No Country for Young People? The Rise of Anti-immigration Populism in Ageing Societies

Dotti, Valerio

Washington University in St. Louis

14 April 2020

No Country for Young People? The Rise of Anti-immigration Populism in Ageing Societies[*]

Valerio Dotti[†]

14th April 2020[‡]

Abstract

We investigate the effects of (i) population ageing and (ii) rising income inequality on immigration policies using an overlapping-generations model of elections with endogenous political parties. In each period, young people work and pay taxes while old people receive social security payments. Immigrants are generally young, meaning they contribute significantly to financing the cost of public services and social security. Among natives, the elderly and the poor benefit the most from public spending. However, because these two types of voters do not fully internalize the positive fiscal effects of immigration, they have a common interest in coalescing around a *populist party* (or multiple) seeking to curb immigration and increase the tax burden on high-income individuals. Population ageing and rising income inequality increase the size and, in turn, the political power of such parties, resulting in more restrictive immigration policies, a larger public sector, higher tax rates, and lower societal well-being. Calibrating the model to UK data suggests that the magnitude of these effects is large. The implications of this model are shown to be consistent with patterns observed in UK attitudinal data.

JEL classification: D72, C71, J610, H550.

Keywords: Immigration, Ageing, Policy, Voting.

[*]Acknowledgements: I am thankful to Antonio Cabrales, Ian Preston and John Roemer for their invaluable advice. I thank V. Bhaskar, Michele Boldrin, Saumya Deojain, Christian Dustmann, Johannes Horner, Suehyun Kwon, Guy Laroque, Gilat Levy, Konrad Mierendorff, Massimo Morelli, Aureo de Paula, Nikita Roketskiy and Larry Samuelson. I thank seminar and conference participants at UCL and Yale University. I gratefully acknowledge financial support from ESRC and UCL.

[†]Department of Economics, Washington University in St. Louis, One Brookings Dr., MO 63130, United States of America. Email: vdotti@wustl.edu

[‡]First version 26th October 2015. Latest version available at http://valeriodotti.github.io/.

What are the effects of population ageing and rising income inequality on immigration policy? Why are anti-immigration political parties increasingly successful in rapidly ageing countries, which arguably need more legal immigration to mitigate the impact of population ageing on public finances? Should we expect increasing restrictions on immigrant workers inflow in these countries? This paper aims to answer these questions using a theoretical model and providing suggestive empirical evidence.

This study is motivated by four key findings from the empirical literature on migration:

1. Aversion to immigration (Dustmann & Preston, 2007; Facchini & Mayda, 2007; Card et al., 2011) and support for anti-immigration political parties (Becker & Fetzer, 2017; Van der Brug et al., 2000) tend to be strong among the elderly (*Fig. 1A*) and the low-income native citizens relative to people in other sociodemographic groups (*Fig. 1B*).[1]

2. Economic hostility towards immigration is primarily motivated by concerns about its effects on public finances, specifically those related to public spending policies (Dustmann & Preston, 2006, 2007; Boeri, 2010). This suggests a *perception* among natives of competition with immigrants over welfare benefits and the use of crowded-out public services.

3. Immigrants are, on average, net fiscal contributors. The empirical evidence indicates that this is true both in the UK (Dustmann et al., 2010; Dustmann & Frattini, 2014), and the US (Lee & Miller, 2000; Orrenius, 2017), implying that, at least in the long run, immigrants do not directly pull fiscal resources from the natives.[2]

4. Aversion to immigration is a key determinant of the success of the so called *right-wing populist parties* in numerous countries. Examples include the *United Kingdom Independence Party* (UKIP) in the UK (Becker & Fetzer, 2017), the *National Front* in France (Edo et al., 2019), and the *Northern League* in Italy (Barone et al., 2016).[3]

These empirical findings lead to a two-fold puzzle. Firstly, why is hostility towards immigration motivated by concerns about its fiscal effects in countries where these effects are generally positive? Secondly, why are the elderly and the poor—who benefit the most from the fiscal surplus from immigration—more averse to immigration-friendly policies?

[1] For instance, in 2017 61% of the British citizens over 60 wanted less immigration, while just 45.3% of those under 35 years felt the same way (BSA 2017). In the US, the corresponding values for 2016 are 27.8% and 44.1% (General Social Survey, 2016).

[2] The evidence regarding other European countries is heterogeneous (Boeri, 2010). For an extensive survey on the issues involved in evaluating the impact of immigration on public finances, see Preston (2014).

[3] An informal description of the concept of right-wing populism adopted in this paper can be found in the next section; a formal definition is provided in section 3.

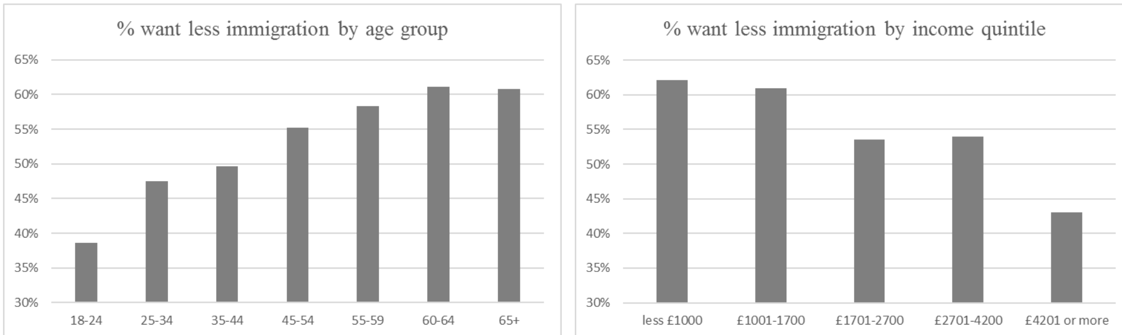

Figure 1: Percentage of respondents wanting the number of immigrants to be reduced by age group (*1A*) and household income quintile (*1B*), British Social Attitude Survey 2017.

We propose a channel that can explain this puzzle. Throughout this paper, we provide theoretical and empirical arguments to argue that it is a *key* channel.

We study an economy in which the native population—the *citizens*—in each period consists of two age groups: young and old. All immigrants are young. Individuals, both citizens and immigrants, live for two periods at most, vary in productivity level, and derive utility from the consumption of private goods and government services. While citizens also have an exogenous common taste for immigration which is meant to capture any non-economic factors affecting their immigration policy preferences,[4] relative policy preferences are entirely driven by economic factors.

In each period the society chooses a two-dimensional policy consisting of an immigration quota and governmental service provisions. The elderly receive an exogenous public pension financed by tax revenues. The government budget is assumed to be balanced in each period. Thus, the policy choice endogenously determines the income tax rate. The age profile of the immigrants implies that they are, on average, net fiscal contributors.

In this setup, voters are able to choose both the immigration policy and how society divides the net fiscal benefits from immigration. This novel feature of the model generates the key trade-off underpinning our results, which is as follows.

Immigration generates a fiscal surplus, which can be employed to (i) increase public spending and/or (ii) reduce taxes. The elderly and the low-income citizens are less affected by income tax changes than are the young and rich citizens. Thus, choice (i) mostly benefits the former sociodemographic groups while choice (ii) favours the latter. This implies that endogenous fiscal policy choices are crucial in

[4]These factors include, among others, the effects of immigration on *compositional amenities* documented in Card *et al.* (2012). These are public goods whose quality depends on the sociodemographic and ethnic composition of the society. They are typically related to the specific religion, traditions, and language of the receiving society.

shaping the attitudes towards immigration of different types of voters.

As a result, the elderly and the low-income citizens (a) support higher public spending than the young and rich and (b) prefer to finance this spending through higher income tax rates rather than through further immigration. Channels (a) and (b) imply that the old and the poor citizens have a common interest. Thus, they coalesce around a political party (or multiple) offering relatively restrictive immigration policies, high public spending, and high taxes. We refer to such coalitions as *right-wing populist parties*.[5] Conversely, young and rich citizens, who support less restrictive immigration policies, a smaller government, and lower taxes, coalesce around *libertarian parties*.

It is important to note that the model generates no actual competition between immigrants and natives over welfare benefits because the fiscal gains from immigration always outweigh its crowding-out effect on public services (with a fixed tax rate). Nevertheless, the political process induces *perceived* competition. Open immigration policies are always endogenously bundled with a relatively small government in libertarian party platforms. The success of these parties results in both more immigration and less public spending. This prompts elderly and poor voters to behave as though they are competing with immigrants over public benefits.[6] That is, they join the right-wing populist parties and support them in the elections.

Demographic shocks tilt the relative power of the two opposing political factions; population ageing and increasing income inequality result in a larger share of elderly and poor voters. Thus, they gain in both size and political power, which fuels the populist parties. This yields an equilibrium policy of low immigration and high public spending. This channel underpins the main analytical results of this paper, which are as follows:

1. A rise in the longevity and/or a fall in the birth rate and/or an increase in income inequality lead to a tightening in immigration policy, an increase in public spending, and a sharp increase in the tax rate. Hence, the political process tends to exacerbate the effects of population ageing on public finances.

2. The policy change is driven by the electoral success of right-wing populist parties, which are characterized by an anti-immigration and pro-public spending platform and are supported by elderly and low-income citizens.

[5] This term is most commonly used in the recent studies, summarized in section 2, to reference parties opposed to both liberal economic policy and immigration. The terms used to refer to this type of party ideology vary across countries. In continental Europe, for instance, they are often referred to as *souverainist* parties to emphasise the role of national sovereignty in their platforms.

[6] This mechanism applies even if the elderly and the poor benefit from public spending financed through immigration as much as or more than the rich.

3. The tightening of immigration policy generates a welfare loss for the entire society, though it mostly harms middle- and upper-class workers as well as future generations.

Moreover, we provide two sets of quantitative results:

1. We calibrate a parametric version of the model to UK data. This exercise reveals that the magnitude of the analytical effects described above may be rather large. For instance, 5 more years of life expectancy at 65 yields a policy allowing for 11.3% fewer working-age immigrants. A 10% decrease in income inequality (measured with the Gini coefficient) yields 11.9% more immigrants.

2. We show that the predictions of the theoretical model are consistent with patterns observed in British Social Attitudes Survey data from the 1995-2017 period. Specifically, age is positively correlated with aversion to immigration and with support for public spending financed through taxes, even after controlling for non-economic factors. Similarly, household income is negatively correlated with those same two attitudinal measures.

These results provide a rationale to explain why ageing countries, which would arguably benefit from more immigration, tend to limit it. Ageing societies tend to disregard the wellbeing of young people–natives and immigrants alike—and future generations. Our analysis suggests that this dynamic, which has widespread economic, demographic, and political consequences, is unlikely to change.

1 Related Literature

The theoretical literature on the political economy of immigration policies is vast.

While some papers focus on immigration policies related to standards, such as skill requirements (Benhabib, 1996; Ortega, 2005), the most common approach, which this paper takes, involves analysing policies that restrict the number of immigrants, such as immigration quotas (see Preston, 2014 for a survey). These studies emphasize the importance of intergenerational aspects related to the pension system (e.g., Razin & Sadka, 1999; Kemnitz, 2003; Leers et al. 2003; Krieger, 2003; Ben-Gad, 2012), and immigrant fertility (Bohn & Lopez-Velasco, 2019) to explain the determinants of political views towards immigration policies. Most of these papers assume a unidimensional policy space. That is, voters choose the immigration quota but not the fiscal policy.

A key finding in the literature is that the assumption of a unidimensional policy space generates inconsistent predictions. This issue is described in Haupt and Peters (1998) and Facchini and Mayda

(2009). These papers study a simple economy characterized by a linear income tax and assume that revenues are provided to all citizens as lump-sum rebates. In this setting, the requirement of unidimensionality can be satisfied in two ways. Either (i) the level of public spending or (ii) the income tax rate must be exogenously determined. According to Facchini and Mayda (2009), these two alternative assumptions correspond, respectively, to the classes of:

1. Tax adjustment models (TAMs; e.g., Scholten and Thum, 1996)
2. Benefit adjustment models (BAMs; e.g., Razin & Sadka, 1999, 2000),

These two model types deliver opposite predictions regarding the relationship between age, pre-tax income, and attitude towards immigration. If immigrants are net fiscal contributors, TAMs show that elderly and low-income citizens are more hostile to immigration than the young and rich citizens; the opposite is true for BAMs.

The intuition that underpins these seemingly contradictory results is as follows. If public spending is exogenously determined, the effect of a rise in the tax base is a fall in the tax rate. Conversely, if the tax rate is unaffected by voter choice, the effect is a rise in public spending per capita. In the former case, immigration mainly benefits young and high-income citizens; in the latter case, the elderly and low-income citizens enjoy the largest share of the gains.

In both models, the endogenous effects of immigration are weakly negative on taxes and weakly positive on public spending when immigrants are net fiscal contributors.[7] Thus, neither of these approaches provides a rationale for the well-documented aversion towards immigration based on its fiscal effects. Preston (2014) argues that the source of this apparent inconsistency lies in how society distributes the gains from immigration and suggests that this puzzle can be addressed by a model that allows for immigration, public spending, and tax policy to be endogenous. Despite this, most studies are based on unidimensional models on account of technical reasons. The traditional literature on voting (e.g., Plott, 1967; Grandmont, 1978) establishes rather restrictive conditions for the existence of a *Condorcet Winner*—a platform that is preferred to any alternative by a majority of voters—if the policy space is multidimensional. This implies that Black's *median voter theorem* (1948) does not typically hold.[8]

[7] Some models in the literature show a negative relationship between immigration and public spending. In Haupt and Peters (1998), for instance, state pensions are decreasing in immigration. This negative relationship is the direct result of an assumption; in our analysis, they are going to be endogenous outcomes of voter choice. Thus, while these models provide interesting predictions, they are unsuitable for the specific questions of this paper.

[8] The use of such conditions is typically restricted to relatively simple problems of redistribution (e.g., Borge and Rattsø 2004; Calabrese, 2007) because they imply strong restrictions on voter preferences. These requirements are generally too restrictive to be satisfied in a dynamic model like the one proposed in this paper. The technical reason for this is illustrated in an online appendix. Alternatives to unidimensional voting models are popular in the literature

Thus, voting models that allow for multiple endogenous policy dimensions require the use of a different solution concept. The downside of most alternative approaches is that they generally do not deliver sharp predictions about the equilibrium policy outcome (see Dillhon (2005) and Dotti, (2019a) for reviews).

To our knowledge, the only attempt to depart from unidimensionality in models in voting models on immigration policies is Razin et al. (2014). They propose an OLG model similar to the one used in this paper, in which the native population consists of skilled workers, unskilled workers, and the elderly. They characterize the political coalitions that can prevail among these three types of voters and derive various interesting predictions. Nevertheless, their approach is unsuitable to answer the questions in this paper, as they assume exogenous tax rates. Thus, the implications of their model, in terms of immigration preferences, are the same as those of a standard BAM.

This paper is based on a successful stream of literature (Levy, 2004, 2005) that exploits the role of political parties in ensuring stability in a multidimensional deterministic voting model. Specifically, we propose a model of electoral competition that extends the static framework from Dotti (2019a, 2020) to a dynamic setting. Under appropriate preferences restrictions, the model delivers sharp predictions about the equilibrium policy outcome and coalition structures; it is, therefore, well-suited to answer the questions in this paper.

Lastly, this paper contributes to the rapidly growing theoretical literature on populism in Western politics. Specifically, it explains the rise of right-wing populism in Europe through a novel channel.

Right-wing populist parties are defined in the literature as those that combine a conservative agenda (e.g., restrictive immigration policy, nationalism, etc.) with anti-elitism (Acemoglu et al. 2013). The anti-elitism factor implies that these parties, in contrast to traditional conservative parties, are generally characterized by vehement rhetoric against income inequality, which they aim to combat by over-inflating the size of government (Dornbusch and Edwards, 1991), regardless of the long-term consequences of such a policy (Guiso et al., 2019).[9]

In line with the literature, we define right-wing populist parties as those that exhibit a typically conservative stance on immigration and a traditionally left-wing fiscal policy platform. This description

but are usually not useful for answering questions about the comparative statics of the equilibrium policy outcome. See Dotti (2019a) for a detailed analysis of the advantages and disadvantages of various theoretical frameworks in the study of comparative statics in voting models.

