated with natural identities, like ethnicity or religion for example.[5] Acknowledging that both ways have their pros and cons, we believe that for the purposes of our experiment it was useful to abstract from any natural identities, in order to avoid as much as possible any correlation of the identity groups with specific political views or ideologies.

A prominent way to create minimal group identities in the lab is the use of the Klee-Kandinsky protocol (e.g. Tajfel et al. (1971), Chen and Li (2009)). Paul Klee and Wassily Kandinsky were two painters who worked roughly during the same time period in the same region of Germany (they even lived in the same street) and who remain relatively unknown to the general public till date. Art historians argue that their patterns are very similar.[6] As a result, it is to be expected that any categorization based on the taste for these painters should be as good as random.

[5] See Lane (2016) for a discussion on the possible explanations.

[6] "... similarities in scale, theme, motifs and even technique suggest a dialogue between the two artists. Their lives intersected at various times between 1911 and 1937, and their art responded to each other's." (Barnett, 2015)

Specifically, our participants were shown five pairs of paintings. The images of these paintings had all the same resolution and surface area. Each pair featured one painting by each author. They were asked to choose the painting they liked most within each pair. They were then categorized as a "Klee" or as a "Kandinsky" according to the majority of their choices. Following Chen and Li (2009) and Landa and Duell (2015), we also implemented an identity enhancement task.[7] Subjects were shown two additional paintings, one by Klee and one by Kandinsky. Members of each group had to identify the painting that belonged to the painter of their group. If the majority of members within a group identified the author correctly they all received 4 additional tokens.

In the model described in the previous section, we introduced in-group favouritism by assuming that identity entered directly in the utility of voters when approving a representative of their same group. However, in-group favouritism could also operate through differential beliefs about the action of the representative upon observing that the project failed. According to the literature, voters may process information differently if they like the politicians involved in a corruption scandal (see for example Rahn (1993), Taber and Lodge (2006), and Anduiza, Gallego, and Muñoz (2013)). In order to disentangle these two channels, we elicited subjects' beliefs about the frequency of dishonest behaviour by representatives at the end of the session. Subjects were paid 4 tokens if their guess was within a band of 5% percentage points around the actual frequency.

At the end of each session, subjects completed a questionnaire on individual characteristics. It included questions on gender, age, field of study, religiosity, race/ethnicity, family income, ideology, volunteering activities, trust, and risk attitudes. In the identity sessions, subjects were also asked about their familiarity with the two painters and about the level of attachment with their group identity during the session.[8]

[7] See Charness, Rigotti, and Rustichini (2007) for the importance of the salience of artificially induced identities.

[8] A descriptive analysis of the answers to the questionnaire can be found in appendix B.

3.4 Results

3.4.1 Identity inducement

According to the MGP, a categorization using artificially induced identities should be random and orthogonal to any personal characteristics. This also means that subjects should be split into two groups of roughly the same size. However, participants in our sample divided into groups of quite unequal size. Subjects who preferred Kandinsky over Klee were 73.4% of the sample. Note however that this does not alter the theoretical predictions derived in section 3.2.3 as $B_2' = 2\overline{\tau}y$ implies an honest equilibrium does not exist regardless of the proportion of subjects in each identity group ζ_i for $i = A, B$.

More worrisome perhaps is that we find evidence that identity correlates with some individual characteristics. The regressions in Table 3.1 report those individual characteristics that subjects reported in the questionnaire which were significantly associated with their choice of identity. Column (1) is a linear probability model on the likelihood of a subject being a "Kandinsky", while specifications (2) and (3) report results of a probit model.

STEM is a dummy variable that takes the value 1 if participant's degree is in Maths, Physics, Engineering, Medicine, Chemistry, or Biology. Ideology is another dummy taking the value 1 if the subject reports that their ideology is between 4 and 7 on a Likert ideology scale between 1 (extreme left) and 7 (extreme right).

Estimates from these regressions yield a very significant association between a taste for the paintings by Wassily Kandinsky and being male, younger, leftist, and studying a non-STEM degree. The reasons behind these associations are of no interest to us but they mean that we must include these controls in all our main regressions. Otherwise, the correlation between personal characteristics and the "treatment" received might bias our estimates.

Kandinsky	(1) LPM	(2) Probit	(3) margins (dy/dx)
Female	-0.176*	-0.587**	-0.167**
	(0.072)	(0.268)	(0.070)
$\geq$ 21 years	-0.221***	-0.698***	-0.199***
	(0.051)	(0.161)	(0.046)
STEM Studies	-0.190***	-0.597***	-0.170***
	(0.042)	(0.191)	(0.039)
Centre-Right	-0.159**	-0.442**	-0.126**
	(0.049)	(0.171)	(0.052)
Observations	94	94	94

Robust errors clustered by session. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.1: Associations between personal characteristics and identity

3.4.2 Approval decisions

In this section we discuss the approval rates according to the identity of the players in a matching. Figure 3.2 shows the approval rates by identity matches (control, different, and same) and by tokens received by the citizen. When citizens receive zero tokens, they approve more often representatives sharing their identity, and slightly less often the ones with a different identity.

Also, we observe that citizens tend to approve less often representatives with a different identity compared to the control even when the project is successful and they can be sure the representative was honest. However this difference is not statistically significant.[9]

[9]Table B.5 in the appendix provides the output of the relevant regressions.

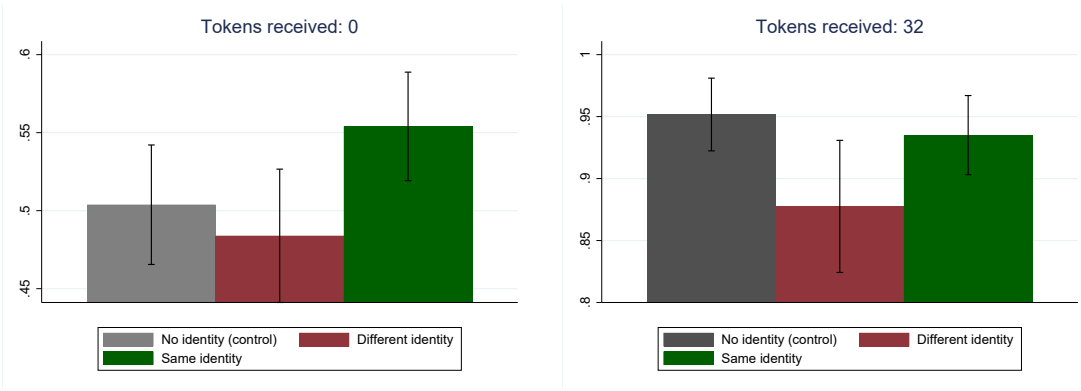


Table 3.2: Average approval by tokens received, treatment, and matchings

Focusing on the difference on the approval decisions when the project failed, we run a linear probability model on the choice of the citizens to the different treatments and matches.[10] Table 3.3 shows the result of the comparison between same and different matches, and between same matches and control.[11] We use random effects panel models with robust errors clustered by session to account for session heterogeneity.

Specification (1) includes the main independent variables, a control for period, and a dummy for the location of the session (Birmingham or Leicester). Specification (2) adds the field of study as an additional control, and if the participant started the game by playing as representative. Finally, specification (3) includes all individual characteristics recorded in the questionnaire that correlate with the choice of identity. We see that citizens sharing their identity with their representatives approve them around 11.1% more often compared to the control treatment with the coefficient being statistically significant in both (1) and (2). The effect becomes slightly smaller in column (3) but it remains statistically significant at a different level.

[10]Probit models are presented in the appendix for robustness.

[11]Table B.6 in the appendix replicates the same specification but using the comparison between control sessions and different identity matches. We do not see any statistically significant difference in citizens' behaviour between the control sessions and the out-group matches in the identity sessions.

Approval choice (0 tokens)	(1)	(2)	(3)
Same Identity vs. Control	0.111**	0.111**	0.098*
	(0.052)	(0.051)	(0.051)
Different vs. Same Identity	-0.067**	-0.069**	-0.068**
	(0.029)	(0.029)	(0.029)
Round	-0.002*	-0.002*	-0.002*
	(0.001)	(0.001)	(0.001)
Birmingham	0.324***	0.323***	0.331***
	(0.033)	(0.032)	(0.035)
Start as Representative		0.056	0.044
		(0.060)	(0.071)
STEM Studies		0.088***	0.089***
		(0.025)	(0.022)
Centre-Right Ideology			-0.031
			(0.062)
$\geq$ 21 years			0.046
			(0.054)
Female			-0.045
			(0.047)
Observations	1,971	1,971	1,971

Robust errors clustered by session. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.3: Panel Linear Probability Models on the decision of the citizen (project failed)

Moreover, we see that citizens who have different identity than the representative approve them almost 6.8% less often compared to the same identity matches.

Notice that this coefficient varies little across specifications, in magnitude and in significance. We also observe that the field of study has a positive and significant coefficient. It seems that students in STEM degrees approve politicians more often compared to the rest of participants. Interestingly, the Birmingham dummy is highly statistically significant showing that subjects in the Birmingham sessions approved on average more often than the ones in Leicester. Even if we do not control for this effect the differential coefficient remains statistically significant at 5%.

The findings so far can be summarized as follows:

Result 1 (Approval rates - Same vs Different)**.** *After receiving 0 tokens, citizens approve less often a representative when they do not share identity compared to when they share identity.*

Result 2 (Approval rates - No identity (control) vs Same)**.** *After receiving 0 tokens, citizens approve more often a representative when they know that they share identity compared to when they do not know their representative's identity.*

3.4.3 Sending rates

Control treatment

According to the theoretical predictions, if the difference in approval probabilities a_H and a_L is large enough, representatives should be honest and fund the project (Proposition 3.1). In particular, Prediction 1 states that under the values used in our experiment, an honest equilibrium exist if and only if citizens always approve the politician when the project succeeds and never if the project fails. For any other approval rates, the representative should pocket the funds.

Figure 3.2 shows that the observed approval rates do not match the values required to incentivize the representatives to behave honestly. Citizens approve far too often than they should after observing that they received no tokens, and less than they should after receiving the tokens.

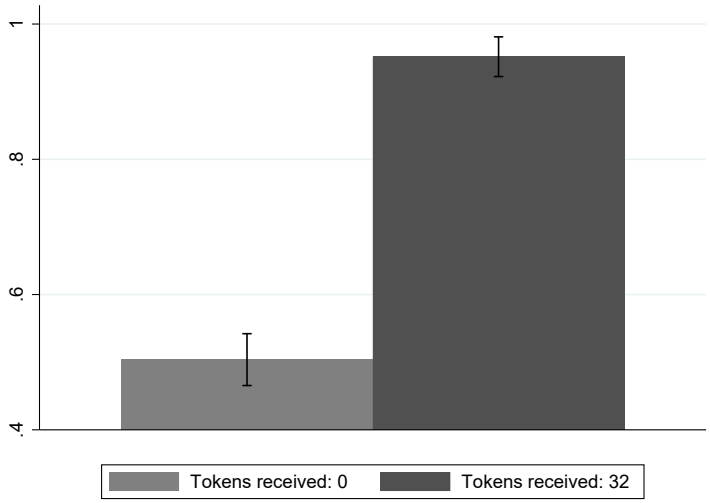


Figure 3.2: Average approval by tokens received in the control sessions

More specifically, the approval rate when they see 0 tokens is 50.38%, whereas it is 95.17% when they receive the 32 tokens. Even though 95.17% seems quite close to 100%, after performing a proportion test, we see that it is actually statistically different than full approval (p-value=0.00).[12] Given these approval rates, representatives should always behave dishonestly. However, representatives' sending rate was 45.37% across the control sessions. This is a remarkable result which shows a clear preference for honesty:

Result 3 (Sending rates - Honesty)**.** *Given citizens' actual approval rates, representatives behave honestly more often than predicted.*

Given the observed approval rates, representatives who sent the tokens were losing an average of 8.46 tokens, which represents a 29.9% of subjects' average earnings.[13]

Identity treatment

Prediction 2 states that, under the parameters we implement in the experiment, an honest equilibrium cannot exist if citizens show in-group favouritism and regardless

[12] As we cannot run a proportion test if the proportion lies on the boundaries of the set $(0, 1)$, we performed the test for 99.99%, used as an approximation.

[13] Section 3.4.3 describes the corresponding calculations taking into account all sessions.

of the identity composition of the set of potential voters. As a result, in the identity treatment we should observe lower sending rates than in the control. However, Figure 3.3 shows that the sending rates in the two treatments are almost identical, i.e. representatives seem to play the same strategy regardless of the presence of identity. It would seem that representatives do not actually expect that in-group/out-group matches change citizens' behaviour.

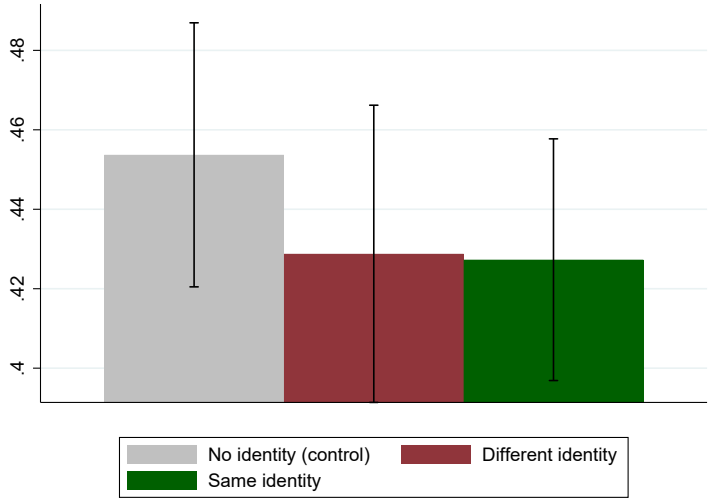


Figure 3.3: Average sending by treatment and group matching

More specifically, in the identity treatment representatives keep the money 57.21% of the occasions, whilst in 54.63% of the rounds in the control. These proportions are not significantly different at any conventional level with the p-value of the two-sided proportion test being 0.11. Moreover, recall that in the identity treatment, citizens know the identity of the representatives they are matched with, but not vice versa. As a result, we expected no difference between the different and same matches, which verifies also the internal validity of our experiment.

Table 3.4 below presents the regression results showing that there is indeed no statistically significant difference in the average sending neither across treatments nor between same and different matches.

Sending choice	(1)	(2)
Same Identity vs. Control	-0.037	-0.033
	(0.042)	(0.045)
Different vs. Same Identity	-0.009	-0.008
	(0.011)	(0.011)
Round	-0.003**	-0.003**
	(0.001)	(0.001)
Birmingham	-0.081	-0.081
	(0.053)	(0.057)
Start as Representative	0.071**	0.060**
	(0.031)	(0.031)
STEM Studies		-0.023
		(0.064)
Centre-Right Ideology		-0.088*
		(0.049)
$\geq$ 21 years		-0.020
		(0.043)
Female		0.078
		(0.060)
Observations	2,556	2,556

Robust errors clustered by session. * $p < 0.1$, ** $p < 0.05$

Table 3.4: Panel Linear Probability Models on the decision of the representative

The first specification in Table 3.4 includes the main independent variables controlling for the period, the sessions ran at the experimental lab in Birmingham,

and the effect of playing in the first 6 rounds as a representative.[14] In specification (2) we add those individual characteristics that played a role in the identity choice. We see that in both cases there is no difference in the choice of representatives among treatments. To be more specific, in both specifications the coefficient of the same identity matches compared to the base, which is the control, is negative, but far from significant. Moreover, the coefficient of the differential between the different and the same matches is again not statistically significant. There seems to exist some learning, as the more periods subjects play the more they keep the money, getting closer to the theoretical prediction. Also, we see that subjects who played as representatives in the first 6 rounds tend to send more money compared to those starting as citizens. This may be due to the fact that, once subjects play as citizens, they update their beliefs about the approval rates and keep the tokens more often when they later play as representatives. Students' field of study does not play a role in their decision when they play as representatives, but their ideology do. Participants who self-report to be conservative are more likely to keep the tokens. However, including these individual characteristics does not affect our main result about representatives' behaviour:

Result 4 (Sending rates)**.** *There is no difference in representatives' sending behaviour across treatments and matches.*

Representatives' earnings

We see that in both treatments, conditional on the actual approval rates of the citizens, representatives forgo earnings behaving honestly. That would suggest that representatives suffer a cost of dishonesty. In order to measure this cost, we next calculate the loss in representatives earnings due to honesty by using the observed approval frequencies.

[14] As before, the appendix includes the corresponding Probit models for robustness.

Rep's choice	**Citizen's choice**		
	Not Approved (0)	Approved (32)	Total
Keep (16)	691	749	1,440
Send (0)	301	815	1,116
Total	992	1,564	2,556

Table 3.5: Representatives' choices by citizens' decisions

Table 3.5 presents the frequencies of the four possible combinations of representatives' and citizens' choices (keep/send by approve/not approve) including all sessions (recall representatives were not aware of the citizen's identity in the Identity treatment). The numbers in parentheses are the amount of tokens a representative earns with that choice; their combination yields his final earning. For instance, if a representative kept the funds and the citizen approved him, the representative earned $16 + 32 = 48$ tokens. Using this information, we can calculate the expected earnings of a representative who keeps or sends the money. If a representative kept the money (16 tokens), he received on average $(1,440 \times 16 + 749 \times 32)/(1,440) = 32.64$ tokens, whereas he received $(815 \times 32)/(1,116) = 23.4$ tokens in expectation if he sent the funds. In other words, an honest representative forgoes 9.24 tokens or £4.62, which represents 28.3% less than the expected earning of a dishonest representative.

3.4.4 Beliefs

One remaining question is why citizens, upon observing a failed project, approve politicians of their same identity significantly more often than politicians of different identity. There are two potential reasons for this. The first one is that participants' in group favouritism operates through social preferences (Chen and Li (2009)). Approving the representative awards him a positive payoff, so if citizens value the payoff of their fellow group members, they will be more inclined to approve in-group representatives. On the other hand, it may just be that par-

ticipants believe that in-group representatives are more honest and approve them more frequently to reward their honesty. In order to check whether the second channel might have been relevant, we elicited beliefs about honesty rates by asking participants the following question at the end of the experiment:

> "From all the rounds in which you played as a citizen, the representative was a Klee (Kandinsky), and you received 0 tokens, what is the percentage of occasions in which you believe the representative kept the tokens for themself?"

This question was incentivized. Participants earned 4 additional tokens if their answers were in a 5 percentage points band around the true percentage of dishonest representatives.

Among sessions, we randomised the identity group of the representatives whose honesty we asked participants about. That meant that a fraction of subjects answered the question for politicians of their same identity whereas the rest did for out-group politicians. With this we forwent some data but we avoided any experimenter demand effect that could have arose from asking about the honesty of both types of representatives.

Recall that representatives did not know the citizen's identity and this was public knowledge. Hence, this question should elicit pure priors about the (dis)honesty of same and different identity participants.

Table 3.6 below presents descriptive statistics for these elicited beliefs. It also includes the actual average value of the representatives' decisions when the outcome was 0 by treatment and matches.

	Obs.	Average Beliefs (%)	Average Dishonesty (%)	Belief ≠ Dishonesty (t-test)
Same	47	58.08	74.08	0.0000
Different	47	65.72	73.05	0.0387
Control	48	61.44	71.84	0.0021

Table 3.6: Beliefs vs actual dishonesty rates when project fails

The first observation from this table is that participants believed that representatives were significantly more honest than they actually were on average across treatments and matches. From the p-values of the t-tests we conclude that in all three cases we can reject the null hypothesis that average beliefs and the actual average decision of the representatives are the same. This helps to explain the high approval rates we observe when citizens received zero tokens compared to the equilibrium prediction.

