CESifo WORKING PAPERS

8539
2020

Original Version: September 2020
This Version: September 2021

Economic Shocks and Populism: The Political Implications of Reference-Dependent Preferences

Fausto Panunzi, Nicola Pavoni, Guido Tabellini

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 8539*

Economic Shocks and Populism:
The Political Implications of Reference-Dependent Preferences

Abstract

This paper studies electoral competition over redistributive taxes between a safe incumbent and a risky opponent. As in prospect theory, economically disappointed voters become risk lovers, and hence are attracted by the more risky candidate. We show that, after a large adverse economic shock, the equilibrium can display policy divergence: the intrinsically more risky candidate proposes lower taxes and is supported by a coalition of very rich and very disappointed voters, while the safe candidate proposes higher taxes. This can explain why new populist parties are often supported by economically dissatisfied voters and yet they run on economic policy platforms of low redistribution. We show that survey data on the German SOEP are consistent with our theoretical predictions on voters' behavior.

JEL-Codes: H000, D700, D900.

Keywords: populism, prospect theory, behavioural political economics.

Fausto Panunzi
Department of Economics and IGIER
Bocconi University / Milan / Italy
fausto.panunzi@unibocconi.it

Nicola Pavoni
Department of Economics and IGIER
Bocconi University / Milan / Italy
nicola.pavoni@unibocconi.it

Guido Tabellini
Department of Economics and IGIER
Bocconi University / Milan / Italy
guido.tabellini@unibocconi.it

September 2021
We are grateful to Jeff Frieden, Edoardo Grillo, Salvo Nunnari, Mattias Polborn, Ronny Razin, and participants at workshops at CEPR, CIFAR, CSEF, IMF, LUISS University, University ”Federico II” Naples, UPF, the 2018 Summer Symposium on Financial Markets (Gerzensee) for helpful comments, to Giampaolo Bonomi, Sergio Cappellini, Niccoló Cattadori, Silvia Farina and Francesco Filippucci for excellent research assistance and to ERC grant 741643 for financial support.

1 Introduction

The party systems of several Western democracies have been transformed by the rise of populism. Mainstream centrist and socialdemocratic parties have lost support, while new populist politicians have gained popularity and in some countries they have shaped government policy. Several empirical papers have shown that this transformation is associated with adverse economic shocks and economic insecurity. Survey evidence shows that individuals who feel economic insecure or fear a loss of social status are more likely to vote for populist parties (Guiso et al. 2017, Gidron and Hall 2017, Dal Bo et al. 2018). Communities that have suffered large employment losses due to import competition from China support populist parties and have become more polarized (Colantone and Stanig 2018, Autor 2020). These recent phenomena are also consistent with historical evidence showing that large recessions induced by financial crisis are associated with increased polarization and a shift towards new radical right wing parties (Mian et al. 2014, Funke et al. 2016). The survey by Guriev and Papaioannou (2020) cites other relevant contributions.

This association between populism and economic adversity is puzzling, because several populist parties or politicians support right wing policy platforms of tax cuts and welfare state retrenchments that benefit the rich but seem to run counter the interests of the weakest segments of society. And yet, populist parties have gained vote shares irrespective of their stance on redistribution - see Appendix Figure 6. In countries such as the US, the UK or Italy, populist politicians and populist policy platforms are often supported by an unwieldy coalition of rich constituencies and of angry and disappointed low-middle class voters. Why is this happening, and what are the implications of these phenomena?

In this paper we answer this question by taking the insights of prospect theory to political economics. A feature of the new populist politicians is that they are much more risky than their mainstream counterparts. This is not just because they are new and untested. Populist politicians often reject external constraints and checks and balances limiting government discretion (Mueller 2017, Rode and Revuelta 2015), and support more radical and unconventional policies.[1] This intrinsic riskiness makes populist politicians partic-

[1]In emerging countries, populist governments often enacted overly expansionary macoeconomic policies that ultimately led to balance of payments crisis and economic collapse (Dornbusch and Edwards 1991, Edwards 2019, Herrera et al. 2019). Only a few Western democracies have had populist governments, but in Italy the uncertainty associated with the rise of populist parties was reflected in Credit Default Swaps

ularly attractive to disappointed voters, who have already lost of a lot relative to their expectations and therefore attach a large value to the chance of getting back towards their reference point. For the same reason, instead, they are opposed by voters whose income is close to what they expected, and who are particularly fearful of risk because of their loss aversion.

This intuition is illustrated in Figure 1, that depicts the typical shape of preferences over consumption implied by prospect theory. Point x on the horizontal axis is the reference point against which consumption is evaluated. For sufficiently high levels of consumption, the agent is risk neutral.[2] If consumption is in the neighborhood of the reference point, the agent becomes risk averse, because he is averse to losses that take him below the reference point. Finally, if consumption is much below the reference point, the agent becomes risk lover. In the paper we assume that preferences for risk increase with disappointment, i.e., the lower is c relative to x, the more risk is loved.[3]

The idea that populist politicians are attractive to some voters also because they are risky and untested is supported by survey evidence. Exploiting data from the German Socio Economic Panel, in section 2 we show that individuals who become highly disappointed with their economic situation also become very risk lovers and tend to switch towards populist and radical parties.

The rest of this paper then explores the implications of this idea in a simple model of electoral competition, where voters' preferences take the shape of Figure 1. A safe (moderate) and a risky and less efficient (populist) politician compete in elections over redistributive taxation. Suppose that the economy is hit by a large negative aggregate shock, so that several voters fall below their consumption reference point. These very disappointed voters become risk loving and (for the same policy) they prefer the risky politician to the safe one, while voters close to their reference point react in the opposite way (they become risk averse and lean towards the moderate candidate). How do these

(Balduzzi et al. 2019) and in comments in the press: *"Italy is on the brink of installing the most unconventional, inexperienced government to rule a western European democracy since the EU's founding Treaty of Rome in 1957"*, Rome opens its gates to the modern barbarians, Financial Times, May 14, 2018. *"By abandoning traditional parties, unable to meet its needs, Southern Italy has declared its willingness to experiment something that it is not known yet and that could be better than the status quo. The opposite of risk aversion"*. Drago and Reichlin, Corriere della Sera, March 13, 2018.

[2]This is done for simplicity. As discussed in subsection 4.5, we could assume decreasing risk aversion for consumption levels above the reference point, and all relevant results in the paper would still hold.

[3]We neglect another implication of Prospect theory, namely probability weighting (decision weights distort probabilities towards 1/2).

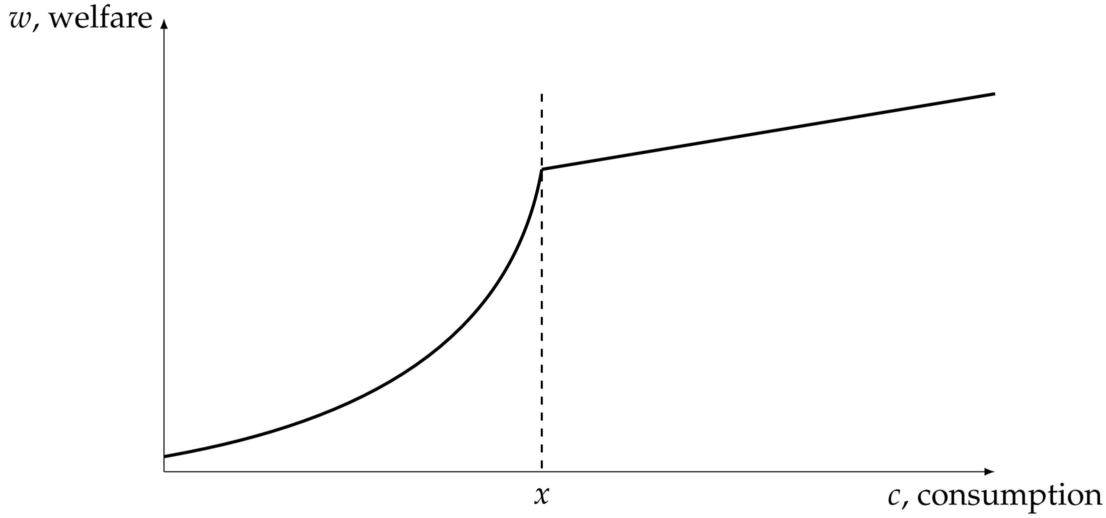

Figure 1: An example of the voters' preferences over consumption. Point x on the horizontal axis is the reference point against which consumption is evaluated. Welfare changes linearly to the right of x; to the left of the reference point, preferences become convex.

office seeking politicians respond to this situation, and what policy platforms do they announce?

The answer is not straightforward, because heterogeneous candidates and reference dependent preferences add a second dimension of political conflict, on the intrinsic riskiness of alternative candidates, besides the traditional conflict over redistribution. This generates a particular non-monotonicity in how voters with different levels of income evaluate their voting options, and this matters for the redistributive platforms that are announced in equilibrium. We show that political equilibria can take two forms. One possibility is an equilibrium with policy convergence, where both parties compete for the same marginal voter and choose the same tax policy. But if the economy is hit by a large negative shock, a second possibility is that the equilibrium displays policy divergence. The populist candidate runs on a policy platform of lower taxation and lower redistribution, compared to the moderate candidate, and he is supported by a coalition of rich voters and of poorer and very disappointed voters.

Equilibrium policy divergence is driven by the fact that the populist candidate has a stronger incentive to cut taxes below the level set by his opponent, compared to the moderate candidate. The reason has to do with the intrinsic riskiness of the populist candidate. Consider a unilateral deviation from the same hypothetical tax rate. A lower tax rate, by making poor voters worse off, also increases their disappointment, and this makes them lean towards the more risky candidate. Hence, poor voters do not punish the populist candidate for a tax cut by as much as they would punish the safe candidate for the same deviation.

The equilibrium is more likely to entail policy divergence if a large negative shock pushes a significant fraction of the electorate below its reference point. In this case, many middle-class voters become moderately disappointed, and only a sufficiently large redistributive policy allows the safe candidate to retain them. This in turn induces the populist politician to propose a lower tax rate, building the unwieldy coalition of richer voters and most disappointed voters.

In a concluding section we discuss additional political implications of reference dependent preferences.

Literature. Our paper is related to a rapidly growing literature on populism, recently reviewd by Guriev and Papaioannou (2020) - see also Levy et al. (2020). This literature too seeks to explain why economic distress is associated with support for populist parties, and why voters hurt by adverse economic shocks support right wing policy platforms. A common explanation is that voters care about a second policy dimension, such as cultural issues or immigration policy (eg. Norris and Inglehart 2019, Bonomi et al. 2020, Huber 2017). They are thus prepared to accept less redistribution, in order to get the cultural or immigration policy that they prefer. In our setting, instead, voters only care about future consumption and the policy space is one-dimensional. Reference dependent preferences create a second dimension of heterogeneity in voters' preferences, however, over the risk profile of consumption. This second dimension of political conflict, together with the intrinsic differences in risk between opposing candidates, is responsible for equilibrium policy divergence.

The idea that voters trade off their policy preferences against the intrinsic features of competing candidates has also been explored in a series of papers on differentiated candidates. Groseclose (2001) and Aragones and Palfrey (2002) studied electoral competition between candidates with different valence. Voters have standard policy preferences, and they all prefer the candidate with higher valence. In equilibrium the advantaged candidate moves to the center, while the disadvantaged candidate chooses a more extreme policy.[4] More recently, Krasa and Polborn (2014), (2010), (2012) have shown that policy divergence occurs in equilibrium in a setting where two candidates with different abilities choose a one dimensional policy in order to maximize the probability of winning. Their key assumption is that candidates' abilities are complements to the policy over which they compete. In our paper, instead, candidates differ in their intrinsic risk, and voters' risk preferences vary non-monotonically with their income.

Since the seminal work of Kahneman and Tversky (1979), a very large literature has explored the insights of prospect theory in a variety of economic settings (see the survey by Barberis 2013). Quattrone and Tversky (1988) provide experimental evidence that prospect theory may also explain voters' behavior. Among other observations, they (informally) point out that adverse economic circumstances may induce voters to seek political risks, and this may benefit lesser known challengers and hurt mainstream incumbents. We are not aware of any paper that formally explores the implications of this idea in a setting of electoral competition, where candidates who differ in their intrinsic risk profiles also make other policy choices.[5]

Some recent papers have studied the implications of reference dependent preferences for political behavior. Alesina and Passarelli (2019) and Lockwood and Rockey (2015) study electoral competition when voters are loss averse. Passarelli and Tabellini (2017) study political protests in a setting where citizens become angry if they feel that they are treated unfairly relative to their (endogenous) reference point. Grillo and Prato (2020) and Besley and Persson (2019) assume that endogenous political institutions shape citizens' reference points, and show how this in turn affects political incentives to strategically

[4]Models of valence with office-motivated candidates rarely have pure strategy equilibria. The intuition being that the 'better' candidate wants to copy the worse candidate's policy, while the disadvantaged candidate tries to distance himself from the disadvantaged candidate. The disadvantaged must hence mix in order to be unpredictable. Aragones and Palfrey show that as the policy space becomes large the equilibrium approaches the standard case, with policy convergence towards the (estimated) median voter's ideal point. Groseclose obtain existence of a pure strategy equilibria by assuming that the candidates also differ in their intrinsic preferences over certain policies.

[5]An older literature discusses under what conditions candidates have an incentive to announce risky policy platforms. In particular, Shepsle (1972) pointed out that, if voters are risk loving, then office seeking candidates have an incentive to randomize over extreme policy positions. Alesina and Cukierman (1990) and Aragones and Neeman (2000) show that, if candidates have partisan policy preferences, then they find it optimal to announce ambiguous policy positions, even if voters are risk averse and dislike ambiguity, because this increases their degrees of freedom over subsequent policy choices. More recently, Karakas and Mitra (2020) have studied electoral competition between an "establishment" and an "outsider" candidate that differ in their willingness to change the status quo, but voters' risk preferences do not play a role in their analysis.

manipulate institutions. Binzell and Carvalho (2016) show that economic disappointment (modelled as reference dependence) has led to an Islamic revival in Egypt. None of these papers considers attitudes towards risk.

Reference dependent preferences in choices under uncertainty have been studied in several economic applications, following the seminal work by Koszegi and Rabin (2006), (2007). Our model of voters' behavior uses a similar approach, although rational and self-fulfilling expectations do not play a central role in our static setting.

A large empirical literature has studied the determinants of attitudes towards risk, using survey or experimental evidence (see for instance Dohmen 2015, Dohmen et al. 2011). This literature often finds that risk aversion tends to increase during economic downturns, but there is also evidence that becoming unemployed is associated with increased preference for risk (Dohmen et al. 2016).

Finally, limited liability may have similar implications for risk taking as prospect theory: if losses have a lower bound while upside risk is fully internalized, this may induce a preference for more risk. This idea has been studied in corporate finance (e.g, Jensen and Meckling 1976). We are not aware of applications of limited liability to a political setting, although by the same logic, individuals who have nothing more to loose (and perhaps are cushioned by the welfare state) are induced to take political risks. Some of our general considerations might hence apply to this case as well. At the same time, the notion of reference dependent preferences that we study in this paper differs from limited liability as preferences for risk depend on losses relative to a reference point. Hence, even middle or high income individuals may turn risk loving, if they are economically very disappointed.

The outline of the paper is as follows. In Section 2 we study the German SOEP. Section 3 lays out the theoretical model of reference dependent preferences in a political setting. Section 4 studies electoral competition between a risky and a safe candidate over a redistributive tax, and derives the conditions for an equilibrium where the more risky candidate promises less redistribution and is supported by a coalition of disappointed and rich voters. Section 5 discusses other implications of our results.

2 Empirical Motivation

In this section we provide some empirical evidence that economically very disappointed voters and highly risk loving voters are more likely to vote for a populist politician. In a nutshell, the theory developed in the remaining sections is built on the following implications of prospect theory. An individual who suffers an unexpected and large income loss is forced to consume below his reference point (in this sense he is disappointed). This makes him become risk loving and as a result he leans towards populist candidates who are perceived as more risky. Note that this prediction applies to individuals who are far below their reference point. Individuals who are close to their reference point (on both sides) are actually opposed to the populist candidate, because loss aversion makes them particularly fearful of risk (cf. Figure 1). We now show that survey evidence is consistent with these predictions.

2.1 Data

The data come from the core sample of the German Socio-Economic Panel (SOEP). Although political populism is not as prominent in Germany as in other advanced democracies, the SOEP is one of the few longitudinal surveys that includes questions on risk attitudes, political preferences and satisfaction with one's economic situation. We consider observations from 2008 to 2016 (the measure of risk loving that is the focus of this paper is missing for the year 2007 and we discard earlier years when populism was not so widespread). Individuals younger than 18 years of age are not included in our sample.

Populist preferences Our dependent variable of interest is the dummy variable *populist*, that equals 1 if the respondent leans towards a populist party, and 0 otherwise - individuals who do not lean towards any party (about half of the sample) are coded as 0.[6] We follow the literature (Inglehart and Norris 2016 and von Beyme 1988) and define as populist the following (right wing) parties: Alternative fur Deutschland (AfD), Piratenpartei, and the coalition of NPD, DVU, die Republikaner and die Rechte. Note that only a small percentage of voters declares to be leaning towards populist parties (cf. the summary statistics reported Table 1). Support for populist parties is initially below 1%, and it is

[6]We do exclude from the sample respondents answering to a survey in which the question on party preference is missing.

captured only by the far right. The emergence of the Piratenpartei and of AfD increases the share of populist support, that in our sample reaches 2.6% in 2016. As a fraction of those who declare a political preference, these numbers are doubled, but still low relative to observed vote shares.

Income dissatisfaction In the theoretical model, a key determinant of risk preferences is whether income is above or below an individual's reference point. We don't observe individual reference points, but a question in the survey asks how satisfied is the respondent with his household income. Possible answers range from 0 (completely dissatisfied) to 10 (completely satisfied). We define the variable *income dissatisfaction* as 10 minus the answer to this question. The theory predicts a non-monotonic effect of disappointment: mildly disappointed individuals who are close to their reference point are risk averse, due to loss aversion, while very dissatisfied individuals are risk loving. To capture this non-monotonicity, we thus define a dummy variable *extreme income dissatisfaction* that equals 1 if income dissatisfaction is equal to or above 7 (results are very similar if the threshold for being extremely dissatisfied is higher, or if extreme dissatisfaction, rather than being a dummy variable, is defined as 0 below the threshold, and equal to income dissatisfaction at or above the threshold).

Attitudes towards risk As a measure of risk attitudes, we use the question concerning self-reported risk aversion, namely: "Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks? Please tick a box on the scale, where the value 0 means: "unwilling to take risks" and the value 10 means: "fully prepared to take risk". We call this variable *risk_love*. This question is carefully validated by Dohmen et al. (2011), who find that it is a relevant predictor of actual risk-taking behavior, both in incentivized experiment and in different unincentivized questions on hypothetical lottery participation and on risk-taking in different life domains such as car driving, financial matters, sports and leisure, career, and health. Here too, to account for possible non-linearities in the data and isolate individuals with a high propensity to take risks, we also define the variable *extreme risk love*, a dummy variable that takes value of 1 if the individual is very risk loving, meaning that the variable *risk love* equals either 9 or 10 (the two top values). Again, results are similar if instead extreme risk love is defined as 0 below the threshold, and as equal to risk love at or above the threhold of 9.

Other regressors Income data (*Income*) refer to the household and are measured in logs. Because raw income data report several extreme values, the data are trimmed at 99% and 1%. From the income data, we also construct a dummy variable that identifies individuals who have suffered very large income losses (again, only negative income shocks that are sufficiently large to take the respondent much below his reference point are predicted to increase the propensity to vote populist and to take risks). Specifically, the dummy variable equals 1 if current individual income is at least 26% below the three year moving average of lagged own individual income. The threshold 26% corresponds to the 80th percentile of income losses (defined as log of current income minus the log of the three year moving average of lagged individual income). We define this dummy variables as *Large income loss*.

In several specifications we also control for being male, two dummy variables for age class (under 28 and over 45 years of age, baseline is 28-45), two dummy variables for having reached lower-secondary school and tertiary education as the maximum level of education (the baseline is having completed upper secondary, the median level of education in Germany). In addition, we control for being unemployed and for being out of the labor force (mainly pensioners, inactives and students): the baseline case is thus a partly or fully employed worker. Finally we include dummy variables for being a first and a second generation immigrant and for being a resident of East Germany. Summary statistics of all variables are displayed in Appendix Table A1. In the Appendix Table A2 we display the correlation coefficients amongst the variables of main interest (pooling together all observations); correlation coefficients tend to be low for most variables.

2.2 Estimation strategy

With the help of the variables defined above, we now explore whether survey data are consistent with the following causal chain:

$$\text{Income loss} \Rightarrow \text{Income dissatisfaction} \Rightarrow \text{Risk love} \Rightarrow \text{Populist support}$$

Although we observe all links in this chain, we don't have random exogenous variation in all its elements. Moreover, we cannot use earlier elements of this chain as instruments for later elements, because the exclusion restrictions would not be credible. For instance, income dissatisfaction could have a direct effect on political preferences, besides the effect induced through risk attitudes. At the same time, income dissatisfaction could be an imperfect measure of being below one's reference point, so that large income losses could have a direct impact on risk attitudes, beyond the observed effect on income dissatisfaction. Hence, we cannot estimate causal effects. We can only ask whether observed correlations in the data are consistent with our theoretical hypothesis. Nevertheless, controlling for individual fixed effects allows us to infer correlations from yearly variation at the individual level, which reduces the scope of omitted variables. Moreover, this can be done for all links in the hypothetical causal chain.

We thus estimate regressions of the following type:

$$y_{it} = \alpha_i + \delta_t + \beta T_{it} + \gamma' \mathbf{X}_{it} + \varepsilon_{it}$$

where α_i and δ_t are individual (i) and year (t) fixed effects, y is the outcome of interest, T is the "treatment" of interest, namely one of the prior elements in the causal chain, and $\mathbf{X}$ is a vector of control variables.

We start with the last link of the causal chain, so that y is the dummy variable for leaning towards a populist party, and T is being risk loving and extremely risk loving. We then move to the immediately preceding causal link. Here the treatments T are income dissatisfaction and extreme income dissatisfaction, while the outcome of interest y is attitudes towards risk, or (in a reduced form implication) again the *populist* dummy variable. Note that here the theory predicts non-monotone effects. Being very dissatisfied increases the propensity to vote populist and makes individuals more risk loving, while mild dissatisfaction or moderate satisfaction has the opposite effects due to loss aversion. Finally we estimate the correlations implied by the first link in the causal chain, where the treatment T is having experienced a large income loss, and the outcome variables are all of the subsequent elements of the causal chain, namely being extremely dissatisfied, being extremely risk loving and leaning towards a populist party. In order to isolate the effects of large income shocks from other sources of variation, we always define the outcome of interest y to be an extreme individual feature, such as being extremely risk lover or extremely dissatisfied. Hence the dependent variable y is always a dichotomous dummy variable, irrespective of how it is defined. Estimation is thus by conditional logit throughout, and standard errors are always clustered by individual. The tables also report average marginal effects in square brackets.

2.3 Results

Risk loving and support for populism We start with the last link in the causal chain, and ask how risk attitudes correlate with support for populist parties. Table 1 reports the estimates of logit regressions where the dependent variable is leaning towards a populist party (*populist*). Throughout we include year fixed effects.

Columns 1-4 are estimated on pooled data (over individuals and years) without individual fixed effects. Column (1) illustrates a positive correlation between being risk loving and leaning towards a populist party. The correlation is even stronger for individuals who are extremely risk loving (column 2). Controlling for individual features dampens the correlation with risk loving, but the estimated coefficient on *extreme risk love* increases further and remains highly significant (columns 3 and 4). In the specification of column (4), on average extreme risk lovers are 0.6 percentage points more likely to lean towards a populist parties (on top of the positive effect associated with being risk loving). Recalling that populist supporters are only about 1% of this sample on average, this is a large effect. Note that higher income is associated with not leaning towards populism, while being unemployed has a positive association with populism.

Columns 5 and 6 add individual fixed effects, and hence come closer to estimating a causal effect of attitudes towards risk on populist sympathies. Although the estimated coefficient on *risk love* is no longer statistically different from 0, the estimated coefficient on the dummy variable *extreme risk love* remains positive and significant, and not much smaller than in the pooled regressions. The average marginal effect rises by an order of magnitude. According to the specification in column 6, on average, becoming extremely risk loving increases the probability of leaning towards a populist party by over 7 percentage points (gross of the negative but insignificant effect associated with risk loving).[7] Note that the sample size is much smaller when individual fixed effects are included, and this is responsible for the changes in the estimated coefficients. Estimating the pooled regressions (without individual fixed effects) on the smaller sample also yields large average marginal effects on *extreme risk love*, and similar estimated coefficients to those reported in columns 5 and 6, when fixed effects are included (results available upon request). This remark also applies to the regressions reported below.

[7] The average marginal effect is estimated conditional on the individual fixed effect being 0, i.e. for individuals who on average do not lean towards populism.

Table 1: **Risk Love and Populism**

Dep. var.	Populist Dummy					
	(1)	(2)	(3)	(4)	(5)	(6)
Risk Love	0.1168***	0.0897***	0.0560***	0.0261*	-0.0216	-0.0213
	(0.013)	(0.013)	(0.014)	(0.014)	(0.021)	(0.022)
	[0.0013]	[0.0010]	[0.0006]	[0.0003]	[-0.0038]	[-0.0041]
Extreme Risk Love		0.4769***		0.5414***	0.3210**	0.3340**
		(0.100)		(0.107)	(0.152)	(0.155)
		[0.0054]		[0.0061]	[0.0569]	[0.0636]
Unemployed			0.3879***	0.3741***		-0.0395
			(0.098)	(0.098)		(0.167)
			[0.0043]	[0.0042]		[-0.0075]
Income			-0.2739***	-0.2655***		-0.1779
			(0.051)	(0.051)		(0.119)
			[-0.0031]	[-0.0030]		[-0.0339]
Observations	206,751	206,751	197,138	197,138	7,720	7,520
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Other Controls	No	No	Yes	Yes	No	No
Individual FE	No	No	No	No	Yes	Yes

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered at the individual level, marginal effects in square brackets. In columns 3, 4 and 6 we also control for being in or out the labor force. Other individual controls (where included) are dummy variables for gender, age group, immigrant status, education level, and macroregion. Estimation is by logit in columns 1-4, by conditional logit in columns 5 and 6. Source: SOEP.

Income dissatisfaction and support for populism According to prospect theory, being risk loving is triggered by income losses sufficiently below a reference point - e.g. Koszegi and Rabin (2009) and Pagel (2017). We don't observe the reference point, but we can use the question on income dissatisfaction as a measure of economic disappointment.

We start by estimating a semi-reduced form, and ask whether support for populism is associated with *income dissatisfaction* and *extreme income dissatisfaction*. The estimates are displayed in Table 2, columns 1-4. In column 1 we don't include individual fixed effects. More dissatisfied individuals are more likely to lean towards populist parties, and the two income dissatisfaction variables absorb the effect of the two economic variables (unemployment and income) that unlike in Table 1 are no longer statistically significant. In column 2 we add individual fixed effects. Being extremely dissatisfied remains highly significant, and the marginal effect is large: becoming extremely dissatisfied increases the probability of support for populism by almost 7 percentage points. Income dissatisfaction on its own is no longer statistically significant. Note that, since we are controlling for being extremely dissatisfied, the estimated coefficient on *income dissatisfaction* captures the effect of being only moderately dissatisfied or being satisfied. Hence, the expected coefficient on this variable is indeed 0 or even negative, since individuals close to their reference are predicted to be loss averse.