[9] Guiso et al. (2019) describe two identifying features of populism: (1) the claim to be "on the side of the people against the elite"; (2) the promotion of "policies without regard to the long-term or indirect consequences". These two characteristics imply supply-oriented solutions, which they call economic populism. Economic populism typically entails a policy platform that is both strongly redistributive (e.g., a minimum income), and myopic in its goals (e.g., reduced immigration, higher public debt, lower retirement age, etc.). Similarly, Mudde (2007) describes right-wing populism as "opposing neoliberalism and immigration".

is consistent with the political trajectories of several parties typically labelled right-wing populist parties in Europe, such as the *National Front* in France, the *United Kingdom Independence Party*[10] and the *Northern League* in Italy. Over the last two decades, these parties have shifted away from their early libertarian economic positions to strong interventionist views (Minkenberg, 2000; Mudde, 2007), especially regarding certain provisions, including healthcare, social services, and elderly care.[11]

2 The Model

This section consists of two parts: (1) the economic model of immigration and public spending and (2) a description of the political process.

2.1 The Economic Model

We propose an infinite-horizon overlapping-generations model of immigration and public spending akin to those in the literature, particularly the model in Razin and Sadka (1999). Unlike their model, however, both public spending and immigration are endogenous in our model.

2.1.1 Demographic Structure

Each period t has length normalized to 1 and features a continuum of individuals[12] divided into two generational groups: the working-age population (Y) and the elderly (O). All immigrants are all in the working-age group. Within the working-age population are n_t natives and m_t immigrants. The elderly population has size o_t, which includes those who were immigrants in period $t-1$. The size of each group is summarized in *Fig. 2*.

Natives and immigrants potentially have different exogenous expected fertility rates, which are denoted by σ and σ^m, respectively. Let $\Delta \equiv \sigma^m - \sigma$. We assume that $\Delta \geq 0$[13] and that the supply of potential

[10] While initially labelled a libertarian party advocating a smaller state, UKIP has consistently proposed a policy platform characterized by a substantial increase in public spending. For instance, the party manifesto for the 2015 national elections pledged "an extra £3bn a year into the NHS in England" and "a commitment to spend 2% of GDP on defence initially, looking to increase it substantially after that". These figures far exceeded the pledges of their main rivals, the Conservatives and the Labour Party. See: Curtice, 2012.

[11] Other examples include the Freedom Party in Austria, the Danish People's Party in Denmark, and Fidesz in Hungary. In some cases, right-wing populist parties stem from the evolution of pre-existing far-right statist groups (e.g., the Finns Party in Finland, the Sweden Democrats in Sweden, and Brothers of Italy). Other are entirely new parties founded with their modern populist platforms sometime in the last 20 years (e.g., Independent Greeks in Greece, the Party for Freedom in the Netherlands, and Law and Justice in Poland).

[12] This assumption represents a society with a large number of citizens, and is common in the literature (e.g. Razin and Sadka, 1999).

[13] We assume Δ to be constant for all periods $t = 1, 2, ..T$, with $T \to +\infty$. Thus, any change in Δ should be interpreted as an unanticipated shock. We assume $\Delta = 0$ for any period $t > T$. The latter assumption is imposed for technical reasons.

$t-1$			t			$t+1$	
$\lambda(n_{t-2}+m_{t-2})$	(O)	$\rightarrow$	$\times$				
$n_{t-1}+m_{t-1}$	(Y)	$\rightarrow$	$\lambda(n_{t-1}+m_{t-1})$	(O)	$\rightarrow$	$\times$	
		born	n_t+m_t	(Y)	$\rightarrow$	$\lambda(n_t+m_t)$	(O)
				born	$\rightarrow$	$n_{t+1}+m_{t+1}$	(Y)

Figure 2: Size of each generation.

time	$t-1$			t			$t+1$	
born $t-2$	OLD	(O)	$\rightarrow$	$\times$				
born $t-1$	NATIVE (n_{t-1}) Immigrant (m_{t-1})	(Y)	$\rightarrow$	OLD	(O)	$\rightarrow$	$\times$	
born t				NATIVE (n_t) Immigrant (m_t)	(Y)	$\rightarrow$	OLD	(O)
born $t+1$							NATIVE (n_{t+1}) Immigrant (m_{t+1})	(Y)

Figure 3: Structure of Overlapping Generations and Voting Rights.

immigrants is large.[14] At the end of each period, the immigrants and their children are fully assimilated to the native population (i.e. they become identical to natives of the same age group).[15] Under these assumptions, the size of the working-age native population is given by the formula $n_{t+1} = \sigma n_t + \sigma^m m_t$. A young individual at time t survives to period $t+1$ with probability $\lambda \in [0, 1]$. Thus, life expectancy at birth is $1 + \lambda \leq 2$ and the size of the elderly population is $o_t = \lambda(n_{t-1} + m_{t-1})$. Note that o_t is an increasing function of life expectancy.

2.1.2 Citizenship and Voting Rights

Fig. 3 summarizes the structure of voting rights. We assume that only the $n_t + o_t$ *citizens* (young natives and the elderly, marked by capital letters in *Fig. 3*) vote—immigrants do not. Immigrants can acquire voting rights by residing in the country for one period (i.e., citizenship through *naturalization*). Their children are considered to be native citizens with full rights (*ius soli*). This is consistent with the rules to acquire citizenship in several Western countries.[16] In Section 3.7, we show that the results carry over under alternative assumptions regarding the acquisition of voting rights.

[14] This assumption ensures that any immigration quota adopted by the government within the range of available policies is binding, meaning that the number of immigrants is always exactly equal to the quota in each t.

[15] Assimilation applies to more features than just fertility behaviour (e.g., immigration preferences; see section 2.1.3).

[16] *Ius soli* is the citizenship rule in US and Canada but not in European countries. Nevertheless, virtually all second-generation immigrants in Europe acquire citizenship by the time they reach legal voting age. Thus, for the purpose of this analysis, this is a reasonable assumption. In some countries, such as Japan, the immigrants and their children do not automatically become citizens unless they have a native parent (*ius sanguinis*).

2.1.3 Individual Preferences

A citizen i of group (Y) in period t has preferences over consumption of private goods C_t^i, the extent of government services G_t, and the share of immigrants in society $M_t = m_t/(m_t + n_t)$, represented by the following utility function:

$$U_t^{i,Y}\left(\left\{C_{t+r}^i, M_{t+r}, G_{t+r}\right\}_{r=0}^{1}\right) = C_t^i + b(G_t) + c(M_t) + \beta\lambda\left[C_{t+1}^i + d(G_{t+1}) + c(M_{t+1})\right] \tag{1}$$

where β captures how an individual discounts future utility and the functions b, d, and c are strictly concave C^∞ functions. The function c represents an exogenous taste for immigration and is the same for all citizens.[17] It captures all non-economic factors affecting voter preferences regarding immigration. Its domain is $[0, \overline{M}]$, where $\overline{M} < 1$ is the level corresponding to fully unregulated immigration. We do not restrict the sign of c and c' for interior values of M_t but we assume $c'(0) \geq 0$ and $c'(\overline{M}) = -\infty$ (i.e. citizens are strongly averse to fully unregulated immigration but don't mind a small number of immigrants). The presence of c in the utility function facilitates the derivation of the results and does not affect the mechanisms underpinning the results of the paper (details in Appendix B). Additionally, $G_t \in [0, \overline{G}]$ and b satisfies $b'(0) = +\infty$ and $b'(\overline{G}) = 0$.

For retired individuals in period t the direct utility $U_t^{i,O}$ is constructed in a similar manner, except it is solely a function of consumption, government services and immigration in the current period:

$$U_t^{i,O}\left(C_t^i, M_t, G_t\right) = C_t^i + d(G_t) + c(M_t) \tag{2}$$

where the features of c and d are illustrated in the previous paragraphs. Lastly, immigrants consume both private goods and government services in the same way natives do; however, their preference specification is irrelevant for electoral outcomes, as they do not vote.[18]

[17] This assumption can be easily relaxed as long as the marginal effect of a rise in M_t on the taste component is nondecreasing in income. Similarly, I can allow for young and old citizens to have different exogenous tastes for immigration as long as the marginal effect of a rise in M_t on the taste component of old individuals is lower than (or equal to) that of any young citizen.

[18] Immigrant preferences play a role in welfare analysis; this aspect is illustrated in Section 3.5.

2.1.4 Production

Each working-age individual can be employed either in the private or public sector at a wage rate equal to their productivity. The private sector produces the consumption good using a linear technology and labour as input. Thus, the total production of consumption goods equals the total gross income of private-sector workers. The public sector produces government services. The per-capita provision of government services $G_t \in [0, \overline{G}]$ is assumed to be proportional to the share of total effective labour employed in the public sector.[19]

Each individual is endowed with 1 unit of time, and their labour supply is perfectly inelastic. This assumption simplifies the analysis and does not drive the trade-offs that underpin this paper's predictions.[20] In particular, all the results hold if the wage elasticity of labour supply is positive for all productivity levels (details in the online appendix).

Let $y_t^i = \xi \theta_t^i$ denote the income of citizen i in period t, where ξ is an aggregate component and θ_t^i is i's productivity type. The type θ_t^i is continuously distributed with time-invariant c.d.f. $Q(\theta_t; \rho)$ and p.d.f. $q(\theta_t; \rho)$. The parameter ρ captures the degree of income inequality. Specifically, the c.d.f. $Q(\theta_t; \rho)$ with $\rho \in [0, 1]$ is constructed as follows:

$$Q(\theta_t; \rho) = \rho Q_2(\theta_t) + (1 - \rho) Q_1(\theta_t) \tag{3}$$

where the functions Q_1 and Q_2 satisfy (i) $Q_1(\theta_t) < Q_2(\theta_t)$ for all $0 < \theta_t < \hat{\theta}$ and $Q_1(\theta_t) \geq Q_2(\theta_t)$ for all $\theta_t \geq \hat{\theta}$ for a given threshold $\hat{\theta}$ that satisfies $Q_1(\hat{\theta}) \geq .5$, and (ii) $\int \theta_t dQ_1(\theta_t) = \int \theta_t dQ_2(\theta_t) = 1$. These conditions mean that the distribution of θ_t under Q_2 is a mean-preserving spread of that under Q_1. A marginal increase in the parameter ρ is equivalent to a marginal increase in the variance of income distribution at constant mean income $\bar{y} = \xi$; thus, it can be interpreted as an increase in income inequality.[21]

We assume that the expected productivity of immigrants is independent of current policies. While such an assumption is admittedly restrictive, it is intended to describe an economy facing a large

[19]Thus, government services are a partially congested public good enjoyed by both young and old individuals. This assumption ensures that the tax-price of G_t (i.e., the marginal effect of an increase in G_t on the tax paid by a given citizen) does not move to zero in the long run as a consequence of population growth. Examples of public services that display these features include public transportation, public offices, and the police.

[20]The assumption of a production function that is linear in labour is common in related literature (e.g., Razin & Sadka, 2000). It is justified if one considers that, in a more complex economy, these effects tend to be offset by adjustments in the stock of capital (not explicitly assumed in this analysis) that occurs over the relatively long framework of a generation. This mechanism is considered to be particularly effective for offsetting long-run effects of immigration on wages if firms have access to international capital markets (see: Ben-Gad, 2012).

[21]A similar comparative statics exercise is performed in Dotti (2020) for a model of redistribution.

supply of potential immigrants that cannot effectively select immigrants on the basis of observable characteristics.[22] Upon arrival, the immigrants do not possess the host country-specific skills required by its production environment (e.g., language proficiency, see: Preston, 2014). Thus, to fill this gap the government initially faces an average cost per immigrant equal to a share l of per capita income.[23] Net of such a cost, immigrants and natives are assumed to possess the same expected productivity.

2.1.5 Social Security

We assume the existence of a public pension system. Specifically, in each period t, the government transfers a net amount $p_t^i = p_t(y_{t-1}^i, g_t)$ to each retiree, where p_t is weakly increasing in y_{t-1}^i and $g_t = o_t/n_t$ denotes the *citizens' old-age dependency ratio* (i.e. the old-age dependency ratio measured before immigration takes place in period t).[24]

The total amount allocated to social security for the elderly in period t is equal to a fixed share γ of the expected total income of the working-age citizens (i.e., p_t satisfies $\left[\int p_t(\xi\theta, g_t)q(\theta,\rho)d\theta\right] \times o_t = \gamma\bar{y}n_t$). This means that (i) the pension system features an *automatic balance mechanism,*[25] and (ii) the cost of the pension system per worker $\gamma\bar{y}\frac{n_t}{n_t+m_t} = \gamma\bar{y}\left(1 - M_t\right)$ is decreasing in the share of immigrants to young residents M_t.

These two features have two key consequences for the pension system: (i) its total cost adjusts according to anticipated changes in life expectancy and fertility in order to preserve its sustainability in the long run and ensure that changes in the size of government are solely driven by endogenous political choices; (ii) its tax-cost per worker can be reduced by widening the tax base through a less restrictive immigration policy.

While we assume that p_t^i is exogenous to current voter choices, the existence of a public pension system can be justified in infinite-horizon OLG models as the outcome of a self-enforcing intergenerational agreement, as shown in in Rangel and Zeckhauser (2001) and Boldrin and Montes (2005). These

[22] This also means that we assume away the possibility of endogenous self-selection of welfare-dependent immigrants based on the amount of public benefits provided by the receiving country. While this is a theoretically important concern (Borjas, 1999), the size of this effect is generally deemed to be fairly small (Preston, 2014). Thus, for the purpose of this study, our assumption is a reasonable approximation.

[23] Alternatively, l can be interpreted as a reduced-form expression for the forgone tax revenue due to the initial lack of such country-specific skills or to a lower average productivity among immigrants relative to the natives.

[24] The outlined pension system is very flexible. It is consistent with both a Beveridgean (if p_t is constant in y_{t-1}^i) and a Bismarckian (if p_t is increasing and linear in y_{t-1}^i) pension system, as well as with a combination of the two. Alternatively, one can assume that the government collects pension contributions from each worker and commits to pay a pension which is increasing the worker's contribution.

[25] Automatic balance mechanisms are becoming increasingly common. These mechanisms consist of a formula that translates a change in average life expectancy into a change in monthly pension payments (e.g., Finland after the 2005 reform), or into a change in the retirement age (e.g., Italy after reforms were introduced in 2010 and extended in 2011). Automatic balance mechanisms are also embedded in the pension systems of Canada, Germany, Japan and Sweden.

studies indicate that the extent at which working-age people can reduce their net transfers to the elderly through taxation is limited, as the long-run sustainability of the intergenerational agreement depends upon the *net* benefits young expect to receive in old age. In line with this interpretation, p_t^i denotes the promised pension *net of taxes* on social security benefits.[26]

While these assumptions are admittedly restrictive, they are consistent with the principle embedded in most social security systems—individual pension benefits are determined at the time of retirement or earlier and are mostly unaffected by policy changes that occur afterwards.

2.1.6 Public Finances

The public sector raises revenue through a linear tax τ_t on labour income and spend it on public sectors salaries, immigration costs, and pensions for the elderly.

We do not consider public debt and assume that the government budget is balanced in every period. This assumption simplifies the analysis but does not affect voter trade-offs. Using the aforementioned formulas for the total public sector salary costs $G_t\bar{y}(m_t + n_t)$, total pension expenditures $\gamma\bar{y}n_t$ and immigration costs $l\ m_t\bar{y}$, the government budget constraint is constructed as follows:

$$\tau_t \geq \frac{G_t\left(n_t + m_t\right)\bar{y} + \gamma\bar{y}n_t + l\ m_t\bar{y}}{\left(n_t + m_t\right)\bar{y}} = \frac{Total\ Spending_t}{Total\ Output_t} \tag{4}$$

The right-hand side of (4) is the size of the government, defined as the ratio of total public spending to total output. We define the degree of *economic liberalism* (vs. *interventionisms*) of the fiscal policy, denoted by L_t, as the difference between the maximum size of government in period t at given immigration policy M_t and its actual level. Given the assumptions made on the social security system, it is easy to show that $L_t = \overline{G} - G_t$ (i.e., it depends solely on endogenous public spending). This variable change is just a matter of convenience—the reasons behind it is made clear in Section 3.