Participants overestimated the honesty of representatives to the largest extent in the same identity matches. Actual dishonesty rates were very similar across matches but citizens trusted representatives of their same identity significantly more compared to representatives with different identity. In order to see if this effect is significant we run a number of regressions and present the results in Table 3.7. In all three specifications the dependent variable is the elicited beliefs measured in percentage. We see that the coefficient of the differential change between the average beliefs for a same identity representative versus a different identity one is statistically significant at 10% after controlling for ideology, gender, and family income.

In specification (3) we add three variables to control for participants' characteristics and possible order effects; the percentage of cases in which the individual sent the tokens as representative (*Average Honesty*), the percentage of occasions in which the participant approved the representative after receiving zero tokens (*Average Approval*) and a dummy taking the value of one when the participant started as representative. The relationship between individual honesty and their belief on the dishonesty of representatives is highly significant and negative. That is, the more honest a subject was when playing as representative the more honest they believed other representatives were too.[15] There is no evidence that participants who approve more often or started as representatives had different beliefs.

[15]Because the choice of identity depends on personal characteristics, as shown in table 3.1, we ran some additional regressions to see if politicians of one identity are on average more honest than politicians of the other identity. Table B.11 in the appendix shows that there is no statistically significant difference in honesty levels between the two different identities.

Dishonesty Beliefs (%)	(1)	(2)	(3)
Different Identity vs. Control	4.566	3.502	2.838
	(3.936)	(4.130)	(2.791)
Same vs. Different Identity	-6.875	-7.581*	-9.049**
	(3.976)	(3.753)	(3.473)
Birmingham	3.585	3.266	-0.495
	(4.645)	(4.663)	(3.416)
Centre-Right Ideology		-4.260	-8.074**
		(3.463)	(2.646)
Female		-4.485	-1.084
		(3.926)	(4.716)
Average Income		7.653*	3.519
		(3.822)	(3.985)
Average Honesty (%)			-0.366***
			(0.059)
Average Approval (%)			0.023
			(0.068)
Start as Representative			-3.641
			(2.243)
Observations	142	142	142

Robust errors clustered by session. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.7: Regressions on elicited beliefs

To sum up, elicited beliefs suggest that the differences we observe in approval rates are at least partially due to different priors about honesty. Upon observing a

failed project, participants believe that the representatives of their same identity are more honest than the representatives from the other identity group. Citizens might have wanted to reward representatives they believed were honest by approving them. Because they expected representatives of their same identity to be more honest, citizens approved them more often. The result below summarizes the analysis.

Result 5 (Beliefs)**.** *Citizens who share identity with the representative believe that it is less likely that the representative kept the money when they received 0 tokens compared to citizens who do not share identity with the representative.*

3.5 Conclusion

There is substantial evidence showing that voters are often reluctant to remove corrupted politician from office. The evidence also suggests that in-group favouritism is one important driver of this phenomenon. In this chapter, we brought to the lab a pure moral-hazard model of electoral accountability with stochastic public good provision. We found that social identity is an important factor in voters' reluctance to vote out possibly corrupt politicians. Specifically, voters, who observe an unsuccessful project tended to approve more often politicians who belonged to their same identity group than those who did not or when identity was absent.

This result is in line with the empirical evidence that partisanship plays an important role in corruption. We corroborate that in-group favouritism operates independently of ideological or reciprocity (quid pro quo) considerations. We have shown that, even when identity is artificially induced, citizens tend to believe politicians of their same identity are relatively more honest and retain them in office more often.

Chapter 4

Welfare Implications of Median Voter Models

4.1 Introduction

The median voter model of electoral competition has been used extensively in the literature to describe the way politicians are elected in a democratic society. The standard model, influenced by the initial works of Hotelling (1929), Black (1948), and Downs (1957), predicts that in an election, where two office-motivated politicians compete by proposing policy platforms, the median voter is pivotal. Therefore, the result of the competition is the optimal policy of the median voter and the two politicians tie in the election by proposing this policy.[1]

In a democracy elected politicians are responsible for the amount of taxation and the provision of public goods and services. As a result, the median voter model offers the workhorse environment to understand how political competition works and what the implications for the size (and composition) of the government are. Meltzer and Richard (1981) were one of the first to use an electoral competition model in order to describe the democratic choice of taxation and redistribution.[2] The model assumes linear taxation and the existence of a lump-sum transfer financed by public budget. Voters' preferences are decomposed in consumption and leisure and all voters benefit equally from public expenditure. Applying the median voter theorem in this framework, the equilibrium result is the optimal tax rate of the median income voter.

[1] It is worth mentioning here that such an equilibrium does not exist under the assumption of linear transportation costs (see D'Aspremont, Gabszewicz, and Thisse (1979)).

[2] They adapted and extended the model proposed by Romer (1975) and Roberts (1977).

Subsequently, the basic framework of this model has been used widely to discuss the optimal choice of public expenditure in a democracy. In most of the cases, however, there is an additional assumption on the quasi-linearity of preferences between private and public good consumption, as it simplifies significantly the analysis on electoral competition.[3] Taking into account the welfare implications of this model, we see that they seem to be quite rigid. Let us consider a utilitarian social planner who maximizes the average utility of the society. Under the assumptions of quasi-linearity and equal benefits from public policy for all voters, the utilitarian benchmark reflects the preferences of the average income voter. Therefore, comparing the equilibrium result with the first best outcome, the welfare implications of the model depend only on the relative position of the median compared to the average voter. If the median and average income voters are identical, i.e. the income distribution is symmetric, then the equilibrium result is always efficient, as it coincides with the social optimum. However, if the median voter is poorer than the average voter (assuming an empirically relevant income distribution), competition always over-provides public goods. The reason is straightforward; the benefits from taxation are the same for all voters, but the median voter faces a lower opportunity cost, in terms of withdrawn private consumption, compared to the average voter.

Similar forces drive also the implications of the model regarding the relationship between redistribution and inequality. If we define inequality as the difference between median and average income, then the model predicts that the higher the inequality in society, the higher the redistribution. The mechanism behind this result is based again on the fact that the median voter is pivotal and all voters receive equally symmetric public benefit. Therefore, the lower the opportunity cost of the median voter, the higher the optimal median tax rate. However, the empirical evidence is inconclusive about this relationship. Some studies verify the positive relationship between inequality and redistribution, like Meltzer and Richard (1983) or Milanovic (2000), while others are indecisive about the effect,

[3]For a summary and different representations of this model see Persson and Tabellini (2002) and Acemoglu and Robinson (2005).

such as Persson and Tabellini (1994). More recent studies actually find that income inequality and redistribution are negatively related (see Gouveia and Masia (1998) and Razin, Sadka, and Swagel (2002) among others).

As a result, interesting questions arise. How will the welfare implications of the model change if the benefits received from public expenditure are not the same for everyone and people perceive them differently according to their income? Also, what if we remove the assumption of quasi-linearity in private consumption, and what is the role of distortive taxation in this framework? The assumption of equal benefits from public good provision for every person in a society is quite restrictive. There are many cases, where we expect poor and rich people to enjoy public provision disproportionately. Recall for an example the model introduced in the first chapter of the thesis. People can disagree on the perceived ability of a politician to govern after a shock in the economy. It might be the case that poorer voters trust the politician more, while rich voters less, or vice versa. This heterogeneity is reflected in their different perceived benefits from public expenditure. Moreover, we can think about the case of different public goods. A rich voter might gain less from the provision of public transport, as they have the option to use their own vehicle. On the other hand, they might benefit more from other types of public goods, like policing for an example, compared to the poorer section of the society.

In order to address these points, we propose a model of linear taxation and public good provision, where voters, apart from being heterogeneous in income, have also different valuations of the public good. We abstract from the assumption of quasi-linearity and we allow for taxation to distort income. We discuss the important welfare implications that the model can bring to the existing literature. More specifically, we elaborate on how under certain conditions of voters' different valuations, the result of the competition might be inefficient even with a symmetric income distribution, and how even with a rightly-skewed distribution, competition might end up under-providing public goods.

The rest of the chapter is as follows. Section 4.2 provides an overview of the

literature. Section 4.3 describes the primitives of the model, the equilibrium result, and the welfare benchmark. Section 4.4 discusses the different channels and the welfare implications of the model, while section 4.5 concludes.

4.2 Literature Review

Several papers in the literature have identified cases of under-provision of public goods, incorporating different considerations. This section provides a discussion of the relevant literature.

Wright (1986) analyzes the provision of unemployment insurance in a dynamic stochastic framework. The main focus of the paper is to abstract from any incentive effects of unemployment insurance and characterize the median voter equilibrium in a society, where voters are heterogeneous in terms of their stochastic employment opportunities or/and their current conditions. In other words a society, where agents have different benefits from unemployment insurance conditional on their status. Because of the dynamic nature of the model, even if voters are intrinsically identical, if the median voter is currently employed, the political equilibrium under-provides unemployment insurance compared to the optimal benchmark, which requires complete insurance. The reason is that the presence of a discount rate makes the unemployment insurance benefits of a current employed individual to be less than the costs that they incur by financing the insurance of the currently unemployed.

Another case where under-provision of public goods has been discussed in the literature is in the provision of local public goods (see for example Coughlin and Nitzan (1981)). If we assume a scenario of full decentralization and no externalities, where each district could pay only the taxes to finance the public good that its voters demand, the welfare implications of the competition would be the same as the ones of the standard model. However, in a centralized system the different public goods are financed by a common pool of taxes. It follows naturally that the voters of a specific district prefer over-provision of their own good and under-provision of public goods preferred by other districts, as in this case they incur

the cost (taxation) but not the benefits of the public good.[4] In other words these models consider a local redistribution of the benefits from centralized taxation. This effect works in the opposite direction of the standard framework, where there is redistribution of taxation of a common public expenditure pool. As a result, in this environment, we should expect under-provision of public goods in some districts depending mainly on the political power of each of them.

A different part of the literature has added additional channels in how voters' preferences may differ and how this might affect the amount of public good provision. Roemer (1998) analyzes the political equilibrium outcome in a model with voters, who differ, apart from income, also in another characteristic orthogonal to income; let's say religion. In this case there is a trade-off between competition on religious issues and competition for redistribution. As a result if the average income of the group of voters with median religious ideas is higher than the average income of the population, then the equilibrium amount of redistribution might be lower (compared to the case without the second dimension) even if poorer voters prefer higher taxation. More recently, Austen-Smith and Wallerstein (2006) consider the case of individuals of different race. Even though the paper has a more general contribution of studying the effects of affirmative action, the authors also show that if there is redistribution by race, then redistribution by income can be reduced.

Other studies have analysed the role of different institutions in the provision of public goods. For an example, Persson and Tabellini (1999) and Persson and Tabellini (2002) discuss the differences between the effects of majoritarian versus proportional elections and of presidential versus parliamentary ones. They show that in majoritarian and presidential regimes the provision of public goods is less compared to the other institutions. Moreover, Lizzeri and Persico (2001) consider a model of redistribution and public good provision, where politicians can use cash redistribution to target transfers to subsets of voters or local projects (pork barrel spending). Therefore, politicians have incentives to under-provide public goods, as

[4]Lockwood (2002) and Besley and Coate (2003) analyze the trade-off between centralization and decentralization in this framework.

there is a trade-off between efficiency and targetability, and public goods benefits are not easily targeted to specific voters. The authors compare different institutions, namely a winner-take-all-system (office spoils received only by the winner), a proportional system (spoils divided according to vote share), and the electoral college system. They show that, if the public good is desirable, the first provides less often public goods compared to the second, and vice versa. The last is always less efficient compared to the rest. It is important to mention that in the model voters are assumed to be identical. However, the authors comment that even if they include some type of voter heterogeneity, such as heterogeneous valuations for the public good, the main conclusions remain the same. The only additional implication is that public good becomes more likely to be provided, which contradicts the standard result of the median voter models, in which competition becomes efficient when the income distribution is symmetric. Furthermore, Levy (2004) studies a model of electoral competition with party formation, concluding that in a multidimensional space there might be less taxation and less redistribution compared to a system without parties.

As discussed also in the introduction, an increased part of the literature considers the inequality implications of these models. Following the contradictory empirical results, several papers provide different theoretical explanations of a possible negative relationship between inequality and redistribution. For an example, Benabou (2000) discusses the role of capital market imperfection, where Levy and Razin (2015) study the implications of positive income sorting.[5]

Other papers have incorporated other regarding preferences in voters' utility. One of the most prominent papers in this literature is by Alesina and Angeletos (2005), who study a median voter model of redistribution incorporating fairness considerations. They make a distinction between justifiable inequality (a result of bad luck or lack of connections) and non-justifiable inequality (lack of effort), introducing tax distortions in the model.[6] In the presence of fairness the model

[5]For an empirical investigation on the role of capital market imperfections see De Mello and Tiongson (2006).

[6]Galasso (2003) includes also fairness in the standard redistribution model, but without this distinction, and he shows that the standard result is actually strengthened.

has multiple equilibria; if voters anticipate high taxes, they will exert low effort, which means that inequality is due to bad luck and it is optimal ex-post to receive high redistribution matching voters' expectations. However, if they anticipate low taxes, they will exert high effort, which means that inequality is due to low effort or talent and it is optimal ex post to have low redistribution. Their model explains the existence of both high inequality-high redistribution societies (e.g. Europe) and high inequality-low redistribution (e.g. US) at the same time.

The latest paper in this literature is by Dotti (2020). The author extended the model presented by Meltzer and Richard (1981) including both a lump sum transfer and the provision of public goods. Assuming a multidimensional policy space, where voters can choose how many and which public goods they prefer, higher inequality might decrease the size of the government, but it might be positively correlated with the progressivity of the tax system. In other words, the composition of public spending plays a role in the decision about the amount of public expenditure. Because the government's budget is split between the transfer and the public good, if the public good is a normal good, higher inequality implies a lower preferred level of public good provision. The author refers also briefly to the welfare implications of the model, concluding that the equilibrium tax rate might be lower than the social optimal under the assumption of a social planner, who is inequality averse.

4.3 The Model

We consider a model of electoral competition over alternative taxation platforms in view of the different benefits voters expect to derive from aggregate public expenditure. Taxation is proportional and the tax rate is denoted by $\tau \in [0, 1]$. We assume two office-motivated political candidates and a unit mass of voters, i, distributed on the interval $[0, 1]$ according to the $pdf\, f(i)$.

Voters differ in their pre-tax (gross) income, $y_i(\tau)$, which is a continuous and differentiable function of the voter's identifier index, $i \in [0, 1]$, and of the propor-

tional tax rate, $\tau \in [0,1]$, charged by the politician in power, i.e.

$$y_i(\tau) = y(i,\tau) \text{ with } i \in [0,1] \text{ and } \tau \in [0,1].$$

We denote by g the aggregate public expenditure to supply public goods and services to the voters, and with T the total tax revenue. The level of g is automatically set by the government's balanced budget constraint:

$$g = T,$$

where

$$T = \tau \int_{y_0(\tau)}^{y_1(\tau)} y(i,\tau) f(i) di = \tau \overline{y}(\tau).$$

By allowing the tax rate to affect individuals' before-tax income, we include in the model distortionary taxation as a determinate factor in voters' decisions.[7] We assume that, without taxation (i.e., for $\tau = 0$), individual's pre-tax income strictly increases with the individual index i:[8]

$$\frac{\partial y(i,0)}{\partial i} > 0, \forall i \in [0,1].$$

Marginal tax distortions negatively affect individual's pre-tax income, i.e.,

$$\frac{\partial y(i,\tau)}{\partial \tau} \leq 0, \forall i \in [0,1] \text{ and } \tau \in [0,1]$$

and we allow the intensity of this effect to vary across individuals, i.e.,

$$\frac{\partial^2 y(i,\tau)}{\partial \tau \partial i} \neq 0.$$

[7] Acemoglu and Robinson (2005) use a different specification of the standard model, including a general dead-weight cost of taxation to depict the distortive nature of taxation. We depart from this interpretation of distortion, allowing the tax rate to affect directly voters' pre-tax income, i.e. we assume that taxation distorts private incentives.

[8] We include this assumption in order to avoid tax distortions to change the ordering of individuals' income. Thus, in several parts of the remaining of the chapter we carry out the analysis with respect to y_i, hence also to i.

However, we constrain this cross-derivative to preserve the ranking of individuals' before-tax income at any level of taxation, i.e.

$$\frac{\partial y(i,\tau)}{\partial i} > 0, \forall i \in [0,1] \text{ and } \tau \in [0,1].$$

Voters' utility is decomposed in private consumption, c_i, and effective public consumption as perceived by each individual i, g_i, i.e.,

$$U_i = U(c_i, g_i),$$

where private consumption equals voter's after tax income:

$$c_i = (1-\tau)y_i(\tau).$$

The perceived effective consumption of public goods and services, g_i, measures the individual belief about the effective benefit the individual derives from total public expenditure, g. Specifically, it is formalized as a function of the total public expenditure and the individual income,

$$g_i = B(g, y_i(\tau)),$$

being increasing and concave in g, as in the standard models in the literature. We discuss more about the different interpretations of g_i in section 4.4, where we discuss the welfare implications of the model.[9]

4.3.1 Median voter result

In this section we state the conditions to obtain a median voter result in our model.

The first requirement is the individual preferences to be single peaked, i.e. concave in the policy choice:

[9]Our framework could incorporate voters with ideological preferences on the size of the public sector, g, i.e. voters with different utilities for public consumption according to their income. This would be done by $U(\cdot)$ becoming individual specific with g entering the function directly, and/or income being one of the direct arguments. This extension has been left for future work.

$$\frac{d^2U_i}{d\tau^2} = \frac{d}{d\tau}\left[\frac{\partial U}{\partial c_i}\left(\frac{\partial c_i}{\partial \tau} + \frac{\partial c_i}{\partial y_i}\frac{\partial y_i}{\partial \tau}\right) + \frac{\partial U}{\partial g_i}\left(\frac{\partial g_i}{\partial \tau} + \frac{\partial g_i}{\partial y_i}\frac{\partial y_i}{\partial \tau}\right)\right] < 0. \tag{4.1}$$

Then, we need to guarantee the monotonicity of preferences, i.e. voters' optimal tax rate to be inversely ordered by income:

$$\frac{d\tau_i}{di} = -\frac{\frac{d}{di}\left(\frac{dU_i}{d\tau}\right)}{\frac{d^2U_i}{d\tau^2}} < 0. \tag{4.2}$$

From inequality (4.1) we know that the denominator of the fraction is required to be negative, so the numerator should also be negative for the decreasing monotonicity to hold, i.e.

$$\frac{d}{di}\left(\frac{dU_i}{d\tau}\right) = \frac{d}{di}\left[\frac{\partial U}{\partial c_i}\left(\frac{\partial c_i}{\partial \tau} + \frac{\partial c_i}{\partial y_i}\frac{\partial y_i}{\partial \tau}\right) + \frac{\partial U}{\partial g_i}\left(\frac{\partial g_i}{\partial \tau} + \frac{\partial g_i}{\partial y_i}\frac{\partial y_i}{\partial \tau}\right)\right] < 0. \tag{4.3}$$

In other words we require the cross derivative of the preferences for taxation across voters to be decreasing in income.[10,11]

If the above conditions are satisfied then the winning taxation platform, τ_m, is the optimal rate of the median voter:

$$\tau_m = \underset{\tau}{\operatorname{argmax}}\, U_m(c_m, g_m)$$

with c_m and g_m denoting the private and the effective public consumption of the median voter respectively. In other words, τ_m is the platform that makes the marginal cost of public consumption for the median voter equal to her marginal benefit:

$$\frac{dU_m}{d\tau} = 0 \Rightarrow U'_{c_m}\frac{dc_m}{d\tau} + U'_{g_m}\frac{dg_m}{d\tau} = 0 \Rightarrow U'_{g_m}\frac{dg_m}{d\tau} = -U'_{c_m}\frac{dc_m}{d\tau}. \tag{4.4}$$

[10] Both inequalities (4.1) and (4.3) are discussed in more detail for every specification we use in section 4.4.