In columns 3 and 4 we add the two indicators of risk loving to the right hand side, with and without individual fixed effects. Both *extreme income dissatisfaction* and *extreme risk love* remain statistically significant, although with smaller coefficients than when they are entered in isolation (compare columns 4 and 2 in Table 2, and column 4 in Table 2 with column 6 in Table 1). This suggests that becoming very dissatisfied has a direct association with support for populism, besides the possible induced effect on risk attitudes.

Last, we ask whether extreme income dissatisfaction is associated with being risk loving, as predicted by the intermediate link in the causal chain depicted above. The answer is provided by columns 5 and 6 of Table 2, where the dependent variable is being extremely risk lover. The relationship between *extreme risk love* and income dissatisfaction is non-linear, as predicted by the theory, both with and without individual fixed effects. Extreme risk lovers tend to be more satisfied with their income, but also more extremely dissatisfied. In particular, becoming extremely dissatisfied increases the probability of becoming an extreme risk lover by almost 5 percentage points. This non-monotonic relationship between income satisfaction and risk preferences is consistent with the notion of prospect theory, that being far below vs close to (or above) one's reference point has opposite effects on risk preferences.

Income losses and populism Finally, we consider the first link in the causal chain, triggered by large income losses (relative to a three year moving average of past household income). Recall that we control for income and being unemployed throughout, so the estimated coefficient on *large income loss* only captures the effects of large negative income shocks. Again, we start with the reduced form implications, and we ask whether expe-

Table 2: **Income Dissatisfaction and Populism**

Dep. var.	Populist Dummy				Extreme Risk Love	
	(1)	(2)	(3)	(4)	(5)	(6)
Income Dissatisfaction	0.0813***	-0.0110	0.0866***	0.0010	-0.0821***	-0.0535***
	(0.020)	(0.028)	(0.020)	(0.029)	(0.012)	(0.014)
	[0.0009]	[-0.0022]	[0.0010]	[0.0002]	[-0.0027]	[-0.0076]
Extreme Income Dissatisfaction	0.4148***	0.3685**	0.3916***	0.3385**	0.7966***	0.3436***
	(0.102)	(0.149)	(0.103)	(0.151)	(0.062)	(0.082)
	[0.0046]	[0.0735]	[0.0044]	[0.0665]	[0.0262]	[0.0486]
Risk Love			0.0279*	-0.0213		
			(0.014)	(0.022)		
			[0.0003]	[-0.0042]		
Extreme Risk Love			0.4873***	0.3132**		
			(0.110)	(0.157)		
			[0.0055]	[0.0615]		
Unemployed	0.1494	-0.0476	0.1397	-0.0929	0.2660***	-0.0101
	(0.100)	(0.171)	(0.102)	(0.174)	(0.059)	(0.090)
	[0.0017]	[-0.0095]	[0.0016]	[-0.0183]	[0.0087]	[-0.0014]
Income	-0.0900*	-0.1600	-0.0748	-0.1678	-0.1302***	-0.0993
	(0.054)	(0.119)	(0.055)	(0.120)	(0.034)	(0.062)
	[-0.0010]	[-0.0319]	[-0.0008]	[-0.0330]	[-0.0043]	[-0.0140]
Observations	194,457	7,380	189,201	7,196	196,175	19,596
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Other Controls	Yes	No	Yes	No	Yes	No
Individual FE	No	Yes	No	Yes	No	Yes

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered at the individual level, marginal effects in square brackets. The dependent variable is voting for a populist party in columns 1-4, being an extreme risk lover in columns 5 and 6. In all columns, we also control for being in or out the labor force. Other individual controls (where included) are dummy variables for gender, age group, immigrant status, education level, and macroregion. Estimation is by logit in columns 1, 3 and 5, by conditional logit in the remaining columns. Source: SOEP.

riencing a large income loss is associated with support for populist parties. Columns 1 and 2 of Table 3 show that this is so: leaning towards populist parties is positively influenced by having experienced a large income loss, although magnitudes are small and statistical significance is only at 10%. In column 1, the specification is as in column 4 of Table 1, except that the two variables capturing risk attitudes are replaced by a dummy variable for having experienced large income losses. The dummy variable *large income loss* has a positive and significant estimated coefficient, although the average marginal effect is small: individuals who experienced a large income loss are more likely to lean towards populism by 0.2 percentage points. Column 2 adds individual fixed effects, and thus only captures variation over time by the same individuals; the estimated coefficient of *large income loss* rises, although the average marginal effect is about the same.

The other columns of Table 3 consider the remaining links. In columns 3-6 the dependent variable is being an extreme risk lover. Experiencing a large income loss is associated with extreme risk loving, as expected, irrespective of whether we also include income dissatisfaction as a regressor. Here too, however, the magnitude is rather small (large income losses increase the probability of extreme risk loving by about one percentage point), and the relationship between income dissatisfaction and extreme risk loving remains non-linear and highly significant, as in Table 2. This suggests that the variable *large income loss* captures only a small part of the effect of economic disappointment on risk preferences.

Finally, in columns 7-8 of Table 3 we consider the effect of *large income loss* on *extreme income dissatisfaction* (the first link in the causal chain). Again, the estimated coefficient of *large income loss* is positive and significant, as expected, but the magnitude is small, particularly when individual fixed effects are included. Note however that here income and being unemployed are always highly significant and with the expected sign.

As noted above, about half of the respondents are classified as not leaning towards a populist party because they don't lean towards any party or because they did not answer the relevant question. Appendix Tables A3 and A4 investigate whether there is a pattern in these non-responses. We replicate the relevant columns in Tables 1-3 above, but replace the dependent variable *populist* with a dummy variable that equals 1 if the respondent did not lean towards any political party or did not answer the political question. Non-responses are more likely amongst individuals who are extreme risk lovers and *less* likely amongst those who are extremely dissatisfied. But once individual fixed effects are included, the treatment variables of interest (extreme risk love, extreme income dissatisfaction and large income loss) are no longer statistically significant in these regressions. The only exception is column (2) of Appendix Table A4, where becoming extremely dissatisfied reduces the likelihood of a non-response, with statistical significance below 0.1. We conclude that, once individual fixed effects are included, the pattern of missing observations is generally not significantly correlated with the treatment variables of interest, suggesting that our results are not driven by non-responses.

All in all, these empirical findings suggest that risk attitudes play a relevant role in explaining support for populism, and that large economic disappointment can induce support for populist and radical parties by inducing a preference for political risk.

Table 3: **The Effect of Income Losses**

Dep.var.	Populist Dummy		Extreme Risk Love				Extreme Income Dissatisfaction	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Large Income Loss	0.1886*	0.3036	0.3623***	0.1589	0.3564***	0.1840*	0.1547***	0.1008*
	(0.104)	(0.185)	(0.062)	(0.098)	(0.063)	(0.099)	(0.034)	(0.0532
	[0.0022]	[0.0020]	[0.0093]	[0.0140]	[0.0091]	[0.0109]	[0.0126]	[0.0000]
Income Dissatisfaction					-0.1110***	-0.0756***		
					(0.017)	(0.021)		
					[-0.0028]	[-0.0045]		
Extreme Income Dissatisfaction					0.9433***	0.3762***		
					(0.088)	(0.123)		
					[0.0241]	[0.0222]		
Unemployed	0.4926***	-0.1292	0.3946***	-0.2224	0.3440***	-0.2028	1.1786***	1.0866***
	(0.122)	(0.236)	(0.085)	(0.143)	(0.089)	(0.146)	(0.045)	(0.070)
	[0.0057]	[-0.0008]	[0.0101]	[-0.0196]	[0.0088]	[-0.0120]	[0.0958]	[0.0003]
Income	-0.1891**	0.4120*	-0.0606	-0.1758	-0.0844	-0.2004	-1.3962***	-0.8108***
	(0.075)	(0.234)	(0.055)	(0.132)	(0.056)	(0.1329)	(0.032)	(0.079)
	[-0.0022]	[0.0027]	[-0.0015]	[-0.0155]	[-0.0022]	[-0.0118]	[-0.1135]	[-0.0002]
Observations	134,270	4,808	133,831	10,543	132,657	10,373	137,164	32,126
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Other Controls	Yes	No	Yes	No	Yes	No	Yes	No
Individual FE	No	Yes	No	Yes	No	Yes	No	Yes

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered at the individual level, marginal effects in square brackets. The dependent variable is voting for a populist party in columns 1-2, being an extreme risk lover in columns 3-6, and a dummy for extreme dissatisfaction with household income in columns 7 and 8. In all columns, we also control for being in or out the labor force. Other individual controls (where included) are dummy variables for gender, age group, immigrant status, education level, and macroregion. Estimation is by logit in columns 1, 3, 5 and 7, by conditional logit in the remaining columns. Source: SOEP.

3 Voting Over Safe and Risky Candidates

In this section we show that large economic downturns create economic uncertainty through the political system. More precisely, we consider voters with reference-dependent preferences who face the choice between two possible candidates, a moderate and a populist. The moderate is safe (i.e. it entails no uncertainty), while the populist is risky and less efficient than the populist. We model this with the assumption that, when the populist is in office, each voter has a stochastic income with a lower expected value compared to his certain income when the moderate is in office. For now these are the only differences between candidates (redistributive policies chosen by each candidate are considered in the next section). These intrinsic differences between the two candidates reflect policies that will be chosen after the election, and the two politicians are expected to select different types of policies. In particular, populist politicians, being anti-establishment and more radical, are expected to enact more risky and on average less efficient policies, or to react to future shocks in more unpredictable ways (e.g. by abandoning the Euro, or by trying untested and unconventional policies).

With standard preferences, all voters would be in favor of the moderate candidate. This is no longer true with reference-dependent preferences. Voters whose income is much below their reference point become risk-lover. For them, the risk associated to the populist candidate may be appealing and they may vote for him, provided that the difference in efficiency is not too large. We characterize who benefits from and is most opposed to economic uncertainty: the very disappointed voters are the beneficiaries, the mildly disappointed (or close to disappointed) are the most opposed, while the very well off (relative to their expectations) are mildly opposed.

Although the relative preferences of voters between the two parties are non-monotone, we show that this simple model enjoys a single crossing property. If the median voter is sufficiently disappointed, the populist can prevail in the elections.

3.1 Model

Populist vs Moderate Candidates. There are two candidates and a continuum of voters that differ only in their income, $\theta \geq 0$. Let $G(\cdot)$ denote the cumulative distribution of income, θ, over the population. For an agent with income θ, the moderate (M) candidate offers the status quo, namely income θ. The populist candidate (P) instead is more risky and inefficient. For an agent who would have income θ with the moderate, the populist generates income $\theta - z + \eta$ where $z > 0$ and η is a random variable distributed over $[-\varepsilon, \varepsilon]$ with density h so that $\mathbf{E}(\eta) = \int_{-\varepsilon}^{+\varepsilon} \eta h(\eta) d\eta = 0$. The populist is therefore riskier than the moderate (because of the presence of the random shock η) and less efficient. The parameter z is a measure of the inefficiency of the populist compared to the moderate. We assume $\varepsilon > z$ so that with some probability the populist can generate a higher income than the moderate for all voters. Note that the difference between the populist and moderate candidates only concerns aggregate outcomes and it is the same for all voters - although as shown below different voters have different evaluations of these intrinsic features of candidates. As discussed in subsection 4.5 below, our main results continue to hold even if $z = 0$ (i.e. if on average the populist and the moderate candidate are equally efficient).

Voters Preferences. Voters have reference-dependent preferences. In particular, for a given distribution of consumption F_c they enjoy utility:[8]

$$U(F_c, x) = \mathbb{E}[c + \mu(c - x)],$$

where c is consumption, $\mu(\cdot)$ is a negative valued, increasing, and convex function that penalizes the voter whenever $c < x$. We will refer to $x > 0$ as 'the reference point' and we assume that it is the same for everyone, irrespective of his actual income. In subsection 4.5 we discuss how to relax this assumption. We also take x as exogenously given, although it could be endogenously derived, for instance assuming that agents are rational and forward looking, as in Koszegi and Rabin (2009).

Let $d = c - x$. We summarize our assumptions on μ below:

ASSUMPTION 1. *The function μ is continuous over the whole domain, with $\mu(d) = 0$ for all $d \geq 0$. Moreover, for $d < 0$, μ is at least twice continuously differentiable and the following properties hold:*
(i) $\mu(d) < 0$; (ii) $\mu'(d) > 0$; (iii) $\mu''(d) > 0$; (iv) $\mu'''(d) \leq 0$; (v) for $d = 0$, μ admits left derivatives at least till the third degree $\mu^{\underline{n}}_{-}(0)$, $n = 1, 2, 3$, which are compatible with the natural extensions of (ii)-(iv).

[8]With an abuse of notation, we denote by c both the random variable and a particular realisation of it.

An example of such preferences is represented in Figure 1 above.[9] Conditions (i)-(iii) are standard in models of reference-dependent preferences (e.g., Koszegi and Rabin (2009)). Point (iv) of Assumption 1 is less standard and plays an important role in our results. In essence, a negative third derivative (i.e., a concave first derivative) is equivalent to having preferences for moving risk from high to low income levels. This assumption hence implies that more disappointed agents are more likely to favor the Populist candidate because of its intrinsic riskiness.[10]

Timing of Events Agents start with a given reference point x. After that, but before elections, their income level θ is realized. Then elections are held and the winner is elected. If the populist candidate wins, income shocks are realized and agents consume. Hence, if the moderate wins, agent with income θ consumes $c_M(\theta) = \theta$ and enjoys utility:

$$w_M(\theta, x) = \theta + \mu(\theta - x);$$

If instead the populist is elected, the same agent has an expected consumption equal to $\mathbf{E}(c_P(\theta)) = \theta - z$ and enjoys an expected utility

$$w_P(\theta, x) = \theta - z + \int_{-\varepsilon}^{\varepsilon} \mu(\theta - z + \eta - x) h(\eta) d\eta.$$

[9] As an example consider the cubic polynomial form for μ and consider nonnegative consumption. For $c \geq x$, $\mu(c - x) \equiv 0$. While for $c - x \leq 0$, $c \geq 0$:

$$\mu(c - x) = a(c - x) + \frac{b}{2}(c - x)^2 + \frac{q}{3}(c - x)^3 .$$

with $a > 0$, $b > 0$, $q \leq 0$, and $a - bx + qx^2 > 0$. Note that $\mu(0) = 0$ and, for $0 \leq c < x$, $\mu'(c - x) = a + b(c - x) + q(c - x)^2 > 0$, $\mu''(c - x) = b + 2q(c - x) > 0$ and $\mu'''(c - x) = 2q \leq 0$. The slope of μ is maximal at $c = x$ where it takes the value of a and it decreases to $a - bx + qx^2$ for $c = 0$.

[10] Consider indeed the difference between the expected value of μ of a zero mean gamble and the no gamble value evaluated with μ at a given t:

$$\mathbb{E}\left[\mu(t + \eta)\right] - \mu(t),$$

where η is a zero mean random variable (mean preserving spread). If we derive the previous expression with respect to t we get

$$\mathbb{E}\left[\mu'(t + \eta)\right] < \mu'(t),$$

where the inequality comes from the Jensen's inequality and our assumption of concave first derivative. This means - as claimed - that the (positive) gap between the expected gain of a zero mean gamble and the no gamble situation decreases with t.

Non-monotone support with Single Crossing. Thanks to the linearity of the baseline preferences, for each given z, the difference in expected utilities between the two candidates for voter with income θ is a function of the difference $\theta - x$ alone, namely:

$$w_M(\theta, x) - w_P(\theta, x) := \Delta(\theta - x) = z + \mu(\theta - x) - \int_{-\varepsilon}^{\varepsilon} \mu(\theta - x - z + \eta)h(\eta)d\eta.$$

We show next that the functions w_P and w_M satisfy a single crossing property. We start with a lemma. Consider a given z, and let $t := \theta - x$.

LEMMA 1. *Under assumption 1: (i) for $t > 0$, $\Delta(\cdot)$ is strictly positive and (weakly) decreasing, while (ii) for $t < 0$, $\Delta(\cdot)$ is a strictly increasing function.*

Lemma 1 formalizes an important property of our framework. The lemma immediately implies that Δ is maximal at $t = 0$, that is, the voter who most prefers the moderate is $\theta = x$, while at the extremes of the support we have the most mildly opposed to P. This feature of the model is consistent with the empirical evidence in Table 2, and it plays an important role in the political equilibrium analyzed in the next section.

Lemma 1 has another interesting implication, which is stated in the next proposition.

PROPOSITION 2. *Assume Assumption 1. Then $w_M(\theta, x)$ and $w_P(\theta, x)$ can cross at most once, at the income level $\hat{\theta} < x$.*

An example of the (single) crossing of the two curves is given in Figure 2, where we indicate with $\hat{\theta}$ be the value of income at which the crossing occurs (assuming it exists, otherwise the moderate gets 100% of votes because of his superior efficiency). Voters with income $\theta > \hat{\theta}$ prefer candidate M, while those with income $\theta < \hat{\theta}$ prefer candidate P.

Importantly, what matters in the figure, is not the value of income θ per se but its relation with the reference point x. The same income distribution can lead to different political outcomes for different values of the reference point. In particular, large economic downturns may produce a political demand for risky policy choices. If the downturn is deep enough, so that the fraction of agents below $\hat{\theta}$ increases, the likelihood that the populist is elected increases. Recall that by Assumption 1, the more disappointed an agent is - i.e., the farther left he is from the reference point - the more he benefits from policy uncertainty.

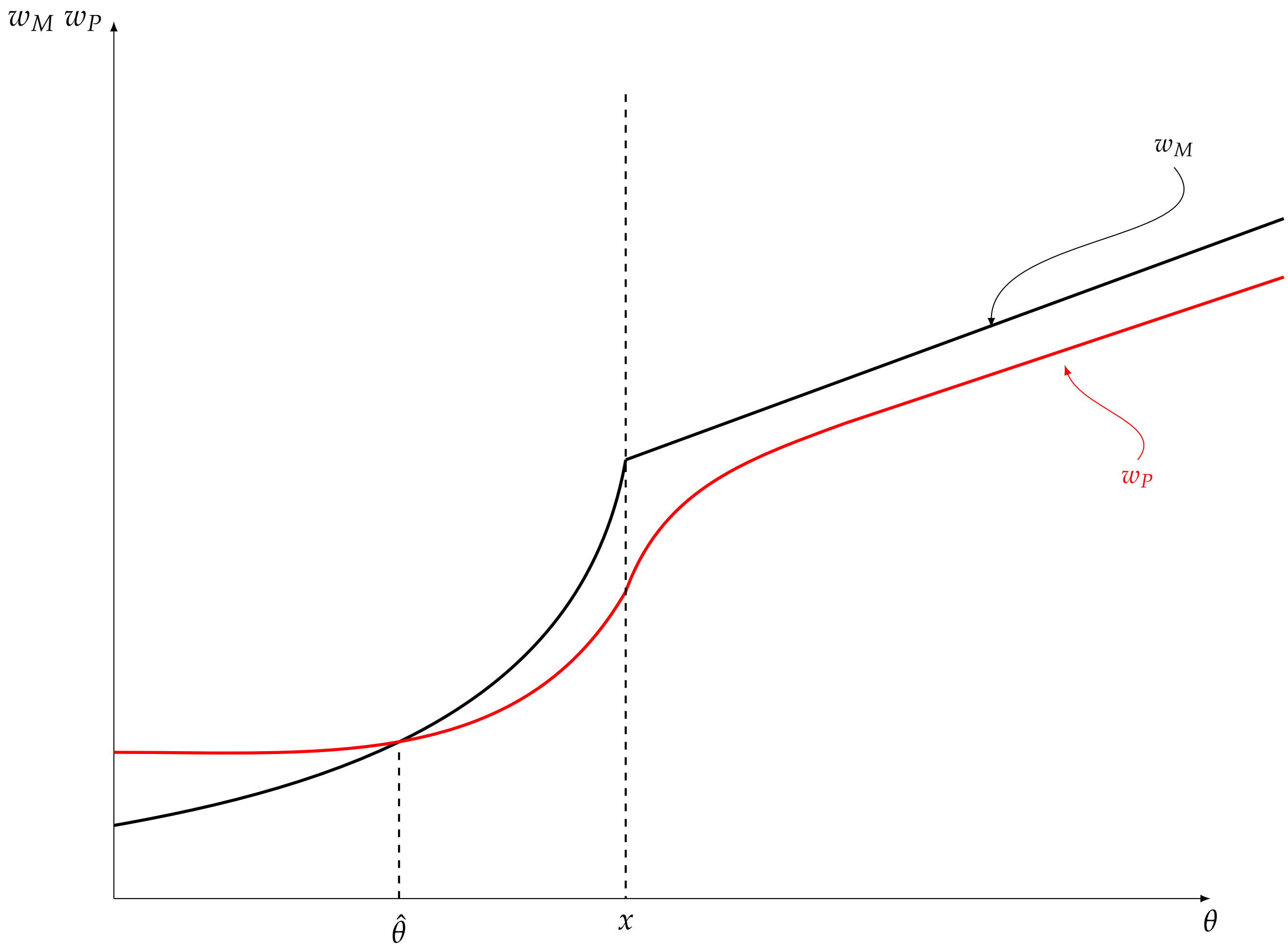

Figure 2: An Example of the single crossing in the space of θ.

4 Political Competition over Redistributive Taxation

In this section, we allow the moderate and the populist candidates to compete over income tax rates $0 \leq \tau_M, \tau_P \leq 1$, assuming linear but distorting taxes (subscripts refer to candidates). The tax proceeds are distributed as a lump sum to every agent. Under the moderate candidate, the consumption of agent θ equals:

$$c_M(\theta, \tau) = (1 - \tau_M)\theta + f(\tau_M),$$

while, under the populist, consumption is random and equals:

$$c_P(\theta, \tau) = (1 - \tau_P)\theta - z + f(\tau_P) + \eta,$$

where, recall, the random variable η is distributed between $-\varepsilon$ and $+\varepsilon$ and both η and z affect all agents equally. Note that since $-z + \eta$ only affects aggregate income and redistribution is lump sum, effectively only the fixed and idiosyncratic income component θ is taxed. The function $f(\tau)$ solves the government budget constraint and embeds inefficiencies due to distortionary taxation. Specifically, we assume:

$$f(\tau) = \tau \mathbf{E}\theta - i(\tau), \tag{1}$$

with $i(\cdot) \geq 0$ an increasing and convex function representing tax distortions, and $\mathbf{E}\theta$ representing the cross-sectional average of θ.[11] Throughout we assume that the properties of $i(\cdot)$ satisfy the following:

ASSUMPTION 2. *The function f as defined in* (1) *takes a maximal value at tax rate* $0 \leq \tau_0 < 1$ *and the function $i(\cdot)$ is such that $i(0) = 0 = i'(0)$.*

Thus, if f is differentiable, at τ_0 we have $f'(\tau_0) = \mathbf{E}\theta - i'(\tau_0) \leq 0$, with equality if $\tau_0 > 0$. It is easy to verify that the optimal tax rate for an agent with income θ, denoted by τ^θ, satisfies the optimality condition:$f'(\tau^\theta) \leq \theta$, with equality if $\theta \leq \mathbf{E}\theta$ and strict inequality otherwise, irrespective of which type of agent is in office (this follows from the assumption that the shock associated with the populist is aggregate). Hence, τ^θ, is a decreasing function of θ if $\theta < \mathbf{E}\theta$ and $\tau^\theta = 0$ if $\theta \geq \mathbf{E}\theta$. Moreover, by definition of τ_0, $\tau^\theta \leq \tau_0$. No voter with $\theta > 0$, no matter how poor, would ever want to have a tax rate on the wrong side of the Laffer curve.

Despite these well behaved policy preferences, however, voters' preferences for the package of a politician's type (moderate or populist) and his associated tax rate do not satisfy the single crossing property. That is, for some tax rates τ_P and τ_M, the utility functions $w_P(\theta, \tau_P)$ and $w_M(\theta, \tau_M)$ may intersect at more than one value of θ. This can be seen from Figure 3. If $\tau_P = \tau_M$, then the indirect utility functions $w_P(\theta, \tau_P)$ (dotted curve) and $w_M(\theta, \tau_M)$ of voter θ under each candidate remain as depicted in Figure 2, and the

[11] To avoid confusion, we use the 'bold' notation $\mathbf{E}$ for cross-sectional averages and the 'math' notation $\mathbb{E}$ when the integration is taken over the shock $\eta \in [-\varepsilon, \varepsilon]$.

two utility functions intersect at most once to the left of x, at income level $\hat{\theta}$. But suppose that τ_P is lowered below τ_M. This steepens the slope of $w_P(\theta, \tau_P)$ (solid curve) relative to that of $w_M(\theta, \tau_M)$. Hence the linear component of $w_P(\cdot)$ (where $c^P > x$) intersects the linear component of $w_M(.)$ for $\theta < \infty$. At the same time, by the argument discussed in the previous section, the convex component of $w_P(.)$ could remain flatter than that of $w_M(\cdot)$ for $c^P < x$ and for small realizations of θ. Hence the two curves can intersect in at least two points, one to the left (denoted as $\underline{\theta}$) and one to the right (denoted as $\overline{\theta}$) of x, as shown in Figure 3.

The absence of single crossing implies that a Condorcet winner may not exist. In such a case, a simple model of Downsian electoral competition, where the two candidates compete in the election by simultaneously committing to a tax rate, does not have a pure-strategy Nash equilibrium, because of the non convexity of the objective functions.

To get around this problem, we assume instead that the two candidates move sequentially and that they maximize their vote share.[12] Specifically, we assume the following timing. First, the moderate candidate commits to a tax rate τ_M. Then the populist candidate commits to a tax rate τ_P. Finally, voters observe both tax rates and, knowing the risk properties of each candidate, they vote. This timing assumption can be interpreted as saying that the moderate candidate is the incumbent and this makes him less flexible in his policy announcements. Although we will maintain such timing in our narrative, this can be considered non-consequential, for at least two reasons. First, since the game played is zero-sum, whenever the game admits a Nash equilibrium in pure strategies (i.e., a saddle of the payoff function), Stackelberg and Nash equilibria generate the same payoffs and the same tax rates. Second, in Proposition 11 in Appendix C.1, we show that the key qualitative properties of the equilibrium set - as reported in Proposition 5 for the benchmark timing - are independent on the identity of the first mover.[13]

[12]The assumption that candidates maximize the vote share can be interpreted as saying that (exogenous) political rents are an increasing function of the vote share. Alternatively, the outcome of the election could be random and determined by the realization of an aggregate and exogenous popularity shock, making the probability of victory an increasing function of the vote share.

[13]Since the choice variables lie in the compact set $[0, 1]$ the existence of a (Stackelberg, sequential) equilibrium (in pure strategies) of the dynamic game is guaranteed whenever the payoff functions of both players are continuous. In our context, it is easy to show that continuity of the payoff functions is guaranteed whenever the net distortion and the loss functions - f and μ - and are continuous and the distribution of over θ admits a density (in particular, it has no mass points, perhaps with the exclusion of the extremes).