Assuming that the government budget constraint is always satisfied with equality,[27] we can solve (4) for τ_t and define the tax rate function $\tau\left(M_t, L_t\right)$ as follows:

$$\tau\left(M_t, L_t\right) = \gamma + \overline{G} - L_t - \left(\gamma - l\right)M_t \tag{5}$$

[26] Of course, in reality net pension benefits may be affected by changes in the tax policy. Nevertheless, for the reasons summarized in this section, the tax rate on social security benefits is unlikely to be very responsive to endogenous policy changes. Moreover, social security for the elderly often includes benefits that are exempt from taxes, such as public health insurance (e.g., Medicare in the US) and subsidized home services.

[27] This must be true at any equilibrium of the voting game. This becomes clear after the equilibrium concept is described in Section 3.2.

where we assume $0 < \overline{G} < 1 - \gamma$ to ensure that $0 < \tau(M_t, L_t) < 1$ for all $(M_t, L_t) \in X$.[28] The consequences of relaxing this assumption are illustrated in the online appendix.

Note that formula (5) suggests that working-age citizens can ease their tax burden by voting for more open immigration policies so long as $\gamma > l$ (i.e., so long as immigration costs are not too large). Formula (5) also illustrates that the tax rate $\tau(M_t, L_t)$ is also equal to the size of the government.

Under these assumptions, a working-age individual's private goods consumption is given by her post-tax income, such that $C_t^i = [1 - \tau(M_t, L_t)]\, y_t^i$.

2.1.7 Policy Space

Voters face a two-dimensional policy space in each period t. Policy platforms consist of an immigration quota M_t and a degree of economic liberalism L_t. Let $\underline{M} \in \left[0, \overline{M}\right]$ be the level of M_t that satisfies $c'(\underline{M}) = 0$.[29] The policy space is the set $X \equiv \left[\underline{M}, \overline{M}\right] \times \left[0, \overline{G}\right]$ with typical element (M_t, L_t).[30]

2.1.8 Citizens' Objective Function

Let $\varphi = (\beta, \gamma, \lambda, \mathit{\Delta}, \sigma^m, \xi, l, \rho)$ be a vector of common parameters. Using the formula for p_t^i into the utility function of an elderly citizen in (2) I obtain their indirect utility, which is shown through $U_t^{i,O}\left(p_t(y_{t-1}^i, g_t), M_t, \overline{G} - L_t\right) = p_t(y_{t-1}^i, g_t) + d(\overline{G} - L_t) + c(M_t)$. This formula shows that an old citizens' preferences over (M_t, L_t) in period t are independent of their pension levels, income when young, expectations of future policies (M_{t+1}, L_{t+1}), and history up to period t. Thus, all elderly have the same policy preferences; as such, they can be represented by a function $u_t^{i,O} = u\left((M_t, L_t), (M_{t+1}, L_{t+1}); -1, \varphi, g_t\right)$, which has the formula:

$$u_t^{i,O} = d(\overline{G} - L_t) + c(M_t), \tag{6}$$

where $\theta_t = -1$ denotes the preference type of *all* old citizens.

This representation dramatically simplifies the analysis. Since $u_t^{i,O}$ is independent of θ_{t-1}^i and $Q(\theta_t; \rho)$ is time-invariant, the dynamic framework essentially becomes equivalent to one in which an entirely new population of citizens replaces the previous one at the end of each period, such that the age distribution of the "new" fictitious population is determined solely by the citizens' old-age dependency

[28] This restriction is crucial for the results in the next section to hold. If the tax rate hits the upper bound, the predictions of the model become those of a standard benefit adjustment model, as illustrated in the online appendix.

[29] Such value always exists in $\left[0, \overline{M}\right]$ given the assumptions established in Section 2.1.3.

[30] Note that this definition excludes from X all the policies with $M_t < \underline{M}$ (if any exists in X). This assumption is innocuous because none of these policies are Pareto efficient, meaning they cannot be credibly proposed by any party.

ratio g_t. Thus, the model exhibits no dynamic features typical of an OLG model aside from the evolution of the unique aggregate state g_t.[31]

A young citizen has preference type equal to their productivity parameter $\theta_t \in [0, +\infty)$. I construct the distribution of citizen types θ_t in period t (conditional on history h_t), which possesses the following c.d.f.:

$$F_t(\theta_t; \rho \mid h_t) = \begin{cases} 0 & if \quad \theta_t < -1 \\ \frac{g_t(h_t)}{1+g_t(h_t)} & if \quad -1 \leq \theta_t < 0 \\ \frac{g_t(h_t)+Q(\theta_t;\rho)}{1+g_t(h_t)} & if \quad \theta_t \geq 0 \end{cases} \tag{7}$$

where $g_t(h_t)$ denotes the citizens' old-age dependency ratio prevailing after history $h_t \in H_t$. Lastly, I use (7) to define the totally ordered set of citizens' types at time t as $\Theta_t := \{-1\} \cup [0, +\infty)$ (i.e., the set of types that possess non-zero probability density [or mass]).

In a similar manner, we derive a young citizen's preferences over policies at time t as their expected indirect utility $u_t^{i,Y} = u\left((M_t, L_t), (M_{t+1}, L_{t+1}); \theta_t^i, \varphi, g_t\right)$. Using formula (1), this becomes:

$$u_t^{i,Y} = \begin{array}{c} \left[1 - \tau\left(M_t, L_t\right)\right]\xi\theta_t^i + b(\overline{G} - L_t) + c(M_t) + \\ +\beta\lambda E_t\left[p_{t+1}(\xi\theta_t^i, g_{t+1}) + d(\overline{G} - L_{t+1}) + c(M_{t+1}) \mid (M_t, L_t), h_t\right] \end{array} \tag{8}$$

where the second term depends upon the expectations regarding future policies (M_{t+1}, L_{t+1}) given the history h_t and current policy choices (M_t, L_t).

Formulas (6) and (8) demonstrate that low-income and elderly citizens always prefer, given a certain level of public spending, a policy that finances it through higher income taxes rather than through a larger number of immigrants. This tradeoff holds true despite the net positive fiscal contribution of immigrants, of which the elderly and the poor are net beneficiaries.

2.2 The Political Process

In this section, we provide informal descriptions of the political process and the equilibrium concept. A formal description is provided in Appendix A.1.

[31] Because each old citizen i's objective function in period t is independent of θ_t^i, the state of the economy is fully summarized by the aggregate state g_t. Thus, we can construct a single collective agent in each period t (details provided in the following section)without needing to include each citizen's productivity type in the state space.

We propose a model of electoral competition among endogenous political parties that incorporates the key insights from Dotti (2019a, 2020) into a dynamic setting based on Maskin and Tirole (2001). In each period t, a political process translates individual preferences into a policy outcome $x_t^* = (M_t^*, L_t^*) \in X \cup \{x^0\}$, where x^0 is the default policy.[32] Each type of citizen is member of a single party $P_t^j \in \mathbb{P}_t$, where $\mathbb{P}_t = \{P_t^1, P_t^2, ..., P_t^J\}$ is a partition of Θ. Each party P_t^j selects one type $\theta^j \in P_t^j$, who can either run in elections as a candidate or remain inactive. Each candidate credibly commits to a platform $a_t^j \in X$. The citizens choose the winning candidate using the method of majority rule.

Let $\left\{a_t^j\right\}_{P_t^j \in \mathbb{P}_t}$ be the collection of all party platforms in period t, where $a_t^j = \emptyset$ denotes an inactive party, and $A_t(\mathbb{P}_t)$ be the corresponding set of policy platforms available to voters in given party structure $\mathbb{P}_t$.

The equilibrium concept builds on Maskin and Tirole's (2001) definition of Markov perfect equilibrium. The definition consists of two parts.

First, we define a *collective strategy* s as a function that selects a political outcome at each point in time t and after any history $h_t \in H_t$, where H_t denotes the set of all possible histories up to period t. A political outcome in period t is a tuple consisting of a partition of the set of citizens into parties $\mathbb{P}_t$, a set of party platforms $A_t(\mathbb{P}_t)$, and an equilibrium policy x_t^*.

In line with Maskin and Tirole (2001), we assume that the society plays an equilibrium collective strategy in every period and after any history and that this collective strategy is Markovian. In other words, the equilibrium concept treats the population of citizens in each period, with respect to the definition of Markov perfection, as a collective agent.[33]

We depart from Maskin and Tirole (2001) in the type of equilibrium that is played in every period. They adopt a non-cooperative solution concept: the agents are required to play Nash equilibrium after any history. In contrast, we define subgame perfection with respect to a cooperative concept that we refer to as *stable political outcome (SPO)*. The *SPO* is a simplified version of the static equilibrium concept in Dotti (2019a), which is itself an evolution of Levy (2004).

Informally, a political outcome $\langle \mathbb{P}_t, A_t(\mathbb{P}_t), x_t^* \rangle$ in period t is a *SPO* given h_t and s if satisfies:

(i) *Majority rule (MR)*: the equilibrium policy x_t^* is selected from $A_t(\mathbb{P}_t)$ through majority rule whenever such method delivers an outcome; otherwise, it is the default policy x^0.[34]

[32] We assume that the default policy x^0 satisfies $u\left(x^0, x_{t+1}\theta_t, \varphi, g_t\right) = -\infty$ for all $\theta_t \in \Theta$ and all $\varphi \in \Phi$.This assumption is common in similar models of elections, such as Levy (2004, 2005) and Dotti (2020).

[33] This simplification is possible because the dynamic problem features a unique aggregate state of the economy, as illustrated in the previous section.

[34] We adopt the traditional social choice definition of such a method. This differs from the definition of majority rule as a voting method in elections (see: May, 1952). This assumption is equivalent to choosing the winning candidate using

(ii) *Citizen-candidates (CC)*: each party's platform is *credible* only if it is an ideal policy of a party member, as in Osborne and Slivinski (1996) and Besley and Coate (1997).[35]

(iii) *Partisan Membership (PM)*: a citizen is member of a party only if they weakly prefer the party platform to that proposed by any other party.[36]

(iv) *Core Stability (CS)*: a partition is stable if and only if there is no citizen coalition that—independently of the actions and the party membership of other citizens—can obtain a strictly better outcome for all its members by forming a new party and selecting a candidate among its members.[37]

Second, we define a *Markov-perfect party equilibrium (MPPE)* as a Markovian collective strategy that forms a *SPO* after any history h_t in each period t, for any $t = 1, 2, ...$ The *MPPE* is generally not unique; however, we can derive a useful characterization of all equilibria under relatively mild restrictions.

Third, we define the *pivotal citizen* in period t (denoted by θ_t^p) as a type of citizen—if one exists—for which the equilibrium policy x_t^* in period t given history h_t for all s_t that form a *MPPE* is the ideal policy—i.e., $x_t^* = x_t^p$ with $x_t^p \in \arg\max_{x_t \in X} E_t \left[u\left(x_t, x_{t+1}^*; \theta_t^p, \varphi, g_t\right) \middle| s_t, h_t \right]$.

3 Results

In this section, we present the main results, namely the existence and characterization of the voting equilibrium and steady-state as well as the analytical comparative statics results.

3.1 Equilibrium Existence and Characterization

The model presented in the previous section, which underpins the results in this section, exhibits the following properties:

(i) The policy space X is a compact set and the partially ordered set $(X, \leq)$ is a complete sublattice of $(\mathbb{R}^2, \leq)$

(ii) The set of citizen types Θ is a totally ordered set.

a Condorcet method. This assumption allows us to abstract from the issues that typically arise in electoral games with strategic voters. For instance, it is well known that most voting methods (including plurality rule and majority rule) yield equilibria in which the Condorcet winner is not selected even if preferences are single-peaked. Exceptions are the Condorcet procedure and the multi-stage runoff (see: Bag et al., 2009).

[35] Condition *CC* captures the idea that, in the absence of reputation concerns, a policy-motivated candidate cannot credibly commit to any platform other than their most preferred one.

[36] Condition *PM* closely resembles Levy's (2005) concept of *Partisan equilibrium.*

[37] *CS* represents a fairly weak stability condition. All the results hold true for a range of more restrictive conditions, see Appendix B. For a detailed description of the possible assumption regarding profitable deviations in coalition games in partition function form, see Ray and Vohra (2011).

(iii) Citizens' preferences given history h_t and collective strategy s satisfy *Quasisupermodularity (QSM)* in (M_t, L_t) and the *Strict Single Crossing Property (SSC)* in $(M_t, L_t; \theta_t)$ (proof in Appendix B)

The conditions *QSM* and *SSC* are essentially those defined in Milgrom and Shannon (1994). They are widely used in many subfields of Economic Theory. Formal definitions of *QSM* and *SSC* are provided in Appendix A.2.

It is important to highlight that *SSC* is a much less restrictive assumption than both the unidimensional single crossing condition in Gans and Smart (1996) and single peakedness in Black (1948). As a result, this condition alone is insufficient to ensure the existence of a Condorcet winner. In fact, in our economic model, voter preferences satisfy *QSM* and *SSC* but, typically, neither single peakedness nor unidimensional single crossing over X. Thus, a Condorcet winner over X generally does not exist.[38]

Given these three properties, I can show that a stationary Markov-perfect party equilibrium (*MPPE*) always exists and derive sufficient conditions for uniqueness of the equilibrium policy outcome x_t^*. These results are formalized in the following proposition.

Proposition 1. *(i) A stationary MPPE always exists. (ii) In any MPPE the policy outcome x_{t+r}^* for $r = 0, 1, 2, ...$ is an ideal policy of the pivotal citizen θ_{t+r}^p. (iii) The pivotal citizen's type θ_{t+r}^p is weakly decreasing in g_{t+r}. (iv) There exists $\hat{\Delta} > 0$, such that if $\Delta \in [0, \hat{\Delta})$, then the equilibrium policy outcome x_t^* in each period t is unique given state g_t.*

Proof. See Appendix B.1.

Proposition 1(iv) implies that (1) if $\Delta \geq \hat{\Delta}$, then there may be more than one policy outcome that is consistent with *MPPE*,[39] and (2) even if $\Delta \in [0, \hat{\Delta})$ holds, the *MPPE* may not be unique. however, in any *MPPE,* the policy outcome in each period t is the same. This is in line with the fact that in each period t, there are typically several *SPO*s featuring different party structures $\mathbb{P}_t$ for any given state g_t that all support the same policy outcome $x_t^* = x_t^p$.

[38] A detailed description of these conditions and and proof of non-existence of a Condorcet winner in our setup are provided in the online appendix.

[39] In such a case there may be multiple future equilibrium trajectories which are consistent with perfect foresight.

3.2 Intuition

The intuition underpinning the existence of a voting equilibrium in each period is simple. For illustrative purposes, suppose $\Delta \in [0, \hat{\Delta})$,[40] and consider a period t. First, the voter objective function is strictly concave[41] and X is a compact set. Thus, each citizen possesses a unique ideal policy. Second, voter preferences satisfies QSM in (M_t, L_t) and SSC in $(M_t, L_t; \theta_t^i)$ over a complete lattice. Thus, theorem 4 in Milgrom and Shannon (1994) implies that citizens' ideal policies are totally ordered in X, as illustrated in Figure 4.[42]. Third, only ideal policies can be credibly proposed under the citizen-candidate condition (CC), meaning that any credible platform profile A_t is also a totally ordered set.

Voter preferences over the multidimensional choice domain X do not generally satisfy single peakedness[43] (Plott, 1967; Grandmont, 1978). Nevertheless, since any credible platform profile A_t is totally ordered, the restricted policy space A_t is equivalent to a unidimensional choice domain given voter preferences, which means that, in turn, single peakedness holds true over any such sets. In other words, the multidimensional majority voting game over X is reduced to a unidimensional one over A_t.