[11] Here we could also assume that the utility function is supermodular, which will imply the monotonicity condition.

4.3.2 Welfare

Our welfare benchmark is the utilitarian social welfare function that aggregates the preferences of the population as described below:

$$U_s = \int_i U(c_i, g_i) f(i) di.$$

Then the socially optimal taxation rate, τ_s, follows from the maximization of the above integral, i.e.,

$$\tau_s = \underset{\tau}{\operatorname{argmax}}\, U_s$$

with FOC as follows:

$$\frac{dU_s}{d\tau} = 0 \Rightarrow \int_0^1 \left(U'_{c_i} \frac{dc_i}{d\tau} + U'_{g_i} \frac{dg_i}{d\tau} \right) f(i) di = 0 \Rightarrow$$
$$\Rightarrow \int_0^1 \left(U'_{g_i} \frac{dg_i}{d\tau} \right) f(i) di = - \int_0^1 \left(U'_{c_i} \frac{dc_i}{d\tau} \right) f(i) di. \quad (4.5)$$

The welfare implications of the model depend on the comparison between the equilibrium tax rate, i.e., the preferred tax rate of the median voter, τ_m, and the optimal tax rate of the social planner, τ_s. We see that this comparison depends on several different channels, i.e., the differences in the marginal utility of private consumption across the population, the different perceptions of voters regarding effective public consumption, and the effects of distortive taxation. In the next section we discuss these channels separately.

4.4 The different channels of the welfare effects

In this section we specialize the model in order to isolate the effects of the aforementioned different channels and discuss their implications.

We start by illustrating the rigid welfare predictions of the standard model used in the literature, as recovered from the general model presented in section 4.3 by assuming quasi-linear preferences, identical benefits from the general provision of public goods for all voters, and non-distortionary taxation (in section 4.4.1). Thereafter, each of the following sections extend the standard model to incorporate each one of the additional channels of welfare effects which would operate in the general model of section 4.3. Proceeding in this way will enable us to assess the potential strength of each of these extra welfare effects without resorting to specific functional forms. Of course, the cost we pay relative to solving the general model with specific functional forms is to hide any relevant interaction between these effects. We first discuss the welfare implications of the marginal utility of private consumption varying across voters by abstracting from the assumption of quasi-linear preferences. Next, we reinstate quasi-linearity of preferences but remove the assumption of identical benefits from public good provision for all voters. Finally, we allow for distortive taxation and we discuss the welfare consequences of this channel.

As we will see, in all cases the welfare predictions of the model stop depending exclusively on the relative income positions of the median and the average voters. Moreover, in all cases under-provision of public goods can arise under empirically relevant income distributions. Interestingly, a general pattern will emerge in all comparisons of the political competition equilibrium and the utilitarian benchmark. Any of them will contrast a marginal effect for the median voter with the average of the same marginal effect for the entire population, and under- or over-provision of public good will crucially depend on the convexity or concavity of the relevant marginal effect across the population.

4.4.1 Standard model

From the general model discussed above we can recover the simpler model used also as a benchmark to discuss the welfare implications of this type of models. We can do this by assuming that the individual benefit from public consumption, g_i, is the same for all voters, the utility function $U(\cdot)$ is quasi-linear in private consumption, and taxation is not distortive, so that the pre-tax individual incomes depend only on the voters identifier index i. Formally:

$$U_i = c_i + g_i,$$

where

$$c_i = (1 - \tau)y_i$$

and

$$g_i = B(g)$$

with B being an increasing and concave function of public consumption, and $g = \tau\overline{y}$.

In this case, the FOC (4.4), which determines the equilibrium platform τ_m in the general model, reduces to:

$$U'_{g_m}\frac{dg_m}{d\tau} = -U'_{c_m}\frac{dc_m}{d\tau} \Rightarrow \frac{\partial B}{\partial g}\frac{\partial g}{\partial \tau} = -(-y_m) \Rightarrow B'_g = \frac{y_m}{\overline{y}}, \tag{4.6}$$

while the FOC (4.5), which determines the socially optimal platform τ_s, becomes:

$$\begin{aligned}\int_0^1 \left(U'_{g_i}\frac{dg_i}{d\tau}\right) f(i)di = -\int_0^1 \left(U'_{c_i}\frac{dc_i}{d\tau}\right) f(i)di \Rightarrow \\ \Rightarrow \int_0^1 \left(\frac{\partial B}{\partial g}\frac{\partial g}{\partial \tau}\right) f(i)di = -\int_0^1 (-y_i)\, f(i)di \Rightarrow \\ \Rightarrow B'_g = \frac{\overline{y}}{\overline{y}} = 1. \end{aligned} \tag{4.7}$$

Comparing conditions (4.6) and (4.7), it is clear that the two conditions coincide if the income distribution is symmetric (i.e., the median and the average income

coincide). In this case, the equilibrium provision of public goods is socially efficient.

If we however assume an empirically relevant, rightly-skewed distribution of income (i.e., with average greater than median income), then competition always over-provides public goods. Notice that, in this case, $\frac{y_m}{\bar{y}} < 1$. As a result, the right hand side (RHS) of equation (4.6) is lower than the RHS of equation (4.7). Recall also that B is concave in the general provision of public goods, g. Thus, if we evaluate both equations at the amount of g that solves equation (4.7), we see that equation (4.6) will require a higher amount of g in order to be satisfied. In other words, g, and hence τ must be greater in the political competition equilibrium than in the social optimum.

Therefore, we see that there exists a redistribution effect that is conditional only to the shape of the income distribution, i.e., the difference between the voters with median and average income. This result is a consequence of the quasi-linearity of preferences combined with a public good function that is identical for all voters. Because all voters receive the same benefit from taxation, the only channel that matters is the difference in their opportunity costs. Since the opportunity cost of taxation for poor voters is lower than the one for the richer part of the population, the median voter faces a lower marginal cost of taxation compared to the average voter. As voters' opportunity cost is only their income (in the absence of any tax distortions) the welfare result depends only on the shape of the income distribution.

4.4.2 Marginal utility of private consumption

We discussed above the welfare implications of the model assuming that voters' preferences are linear in private consumption. In this section we show that just removing this restrictive assumption will significantly enrich the welfare predictions of the model under empirically relevant income distributions.

To see this, we modify the model by just replacing the quasi-linear specification of the voters' preferences with the general functional specification adopted in section 4.3. We will keep assuming that all voters receive the same benefits from the total public expenditure, and that there are no tax distortions.

More specifically we consider the following utility function:

$$U_i = U(c_i, g_i) = U(c_i) + g_i,$$

where $c_i = c(\tau, y_i) = (1-\tau)y_i$ and $g_i = B(g)$ with $g = \tau\overline{y}$.

For this specification of the utility function the two conditions required for the median voter theorem to hold (inequalities (4.1) and (4.3)) become:

$$\frac{d^2U_i}{d\tau^2} < 0 \Rightarrow \frac{d}{d\tau}\left[\frac{\partial U}{\partial c_i}\left(\frac{\partial c_i}{\partial \tau}\right) + \frac{\partial U}{\partial g_i}\left(\frac{\partial g_i}{\partial \tau}\right)\right] < 0 \Rightarrow \frac{\partial^2 U}{\partial c_i^2}y_i^2 + \frac{\partial^2 B}{\partial g^2}\overline{y}^2 < 0. \quad (4.8)$$

$$\frac{d^2U_i}{didi\tau} < 0 \Rightarrow \frac{d}{di}\left[\frac{\partial U}{\partial c_i}\left(\frac{\partial c_i}{\partial \tau}\right) + \frac{\partial U}{\partial g_i}\left(\frac{\partial g_i}{\partial \tau}\right)\right] < 0 \Rightarrow \frac{d}{di}\left[\frac{\partial U}{\partial c_i}(-y_i)\right] < 0 \Rightarrow$$
$$\Rightarrow -\frac{\partial y(i)}{\partial i}\left[y_i(1-\tau)\frac{\partial^2 U}{\partial c_i^2} + \frac{\partial U}{\partial c_i}\right] < 0. \quad (4.9)$$

The second term of inequality (4.8) is negative by concavity of the B function. Therefore, a sufficient condition for the inequality to hold is the marginal utility of consumption to be decreasing. Inequality (4.9) is equivalent to the function $\frac{\partial U}{\partial c_i}(y_i)$ being increasing in i (and hence, in y_i), as the marginal utility of public goods does not depend on i in this specification of the utility function. In other words the marginal cost of taxation, i.e., the product of voters' marginal utility of private consumption and their income, should be increasing for the monotonicity condition to hold.

In this case, the equilibrium and the social optimum conditions (4.4) and (4.5) become:

$$U'_{g_m}\frac{dg_m}{d\tau} = -U'_{c_m}\frac{dc_m}{d\tau} \Rightarrow \frac{\partial B}{\partial g}\frac{\partial g}{\partial \tau} = -U'_{c_m}(-y_m) \Rightarrow B'_g = \frac{1}{\overline{y}}U'_{c_m}y_m \quad (4.10)$$

and

$$\int_0^1 \left(U'_{g_i} \frac{dg_i}{d\tau} \right) f(i)di = - \int_0^1 \left(U'_{c_i} \frac{dc_i}{d\tau} \right) f(i)di \Rightarrow$$

$$\Rightarrow \int_0^1 \left(\frac{\partial B}{\partial g} \frac{\partial g}{\partial \tau} \right) f(i)di = - \int_0^1 \left(-U'_{c_i} y_i \right) f(i)di \Rightarrow$$

$$\Rightarrow B'_g = \frac{1}{\overline{y}} \int_0^1 \left(U'_{c_i} y_i \right) f(i)di. \quad (4.11)$$

We see from equations (4.10) and (4.11) that the the shape of U'_{c_i} plays an important role in the determination of the preferred tax rate of the median voter and the social planner. In the remaining of this section we discuss the conditions required for competition to underprovide public goods in this case.

We start by considering a symmetric income distribution with income being almost identical across voters. We also assume that the marginal utility of private consumption is strictly concave in income, i.e., by Jensen's inequality the marginal utility of private consumption of the average voter is higher than the average marginal utility of private consumption of the population. Then, if we combine both assumptions and evaluate both equations at the same tax rate, we see that the RHS of equation (4.10) would be higher than the RHS of equation (4.11). As a result, by the concavity of the B function competition would require a lower tax rate than the social optimum. In other words strict concavity of the marginal utility of private consumption is a necessary condition for the under-provision result to hold in this case.

Let us now specifically assume an empirically relevant, rightly-skewed distribution of income. In this case if the position of the median voter is not very distant to the position of the average voter, then the under-provision result still holds. In other words, the RHS of equation (4.10) can still be higher than the RHS of equation (4.11). Therefore, by the concavity of the B function competition would under-provide public goods compared to the first best solution. Notice that these requirements (strict concavity of MU and median income close to the average income) are consistent with the two conditions for the median voter theorem to hold (inequalities (4.8) and (4.9)).

As we see from the above discussion, the direction of the welfare inefficiency depends heavily on the characteristics of the utility function. If the marginal utility of private consumption is strictly concave, this leads to the social planner demanding higher public expenditure compared to the median voter, as the latter faces a higher opportunity cost of public provision in terms of withdrawn private consumption. So, we see that in this case the common over-provision welfare result might not hold any more.

4.4.3 Effective public consumption

In this section we focus on the welfare implications of the heterogeneous benefits of taxation, abstracting from any other possible channels that might affect the welfare result. Therefore we assume that preferences are linear in private consumption and that voters' pre-tax income is fixed.

More specifically we consider the following utility function:

$$U_i = U(c_i, g_i) = c_i + g_i,$$

where $c_i = c(\tau, y_i) = (1-\tau)y_i$ and $g_i = B(g, y_i)$ with $g = \tau\overline{y}$.

For this specification of the utility function the two conditions required for the median voter theorem to hold (inequalities (4.1) and (4.3)) become:

$$\frac{d^2U_i}{d\tau^2} < 0 \Rightarrow \frac{d}{d\tau}\left[\frac{\partial U}{\partial c_i}\left(\frac{\partial c_i}{\partial \tau}\right) + \frac{\partial U}{\partial g_i}\left(\frac{\partial g_i}{\partial \tau}\right)\right] < 0 \Rightarrow \frac{\partial^2 B}{\partial g^2}\overline{y}^2 < 0. \tag{4.12}$$

$$\frac{d^2U_i}{didi\tau} < 0 \Rightarrow \frac{d}{di}\left[\frac{\partial U}{\partial c_i}\left(\frac{\partial c_i}{\partial \tau}\right) + \frac{\partial U}{\partial g_i}\left(\frac{\partial g_i}{\partial \tau}\right)\right] < 0 \Rightarrow$$
$$\Rightarrow \frac{\partial y(i)}{\partial i}\left[-1 + \frac{\partial^2 B}{\partial y_i \partial g}\overline{y}\right] < 0. \tag{4.13}$$

Inequality (4.12) holds by concavity of the B function, while a sufficient condition for inequality (4.13) to hold is that the function $\frac{\partial g_i}{\partial \tau}$ is decreasing in i (and hence, in y_i).

As presented above and shown also in section 4.3, we model each individual voter's expected effective benefit from public consumption, g_i, as a function of the aggregate public expenditure, g, and the voter's income. This individual specific public consumption can be interpreted in different ways. Voters might have different beliefs about the efficiency of the government. In other words they may have different perceptions, according to their income, about the effective amount of public goods and services a ruling politician would be able (or would like) to supply out of the overall tax revenues T (and hence, by the public budget balance constraint, out of the overall public expenditure g). For an example, along the lines of the first chapter of the thesis, voters could attach different weights, depending on their income, to the expected benefits promised by a ruling politician.

Another interpretation would rely on voters' expected differentiated access to and benefit from the overall provision of public goods and services as a function of their income. Different types of public expenditure can be disproportionately enjoyed by poor or by rich voters. For example, rich voters may enjoy relatively more public goods and services like police, security, and prisons, while poor voters may enjoy relatively more public schools and unemployment benefits. An overall composition of the public expenditure weighting more the first category of goods and services than the second, will therefore benefit more the rich than the poor voters. Of course, interpreting g_i in this way immediately poses important questions about the dimensions of the political competition among the contesting politicians.[12] Recall, for example, the model of the first chapter of the thesis. Voters not only attach different weights, depending on their income, to the expected benefits promised by one politician, but they also attach different weights to different politicians. This cross-candidates difference can be thought as voters' different valuation of the parts of the candidates' platforms which specify the mix of public goods and services they promise to deliver. Though, unless candidates are able to affect these weights, they are not competing in the mix of the public goods and

[12]Dotti (2020) discusses both the size and the composition of government in a model with redistribution and public good provision, concluding that the size of the government might decrease with inequality but the tax system might become more progressive.

services. However, the lack of competition in this policy dimension may be motivated by rigid ideological positions voters perceive for each candidate. This would make any strategic attempt to commit to a different composition of the public expenditure than the one voters expect from them non credible. However, our general framework could incorporate also multi-dimensional platform competition in both the size of the government and the different policies.[13] Our future work will address this different dimension and the potential implications on the welfare analysis.

After discussing the different interpretations of the g_i function we now concentrate on the welfare results. We see that in this case the two FOCs, (4.4) and (4.5), become:

$$U'_{g_m}\frac{dg_m}{d\tau} = -U'_{c_m}\frac{dc_m}{d\tau} \Rightarrow \frac{dg_m}{d\tau} = y_m, \tag{4.14}$$

and

$$\int_0^1 \left(U'_{g_i}\frac{dg_i}{d\tau}\right) f(i)di \;=\; -\int_0^1 \left(U'_{c_i}\frac{dc_i}{d\tau}\right) f(i)di \;\Rightarrow\; \int_0^1 \frac{dg_i}{d\tau} f(i)di \;=\; \overline{y}. \tag{4.15}$$

We start our analysis by assuming a symmetric income distribution (i.e., the median and the average income coincide). In this case it is clear that the RHS of equation (4.14) coincides with the RHS of equation (4.15). Therefore, the comparison between the equilibrium tax rate and the social tax rate depends on the difference between the LHS of the two equations. Suppose that the marginal effective public consumption ($\frac{dg_i}{d\tau}$) is strictly convex in income. Then, by Jensen's inequality, the average marginal effective public consumption of the society (LHS of equation (4.15)) is greater than the marginal effective public consumption of the average voter (LHS of (4.14)). Moreover, because of the concavity of the B function, the marginal effective public consumption is decreasing in τ. Therefore, condition (4.15) is satisfied at a higher tax rate compared to condition (4.14), i.e., competition under-provides goods and services. In other words the average voter

[13] Even though in general the median voter theorem fails to hold in a multi-dimensional electoral competition framework, we believe that under our specification of preferences, a median voter result might hold, proposing a tractable framework to study multi-dimensional competition.

faces the same marginal cost of taxation (in terms of withdrawn private consumption) as a utilitarian social planner, but lower marginal benefit of taxation. As a result, the average voter demands less taxation than the tax rate chosen by the social planner. On the other hand, if the voter with average income benefits more from public provision than the society on average (i.e., $\frac{dg_i}{d\tau}$ is strictly concave in income), then the welfare result will be the standard over-provision result.

Consider now a rightly-skewed income distribution. In this case the RHS of equation (4.14) is lower than the RHS of equation (4.15). The comparison of the LHS of both equations depends now also on the type of monotonicity of the marginal effective public consumption with respect to income. If the marginal benefit is strictly convex and increasing in income, then the marginal benefit of the median voter (LHS of equation (4.14)) is lower than the marginal benefit of the average voter, which (by Jensen's inequality) is lower than the average marginal benefit of the population (LHS of equation (4.15)). On the other hand, if it is decreasing, then the marginal benefit of the median voter (LHS of equation (4.14)) is greater than the marginal benefit of the average voter and it might also become greater than the average marginal benefit of the population (LHS of equation (4.15)). In both cases if y_m is arbitrarily close to $\overline{y}$, then still equation (4.15) is satisfied at a higher tax rate compared to (4.14).[14] However, in case where $\frac{dg_i}{d\tau}$ is increasing in income, even with a greater distance between the RHS of the two equations, competition might still under-provide public goods. In other words now the median voter faces a lower marginal cost, y_m, compared to the social planner, $\overline{y}$. If at the same time she receives also a higher marginal benefit than the average voter (i.e., marginal benefit decreasing in income) then both forces move the preferred median tax rate closer to the social optimum and dampen the under-provision result. However, if the function is increasing in income, then the median voter suffers from lower marginal benefit of taxation compared to the average voter, and therefore this effect moves the tax rate towards the opposite

[14]Notice that the under-provision result depends only on the convexity of the relevant function and on the distance between median and average income. Therefore, the conditions for the result are consistent with the requirements for the median voter theorem to hold.

direction of the effect of the marginal cost difference. So, in a society where the median voter receives lower marginal benefit than the average voter, we might have under-provision of public goods and services, even when income inequality (measured as the distance between the median and average income) is quite high.

Recall for example the model of the first chapter. If the society on average perceives a politician as more capable to provide public goods than the middle class (average voter), then the first best solution requires higher public expenditure than the competition result. On the contrary, if the middle class trusts the politician more than the population on average, then competition requires higher public goods provision from this government compared to the utilitarian solution. Notice that the result will be the same for a society where poorer voters trust the politician more and for one where richer voters trust him more, i.e., it does not depend on the type of monotonicity of the marginal benefit of effective public consumption. We could also think of the discussion above in terms of a mix of different public goods and policies. For example, imagine a case where taxation finances mainly public transportation expenses. In this case we can think that poorer voters benefit (in marginal terms) more than richer voters who might receive quite a lower marginal benefit of taxation and at a higher rate. As a result the median voter might face a lower marginal benefit than the average marginal benefit of the population and therefore this public good might be under-provided in the society compared to the first best. A different example could be financing mainly police expenses or public prisons. We can argue that poorer parts of the population care less about spending in policing compared to the middle class and the richer part of the population, as they would potentially prefer the government to publicly invest in other types of public goods as the ones mentioned in the previous example. Therefore it might be the case that the median voter receives higher marginal benefit from this type of public consumption compared to the marginal benefit of the population on average. In this case, electoral competition will over-provide this type of public goods.