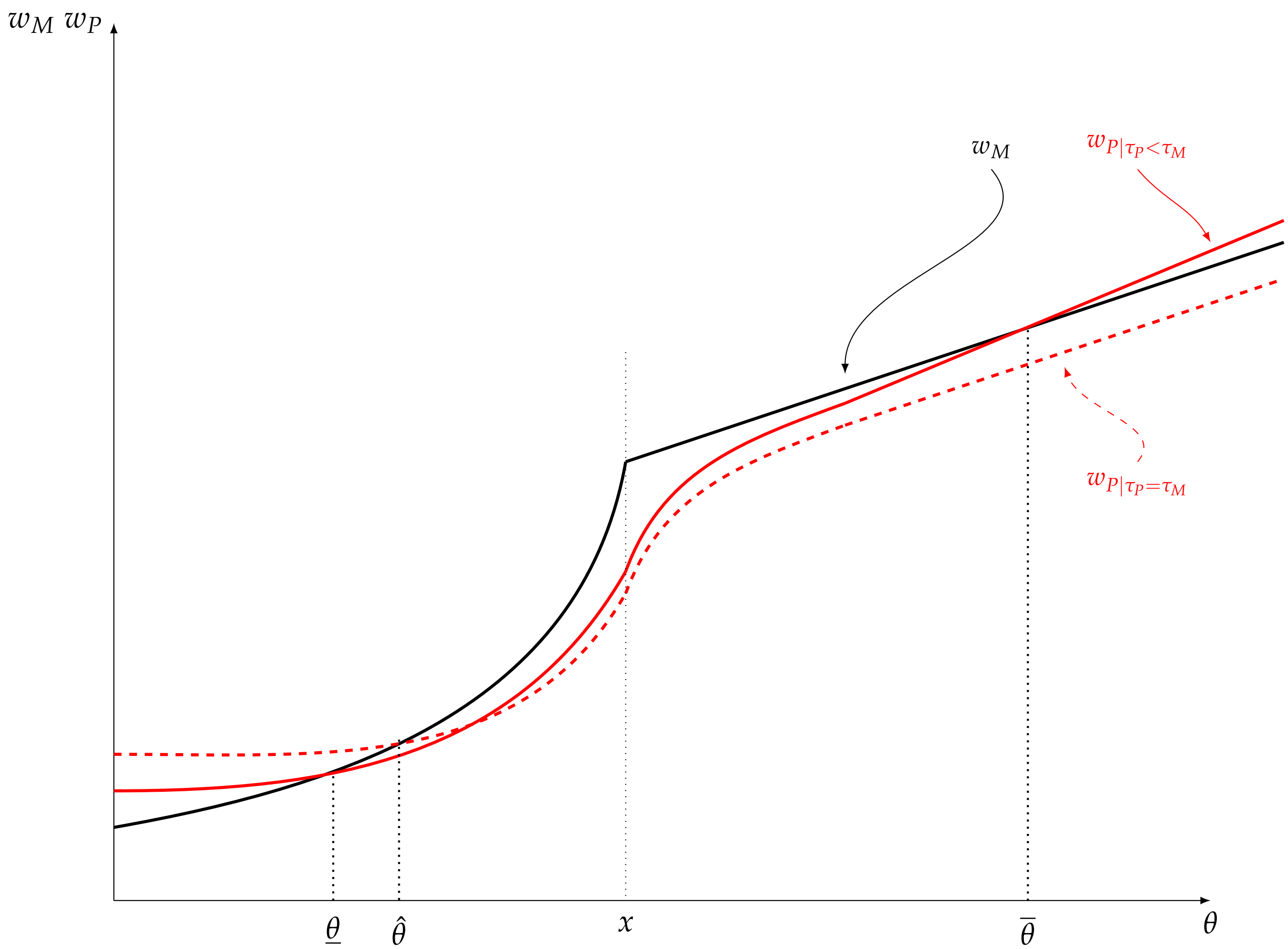

Figure 3: The black solid and the red dashed curves, representing $w_M(\cdot)$ and $w_P(\cdot)$ respectively, are drawn under the assumption that $\tau_M = \tau_P$. Similarly to Figure 2, in this case, the curves intersect only once at income level $\hat{\theta}$. If τ_P is reduced, while τ_M remains constant, the (red dashed) curve $w_P(\cdot)$ rotates to the red solid curve, and it can intersect $w_M(\cdot)$ at two income levels, $\underline{\theta}$ and at $\bar{\theta} > x$. Voters with income with $\theta < \underline{\theta}$ and with $\theta > \overline{\theta}$ prefer candidate P to M. Those in between prefer M. In absence of distortions (i.e., $i \equiv 0$), the rotation would have center the mean income $\mathbf{E}\theta$. With distortion, a reduction in τ_P also shift the curve w_P upwards.

Socially optimal policy. What is the optimal tax policy in this economy? The answer is not straightforward, because the marginal utility of income varies depending on whether voters are above or below their reference point. Moreover, very disappointed voters are risk loving, and so they may prefer not to be insured against income risk. The appendix characterizes the tax rate that solves the utilitarian optimum assuming that income is not stochastic (or, equivalently, that the moderate is elected), and shows that it is strictly positive if $\mathbf{E}\theta > x$ and there is a positive measure of agents with $\theta < x$. In this case, all disappointed agents are poorer than the average, and hence they benefit from redistribution. Since disappointed voters have higher marginal utility of income due to their loss aversion, a positive tax rate is optimal despite the tax distortions. In particular, at the social optimum the marginal tax distortions are equated to the marginal social benefit of redistributing in favor of the disappointed agents.

If $\mathbf{E}\theta < x$, however, this is no longer true, since some individuals with particularly high marginal utility are also hurt by redistribution. In this case a positive tax rate is socially optimal only under an additional condition stated in the appendix. This condition is more likely to be satisfied if: (i) there are not too many individuals with income between x and $\mathbf{E}\theta$; (ii) the function $\mu(\cdot)$ is not too convex.

We now turn to the analysis of the Political equilibrium. For this purpose, it is convenient to separately discuss two cases, $\tau_M \leq \tau_P$ and $\tau_M \geq \tau_P$.

4.1 Single crossing

We first consider the case $\tau_P \geq \tau_M$. The next Lemma states a sufficient condition under which the single crossing property continues to hold over this range of tax rates.

LEMMA 3. *Suppose that Assumptions 1 and 2 hold, and also assume that* $\mathbf{E}\theta \leq x - \varepsilon + z$. *Then, for* $\tau_P \geq \tau_M$ *we have: (i) the expected utilities* $w_M(\theta, \tau_M)$ *and* $w_P(\theta, \tau_P)$ *cross at most once, and (ii) this may only happen for* $\theta = \hat{\theta}$ *such that* $\mathbb{E}c_P(\hat{\theta}, \tau_P) < c_M(\hat{\theta}, \tau_M) < x$. *(iii) Moreover,* $w_M(\theta, \tau_M) \geq w_P(\theta, \tau_P)$ *(resp.* $w_M(\theta, \tau_M) \leq w_P(\theta, \tau_P)$*) if* $\theta \geq \hat{\theta}$ *(resp.* $\theta \leq \hat{\theta}$*).*

Note that $z - \varepsilon < 0$. The (sufficient) condition $\mathbf{E}\theta \leq x - \varepsilon + z$ can hence be interpreted as saying that the economy has been hit by such a large negative shock that even the average voter (whose income is $\mathbf{E}\theta$) always remains below his reference point when the populist is elected. This assumption also implies that any voter with income $\theta > x$ also has $\theta > \mathbf{E}\theta$, and hence he/she opposes redistribution.[14]

The threshold level $\hat{\theta}(\tau_M, \tau_P)$ for which $w_P(\hat{\theta}, \tau_P) = w_M(\hat{\theta}, \tau_M)$, if it exists, is defined implicitly by the following equation

$$c_M\left(\hat{\theta}(\tau_M, \tau_P), \tau_M\right) + \mu\left(c_M\left(\hat{\theta}(\tau_M, \tau_P), \tau_M\right) - x\right) = \mathbb{E}c_P\left(\hat{\theta}(\tau_M, \tau_P), \tau_P\right) + \mathbb{E}\mu\left(c_P(\hat{\theta}(\tau_M, \tau_P), \tau_P) - x\right) \tag{2}$$

Single crossing implies that the problem of the populist party restricted to choosing $\tau_P \geq \tau_M$ is very simple. By Lemma 3, all voters with $\theta > \hat{\theta}(\tau_M, \tau_P)$ vote for M, while all voters with $\theta < \hat{\theta}(\tau_M, \tau_P)$ vote for P. Then, to maximize his vote share, the populist candidate sets τ_P to maximize $\hat{\theta}(\tau_M, \tau_P)$, taking τ_M as given and subject to $\tau_P \geq \tau_M$. We denote with $\tau_P^*(\tau_M)$ the solution to the populist optimization problem over this range. The function $\tau_P^*(\tau_M)$ is defined implicitly by $\frac{\partial \hat{\theta}}{\partial \tau_P} \leq 0$, with equality for $\tau_P^*(\tau_M) > \tau_M$, and strict inequality if the constraint $\tau_P \geq \tau_M$ is binding on P (with positive multiplier).

By symmetry of the (zero sum game) problem, the moderate aims at minimizing $\hat{\theta}(\tau_M, \tau_P^*(\tau_M))$. By the envelope property, the effect of τ_M on $\hat{\theta}$ has opposite sign with respect to that of τ_P and it is larger in absolute value. In particular, taking into account the reaction function of the populist candidate, the threshold $\hat{\theta}$ is strictly increasing in τ_M in the range where $\tau_P^*(\tau_M) > \tau_M$. More formally:

LEMMA 4. *Consider the case where the populist is restricted to set $\tau_P \geq \tau_M$. Under the assumptions of Lemma 3 (so that we have single crossing over the relevant range):*
i) For each τ_M admitting an interior solution to the populist optimization problem under the constraint $\tau_P \geq \tau_M$, $\tau_P^*(\tau_M)$ is the most preferred tax rate for voters with income $\hat{\theta}(\tau_M, \tau_P^*(\tau_M))$.
ii) The threshold $\hat{\theta}(\tau_M, \tau_P^*(\tau_M))$ is strictly decreasing in τ_M for $\tau_P^*(\tau_M) > \tau_M$; it is constant in τ_M if $\tau_P^*(\tau_M) = \tau_M$ and the constraint $\tau_P \geq \tau_M$ is not binding for candidate P; it is strictly increasing in τ_M if $\tau_P^*(\tau_M) = \tau_M$ and the constraint $\tau_P \geq \tau_M$ is binding (with positive multiplier) for candidate P.

Intuitively, (i) says that under single crossing an unconstrained populist candidate sets the tax rate that is optimal for the individual who is just indifferent between M and P, since this maximizes the area to the left of the indifferent individual; while (ii) formalizes the fact that - when $\tau_P \geq \tau_M$ - M and P compete for the same voter from opposite sides.

[14] This assumption is sufficient but not necessary, and it can be weakened. In the proof, this assumption guarantees that non-disappointed voters (those with $\theta > x$) cannot be attracted (by the Populist) with higer taxes (part A). It is also used to guarantee that, for disappointed voters, $\mathbb{E}c_P(\theta, \tau_P) \geq c_M(\theta, \tau_M)$ implies $w_P(\theta, \tau_P) > w_M(\theta, \tau_M)$ (part B2).

Thus, anticipating the populist behavior in the range $\tau_P \geq \tau_M$, candidate M finds it optimal to keep increasing its tax rate, up to the point where $\tau_P^*(\tau_M) = \tau_M$ is an unrestricted local optimum for his populist opponent. Note in particular that, if the populist finds it optimal to choose $\tau_P^* > \tau_M$, then the moderate would be able to decrease $\hat{\theta}$ further (and hence increase his vote share) by raising τ_M.

4.2 Political Equilibrium

We are now ready to study the full equilibrium. To simplify the exposition, throughout we assume that the distribution G has (full) support over $\mathbb{R}_+$. Let $V_P^+(\tau_M)$ denote the vote share of the populist candidate as a function of τ_M, given $P's$ best response to τ_M in the range $\tau_P \geq \tau_M$. By Lemma 4 and the discussion above, we have:

$$V_P^+(\tau_M) := \max_{\tau_P \geq \tau_M} \; G(\hat{\theta}(\tau_M, \tau_P)), \tag{3}$$

where $\hat{\theta}(\tau_M, \tau_P)$ has been defined in (2). Let $\tau_P^+(\cdot)$ be the (restricted) policy correspondence representing the set of solutions to (3) for ech τ_M. Similarly, let $V_P^-(\tau_M)$ denote the vote share of the populist candidate as a function of τ_M, given $P's$ best response to τ_M in the range $\tau_P \leq \tau_M$. Since over this range $w_P(\theta, \tau_P^-(\tau_M))$ and $w_M(\theta, \tau_M)$ can cross more than once, we do not attempt to characterize $V_P^-(\tau_M)$ at this stage. Let $\tau_P(\cdot)$ be the unrestricted policy correspondence for the populist.

The moderate candidate's sets τ_M so as to solve the following Minimax problem:

$$\min_{\tau_M \in [0,1]} [\max\{V_P^+(\tau_M), V_P^-(\tau_M)\}]. \tag{4}$$

We denote by τ_M^* any solution to this problem. Intuitively, M knows that, once it has set τ_M, his opponent will respond by choosing the optimal tax rate in the range ($\tau_P \geq \tau_M$ or $\tau_P \leq \tau_M$) that maximizes $P's$ vote share. Thus, M chooses τ_M to minimize his opponent's best alternative.

There are two possibilities. First, M sets τ_M so as to minimize $V_P^+(\tau_M)$, subject to $V_P^+(\tau_M) \geq V_P^-(\tau_M)$. By Lemma 4, $V_P^+(\tau_M)$ is decreasing in the range where $V_P^+(\tau_M) \geq V_P^-(\tau_M)$, strictly so if $\tau_P^+(\tau_M) > \tau_M$. Hence, in this case the equilibrium is found at the point where $V_P^+(\tau_M^*) = V_P^-(\tau_M^*)$, namely P is left indifferent between $\tau_P \geq \tau_M^*$ and $\tau_P \leq$

τ_M^* (of course policy convergence at the point $\tau_P(\tau_M^*) = \tau_M^*$ is also a possibility). Note that we cannot have a situation where $V_P^+(\tau_M) > V_P^-(\tau_M)$ and $\tau_P(\tau_M) > \tau_M$, because M would then find it optimal to raise τ_M.

The second case is where M sets τ_M so as to minimize $V_P^-(\tau_M)$, subject to $V_P^-(\tau_M) \geq V_P^+(\tau_M)$. This case could either result in a binding constraint, where $V_P^+(\tau_M^*) = V_P^-(\tau_M^*)$ as above, or else the constraint is not binding, and the solution is found by minimizing $V_P^-(\tau_M)$ with respect to τ_M at a point such that in equilibrium $V_P^-(\tau_M^*) > V_P^+(\tau_M^*)$. By definition of $V_P^-(\tau_M)$ then, in this case $\tau_P(\tau_M^*) < \tau_M^*$. This in turn implies that single crossing is violated (since single crossing always implies policy convergence, i.e. $\tau_P(\tau_M^*) = \tau_M^*$). Hence, in this second case the functions $w_P(\theta, \tau_P(\tau_M^*))$ and $w_M(\theta, \tau_M^*)$ cross more than once (i.e for at least two different values of θ).

We summarize this discussion in the following:

PROPOSITION 5. *Under the assumptions of Lemma 3, a political equilibrium outcome can be of two sorts:*
i) Indifference: The populist candidate is left indifferent between the solution when choosing $\tau_P \geq \tau_M$ and the solution(s) to the problem restricted to $\tau_P \leq \tau_M$;
ii) Populist tax cuts: the populist candidate strictly prefers $\tau_P(\tau_M^*) < \tau_M^*$, and the schedules $w_P(\theta, \tau_P(\tau_M^*))$ and $w_M(\theta, \tau_M^*)$ cross at more than one income level θ.

Of course, the case when the populist is left indifferent (point (i)) includes the solution with policy convergence, $\tau_P(\tau_M^*) = \tau_M^*$. In the next Corollary we state that this is the sole possibility if the optimal policy for the populist is single-valued, delivering single crossing. However, we can also have the situation where, at τ_M^*, the populist has two global maxima amongst which he is indifferent: one where the solution is a tax rate τ_P^+ larger than τ_M^*, and another where the optimal tax rate τ_P^- is smaller than τ_M^*.

COROLLARY 6. *(Policy Convergence) Assume that in the political equilibrium of type (i) in Proposition 5 the populist has a single-valued optimal policy (i.e., a unique optimal solution). Then the equilibrium displays single crossing and we have policy convergence at the bliss point of the voter who is just indifferent between M and P at the equilibrium tax rate. That is: $\tau_P^* = \tau_M^* = \tau^{\hat{\theta}}$.*

The intuition for such outcome is as follows. In the range $\tau_P \geq \tau_M$, both candidates would like to flatten the curves of voters' welfare w_M and w_P in Figure 2, so as to shift the threshold $\hat{\theta}$ in their desired direction. This tendency is mitigated by the fact that higher taxes also entail higher tax distortions, however. Higher tax distortions shift the expected utility downwards, and hence move the threshold $\hat{\theta}$ in the opposite direction. In this case, these two forces are balanced. Not surprisingly, single crossing implies policy convergence at the bliss point of the voter who is just indifferent between the two candidates (not necessarily the median voter).

The equilibrium of case (ii) in Proposition 5 is more interesting. Here we do not have single crossing (i.e., the voters' utility functions intersect more than once), and in equilibrium there is no policy convergence: the populist candidate announces a lower tax rate than the moderate. By Proposition 5, we know that there must be at least another intersection point, possibly more than one. As in the previous discussion of Figure 3, let $\bar{\theta}$ denote the highest level of income for which $w_P(\theta, \tau_P(\tau_M^*))$ and $w_M(\theta, \tau_M^*)$ intersect. Since $\tau_P(\tau_M^*) < \tau_M^*$, it must be the that $w_P(\bar{\theta}, \tau_P(\tau_M^*))$ intersects $w_M(\bar{\theta}, \tau_M^*)$ from below.[15] Hence, all voters with income $\theta > \bar{\theta}$ prefer candidate P to candidate M. Let $\underline{\theta}$ denote the first intersection point to the left of $\bar{\theta}$, so that at $\underline{\theta}$ the function $w_M(\underline{\theta}, \tau_M^*)$ intersects $w_P(\underline{\theta}, \tau_P(\tau_M^*))$ from below (i.e. w_M is steeper than w_P). Hence, all voters with income $\theta < \underline{\theta}$ prefer candidate P to candidate M. Since $\tau_P(\tau_M^*) < \tau_M^*$ and $\mu'(\cdot) > 0$, this second intersection point can only occur in the region where w_M is non-linear - i.e., where $c_M(\underline{\theta}, \tau_M^*) < x$. We thus have:

COROLLARY 7. *In case (ii) of Proposition 5 the populist candidate is supported by a coalition that includes the richest voters (i.e. all voters with $\theta > \bar{\theta}$) and disappointed voters (i.e. voters with $c_M < x$ in equilibrium).*

To illustrate this corollary in a case with double crossing of the functions $w_M(\theta, \tau_M^*)$ and $w_P(\theta, \tau_P(\tau_M^*))$, we can refer to Figure 3. This is the case illustrated by the solid curve. All voters with income $\theta > \bar{\theta}$ and with income $\theta < \underline{\theta}$ prefer candidate P to candidate M. The populist candidate is supported by a coalition of the extremes, the rich voters and the poor and disappointed voters. With more than double crossing, the poorest voters could support the moderate incumbent (who promises higher taxes), but it would remain true that the populist candidate draws the support of the richest voters and of some disappointed voters. With double (or more) crossing, it is also true that the moderate candidate always receives the support of some voters with intermediate levels of income, i.e. with

[15] The fact that w_P intersects w_M from below can be proved by contraddiction. Suppose that at $\bar{\theta}$ w_P intersects w_M from above. Then $\bar{\theta}$ cannot be the right-most intersection point, because: (i) $\tau_P(\tau_M^*) < \tau_M^*$ implies $\frac{\partial w_P}{\partial \theta} > \frac{\partial w_M}{\partial \theta}$ in the linear part of the utility functions: (ii) G has (full) support over $\mathbb{R}_+$.

$\underline{\theta} < \theta < \bar{\theta}$. These are the voters who fear the risky populist politician the most because, being close to their reference point, they could suffer a lot if the populist politician has a bad draw, because of loss aversion.

In other words, the populist candidate knows that he appeals the most to the risk loving and disappointed voters. He can thus afford to choose a tax rate that is too low for these voters, knowing that they would be reluctant to vote for the moderate politician that they dislike. Setting a lower tax rate, enables the populist candidate to also gain the vote of the richest and non-disappointed individuals, who only care about the policy platforms and not about the intrinsic features of the two politicians. Thus, in equilibrium the populist politician sets a lower tax rate and is supported by a coalition of rich and disappointed voters.

Finally, who wins the election? If in equilibrium $\tau_P(\tau_M^*) \geq \tau_M^*$, then by Lemma 3 we have single crossing at the level of income $\hat{\theta}(\tau_P(\tau_M^*), \tau_M^*)$. Voters to the right of $\hat{\theta}$ prefer the moderate candidate, those to the left prefer the populist candidate. Hence candidate M wins if the median level of income exceeds $\hat{\theta}$, while candidate P wins in the opposite case. Note that, if in equilibrium $w_M(\theta, \tau_M^*)$ and $w_P(\theta, \tau_P(\tau_M^*)$ cross only once at the point $\hat{\theta}$, although there is policy convergence, the equilibrium tax rates of either candidate are not attracted by the median voter bliss point. If in equilibrium $\tau_P(\tau_M^*) < \tau_M^*$, then we may have double crossing or more, we have no policy convergence, and to determine who wins we have to sum the coalitions of voters in favor of one or the other candidate. For instance, in the case of double crossing, the size of the voting coalition in favor of the moderate candidate is $G(\bar{\theta}(\tau_P(\tau_M^*)\, \tau_M^*)) - G(\underline{\theta}(\tau_P(\tau_M^*)\, \tau_M^*))$. If this expression exceeds $1/2$ then M wins, otherwise P does. Here too, the median voter bliss points does not pin down equilibrium tax rates.[16]

4.3 Double crossing

We now consider the case in which the value functions $w_M(\theta, \tau_M)$ and $w_P(\theta, \tau_P)$ cross at most twice for $\tau_P < \tau_M$. In Appendix D, we provides a set of sufficient conditions for this to happen. These conditions amount to assuming - roughly - that: (i) for any $\tau_P < \tau_M$, the right-most crossing point occurs sufficiently far from the reference point

[16] If the election outcome was determined by the realization of an aggregate and exogenous popularity shock, as discussed in footnote 11, then the statements in the text would have to be interpreted as affecting the probability of winning, but not who is ultimately appointed.

x , and that (ii) the function $\mu(\cdot)$ is sufficiently well behaved that $w_M(\theta, \tau_M)$ remains below $w_P(\theta, \tau_P)$ for all levels of income below the first crossing point to the left of x. The latter is obviously the crucial assumption, while property (i) only serves to formally state property (ii). Under these conditions, if $\tau_P < \tau_M$ then we have at most double crossing. As a result, by Proposition 5, in equilibrium candidate P is supported either by a group of poor and disappointed voters (single crossing equilibrium), or by a coalition of poor and disappointed voters and of rich voters, while voters with consumption close to the reference point vote for M (double crossing equilibrium).

Consider how both candidates behave in this second type of equilibrium, assuming that at this point P has a single valued optimal policy so that we can use the envelope theorem. Since there is double crossing, the optimal tax rate for P maximizes the area to the right of $\bar{\theta}$ plus the area to the left of $\underline{\theta}$. His optimality condition, taking τ_M as given, is:

$$g(\underline{\theta})\frac{\partial \underline{\theta}}{\partial \tau^P} \leq g(\bar{\theta})\frac{\partial \bar{\theta}}{\partial \tau^P},$$

with equality if $\tau_P > 0$, and where $g(\cdot)$ is the density of the income distribution $G(\cdot)$. This condition says that, at an interior optimum, the marginal votes gained by P amongst the poor when raising τ_P are equal to the votes lost amongst the rich.[17] By the envelope theorem, the optimality condition for M, taking into account how τ_P responds to τ_M, is:

$$g(\underline{\theta})\frac{\partial \underline{\theta}}{\partial \tau^M} = g(\bar{\theta})\frac{\partial \bar{\theta}}{\partial \tau^M},$$

which has the same interpretation, namely M also equates the marginal votes gained and lost on the opposite sides of the income distribution, evaluated at exactly the same thresholds, $\underline{\theta}$ and $\bar{\theta}$. Note that M is at an interior optimum.

For simplicity, suppose that the right-most intersection point $\bar{\theta}$ occurs in the region where both $w_M(\theta, \tau_M)$ and $w_P(\theta, \tau_P)$ are linear, as in Figure 3. Then these optimality conditions imply that in equilibrium $\tau_P(\tau_M^*) < \tau_M^*$. This is because, as shown by equation (12) in the proof of Lemma 4 in appendix B, at $\tau_M = \tau_P$ we have:

$$\left|\frac{\partial \underline{\theta}}{\partial \tau^M}\right| > \frac{\partial \underline{\theta}}{\partial \tau^P} \quad \text{and} \quad \frac{\partial \bar{\theta}}{\partial \tau^M} = \left|\frac{\partial \bar{\theta}}{\partial \tau^P}\right|.$$

[17] Note indeed, that in order to have an equilibrium of this type, it must be that $\underline{\theta} < \mathbf{E}\theta$.

Namely, the lower income threshold $\underline{\theta}$ is more sensitive to changes in τ_M than in τ_P, while the upper threshold is equally sensitive at $\tau_M = \tau_P$.[18] In other words, M gains more votes amongst the poor when it raises its tax rate, compared to what happens when P does the same. The reason is that more redistribution makes poor agents better off, and hence less disappointed, so that their consumption gets closer to x. Since we assume $\mu'''(.) < 0$, they thus become less risk loving, and this increases the number of voters who lean towards M. This effect is absent at the upper threshold $\bar{\theta}$, where swing voters are risk neutral. Hence, the equilibrium must be found at a point where $\tau_P(\tau_M^*) < \tau_M^*$. At this point, tax distortions are higher under M, and the concavity of f implies that $\frac{\partial\bar{\theta}}{\partial\tau^M} > |\frac{\partial\bar{\theta}}{\partial\tau^P}|$. Namely, the upper threshold is also more sensitive to changes in τ_M than in τ_P when $\tau_P(\tau_M^*) < \tau_M^*$, because of the larger tax distortions associated with τ_M. If $\bar{\theta}$ was in the non-linear part of w_M or of w_P, then the same result would also hold, although the expressions for $\frac{\partial\bar{\theta}}{\partial\tau^M}$ and $\frac{\partial\bar{\theta}}{\partial\tau^P}$ would entail additional terms.[19]

In other words, under the stated assumptions the equilibrium is such that, at both margins, the elasticity of votes gained or lost is higher for M than for P. This in turn follows from the intrinsic difference between the two candidates. Because P is risky, it has less to gain from offering redistribution to the poor. The reason is that, by making the poor voters better off, P also reduces its intrinsic attractiveness amongst disappointed voters. Not surprisingly, therefore, in equilibrium P sets lower taxes than M.