Black's (1948) theorem states that if voters possess single-peaked preferences, the outcome of the Condorcet method is always the alternative that is preferred by the median voter. In the model proposed in this paper, such an individual—referred to as the pivotal citizen—is one that possesses the median type $\theta_t^p \in \Theta$ in period t with respect to the distribution defined in (7).

Given that the median voter theorem holds over any credible platform profile A_t, it is easy to show that, in any stable political outcome, the pivotal citizen obtains their most preferred policy. Specifically, for any tuple $\langle \mathbb{P}_t, A_t(\mathbb{P}_t), x_t^* \rangle$, such that $x_t^p \notin A_t(\mathbb{P}_t)$ and therefore $x_t^* \neq x_t^p$, all citizens of type θ_t^p can cooperatively deviate by forming a party $\tilde{P}_t^d = \{\theta_t^p\}$ and proposing the platform x_t^p. In such a scenario, the newly formed party wins the elections given any possible credible alternative that other parties can put forward; all of its members are made strictly better off as a result of this deviation. Thus, the proposed political outcome does not satisfy CS, the stability condition. This median voter result is used to derive this paper's main predictions. In particular, it facilitate the derivation of monotone comparative statics results, which are used to derive the main results in Section 3.4.

[40] This condition is not required for Proposition 1 part (i) and (ii) to hold true but facilitates the understanding of the mechanisms underpinning the results.

[41] See Lemma 2 in Appendix B.1.

[42] This example illustrates the simple case in which $\lambda = 0$, $d(\cdot) = 0.2\ln(\cdot)$, and $c(\cdot) = -\frac{\gamma - l}{2}(\cdot)^2$.

[43] They also do not typically satisfy other conditions that ensures the existence of a Condorcet winner, such as the unidimensional single-crossing property (Gans & Smart, 1996).

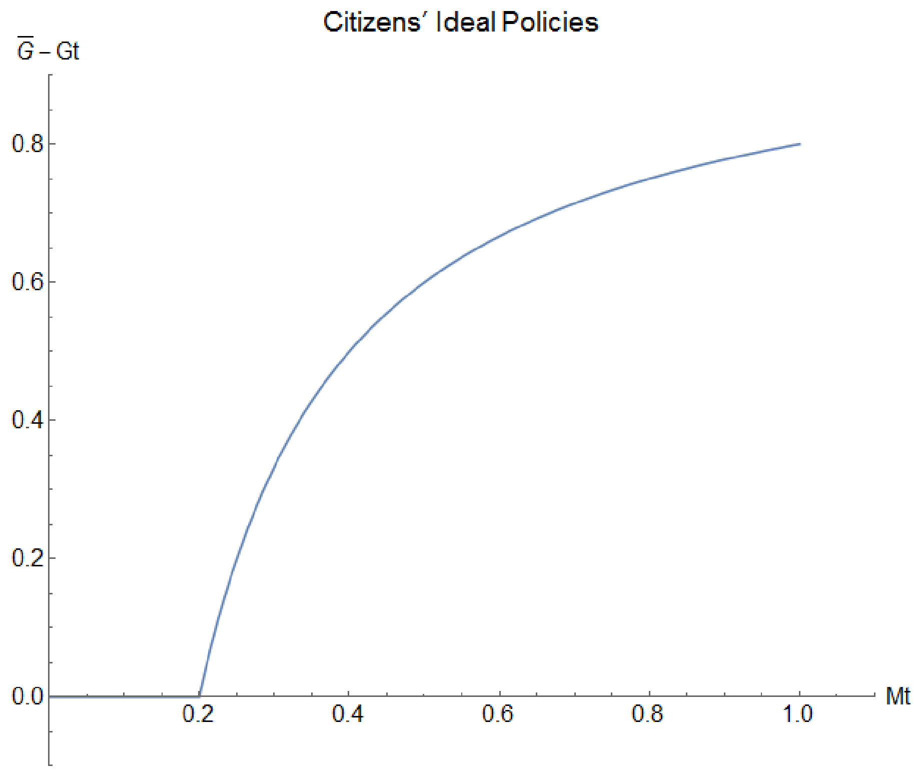

Figure 4: Citizens' ideal policies in the illustrative example.

3.3 Steady-State

An *MPPE* a is steady-state at time t if, in each period $t+r$ with $r = 0, 1, 2, ...$, the state of the economy is constant. That is, $g_{t+r} = g_t = g_{SS}$, where the subscript SS denote the steady-state value of a (state or control) variable. Note that in a steady-state, the equilibrium policy $x_t^* = x_{SS}^*$ is constant over time.[44]

We derive the following result.

Proposition 2. *There exists $\check{\Delta} > 0$ such that if (i) $\Delta \in [0, \check{\Delta})$, then a steady-state exists, which is unique and globally stable. If $\Delta = 0$, then (ii) any MPPE at time t is history-independent and the economy converges immediately to the steady state after any history h_t.*

Proof. See Appendix B.2.

[44] Conversely, $\mathbb{P}_t$ and $A_t\left(\mathbb{P}_t\right)$ do not need to be time-invariant at a steady-state.

Proposition 2 ensures that, as long as the fertility rate among immigrants is sufficiently close to that among the natives (i.e., $\mathit{\Delta} \in [0, \check{\mathit{\Delta}})$), a unique steady-state exists and exhibits several desirable properties that makes it suitable for the analysis in the next section.

3.4 Effect of Ageing, Rising Inequality and Productivity Shocks

In this section, we analyse the long-run effects of permanent shocks on some key parameters in the model. In this section, for simplicity we assume that the condition $\mathit{\Delta} \in [0, \check{\mathit{\Delta}})$ stated in Proposition 2 is satisfied. We relax this assumption in Section 6. I aim to study the effects of four types of shock, which are defined below.

Definition 1.

(a) An *increase in longevity* is a permanent rise in the life expectancy parameter λ.

(b) A *decrease in fertility* is a permanent fall in the birth rate parameter of the native population σ.

(c) An *increase in income inequality* is as a permanent rise in the relative inequality parameter ρ.

(d) An *economic depression* is as a permanent fall in aggregate productivity parameter ξ.

A decrease in fertility in period t is *critical* if $g_t \geq \hat{g}_t$ for some positive threshold $\hat{g}_t \in (0, 1)$[45] (i.e., if the citizens' old-age dependency ratio is sufficiently close to 1 before the shock occurs).

The main result of this paper stems from studying the effects of a shock of type (a), (b), (c), and/or (d) on the steady-state value of the key equilibrium outcomes of this economy. A sociodemographic and/or productivity shock affects citizens' preferences in three possible ways:

(i) it changes the demographic composition of the voting population and, in turn, the identity of the pivotal citizen (*political effect*)

(ii) it directly affects the government budget constraint (e.g., higher cost of pensions, smaller tax base, etc.) (*budget effect*)

(iii) it affects voter expectations regarding future equilibrium policies, both directly and through the effect of changes in current policy choices (*sophisticated farsightedness*).

The assumption that the pension system features an automatic balance mechanism illustrated in Section 2.1.5 implies that effect (ii) is zero during the period in which the shock occurs. Moreover, the assumption $\mathit{\Delta} \in [0, \check{\mathit{\Delta}})$ implies that the effect of changes in the immigration policy on the composition of the voting population in the following period is small, such that effect (iii) roughly corresponds

[45]The exact formula for the threshold $\hat{g}_t$ is provided in Appendix B.3.

to the direct effect of the shock on the citizens' old-age dependency ratio g_t. Thus, the results are primarily driven by the political effect (i) in the short run and by changes in the state g_t in the long run. The findings are illustrated in the following statement.

Proposition 3. *(Effect of population ageing, increasing inequality, and economic depression). There exists $\hat{\Delta} > 0$ such that, if $\Delta \in [0, \hat{\Delta})$, then (i) an increase in longevity and/or (ii) an increase in income inequality and/or (iii) a critical decrease in fertility, and/or (iv) an economic depression translate to (1) a less open immigration policy M_{SS}, a less liberal economic policy L_{SS}, (3) a larger size of government τ_{SS}, and (4) a higher citizens' old-age dependency ratio g_{SS}.*

Proof. See Appendix B.3.

The intuition underpinning this result is simple; it is illustrated in *Figs. 5–6*, in which individual productivity possesses a log-logistic distribution.[46]

This example shows that both population ageing (*Fig. 5*) and rising income inequality (*Fig. 6*) result in a decrease in the median type θ_t^p over Θ. A low-income citizen is less affected by income taxes than a high-income citizen but both types equally benefit equally from public spending and restricted immigration. This suggests that the ideal policy of the former features higher public spending and less immigration than that of the latter. The multidimensional median voter theorem indicates that the citizen featuring the median type over Θ is pivotal. Thus, as the pivotal citizen becomes a less productive individual, the equilibrium policy shifts in favour of elderly and welfare-dependent citizens, penalising, in turn, the young and most productive part of the society.

Regarding the effect of a decrease in ξ, the idea is that a negative productivity shock reduces the total fiscal gains from immigration, leading to a lower preferred immigration level for all citizens.

[46] In this illustrative example, the distribution of the productivity parameter θ_t (conditional on $age = Y$) is log-logistic with c.d.f. $Q(\theta_t; \rho) = \left(1 + (\theta_t/a(\rho))^{-s(\rho)}\right)^{-1}$, where $s(\rho) = \frac{2-\rho}{1-\rho}$ and $a(\rho) = \frac{s(\rho)}{\pi} \sin\left(\frac{\pi}{s(\rho)}\right)$. Thus, the distribution of citizens' types is $F_t(\theta_t; \rho \mid h_t) = \left[g_t + \left(1 + (\theta_t/a(\rho))^{-s(\rho)}\right)^{-1}\right] / (1 + g_t)$. The log-logistic distribution implies that the Gini coefficient of the productivity distribution of young citizens is given by $Gini = 1/s(\rho)$. The proposed parametrization ensures that $E[\theta_t] = 1$ for all $\rho \in [0, 1)$ and that $Gini \in [0, 1/2]$ for all $\rho \in [0, 1)$. In particular, an increase in ρ implies a mean-preserving distribution spread.

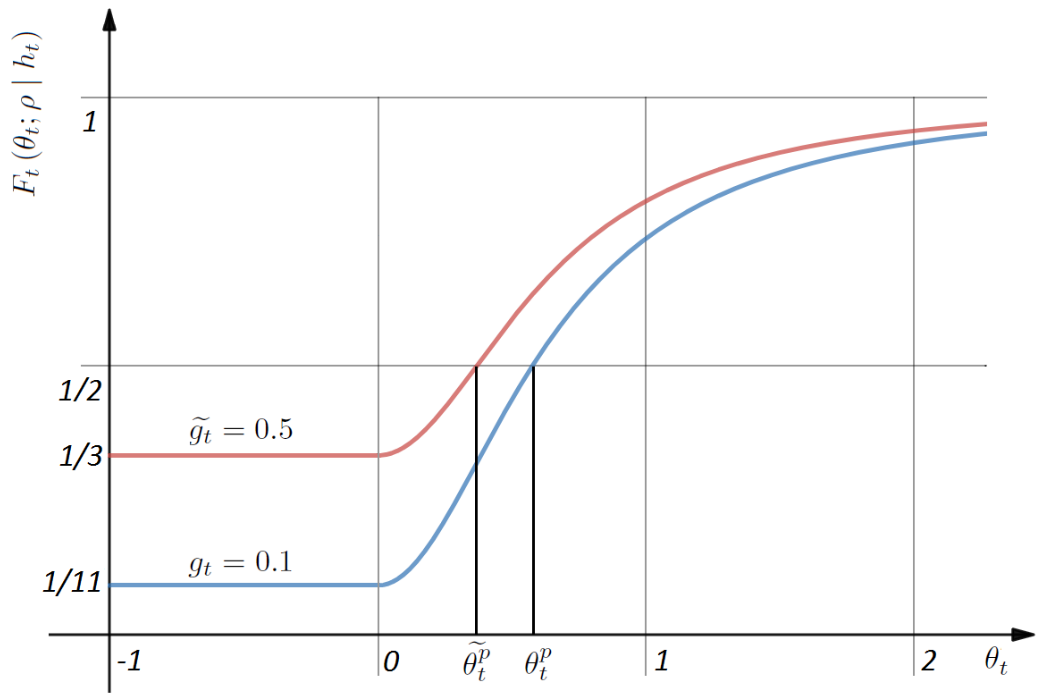

Figure 5: Pivotal citizen's type for different values of the old-age dependency ratio: $g_t = 0.1$ vs. $\widetilde{g}_t = 0.5$ implying $\theta_t^p \simeq 0.37$ and $\widetilde{\theta}_t^p \simeq 0.52$.

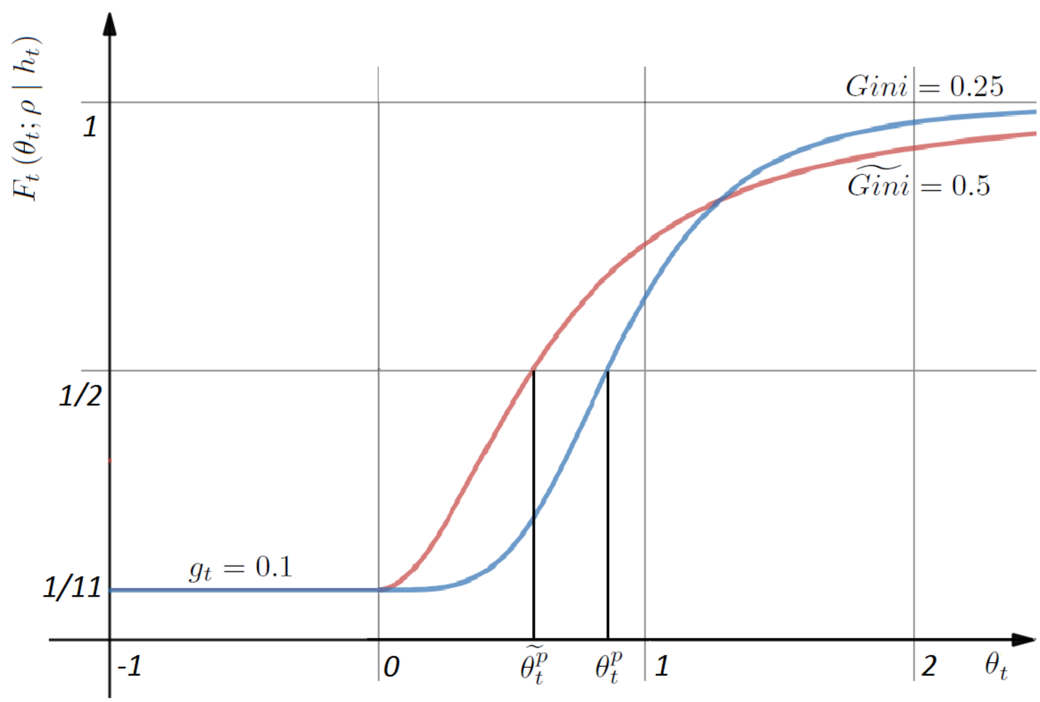

Figure 6: Pivotal citizen's type for different values of income inequality: $Gini = 0.25$ ($\rho = 2/3$) vs. $\widetilde{Gini} = 0.5$ ($\rho = 0$) implying $\theta_t^p \simeq 0.52$ and $\widetilde{\theta}_t^p \simeq 0.86$.

3.5 Equilibrium Party Structure

In this section, we characterize the type of party structures that prevail in any *MPPE*. To do so, we propose an ideological characterization of a party structure $\mathbb{P}_t$ and of each party $P_t^j \in \mathbb{P}_t$, based on the direction of the policy reforms that each party proposes.