4.4.4 Distortive taxation

Let us now add in the standard model only the channel of distortionary taxation. We keep all other assumptions as in the standard model, i.e., quasi-linear preferences and equal benefits from public expenditure across the population.

More specifically we consider the following utility function:

$$U_i = U(c_i, g_i) = c_i + g_i,$$

where $c_i = c(\tau, y_i(\tau)) = (1-\tau)y_i(\tau)$ and $g_i = B(g)$ with $g = \tau\overline{y}(\tau)$.

For this specification of the utility function the two conditions required for the median voter theorem to hold (inequalities (4.1) and (4.3)) become:

$$\begin{aligned}\frac{d}{d\tau}\left(\frac{dU_i}{d\tau}\right) = \frac{d}{d\tau}\left[\frac{\partial U}{\partial c_i}\left(\frac{\partial c_i}{\partial \tau} + \frac{\partial c_i}{\partial y_i}\frac{\partial y_i}{\partial \tau}\right) + \frac{\partial U}{\partial g_i}\left(\frac{\partial g_i}{\partial \tau} + \frac{\partial g_i}{\partial y_i}\frac{\partial y_i}{\partial \tau}\right)\right] < 0 \Rightarrow \\ \Rightarrow -2\frac{\partial y_i(\tau)}{\partial \tau} + (1-\tau)\frac{\partial^2 y_i(\tau)}{\partial \tau^2} + \\ + \frac{\partial^2 B}{\partial g^2}\left[\overline{y}(\tau)^2 + 2\tau\overline{y}(\tau)\frac{\partial \overline{y}(\tau)}{\partial \tau} + \overline{y}(\tau)^2\left(\frac{\partial \overline{y}(\tau)}{\partial \tau}\right)^2\right] + \\ + \frac{\partial B}{\partial g}\left[2\frac{\partial \overline{y}(\tau)}{\partial \tau}\tau\frac{\partial^2 \overline{y}(\tau)}{\partial \tau^2}\right] < 0. \quad (4.16)\end{aligned}$$

$$\begin{aligned}\frac{d}{di}\left(\frac{dU_i}{d\tau}\right) = \frac{d}{di}\left[\frac{\partial U}{\partial c_i}\left(\frac{\partial c_i}{\partial \tau} + \frac{\partial c_i}{\partial y_i}\frac{\partial y_i}{\partial \tau}\right) + \frac{\partial U}{\partial g_i}\left(\frac{\partial g_i}{\partial \tau} + \frac{\partial g_i}{\partial y_i}\frac{\partial y_i}{\partial \tau}\right)\right] < 0 \Rightarrow \\ \Rightarrow -\frac{\partial y_i(\tau)}{\partial i} + (1-\tau)\frac{\partial^2 y_i}{\partial i \partial \tau} < 0. \quad (4.17)\end{aligned}$$

Both conditions depend on the relative magnitude of the different terms. More specifically, we see that inequality (4.17) depends on the magnitude of two partial derivatives of income; the cross derivative and the derivative with respect to income.

In this case, the two first order conditions (equations (4.4) and (4.5)) become:

$$U'_{g_m}\frac{dg_m}{d\tau} = -U'_{c_m}\frac{dc_m}{d\tau} \Rightarrow \frac{\partial B}{\partial g}\frac{\partial g}{\partial \tau} = y_m(\tau) - (1-\tau)\frac{\partial y_m}{\partial \tau} \Rightarrow$$
$$\Rightarrow B'_g = \frac{y_m(\tau) - (1-\tau)\frac{\partial y_m}{\partial \tau}}{\overline{y}(\tau) + \tau\frac{\partial \overline{y}}{\partial \tau}} \quad (4.18)$$

and

$$\int_0^1 \left(U'_{g_i}\frac{dg_i}{d\tau}\right) f(i)di = -\int_0^1 \left(U'_{c_i}\frac{dc_i}{d\tau}\right) f(i)di \Rightarrow$$
$$\Rightarrow \frac{\partial B}{\partial g}\frac{\partial g}{\partial \tau} = -\int_0^1 \left((1-\tau)\frac{\partial y_i}{\partial \tau} - y_i(\tau)\right) f(i)di \Rightarrow$$
$$\Rightarrow B'_g = \frac{\overline{y}(\tau) - (1-\tau)\int_0^1 \left(\frac{\partial y_i}{\partial \tau}\right) f(i)di}{\overline{y}(\tau) + \tau\frac{\partial \overline{y}}{\partial \tau}}. \quad (4.19)$$

Suppose that marginal income ($\frac{\partial y_i}{\partial \tau}$) is strictly convex in voters' identifier index i and the distribution of income is symmetric. By Jensen's inequality the marginal income of the average voter is lower than the average marginal income of the population. Recall that income is decreasing in taxation for all voters. If we combine all of this and we evaluate both conditions at the same tax rate, we see that the RHS of equation (4.18) is higher than the RHS of equation (4.19):

$$\frac{\partial \overline{y}}{\partial \tau} < \int_0^1 \frac{\partial y_i}{\partial \tau} f(i)di \Rightarrow \frac{\overline{y}(\tau) - (1-\tau)\frac{\partial \overline{y}}{\partial \tau}}{\overline{y}(\tau) + \tau\frac{\partial \overline{y}}{\partial \tau}} > \frac{\overline{y}(\tau) - (1-\tau)\int_0^1 \left(\frac{\partial y_i}{\partial \tau}\right) f(i)di}{\overline{y}(\tau) + \tau\frac{\partial \overline{y}}{\partial \tau}}. \quad (4.20)$$

Therefore, because of the concavity of the B function, equation (4.18) will be satisfied at a lower tax rate than equation (4.19), i.e. competition under-provides public goods. Notice that we assume that the denominator, which is the same in both equations, is positive, i.e., $\frac{\partial g}{\partial \tau} > 0$. If that denominator is negative, then the tax rate is so high and the tax distortion so strong that public expenditure, and hence the total tax revenue becomes decreasing in the tax rate. In other words, we would be on the decreasing portion of a Laffer curve. But then, decreasing the tax rate will increase public consumption (i.e., tax revenue) and income (and hence private consumption) for all voters. As a result, such a tax rate cannot be

an equilibrium, as if one politician sets that high of a tax rate, the other will just undercut it and win everybody's vote. Also, such a tax rate will not be preferred by the social planner, as decreasing it would be a Pareto improvement. Hence, that denominator cannot be negative neither at the social optimum nor at the political competition equilibrium.[15]

Consider now the case of an empirically relevant, rightly-skewed income distribution. Under the reasonable assumption that the marginal distortion increases with income, we see that the marginal cost of the median voter (RHS of condition (4.18)) is lower than the marginal cost of the average voter, if they are evaluated at the same tax rate. Thus, the under-provision result can be dampened. However, competition will still under-provide public goods, if the distance between the median and the average income voter is sufficiently small.[16] In other words, with low inequality (measured as the difference between median and average income) the result can still exist. Notice also that with increasing marginal distortion richer voters suffer more from taxation and thus the difference between the average and median income will be shrinking, working in favour of under-provision.

4.5 Conclusion

Median voter models have been used extensively in the literature to describe the main forces of electoral competition and the democratic choice of public expenditure and redistribution. However, some of the main assumptions widely used, i.e., equal benefits from public good provision and quasi-linear preferences between private and public consumption, make the welfare implications of the model quite rigid. In this chapter we provide a general model, where we relax these assumptions, and discuss the interesting welfare implications that this general framework allows for. If people perceive the benefits of public expenditure differently accord-

[15] In general, economies being on the downward sloping part of the Laffer curve have been disputed based on empirical evidence. Our discussion could be providing a theoretical argument for these findings.

[16] Notice that the relevant conditions for under-provision of public goods are consistent with the conditions required for the median voter theorem to hold.

ing to their income, then competition might actually redistribute less than the utilitarian social optimum, even if the distribution of income is rightly-skewed. The result depends on the shape of the function of marginal benefit of effective public consumption with respect to income. If we think for example, a situation, where the middle class perceives the marginal benefit received from a politician in office as less valuable than the rest of the population on average, then the middle class might optimally choose under-provision of public goods and services. The same situation might arise if the middle class benefits less from a specific type of public good compared to the rest of the population on average. Our planned future work includes the analysis of an electoral competition model where politicians compete also on the provision of different public goods, i.e. a multi-dimensional competition on both the size and the composition of the government.

We also elaborate on the effects of heterogeneous marginal utility of private consumption and distortive taxation. We discuss the underlying comparison for different channels between the marginal effect of the median voter and the average marginal effect of the population. We saw that the welfare implications of the model depend eventually on the convexity or concavity of the marginal benefit and cost functions. Even though, a formal discussion of the interaction of these three different channels remains to be addressed, we should expect that the result will depend on the sign and the magnitude of the different forces. If for example, the median voter benefits less from a specific public good and if she also faces higher marginal tax distortion than the population on average, then this last effect should strengthen the under-provision result. The same can happen if at the same time her marginal utility of private consumption is higher than the average societal marginal utility. We believe that this general framework provides an appropriate environment to study the welfare implications of median voter models bringing together different and important channels both in the cost and in the benefit side of public expenditure.

Chapter 5

Conclusion

This thesis studied in three chapters the ways electoral decisions can be affected by heterogeneous beliefs and the presence of group identity considerations.

In the second chapter we showed how electoral competition may lead to over-austerity after a negative economic shock. This result is driven by the different beliefs voters form about the competence of politicians to rule after the shock. More specifically, if the population on average trusts more an outsider challenger than the middle class, then competition under-provides public goods and vice versa. However, it is worthwhile noticing that while small shocks may not have drastic effects on voters' beliefs on politicians' competence to deliver public good, economic crises, like the great recession, can create huge and possibly quite persistent fluctuations in the existing political system. Our planned future work will focus on the dynamic analysis of political instability due to economic shocks.

The third chapter of this thesis explored the effects of identity attachment in voters' decision to re-elect possibly corrupt politicians. We built a model of electoral accountability with pure moral hazard and tested it in the lab. Our main result showed that, when voters observed a failed project, they approve politicians of their same identity group significantly more often than in the control and compared to politicians of a different group. This is partially driven by voters believing that same-identity politicians were more honest. We also saw that subjects acting as politicians are much more honest than expected by the equilibrium prediction.

Lastly, generalising the idea of the second chapter, the fourth chapter discussed the welfare implications that can arise once the standard assumptions of the median voter models of redistribution are relaxed. We considered a general electoral competition model without quasi-linear preferences, with voters' differ-

ent perceptions of public expenditure conditional on income, and with distortive taxation. We argued successively that in each of the above cases, the competition can under-provide public goods and services, even if the income distribution is rightly-skewed. This contradicts the standard welfare implications of these type of models. More specifically, we saw that the welfare result of the model depends mainly on the convexity or concavity of the relevant marginal effects of taxation across the population. Our future work will consider a multidimensional electoral competition framework to allow for politicians to compete for the composition of public policies.

Appendices

Appendix A

A.1 Omitted proofs

Proof of Lemma 2.1. Using equation (2.2), the derivative of the expected utility of a voter regarding the challenger with respect to the taxation platform is:

$$\frac{\partial EU_i^c}{\partial \tau} = -y_i + \phi(y_i)\, W'(\tau\, \overline{y})\, \overline{y} \equiv G(\tau, y_i).$$

Then, applying the Implicit Function Theorem, we obtain:

$$\frac{\partial \tau_i}{\partial y_i} = -\frac{G'_{y_i}}{G'_{\tau_i}} = -\frac{-1 + \phi'(y_i)\, W'(\tau_i\, \overline{y})\, \overline{y}}{\phi(y_i)\, W''(\tau_i\, \overline{y})\, \overline{y}^2}.$$

□

Proof of Lemma 2.2. The challenger wins the competition iff the expected utility regarding him is higher than the expected utility regarding the ruler for the majority of the population:

$$\mathrm{EU}_i^c(\tau^c, y_i) > \mathrm{EU}_i^r(\tau^r, y_i) \Leftrightarrow y_i\,(\tau^r - \tau^c) + \phi(y_i)\, W(\tau^c\, \overline{y}) - \lambda\, W(\tau^r\, \overline{y}) > 0. \quad \text{(A.1)}$$

For any $\tau^r \in (0, 1]$ the challenger can choose $\tau^c = \tau^r = \tau$ and win. Notice that in this case inequality (A.1) becomes

$$\mathrm{EU}_i^c(\tau, y_i) > \mathrm{EU}_i^r(\tau, y_i) \Leftrightarrow (\phi(y_i) - \lambda)\, W(\tau\, \overline{y}) > 0. \quad \text{(A.2)}$$

Inequality (A.2) holds since $\phi(y_i) > \lambda$ for any $y_i \in [0, 1]$ (assumption 2.3) and $W(\tau\, \overline{y}) > 0$ for any $\tau \in (0, 1]$. For $\tau = 0$ all voters are indifferent between the two candidates, but the challenger has a profitable deviation. There exists an $\varepsilon > 0$ and small enough such that, if $\tau^c = \varepsilon$, the challenger collects all the votes:

$$\mathrm{EU}_i^c(\varepsilon, y_i) > \mathrm{EU}_i^r(0, y_i) \text{ for all } y_i \in [0, 1].$$

To see this notice that by assumption 2.2:

$$\lim_{\tau^c \to 0} \frac{\partial EU_i^c}{\partial \tau^c} = \lim_{\tau^c \to 0} [-y_i + \phi(y_i)\, \overline{y} W'(\tau^c \overline{y})] \to +\infty.$$

□

Appendix B

B.1 Additional Results

Sessions summary

Lab	Date	Treatment	Subjects	Average earnings	Min	Max
LExEcon	13/3/19	Identity	16	11.75	4	28
LExEcon	13/3/19	Control	16	12.25	4	30
LExEcon	14/3/19	Identity	16	12.75	6	32
LExEcon	14/3/19	Identity	16	16.5	6	32
LExEcon	20/3/19	Control	16	11.5	4	28
LExEcon	20/3/19	Identity	16	19.25	6	30
LExEcon	21/3/19	Identity	16	15.5	6	30
BEADS	11/11/19	Identity	14	16	6	30
BEADS	11/11/19	Control	16	18.625	4	28

Table B.1: Summary information per session

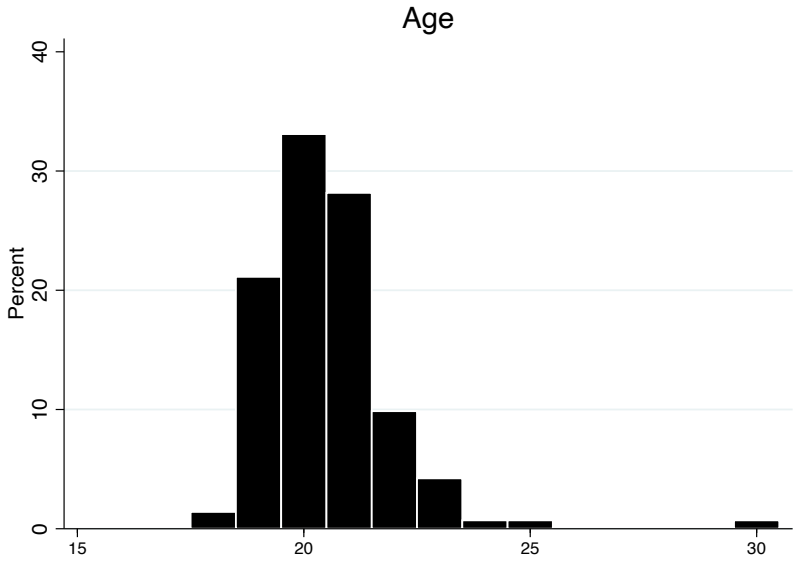
Questionnaire-Descriptives

"What is your gender?"

Gender	Percent
Female	57.75
Male	42.25

Table B.2: Gender

"What is your age?"

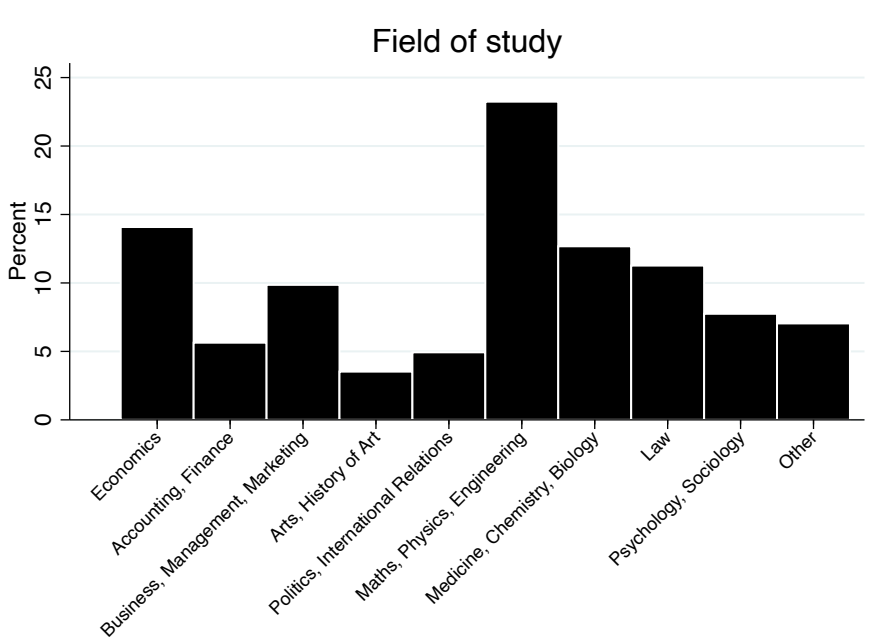


"What is the title of your degree? "

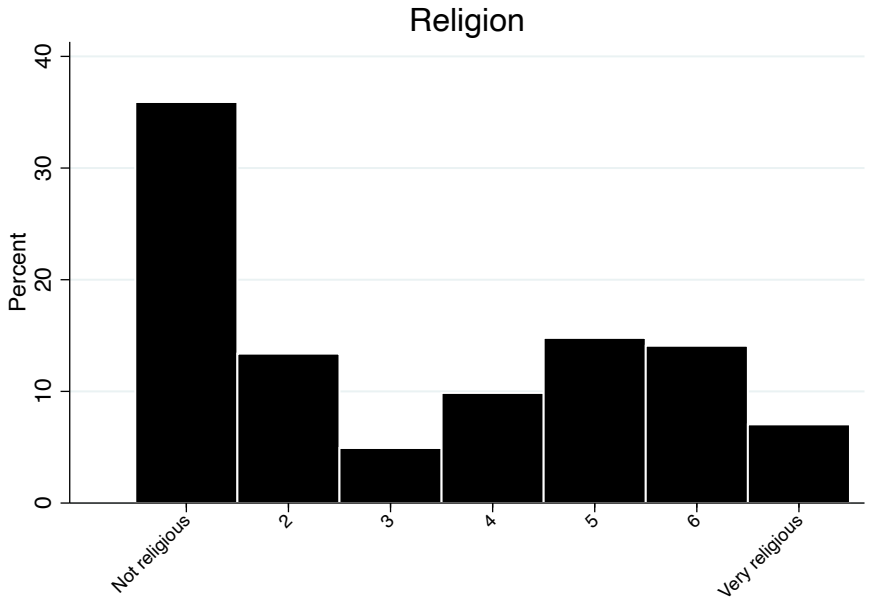
Title	Percent
Bachelor	86.62
Master	13.38

Table B.3: Degree title

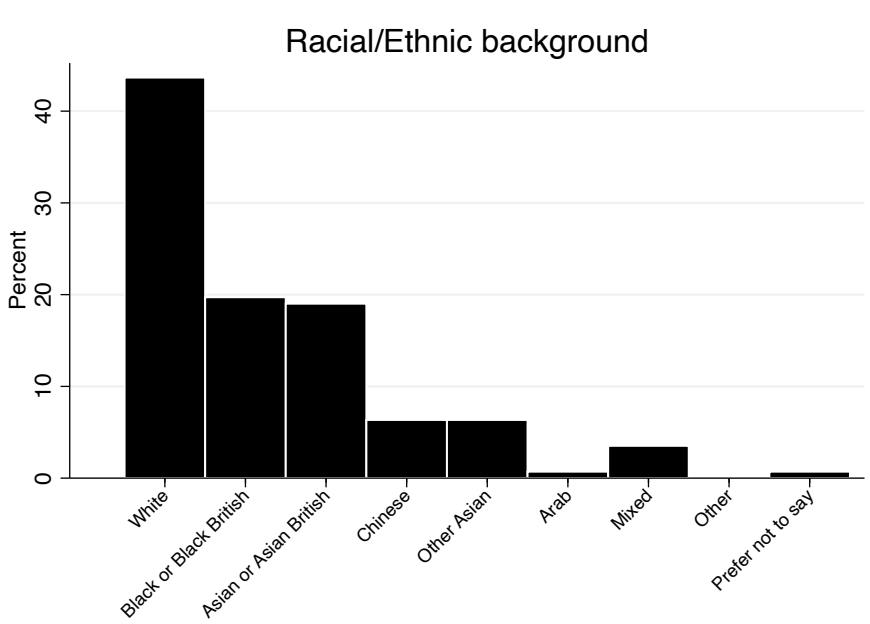
"Please give us the main field of your study or the one closer to it."