[18]Specifically, we have:

$$\frac{\partial\underline{\theta}}{\partial\tau_P} = \frac{f_\tau(\tau_P) - \underline{\theta}(\tau_M, \tau_P)}{(1-\tau_M)R(c_M, c_P) - (1-\tau_P)},$$

while

$$\frac{\partial\underline{\theta}}{\partial\tau_M} = \frac{R(c_M, c_P)[\underline{\theta}(\tau_M, \tau_P(\tau_M)) - f_\tau(\tau_M)]}{(1-\tau_M)R(c_M, c_P) - (1-\tau_P^*(\tau_M))}.$$

where $R(c_M, c_P) := (1 + \mu'(c_M - x))/(1 + \mathbb{E}\mu'(c_P - x))$, and abusing notation $\mathbb{E}\mu'(c_P - x) := \int_{-\varepsilon}^{x-\mathbb{E}c_P} \mu'(\mathbb{E}c_P + \eta - x)h(\eta)d\eta$. Since at the crossing point we have $\mathbb{E}c_P < c_M$ and $\mu'(.)$ is positive and concave, we have $R(c^M, c^P) > 1$.

[19]More formally, if $\bar{\theta}$ is in the linear part of w_M and w, we have

$$\frac{\partial\bar{\theta}}{\partial\tau_P} = \frac{\underline{\theta}(\tau_M, \tau_P) - f_\tau(\tau_P)}{\tau_M - \tau_P},$$

while

$$-\frac{\partial\bar{\theta}}{\partial\tau_M} = \frac{\underline{\theta}(\tau_M, \tau_P) - f_\tau(\tau_M)}{\tau_M - \tau_P}.$$

Whenever f is strictly concave, $\tau_P < \tau_M$ implies $f_\tau(\tau_M) < f_\tau(\tau_P)$. If $\bar{\theta}$ was in the non-linear part of w_M or of w_P, then these expressions would contain additional terms, which we do not report here for ease of exposition, but the argument would contiune to hold.

4.4 Numerical Example

We now illustrate the intuitions of the equilibrium with double crossing with a numerical example. The example is documented in Figure 4 below and Figures 7, 8, 9, 10, 11 and 12 in Appendix E. Appendix Figure 10 illustrates the populist reaction function and displays the discontinuity in $P's$ reaction function: if τ_M is sufficiently low, $\tau_P(\tau_M) > \tau_M$ and it is increasing in τ_M. But once τ_M exceeds a critical value (0.12 in this numerical example), P finds it optimal to jump to a lower tax rate $\tau_P(\tau_M) > \tau_M$. Candidate M in turn finds it optimal to accept this situation, and the Stackelberg equilibrium is found at $\tau_M^* = .22$ and $\tau_P(\tau_M^*) = .18$, where we have double crossing and both candidates are at an interior optimum.

The intuition for the result is similar to that described in the example in the next subsection and in Figure 5. The equilibrium with double crossing emerges because M does not want to loose the support of moderately disappointed voters. In order to get their vote, M must offer sufficiently high redistribution, otherwise P could attract them with a tax rate close to their bliss point. By setting a large τ_M, however, the moderate creates an opportunity for the populist to reduce the tax rate and attract the richest voters. This is indeed what P does in the political equilibrium with double crossing: given that attracting the numerous poor agents close to their reference point is impossible for P, the best move for the populist is to give up some poor voters in exchange for a sizable fraction of the rich.

This numerical example has the same qualitative properties, including $\tau_P^* < \tau_M^*$, if roles are reversed and P is the Stackelberg leader. Appendix Figure 12 illustrates the reaction functions $\tau_P(\tau_M)$ (blue solid line) and $\tau_M(\tau_P)$ (red dashed line). Both reaction functions are discontinuous and they never cross, indicating that a Nash equilibrium does not exist. Appendix Figure 12 also reports the Stackelberg equilibria when P is moving first (the two red squares). These are both at a point where $\tau_P^* < \tau_M(\tau_P^*)$. Proposition 11 in the Appendix C.1 proves that this is not a coincidence, since the key qualitative features of Proposition 5 are preserved even if the timing assumptions are reversed and P moves first. This is not surprising, since with $\tau_P \geq \tau_M$ we have single crossing by Lemma 3. Hence, if the equilibrium was in the region $\tau_P \geq \tau_M$, either the follower (here M) would have to be indifferent between two points, one above and another below the 45 degree line, or there would be policy convergence in equilibrium. Even if P moves first, there cannot be an equilibrium where the follower (M) strictly prefers $\tau_M > \tau_P$.

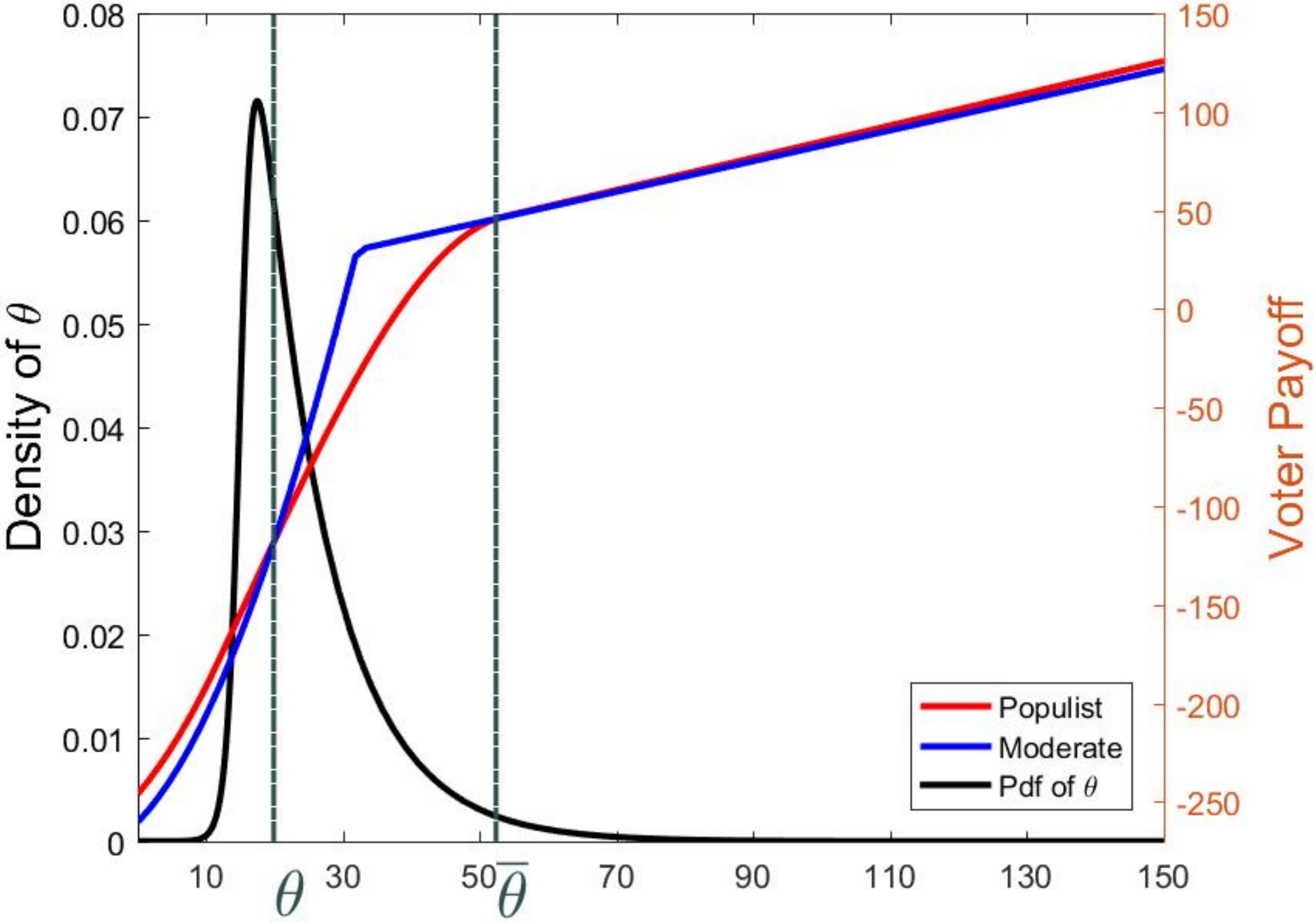

Figure 4: **A numerical example of double crossing.** The figure combines figures 8 and 9 in Appendix E. It reports both the distribution of θ (the left scale) and voters' welfare (reported in the right scale) under the two parties - w_P (in red) and w_M (in blue) - as a function of θ for the equilibrium level of taxes. The vertical dashed lines represent the equilibrium crossing points. Voters prefer M between the two points and P for θ outside the two vertical lines.

4.5 Extensions

In this model there are two dimensions of political conflict. One is the traditional redistributive conflict over tax policy. The other is disagreement over the desired risk properties of candidates. Both dimensions are related to individual income, relative to average income for redistribution, and relative to a common reference point for risk. Preferences for high taxes are decreasing in individual income, θ. Similarly, for any tax rate $\tau \in [0, 1)$, the difference between net income (consumption) and the common reference point x increases with θ. The shape of μ implies that the risk attitude of voters as a function of income is non-monotone, with voters who consume close to the reference point being most averse to risk. Our results generalize to extensions of the model that preserve these qualitative features. In this section we discuss three such extensions.

Risk aversion We assumed that the direct utility of voters is linear in consumption. All our results would still hold if we assumed risk averse voters, namely if $w^M(c, x) = u(c) + \mu(c - x)$ with u strictly increasing and concave. In this case, however, we need two additional 'regularity' assumptions. First, in order to keep individuals with high income not very averse to risk (and hence potentially willing to vote for the populist), u should display Decreasing Absolute Risk Aversion (DARA). Risk aversion penalizes the populist, and DARA implies that such penalty decreases with income. Second, more disappointed individuals (i.e. those with lower $c - x$) should be more likely to prefer the populist candidate in the absence of taxation (as in Lemma 1). For this to be the case, we need to assume that for $c < x$ the function w^M satisfies Assumption 1. In particular, the convexity of μ must dominate the concavity of u - i.e., $u''(c) + \mu''(c - x) > 0$ and $u'''(c) + \mu'''(c - x) \leq 0$ for all $c < x$.

Efficiency The assumption that the populist candidate is less efficient on average (i.e. that $z > 0$), although not implausible, is not necessary, and the main results would continue to hold with suitable adjustments if $z = 0$. If voters were risk averse (with DARA as described above), for the same tax rates all voters with consumption above their reference point would continue to prefer the moderate to the populist even if $z = 0$, because the moderate is safer. Hence, Figures 2 and 3 would maintain their key properties, except that above the reference point utilities would display decreasing concavity, and the proof of Proposition 5 would continue to hold.

If voters were risk neutral and $z = 0$, then voters with $c \geq x + \varepsilon$ would be exactly indifferent between the two candidates if they set the same tax rates - i.e. the functions w_M and w_P in Figure 2 would coincide for $c \geq x + \varepsilon$. Suppose that indifferent voters randomize between the two candidates, so that each candidate gets 1/2 of the indifferent voters. Then it would still be true that we cannot have an equilibrium where candidate P strictly prefers $\tau_P > \tau_M-$ i.e. in equilibrium he is either left indifferent between $\tau_P > \tau_M$ vs $\tau_P < \tau_M$, or he strictly prefers $\tau_P < \tau_M$. However, the equilibrium can no longer have policy convergence ($\tau_P = \tau_M$) at a positive tax rate. The reason is that now an infinitesimal tax cut by P below τ_M would allow him to get 1/2 of the previously indifferent rich voters, and it would cost him only an infinitesimal loss amongst the poor and disappointed voters. Hence policy convergence cannot be an equilibrium. Except for this, the equilibrium would retain all the qualitative properties described above.

Heterogeneous Reference Points We can also relax the assumption that all voters have the same reference point, as long as we preserve the feature that individuals with lower income are more disappointed (i.e., they are further below their idiosyncratic reference point). For example, all our results still hold if the reference point of individuals with income θ is given by: $x(\theta) = (1-\rho)\theta + \rho x_0$ for $1 \geq \rho > \tau_0$ (where τ_0 is the tax rate that maximizes tax revenue), and with x_0 the same for all agents.[20]

As a further generalisation, suppose that x_0 in the previous formulation also varies across individuals. Each agent is now identified by two elements, (θ, x_0), with $x_0 \in [x_{min}, x_{max}]$. As shown in Proposition 12 (ii) in Appendix C.2, without further restrictions we cannot exclude the possibility of an equilibrium where the populist candidate

[20] Here is the outline of the proof. Lemma 1 and Proposition 2 only depend on the difference $\theta - x$, which in this case becomes $\theta - x(\theta) = \rho(\theta - x_0)$. Lemma 3 holds as long as in the initial assumption x is replaced by x_0, that is, we assume $\mathbb{E}\theta \leq x_0 - \varepsilon + z$. The assumption $\rho > \tau_0$ guarantees that even when $\tau = \tau_0$ the difference between disposable income (consumption) and the reference point increases with θ. A formal proof of such crucial result in the even more general case of heterogeneous x_0 is presented in in Appendix C.2, Proposition 12 (i). Lemma 4 also holds: point (i) of the lemma is the consequence of the envelope condition and it can be checked directly that Lemma 4 (ii) holds as well. Propositions 5 and 7 are direct derivations from Lemma 3 and Lemma 4 so they follow from it. Finally, it can be checked directly that the double crossing result of Proposition 13 in the appendix still holds, as long as Assumptions 4 and 5 are kept as stated, and Assumption 3 is replaced by:

$$\frac{f(\tau_0) + z}{\tau_0} - \frac{z - \varepsilon}{\rho} > x_0.$$

strictly prefers $\tau_P > \tau_M$. On point (i) of Proposition 12 however, we show that if we strengthen the assumption made in Lemma 3 to $\mathbf{E}\theta \leq x_{min} - \varepsilon + z$, then $\tau_P > \tau_M$ implies single crossing.[21] In other words, if the populist candidate strictly prefers $\tau_P > \tau_M$, then in equilibrium he is supported by all voters with income below a given threshold. This result replaces Proposition 5 (i), implying that we lose the policy convergence result of Corollary 6. For $\tau_P < \tau_M$, a generalized version of Corollary 7 instead survives, where the single crossing property for a subset of x_0-agents co-exists with a double or more crossing for another x_0-group.[22]

The previous arguments are not based on any distributional assumption of x_0 conditional θ. In particular, we could have a distribution of reference points $x(\theta, x_0)$ that displays a non-monotone (stochastic) relationship with individual income. In this case, we could have equilibria with $\tau_P < \tau_M$ where the populist is supported by a coalition dominated by rich voters and disappointed middle class voters, while most of the poorest voters are not very disappointed and hence support the moderate candidate because they are attracted by its more redistributive policies. Among other results, the next subsection illustrates this possibility with a simple example.

4.6 An example with three income groups

In this section, we analyze a simple example with three income groups. The example serves two purposes. First, it provides an intuition of the conditions under which, in equilibrium, the populist sets a lower tax rate than the moderate and is supported by a coalition of rich voters and of poor disappointed voters. Second, it shows that, when reference points are heterogeneous, it can remain true that in equilibrium $\tau_P < \tau_M$, but now the populist is supported by a different coalition, namely the rich and the disappointed middle class.

[21] Recall that the assumption guarantees that all non-disappointed voters have above average income, and hence cannot be attracted by the populist if he proposes a higher tax rate. We can hence follow the arguments described in footnote 20 and see that the proofs of Lemma 1, Proposition 2, and Lemma 3 apply line by line, *conditional on* x_0, for each $x_0 \in [x_{min}, x_{max}]$. We however lose Lemma 4.

[22] It we want to guarantee not more than two crossings for each x_0 (Proposition 13), we would need to strengthen Assumptions 4, 5 to hold for every $x_0 \in [x_{min}, x_{max}]$, and Assumption 3 to be replaced by

$$\frac{f(\tau_0) + z}{\tau_0} - \frac{z - \varepsilon}{\rho} > x_{max}.$$

.

Throughout we maintain a simplifying assumption, namely that the deadweight loss function $i(\tau)$ takes the following form: $i(\tau) = 0$ for all $\tau \in [0, \tau_0]$ and $i(\tau) = +\infty$ for $\tau > \tau_0$. This assumption implies that all voters below (resp. above) $E[\theta]$ have as most preferred tax rate $\tau = \tau_0$ (resp. $\tau = 0$).

Suppose that there are only three income groups θ_i, $i = 1, 2, 3$. Individuals within each group are identical. The following inequalities hold:

$$\theta_1 < \theta_2 < \mathbf{E}\theta \leq \mathbf{E}\theta - z + \varepsilon < x < \theta_3.$$

This assumption has the following implications. First, since $\theta_1 < \theta_2 < \mathbf{E}\theta < \theta_3$, the ideal tax rate for groups 1 and 2 is $\tau = \tau_0$, whereas voters in group 3 would like to set $\tau = 0$. Second, as $\mathbf{E}\theta \leq \mathbf{E}\theta - z + \varepsilon < x$, groups 1 and 2 are forced to consume below their reference point x, no matter how high the redistributive tax rate is. The voters in these groups are therefore risk-loving. We want to find conditions under which $\tau_M = \tau_0$ and $\tau_P = 0$ is an equilibrium, i.e., the populist sets a lower tax rate than the moderate. Suppose that θ_1 is so far below the reference point that voters belonging to this group prefer to vote for P even when $\tau_M = \tau_0$ and $\tau_P = 0$, as they are attracted by the riskiness of the populist.[23] Then the relevant competition is for the votes of groups 2 and 3. Assume that voters in group 2 prefer M to P when the candidates propose the same tax rate.[24] For them, the inefficiency of the populist prevails over its riskiness. Let $\underline{\tau} > 0$ be the minimum tax rate set by M that prevents P from obtaining the support of group 2.[25] In other words, if $\tau_M > \underline{\tau}$, voters in group 2 prefer M to P irrespective of τ_P. Suppose also that if M sets $\tau_M \geq \underline{\tau}$, then P can obtain the votes of group 3 by setting $\tau_P = 0$.[26] Essentially, we are assuming that the redistributive conflict between the two groups is so strong that there is no tax rate that allows M to secure the votes of both group 2 and group 3, despite its superior efficiency. Finally, suppose that the mass of voters in group 2 is larger than that of group 3. See Figure 5 for a graphical representation of the described scenario.[27]

To find the equilibrium, we work backwards. We start with P's best response corre-

[23] Formally, this is equivalent to assuming $\theta_1 < \underline{\theta}$, where $\underline{\theta}$ is such that $w_M(\underline{\theta}, \tau_0) = w_P(\underline{\theta}, 0)$.

[24] Formally, this is equivalent to assuming $\theta_2 > \hat{\theta}$, where $\hat{\theta}$ is such that $w_M(\hat{\theta}, 0) = w_P(\hat{\theta}, 0)$. It can be shown that the (unique) crossing point $\hat{\theta}(\tau)$ such that $w_M(\hat{\theta}(\tau), \tau) = w_P(\hat{\theta}(\tau), \tau)$ increases with τ.

[25] Formally, this is equivalent to assuming $w_M(\theta_2, \underline{\tau}) = w_P(\theta_2, \tau_0)$.

[26] Formally, this is equivalent to assuming $\theta_3 > \bar{\theta}_2$ where $\bar{\theta}_2$ is such that $w_M(\bar{\theta}_2, \underline{\tau}) = w_P(\bar{\theta}_2, 0)$.

[27] Existence of such thresholds values for an appropriate choice of parameters can be guaranteed, and they are available upon request.

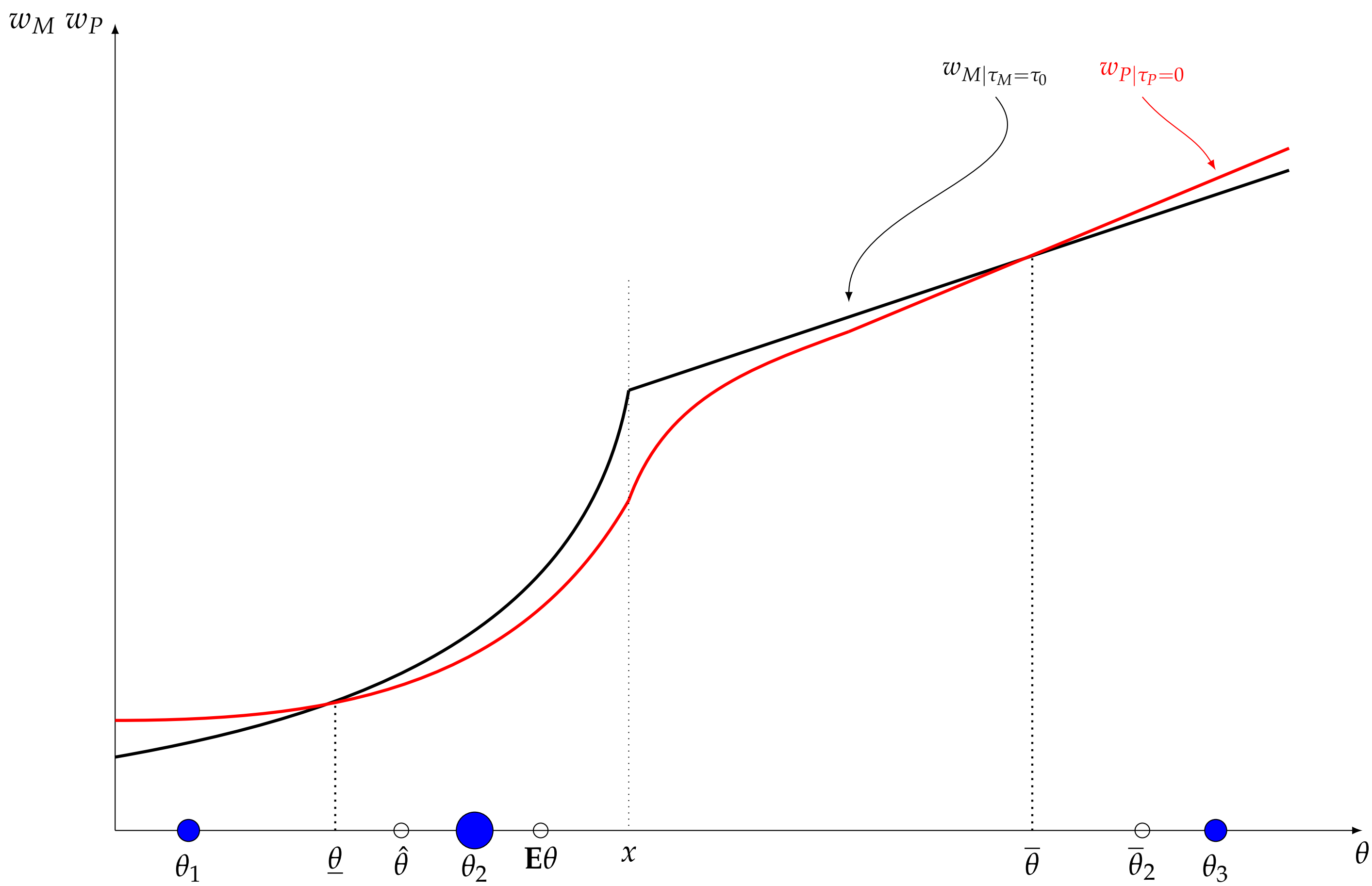

Figure 5: **Example of double crossing equilibrium.** In the picture above, only blue bullets have positive mass and the size of the bullets is proportional to the mass of voters at the point. The levels $\underline{\theta}$ and $\overline{\theta}$ correspond, respectively, to the lower and upper thresholds when $(\tau_M, \tau_P) = (\tau_0, 0)$, our equilibrium outcome. The level θ_2 also corresponds to the lower crossing point for the pair of policies $(\tau_M, \tau_P) = (\underline{\tau}, \tau_0)$, while $\overline{\theta}_2$ corresponds to the upper threshold in the case where $(\tau_M, \tau_P) = (\underline{\tau}, 0)$. Finally, $\hat{\theta}$ refers to the (unique) crossing for $(\tau_M, \tau_P) = (0, 0)$ (all crossing with equal taxes lie below $\hat{\theta}$). It can be verified, that, given the setup, $\tau_P = \tau_0$ belongs to the P's best response for $\tau_M < \underline{\tau}$ while $\tau_P = 0$ is optimal whenever $\tau_M > \underline{\tau}$. Since the mass of points at θ_2 is larger than that at θ_3, the moderate will choose $\tau_M > \underline{\tau}$ and the populist will optimally set a lower tax rate. In particular, the pair $(\tau_M, \tau_P) = (\tau_0, 0)$ is an equilibrium outcome.

spondence. For $\tau_M < \underline{\tau}$, P can gain the votes of group 2 by setting $\tau_P = \tau_0$. Given our assumption on the mass of the two groups, this is optimal for P.[28] For $\tau_M > \underline{\tau}$ instead P cannot obtain the votes of group 2. However, by setting $\tau_P = 0$, he can obtain the votes of group 3. Hence $\tau_P = 0$ is optimal in this case for P (in case of indifference, we assume votes are equally split across the two parties).[29] Now consider M's problem. Obviously, M will target group 2. So M sets a tax $\tau_M > \underline{\tau}$ (because of the equal split assumption, M will never chose $\tau_M = \underline{\tau}$). P will optimally reply by setting $\tau_P = 0$ to attract the votes of group 3. Therefore the pair $(\tau_M, \tau_P) = (\tau_0, 0)$ is an equilibrium outcome.

Intuitively, in the example, candidate M has an advantage with both groups 2 and 3, because he is more efficient than P and the risk loving preferences of group 2 are not sufficiently strong to compensate for the inefficiency. At the same time, the redistributive conflict between θ_2 and θ_3 is so strong that there is no tax rate that allows candidate M to receive the votes of both 2 and 3. M has to choose which group to target. In our example, group 2 is larger, hence M sets $\tau_M = \tau_0$ and P responds with $\tau_P = 0$ to attract group 3.[30] Candidate P is then supported by a coalition of the extremes if there is a group of very disappointed voters who prefer P because of its riskiness, despite a policy of low redistribution.

The example shows some of the key features under which in equilibrium the populist sets a lower tax rate than the moderate and is supported by a coalition of disappointed voters, group 1, and rich voters, group 3. First, in our example the interest of groups 2 and 3 are sufficiently divergent so that the moderate cannot obtain for sure the votes of both groups with the same tax rate. In other words, the redistributive conflict between the "disappointed middle class" and the rich" must be noticeable: for instance, the income of group 3 must be well above the average. At the same time, none of these two groups must be "captive" for the moderate. This implies that the inefficiency of the populist and the disappointment of the middle class should not be very large. If these conditions are not satisfied, the moderate can always set a tax rate that will induce both groups to vote for him. Second, group 2 must be larger than group 3. If not, the moderate would set a low tax rate to target the rich and the populist would have no incentive to set an even lower

[28]The solution is not unique, however τ_0 is always in the P's best response correspondence for $\tau_M < \underline{\tau}$.