Definition 2. A party $P_t^j \in \mathbb{P}_t$ proposing a platform (M_t^j, L_t^j) in period t is said to be:

1. *progressive* if it proposes a more permissive immigration policy and a more interventionist fiscal policy: $M_t^j \geq M_{t-1}^*$ and $L_t^j \leq L_{t-1}^*$

2. *conservative* if it proposes a more restrictive immigration policy and a more liberal fiscal policy: $M_t^j \leq M_{t-1}^*$ and $L_t^j \geq L_{t-1}^*$;

3. *libertarian* if it proposes a more permissive immigration policy and a more liberal fiscal policy: $M_t^j \geq M_{t-1}^*$ and $L_t^j \geq L_{t-1}^*$;

4. *right-wing populist* if it proposes a more restrictive immigration policy and a more interventionist fiscal policy: $M_t^j \leq M_{t-1}^*$ and $L_t^j \leq L_{t-1}^*$.

An ideological position of type 1, 2, 3, or 4 is *strict* if $(M_t^j, L_t^j) \neq (M_{t-1}^*, L_{t-1}^*)$.

Definition 3. A tuple $\langle \mathbb{P}_t, A_t(\mathbb{P}_t) \rangle$ is said to be:[47]

1. *a left vs. right party system* if $\nexists\ P_t^j,\ P_t^k$ in $\mathbb{P}_t$ such that (i) $(M_t^j, L_t^j) \neq (M_t^k, L_t^k)$ and (ii) either $M_t^j \geq M_t^k$ and $L_t^j \geq L_t^k$, or $M_t^j \leq M_t^k$ and $L_t^j \leq L_t^k$;

2. *a populist vs. libertarian party system* if $\nexists\ P_t^j,\ P_t^k$ in $\mathbb{P}_t$ such that (i) $(M_t^j, L_t^j) \neq (M_t^k, L_t^k)$ and (ii) either $M_t^j \geq M_t^k$ and $L_t^j \leq L_t^k$, or $M_t^j \leq M_t^k$ and $L_t^j \geq L_t^k$.

In words, in a left vs. right party system parties can be ordered from the most progressive to the most conservative, while in a populist vs. libertarian party system parties can be ordered from the most libertarian to the most right-wing populist.

Given the above definitions, I can state the following result.

Proposition 4. *(Equilibrium party system). If $\Delta \in [0, \hat{\Delta})$, then (i) in any MPPE, the tuple $\langle \mathbb{P}_{t+r}, A_{t+r}(\mathbb{P}_{t+r}) \rangle$ is a populist vs. libertarian party system in each period $t+r$ for all $r = 1, 2, ...$;*

[47] Note that a party system with $\mathbb{P}_t = \{P_t^j\}$ satisfies both Definition 3.1 and 3.2.

(ii) each member of any right-wing populist party is weakly older and weakly lower-income than any member of any strictly libertarian party; (iii) if a marginal shock of type (a), (b), (c), and/or (d) occurs in period t, then in any period $t+r$ with $r = 0, 1, 2, ...$ the winning party is right-wing populist; in particular, (iv) there is an MPPE such that, in each period $t+r$, the tuple $\langle \mathbb{P}_{t+r}, A_{t+r}(\mathbb{P}_{t+r}), x_{t+r} \rangle$ is a stable two-party structure in which the winning party is right-wing populist and includes all citizens with $\theta^i_{t+r} \le \theta^p_{t+r}$ (i.e., the old and the low-income citizens).

Proof. See Appendix B.4.

Proposition 4(i) states that in the presence of a salient immigration policy dimension, the electoral competition tends to be between libertarian and right-wing populist parties. Moreover, Proposition 4(ii) states that populist parties tend to be coalitions comprised of poor[48] and old citizens.

The power of the right-wing populist parties grows if the economy is hit by a sociodemographic shock, be it an increase in longevity, fall in fertility, or increase in income inequality. In such a case, Proposition 4(iii) implies that the transition to more restrictive immigration policies, a larger public sector and higher taxes (implied by Proposition 2) is driven by the electoral success of this kind of party. In other words, the policy change is the result of an increase in the political power of the coalition of old and low-income citizens relative to that of middle- and upper-class young individuals.

3.6 Welfare Analysis

The findings in Proposition 3 do not necessarily indicate that the predicted policy changes are desirable among society as a whole.

In this section, we present a welfare analysis demonstrating that in ageing societies, a marginal tightening in immigration policy from its equilibrium level is typically unambiguously harmful. I use a social welfare function (SWF) as a measure of the societal well-being. The SWF is a weighted average of the utility of citizens at time t and the expected utility of future generations. Let $\mu_{t+r}(\theta_{t+r})$ denote the Pareto weight assigned to individuals of type θ_{t+r} in period $t+r$.[49]

[48] Specifically, they are those in the ζ_t-percentile of the income distribution, with $\zeta_t = (1 - g_t) \times 50 < 50$.

[49] We do not account for the welfare of current potential immigrants. This allows us to abstract from a full description of their utility function. Nevertheless, if immigration choices are endogenous, any potential immigrant should be weakly better off if able to immigrate, because they still have the choice between remaining in their country of origin or to emigrating to a different country. Thus, whenever a tightening in the immigration policy is harmful to citizens, this result should hold true if we account for the welfare of potential immigrants.

The SWF is constructed as follows:

$$SWF\left((M_t, L_t); \varphi \mid h_t, s_t\right) = \qquad E_t\left[\int_{-1}^{+\infty} u_t\left((M_t, L_t), x_{t+1}^*; \theta_t, \varphi, g_t(h_t)\right) \mu_t(\theta_t) d\theta_t + \right.$$
$$\left. + \sum_{r=1}^{\infty} \int_0^{+\infty} u_{t+r}\left((M_{t+r}^*, L_{t+r}^*), x_{t+r+1}^*; \theta_{t+r}, \varphi \mid h_{t+r}, s_{t+r}\right) \mu_{t+r}(\theta_{t+r}) d\theta_{t+r} \middle| h_t, s_t\right] \tag{9}$$

We study the effect of a marginal change in M_t evaluated at $M_t = M_t^*$ on the above measure of aggregate well-being. The idea underpinning this exercise is simple: if at an equilibrium point at which $M_t^* < \overline{M}$ the marginal effect of an increase in M_t on the SWF is greater than that on the utility of the median voter, there exists a policy (M_t', L_t^*) with $M_t' > M_t^*$ which is welfare-improving.

This means that, in turn, if the immigration policy in equilibrium is such that $M_t^* < M_{t-1}^*$ as a consequence of a marginal demographic shock, the society benefits, *ceteris paribus*, from moving back towards the previous level M_{t-1}^*. In other words, the society is harmed by the change in the immigration policy at the margin. from this, we can state the following result.

Proposition 5. *For any Social Welfare Function $SWF((M_t, L_t); \varphi \mid h_t, s_t)$ that assigns a strictly positive weight to each native individual of working age, there exist thresholds $\breve{g}_t \in [0, 1)$ and $\check{\Delta} > 0$ such that if $g_t \in [\breve{g}_t, 1)$ and $\Delta \in [0, \check{\Delta})$, then a marginal loosening in the immigration policy is welfare-enhancing.*

Proof. See Appendix B.5.

The intuition underpinning this result is as follows. On one hand, the marginal fiscal benefit from immigration for a working-age individual is constant in M_t. On the other hand, the marginal taste cost of immigration tends to zero as M_t approaches $\underline{M}$. The value of θ_t^p tends to 0 as g_t approaches 1 (i.e., as the pivotal citizen approaches zero taxable income and, in turn, is less affected by changes in the income tax caused by immigration).[50]

Thus, as $g_t \to 1$ the equilibrium quota M_t^* tends to $\underline{M}$. This implies that at the equilibrium, the marginal aggregate fiscal gains from immigration for the working-age citizens are very large relative to its marginal social costs due to taste. As a result, provided that the social welfare function assigns a positive—even if small—weight to young citizens, if the citizens' old-age dependency ratio is sufficiently

[50] This is a sensible scenario if one considers a more realistic tax system in contrast to the simple tax schedule described in Section 3. For instance, if the tax system features a personal allowance, as in the UK, the zero taxable income threshold must be adjusted accordingly.

close to 1, a marginal increase in immigration from its equilibrium level always results in higher social welfare.[51]

Proposition 5 suggests that societies characterized by a high old-age dependency ratio are likely to implement excessively restrictive immigration policies. Moreover, it implies that a marginal tightening in the immigration policy caused, for instance, by population ageing is harmful to society. This result is suggestive in the light of the increasingly controversial restrictions to immigration that have been progressively implemented in countries characterized by rapidly ageing populations, such as the UK and Italy.

4 Calibration and Simulated Counterfactuals

The analytical predictions in Section 3 are purely qualitative. As such, they do not provide any insight into the magnitude of the effects. Thus, in this section, we parametrise the model and calibrate it to UK data. We then use the calibrated model to simulate key counterfactuals. While the exact quantitative predictions of this numerical exercise should be viewed as purely illustrative, they suggest that the effect of population ageing and rising inequality on immigration policies may be rather large in magnitude. The results are summarised in this section and extensively presented in the online appendix to this paper.

I use the following utility functions:

$$
\begin{gathered}
U_t^{i,Y}\left(C_t^i, M_t, G_t\right) = C_t^i + \delta_1 \ln\left(G_t\right) - \delta_2 M_t^2 + \beta\lambda\left[C_{t+1}^i + \delta_1 \ln\left(G_{t+1}\right) - \delta_2 M_{t+1}^2\right] \\
U_t^{i,O}\left(C_t^i, M_t, G_t\right) = C_t^i + \delta_1 \ln\left(G_t\right) - \delta_2 M_t^2
\end{gathered}
\tag{10}
$$

for young and old citizens, respectively. We assume that the pre-tax equivalised income of UK households (among non-retired individuals) possesses a Dagum distribution (generalized log-logistic) and we calibrate the parameters to fit the mean, median, and Gini coefficient in the 2017-2018 UK population (Office of National Statistics, 2019). The parameters capturing demographics such as life expectancy at 65 and the fertility rates of natives and immigrants are all consistent with the corresponding values

[51] Conversely, even if g_t is close to zero, a marginal increase in M_t at the equilibrium does not necessarily harm social welfare. Specifically, a threshold $\breve{g}_t \in [0,1]$ such that if $g_t \leq \breve{g}_t$ the society would benefit from a marginally more restrictive immigration policy may not exist for all the possible SWFs that satisfy the conditions stated above and that assign a strictly positive weight to the elderly. Nevertheless, such threshold $\breve{g}_t$ exists for some specific functional forms, such as the utilitarian SWF.

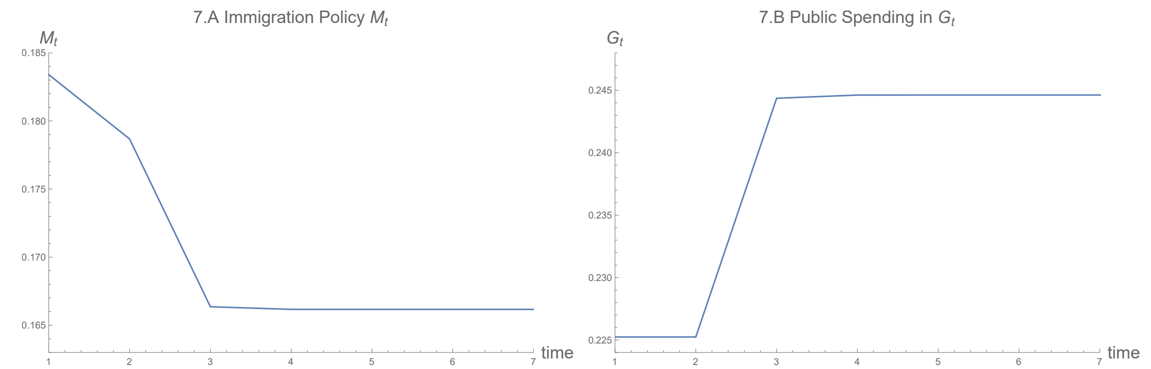

Figure 7: Effect of an increase of 5 years in life expectancy at 65 on the immigration policy M_t (left) and on public spending G_t (right).

of 2017-2018 (ibidem). Lastly, the parameters of the utility function and the pension rate γ are all calibrated using data about public spending in the UK from the HM Treasury's Public Expenditure Statistical Analysis (PESA) 2018 report.

I use the calibrated model to simulate the effects of:

1. a permanent increase in life expectancy at 65 (+5 years)

2. a decrease in the Gini coefficient of equivalized pre-tax income of non-retired households (-10%)

The simulated counterfactuals imply that, in the UK, an increase of 5 years in life expectancy at 65 years old translates to a new steady-state policy featuring $866,768$ less resident immigrants of working age—equal to 11.27% of the foreign-born working-age population in the UK in 2017–2018 (*Fig. 7.A*), and a 8.6% increase in (non-pension) public spending per working-age individual (*Fig. 7.B*). Similarly, a decrease of 10% in income inequality—measured as the Gini coefficient of equivalised pre-tax income of non-retired households—translates to a new policy allowing for $913,800$ (+11.88%) additional working-age resident immigrants (*Fig. 8.A*) and a 9.26% reduction in (non-pension) public spending per individual of working age (*Fig. 8.B*).

It is important to contextualise these results. In the UK, life expectancy at 65 years old has increased by approximately 6.8 years between 1980 and 2018, and the pre-tax equivalised Gini coefficient for non-retired households has risen by 33.2% over the same period (Office for National Statistics, 2019). As already stated, the magnitude of our quantitative predictions should be considered to be merely illustrative. Nevertheless, our results suggest that population ageing and rising inequality in the UK played a substantial role in shaping the increased levels of aversion towards an open immigration policy over the last few decades.

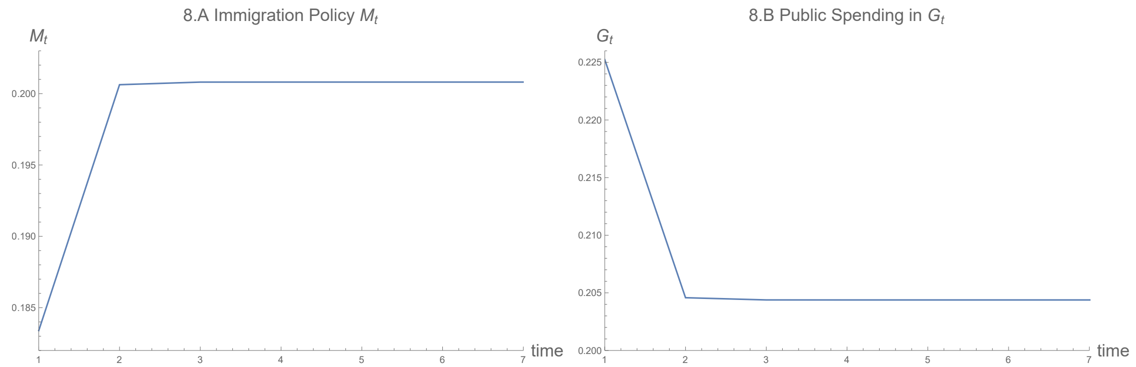

Figure 8: Effect of a 10% decrease in the Gini coefficient of equivalised income of non-retired households on the immigration policy M_t (left) and on public spending G_t (right).

The results of other numerical counterfactual exercises, such as a native fertility rate shock, are described in detail in the online appendix. In Section 6, I use the calibrated model to perform a number of robustness checks on our main results.

5 Empirical Evidence

In this section, we investigate the determinants of British adult residents' attitudes towards immigration and public spending using data from the British Social Attitude Survey (BSA).[52] Specifically, we use its 1995, 2003, 2008, 2011, 2013 and 2017 rounds, which include a specific question about attitudes towards immigration.

This kind of empirical analysis is not novel. Dustmann and Preston (2007), for instance, used earlier rounds of this survey (1983–1990) to quantify how racial and economic factors shape British attitudes towards immigration. The key difference between past research and this study—which is more limited in scope—lies in the goal of the analysis. While we do not claim to prove the existence of a causal relationship, we aim to provide suggestive evidence for four key implications of the theoretical model.