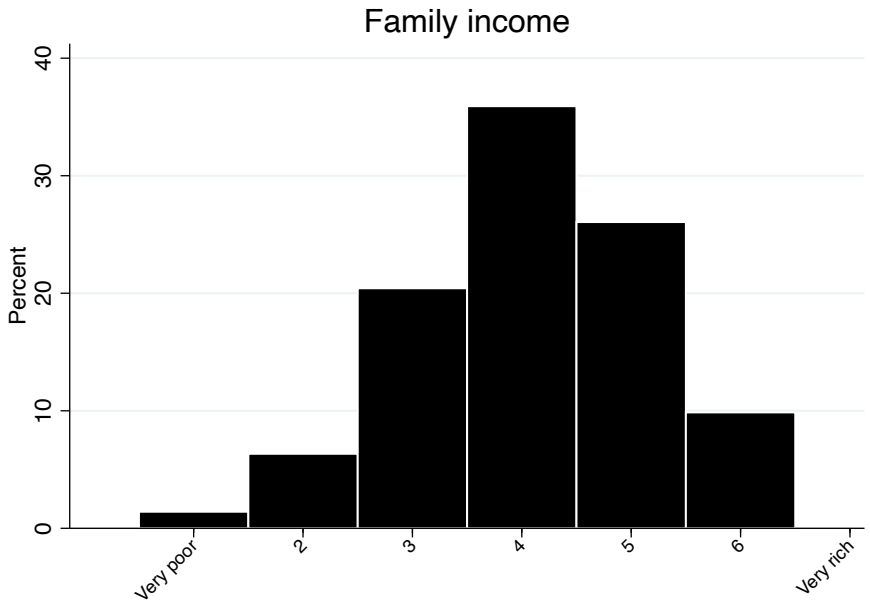
"How religious would you describe yourself?"



"Which of the following best describes your racial or ethnic background?"



"Taking everything into account, how would you characterise the standard of living of the family in which you were raised?"

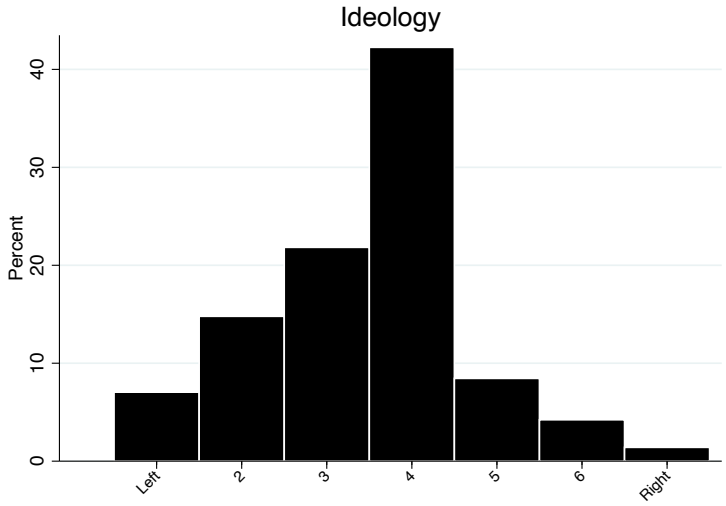


"Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?"

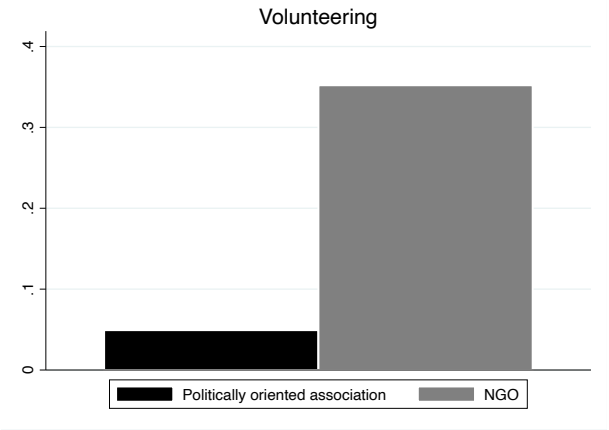
Trust	Percent
Can't be too careful	66.2
Most people can be trusted	33.8

Table B.4: Trust

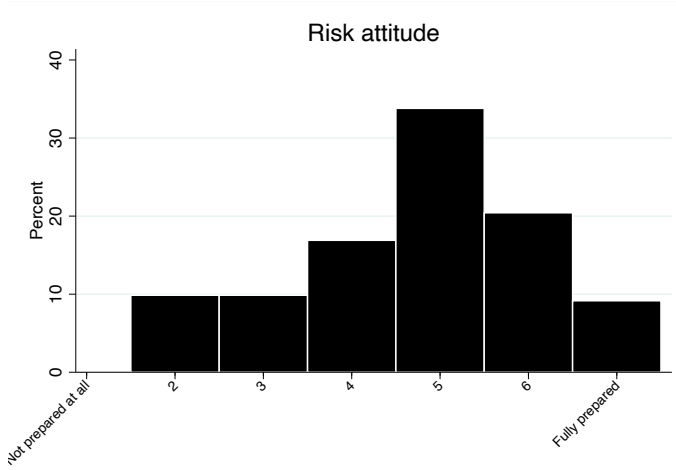
"In political matters people talk of "the left" and "the right". What is your position?"



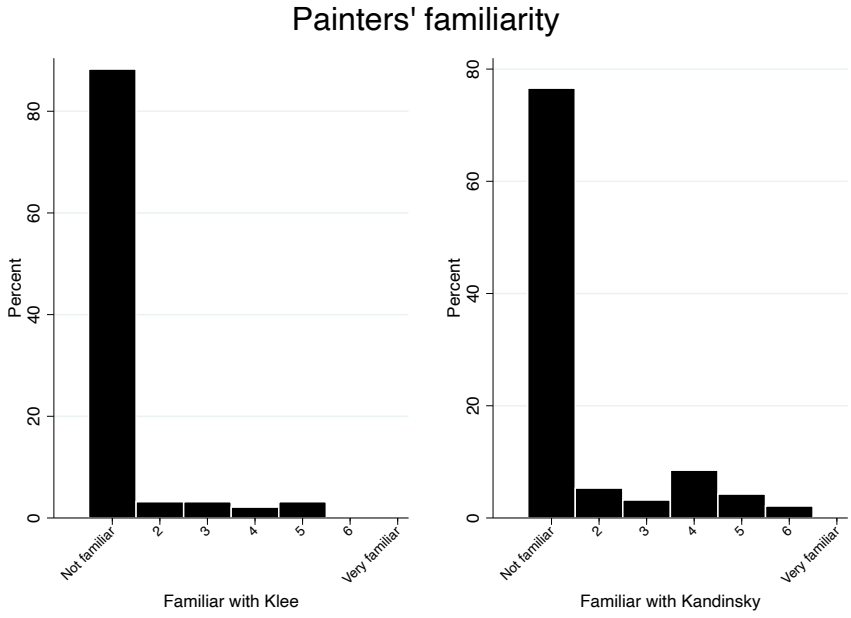
"In the last 12 months have you volunteered for any politically oriented association/charity or NGO?"



"How do you see yourself? Are you a person who is fully prepared to take risks or do you try to avoid taking risks?"[1]

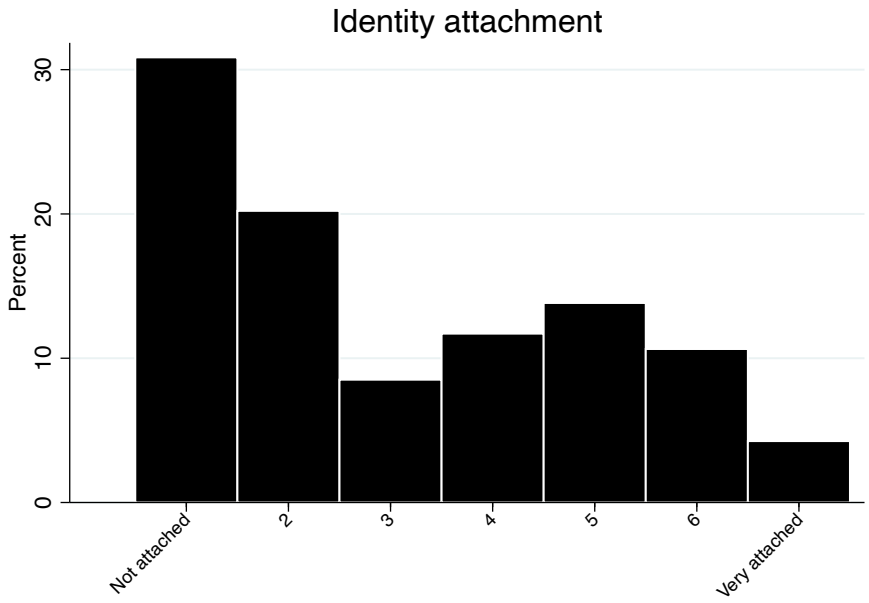


"Please rate how familiar you were with the paintings by Klee/Kandinsky before this experiment."



[1]Dohmen et al., 2011 have validated this question as a predictor of risky behaviour.

"How closely attached you felt to your identity (as a Klee or as a Kandinsky) throughout the experiment?"



Additional Regressions on Approval rates

Approval choice (32 tokens)	(1)	(2)	(3)
Different Identity vs. Control	-0.0615	-0.0623	-0.0631
	(0.0442)	(0.0452)	(0.0461)
Same vs. Different Identity	0.0374	0.0396	0.0420
	(0.0279)	(0.0304)	(0.0306)
Round	-0.0026**	-0.0029**	-0.0029**
	(0.0013)	(0.0013)	(0.0013)
Birmingham	0.0417*	0.0409	0.0464
	(0.0246)	(0.0272)	(0.0295)
Start as Representative		0.0436**	0.0372*
		(0.0218)	(0.0201)
STEM Studies		0.0488***	0.0604***
		(0.0152)	(0.0181)
$\geq$ 21 years			0.0313
			(0.0256)
Centre-Right Ideology			-0.0303
			(0.0362)
Female			0.0386
			(0.0292)
Observations	585	585	585

Robust errors clustered by sessions. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.5: Panel Linear Probability Models on the decision of the citizen (project succeeded)

Approval choice (0 tokens)	(1)	(2)	(3)
Different Identity vs. Control	0.0435	0.0418	0.0295
	(0.0434)	(0.0423)	(0.0409)
Same vs. Different Identity	0.0674**	0.0684**	0.0685**
	(0.0293)	(0.0294)	(0.0291)
Round	-0.0020*	-0.0020*	-0.0021*
	(0.0012)	(0.0012)	(0.0012)
Birmingham	0.3238***	0.3228***	0.3310***
	(0.0328)	(0.0321)	(0.0345)
Start as Representative		0.0555	0.0443
		(0.0603)	(0.0708)
STEM Studies		0.0831***	0.0892***
		(0.0280)	(0.0219)
$\geq$ 21 years			0.0462
			(0.0537)
Centre-Right Ideology			-0.0307
			(0.0621)
Female			-0.0448
			(0.0466)
Observations	1,971	1,971	1,971

Robust errors clustered by sessions. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.6: Panel Linear Probability Models on the decision of the citizen (project failed-different identity vs. control)

Panel Probit models

Approval rates

	Approval choice (0 tokens)			Margins (dy/dx)		
	(1)	(2)	(3)	(1)	(2)	(3)
Same Identity vs. Control	0.4346**	0.4333**	0.3869*	0.1076**	0.1073**	0.0958**
	(0.2004)	(0.1966)	(0.1953)	(0.0496)	(0.0484)	(0.0480)
Different vs. Same Identity	-0.2701**	-0.2743**	-0.2715**	-0.0668**	-0.0679**	-0.0672**
	(0.1132)	(0.1138)	(0.1131)	(0.0301)	(0.0302)	(0.0298)
Round	-0.0079*	-0.0079*	-0.0080*	-0.0019*	-0.0019*	-0.0019*
	(0.0047)	(0.0048)	(0.0047)	(0.0011)	(0.0011)	(0.0011)
Birmingham	1.2621***	1.2607***	1.2846***	0.3124***	0.3124***	0.3183***
	(0.1352)	(0.1384)	(0.1574)	(0.0291)	(0.0306)	(0.0335)
Start as Representative		0.2578	0.2122		0.0639	0.0526
		(0.2583)	(0.3023)		(0.0650)	(0.0758)
STEM Studies		0.3344***	0.3545***		0.0828***	0.0878***
		(0.1080)	(0.1076)		(0.0231)	(0.0238)
Centre-Right Ideology			-0.1967			-0.0487
			(0.2370)			(0.0589)
$\geq$ 21 years			0.1276			0.0316
			(0.2165)			(0.0533)
Female			-0.1418			-0.0351
			(0.1730)			(0.0423)
Observations	1,971	1,971	1,971	1,971	1,971	1,971

Robust errors clustered by sessions. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.7: Panel Probit Models and margins on the decision of the citizen (project failed)

	Approval choice (0 tokens)			Margins (dy/dx)		
	(1)	(2)	(3)	(1)	(2)	(3)
Different Identity vs. Control	0.1645	0.1589	0.1153	0.0407	0.0393	0.0285
	(0.1608)	(0.1568)	(0.1538)	(0.0388)	(0.0378)	(0.0373)
Same vs. Different Identity	0.2701**	0.2743**	0.2715**	0.0668**	0.0679**	0.0672**
	(0.1132)	(0.1138)	(0.1131)	(0.0301)	(0.0302)	(0.0298)
Round	-0.0079*	-0.0079*	-0.0080*	-0.0019*	-0.0019*	-0.0019*
	(0.0047)	(0.0048)	(0.0047)	(0.0011)	(0.0011)	(0.0011)
Birmingham	1.2621***	1.2607***	1.2846***	0.3124***	0.3124***	0.3183***
	(0.1352)	(0.1384)	(0.1574)	(0.0291)	(0.0306)	(0.0335)
Start as Representative		0.2578	0.2122		0.0639	0.0526
		(0.2583)	(0.3023)		(0.0650)	(0.0758)
STEM Studies		0.3344***	0.3545***		0.0828***	0.0878***
		(0.1080)	(0.1076)		(0.0231)	(0.0238)
Centre-Right Ideology			-0.1967			-0.0487
			(0.2370)			(0.0589)
$\geq$ 21 years			0.1276			0.0316
			(0.2165)			(0.0533)
Female			-0.1418			-0.0351
			(0.1730)			(0.0423)
Observations	1,971	1,971	1,971	1,971	1,971	1,971

Robust errors clustered by sessions. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.8: Panel Probit Models and margins on the decision of the citizen (project failed-different identity vs. control)

	Approval choice (32 tokens)			Margins (dy/dx)		
	(1)	(2)	(3)	(1)	(2)	(3)
Different Identity vs. Control	-0.5317	-0.5705	-0.5664	-0.0421	-0.0457	-0.0458
	(0.3881)	(0.3999)	(0.4112)	(0.0388)	(0.0378)	(0.0402)
Same vs. Different Identity	0.2898	0.3333	0.3650	0.0229	0.0267	0.0257
	(0.2329)	(0.2594)	(0.2631)	(0.0219)	(0.0249)	(0.0257)
Round	-0.0317*	-0.0347**	-0.0349**	-0.0025**	-0.0027**	-0.0028*
	(0.0164)	(0.0163)	(0.0161)	(0.0012)	(0.0012)	(0.0013)
Birmingham	0.6488*	0.6575**	0.6171*	0.5146*	0.0527*	0.0499*
	(0.3387)	(0.3345)	(0.3384)	(0.0306)	(0.0305)	(0.0303)
Start as Representative		0.5803**	0.5180**		0.0465**	0.0419**
		(0.2410)	(0.2206)		(0.0214)	(0.0176)
STEM Studies		0.6723***	0.7501***		0.0539**	0.0607**
		(0.2606)	(0.2389)		(0.0241)	(0.0242)
Centre-Right Ideology			-0.4631			-0.0375
			(0.4216)			(0.0401)
$\geq$ 21 years			0.2074			0.0168
			(0.3200)			(0.0260)
Female			0.3339			0.0270
			(0.3421)			(0.0263)
Observations	585	585	585	585	585	585

Robust errors clustered by sessions. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.9: Panel Probit Models and margins on the decision of the citizen (project succeeded)

Sending rates

	Sending choice		Margins (dy/dx)	
	(1)	(2)	(1)	(2)
Same Identity vs. Control	-0.1569	-0.1436	-0.0382	-0.0349
	(0.216)	(0.2280)	(0.0509)	(0.0539)
Different vs. Same Identity	-0.0271	-0.02408	-0.0066	-0.0058
	(0.0507)	(0.0501)	(0.01207)	(0.0119)
Round	-0.0115**	-0.0115**	-0.0028**	-0.0028**
	(0.0053)	(0.0054)	(0.00128)	(0.00129)
Birmingham	-0.4945	-0.4827	-0.1204*	-0.1176
	(0.3128)	(0.3339)	(0.0726)	(0.0773)
Start as Representative	0.2777*	0.2235	0.0676*	0.0544
	(0.1623)	(0.1415)	(0.0397)	(0.0353)
STEM Studies		-0.1185		-0.0288
		(0.29211)		(0.0708)
Centre-Right Ideology		-0.3759*		-0.0915*
		(0.2067)		(0.05006)
$\geq$ 21 years		-0.0358		-0.0087
		(0.1726)		(0.0419)
Female		0.3368		0.082
		(0.2496)		(0.0595)
Observations	2,556	2,556	2,556	2,556

Robust errors clustered by sessions. * $p < 0.1$, ** $p < 0.05$

Table B.10: Panel Probit Models and margins on the decision of the representative