[29]Again, the optimal choice for P is not unique. Note that, since $\hat{\theta}(\tau) < \theta_2 \ \forall \tau$, P's best response correspondence only includes tax rates *strictly lower* than τ_M.

[30]If group 3 were larger, then we would have the opposite situation: candidate M sets a low tax rate and receives the votes of group 3, while P sets a high tax rate and obtains the votes of groups 1 and 2.

tax rate. Third, we have assumed in the example that voters in group 1 are 'captive' for the populist. This assumption could be relaxed. The crucial point is that, in case the moderate and the populist candidate propose the same tax rate, the "poor" prefers the populist thanks to his riskiness. Finally, the rich should not be very disappointed. Otherwise, the populist could attract their votes without setting a lower tax rate than the moderate.

In this example, the populist is supported by the rich, who are attracted by the lower tax rate, and by the very disappointed poor, who are attracted by its riskiness. Nevertheless, there is evidence that support for populist parties is often higher amongst low-middle class individuals who have fallen behind in the social scale, and not necessarily amongst the very poor (see for instance Inglehart and Norris 2016 and Gidron and Hall 2017). A modification of the example that allows for heterogeneous reference points among groups can explain this empirical finding. Suppose now that the middle income group is more disappointed than the group with lowest income (i.e., $0 < x_1 - \theta_1 < x_2 - \theta_2$ and $\theta_2 > \theta_1$).[31] Then, under suitable assumptions on parameter values, using the same logic of the previous example, one can show that the equilibrium strategies are the same as before (i.e. the populist candidate sets a lower tax rate than the moderate), but the populist candidate is now supported by a coalition composed of groups 2 and 3, whereas the moderate obtains the votes of group 1. In other words, if disappointment is no longer an increasing function of individual income, then the voters who are attracted by the populist because of its riskiness are not necessarily those at the bottom of income distribution. In particular, the middle class (group 2, in our example), when very disappointed, can be more attracted by the populist's riskiness than the poor, who in turn support the moderate because he proposes a higher redistribution.

5 Discussion

The implications of prospect theory have been extensively explored in finance and in economics, but not in politics. Yet, the idea that economically disappointed voters are attracted by risk is as important in politics as in economics (e.g., see Quattrone and Tversky 1988). In this paper we have argued that this idea can contribute to explain the success

[31] This case can be contemplated by the extension of Section 4.5 with heterogeneous reference points where $\rho = 1$, $x_{min} \leq x_1 < x_2 \leq x_{max}$ and the conditional distribution of x_0 is degenerate on those points.

of new radical and populist parties after large negative economic shock. We have also shown that electoral competition between a safe and a risky opponent can lead to an equilibrium where the more risky candidate runs on a platform of lower redistributive taxes, and is supported by an ”unwieldy” coalition of very disappointed and very rich voters. This happens for two reasons. First, prospect theory implies that voters at the extremes are the least opposed to risky candidates. Second, the more risky candidate knows that he can retain several disappointed supporters even if he pursues a policy of low redistribution that is suboptimal for them. If the safe incumbent were to pursue a similar policy of low redistribution, he would face a steeper electoral loss because risk loving and low income voters would be more ready to abandon him. In other words, it is the intrinsic preference for risk of low income and disappointed voters that induces policy divergence and gives rise to the ”unwieldy” coalition supporting the populist candidate.

Note that not all economic downturns are predicted to have this effect. Disappointment is due to the contrast between current and planned well being. A gradual and predictable downturn need not generate disappointment and the associated rise in risk preferences, if it also induces a decline in the reference point. This may explain why support for extremist and populist parties is often associated with large financial crises, rather than with more regular and predictable business downturns (Funke et al. 2016). The extent to which the welfare system and the labor market shields voters from adverse shocks, or helps them to find other economic opportunities, could also have an impact on the phenomena studied in this paper. Economies were larger groups of workers are more vulnerable to harsh and prolonged economic disappointment are more likely to support radical and unconventional parties in times of crisis.

The same logic also suggests that the support for populist parties after a deep recession could be short lived. In this paper we have not explored the dynamic implications of prospect theory. But reference points are unlikely to be immutable, and individuals could adjust their expectations to persistent economic conditions. If so, economic disappointment could be mitigated by the passage of time, which in turn would reduce the attractiveness of more risky candidates. Exploring the dynamic implications of time varying reference points in a political setting is an interesting area of future research.

Financial crises differ from regular business fluctuations not only because they are large and unexpected, but also because often they can be blamed on the economic and political establishment. This points to another behavioral feature of populist parties: they appeal to popular resentment against economic and political élites that appear to have let down the common man. A natural question for future research is whether our theoretical results on policy divergence and the ”unwieldy coalition” also extend to a setting where two competing candidates are intrinsically different in this dimension of resentment or blame by disappointed voters, and not only in their risk properties.

Most analyses of populism presume or imply that it is a bad thing. In our model, the intrinsic riskiness of the populist candidate is a desirable feature, at least for some voters. This raises the question of why both candidates cannot tailor the risk taking implications of their policies to voters' demands. Indeed, some papers have asked what may induce politicians to undertake more or less risky policies (eg. Aghion and Jackson 2016). Not all policies are chosen in the course of electoral competition, however, and the propensity to undertake new and untested policies, or to make risky future decisions, is an important intrinsic feature of political representatives. This in turn raises the question of how political candidates are selected by parties, and what may induce the selection of safer or more risky political candidates.

Finally, risky policies have an impact not only on domestic voters, but also on other countries or international organizations. Is there something that other countries or international organizations can do to prevent the rise of political risk in neighboring countries? Our model suggests that risky and radical governments are more likely to win office after a large negative shock. Measures that mitigate the impact of the negative shock may thus help moderate and safer candidates, also to the benefit of other countries and international organizations negatively affected by risky policies

To fix ideas, take the perspective of an external creditor. The creditor may be harmed if the more risky candidate is elected, because in case of bad outcomes the country may be unable to fully repay its external debt. Risk is disliked by creditors, because the upside benefit is captured by the debtor, while the downside risk is borne by the creditor in the event of default. Interestingly, a creditor can influence the election outcome in a country in its preferred direction, i.e., against the more risky candidate, by unilaterally forgiving part of the debt. The key insight is that voters' risk appetite is decreasing in their level of consumption. By unilaterally reducing the debt, the creditor increases voter's expected consumption and therefore favors the election of the less risky candidate. Obviously, the creditor trades-off the cost of reducing the debt vs the lower repayment expected if the risky candidate is elected. In some circumstances, the cost for the creditor of forgiving part of the debt may be lower than the cost of a partial repayment when the risky candidate is elected. It is worth mentioning that the perspective of debt default, even only a partial default, may by itself induce voters to elect the populist, as their expected consumption may increase. This is the well-known risk-shifting effect analyzed in corporate finance by Jensen and Meckling (1976). Debt forgiveness may be an optimal response of the creditor even when voters preferences are not reference-dependent. On the other hand, with reference-dependent preferences unilateral debt forgiveness can be optimal even in the absence of a direct benefit for the the voters, as long as the election of a risky candidate harms the creditor.

This lesson is more general than the example of debt repayment, and applies also to the design of conditionality in granting assistance loans. If conditionality imposes harsh and unexpected pain to a large number of voters, it may backfire. Disappointed voters may react by choosing the more risky political option, even if it is economically less efficient, because it is the only alternative that may give them a chance to get back to their expected level of well being. The recent European Recovery Fund and the Marshall Plan are examples of international aid programs that may have been designed also taking into account this general insight.

References

[1] Aghion, Philippe and Matthew Jackson (2016) ”Inducing Leaders to Take Risky Decisions: Dismissal, Tenure, and Term Limits”, *American Economic Journal, Microeconomics*, Vol 8, N.3 pp 1-38

[2] Alesina, Alberto and Francesco Passarelli (2019), “Loss Aversion in Politics”, *American Journal of Political Science*, 936-947

[3] Alesina Alberto and Alex Cukierman (1990), ”The Politics of Ambiguity”, *Quarterly Journal of Economics*, vol CV, 4

[4] Apostol, T. M., 1957, *Mathematical analysis: A modern approach to advanced calculus*. Addison-Wesley series in mathematics. Reading, Massachusetts: Addison Wesley.

[5] Aragones Enriqueta and Zvika Neeman (2000) ”Strategic Ambiguity in Electoral Competition”, *Journal of Theoretical Politics*

[6] Autor, David, David Dorn, Gordon Hanson and Kaveh Majlesi (2020), Importing Political Polarization? The Electoral Consequences of Rising Trade Exposure, *American Economic Review*, October, vol 110 n.10

[7] Barberis, Nicholas C. (2013), “Thirty Years of Prospect Theory in Economics: A Review and Assessment”, *Journal of Economic Perspectives*, 27 (1) 173-96

[8] Balduzzi, P, E Brancati, M Brianti, and F Schiantarelli (2020), “Populism, political risk and the economy: Lessons from Italy”, IZA Discussion Paper 12929.

[9] Besley, Tim and Torsten Persson (2019) ”Democratic Values and Institutions”, *American Economic Review Insights* 1, 59-76.

[10] Beyme, von Klaus (1988), “Right-Wing Extremism in Post-War Europe”, *West European Politics,* volume 11

[11] Binzel, Christine and Jean-Paul Carvalho (2016) ”Education, Social Mobility and Religious Movements: The Islamic Revival in Egypt”, *The Economic Journal*

[12] Bonomi, Giampaolo, Nicola Gennaioli and Guido Tabellini (2020), “Identity, Beliefs, and Political Conflict”, mimeo

[13] Colantone, Italo and Piero Stanig (2018), “The Trade Origins of Economic Nationalism: Import Competition and Voting Behavior in Western Europe”, *American Journal of Political Science*, 62 (4) 936-953

[14] Dal Bo, Ernesto, Frederico Finan, Olle Folke, Torsten Persson, and Johanna Rickne (2018), “Economic Losers and Political Winners: Sweden’s Radical Right”, mimeo, University of Stockholm

[15] Dieudonné, J., 1969, ‘Foundations of modern analysis.’ Number 10-I in *Pure and Applied Mathematics*. New York: Academic Press. Volume 1 of Treatise on Analysis

[16] Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, J. Schupp, and Gert. G. Wagner (2011), “Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences”, *Journal of the European Economic Association*, 9(3) 522-550

[17] Dohmen, Thomas, Hartmut Lehmann and NorbertocPignatti (2016), “Time-varying individual risk attitudes over the Great Recession: A comparison of Germany and Ukraine”, *Journal of Comparative Economics* 44, pp. 182-200

[18] Dornbusch, R. and S. Edwards (1991) *The Macroeconomics of Populism in Latin America*, University of Chicago Press, Chicago

[19] Edwards, S. (2019) ”On Latin American Populism, and Its Echoes around the World”, *Journal of Economic Perspectives* 33(4), pp. 76-99

[20] Funke, Manuel, Moritz Schularick, and Christoph Trebesch (2016). ”Going to Extremes: Politics after Financial Crises, 1870-2014” *European Economic Review*.

[21] Gidron, Norman, and Peter A. Hall (2017), “The politics of social status: economic and cultural roots of the populist right”, *British Journal of Sociology*, 68 (S1), S57-S84.

[22] Grillo Edoardo and Carlo Prato (2020), ”Reference Points and Democratic Backsliding”, mimeo

[23] Guriev, Sergei and Elias Papaioannou (2020) ”The Political Economy of Populism”, *Journal of Economic Literature*, forthcoming

[24] Groseclose, Tim (2001), ”A Model of Candidate Location When One Candidate Has a Valence Advantage”, *American Journal of Political Science*, Oct. vol. 45 N.4, pp. 862-86

[25] Guiso, Luigi, Helios Herrera, Massimo Morelli, and Tommaso Sonno (2017), "Demand and Supply of Populism", EIEF Working Papers Series 1703, Einaudi Institute for Economics and Finance (EIEF).

[26] Herrera, H., Ordonez, G. and C. Trebesch (2019)"Political Booms, Financial Crises", *Journal of Political Economy*, 128(2): 507-43

[27] Huber, John (2017) *Exclusion by Elections: Inequality, Ethnic Identity and Democracy*, Cambridge University Press, Cambridge

[28] Krasa, Stefan, and Mattias Polborn (2010), "Competition between Specialized Candidates" *American Political Science Review* 104 (4): 745–65.

[29] Krasa, Stefan, and Mattias Polborn (2012), "Political Competition between Differentiated Candidates." *Games and Economic Behavior* 76 (1): 249–71.

[30] Krasa, Stefan, and Mattias Polborn (2014), "Social ideology and Taxes in a Differentiated Candidates Framework", *American Economic Review* 2014, 104(1): 308–322

[31] Inglehart, Ronald F., and Pippa Norris (2016), "Trump, Brexit, and the Rise of Populism: Economic Have-Nots and Cultural Backlash", HKS Working Paper No. RWP16-026

[32] Jensen, Michael C., and William H. Meckling (1976), "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure", *Journal of Financial Economics*, 3(4) 305-360

[33] Kahneman, Daniel, and Amos Tversky (1979), "Prospect Theory: An Analysis of Decision under Risk", *Econometrica*, 47 (2) 263-291

[34] Karakas, Leyla D. and Devashish Mitra (2020) "Inequality, Redistribution and the Rise of Outsider Candidates", mimeo

[35] Koszegi, Botond, and Matthew Rabin (2006), "A Model of Reference-Dependent Preferences", *The Quarterly Journal of Economics,* 121(4) 1133-1165

[36] Koszegi, Botond, and Matthew Rabin (2007), "Reference-Dependent Risk Attitudes", *American Economic Review*, 97(4) 1047-73

[37] Koszegi, Botond, and Matthew Rabin (2009), “Reference-Dependent Consumption Plans”, *American Economic Review*, 99(3) 909-36

[38] Levy, Gilat, Ronny Razin and Alwyn Young (2020), ”Misspecified Politics and the Recurrence of Populism”, mimeo

[39] Lockwood, Ben, and James Rockey (2015), “Negative Voters: Electoral Competition with Loss-Aversion”, Working Paper No. 15/15, University of Leicester

[40] Mian Atif, Amir Sufi and Francesco Trebbi, (2014), ”Resolving Debt Overhang: Political Constraints in the Aftermath of Financial Crises”, *American Economic Journal: Macroeconomics,* 6 (2): 1-28.

[41] Milgrom, P. and I. Segal, (2002): “Envelope Theorems for Arbitrary Choice Sets,” *Econometrica*, 70, 583?601.

[42] Mueller, J.W. (2017) *What is Populism?* Penguin UK

[43] Norris, Pippa. *The Global Party Survey*, 2019. V1.0 www.GlobalPartySurvey.org

[44] Norris, Pippa, and Ronald Inglehart (2019), *Cultural Blacklash: Trump, Brexit and, Authoritarian Populism*, Cambridge University Press

[45] Pagel, Michaela (2017), ”Expectations-Based Reference-Dependent Life-Cycle Consumption.” *The Review of Economic Studies* 84, no. 2: 885-934.

[46] Passarelli, Francesco and Guido Tabellini (2017), “Emotions and Political Unrest”, *Journal of Political Economy*, 125(3) 903-946

[47] Rode, M and Revuelta, J. (2015), ”The Wild Bunch! An Empirical Note on Populism and Economic Institutions”, *Economics of Governance* 16(1): 73-96

[48] Quattrone, Gerge A. and Amos Tversky (1988), “Contrasting Rational and Psychological Analyses of Political Choice”, *American Political Science Review,* 82(3) 719-736

[49] Shepsle, Kenneth (1972), ”The Strategy of Ambiguity: Uncertainty and Electoral Competition”, *The American Political Science Review* , Vol. 66, No. 2 pp. 555-568

Appendix

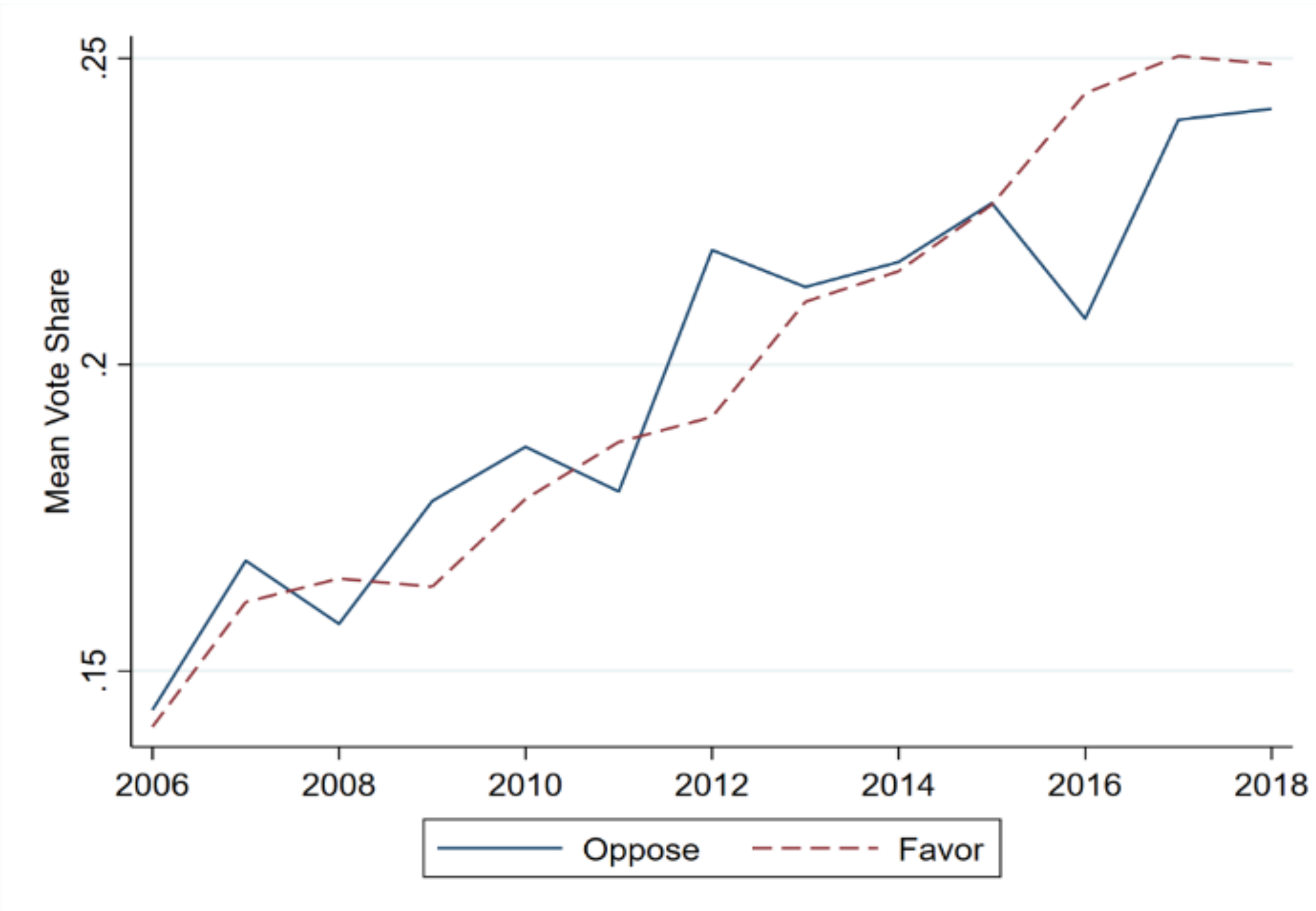

Figure 6: **Vote shares of Populist parties, by their stance on redistribution.** The figure plots the average vote share of 64 populist parties (as defined by Norris, 2019) across 26 European countries between 2006 and 2018. For each country and for each year we aggregate the vote share of all populist parties obtained in the latest National Elections. Parties are divided according to their stance on redistribution as measured in the 2014 Chapel Hill Expert Survey by the variable Redistribution. This variable measures the party position on redistribution of wealth from the rich to the poor, and ranges from 0= Strongly favors redistribution to 10=Strongly opposes redistribution. The solid blue (resp. dashed red) line represents parties with a value on the redistribution scale strictly greater (resp. strictly lower) than 5. European countries are: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Romania, Slovakia, Slovenia, Spain, Sweden, United Kingdom.

A More on the Empirical Analysis

The German Socio Economic Panel is a large longitudinal survey of private households in Germany, administered by the German Institute for Economic Research, DIW Berlin. SOEP targets the whole German population, and is composed of several sub-samples, selected as multi-stage regionally-clustered random samples, in which the respondents (households) are selected via a random-walk. From there on, samples are re-interviewed across each wave, undergoing attrition. The questionnaire is physically run as a face-to-face interview with all members of a given survey household aged 16 years and over. If the face-to-face interview is refused a telephone interview or e-mail is admitted. In practice, the respondent is not monetarily incentivized: she only receives catchy information on the possible use of the questionnaire, a letter of thanks after the questionnaire, a small gift (worth 5 to 10 DM, or 3 to 6 Euros) and a ticket for a famous lottery. One person in the household (the self-selected "head" of the household) is also asked some questions covering the whole household (e.g. dwelling, total household income). At the end of the survey, some variables are imputed by the SOEP staff, based on objective measurement by the interviewer (e.g. region, dwelling charateristics) or on statistical imputation procedures.

The survey question for individual income directly asks each respondent for her income and working hours, while the survey question for household aggregate net income is asked only to the household head: "If you take a look at the total income of all members of the household: How much is the monthly household net income today?". In the analysis, we use the household post-government income variable produced for the CNEF in the Cross-National Equivalent Files PEQUIV, which aggregates individual answers in the household, and is adjusted for inconsistent estimates and non-responses.

Table A1: **Summary Statistics**

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
Populist Dummy	212,390	0.011	0.106	0	1
Risk Love	218,145	4.593	2.396	0	10
Extreme Risk Love	218,145	0.039	0.192	0	1
Income Dissatisfaction	225,217	3.445	2.288	0	10
Extreme Income Dissatisfaction	225,217	0.114	0.318	0	1
Large Income Loss	140,581	0.077	0.266	0	1
Male	238,597	0.463	0.499	0	1
Age Under 28	238,597	0.135	0.342	0	1
Age Over 45	238,597	0.517	0.500	0	1
Direct Immigrant	238,597	0.169	0.375	0	1
Second-Generation Immigrant	238,597	0.072	0.259	0	1
Low Education Level	233,219	0.162	0.369	0	1
Medium Education Level	233,219	0.548	0.498	0	1
High Education Level	233,219	0.290	0.454	0	1
Income	228,540	10.40	0.568	8.587	11.66
Resident in East Germany	238,597	0.223	0.416	0	1
Unemployed	238,597	0.051	0.220	0	1
Out of Labor Force	238,597	0.358	0.479	0	1

Table A2: **Correlations**

	Populist (Dummy)	Risk Love	Extreme Risk Love	Income Dissat.	Extr. Income Dissat.	Large Income Loss	Unemployed	Income
Populist (Dummy)	1							
Risk Love	0.0356 0.0000	1						
Extreme Risk Love	0.0339 0.0000	0.4043 0.0000	1					
Income Dissatisfaction	0.0377 0.0000	-0.0440 0.0000	0.0143 0.0000	1				
Extreme Income Dissatisfaction	0.0372 0.0000	-0.0279 0.0000	0.0330 0.0000	0.7093 0.0000	1			
Large Income Loss	0.0054 0.0443	0.0032 0.2377	0.0184 0.0000	0.1005 0.0000	0.0836 0.0000	1		
Unemployed	0.0252 0.0000	0.0111 0.0000	0.0354 0.0000	0.2406 0.0000	0.2169 0.0000	0.0691 0.0000	1	
Income	-0.0130 0.0000	0.0898 0.0000	-0.0127 0.0000	-0.3857 0.0000	-0.2551 0.0000	-0.2258 0.0000	-0.1947 0.0000	1

Table A3: **Pattern of Missing Observations, Panel A**

Dep. var.	Missing Observation Dummy					
	(1)	(2)	(3)	(4)	(5)	(6)
Risk Love	-0.0063**	-0.0172***	-0.0124***	-0.0163***	-0.0022	-0.0027
	(0.0031)	(0.0033)	(0.0033)	(0.0035)	(0.0054)	(0.0055)
Extreme Risk Love		0.3639***		0.1371***	-0.0204	-0.0078
		(0.0351)		(0.0375)	(0.0551)	(0.0562)
Unemployed			0.1610***	0.1590***		0.0352
			(0.0356)	(0.0356)		(0.0558)
Income			-0.2844***	-0.2831***		0.0581*
			(0.0168)	(0.0168)		(0.0331)
Observations	206,751	206,751	197,138	197,138	86,521	83,658
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Other Controls	No	No	Yes	Yes	No	No
Individual FE	No	No	No	No	Yes	Yes

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered at the individual level. The dependent variable is a dummy equal to one if the individual does not answer to the question on party leaning in SOEP survey. The sample does not include individuals answering to a questionnaire in which such a question is not available. In all columns, we also control for being in or out the labor force. Other individual controls (where included) are dummy variables for gender, age group, immigrant status, education level, and macroregion. Estimation is by logit in columns 1, 2, 3 and 4, by conditional logit in the last two columns. Source: SOEP.

Table A4: **Pattern of Missing Observations, Panel B**

Dep. var.	Missing Observation Dummy					
	(1)	(2)	(3)	(4)	(5)	(6)
Income Dissatisfaction	0.0704***	0.0077	0.0723***	0.0080		
	(0.0048)	(0.0076)	(0.0049)	(0.0077)		
Extreme Income	-0.1676***	-0.0725*	-0.1727***	-0.0569		
Dissatisfaction	(0.0289)	(0.0429)	(0.0294)	(0.0438)		
Risk love			-0.0141***	-0.0019		
			(0.0036)	(0.0056)		
Extreme Risk Love			0.1387***	-0.0224		
			(0.0383)	(0.0579)		
Large Income Loss					-0.0429	0.0229
					(0.0261)	(0.0449)
Unemployed	0.0746**	-0.0043	0.0902**	0.0208	0.1652***	0.1464*
	(0.0356)	(0.0558)	(0.0367)	(0.0581)	(0.0484)	(0.0768)
Income	-0.2052***	0.0600*	-0.2036***	0.0544	-0.3720***	0.0038
	(0.0173)	(0.0338)	(0.0176)	(0.0346)	(0.0240)	(0.0660)
Observations	194,457	82,318	189,201	79,919	134,270	55,793
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Other Controls	Yes	No	Yes	No	Yes	No
Individual FE	No	Yes	No	Yes	No	Yes

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered at the individual level. The dependent variable is a dummy equal to one if the individual does not answer to the question on party leaning in SOEP survey. The sample does not include individuals answering to a questionnaire in which such a question is not available. In all columns, we also control for being in or out the labor force. Other individual controls (where included) are dummy variables for gender, age group, immigrant status, education level, and macroregion. Estimation is by logit in columns 1, 3 and 5, by conditional logit in the remaining columns. Source: SOEP.

B Proofs

Proof of Lemma 1

Proof. We take the derivative of $\Delta(\cdot)$ with respect to its argument in different points. Even though some points are not differentiable, the function is continuous as a difference of continuous functions, so the finitely many points of non-differentiability have zero effect on the level of Δ.