The theoretical model implies that the heterogeneity in attitudes towards an open immigration policy across socioeconomic groups is driven by the varying extent at which groups internalize the positive fiscal effects of immigration. If this key implication is true, then the correlations between age, income, and attitudes towards immigration and public spending, which are extensively documented by existing empirical work, should (i) survive after controlling for non-economic factors, such as cohort

[52]The BSA does not cover the entirety of the United Kingdom because it does not include respondents from Northern Ireland.

effects, education levels, religious beliefs, etc., which are unlikely to be affected by the fiscal policy, and (ii) exhibit the sign implied by the order of types in Θ, which corresponds to testing the following hypotheses:

1. The preferred number of immigrants is negatively correlated with age (*H1*)
2. The preferred number of immigrants is positively correlated with income (*H2*)
3. The preferred level of taxation to finance public spending is positively correlated with age (*H3*)
4. The preferred level of taxation to finance public spending is negatively correlated with income (*H4*).

The next section details the data, methodology, and results of this analysis.

5.1 Data and Methods

The dataset accounts for a total of 20,460 observations.[53] The explanatory variables are respondent age (*RAge*) and household income decile (*HHIncD*).[54]

We control for the highest educational qualification attained by the respondent (*HEdQual*), on a scale from 1 (graduate degree) to 7 (no qualification). Dummy variables capture whether the household includes children (*ChildHh*), the sex of the respondent (*RSex*), if they live in rural areas (*ResPres*), if they are born abroad (*BornAbr*), if they are religious (*Religion*), and if they are unemployed (*Unempl*).[55]The dummy variable *Brexit* corresponds to the year 2017 (i.e., the only included survey round that was conducted after the referendum on EU membership).

The first outcome variable *LessImmigr* captures the respondent's attitude towards further immigration. The question is "Do you think the number of immigrants to Britain nowadays should be increased a lot, increased a little, remain the same as it is, reduced a little or reduced a lot?" The respondent selects a value on a discrete scale from 1 ("increased a lot") to 5 ("reduced a lot").[56] Thus, the variable

[53]Only 13,398 observation include information on attitudes towards immigration; only 17,895 observations include information about attitudes towards public spending financed through taxes.

[54]The use of household income instead of individual income is justified because the effect of taxes on individual consumption levels typically depends on household income. For instance, for a household in which only one member has positive income, the consumption levels of other family members depends on the income tax rate, even if they do not directly pay an income tax.

[55]The 1995 round of the survey does not include information regarding the respondents' country of birth or the presence of children in the household. Thus, data from that round are only used in specification (2) in Table 2.

[56]For the 2017 round of BSA, the question changed to "Once Britain has left the EU, do you think immigration into Britain should be increased, reduced, or stay at more or less the same level as now?" Due to this change, we control for the dummy *Brexit* in specifications (1) - (2) - (3) and exclude the most recent data round (2017) in specification (4).

Variable	Obs	Mean	Std. Dev.	Min	Max
LessImmigr	13,398	4.162636	.9659662	1	5
IncreaseTax	17,895	2.439285	.6045544	1	3
RAge	23,094	50.43306	18.31855	17	99
HHIncD	19,177	5.422016	3.352886	1	10

Table 1: Summary Statistics.

LessImmigr measures the degree of aversion towards open immigration policies. The majority of respondents in all periods exhibit a strong aversion to further immigration.

The second outcome variable *IncreaseTax* measures the respondent's attitude towards public spending financed through taxation. The question is "Suppose the government had to choose between the three options on this card: reduce taxes and spend less on health, education and social benefits, keep taxes and spending on these services at the same level as now, increase taxes and spend more on health, education and social benefits. Which do you think it should choose?" The respondent selects a value on a discrete scale from 1 ("spend less") to 3 ("spend more"). Summary statistics are shown in Table 1.

It is well known that it is not generally possible to separately identify age, cohort, and period effects in linear models (Heckman & Robb, 1985). I address this problem by imposing various restrictions on the nature of cohort and/or period effects, each corresponding to an empirical specification, all of which are detailed in the next section. All the results are very robust across various specifications.

I use a standard ordered logit model because of the discrete and ordered nature of each outcome variable. The outcome variable $LessImmigr$ can take values $j \in \{1,2,3,4,5\}$. A latent variable $LessImmigr^*$ is assumed through:

$$LessImmigr^*_{it} = \beta_1 RAge_{it} + \beta_2 HHIncD + \beta_3 HEdQual + \ldots + \epsilon_{it}$$

The probability of observing the outcome $LessImmigr_{it} = j$ conditional on covariates is:

$$Prob\left(LessImmigr_{it} = j \mid X_{it}\right) = F\left(\alpha_j - LessImmigr^*_{it}\right) - F\left(\alpha_{j-1} - LessImmigr^*_{it}\right)$$

where X_{it} is the vector of explanatory variables and α_{j-1}, α_j are the endogenous thresholds on the value of the latent variable that correspond to a switch from choice $j-1$ to j and from choice j and $j+1$, respectively. The robust standard errors are clustered at regional level.[57]

[57] Clustering for specifications (1) - (3) - (4) is based on a twelve-region partition. For specification (2), which includes data from the 1995 survey round, clustering is based on a six-region partition due to a different classification used prior to 2003.

The same specification is adopted for the second regression, which uses *IncreaseTax* as outcome variable, except that *IncreaseTax* can take values $k \in \{1, 2, 3\}$.

5.2 Determinants of the Preferred Number of Immigrants

Table 2 presents the results of the ordered logit regression with standard errors in parentheses. Table 3 shows the average marginal effects of the regressors of interest with respect to the outcome *LessImmigr* = 5 (i.e., that which corresponds to the strongest hostility towards immigration).

In line with the prediction of the theoretical model, respondent age exhibits a significant positive relationship with the aversion towards immigration, providing support to hypothesis *H1*. Specifically, an additional year of age results in an approximate average increase of 1 percentage point in the probability of outcome *LessImmigr* = 5. Moreover, the parameter on household income decile and the corresponding marginal effect are negative in all specifications and statistically significant in most, meaning that high-income respondents tend to be less averse to immigration relative to low-income respondents. This is consistent with hypothesis *H2*.

Specifications (1) and (2) include time trends and dummies for the respondent's cohort.[58] Note that the negative relationship between age and attitude towards immigration suggested by the model is supported by this analysis even after controlling for the presence of cohort effects. Specifications (3) and (4) includes cohort trends and dummies for the survey year. The coefficient on the dummy *Brexit* is negative and statistically significant in all specification that include this variable.

For illustrative purposes, we simulate the probability of response *LessImmigr* = 5 by an employed, male, UK-born individual in 2017 evaluated at different ages. *Fig. 9* plots the effect of age on the probability of a *LessImmigr = 5* response from a fictitious individual constructed using the estimates in Table 2. Specifically, *Fig. 9.A* illustrates the effect of age for three different cohorts (1906-1915, 1936-1945 and 1986-1995) and shows that more recent cohorts are, on average, more averse to immigration. *Fig. 9.B* plots the effect of the dummy *Brexit* on the same fictitious individual; it shows that attitudes towards immigration have improved in 2017, possibly due to the referendum result.

[58] We group the cohorts using intervals of 10 years (e.g., 1906-1915, 1916-1925, etc.).

Ordered logit with cohort dummies (1) – (2) or year dummies (3) – (4)

VARIABLES	(1) LessImmigr	(2) LessImmigr	(3) LessImmigr	(4) LessImmigr
RAge	0.0459***	0.0452***	0.0449***	0.0415***
	(0.0108)	(0.00898)	(0.00606)	(0.00624)
HHIncD	-0.0233**	-0.0156*	-0.0232**	-0.0127
	(0.00943)	(0.00903)	(0.00918)	(0.00913)
HEdQual	0.221***	0.217***	0.221***	0.243***
	(0.0201)	(0.0137)	(0.0204)	(0.0249)
RAge_2	-0.000157	-9.83e-05	-0.000153**	-9.21e-05
	(0.000117)	(9.93e-05)	(7.67e-05)	(7.12e-05)
RSex	0.0264	0.0595	0.0275	0.0405
	(0.0549)	(0.0375)	(0.0548)	(0.0668)
Unempl	-0.148	-0.164**	-0.137	-0.140
	(0.0931)	(0.0698)	(0.0904)	(0.116)
Religion	-0.000976	-0.0936***	-0.00295	-0.0437
	(0.0356)	(0.0265)	(0.0352)	(0.0434)
ResPres	0.178***		0.178***	0.175***
	(0.0543)		(0.0540)	(0.0563)
BornAbr	-0.813***		-0.817***	-0.927***
	(0.0937)		(0.0935)	(0.119)
ChildHh	-0.179***		-0.189***	-0.182***
	(0.0430)		(0.0435)	(0.0537)
North_En	0.285	0.218	0.285	0.312*
	(0.190)	(0.205)	(0.188)	(0.177)
South_En	0.0939	-0.00859	0.0947	0.161
	(0.194)	(0.262)	(0.193)	(0.185)
year	-0.000373	-0.0436***		
	(0.00916)	(0.00456)		
Brexit	-1.254***		-1.290***	
	(0.0794)		(0.108)	
Cohort10			0.235***	0.245***
			(0.0531)	(0.0684)
Observations	9,407	10,303	9,407	7,136
Pseudo R2	0.0774	0.0368	0.0772	0.0556

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 2: Preferred number of immigrants (BSA 1995-2017).

Marginal Effects on Outcome LessImmigr = 5				
	(1)	(2)	(3)	(4)
	marginal_eff	marginal_eff	marginal_eff	marginal_eff
VARIABLES	y=5	y=5	y=5	y=5
RAge	0.00991***	0.0104***	0.00972***	0.00922***
	(0.00234)	(0.00214)	(0.00132)	(0.00136)
HHIncD	-0.00502**	-0.00360*	-0.00501**	-0.00282
	(0.00204)	(0.00208)	(0.00199)	(0.00203)
RAge_2	-3.40e-05	-2.27e-05	-3.31e-05**	-2.05e-05
	(2.52e-05)	(2.30e-05)	(1.65e-05)	(1.58e-05)
Observations	9,407	10,303	9,407	7,136

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 3: Preferred number of immigrants: marginal effects.

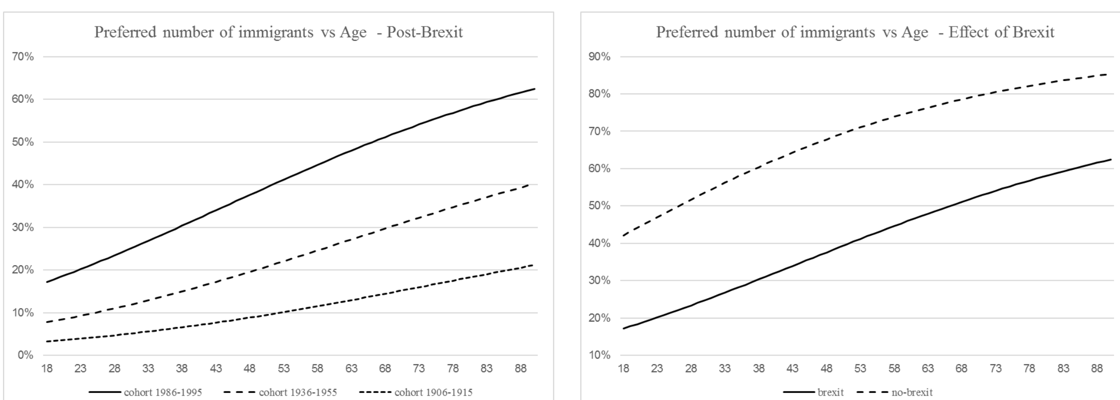

Figure 9: Probability of *LessImmigr* = 5 vs Age: simulated probabilities. Effect of cohort (*9.A*) and Brexit (*9.B*).

5.3 Determinants of the Preferred Level of Taxation and Public Spending

The outcome variable is *IncreaseTax*. The controls of each specification are the same as those for the corresponding specification in Section 5.2. Table 4 shows the results of the ordered logit regression with standard errors in parentheses. Table 5 shows the average marginal effects of the regressors of interest with respect to the outcome $IncreaseTax = 3$ (i.e., that which corresponds to the strongest support for a large government).

The relationship between the outcome variable *IncreaseTax* and both age and income decile are significant; the signs are consistent with hypotheses *H3* and *H4* of the model as well as with the findings in the recent literature. The magnitude of the marginal effects is relatively modest.

This analysis provides strong support for the four key implications of the model regarding voter preferences. The next section discusses the robustness of these findings.

6 Discussion, Robustness, and Extensions

In this section, we discuss the empirical findings from Section 5 and provide analytical and numerical robustness results.

6.1 Empirical Findings: Discussion

The results in Section 5.2 are consistent with those in similar studies that use alternative dataset and analyse other countries or regions. For instance, Dustmann and Preston (2007), Facchini and Mayda (2007) and Card et al. (2011), using data from the British Social Attitude Survey, the International Social Survey Programme and the European Social Survey, respectively, all support hypotheses *H1* and *H2*. Together, these studies constitute substantial empirical evidence in support of the two key implications of the model proposed in this paper.

A much more demanding empirical question is whether population ageing and/or income inequality have an impact on immigration policy and, if so, to what extent this is due to a causal link. An attempts to answer this question has been carried out by Boeri and Brucker (2005) for 15 European countries countries using a variety of data sources and approaches. Their results are mostly in line with the predictions of our model. However, due to the limitations of the existing literature, this remains an open and challenging question for future research.

Ordered Logit with cohort dummies (1) – (2) or year dummies (3) – (4)

VARIABLES	(1) IncreaseTax	(2) IncreaseTax	(3) IncreaseTax	(4) IncreaseTax
RAge	0.0362**	0.0268***	0.0294***	0.0258**
	(0.0161)	(0.00666)	(0.0111)	(0.0124)
HHIncD	-0.0264**	-0.0139	-0.0275**	-0.0332***
	(0.0119)	(0.00960)	(0.0119)	(0.0111)
HEdQual	-0.0608***	-0.0536***	-0.0587***	-0.0511***
	(0.0104)	(0.00833)	(0.0106)	(0.0115)
RAge_2	-0.000480***	-0.000380***	-0.000418***	-0.000438***
	(0.000136)	(5.57e-05)	(7.18e-05)	(7.74e-05)
RSex	0.182***	0.191***	0.182***	0.182***
	(0.0298)	(0.0429)	(0.0299)	(0.0342)
Unempl	0.223**	0.203	0.215**	0.151
	(0.104)	(0.127)	(0.106)	(0.107)
Religion	-0.186***	-0.235***	-0.178***	-0.190***
	(0.0651)	(0.0499)	(0.0663)	(0.0610)
ResPres	-0.0157		-0.0141	0.00116
	(0.0197)		(0.0198)	(0.0254)
BornAbr	-0.370***		-0.367***	-0.199***
	(0.0566)		(0.0571)	(0.0469)
North_En	-0.185***	-0.0996	-0.191***	-0.225***
	(0.0548)	(0.0968)	(0.0558)	(0.0547)
South_En	-0.222***	-0.177***	-0.224***	-0.270***
	(0.0446)	(0.00646)	(0.0437)	(0.0431)
year	-0.0443***	-0.0105*		
	(0.00895)	(0.00589)		
Brexit	1.350***		0.721***	
	(0.0724)		(0.0914)	
Cohort10			-0.227***	-0.272***
			(0.0462)	(0.0614)
Observations	10,380	14,358	10,380	8,087
Pseudo R2	0.0368	0.0162	0.0368	0.0201

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 4: Preferred level of taxation and public spending (BSA 1995-2017).

Marginal Effects on Outcome IncreaseTax = 3

VARIABLES	(1) marginal_eff y=3	(2) marginal_eff y=3	(3) marginal_eff y=3	(4) marginal_eff y=3
RAge	0.00848**	0.00653***	0.00689***	0.00610**
	(0.00378)	(0.00160)	(0.00259)	(0.00290)
HHIncD	-0.00617**	-0.00338	-0.00644**	-0.00785***
	(0.00278)	(0.00234)	(0.00278)	(0.00261)
RAge_2	-0.000112***	-9.24e-05***	-9.78e-05***	-0.000104***
	(3.18e-05)	(1.34e-05)	(1.68e-05)	(1.80e-05)
Observations	10,380	14,358	10,380	8,087

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 5: Preferred level of taxation and public spending: marginal effects.