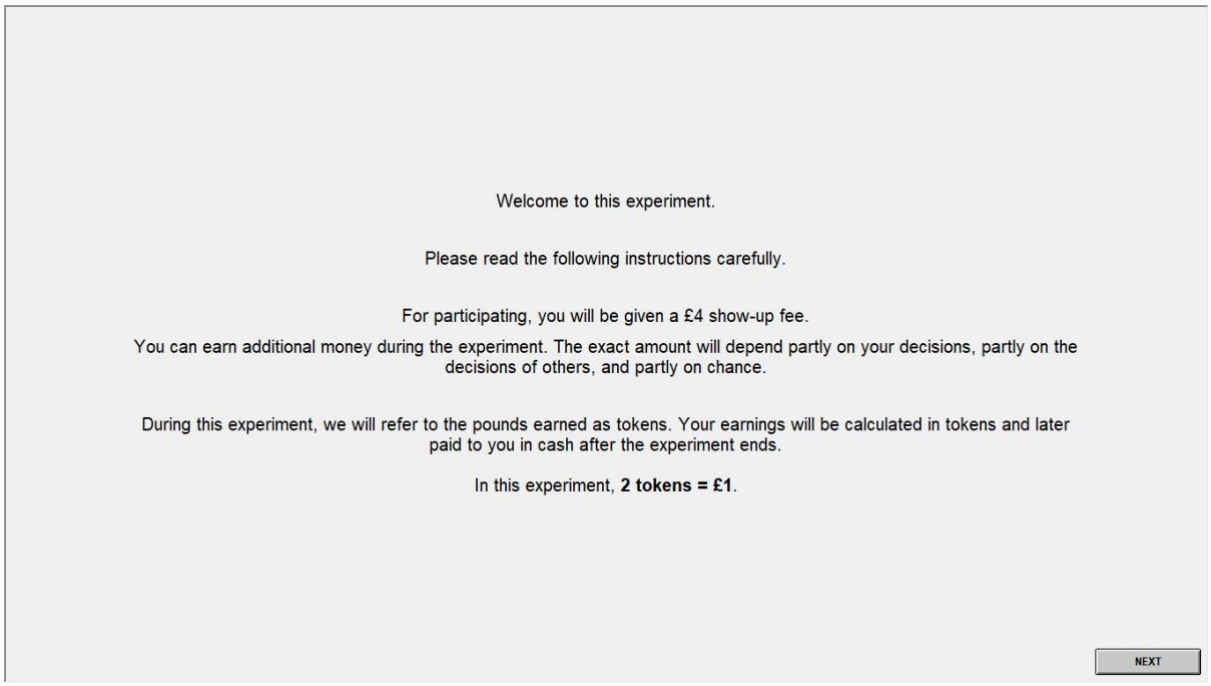
Sending choice	(1)	(2)
Kandinsky	-0.00611	-0.0401
	(0.0889)	(0.0732)
Round	-0.00348**	-0.00349**
	(0.00150)	(0.00149)
Birmingham	-0.157***	-0.164***
	(0.0131)	(0.0301)
STEM Studies		-0.105
		(0.0892)
Centre-Right Ideology		-0.125***
		(0.0473)
$\geq$ 21 years		0.000379
		(0.0336)
Female		0.0644
		(0.0686)
_cons	0.520***	0.616***
	(0.0893)	(0.0915)
Observations	1692	1692

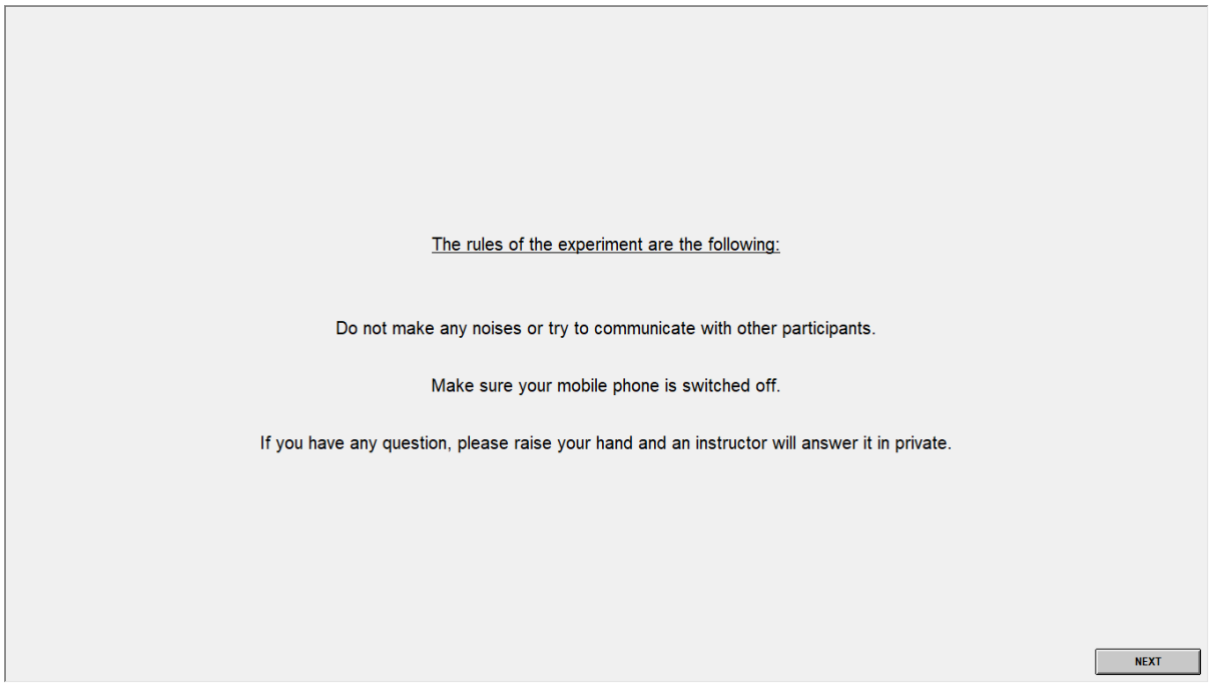
Standard errors in parentheses

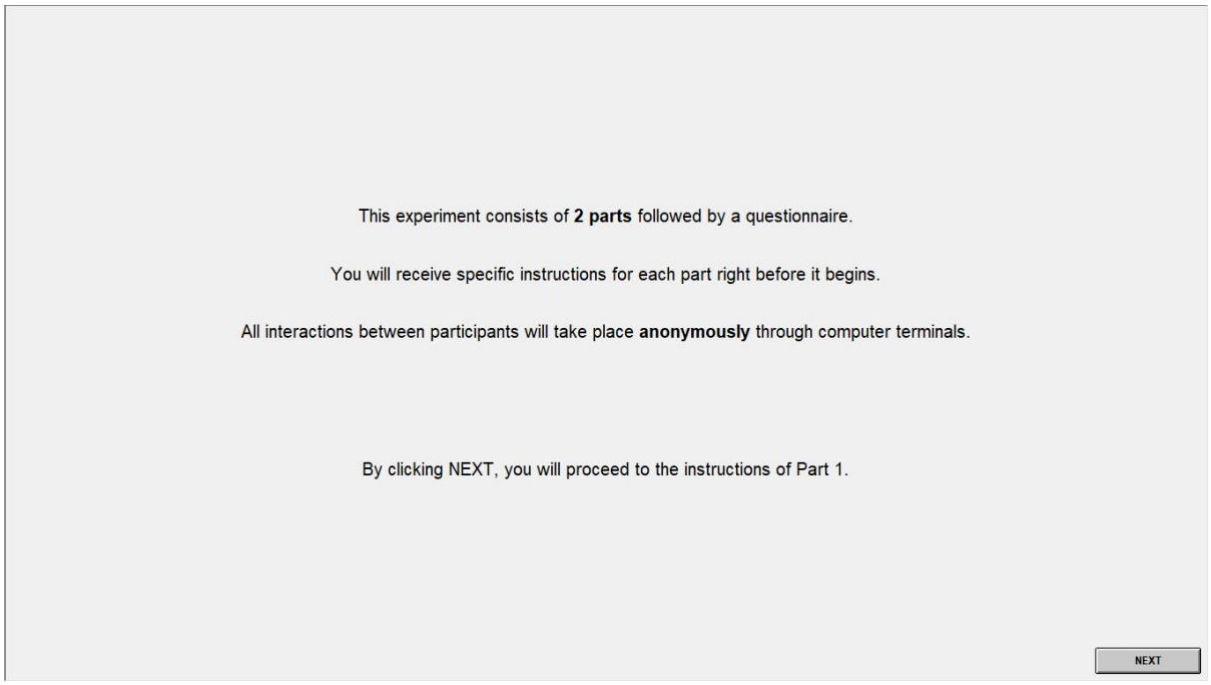
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

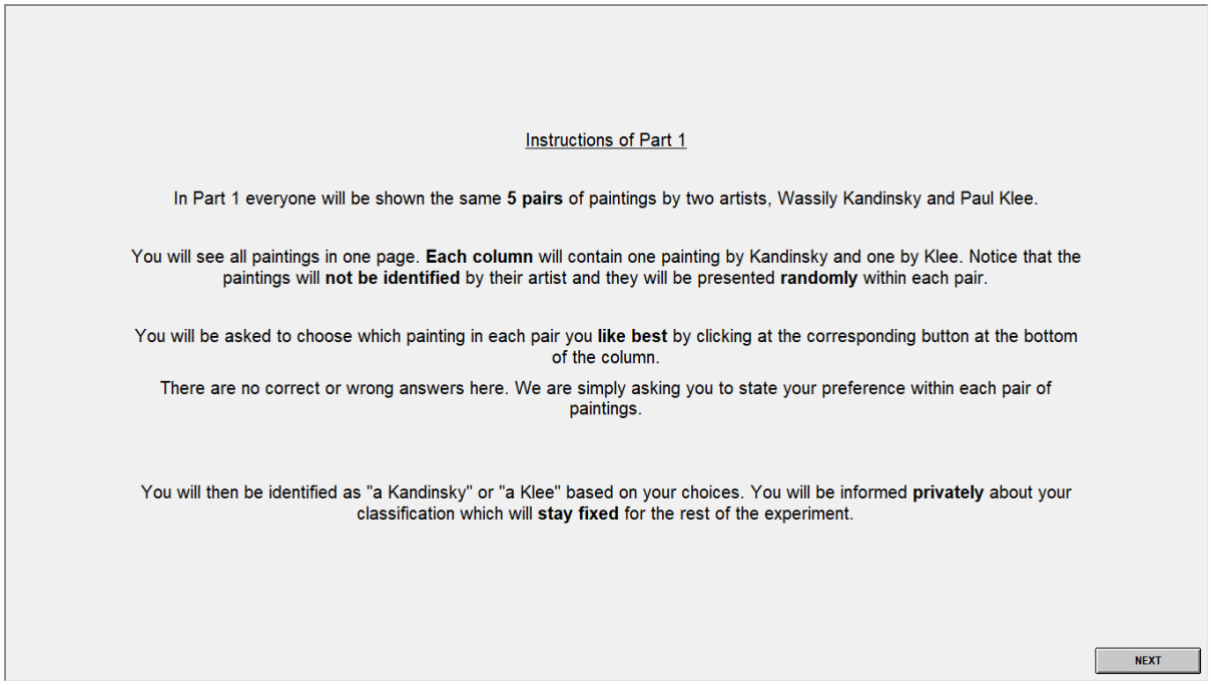
Table B.11: Panel Linear Probability Models on the decision of the representative controlling for identity

Instructions






This experiment consists of **2 parts** followed by a questionnaire.
You will receive specific instructions for each part right before it begins.
All interactions between participants will take place **anonymously** through computer terminals.
By clicking NEXT, you will proceed to the instructions of Part 1.
NEXT


Instructions of Part 1
In Part 1 everyone will be shown the same **5 pairs** of paintings by two artists, Wassily Kandinsky and Paul Klee.
You will see all paintings in one page. **Each column** will contain one painting by Kandinsky and one by Klee. Notice that the paintings will **not be identified** by their artist and they will be presented **randomly** within each pair.
You will be asked to choose which painting in each pair you **like best** by clicking at the corresponding button at the bottom of the column.
There are no correct or wrong answers here. We are simply asking you to state your preference within each pair of paintings.
You will then be identified as "a Kandinsky" or "a Klee" based on your choices. You will be informed **privately** about your classification which will **stay fixed** for the rest of the experiment.
NEXT

Instructions of Part 1

Then, everyone will be shown **2 additional paintings**, one by Kandinsky and one by Klee.

You will be asked to **identify which of the two was painted by Kandinsky if you are a Kandisnky and which was painted by Klee if you are a Klee.**

You will be asked to submit **two answers: your initial guess and your final answer.**

After submitting your initial guess, you will see the initial guesses of you and your fellow "Kandinskys" if you are a Kandinsky or "Klees" if you are a Klee.

Then you will be given the opportunity to **change or confirm** your initial guess and submit your final answer.

To help you with this choice, you will be shown again the screen with the 5 pairs of paintings but with their artist now identified.

NEXT

Instructions of Part 1

Your payment in Part 1 of the experiment will be determined as follows.

If you are a Kandinsky and a half or more of Kandinskys give a correct final answer, then regardless of whether your own final answer was correct or incorrect, you and each of your fellow Kandinskys will receive **4 tokens**.

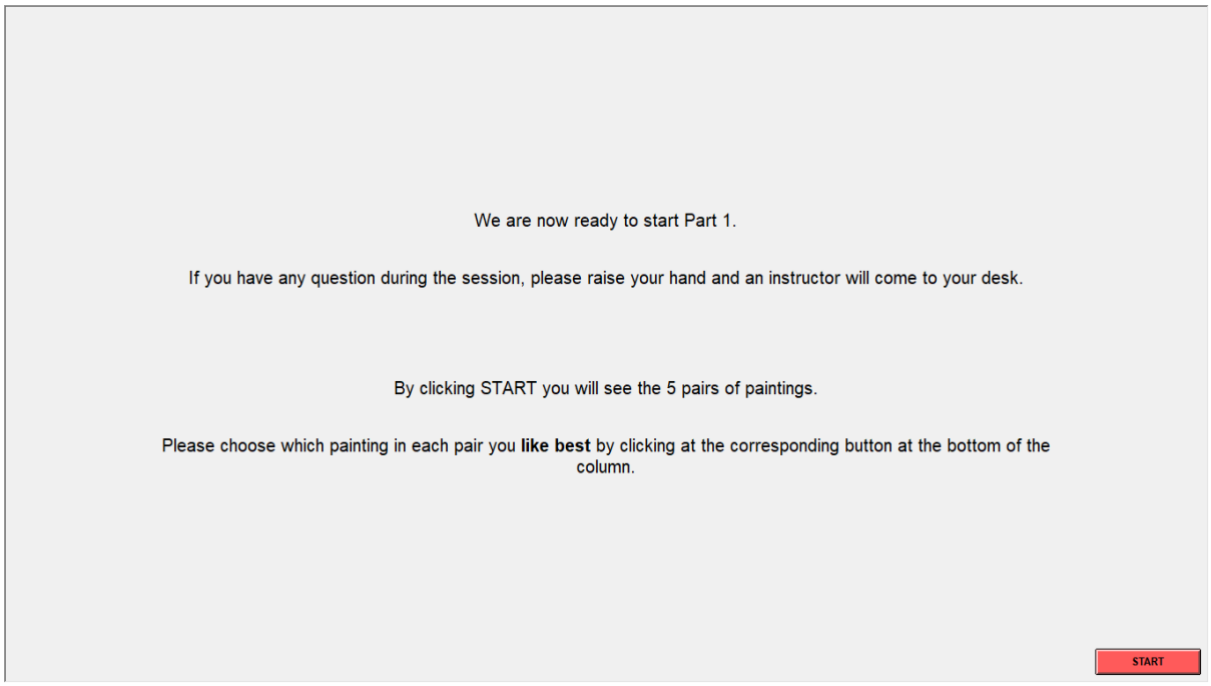
Similarly, if you are a Klee and a half or more of Klees give a correct final answer then regardless of your own final answer, each of the Klees, including you, will receive **4 tokens**.

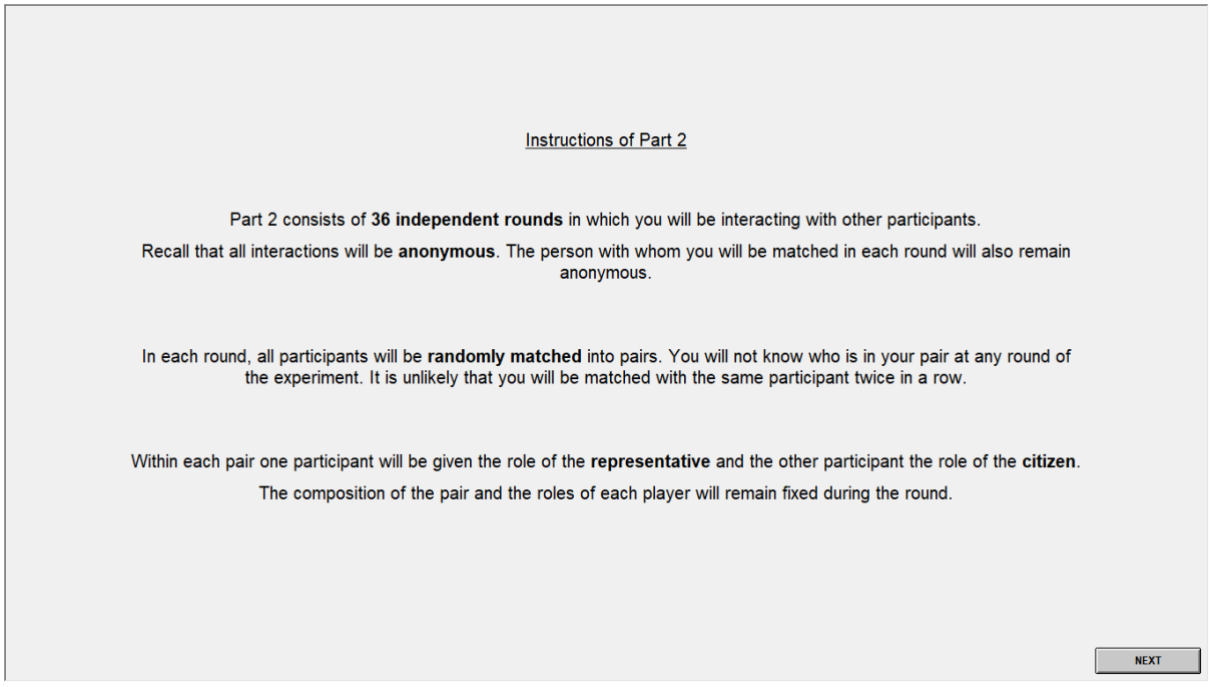
However, if you are a Kandinsky and more than a half of Kandinskys give an incorrect final answer, then, regardless of whether your own final answer was correct or incorrect, you and each of the Kandinskys will receive **0 tokens**.

And similarly, if you are a Klee and the final answers from more than a half of Klees were incorrect, then you and each of your fellow Klees will receive **0 tokens** regardless of what answer he/she gave personally.

Remember that the exchange rate is **2 tokens = £1**. You will find out your payment from Part 1 at the end of the experiment.

NEXT


We are now ready to start Part 1.
If you have any question during the session, please raise your hand and an instructor will come to your desk.
By clicking START you will see the 5 pairs of paintings.
Please choose which painting in each pair you **like best** by clicking at the corresponding button at the bottom of the column.
START


Instructions of Part 2
Part 2 consists of **36 independent rounds** in which you will be interacting with other participants.
Recall that all interactions will be **anonymous**. The person with whom you will be matched in each round will also remain anonymous.
In each round, all participants will be **randomly matched** into pairs. You will not know who is in your pair at any round of the experiment. It is unlikely that you will be matched with the same participant twice in a row.
Within each pair one participant will be given the role of the **representative** and the other participant the role of the **citizen**.
The composition of the pair and the roles of each player will remain fixed during the round.
NEXT

Instructions of Part 2

You will play with the same role for the first 6 rounds. After that, your role will switch for another block of 6 rounds.