Consider first the range where $t > 0$. In this case, $\Delta(t)$ is strictly positive over the whole range. In fact, for this range, $\mu(t) = 0$, and hence

$$\Delta(t) = z - \int_{-\varepsilon}^{\varepsilon} \mu(t - z + \eta)h(\eta)d\eta > 0$$

as $\mu(\cdot) \leq 0$. Also, for $-\varepsilon > z - t$, $\Delta(t) = z$, whereas for $-\varepsilon < z - t$, $\Delta(t) = z - \int_{-\varepsilon}^{z-t} \mu(t - z + \eta)h(\eta)d\eta > 0$, which is decreasing in t as $\mu' > 0$ and $\mu(0) = 0$.

We now consider the range where $t < 0$. We will show that $\Delta(\cdot)$ always increases in this range. Note first that for $t < z - \varepsilon$ all arguments in μ are negative. This implies that

$$\Delta'(t) = \mu'(t) - \int_{-\varepsilon}^{\varepsilon} \mu'(t - z + \eta)h(\eta)d\eta \geq \mu'(t) - \int_{-\varepsilon}^{\varepsilon} \mu'(t + \eta)h(\eta)d\eta \geq 0.$$

The first inequality is implied by the convexity of μ (increasing first derivative, Assumption 1 (iii)) and $z > 0$. The second inequality is the consequence of the Jensen's inequality and Assumption 1 (iv) (the function μ has concave first derivative). Obviously, when μ is strictly concave, the second inequality will be strict and hence Δ will be strictly increasing in this range. Finally consider the range $z - \varepsilon < t < 0$. In this range, we have

$$\Delta(t) = z + \mu(t) - \int_{-\varepsilon}^{z-t} \mu(t - z + \eta)h(\eta)d\eta,$$

and hence

$$\Delta'(t) = \mu'(t) - \int_{-\varepsilon}^{z-t} \mu'(t - z + \eta)h(\eta)d\eta \geq \mu'(t) - \int_{-\varepsilon}^{z-t} \mu'(t - z + \eta)h(\eta)d\eta - \int_{z-t}^{\varepsilon} h(\eta)d\eta\mu'_{-}(0),$$

where we used $\mu(0) = 0$. The inequality is immediate given $\mu'_{-}(0) > 0$ and $\varepsilon > z - t$. Furthermore, if we consider any convex and continuous extension of μ, it must be that $\mu'_{-}(0) \leq \mu'(t - z + \eta)$

for $\eta \in (z-t, \varepsilon)$.[32] Taking an extension that satisfies Assumption 1 (iv) as well, we hence have

$$\Delta'(t) \geq \mu'(t) - \int_{-\varepsilon}^{z-t} \mu'(t-z+\eta)h(\eta)d\eta - \int_{z-t}^{\varepsilon} \mu'_{-}(0)h(\eta)d\eta \geq \mu'(t) - \int_{-\varepsilon}^{\varepsilon} \mu'(t-z+\eta)h(\eta)d\eta > 0,$$

where the last inequality is implied the Jensen's inequality and by Assumption 1 (iii) (convex μ). Again, if Assumption 1 (iii) is strict, we have strictly decreasing $\Delta(\cdot)$ in this range as well. □

B.1 Socially optimal tax rate

The social planner maximizes the utilitarian optimum, namely:

$$\int_{0}^{\infty} c(\theta,\tau)dG(\theta) + \int_{0}^{\theta_x(\tau)} \mu[c(\theta,\tau)-x]dG(\theta), \tag{5}$$

where, recall, $G(.)$ is the distribution of θ, and $\theta_x(\tau)$ is such that $c(\theta_x(\tau),\tau) = x$.

Taking the derivative with respect to τ, and recalling that $\mu[c(\theta_x(\tau),\tau) - x] = 0$, we obtain:

$$-i'(\tau) + \int_{0}^{\theta_x(\tau)} \mu'[c(\theta,\tau)-x][\mathbf{E}\theta - \theta - i'(\tau)]dG(\theta). \tag{6}$$

Now evaluate this expression at the point $\tau = 0$. Since $i'(0) = 0$, we can write it as:

$$\int_{\check{\theta}}^{x} (\mathbf{E}\theta - \theta)\mu'(\theta - x)dG(\theta). \tag{7}$$

If $\mathbf{E}\theta > x$ and there is a positive measure of agents with $\theta < x$ this expression is clearly strictly positive. Hence $\tau = 0$ cannot be a solution. Intuitively, disappointed agents are loss averse, and hence they have a higher marginal utility of income. Since $\mathbf{E}\theta > x$, all disappointed agents are poorer than the average. Hence, a positive tax redistributes from non-disappointed to disappointed agents, and this increases social welfare. Note that convexity plays no role here, just loss aversion is sufficient for this result.

[32] Consider the linear extension for example.

If $\mathbf{E}\theta < x$ then (7) can be written as:

$$\int_0^{\mathbf{E}\theta} (\mathbf{E}\theta - \theta)\mu'(\theta - x)dG(\theta) + \int_{\mathbf{E}\theta}^{x} (\mathbf{E}\theta - \theta)\mu'(\theta - x)dG(\theta). \tag{8}$$

The first term is positive, the second term is negative, hence this expression has an ambiguous sign. Intuitively, some disappointed agents have above average income, so we can no longer tell whether a positive tax rate is optimal.

Rewriting (8), the condition for $\tau > 0$ to be an optimal solution is:

$$\int_0^{\mathbf{E}\theta} (\mathbf{E}\theta - \theta)\mu'(\theta - x)dG(\theta) > \int_{\mathbf{E}\theta}^{x} (\theta - \mathbf{E}\theta)\mu'(\theta - x)dG(\theta) \tag{9}$$

This condition is more likely to be met if:

(i) there are not too many people between $\mathbf{E}\theta$ and x.

(ii) the function $\mu(\cdot)$ is not too convex.

To see why convexity works against (9), note that $\mu'(\cdot)$ is an increasing function. Hence, the marginal utility of income is higher for the disappointed losers from redistribution than from the disappointed beneficiaries (i.e., μ' is uniformly higher on the RHS of (9). Moreover, the LHS of (9) sees higher values of μ' weighted by smaller values of $\mathbf{E}\theta - \theta$, while the opposite happens on the RHS. In other words, the covariance between $(\mathbf{E}\theta - \theta)$ and $\mu'(\theta - x)$ is negative, and this also works against (9) being satisfied.

We summarize these results in the following:

PROPOSITION 8. *The socially optimal tax rate is positive if $\mathbf{E}\theta > x$ and there is a positive measure of agents with $\theta < x$, or if $\mathbf{E}\theta < x$ and (9) holds.*

Suppose that one of these conditions is met, so that the optimal tax rate is positive, ruling out a corner solution. Then, the socially optimal tax rate is implicitly defined by setting (6) equal to 0. Intuitively, at a social optimum the marginal tax distortions, $i'(\tau) > 0$, are equated to the marginal social benefit of redistributing in favor of the disappointed agents (the second term in (6)).

Proof of Lemma 3.

Proof. (i) First of all, note that $c_M(\cdot, \tau_M)$ is strictly monotone in θ for all $\tau_M < 1$. We will hence be able to span all θ by looking at the whole range for c_M.

A. Consider all θ such that $c_M(\theta, \tau_M) \geq x$. Given our assumptions, it must be that $c_M(\theta, \tau_M) \geq \mathbf{E}\theta$. We now show that in this case the two functions w_M and w_P cannot cross because the populist has higher distortions than the moderate and the loss function μ is not acting for the latter. Formally: $c_M(\theta, \tau_M) \geq \mathbf{E}\theta$ and the monotonicity of consumption in θ implies $\theta \geq \mathbf{E}\theta$. As a consequence, $\tau_P \geq \tau_M$ implies that $c_M(\theta, \tau_P) \leq c_M(\theta, \tau_M)$ and $\mathbb{E}c_P(\theta, \tau_P) = c_M(\theta, \tau_P) - z$. Since under M consumption is above the reference point, while under P consumption might end up being below x in the bad state (and $\mu \leq 0$), the difference $w_P(\theta, \tau_P) - w_M(\theta, \tau_M)$ must be larger than z for all such θ.

B. Consider the complement set: θ such that $c_M(\theta, \tau_M) < x$. We here have two cases to consider.
B1. First, assume in addition that $c_M(\theta, \tau_M) \geq \mathbb{E}c_P(\theta, \tau_P)$. That is, we consider all θ such that we have $x > c_M(\theta, \tau_M) \geq \mathbb{E}c_P(\theta, \tau_P)$. The difference between the two functions w_M and w_P is strictly increasing in θ. They might hence cross at most once in this range. Formally, we have:[33]

$$\frac{\partial}{\partial \theta}[w_M(\theta, \tau_M) - w_P(\theta, \tau_P)] = (\tau_P - \tau_M) + (1 - \tau_M)\mu'(c_M - x) - (1 - \tau_P)\mathbb{E}\mu'(c_P - x) > 0,$$

where the last strict inequality is guaranteed even for $\tau_P = \tau_M$ because of Assumption 1 (iv). In particular, note that for any convex extension of μ we have $\mu'(c_M - x) \geq \int_{-\varepsilon}^{+\varepsilon} \mu'(c_P - x)h(\eta)d\eta$ and in any such extension, for $c_P - x > 0$ we must have $\mu'(c_P - x) \geq \mu'_-(0) > 0$.The inequality is hence satisfied a fortiori when for such values $\mu'(c_P - x) = 0$.
B2. Consider now the alternative case where $c_M(\theta, \tau_M) < x$ and $c_M(\theta, \tau_M) < \mathbb{E}c_P(\theta, \tau_P)$. Note that, since $\tau_P \geq \tau_M$, it must be that $\theta \leq \mathbf{E}\theta$. And hence, under our Assumption of f, we have $(1 - \tau_P)\theta + f(\tau_P) \leq \mathbf{E}\theta$ and hence $\mathbb{E}c_P(\theta, \tau_P) = (1 - \tau_P)\theta + f(\tau_P) - z \leq \mathbf{E}\theta - z \leq x - \varepsilon$, where the last inequality is implied by our assumption. This in turn implies that $c_M(\theta, \tau_M) \leq \mathbb{E}c_P(\theta, \tau_P) < x - \varepsilon$.

We now show that the two conditions $x - \varepsilon \geq c_M(\theta, \tau_M)$ and $\mathbb{E}c_P(\theta, \tau_P) \geq c_M(\theta, \tau_M)$ together imply that $w_P(\theta, \tau_P) > w_M(\theta, \tau_M)$. This will show that the functions w_P and w_M cannot cross in the relevant range of θ in this case. In words, we have $w_P(\theta, \tau_P) > w_M(\theta, \tau_M)$ since for this range of θ the populist enjoys higher average consumption and the riskiness of the convex punishment.

[33]Obviously, for $c_P = x$, μ' should be replaced by its left derivative, however this is of measure zero in the integral and we ignore it to simplify notation.

More in detail, since $\mathbb{E}c_P(\theta, \tau_P) \geq c_M(\theta, \tau_M)$, it suffices to show that:

$$\mu\left(c_M(\theta, \tau_M) - x)\right) \leq \mathbb{E}\mu\left(c_P(\theta, \tau_P) - x\right). \tag{10}$$

To see why (15) holds, note that $c_M(\theta, \tau_M) < x - \varepsilon$, so: (i) $\mu(c_M(\theta, \tau_M) - x) < 0$ and (ii) if we define $\tilde{c}_M(\theta, \tau_M) := c_M(\theta, \tau_M) + \eta$ where η takes the values between $-\varepsilon$ and $+\varepsilon$ with density h, from Jensen's inequality, we have $\mathbb{E}\mu\left(\tilde{c}_M(\theta, \tau_M) - x\right) \geq \mu\left(c_M(\theta, \tau_M) - x\right)$. Now, since $\mathbb{E}c_P(\theta, \tau_P) \geq \mathbb{E}\tilde{c}_M(\theta, \tau_M) = c_M(\theta, \tau_M)$ we obtain the result from the monotonicity of μ.

(ii) Since the only case where we can have a crossing point is **B1**, we have also shown the second part of the proposition. Namely, that at the crossing point $\hat{\theta}$, if any, we have $\mathbb{E}c_P(\hat{\theta}, \tau_P) < c_M(\hat{\theta}, \tau_M) < x$.

(iii) It is sufficient to observe that in region **B1** $\frac{\partial(w_M - w_P)}{\partial\theta} > 0$. □

Proof of Lemma 4. As a preliminary result we would like to state some regularity conditions on the objects of analysis.

LEMMA 9. *Assume $f(\cdot)$ is twice continuously differentiable, and $\mu(\cdot)$ are twice continuously differentiable everywhere except at zero. (i) For $\bar{\tau} := (\bar{\tau}_M, \bar{\tau}_P)$, let $\hat{\theta}$ be a threshold for which $w_P(\hat{\theta}, \bar{\tau}_P) = w_M(\hat{\theta}, \bar{\tau}_M)$. Then $\hat{\theta}(\cdot)$ is twice continuously differentiable as a function of (τ_M, τ_P) in an open neighbour of $\bar{\tau}$ as long as $w_{P,\theta}(\hat{\theta}, \bar{\tau}_P) \neq w_{M,\theta}(\hat{\theta}, \bar{\tau}_M)$. (ii) The condition $w_{P,\theta}(\hat{\theta}, \bar{\tau}_P) \neq w_{M,\theta}(\hat{\theta}, \bar{\tau}_M)$ is satisfied for all $\hat{\theta}$ associated to $\bar{\tau}_P \geq \bar{\tau}_M$.*

Proof. (Sketch) (i) The result uses a basic version of the implicit function theorem (e.g., Apostol 1957, page 146) which can be generalized to higher order derivatives (Dieudonné, 1969, pages 265-273). The assumptions imply the function $H : \theta \times [0,1] \times [0,1]$ defining the zero point:

$$H(\theta, \tau_M, \tau_P) := w_M(\theta, \tau_M) - w_P(\theta, \tau_P),$$

is twice continuously differentiable with respect to (θ, τ_M, τ_P) at any triplet $(\hat{\theta}, \bar{\tau}_M, \bar{\tau}_P)$ compatible with a zero of the function. And since $H = w_P - w_M$, the last assumption guarantees that $H_1(\hat{\theta}, \bar{\tau}_M, \bar{\tau}_P) \neq 0$ as required by the theorem. To show that H is twice continuously differentiable we just need to check the $c = x$ cases. First of all, $\hat{\theta}$ cannot be such that $c_M(\hat{\theta}, \bar{\tau}_M) = x$ as this would be incompatible with a crossing point (i.e., a zero of H). In addition the continuity of μ at zero and the integral definition of w_P guarantee twice continuous differentiability of w_P.

(ii) Finally, we show that whenever $\bar{\tau}_P \geq \bar{\tau}_M$, $H_1(\hat{\theta}, \bar{\tau}_M, \bar{\tau}_P) \neq 0$. This is the single crossing case. The result follows since w_P crosses w_M from below and the derivative is changing continuously.

□

LEMMA 10. *Let $\tau_P(\cdot)$ be the Populist's best response (BR) and be such that $\tau_M \leq \tau_P(\tau_M)$. (i) Under the assumptions of Lemma 4, within this range of taxes, the BR is single valued. (ii) If in addition, $\tau_M \leq \tau_P(\tau_M)$ holds in an open neighborood, then $\tau_P(\cdot)$ is continuously differentiable in τ_M. Clearly, (i) and (ii) hold in particular for the policy function $\tau_P^+(\cdot)$ restricted to $\tau_M \leq \tau_P$.*

Proof. (i) Given the monotonicity and continuity of the cumulate G, any point in the BR solves

$$\hat{\theta}_2(\tau_M, \tau_P(\tau_M)) = 0.$$

We now show that $\tau_P(\tau_M)$ is single valued whenever $\tau_M \leq \tau_P(\tau_M)$. First of all, note that, by single crossing, and the strict monotonicity of G all threshold points that maximize populist objective for the same fixed τ_M must deliver the same threshold. Consider the populist problem (3). From (2), The derivative of the threshold level with respect to τ_P equals

$$\frac{\partial \hat{\theta}}{\partial \tau_P} = \frac{f_\tau(\tau_P) - \hat{\theta}(\tau_M, \tau_P)}{(1-\tau_M)R(c_M, c_P) - (1-\tau_P)} \tag{11}$$

where $R(c_M, c_P) := (1 + \mu'(c_M - x)) / \left(1 + \int_{-\varepsilon}^{x - \mathbb{E}c_P} \mu'(\mathbb{E}c_P + \eta - x)h(\eta)d\eta\right)$. Notice this is well defined by Lemma 9. Moreover, since we have shown in Lemma 3 that at the crossing point we have $\mathbb{E}c_P < c_M$ and $\mu'(\cdot)$ is positive, increasing, and concave, we have $R(c_M, c_P) > 1$. Note indeed that - by the Jensen's inequality - for any concave extension of μ' we would have $\mathbb{E}\mu'(c_P - x) \leq \mu'(c_M - x)$. Moreover, $\mathbb{E}\mu'(c_P - x) \geq \int_{-\varepsilon}^{x-\mathbb{E}c_P} \mu'(\mathbb{E}c_P + \eta - x)h(\eta)d\eta$, because $\mu' \geq 0$. As a consequence, the denominator is positive for $\tau_P \geq \tau_M$. By strict concavity of f the numerator cannot be zero at the same threshold for two value of τ_P. As a consequence the BR is single valued for these tax values.

(ii) We can apply the implicit function theorem to first order condition. Since $\hat{\theta}$ is twice continuously differentiable, $\hat{\theta}_2$ is continuously differentiable, delivering the continuous differentiability of the BR. Recall that the implicit function theorem requires that and at the (unique) threshold $\hat{\theta}(\cdot)$ we have $\hat{\theta}_{22}(\tau_M, \tau_P(\tau_M)) \neq 0$. Given the second order necessary conditions require $\hat{\theta}_{22}(\tau_M, \tau_P(\tau_M)) \leq 0$, the condition in the statement can be satisfied only when $\hat{\theta}_{22} < 0$. The local second-order condition for an interior solution is satisfied with strict inequality as required since for $\tau_P(\tau_M) \geq \tau_M$:

$$\frac{\partial^2 \hat{\theta}}{\partial^2 \tau_P}_{|_{\frac{\partial \hat{\theta}}{\partial \tau_P} = 0}} = \frac{f''(\tau_P(\tau_M))}{(1-\tau_M)R(c_M, c_P) - (1 - \tau_P(\tau_M))} < 0.$$

□

We are now ready to show Lemma 4.

Proof. (i) Consider the populist problem (3). As we argued above in (11), the derivative of the threshold level with respect to τ_P equals

$$\frac{\partial \hat{\theta}}{\partial \tau_P} = \frac{f_\tau(\tau_P) - \hat{\theta}(\tau_M, \tau_P)}{(1-\tau_M)R(c_M, c_P) - (1-\tau_P)} \tag{12}$$

where $R(c_M, c_P) := (1 + \mu'(c_M - x)) / \left(1 + \int_{-\varepsilon}^{x - \mathbb{E}c_P} \mu'(\mathbb{E}c_P + \eta - x)h(\eta)d\eta\right)$. As well, we argued above that the denominator is positive for $\tau_P \geq \tau_M$.

Recall from Lemma 10 (i), that the policy is single valued for this range of taxes, and that we denote by $\tau_P^+(\tau_M)$ the solution to problem (3) at τ_M. The optimality condition defining the solution set (also called best response or BR) becomes

$$f_\tau(\tau_P^+(\tau_M)) \leq \hat{\theta}(\tau_M, \tau_P^+(\tau_M)), \quad \text{with equality for } \tau_P^+(\tau_M) > \tau_M. \tag{13}$$

If the solution is interior, then (13) is an equality and hence $\tau_P^+(\tau_M)$ represents the most preferred tax rate of agent with income $\hat{\theta}(\tau_M, \tau_P^+(\tau_M))$.

(ii) Consider the problem of the moderate (4) when considering a tax delivering $\tau_P^+(\tau_M) > \tau_M$. Recall that within this range the solution for τ_P is unique. As a consequence, under our assumptions, the moderate objective is differentiable (Milgrom and Segal, 2002). By the envelope theorem, the derivative $\frac{\partial \hat{\theta}(\tau_M, \tau_P(\tau_M))}{\partial \tau_M}$ is hence given by

$$\frac{\partial \hat{\theta}}{\partial \tau_M} = \frac{R(c_M, c_P)[\hat{\theta}(\tau_M, \tau_P^+(\tau_M)) - f_\tau(\tau_M)]}{(1-\tau_M)R(c_M, c_P) - (1-\tau_P^+(\tau_M))}. \tag{14}$$

Again, in the range $\tau_P > \tau_M$ and since $R(c_M, c_P) > 1$, the denominator of (14) is positive. Since at an interior populist optimum we have $\hat{\theta} = f_\tau(\tau_P^+(\tau_M))$ and since $f_\tau(\tau)$ is *decreasing* in τ, for $\tau_P^+(\tau_M) > \tau_M$ we have $\frac{\partial \hat{\theta}(\tau_M, \tau_P^+(\tau_M))}{\partial \tau_M} < 0$ as claimed. If instead $\tau_P^+(\tau_M) = \tau_M$ we have $f_\tau(\tau_P(\tau_M)) = f_\tau(\tau_M)$. From the optimality condition in the populist problem (13), we have the desired result. □

Proof of Proposition 5 First of all note that, from the optimality condition of the populist (12), whenever the populist choose its optimal tax under single crossing, P must set the optimal tax at the most preferred rate for the indifferent θ. If the populist best response is single valued, from the optimality conditions of the moderate (14), M must set $\tau_M = \tau_P$.

Proof. There are two possibilities. First, M sets τ_M so as to minimize $V_P^+(\tau_M)$ subject to $V_P^+(\tau_M) \geq V_P^-(\tau_M)$. In this case we have single crossing and hence, as we have shown in Lemma 4, $V_P^+(\tau_M) = G(\hat{\theta}(\tau_M, \tau_P(\tau_M^*))$ is decreasing in τ_M. Then the moderate will increase τ_M up to the point where $V_P^+(\tau_M) = V_P^-(\tau_M)$. This is so since V^+, V^- are continuous and for $\tau_M = \tau_P$ we have $V_P^+(\tau_M) = 0 \leq V_P^-(\tau_M)$. If in equilibrium this equality holds the populist is indifferent between $\tau_P(\tau_M^*) \leq \tau_M^*$ and $\tau_P(\tau_M^*) \geq \tau_M^*$ as claimed. Note that since a solution $\tau_P(\tau_M^*) = \tau_M^*$ is feasible for both the problem defining V^+ and V^-, this case includes that where tax rate convergence is the equilibrium solution.

The second case is where M sets τ_M so as to minimize $V_P^-(\tau_M)$, subject to $V_P^-(\tau_M) \geq V_P^+(\tau_M)$. One possibility is that in this case the constraint is binding $V_P^+(\tau_M^*) = V_P^-(\tau_M^*)$ as above. Then again we are in case (i) and the populist is indifferent between $\tau_P(\tau_M^*) \leq \tau_M$ and $\tau_P(\tau_M^*) \geq \tau_M$. The other possibility is that the constraint is not binding, and in equilibrium $V_P^-(\tau_M^*) > V_P^+(\tau_M^*)$. By definition of $V_P^-(\tau_M)$ then, in this case $\tau_P(\tau_M^*) < \tau_M^*$. This in turn implies that single crossing is violated. Recall indeed that we argued at the beginning of the proof that in equilibrium single crossing would imply policy convergence, i.e. $\tau_P(\tau_M^*) = \tau_M^*$. And hence it would violate $V_P^-(\tau_M^*) > V_P^+(\tau_M^*)$ as $\tau_P = \tau_M$ is also feasible for the problem of minimizing V^+ subject to $V^+ \geq V^-$.

The only remaining possibly is hence $\tau_P(\tau_M^*) < \tau_M^*$ which can be optimal because the function $w_P(\cdot, \tau_P(\tau_M^*))$ and $w_M(\cdot, \tau_M^*)$ cross more than once (i.e. for at least two different values of θ). This is case (ii). □

C Proofs for the Extensions Section

C.1 Populist as First Mover

We now show that even if we invert the order of moves the key qualitative features of the equilibrium set do not change, namely:

PROPOSITION 11. *Assume the populist moves first. In equilibrium either (i) the moderate is indifferent between $\tau_M^*(\tau_P^*) \geq \tau_P^*$ and $\tau_M^*(\tau_P^*) \leq \tau_P^*$ or (ii) $\tau_M^*(\tau_P^*) > \tau_P^*$ and we must have more than one crossing. In case (i) we have policy convergence whenever the best response of the moderate is single valued. In case (ii) we have double crossing whenever at least one of the sufficient conditions used in Proposition 13 below is satisfied.*

Proof. (Sketch) Recall the line of proof for Lemma 4 and Proposition 5. By comparing the first order conditions for P and M under $\tau_M < \tau_P$ (and single crossing), if at $\hat{\theta}$ M reaches its maximum, the (unique) threshold (and hence P payoff) is decreasing in τ_P. As a consequence, P will decrease τ_P till $V_M^+(\tau_P^*) = V_M^-(\tau_P^*)$. If this characterizes the solution, we have - by definition - indifference. Alternatively, we have $V_M^+(\tau_P^*) > V_M^-(\tau_P^*)$, which can only happen for $\tau_M(\tau_P^*) > \tau_P^*$ (otherwise we must have policy convergence due to single crossing violating the inequality $V_M^+(\tau_P^*) > V_M^-(\tau_P^*)$) and more than one crossing. $\square$

C.2 Heterogeneous Reference Points

Recall that each agent is now identified by two elements, (θ, x_0). Let $q(\theta|x_0)\ell(x_0)$ be joint density of those points in the population.

PROPOSITION 12. *Consider the case where the reference point solves: $x(\theta, x_0) = (1-\rho)\theta + \rho x_0$ for $1 > \rho > \tau_0$ and $x_0 \in [x_{min}, x_{max}]$. Suppose that Assumptions 1 and 2 hold, and also assume $\mathbf{E}\theta \leq x_{min} - \varepsilon + z$. Then*
(i) Let $\tau_0 \geq \tau_P \geq \tau_M$. Then for each x_0, we have at most one crossing at $\hat{\theta}(x_0, \tau)$ such that

$$\mathbb{E}c_P\left(\hat{\theta}(x_0,\tau),\tau_P\right) \leq c_M\left(\hat{\theta}(x_0,\tau),\tau_M\right) < x(\theta, x_0).$$

(ii) Even when the populist's policy correspondence $\tau_P^*(\cdot)$ is single valued we typically do not have policy convergence.
(iiI) If, in addition, for all $\tau_P \geq \tau_M$ we have $Cov\left(\frac{f_\tau(\tau_P)-\hat{\theta}(x_0,\tau)}{(1-\tau_M)R(x_0)-(1-\tau_P)}q\left(\hat{\theta}(x_0,\tau)|x_0\right), R(x_0)\right) > 0$ or $Cov\left(R(x_0)\frac{f_\tau(\tau_M)-\hat{\theta}(x_0,\tau)}{(1-\tau_M)R(x_0)-(1-\tau_P)}q\left(\hat{\theta}(x_0,\tau)|x_0\right), \frac{1}{R(x_0)}\right) > 0$ (or both), then it cannot be that in equilibrium the moderate charges a lower tax than the populist.