6.2 Analytical Results: Robustness and Extensions

The main analytical results of this paper are robust to a number of alternative assumptions, some of which are outlined below. A detailed description of these additional results is provided in the online appendix.

1. *Voting rights.* The assumptions on voting rights detailed in Section 3 are consistent with the legal procedures to obtain citizenship in place in several countries including the US, Canada and France. In some countries, however—such as the UK, Japan, Germany and Italy—the legal requirements to obtain citizenship are more demanding. Typically, at least one of the parents must be a citizen in order for a child to obtain citizenship at birth (*ius sanguinis*). It is easy to show that all the results hold true under the alternative assumption that immigrants and their children never obtain voting rights.[59]

2. *Labour supply.* As long as the wage elasticity of labour supply is positive for all workers, all the results carry over if the assumption of perfectly inelastic labour supply is relaxed. This is true, for instance, in the presence of a quadratic utility cost of labour.

3. *Endogenous wages.* As long as the degree of concavity of the production function is sufficiently small, all the results carry over if the assumption of linear production function is relaxed, such that wages decrease with the number of immigrants. This is true, for instance, in an economy with capital and Cobb-Douglas production function under relatively mild restrictions.

4. *Endogenous public investment in education.* Suppose that citizens also vote for the level of uniform public investment in education,[60] which determines the average output level in the following period (e.g., $y^i_{t+1} = \xi(e_t)\theta_{t+1}$, where e_t is the level of per-pupil spending in public education)—all of the main results hold true under this alternative setup. Moreover, a decrease in the fertility rate of the natives also translates into a raise in e^*_t.

5. *Partially funded pension system.* In the baseline model, I assume a pure public pension system financed through general taxation. However, partially funded schemes are becoming increasingly common.[61] All the results hold true if one adds a funded part of the pension system in the form of

[59] In such a case, the analysis is simplified because voters do not have to consider the impact of their current immigration policy choices on the age profile of the voting population in the following periods (i.e. there are no sophisticated effects).

[60] This is a common assumption in political economy models of intergenerational investment in education—for a review of this type of model see Dotti (2019b).

[61] Galasso and Profeta (2004) provide empirical evidence of an increasing size of the funded portion of the pension relative to the state pension in several European countries.

compulsory savings. Moreover, an increase in the funded portion of the pension system relative to the public component results in more restrictive immigration policies and higher public spending levels. The intuition underpinning this result is that a transition towards a private pension system leads to a fall in the cost of the social security system per taxpayer and, therefore, a decrease in the marginal fiscal gains of immigration per native worker.

In this paper, we purposely abstract from some factors that are likely to play a role in shaping voters' immigration policy choices. These aspects deserve further discussion, as they represent a topic for future research. The most important of these aspects are: (i) immigrants' endogenous selection when a generous welfare system attract relatively low-skilled immigrants (Benhabib, 1996; Borjas, 1999), (ii) segmented labour markets when immigrants possess specific labour market skills that are not available to natives workers (Peri & Sparber, 2009) and (iii) non economic drivers, such as cultural and psychological motivations (Brettell & Hollifield, 2007; Card *et al.*, 2011).

6.3 Numerical Results: Robustness

The analytical results presented in Section 3 leave some open questions.

First, the main results in Proposition 3 crucially depend on Δ being sufficiently small in magnitude. This is admittedly a strong restriction, as it contrasts the empirical fact that in several Western countries, immigrants exhibit substantially higher fertility rates than the natives (Bohn & Lopez-Velasco, 2019). Thus, in this section, we use the calibrated model from Section 4 to numerically verify that the main results hold true even with this assumption relaxed.

Second, relaxing such assumption also generates some new implications of the model cannot be characterized analytically.

The numerical exercise shows that, as Δ grows large:

1. The main results in Proposition 2 and Proposition 3 carry over for several parametrizations.

2. The effect of current policy choices on expected future equilibrium outcomes (*sophisticated farsightedness*) may become substantial in magnitude. While this affects the equilibrium policy, it does not typically result in qualitatively different predictions regarding policy response direction to sociodemographic shocks.

3. As Δ grows large, multiplicity may arise,[62] such that the predictions regarding the effects of a

[62] For a large value of Δ the steady-state may not be unique and a shock may cause a transition to a different equilibrium path. Moreover, the conditions in Proposition 1 may be no longer satisfied.

sociodemographic shock depend on the choice of a specific equilibrium path.

4. The speed of convergence to the steady-state after a shock decreases in Δ for all of the parametrizations that generate a unique equilibrium path.

This exercise, as well as other simulation results, are extensively illustrated in the online appendix.

7 Concluding Remarks

This paper investigates the interactions among three key demographic, economic and social processes: ageing, rising inequality and immigration. The aim is to analyse how these processes shape fiscal and immigration policies in democratic countries using the UK as a case study. We detail the effects of increasing life expectancy, decreasing birth rates and rising income inequality on policies choices, the political system, and societal wellbeing.

The key novelty of this analysis is that it allows voters to choose both the immigration and fiscal policy (i.e., not only the number of immigrants but also how society divides costs and benefits of immigration). This choice is shown to generate perceived competition between natives and immigrants over welfare benefits and, in turn, make the most welfare-dependent segments of the voting population—the elderly and the poor—strongly hostile to open immigration policies.

The first finding of this paper is that population ageing and rising income inequality increase the political pressure to restrict the inflow of immigrant workers and inflate the size of government. This finding suggests that the negative effects of population ageing on public finances due to the increasing costs of the social security system may be exacerbated by endogenous political effects. Direct and indirect effects of the ageing phenomenon may affect the long-run fiscal soundness of the public sector.

The second finding is about the political effects of these sociodemographic shocks. We show that ageing and rising inequality can help explain the rise of right-wing populist parties in recent years.

The third finding concerns the welfare effects of the predicted policy changes. We show that the tightening of immigration policy induced by population ageing and rising inequality is generally harmful, though the harm is most severe for young people and future generations.

This analysis delivers a pessimistic prediction regarding the ability of our society to adjust to demographic changes and the consequences of such changes on young generations. Population ageing results in an increase in the power of the elderly to shape public policy according to their needs. As a result, young natives and young potential immigrants pay a price. Young natives must support the financial burden of an increasingly large and long-living elderly population while young potential immigrants are prevented from searching for better employment and life opportunities by excessively restrictive immigration policies.

This worrisome *no country for young people* scenario warrants further research on this topic and constitute a challenge for policy design. It suggests that a key goal of social security reforms in the immediate future should be the promotion of the internalization of the positive fiscal effects of immigration among elderly and low-income citizens. This could be achieved, for instance, by linking the generosity of the social security system to the expected future old-age dependency ratio of the native population. Reforms in this vein have been attempted in in several European countries over the last two decades, such as Finland in 2005 and Italy in 2010.

Appendix

A Equilibrium Concept and Properties

Appendix A includes formal descriptions of the equilibrium concept and of the two key properties of citizens' preferences.

A.1 Markov-Perfect Party Equilibrium

Formally, A *collective strategy* s is a function that, for all periods t and each history $h_t \in H_t$, selects a tuple $\langle \mathbb{P}_t, A_t(\mathbb{P}_t), x_t^* \rangle$. Let S be the set of all possible collective strategies and $S_t(h_t)$ denote the set of *continuation strategies* with typical element s_t (i.e., the set of all collective strategies in the subgame starting after h_t). We define the *objective function conditional on history h_t and strategy s_t* of a citizen of type $\theta_t \in \Theta$, denoted by v_t, as:

$$v_t\left(x; \theta_t, \varphi \mid h_t, s_t\right) \equiv E_t\left[u_t\left(x, x_{t+1}^*; \theta_t, \varphi, g_t(h_t)\right) \mid s_t\right] \tag{11}$$

for $s_t \in S_t(h_t)$ and $x \in X$.[63] We assume that the default policy x^0 satisfies $v_t\left(x^0; \theta_t, \varphi \mid h_t, s_t\right) = -\infty$ for all $\theta_t \in \Theta$ and all $\varphi \in \Phi$.[64]

Given the definitions of the set of voter types Θ, we define, for each period t, the following concepts:

(1) A *party structure* $\mathbb{P}_t$ is a finite partition of the set of voter types Θ.

(2) A *platform profile* $A_t(\mathbb{P}_t)$ is a subset of X such that each element a_t^j is the a platform (if any) proposed by at least one party $P_t^j \in \mathbb{P}_t$: $A_t(\mathbb{P}_t) = \left\{x \in X \mid a_t^j = x \ for\ some\ P_t^j \in \mathbb{P}_t\right\}$.

(3) The *set of ideal policies* for a citizen i of type θ_t^i, with typical element x_t^i, is the set of most-preferred policies in X by such a citizen:

$$I\left(\theta_t^i \mid h_t, s_t\right) := \left\{x \in X \mid v_t\left(x; \theta_t^i, \varphi \mid h_t, s_t\right) \geq v_t\left(x'; \theta_t^i, \varphi \mid h_t, s_t\right)\ \forall x' \in X\right\}$$

(4) The *majority core* $K\left(A_t \mid h_t, s_t\right)$ is the set of platform in A_t that are in the core of the majority

[63] In line with Maskin and Tirole, one can also define the *expected utility conditional on history h_t* of a citizen of type $\theta_t \in \Theta$, denoted by V_t, as: $V_t\left(s_t; \theta_t, \varphi \mid h_t\right) \equiv E_t\left[u_t\left(x_t^*, x_{t+1}^*; \theta_t, \varphi, g_t(h_t)\right) \mid s_t\right]$.

[64] This assumption is common in models of elections, such as Levy (2004, 2005) and Dotti (2020). It is easy to show that all the results hold true if one assumes that the default policy is the status quo x_{t-1}^*, but in such case one has to add x_{t-1}^* as a second element of the state space, i.e. $s_t = \{g_t, x_{t-1}^*\}$. Thus, the assumption in the present paper simplifies the notation.

voting game given voter preferences v_t and history h_t:

$$K\left(A_t \mid h_t, s_t\right) := \left\{x \in A_t \,\middle|\, \int \mathbf{1}\left[v_t\left(x;\theta_t,\varphi \mid h_t, s_t\right) \geq v_t\left(a_t;\theta_t,\varphi \mid h_t, s_t\right)\right] f_t\left(\theta_t \mid h_t\right) d\theta \geq 0.5 \; \forall a_t \in A_t\right\}.$$

Definition A.1. *(Stable political outcome).* (i) A *political outcome* in period t given history h_t and continuation strategy s_t is a tuple $\langle \mathbb{P}_t, A_t(\mathbb{P}_t), x_t^* \rangle$ that satisfies the following conditions:

1. *Majority Rule (MR).* The social choice x_t^* is such that: (a) if $K\left(A_t(\mathbb{P}_t) \mid h_t, s_t\right) \neq \emptyset$, then $x_t^* \in K\left(A_t(\mathbb{P}_t) \mid h_t, s_t\right)$; (b) otherwise, $x_t = x^0$.

2. *Citizen-candidates (CC).* A party platform $a_t^j \neq \emptyset$ is such that $a_t^j \in A_t(\mathbb{P}_t)$ only if $a_t^j \in I\left(\theta_t^i \mid h_t, s_t\right)$ for some $\theta_t^i \in P_t^j$.

3. *Partisan Membership (PM).* For each $P_t^j \in \mathbb{P}_t$, it is true that $\theta_t^i \in P_t^j$ (a) if $v_t\left(a_t^j;\theta_t,\varphi \mid h_t, s_t\right) > v_t\left(a_t^k;\theta_t,\varphi \mid h_t, s_t\right)$ for all $a_t^k \in A_t(\mathbb{P}_t)\backslash\{a_t^j\}$, and (b) only if $v_t\left(a_t^j;\theta_t,\varphi \mid h_t, s_t\right) \geq v_t\left(a_t^k;\theta_t,\varphi \mid h_t, s_t\right)$ for all $a_t^k \in A_t(\mathbb{P}_t)$.

(ii) A political outcome $\langle \mathbb{P}_t, A_t(\mathbb{P}_t), x_t^* \rangle$ is *core stable (CS)* given history h_t—and is referred to as a *stable political outcome (SPO)*—if and only if $\nexists \tilde{P}_t^j \subseteq \Theta$ such that for any political outcome $\langle \mathbb{P}_t', A_t(\mathbb{P}_t'), x_t' \rangle$ that satisfies $\tilde{P}_t^j \in \mathbb{P}_t'$ and $\tilde{a}_t^j \in A\left(\mathbb{P}_t'\right)$ one gets $x_t^* \notin K\left(A\left(\mathbb{P}_t'\right) \mid h_t, s_t\right)$ and $v_t\left(x_t';\theta_t,\varphi \mid h_t, s_t\right) > v_t\left(x_t^*;\theta_t,\varphi \mid h_t, s_t\right)$ for all types $\theta_t \in \tilde{P}_t^j$.

The last part of this condition simply states that the new policy outcome x_t' must be able to strictly defeat the original policy outcome x_t^* in any political outcome that may follow the deviation in which x_t^* is proposed by some candidate. This condition is a tie-break rule for the case in which the ideal policy of the pivotal citizen is not unique. As such, it is irrelevant for all the predictions in proposition 1-2-3-4.

Definition A.2. A *Markov-perfect party equilibrium (MPPE)* is a collective strategy $s^* \in S$ that satisfies the following conditions:

1. *Subgame Perfection.* The strategy s forms a *SPO* after any history h_t in each period t (i.e., for all t and all $h_t \in H_t$ the tuple $\langle \mathbb{P}_t, A_t(\mathbb{P}_t), x_t^* \rangle$ satisfies *MR*, *CC*, *PM*, *CS*);

2. *Markovian strategies.* For all t and any $h_t', h_t'' \in H_t$, $g_t(h_t') = g_t(h_t'')$ implies $s_t^*(h_t') = s_t^*(h_t'')$.

Definition A.2 formalizes some simple insights. First, condition (1) restricts the attention to collective strategies that satisfy subgame perfection (i.e., it rules out empty or incredible threats). Second, condition (2) simply states that the collective choice in each period is the outcome of a political process in which the citizens only play collective Markovian strategies. This assumption implies that political outcomes, and particularly the equilibrium policy outcome x_t^*, are solely functions of the payoff-relevant information contained in the history h_t, which is fully summarized by the citizens' old-age dependency ratio g_t.[65]

A.2 Quasisupermodularity and Strict Single Crossing Property

Following Milgrom and Shannon (1994), we define two desirable properties for the conditional objective function v_t.

Definition A.3. The function v_t in period t for given history h_t satisfies:

1. *Quasisupermodularity (QSM)* in (M_t, L_t) if, for any two $(M_t', L_t'), (M_t'', L_t'') \in X$, one gets:

$$\begin{aligned} & v_t\left((M_t', L_t'); \theta_t, \varphi \mid h_t, s_t\right) - v_t\left((M_t', L_t') \wedge (M_t'', L_t''); \theta_t, \varphi \mid h_t, s_t\right) \geq 0 \\ \rightarrow\ & v_t\left((M_t', L_t') \vee (M_t'', L_t''); \theta_t, \varphi \mid h_t, s_t\right) - v_t\left((M_t'', L_t''); \theta_t, \varphi \mid h_t, s_t\right) \geq 0; \end{aligned} \tag{12}$$

2. *Strict single crossing (SSC)* in $(M_t, L_t; \theta_t)$ if, for any two $(M_t', L_t'), (M_t'', L_t'') \in X'$ with $(M_t'', L_t'') \geq (M_t', L_t')$ and $(M_t'', L_t'') \neq (M_t', L_t')$ and any two $\bar{\theta}_t, \underline{\theta}_t \in \Theta$ with $\bar{\theta}_t > \underline{\theta}_t$, one gets:

$$\begin{aligned} & v_t\left((M_t'', L_t''); \underline{\theta}_t, \varphi \mid h_t, s_t\right) - v_t\left((M_t', L_t'); \underline{\theta}_t, \varphi \mid h_t, s_t\right) \geq 0 \\ \rightarrow\ & v_t\left((M_t'', L_t''); \bar{\theta}_t, \varphi \mid h_t, s_t\right) - v_t\left((M_t', L_t'); \bar{\theta}_t, \varphi \mid h_t, s_t\right) > 0. \end{aligned} \tag{13}$$

QSM and *SSC* over the complete sublattice $(X, \leq)$ are desirable properties because they imply that the set of ideal policies $I\left(\theta_t \mid h_t, s_t\right)$ is monotonic nondecreasing in θ_t over X by theorem 4 in Milgrom and Shannon (1994).