For the remaining 24 rounds your role will be reassigned randomly in blocks of 6 rounds.

For example, if you play the first 6 rounds as representative, you will play the next 6 rounds as citizen. After the 12th round you will play randomly as representative or as citizen. In any case you will play with the same role for at least 6 rounds.

The matching process will ensure that by the end of the experiment you play as representative and as citizen the **same number of rounds**.

Recall that although your role will remain the same in blocks of 6 rounds, you will be randomly rematched with another participant in **every** round.

NEXT

Instructions of Part 2

At the beginning of each round, **each pair receives 16 tokens**. The representative is in charge of administering this amount.

He/She **has to decide** whether to *send* these 16 tokens to the citizen or to *keep* them for themself.

If the representative *keeps* the tokens, the citizen receives **nothing**.

If the representative *sends* the tokens to the citizen, this amount of tokens is **multiplied by 2** with **50% chance** and the citizen receives **32 tokens**. With the remaining **50% chance** the tokens *sent* by the representative are **reduced to 0** and the citizen receives **nothing**.

Notice that the citizen may receive 0 tokens either because the representative decided to *keep* the tokens to themself or because the representative *sent* the tokens but they were reduced to 0 by chance.

NEXT

Instructions of Part 2

When the representative makes his/her choice, he/she **does not know** whether the citizen will receive 32 tokens or nothing if the representative sends the tokens, as this is **randomly determined** by the computer.

The citizen **cannot observe the decision** of the representative, only the outcome, i.e. whether he/she receives 0 or 32 tokens.

Hence, when a citizen receives 0 tokens, they cannot know for certain whether the representative *kept* or *sent* the tokens.

The only other information **the citizen will receive** is whether the representative is a Kandinsky or a Klee. **The representative will not know** whether the citizen is a Kandinsky or a Klee.

NEXT

Instructions of Part 2

After observing the tokens he/she receives **and whether the representative is a Kandinsky or a Klee**, the citizen must decide whether to *approve* or *not approve* the representative.

If the citizen *approves* the representative, the representative earns **32 additional tokens**.

If the citizen *does not approve* the representative, the representative has **no further earnings** in that round.

After the citizen makes this choice **the round ends**.

NEXT

Instructions of Part 2

After the last round you will be asked a **question** related to the game.

If you answer correctly you will receive **4 additional tokens**.

NEXT

Instructions of Part 2

At the end of the experiment, **1 round** will be chosen **randomly** by the computer.

You will receive as a payment the amount of tokens earned in that round. All rounds have the same probability of being chosen.

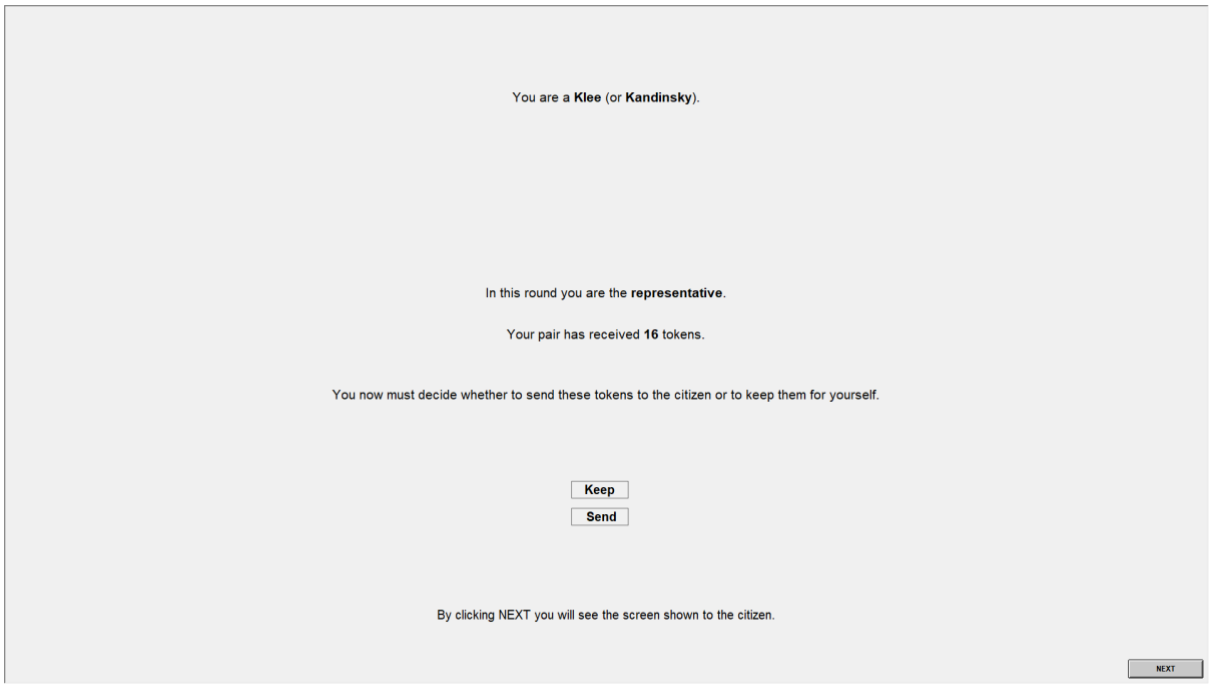
Your **payment from Part 2** will be:

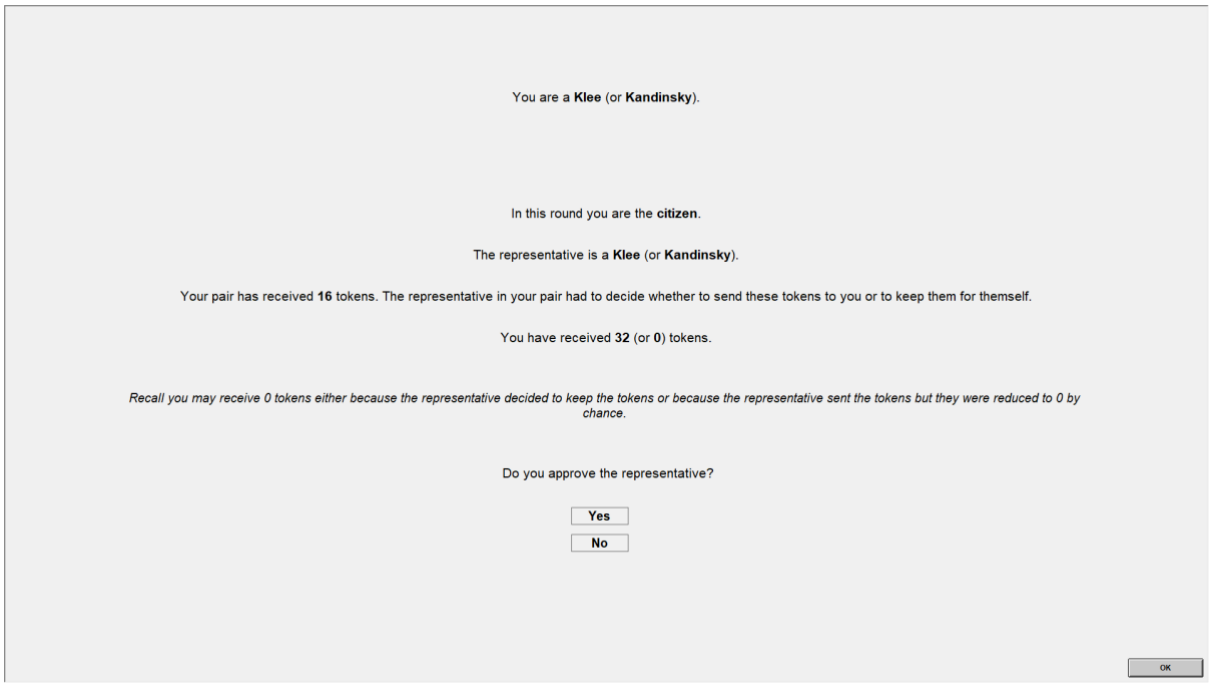
The tokens earned in the randomly chosen round + the tokens earned from answering the last question correctly.

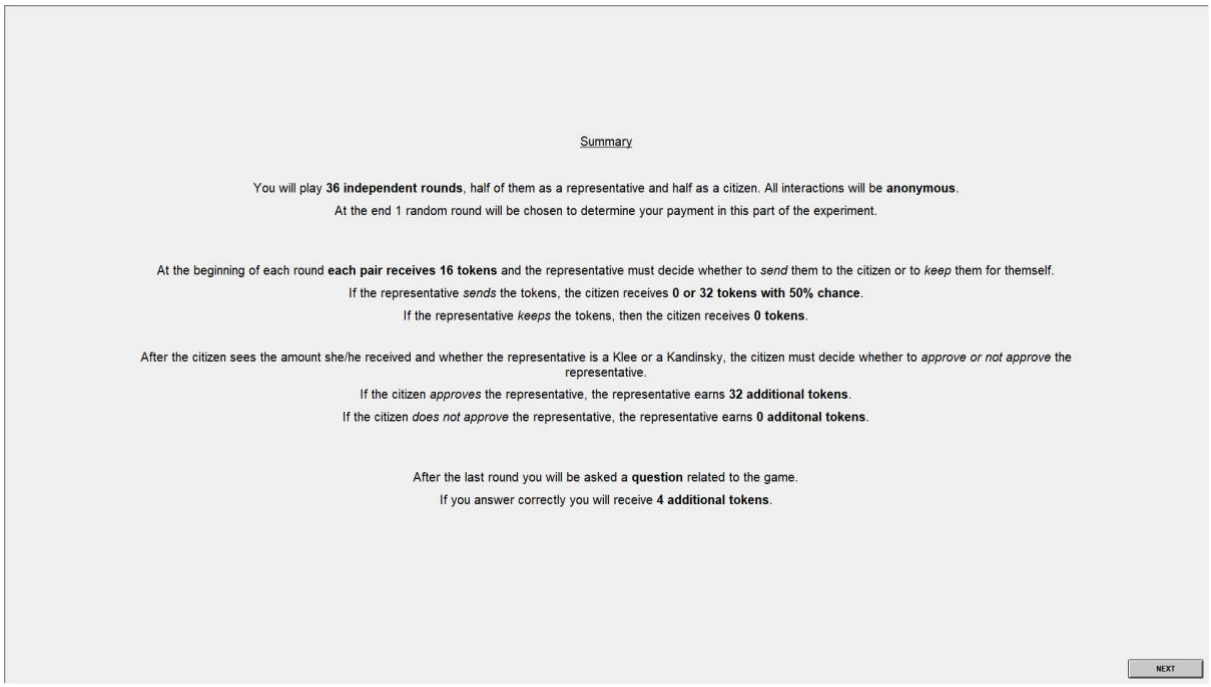
The tokens earned will then be exchanged for money, 2 tokens = £1.

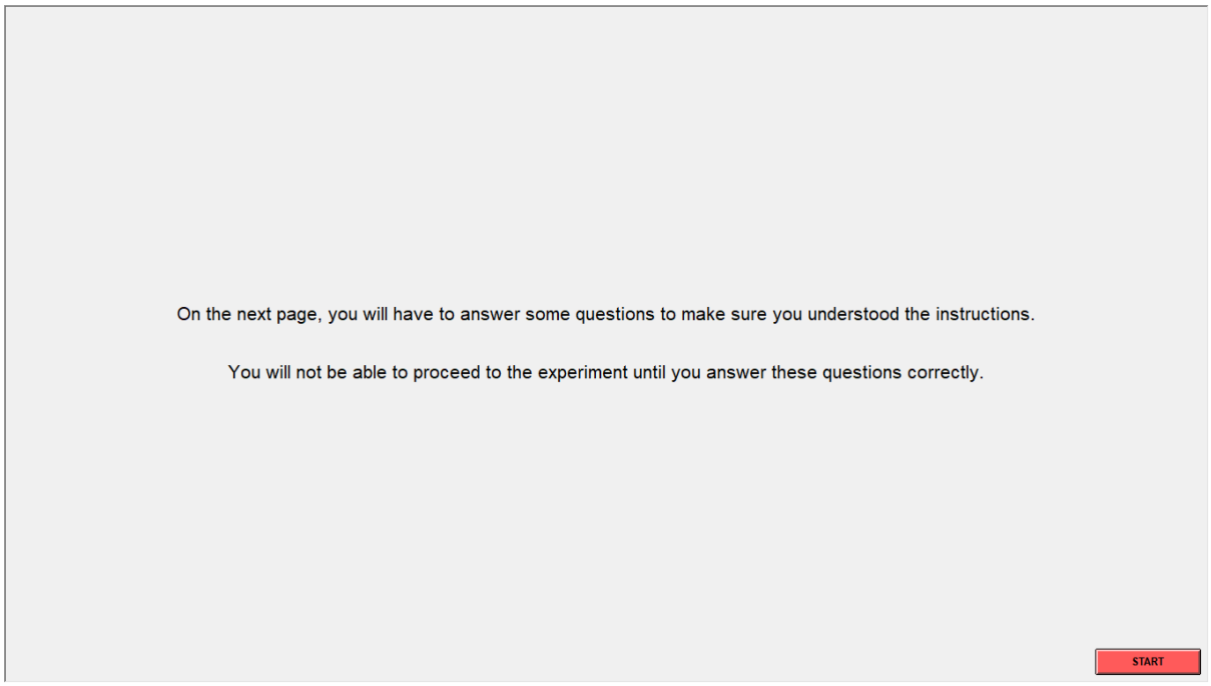
By clicking NEXT, you will see an example in order to familiarise yourself with the interface of Part 2.

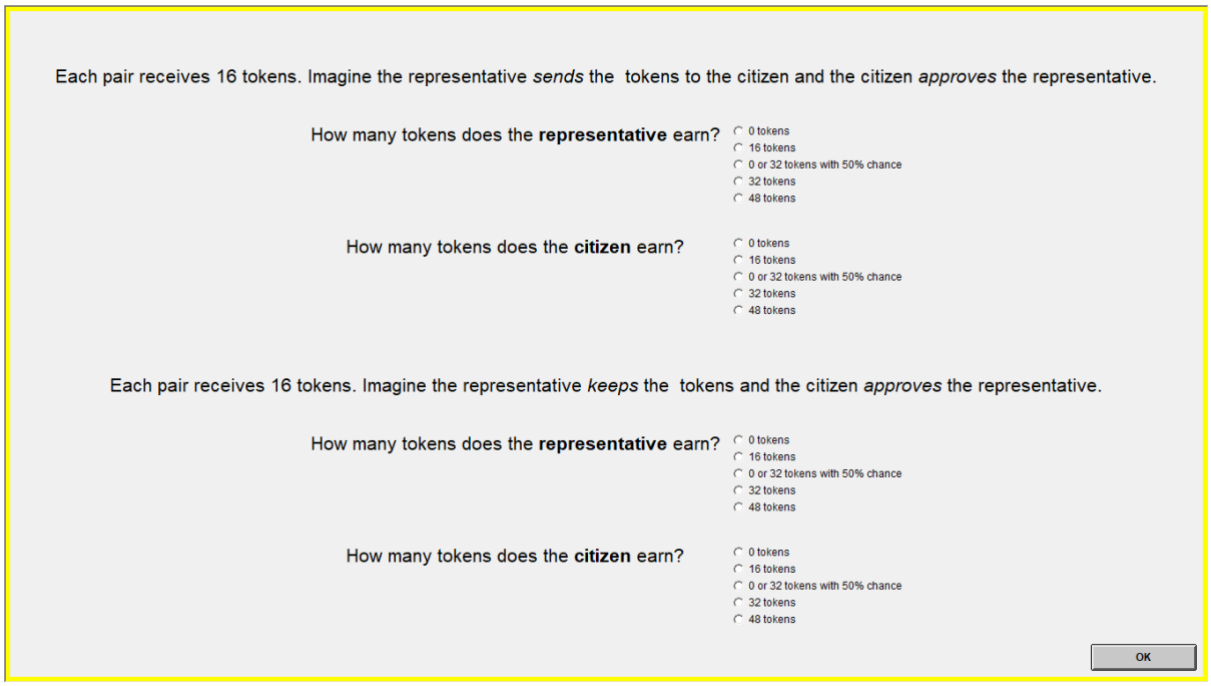
NEXT

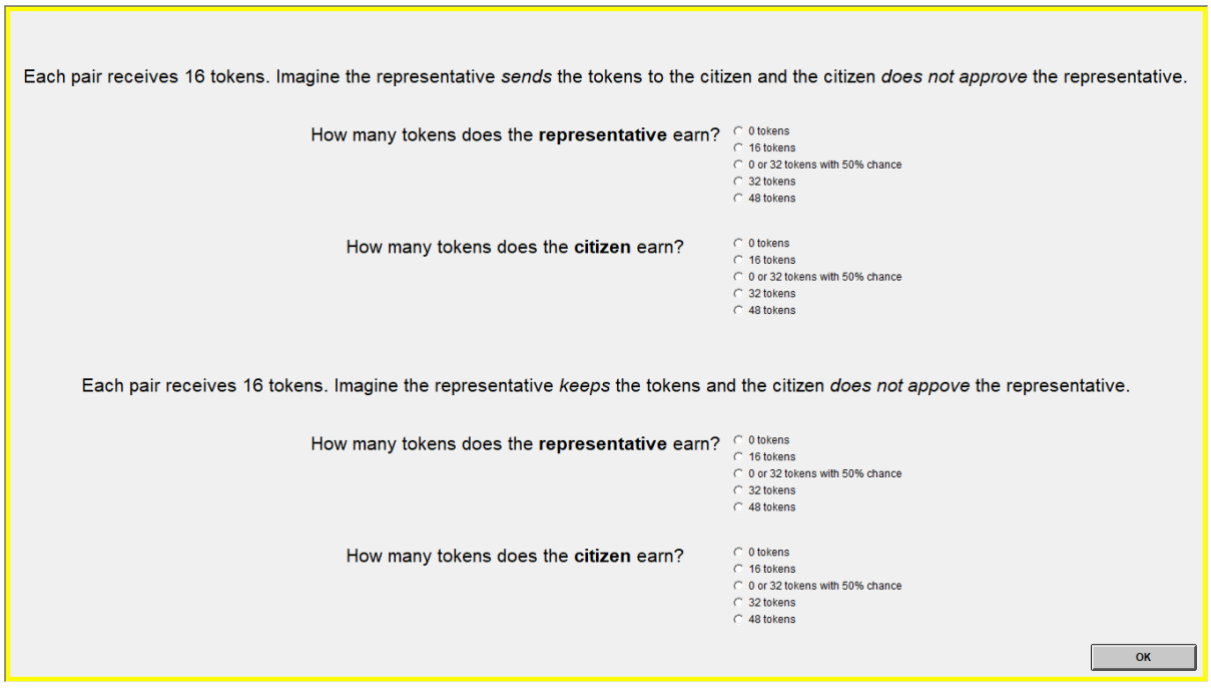

You are a **Klee** (or **Kandinsky**).
In this round you are the **representative**.
Your pair has received **16** tokens.
You now must decide whether to send these tokens to the citizen or to keep them for yourself.
Keep
Send
By clicking NEXT you will see the screen shown to the citizen.
NEXT