Proof. (i) This is the proof of Lemma 3 (for $\tau_0 \geq \tau_P \geq \tau_M$) for this general case with heterogenous reference points. We follow the line of proof of Lemma 3. First of all, note that our assumptions imply that for each given x_0, $c_M(\cdot, \tau_M) - x(\cdot, x_0)$ is strictly monotone in θ for all $\tau_M \leq \tau_0$. We will hence be able to span all ranges of values of θ by looking at the range for $c_M(\cdot, \tau_M) - x(\cdot, x_0)$. Fix $x_0 \in [x_{min}, x_{max}]$ and $\tau_0 \geq \tau_P \geq \tau_M$.
A. Consider all θ such that $c_M(\theta, \tau_M) - x(\theta, x_0) \geq 0$. Given our assumptions, it must be that $c_M(\theta, \tau_M) \geq \mathbf{E}\theta$. To see why this is the case suppose instead that $c_M(\theta, \tau_M) < \mathbf{E}\theta$. If $\theta > \mathbf{E}\theta$ then $\tau_M > 0$ implies $c_M(\theta, \tau_M) < \theta$ but then from the definition of $x(\theta, x_0)$ we have $x(\theta, x_0) \geq x(\theta, x_{min}) > (1-\rho)\theta + \rho\mathbf{E}\theta \geq \mathbf{E}\theta$, which contradicts the fact that $c_M(\theta, \tau_M) \geq x(\theta, x_0)$. If $\theta \leq \mathbf{E}\theta$,

then by the fact that these agent would desire redistribution if it were not distortionary, that $\tau_M \leq \tau_0 < \rho$, and the fact that taxes are in fact distortionary it must be that $c_M(\theta, \tau_M) \leq (1-\tau_0)\theta + \tau_0 \mathbf{E}\theta < x(\theta, x_0)$, again contradicting $c_M(\theta, \tau_M) - x(\theta, x_0) \geq 0$.

We now show that in this case the two functions w_M and w_P cannot cross because the populist has higher distortions than the moderate and the loss function μ is not acting for the latter. Formally: the monotonicity of consumption in θ and $c_M(\theta, \tau_M) \geq \mathbf{E}\theta$ implies $\theta \geq \mathbf{E}\theta$. As a consequence, $\tau_P \geq \tau_M$ implies that $c_M(\theta, \tau_P) \leq c_M(\theta, \tau_M)$ and $\mathbb{E}c_P(\theta, \tau_P) = c_M(\theta, \tau_P) - z$. Since under M consumption is above the reference point, while under P consumption might end up being below x in the bad state (and $\mu \leq 0$), the difference $w_P(\theta, \tau_P) - w_M(\theta, \tau_M)$ must be larger than z for all such θ.

B. Consider the complement set: θ such that $c_M(\theta, \tau_M) < x(\theta, x_0)$. We here have two cases to consider.

B1. First, assume in addition that $c_M(\theta, \tau_M) \geq \mathbb{E}c_P(\theta, \tau_P)$. That is, we consider all θ such that we have $x(\theta, x_0) > c_M(\theta, \tau_M) \geq \mathbb{E}c_P(\theta, \tau_P)$. The difference between the two functions w_M and w_P is strictly increasing in θ. They might hence cross at most once in this range. Formally, we have:[34]

$$\begin{aligned}
&\frac{\partial}{\partial\theta}[w_M(\theta, \tau_M) - w_P(\theta, \tau_P)] = \\
&(\tau_P - \tau_M) + [(1-\tau_M) - (1-\rho)]\mu'(c_M - x(\theta, x_0)) - [(1-\tau_P) - (1-\rho)]\mathbb{E}\mu'(c_P - x(\theta, x_0)) > 0,
\end{aligned}$$

where the last strict inequality is guaranteed even for $\tau_P = \tau_M$ because of Assumption 1 (iv). In particular, note that for any convex extension of μ we have $\mu'(c_M - x) \geq \int_{-\varepsilon}^{+\varepsilon} \mu'(c_P - x)h(\eta)d\eta$ and in any such extension, for $c_P - x > 0$ we must have $\mu'(c_P - x) \geq \mu'_-(0) > 0$.The inequality is hence satisfied a fortiori when for such values $\mu'(c_P - x) = 0$.

B2. Consider now the alternative case where $c_M(\theta, \tau_M) < x(\theta, x_0)$ and $c_M(\theta, \tau_M) < \mathbb{E}c_P(\theta, \tau_P)$. Note that, since $\tau_P \geq \tau_M$, it must be that $\theta \leq \mathbf{E}\theta$. And hence, under our Assumption of f, and the desire of redistribution of this agent in case of no distortion, we have $(1-\tau_P)\theta + f(\tau_P) < (1-\rho)\theta + \rho\mathbf{E}\theta$ and hence $\mathbb{E}c_P(\theta, \tau_P) = (1-\tau_P)\theta + f(\tau_P) - z < (1-\rho)\theta + \rho\mathbf{E}\theta - z \leq x(\theta, x_0) - \varepsilon$, where the last inequality is implied by our assumption. This in turn implies that $c_M(\theta, \tau_M) \leq \mathbb{E}c_P(\theta, \tau_P) < x(\theta, x_0) - \varepsilon$.

We now show that the two conditions $x(\theta, x_0) - \varepsilon \geq c_M(\theta, \tau_M)$ and $\mathbb{E}c_P(\theta, \tau_P) \geq c_M(\theta, \tau_M)$ together imply that $w_P(\theta, \tau_P) > w_M(\theta, \tau_M)$. This will show that the functions w_P and w_M cannot cross in the relevant range of θ in this case. In words, we have $w_P(\theta, \tau_P) > w_M(\theta, \tau_M)$ since for this range of θ the populist enjoys higher average consumption and the riskiness of the convex

[34]Obviously, for $c_P = x$, μ' should be replaced by its left derivative, however this is of measure zero in the integral and we ignore it to simplify notation.

punishment. More in detail, since $\mathbb{E}c_P(\theta,\tau_P) \geq c_M(\theta,\tau_M)$, it suffices to show that:

$$\mu\left(c_M(\theta,\tau_M) - x(\theta,x_0))\right) \leq \mathbb{E}\mu\left(c_P(\theta,\tau_P) - x(\theta,x_0)\right). \tag{15}$$

To see why (15) holds, note that $c_M(\theta,\tau_M) < x(\theta,x_0) - \varepsilon$, so: (i) $\mu(c_M(\theta,\tau_M) - x(\theta,x_0)) < 0$ and (ii) if we define $\tilde{c}_M(\theta,\tau_M) := c_M(\theta,\tau_M) + \eta$ where η takes the values between $-\varepsilon$ and $+\varepsilon$ with density h, from Jensen's inequality, we have $\mathbb{E}\mu\left(\tilde{c}_M(\theta,\tau_M) - x(\theta,x_0)\right) \geq \mu\left(c_M(\theta,\tau_M) - x(\theta,x_0)\right)$. Now, since $\mathbb{E}c_P(\theta,\tau_P) \geq \mathbb{E}\tilde{c}_M(\theta,\tau_M) = c_M(\theta,\tau_M)$ we obtain the result from the monotonicity of μ.

Now, since the only case where we can have a crossing point is **B1**, we have also shown that at the crossing point $\hat{\theta}$, if any, we have $\mathbb{E}c_P(\hat{\theta},\tau_P) < c_M(\hat{\theta},\tau_M) < x(\theta,x_0)$.

(ii) We now take for granted the single crossing property shown above. In particular, note that politicians will never find it optimal to impose a tax larger than τ_0. Using the envelope condition, assuming $\tau_P^*(\cdot)$ is a single valued correspondence, the first order conditions in the case of single crossing for all agents are:

$$\int_{x_{min}}^{x_{max}} \frac{\partial\hat{\theta}(x_0,\tau)}{\partial\tau_i} q\left(\hat{\theta}(x_0,\tau)|x_0\right) \ell(x_0) dx_0 = 0, \qquad i = P, M. \tag{16}$$

Recall that, for each x_0:

$$\frac{\partial\hat{\theta}(x_0,\tau)}{\partial\tau_P} = \frac{f_\tau(\tau_P) - \hat{\theta}(x_0,\tau)}{(1-\tau_M)R(c_M(x_0),c_P(x_0);x_0) - (1-\tau_P)}, \tag{17}$$

where, for each x_0, $c_i(x_0)$, $i = M, P$ indicate the consumption of agents of type x_0 and income equal to that of the crossing point $\hat{\theta}(x_0,\tau)$, and:

$$R(c_M(x_0),c_P(x_0)) := \frac{\left(1+\mu'(c_M(x_0)-x_0)\right)}{\left(1+\int_{-\varepsilon}^{x_0-\mathbb{E}c_P(x_0)} \mu'(\mathbb{E}c_P(x_0)+\eta-x_0)h(\eta)d\eta\right)}.$$

Recalling previous expressions, for the moderate, we have:

$$\frac{\partial\hat{\theta}(x_0,\tau)}{\partial\tau_M} \leq -\frac{\partial\hat{\theta}(x_0,\tau)}{\partial\tau_P} R(c_M(x_0),c_P(x_0);x_0), \text{ with strict inequality if and only if } \tau_P > \tau_M. \tag{18}$$

In addition, as we saw, for $\hat{\theta}(x_0,\tau)$ to be a crossing point, we must have $\mathbb{E}c_p(x_0) < c_M(x_0) < x_0$. As a consequence, as we argued above, we must have $R(c_M(x_0),c_P(x_0)) > 1$ and the denominator of (17) is positive for $\tau_P \geq \tau_M$.

Let us first of all check that we typically lose policy convergence. Since at the optimum for the populist (16) equals zero, we can write the corresponding first order condition for the moderate, in the case where $\tau_M = \tau_P^*(\tau_M)$, as follows:

$$\begin{aligned}
& \int_{x_{min}}^{x_{max}} \frac{\partial\hat{\theta}(x_0,\tau)}{\partial\tau_M} q\left(\hat{\theta}(x_0,\tau_M,\tau_P^*(\tau_M))|x_0\right)\ell(x_0)dx_0 \\
\leq \;& -Cov_\ell\left(\frac{\partial\hat{\theta}(x_0,\tau)}{\partial\tau_P} q\left(\hat{\theta}(x_0,\tau_M,\tau_P^*(\tau_M))|x_0\right), R(c_M(x_0),c_P(x_0);x_0)\right) \\
= \;& -Cov_\ell\left(\frac{f_\tau(\tau_P)-\hat{\theta}(x_0,\tau)}{(1-\tau_M)R(x_0)-(1-\tau_P)} q\left(\hat{\theta}(x_0,\tau)|x_0\right), R(x_0)\right), \qquad (19)
\end{aligned}$$

with strict inequality if and only if $\tau_P > \tau_M$. The first point is easily shown since the covariance in (19) (taken according to the density $\ell(\cdot)$) is generically different from zero.
(iii) We now show that the moderate can never have the necessary optimality conditions satisfied for $\tau_P > \tau_M$. If $\tau_M < \tau_P$ the inequality in (18) and hence the second inequality of (19) are strict. This together with our assumption that the covariance in (19) is positive implies a moderate FOCs strictly negative, violating a necessary condition for optimality. Similarly, since $R(x_0) > 0$, we can write the populist's first order condition at the optimal point for the moderate as follows:

$$\begin{aligned}
& \int_{x_{min}}^{x_{max}} \frac{\partial\hat{\theta}(x_0,\tau)}{\partial\tau_P} q\left(\hat{\theta}(x_0,\tau_M,\tau_P^*(\tau_M))|x_0\right)\ell(x_0)dx_0 \\
\leq \;& -Cov_\ell\left(\frac{\partial\hat{\theta}(x_0,\tau)}{\partial\tau_M} q\left(\hat{\theta}(x_0,\tau_M,\tau_P^*(\tau_M))|x_0\right), \frac{1}{R(c_M(x_0),c_P(x_0);x_0)}\right) \\
= \;& -Cov_\ell\left(R(x_0)\frac{f_\tau(\tau_M)-\hat{\theta}(x_0,\tau)}{(1-\tau_M)R(x_0)-(1-\tau_P)} q\left(\hat{\theta}(x_0,\tau)|x_0\right), \frac{1}{R(x_0)}\right), \qquad (20)
\end{aligned}$$

and again if this covariance is positive we can exclude $\tau_P > \tau_M$ as a possible equilibrium outcome (without indifference). □

To gain intuition on the last covariance term in (19), suppose the conditional density is constant.[35] Then the expression becomes

$$qCov_\ell\left(\frac{f_\tau(\tau_P)-\hat{\theta}(x_0,\tau)}{(1-\tau_M)R(x_0)-(1-\tau_P)}, R(x_0)\right).$$

Let's start with a 'mechanical' interpretation. In the populist first order condition, we must have some entries for $\frac{f_\tau(\tau_P)-\hat{\theta}(x_0,\tau)}{(1-\tau_M)R(x_0)-(1-\tau_P)}$ that are positive and others that are neg-

[35]This holds if the distribution of θ in the relevant range is uniform with identical density for all x_0.

ative.[36] To be zero on average such entries must balance out. Any change in tax of the moderate sees both positive and negative entries amplified as $R(x_0) > 1$. If the covariance is positive, roughly speaking, it means that in the first order conditions of the moderate the positive entries have large $R(x_0)$ while the negative entries have small $R(x_0)$, making impossible to satisfy the first order conditions for a lower τ_M.

The economic intuition can be summarised as follows. If we have an equilibrium where $\tau_P > \tau_M$ it means - in some sense - that the populist is more concerned than the moderate of loosing the disappointed poor. Now suppose that all that matters is the amplification between positive vs negative entries. A positive covariance roughly means that when $f_\tau(\tau_P) - \hat{\theta}(x_0, \tau) > 0$, that is for agents that want more redistribution, $R(x_0)$ is large, that is, the elasticity of votes with respect to τ_M is large; while for voters that want less redistribution, the elasticity of votes with respect to τ_M is small. But this suggests that in fact the moderate should be more concerned than the populist of loosing the votes of the poor, which is exactly the opposite of what is required to get the result.

D Sufficient conditions for having at most double crossing

Here we provide sufficient conditions guaranteeing that the value functions $w_M(\theta, \tau_M)$ and $w_P(\theta, \tau_P)$ cross at most twice for $\tau_P < \tau_M$, and we rule out the possibility of more than two crossing points. For this result, we need the following assumptions.

ASSUMPTION 3. *The function $f(\cdot)$ satisfies the following condition*

$$\frac{f(\tau_0) + z}{\tau_0} - z - \varepsilon > x,$$

with $f(\tau) = \tau \mathbf{E}\theta - i(\tau)$, and $\frac{(1-\tau)z - i(\tau)}{\tau}$ decreasing in τ.

It can be checked directly, that the expression $\frac{(1-\tau)z - i(\tau)}{\tau}$ is decreasing in τ, for example, in the standard quadratic case: $i(\tau) = \frac{1}{2}\tau^2$. Recall that τ_0 denotes the revenue maximizing tax rate. This assumption - together with Assumption 2 - guarantees that, for any $\tau_P < \tau_M$, the right-most crossing point occurs in a region where average consumption under P is weakly lower than consumption under M (in the most typical case, this

[36] If all entries have the same sign they all must equal zero at the optimum. In this (unrealistic) case we are back to the results of the homogeneous x_0 situation. In particular, we might have policy convergence and it is not possibile to have an equilibrium where the populist strictly prefers $\tau_P > \tau_M$.

is because voters have linear utility under both the populist and the moderate candidate and hence the two consumption levels coincide). This in turn implies that at any θ to the left of this point - and in particular at a $\underline{\theta}$ for which we could have a second crossing point between w_M and w_P - average consumption under P is lower than that under M. The latter statement is true since $\tau_p < \tau_M$ implies that net income schedule under the populist $\mathbb{E}c_P$ is steeper than c_M as a function of θ.

ASSUMPTION 4. *The function μ is such that for all admissible $c < x$ we have*

$$\int_{-\varepsilon}^{x-c} \mu''(c+\eta-x)h(\eta)d\eta \geq \mu''(c-x) + h(x-c)\mu'_{-}(0),$$

where, clearly, for all $x - c > \varepsilon$ we have $h(x-c) = 0$. Moreover, the density h is weakly decreasing for positive entries, i.e., for $\eta > 0$.

In the above, $\mu'_{-}(0)$ indicates the left side derivative of $\mu(c-x)$ evaluated at 0 (assumed to always exist). This assumption can be seen as about the concavity of $\mu'(\cdot)$ as it is more likely to be satisfied if μ'' is large for low values. It would be satisfied if μ' is concave enough. A low $\mu'_{-}(0)$ and a low density for large η would also help.

Assumption 4 is unfortunately never satisfied for h uniform. Here we provide a second assumption, which might hold for h uniform but involves a joint condition on the maximal tax rate.

ASSUMPTION 5. *The function μ is such that for all admissible $c < x$ we have*

$$\tau_0 + \int_{-\varepsilon}^{x-c} \mu'(c+\eta-x)h(\eta)d\eta \leq (1-\tau_0)\mu'(c-x),$$

where, clearly, for all $x - c > \varepsilon$ we have $h(x-c) = 0$.

If c is low enough, the condition becomes:

$$\tau_0 + \int_{-\varepsilon}^{\varepsilon} \mu'(c+\eta-x)h(\eta)d\eta \leq (1-\tau_0)\mu'(c-x).$$

Note that the (sufficient) condition is easier to satisfy the more concave is μ'. As well, as c approaches x the condition is less difficult to satisfy, especially if the concavity of μ'

increases with c, namely if $\mu^{iv} \leq 0$. By contrast, for example, if μ is quadratic the condition becomes $\tau_0 \leq -\tau_0 \mu'(c-x)$, which can never be satisfied.

Recalling the mean value theorem, the condition, becomes

$$\tau_0 + \mu'(c+\xi-x) \leq (1-\tau_0)\mu'(c-x),$$

for ξ a negative constant as long as μ' is concave.

PROPOSITION 13 (Double Crossing). *Under Assumptions 1, 2, 3, and either 4 or 5 (or both) we have the following. If $\tau_P < \tau_M$ then the functions $w_M(\cdot,\tau_M)$ and $w_P(\cdot,\tau_P)$ cross at most twice as a function of θ.*

The idea of the proof in case we use Assumption 4 is that - since at the first crossing point to the left of x the slope of w_P must be lower than that of w_M - it suffices to guarantee that to the left of such crossing point the difference of slopes can only increase: the assumption guarantees that w_P is flatter than w_M for all such points.

The idea of the proof in case we use Assumption 5 is that - since at the first crossing point to the left of x, w_P and w_M take the same value (by definition) - it suffices to guarantee that to the left of such crossing point the difference in the levels of such values can only increase: the assumption guarantees that w_P is higher than w_M for all such points.

Proof of Proposition 13

Proof. We start with a preliminary lemma.

LEMMA 14. *Let $\bar{\theta}$ be the largest crossing point for $\tau_P < \tau_M$. Under Assumption 2 and 3, $\mathbb{E}c_P(\bar{\theta},\tau_P) \leq c_M(\bar{\theta},\tau_M)$.*

Proof. We can have two cases.
Case 1: At $\bar{\theta}$ we have $\mathbb{E}c_P(\bar{\theta},\tau_P) < c_M(\bar{\theta},\tau_M)$. There is nothing to show in this case obviously.
Case 2: At $\bar{\theta}$, we have $\mathbb{E}c_P(\bar{\theta},\tau_P) \geq c_M(\bar{\theta},\tau_M)$. In this case we will show that the only possibility is $\mathbb{E}c_P(\bar{\theta},\tau_P) = c_M(\bar{\theta},\tau_M)$. By direct computation, whenever $\mathbb{E}c_P(\bar{\theta},\tau_P) \geq c_M(\bar{\theta},\tau_M)$, the threshold $\bar{\theta}(\tau_P,\tau_M) > 0$ solves:

$$\bar{\theta}(\tau_P,\tau_M) \geq \mathbf{E}\theta + \frac{i(\tau_P) - i(\tau_M) + z}{\tau_M - \tau_P}. \tag{21}$$

Consider now the following sequence of inequalities.

$$\begin{aligned}
\mathbb{E}c_P(\bar\theta(\tau_P,\tau_M),\tau_P) \geq c_M(\bar\theta(\tau_P,\tau_M),\tau_M) &= (1-\tau_M)\bar\theta(\tau_P,\tau_M) + f(\tau_M) \\
&\geq (1-\tau_M)\left[\mathbf{E}\theta + \frac{i(\tau_P)-i(\tau_M)+z}{\tau_M-\tau_P}\right] + \tau_M\mathbf{E}\theta - i(\tau_M) \\
&\geq \mathbf{E}\theta + \frac{1-\tau_M}{\tau_M}z - \frac{i(\tau_M)}{\tau_M} \geq \mathbf{E}\theta + \frac{1-\tau_0}{\tau_0}z - \frac{i(\tau_0)}{\tau_0}.
\end{aligned} \tag{22}$$

The first inequality in the first row is by assumption, the second equality is by definition. The inequality in the second row uses (21), while the inequality in the third row is due to the fact that the expression in the first row increases with τ_P for $\tau_P < \tau_M$. We hence set $\tau_P = 0$ and use Assumption 2. The first inequality in the last row is due to the fact that the expression in the previous row is decreasing in τ_M from Assumption 3. Now, Assumption 3 implies that the very last expression in the chain of inequalities solves:

$$\mathbf{E}\theta + \frac{1-\tau_0}{\tau_0}z - \frac{i(\tau_0)}{\tau_0} = \frac{f(\tau_0)+z}{\tau_0} - z > x + \varepsilon.$$

Recalling the very first term in the chain of inequalities (22), we have $\mathbb{E}c_P(\bar\theta(\tau_P,\tau_M),\tau_P) > x + \varepsilon$. This implies that μ cannot be active for the populist at $\bar\theta$. Since μ is active for the moderate would imply $c_M > \mathbb{E}c_P$ the only possibility is that μ is not active for the moderate either and $\mathbb{E}c_P(\bar\theta(\tau_P,\tau_M),\tau_P) = c_M(\bar\theta(\tau_P,\tau_M),\tau_M)$ as claimed. $\square$

We can now start the core of the proof of the proposition. From Lemma 14, we have $c_M(\bar\theta,\tau_M) \geq \mathbb{E}c_P(\bar\theta,\tau_P) := (1-\tau_P)\bar\theta + f(\tau_P) - z = \int_{-\varepsilon}^{\varepsilon}[(1-\tau_P)\bar\theta + f(\tau_P) + \eta]h(\eta)d\eta - z$. Now, consider the largest crossing point below $\bar\theta$, if it exists (if not then we have single crossing and we are done), and denote it by $\underline{\theta}$. Since $\tau_P < \tau_M$, $c_M(\bar\theta,\tau_M) \geq \mathbb{E}c_P(\bar\theta,\tau_P)$ implies that at any crossing point $\underline{\theta} < \bar\theta$ we must have $c_M(\underline{\theta},\tau_M) > \mathbb{E}c_P(\underline{\theta},\tau_P)$. Clearly, at $\underline{\theta}$ the slope of w_P must be lower than that of w_M, that is, w_P must cross w_M from below. In addition, $c_M(\theta,\tau_M) < x$ otherwise we will not have any second crossing at $\underline{\theta}$. We now show that Assumption 4 guarantees that the slope of w_P to the left of $\underline{\theta}$ will always be lower than that of w_M. This implies that we cannot have another crossing point in this region.

Consider a crossing point to the left of x, with $\tau_P < \tau_M$. We saw that at such $\underline{\theta}$ we have:

(a) (same level): $w_M(\underline{\theta},\tau_M) = w_P(\underline{\theta},\tau_P)$;

(b) (flatter w_P): $w'_M(\underline{\theta},\tau_M) \geq w'_P(\underline{\theta},\tau_P)$;

(c) (lower consumption): $c_M(\theta,\tau_M) \geq \mathbb{E}c_P(\theta,\tau_P)$ for all $\theta \leq \underline{\theta}$.

We want to show that for al $\theta \leq \underline{\theta}$ we have $w_P(\theta, \tau_P) - w_M(\theta, \tau_M) \geq 0$. Note that for $s = M, P$ we have

$$w_s(\theta, \tau_s) = w_s(\underline{\theta}, \tau_s) - \int_{\theta}^{\underline{\theta}} w'_s(t, \tau_s)dt.$$

and hence using, property (a) above, we have:

$$w_P(\theta, \tau_P) - w_M(\theta, \tau_M) = \int_{\theta}^{\underline{\theta}} w'_M(t, \tau_M)dt - \int_{\theta}^{\underline{\theta}} w'_P(t, \tau_P)dt = \int_{\theta}^{\underline{\theta}} [w'_M(t, \tau_M) - w'_P(t, \tau_P)]dt.$$

Part 1. Using Assumption 4. Obviously, if $w'_M(\theta, \tau_M) \geq w'_P(\theta, \tau_P)$ for all $\theta \leq \underline{\theta}$ we would be done. The change in slope of the functions w_P and w_M are given by their second derivatives:

$$w''_M(\theta, \tau_M) = (1 - \tau_M)^2 \mu''(c_M(\theta, \tau_M) - x)$$

and

$$w''_P(\theta, \tau_P) = (1 - \tau_P)^2 \int_{-\varepsilon}^{\{\varepsilon, \mathbb{E}c_P(\theta,\tau_P) - x\}^-} \mu''(\mathbb{E}c_P(\theta, \tau_P) + \eta - x)h(\eta)d\eta - h(x - \mathbb{E}c_P(\theta, \tau_P))(1 - \tau_P)^2 \mu'_-(0).$$

where, to emphasize the properties of μ we used $\{\varepsilon, \mathbb{E}c_P(\theta, \tau_P) - x\}^- := \min\{\varepsilon, \mathbb{E}c_P(\theta, \tau_P) - x\}$. Recall that at all such $\theta < \underline{\theta}$ we have $\mathbb{E}c_P(\theta, \tau_P) < c_M(\theta, \tau_M) < x$. Now, let $c = c_M(\theta, \tau_M)$. We have both:

$$\begin{aligned} (1 - \tau_M)^2 \mu''(c - x) &\leq (1 - \tau_P)^2 \mu''(c - x) \quad \text{and} \\ (1 - \tau_P)^2 \int_{-\varepsilon}^{\mathbb{E}c_P(\theta,\tau_P) - x} \mu''(\mathbb{E}c_P(\theta, \tau_P) + \eta - x)h(\eta)d\eta &\geq (1 - \tau_P)^2 \int_{-\varepsilon}^{x-c} \mu''(c + \eta - x)h(\eta)d\eta. \end{aligned}$$

In the first inequality we simply used $0 \leq \tau_P < \tau_M < 1$; to show the second inequality, we used $\mu''' < 0$ and the fact that $\mathbb{E}c_P(\theta, \tau_P) < c < x$ and $\mu'' \geq 0$. Finally, note that the second statement of Assumption 4 implies $h(x - c) \geq h(x - \mathbb{E}c_P(\theta, \tau_P))$. Now, bringing these results together, and noticing that $(1 - \tau_P)^2 > 0$, the first statement of Assumption 4 implies $w''_P(\theta, \tau_P) \geq w''_M(\theta, \tau_M)$. This is the desired result. It states that for each θ *to the left* of $\underline{\theta}$ the slope of w_P is *lower* than that of w_M.