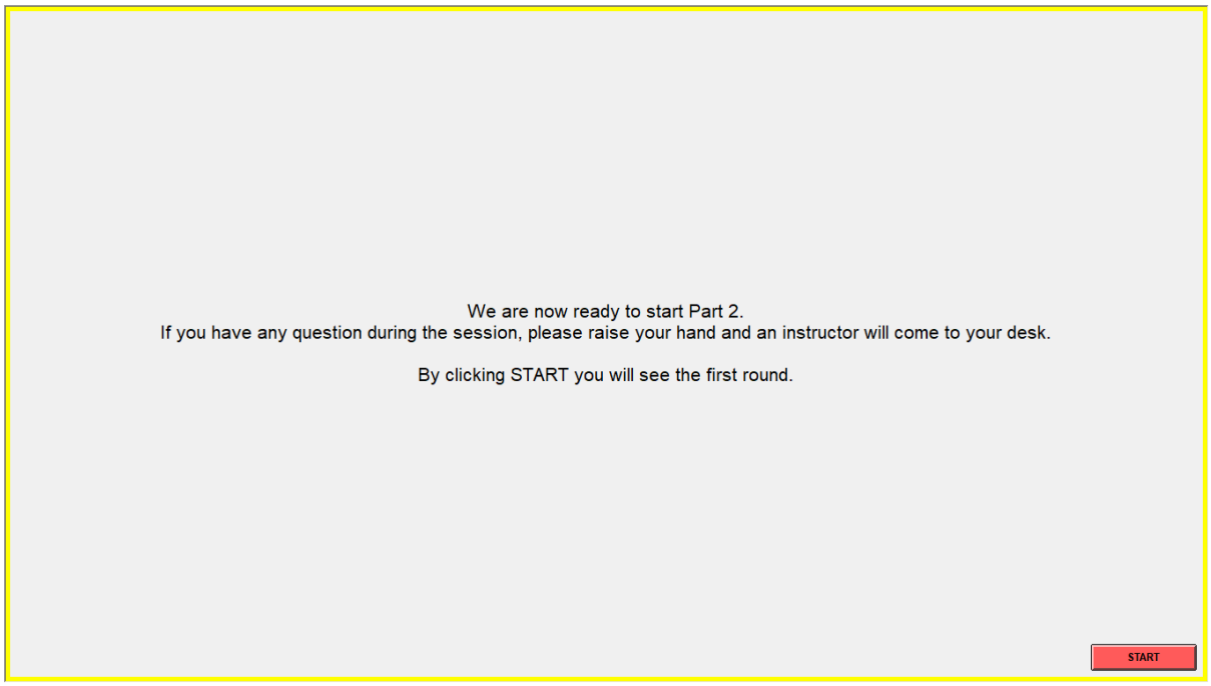

You are a **Klee** (or **Kandinsky**).
In this round you are the **citizen**.
The representative is a **Klee** (or **Kandinsky**).
Your pair has received **16** tokens. The representative in your pair had to decide whether to send these tokens to you or to keep them for themself.
You have received **32** (or **0**) tokens.
Recall you may receive 0 tokens either because the representative decided to keep the tokens or because the representative sent the tokens but they were reduced to 0 by chance.
Do you approve the representative?
Yes
No
OK


Summary
You will play 36 independent rounds, half of them as a representative and half as a citizen. All interactions will be anonymous.
At the end 1 random round will be chosen to determine your payment in this part of the experiment.
At the beginning of each round each pair receives 16 tokens and the representative must decide whether to send them to the citizen or to keep them for themself.
If the representative sends the tokens, the citizen receives 0 or 32 tokens with 50% chance.
If the representative keeps the tokens, then the citizen receives 0 tokens.
After the citizen sees the amount she/he received and whether the representative is a Klee or a Kandinsky, the citizen must decide whether to approve or not approve the representative.
If the citizen approves the representative, the representative earns 32 additional tokens.
If the citizen does not approve the representative, the representative earns 0 additonal tokens.
After the last round you will be asked a question related to the game.
If you answer correctly you will receive 4 additional tokens.
NEXT


On the next page, you will have to answer some questions to make sure you understood the instructions.
You will not be able to proceed to the experiment until you answer these questions correctly.
START


Each pair receives 16 tokens. Imagine the representative sends the tokens to the citizen and the citizen approves the representative.
How many tokens does the representative earn?
0 tokens
16 tokens
0 or 32 tokens with 50% chance
32 tokens
48 tokens
How many tokens does the citizen earn?
0 tokens
16 tokens
0 or 32 tokens with 50% chance
32 tokens
48 tokens
Each pair receives 16 tokens. Imagine the representative keeps the tokens and the citizen approves the representative.
How many tokens does the representative earn?
0 tokens
16 tokens
0 or 32 tokens with 50% chance
32 tokens
48 tokens
How many tokens does the citizen earn?
0 tokens
16 tokens
0 or 32 tokens with 50% chance
32 tokens
48 tokens
OK


Each pair receives 16 tokens. Imagine the representative sends the tokens to the citizen and the citizen does not approve the representative.
How many tokens does the representative earn?
0 tokens
16 tokens
0 or 32 tokens with 50% chance
32 tokens
48 tokens
How many tokens does the citizen earn?
0 tokens
16 tokens
0 or 32 tokens with 50% chance
32 tokens
48 tokens
Each pair receives 16 tokens. Imagine the representative keeps the tokens and the citizen does not appove the representative.
How many tokens does the representative earn?
0 tokens
16 tokens
0 or 32 tokens with 50% chance
32 tokens
48 tokens
How many tokens does the citizen earn?
0 tokens
16 tokens
0 or 32 tokens with 50% chance
32 tokens
48 tokens
OK


We are now ready to start Part 2.
If you have any question during the session, please raise your hand and an instructor will come to your desk.
By clicking START you will see the first round.
START

Bibliography

Acemoglu, D., G. Egorov, and K. Sonin (2013). "A Political Theory of Populism". *Quarterly Journal of Economics* 128.2, pp. 771–805.

Acemoglu, D., G. Egorov, and K. Sonin (2015). "Political Economy in a Changing World". *Journal of Political Economy* 123.5, pp. 1038–1086.

Acemoglu, D. and J. A. Robinson (2005). *Economic Origins of Dictatorship and Democracy*. Cambridge University Press.

Akerlof, G. A. and R. E. Kranton (2000). "Economics and Identity". *Quarterly Journal of Economics* CXV.3, pp. 715–753.

Alesina, A. and G. M. Angeletos (2005). "Fairness and redistribution". *American Economic Review* 95.4, pp. 960–980.

Altomonte, C., G. Gennaro, and P. Francesco (2019). "Collective Emotions and Protest Vote". *CESifo Working Paper* 7463.

Amir, R. (1996). "Cournot oligopoly and the theory of supermodular games". *Games and Economic Behavior* 15.2, pp. 132–148.

Amir, R. (2005). "Supermodularity and complementarity in economics: An elementary survey". *Southern Economic Journal*, pp. 636–660.

Anduiza, E., A. Gallego, and J. Muñoz (2013). "Turning a Blind Eye: Experimental Evidence of Partisan Bias in Attitudes Toward Corruption". *Comparative Political Studies* 46.12, pp. 1664–1692.

Ansolabehere, S. and J. M. Snyder (2000). "Valence Politics and Equilibrium in Spatial Election Models". *Public Choice* 103.3, pp. 327–336.

Aragones, E. and T. R. Palfrey (2002). "Mixed Equilibrium in a Downsian Model with a Favored Candidate". *Journal of Economic Theory* 103, pp. 131–161.

Aragonès, E. and D. Xefteris (2017). "Voters' Private Valuation of Candidates' Quality". *Journal of Public Economics* 156, pp. 121–130.

Austen-Smith, D. and M. Wallerstein (2006). "Redistribution and Affirmative Action". *Journal of Public Economics* 90.10-11, pp. 1789–1823.

Bågenholm, A. (2013). "Throwing the Rascals Out? The Electoral Effects of Corruption Allegations and Corruption Scandals in Europe 1981–2011". *Crime, Law and Social Change* 60.5, pp. 595–609.

Barnett, V. E. (2015). "To My Dear Friend Of Many Years–Klee And Kandinsky's Works On Paper, 1911-1937". In: *Klee & Kandinsky: Neighbors, Friends, Rivals*. Munich, Germany: Prestel Verlag.

Benabou, R. (2000). "Unequal Societies: Income Distribution and the Social Contract". *American Economic Review* 90.1, pp. 96–129.

Berganza, J. C. (2000). "Two Roles for Elections : Disciplining the Incumbent and Selecting a Competent Candidate". *Public Choice* 105.1, pp. 165–193.

Besley, T. and S. Coate (2003). "Centralized versus Decentralized Provision of Local Public Goods: A Political Economy Approach". *Journal of Public Economics* 87.12, pp. 2611–2637.

Black, D. (1948). "On the Rationale of Group Decision-Making". *Journal of Political Economy* 56.1, pp. 23–34.

Brunnermeier, M., A. Simsek, and W. Xiong (2014). "A Welfare Criterion for Models with Distorted Beliefs". *The Quarterly Journal of Economics*, pp. 1753–1797.

Caselli, F. et al. (2014). "The Incumbency Effects of Signalling". *Economica* 81.323, pp. 397–418.

Charness, G. and Y. Chen (2020). "Social Identity, Group Behavior, and Teams". *Annual Review of Economics* 12.

Charness, G., L. Rigotti, and A. Rustichini (2007). "Individual Behavior and Group Membership". *American Economic Review* 97.4, pp. 1340–1352.

Chen, Y. and S. X. Li (2009). "Group Identity and Social Preferences". *American Economic Review* 99.1, pp. 431–57.

Cornaglia, F., M. Drouvelis, and P. Masella (2019). "Competition and the Role of Group Identity". *Journal of Economic Behavior & Organization* 162, pp. 136–145.

Coughlin, P. and S. Nitzan (1981). "Electoral Outcomes with Probabilistic Voting and Nash Social Welfare Maxima". *Journal of Public Economics* 15, pp. 113–121.

D'Aspremont, C., J. J. Gabszewicz, and J.-F. Thisse (1979). "On Hotelling's" Stability in competition"". *Econometrica: Journal of the Econometric Society*, pp. 1145–1150.

De Mello, L. and E. R. Tiongson (2006). "Income Inequality and Redistributive Government Spending". *Public Finance Review* 34.3, pp. 282–305.

Dohmen, T. et al. (2011). "Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences". *Journal of the European Economic Association* 9.3, pp. 522–550.

Dotti, V. (2020). "Income Inequality, Size of Government, and Tax Progressivity: A Positive Theory". *European Economic Review* 121, p. 103327.

Downs, A. (1957). *An Economic Theory of Democracy*. New York: Harper and Row.

Eggers, A. C. (2014). "Partisanship and Electoral Accountability: Evidence from the UK Expenses Scandal". *Quarterly Journal of Political Science* 9.4, pp. 441–472.

Eggers, A. and A. Spirling (2014). "The Advantages and Disadvantages of Incumbency". working paper.

Enelow, J. M. and M. J. Hinich (1982). "Nonspatial Candidate Characteristics and Electoral Competition". *The Journal of Politics* 44.1, pp. 115–130.

Erikson, R. S. (1971). "Advantage of Incumbency in Congressional Elections". *Polity* 3.3, pp. 395–405.

Ferejohn, J. (1986). "Incumbent Performance and Electoral Control". In: *Carnegie Conference on Political Economy*. Vol. 50. 1, pp. 5–25.

Fernández-Vázquez, P., P. Barberá, and G. Rivero (2016). "Rooting Out Corruption or Rooting for Corruption? The Heterogeneous Electoral Consequences of Scandals". *Political Science Research and Methods* 4.2, pp. 379–397.

Fischbacher, U. (2007). "z-Tree: Zurich Toolbox for Ready-Made Economic Experiments". *Experimental Economics* 10.2, pp. 171–178.

Galasso, V. (2003). "Redistribution and fairness: A note". *European Journal of Political Economy* 19.4, pp. 885–892.

Gans, J. S. and M. Smart (1996). "Majority voting with single-crossing preferences". *Journal of Public Economics* 59.2, pp. 219–237.

Gayer, G. et al. (2015). "Pareto Efficiency with Different Beliefs". *Journal of Legal Studies* 43.S2, S151–S171.

Gelman, A. and G. King (1990). "Estimating Incumbency Advantage without Bias". *American Journal of Political Science* 34.4, pp. 1142–1164.

Gilboa, I., L. Samuelson, and D. Schmeidler (2014). "No-Betting-Pareto Dominance". *Econometrica* 82.4, pp. 1405–1442.

Golden, M. A. (2010). "Some Puzzles of Political Corruption in Modern Advanced Democracies". In: *In Democracy and Accountability: Globalized Political Responsibility. Toyko: Fukosha*, pp. 184–199.

Gouveia, M. and N. A. Masia (1998). "Does the Median Voter Model Explain the Size of Government?: Evidence from the States". *Public Choice* 97.1-2, pp. 159–177.

Gründler, K. and N. Potrafke (2019). "Corruption and Economic Growth: New Empirical Evidence". *European Journal of Political Economy* 60, p. 101810.

Guiso, L. et al. (2017). "Populism: Demand and Supply". CEPR Discussion Paper No. DP11871.

Guriev, S. and E. Papaioannou (2020). "The political economy of populism". *Available at SSRN 3542052*.

Hernandez, E. and H. Kriesi (2016). "The Electoral Consequences of the Financial and Economic Crisis in Europe". *European Journal of Political Research* 55.2, pp. 203–224.

Hotelling, H. (1929). "Stability in competition". *Economic Journal* 39.153, pp. 41–57.

Karakas, L. D. and D. Mitra (2020). "Inequality, Redistribution and the Rise of Outsider Candidates". *Games and Economic Behavior* 124, pp. 1–16.

Kartik, N. and R. V. Weelden (2019). "Reputation Effects and Incumbency (Dis)Advantage". *Quarterly Journal of Political Science* 14.2, pp. 131–157.

Kim, J.-H. J. (2012). "Assessing Welfare in Models with Heterogeneous Beliefs". *SSRN Electronic Journal*.

Klor, E. F. and M. Shayo (2010). "Social Identity and Preferences over Redistribution". *Journal of Public Economics* 94.3-4, pp. 269–278.

Kreps, D. (2012). *Microeconomic Foundations I: Choice and Competitive Markets*. Princeton, NJ: Princeton University Press.

Landa, D. and D. Duell (2015). "Social Identity and Electoral Accountability". *American Journal of Political Science* 59.3, pp. 671–689.

Lane, T. (2016). "Discrimination in the Laboratory: A Meta-Analysis of Economics Experiments". *European Economic Review* 90, pp. 375–402.

Levy, G. (2004). "A Model of Political Parties". *Journal of Economic Theory* 115.2, pp. 250–277.

Levy, G. and R. Razin (2015). "Preferences Over Equality in the Presence of Costly Income Sorting". *American Economic Journal: Microeconomics* 7.2, pp. 308–37.

Levy, G., R. Razin, and A. Young (2019). "Misspecified Politics and the Recurrence of Populism". working paper.

Lizzeri, A. and N. Persico (2001). "The Provision of Public Goods under Alternative Electoral Incentives". *The American Economic Review* 91.1, pp. 225–239.

Lockwood, B. (2002). "Distributive Politics and the Costs of Centralization". *The Review of Economic Studies* 69.2, pp. 313–337.

Mauro, P. (1995). "Corruption and Growth". *The Quarterly Journal of Economics* 110.3, pp. 681–712.

Meltzer, A. H. and S. F. Richard (1981). "A Rational Theory of the Size of Government". *The Journal of Political Economy* 89.5, pp. 914–927.

Meltzer, A. H. and S. F. Richard (1983). "Tests of a Rational Theory of the Size of Government". *Public Choice* 41, pp. 403–418.

Méon, P.-G. and K. Sekkat (2005). "Does Corruption Grease or Sand the Wheels of Growth?" *Public Choice* 122.1-2, pp. 69–97.

Milanovic, B. (2000). "The Median-Voter Hypothesis, Income Inequality, and Income Redistribution: An Empirical Test with the Required Data". *European Journal of Political Economy* 16.3, pp. 367–410.

Panunzi, F., N. Pavoniz, and G. Tabellini (2020). "Economic Shocks and Populism: The Political Implications of Reference-Dependent Preferences". CESifo Working Paper.

Persson, T. and G. E. Tabellini (2002). *Political Economics: Explaining Economic Policy*. MIT press.

Persson, T. and G. Tabellini (1994). "Is Inequality Harmful for Growth?" *The American Economic Review* 84.3, pp. 600–621.

Persson, T. and G. Tabellini (1999). "The Size and Scope of Government:: Comparative Politics with Rational Politicians". *European Economic Review* 43.4-6, pp. 699–735.

Rahn, W. M. (1993). "The Role of Partisan Stereotypes in Information Processing about Political Candidates". *American Journal of Political Science*, pp. 472–496.

Razin, A., E. Sadka, and P. Swagel (2002). "The Aging Population and the Size of the Welfare State". *Journal of Political Economy* 110.4, pp. 900–918.

Reed, S. R. et al. (1996). "Political Corruption in Japan." *International Social Science Journal* 48.3, pp. 395–405.

Roberts, K. W. S. (1977). "Voting over Income Tax Schedules". *Journal of Public Economics* 8.3, pp. 329–340.

Roemer, J. E. (1998). "Why the poor do not expropriate the rich: an old argument in new garb". *Journal of Public Economics* 70.3, pp. 399–424.

Romer, T. (1975). "Individual Welfare, Majority Voting, and the Properties of A Linear Income Tax". *Journal of Public Economics* 4.2, pp. 163–185.

Rundquist, B. S., G. S. Strom, and J. G. Peters (1977). "Corrupt Politicians and Their Electoral Support: Some Experimental Observations". *American Political Science Review* 71.3, pp. 954–963.

Solaz, H., C. E. De Vries, and R. A. de Geus (2019). "In-Group Loyalty and the Punishment of Corruption". *Comparative Political Studies* 52.6, pp. 896–926.

Stiglitz, J. (2000). *Economics of the Public Sector*. WW Norton & Company Ltd.

Taber, C. S. and M. Lodge (2006). "Motivated Skepticism in the Evaluation of Political Beliefs". *American Journal of Political Science* 50.3, pp. 755–769.

Tajfel, H. and J. Turner (1986). "The Social Identity Theory of Intergroup Behavior". In: *The Psychology of Intergroup Relations*. Ed. by S. Worchel and W. Austin. Chicago: Nelson Hall: Chicago: Nelson Hall, pp. 7–24.

Tajfel, H. et al. (1971). "Social Categorization and Intergroup Behaviour". *European Journal of Social Psychology* 1.2, pp. 149–178.

Tajfel, H. et al. (1979). "An Integrative Theory of Intergroup Conflict". *Organizational Identity: A Reader* 56, p. 65.

Torcal, M. (2014). "The Incumbent Electoral Defeat in the 2011 Spanish National Elections: The Effect of the Economic Crisis in an Ideological Polarized Party System". *Journal of Elections, Public Opinion and Parties* 24.2, pp. 203–221.

Transparency International (2019). *Corruption Perceptions Index 2019*. Tech. rep.

Uppal, Y. (2009). "The Disadvantaged Incumbents: Estimating Incumbency Effects in Indian State Legislatures". *Public Choice* 138.1-2, pp. 9–27.

Welch, S. and J. R. Hibbing (1997). "The Effects of Charges of Corruption on Voting Behavior in Congressional Elections, 1982-1990". *The Journal of Politics* 59.1, pp. 226–239.

Woon, J. (2012). "Democratic Accountability and Retrospective Voting: A Laboratory Experiment". *American Journal of Political Science* 56.4, pp. 913–930.

Wright, R. (1986). "The redistributive roles of unemployment insurance and the dynamics of voting". *Journal of Public Economics* 31.3, pp. 377–399.

Xiong, W. (2013). "Bubbles, Crises, and Heterogeneous Beliefs". NBER Working Paper Series 18905.