Part 2. Using Assumption 5. Recall again we want to show that for all $\theta \leq \underline{\theta}$:

$$w_P(\theta, \tau_P) - w_M(\theta, \tau_M) = \int_{\theta}^{\underline{\theta}} \left[w'_M(t, \tau_M) - w'_P(t, \tau_P) \right] dt \geq 0.$$

Again, if we show that $w'_M(t, \tau_M) - w'_P(t, \tau_P) \geq 0$ for all $t < \underline{\theta}$ and conceivable τ_P, τ_M, we are done.

If we now write the first derivative of the functions we have

$$w'_M(t, \tau_M) = (1 - \tau_M)[1 + \mu'(c_M(t, \tau_M) - x)];$$

and

$$w'_P(t, \tau_P) = (1 - \tau_P) \left[1 + \int_{-\varepsilon}^{\{\varepsilon, \mathbb{E}c_P(t,\tau_P) - x\}^-} \mu'(\mathbb{E}c_P(t, \tau_P) + \eta - x) h(\eta) d\eta \right],$$

where, to emphasize the properties of μ, we wrote: $\{\varepsilon, \mathbb{E}c_P(t, \tau_P) - x\}^- := \min\{\varepsilon, \mathbb{E}c_P(t, \tau_P) - x\}$ and then we used $\mu(0) = 0$. We hence have:

$$\begin{aligned}
& w'_M(t, \tau_M) - w'_P(t, \tau_P) \\
= \; & (\tau_P - \tau_M) + (1 - \tau_M)\mu'(c_M(t, \tau_M) - x) - (1 - \tau_P) \int_{-\varepsilon}^{\{\varepsilon, \mathbb{E}c_P(t,\tau_P) - x\}^-} \mu'(\mathbb{E}c_P(t, \tau_P) + \eta - x) h(\eta) d\eta \\
\geq \; & -\tau_0 + (1 - \tau_0)\mu'(c_M(t, \tau_M) - x) - \int_{-\varepsilon}^{\{\varepsilon, \mathbb{E}c_P(t,\tau_P) - x\}^-} \mu'(\mathbb{E}c_P(t, \tau_P) + \eta - x) h(\eta) d\eta \\
\geq \; & -\tau_0 + (1 - \tau_0)\mu'(\mathbb{E}c_P(t, \tau_P) - x) - \int_{-\varepsilon}^{\{\varepsilon, \mathbb{E}c_P(t,\tau_P) - x\}^-} \mu'(\mathbb{E}c_P(t, \tau_P) + \eta - x) h(\eta) d\eta \geq 0.
\end{aligned}$$

The first equality is by definition. The inequality in the second row is due to the fact we have replaced the multiplicative terms of taxes with their worst cases possibilities. The inequality in the last row is due to the fact that we replaced c_M for $\mathbb{E}c_P$ and we know that $\mathbb{E}c_P < c_M < x$ and μ is convex. The very last inequality is guaranteed by Assumption 5 using $c = \mathbb{E}c_P$. $\square$

E Parameter Values and Figures from the Simulated Example

In Figures 7 and 8 we report a graphical representation of our loss function μ and distribution of θ with support $\mathbb{R}_+$ and $\mathbf{E}(\theta) = 25$. The inefficiency parameter for the populist equals $z = 1$, while the reference point equals $x = 30$. Finally, we set:

$$f(\tau) = \begin{cases} \tau\mathbf{E}(\theta) - \alpha\frac{\tau^2}{2} & \text{if} \quad \tau \leq \tau_0, \\ -\infty & \text{if} \quad \tau > \tau_0, \end{cases} \quad \text{with } \tau_0 = 0.5, \text{ and } \alpha = 20.$$

$$\eta \sim \mathcal{U}[-\varepsilon, \varepsilon] \quad \text{uniform,} \quad \text{with } \varepsilon = 17.$$

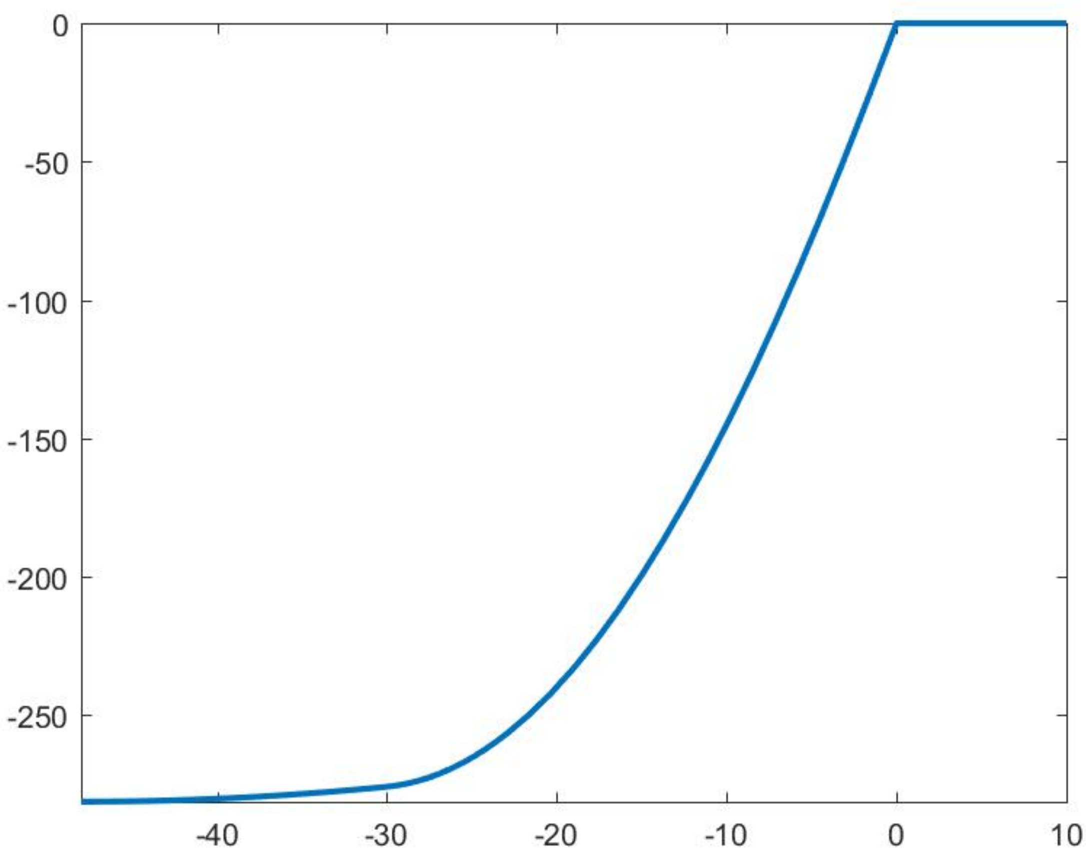

Figure 7: **Voters' preferences.** The figure represents the loss function μ as a function of $c - x$.

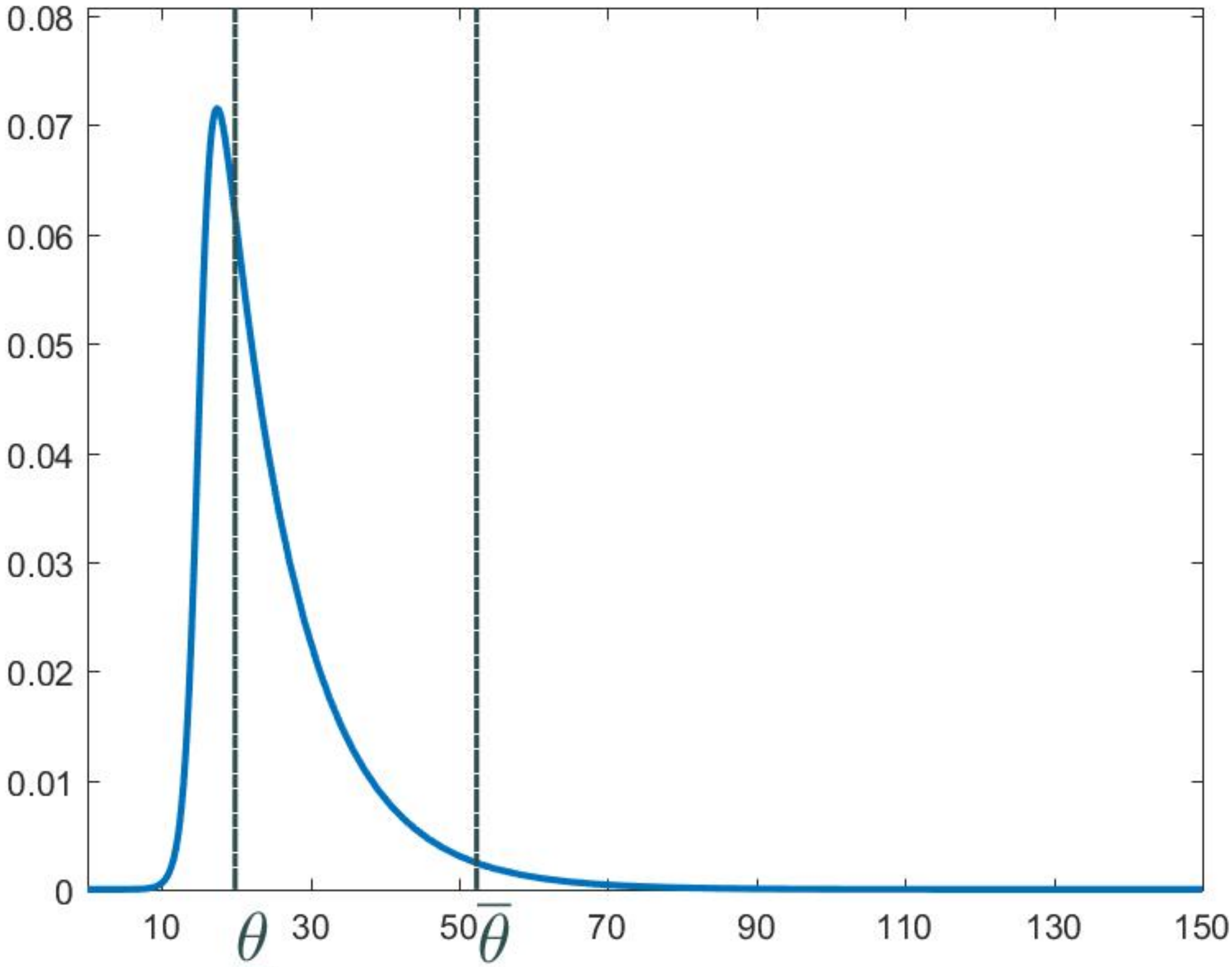

Figure 8: **The distribution of income and the double crossing.** The figure reports the (density) distribution of θ. The vertical dashed lines represent the equilibrium crossing points. Voters prefer M between the two points and P for θ outside the two vertical lines.

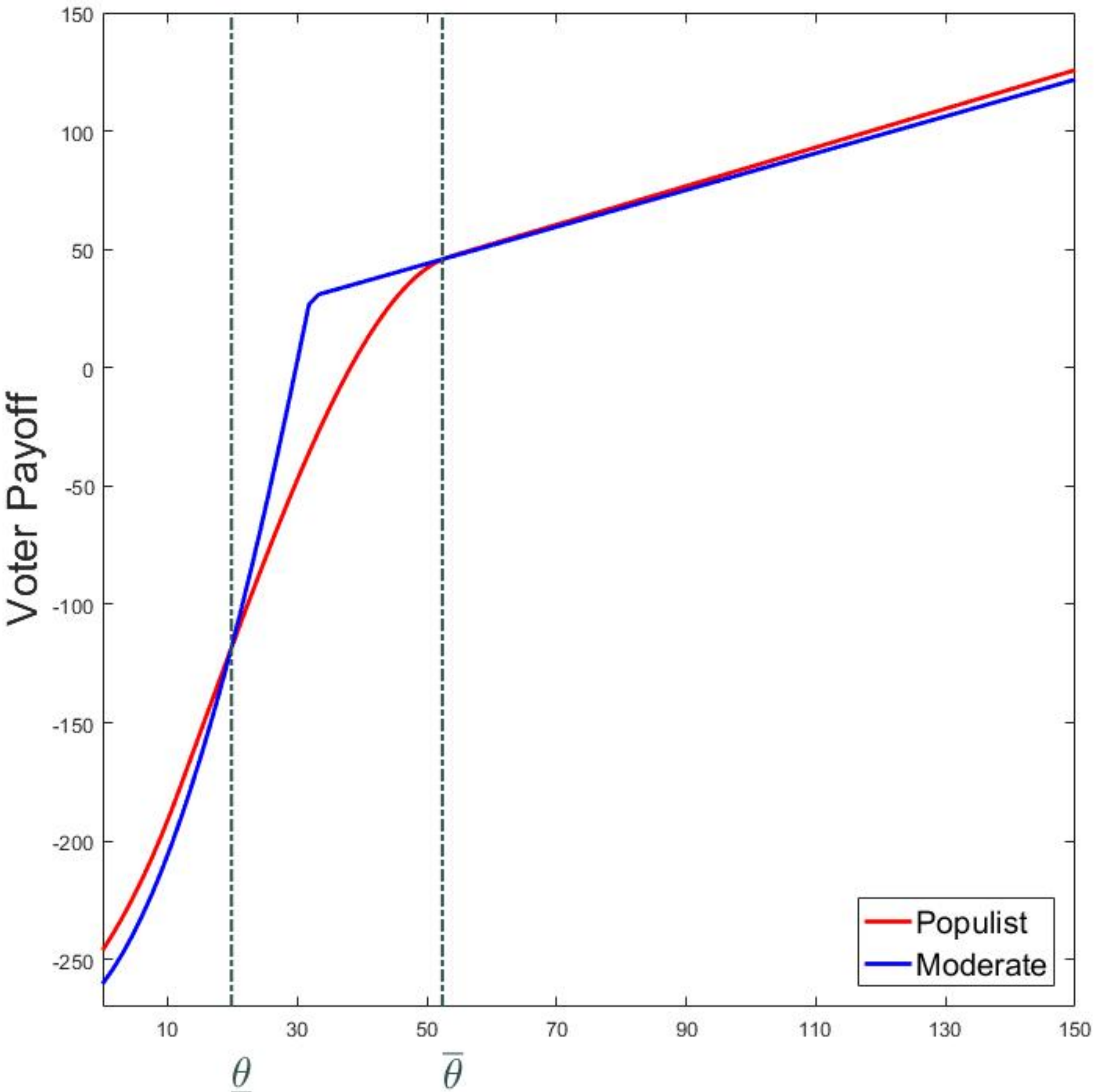

Figure 9: **The crossing in the space of utilities.** The figure reports both w_P (in red) and w_M (in blue) as a function of of θ for the equilibrium level of taxes. The vertical dashed lines represent the equilibrium crossing points. As it can be seen from the figure, voters prefer M between the two points and P for θ outside the two vertical lines.

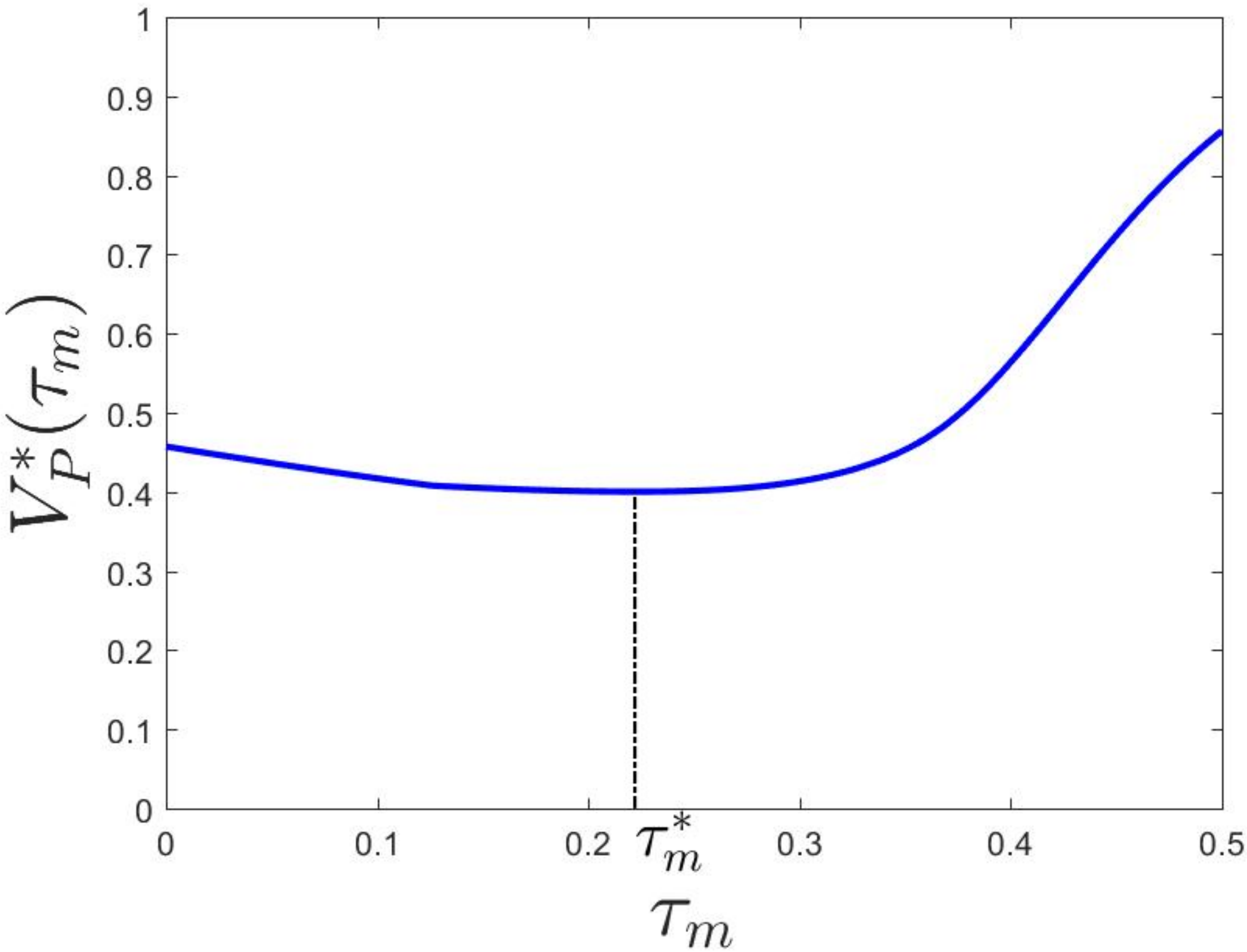

Figure 10: **The share of votes for P.** The figure reports the share of voters that prefer P, that is the populist's payoff, once P reacts optimally to M, as a function of τ_M. Since population is constant, M's payoff equals $1 - V_P^*(\tau_M)$. The vertical dashed line represent the equilibrium value for τ_M (i.e., $\tau_M^* = 0.22$). As it can be seen from the figure, for $\tau_M^* = 0.22$, V_P^* is minimal.

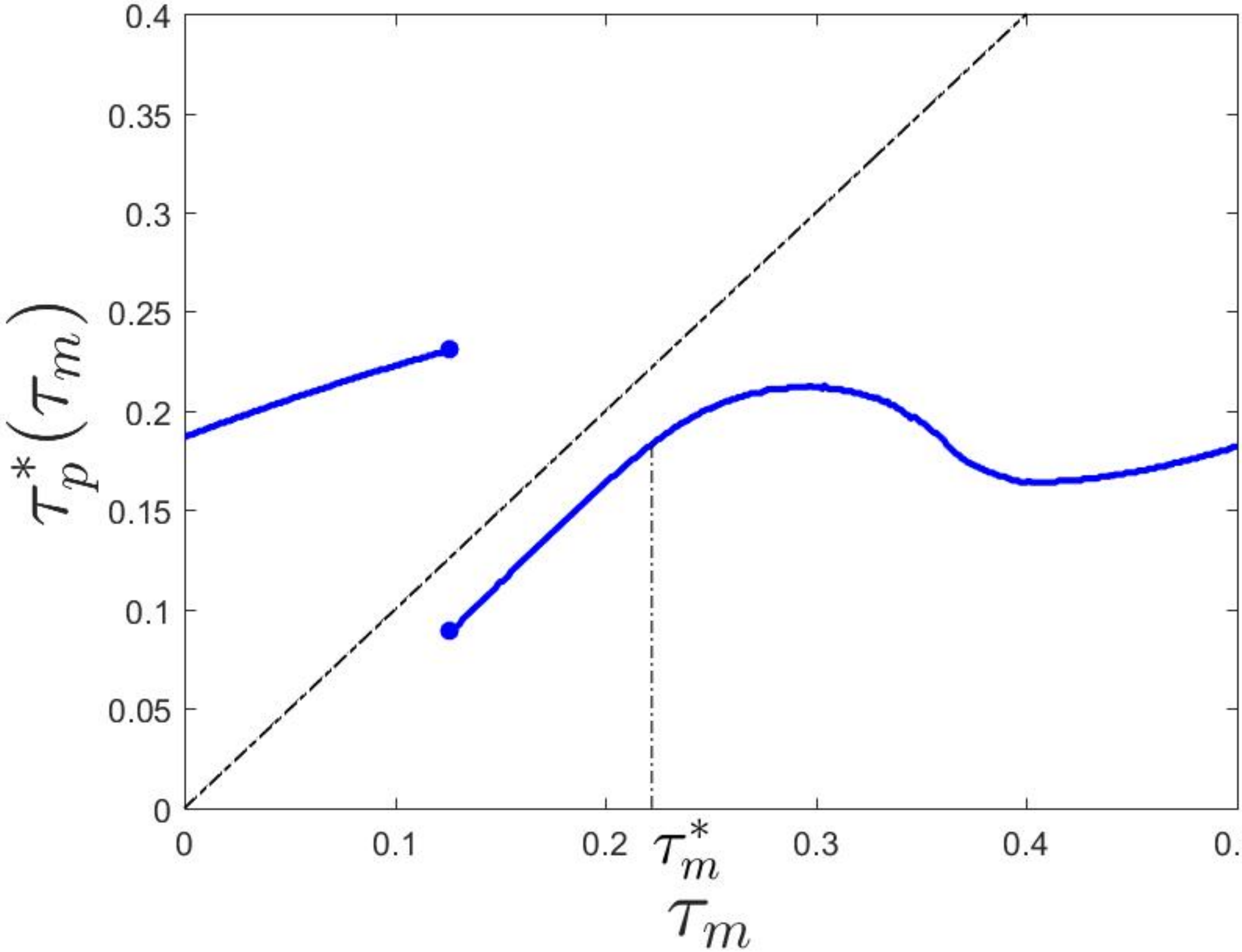

Figure 11: **The populist's Best Response Function.** The figure reports P's optimal choice as a function of τ_M (the best response). The vertical dashed line represent the equilibrium value for τ_M (i.e., $\tau_M^* = 0.22$). As it can be seen from the figure, for $\tau_M > .12$ the populist finds it optimal to set $\tau_P < \tau_M$.

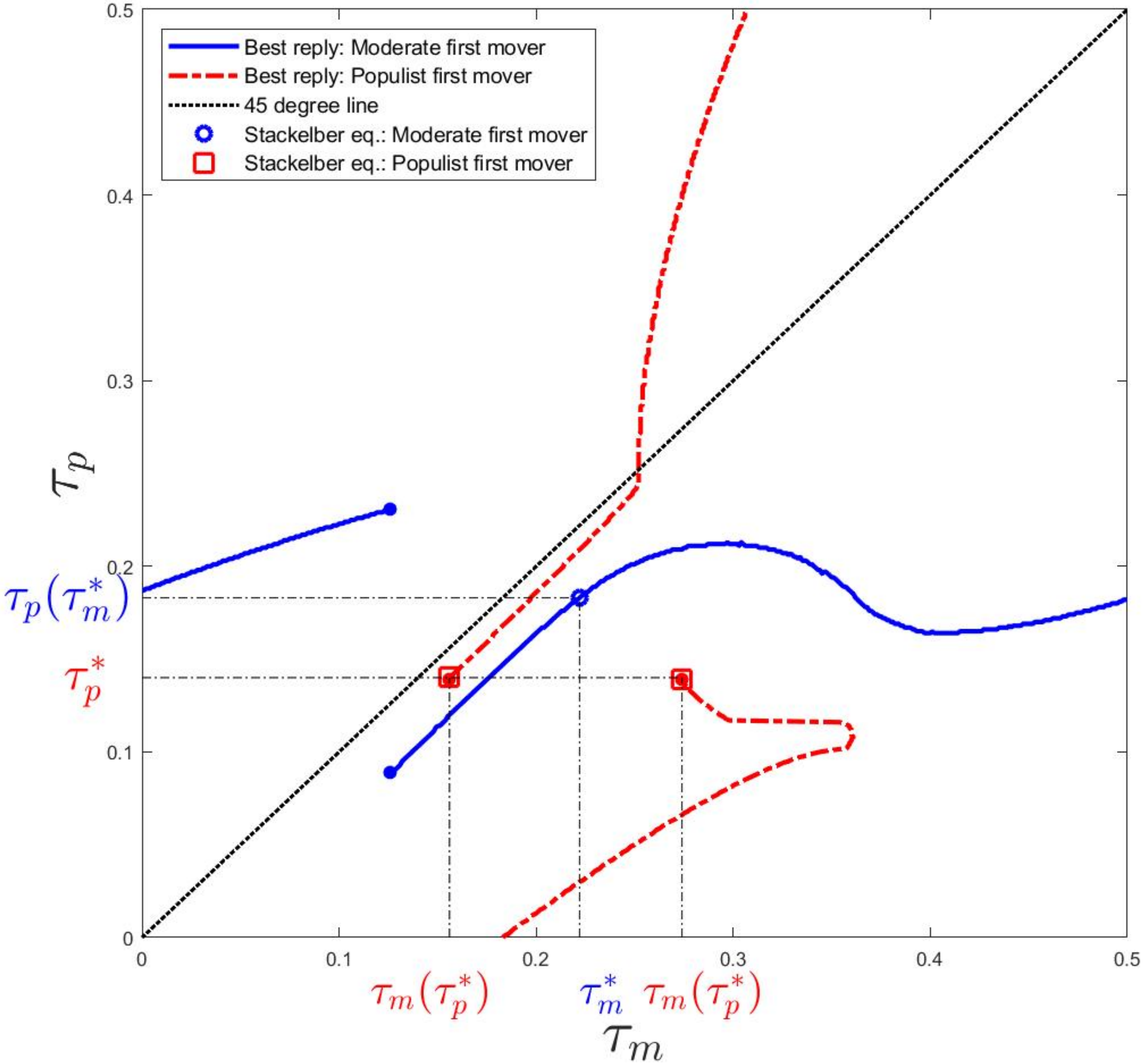

Figure 12: **The moderate and populist's Best Response Functions.** The figure reports both P and M best responses (in solid blue and dashed red, respectively) and indicates the Stackelberg equilibria. The equilibrium for the benchmark timing is indicated on the P best response with a blue circle. The equilibria emerging in the game when P moves first are indicated on the M best response with the two red squares. Note that since the game is of zero sum, indiffernce for M also implies constant payoff for P. In accordance to our theoretical results, in both cases, the populist tax is lower than that of the moderate (they both lie below the 45 degree line). In addition, note that for the parameters used, there is no Nash in pure strategies due to the lack of convexity of the sets in the best response correspondences.