B Proofs

Appendix B includes the proofs to the results of the paper.

[65] The assumptions of perfect foresight and Markovian strategies imply that each citizen's objective function conditional on h_t in each period t satisfies $v_t\left(\cdot; \theta_t, \varphi \mid h_t', s_t\right) = v_t\left(\cdot; \theta_t, \varphi \mid h_t'', s_t\right)$ for all histories h_t', h_t'' such that $g_t(h_t') = g_t(h_t'')$, for all $\theta_t \in \Theta$ and $\varphi \in \Phi$.

B.1 Equilibrium Existence

Preliminaries. Define parameter $\psi \equiv \gamma - l$. First, notice that the structure of the pension system implies $p_{t+1}(\xi\theta_t, g_t) = \xi\tilde{p}_t(\xi\theta_t)/g_t$ for some increasing function $\tilde{p}$ that satisfies $\int \tilde{p}_t(\xi\theta_t)\, q(\theta_t) d\theta_t = \gamma\xi$ for any possible value of ξ. Thus, it must have form $\tilde{p}_t(\xi\theta_t) = \xi\hat{p}_t(\theta_t)$ with $\int \hat{p}_t(\theta_t)\, q(\theta_t) d\theta_t = \gamma$. Using formula (8) and (4), the objective function v_t of a young citizen (i.e. $\theta_t \geq 0$) writes:

$$\begin{aligned} v_t((M_t, L_t); \theta_t, \varphi \mid h_t, s_t) = \quad & \underbrace{\gamma\xi\theta_t + \psi\xi M_t\theta_t + \xi\theta_t L_t + b(\overline{G} - L_t) + c(M_t) + \beta\tilde{p}_t(\xi\theta_t)\,\bar{\sigma}_t}_{A((M_t,L_t);\theta_t,\varphi,g_t(h_t))} + \\ & +\beta\lambda \underbrace{E\left[d\left(\overline{G} - L_{t+1}\right) + c(M_{t+1}) \mid (M_t, \overline{G} - L_t), g_t(h_t)\right]}_{B_{t+1}(M_t,\varphi,g_t(h_t))} \end{aligned} \tag{14}$$

Notice that given M_t and φ the object $\bar{\sigma}_t$ is known, i.e. $\bar{\sigma}_t = \sigma^m - \Delta(1 - M_t)$. Also notice that $B_{t+1}(M_t, \varphi, g_t(h_t))$ is independent of θ_t at time t. Using formula (6), the objective function v_t of an old citizen (i.e. $\theta_t = -1$) writes:

$$v_t((M_t, L_t); -1, \varphi \mid h_t, s_t) = d\left(\overline{G} - L_t\right) + c(M_t) \tag{15}$$

Using formulas (14) and (15) I can state the following results.

Lemma 1. *The function v_t satisfies (i) QSM in (M_t, L_t) and (ii) SSC in $(M_t, L_t; \theta_t)$ for all $\varphi \in \Phi$ and after any history h_t.*

Proof. Part (i). *QSM* in (M_t, L_t). Consider any two elements $(M_t'', L_t''), (M_t', L_t') \in X'$. A sufficient condition for *QSM* is Supermodularity (see Milgrom and Shannon 1994). Thus, for condition (12) to hold true it is sufficient that:

$$\begin{aligned} v_t((M_t'', L_t'') \vee (M_t', L_t'); \theta_t, \varphi \mid h_t, s_t) - v_t((M_t'', L_t''); \theta_t, \varphi \mid h_t, s_t) \geq \\ v_t((M_t'', L_t''); \theta_t, \varphi \mid h_t, s_t) - v_t((M_t'', L_t'') \wedge (M_t', L_t'); \theta_t, \varphi \mid h_t, s_t) \end{aligned} \tag{16}$$

after any history h_t. Let $\check{M}_t = \max\{M_t'', M_t'\}$ and $\hat{M}_t = \min\{M_t'', M_t'\}$, $\check{L}_t = \max\{L_t'', L_t'\}$ and $\hat{L}_t = \min\{L_t'', L_t'\}$, such that $(\check{M}_t, \check{L}_t) = (M_t'', L_t'') \vee (M_t', L_t')$ and $(\hat{M}_t, \hat{L}_t) = (M_t'', L_t'') \wedge (M_t', L_t')$.

Using formula (14), for young citizens the condition above can be written as:

$$
\begin{gathered}
\left(\xi\psi_t\theta_t^i+\beta\tilde{p}_t\left(\xi\theta_t^i\right)\mathit{\Delta}\right)\left(\check{M}_t-M_t''-M_t'+\hat{M}_t\right)+c(\check{M}_t)-c(M_t'')-c(M_t')+c(\hat{M}_t)+\\
+\beta\lambda\left[B_{t+1}\left(\check{M}_t,\varphi,g_t(h_t)\right)-B_{t+1}\left(M_t'',\varphi,g_t(h_t)\right)-B_{t+1}\left(M_t',\varphi,g_t(h_t)\right)+B_{t+1}\left(\hat{M}_t,\varphi,g_t(h_t)\right)\right]+\\
+b(\overline{G}-\check{L}_t)-b(\overline{G}-L_t'')-b(\overline{G}-L_t')+b(\overline{G}-\hat{L}_t)+\xi\theta_t^i\left(\check{L}_t-L_t''-L_t'+\hat{L}_t\right)\geq 0
\end{gathered}
\tag{17}
$$

Firstly, either $\check{M}_t = M_t''$ and $\hat{M}_t = M_t'$, or $\check{M}_t = M_t'$ and $\hat{M}_t = M_t''$ (a). Secondly, either $\check{L}_t = L_t''$ and $\hat{L}_t = L_t'$, or $\check{L}_t = L_t'$ and $\hat{L}_t = L_t''$ (b). Then using results (a) and (b) into formula (17) we get that the left-hand side of (17) always equals zero, which implies that condition (16) is always satisfied for any $\theta_t \geq 0$.

For old citizens, using formula (15) the condition in (16) rewrites:

$$
d(\overline{G}-\check{L}_t)-d(\overline{G}-L_t'')-d(\overline{G}-L_t')+-d(\overline{G}-\hat{L}_t)+c(\check{M}_t)-c(M_t'')-c(M_t')+c(M_t')\geq 0 \tag{18}
$$

Again, using the fact that either either $\check{M}_t = M_t''$ and $\hat{M}_t = M_t'$, or $\check{M}_t = M_t'$ and $\hat{M}_t = M_t''$, and that either $\check{L}_t = L_t''$ and $\hat{L}_t = L_t'$, or $\check{L}_t = L_t'$ and $\hat{L}_t = L_t''$, we get that the left-hand side of (18) equals zero, which implies that condition (16) is also always satisfied for $\theta_t = -1$. Thus, condition (16) is satisfied for all possible types $\theta_t \in \Theta$, which implies that v_t satisfies *QSM* in (M_t, L_t).

Part (ii). *SSC* in $(M_t, L_t; \theta_t)$. I need to show that for any $(M_t'', L_t'') \geq (M_t', L_t')$ in X' with $(M_t'', L_t'') \neq (M_t', L_t')$ and any $\bar{\theta}_t > \underline{\theta}_t$ in Θ the condition in (13) holds true.

First I compare any types of two young citizens, i.e. any two $\bar{\theta}_t > \underline{\theta}_t \geq 0$. A sufficient conditions for (13) to hold true for any two $\bar{\theta}_t > \underline{\theta}_t \geq 0$ is the following.

$$
\begin{array}{c}
v_t\left((M_t'',L_t'');\bar{\theta}_t,\varphi\mid h_t,s_t\right)-v_t\left((M_t',L_t');\bar{\theta}_t,\varphi\mid h_t,s_t\right)> \\
v_t\left((M_t'',L_t'');\underline{\theta}_t,\varphi\mid h_t,s_t\right)-v_t\left((M_t',L_t');\underline{\theta}_t,\varphi\mid h_t,s_t\right)
\end{array},
\tag{19}
$$

which corresponds to the definition of strictly increasing differences in $(M_t, L_t; \theta_t)$ over $\{\bar{\theta}_t, \underline{\theta}_t\}$. Use the formula (14), and notice that $\left(M_{t+1}^*(g_{t+1}), L_{t+1}^*(g_{t+1})\right)$ is independent of θ_t because each type possesses zero probability mass. Then, using formula (14) into condition (19), the latter writes:

$$
\left(\bar{\theta}_t-\underline{\theta}_t\right)\xi\left[\psi_t\left(M_t''-M_t'\right)+\left(L_t''-L_t'\right)\right]+\beta\left[\tilde{p}_t\left(\xi\bar{\theta}_t\right)-\tilde{p}_t\left(\xi\underline{\theta}_t\right)\right]\mathit{\Delta}\left(M_t''-M_t'\right)>0 \tag{20}
$$

which is always true under the assumptions $\psi > 0$ and $\Delta \geq 0$.

Secondly, I compare each type of young citizen with $\bar{\theta}_t \geq 0$ to each old citizen with $\underline{\theta}_t = -1$. For any old individual, using formula (15) I get:

$$\begin{aligned} &v_t\left((M_t'', L_t''); -1, \varphi \mid h_t, s_t\right) - v_t\left((M_t', L_t'); -1, \varphi \mid h_t, s_t\right) = \\ &= d(\overline{G} - L_t'') - d(\overline{G} - L_t') + c(M_t'') - c(M_t') < 0 \end{aligned} \tag{21}$$

where the value of (21) is strictly negative because by assumption d is strictly increasing and c is strictly decreasing for $M_t > \underline{M}$. Thus, condition (13) is always trivially satisfied for any $\bar{\theta}_t, \underline{\theta}_t$ in Θ such that $\bar{\theta}_t \geq 0$ and $\underline{\theta}_t = -1$, because the condition $v_t\left((M_t'', L_t''); -1, \varphi \mid h_t, s_t\right) - v_t\left((M_t', L_t'); -1, \varphi \mid h_t, s_t\right) \geq 0$ is never true. Notice that the fact that (21) is always negative also implies that the corresponding alternative condition for *SSC*: $v_t\left((M_t'', L_t''); \bar{\theta}_t, \varphi \mid h_t, s_t\right) - v_t\left((M_t'', L_t''); \bar{\theta}_t, \varphi \mid h_t, s_t\right) \leq 0$ $\rightarrow v_t\left((M_t'', L_t''); -1, \varphi \mid h_t, s_t\right) - v_t\left((M_t', L_t'); -1, \varphi \mid h_t, s_t\right) < 0$ is also always trivially satisfied, given that the *only if* part of such condition is always true. Lastly, because condition (13) is satisfied for all $\bar{\theta}_t, \underline{\theta}_t$ in Θ, then v_t satisfies *SSC* in $(M_t, L_t; \theta_t)$. Q.E.D.

Proposition 1. *(i) A stationary MPPE always exists. (ii) In any MPPE the policy outcome x^*_{t+r} for $r = 0, 1, 2, ...$ is an ideal policy of the unique pivotal citizen θ^p_{t+r}. (iii) The pivotal citizen's type θ^p_{t+r} is weakly decreasing in g_{t+r}. (iv) There exists $\hat{\Delta} > 0$, such that if $\Delta \in [0, \hat{\Delta})$, then the equilibrium policy outcome x^*_t in each period t is unique given state g_t.*

Proof. Part (i). Suppose a *MPPE* does not exists. Construct a sequence of tuples $\left\{\left\langle \mathbb{P}_{t+r}, A_{t+r}(\mathbb{P}_{t+r}), x^p_{t+r}\right\rangle\right\}_{r=0}^{\infty}$ as follows. In each period $t+r$ for $r = 0, 1, 2, ...$ construct the partition $\mathbb{P}_{t+r} = \left\{P^1_{t+r}\right\}$ and the platform profile $A_{t+r}(\mathbb{P}_{t+r}) = \left\{x^p_{t+r}\right\}$ for some platform $x^p_{t+r} \in I\left(\theta^p_{t+r} \mid h_{t+r}, s_{t+r}\right)$. First, notice that condition *(MR)* in the definition of *MPPE* is trivially satisfied because there is a unique platform in $A_{t+r}(\mathbb{P}_{t+r})$. Secondly, condition *(CC)* is also trivially satisfied because $x^p_{t+r} \in I\left(\theta^p_{t+r} \mid h_{t+r}, s_{t+r}\right)$. Condition *(PM)* is trivially satisfied by $\mathbb{P}_{t+r} = \left\{P^1_{t+r}\right\}$. Lastly, suppose that *(CS)* does not hold. Consider a deviation a political outcome $\left\langle \mathbb{P}'_{t+r}, A_{t+r}(\mathbb{P}'_{t+r}), x'_{t+r}\right\rangle$ that satisfies $\tilde{P}^j_{t+r} \in \mathbb{P}'_{t+r}$ such that all the types θ^j_{t+r} in $\tilde{P}_{t+r} \in \mathbb{P}'_{t+r}$ are made strictly better off relatively to outcome x^p_{t+r}. Firstly, $\theta^p_{t+r} \notin \tilde{P}^j_{t+r}$ because $x^p_{t+r} \in I\left(\theta^p_{t+r} \mid h_{t+r}, s_{t+r}\right)$, thus such type of citizen cannot be made strictly better off. In turn, this implies that one of the possible political outcomes that satisfy $\tilde{P}^j_{t+r} \in \mathbb{P}'_{t+r}$ and $\tilde{a}^j_{t+r} \in A_{t+s}(\mathbb{P}'_{t+r})$ is such that $\left\{\theta^p_{t+r}\right\} \in \mathbb{P}'_{t+r}$ and $x^p_{t+r} \in A\left(\mathbb{P}'_{t+r}\right)$. In such a political outcome, condition *(CC)* and *(CS)* imply that there must exist at least one citizen with $\theta^i_{t+r} \neq \theta^p_{t+r}$ that possesses in his/her set of ideal policies an element $x^i_{t+r} \neq x^p_{t+r}$ such that x^p_{t+r} cannot defeat x^i_{t+r} under the majority rule. Say $x^i_{t+r} \in I\left(\theta^i_{t+r} \mid h_{t+r}, s_{t+r}\right)$ strictly defeats x^p_{t+r}. Recall that Lemma 1 implies that v_{t+r} satisfies (i) *QSM* in (x_{t+r}) and (ii) *SSC* in $(x_{t+r}; \theta_{t+r})$. There are two possible cases.

Case 1. $x^i_{t+r} \geq x^p_{t+r}$ $(x^i_{t+r} \leq x^p_{t+r})$ and $x^i_{t+r} \neq x^p_{t+r}$. Optimality implies $v_{t+r}\left(x^p_{t+r}; \theta^p_{t+r}, \varphi \mid h_{t+r}, s_{t+r}\right) \geq v_{t+r}\left(x^i_{t+r}; \theta^p_{t+r}, \varphi \mid h_{t+r}, s_{t+r}\right)$. *SSC* implies $v_{t+r}\left(x^p_{t+r}; \theta_{t+r}, \varphi \mid h_{t+r}, s_{t+r}\right) > v_{t+r}\left(x^i_{t+r}; \theta_{t+r}, \varphi \mid h_{t+r}, s_{t+r}\right)$ for all $\theta_{t+r} < \theta^p_{t+r}$ $(\theta_{t+r} > \theta^p_{t+r})$. Because θ^p_{t+r} is the median type, the citizens with $\theta_{t+r} \leq \theta^p_{t+r}$ $(\theta_{t+r} \geq \theta^p_{t+r})$ represent at least half of the voting population. Thus, $x^p_{t+r} \in K\left(A_{t+r}(\mathbb{P}'_{t+r}) \mid h_{t+r}, s_{t+r}\right)$, which leads to a contradiction. This results also imply that $x^i_{t+r} \in K\left(A_{t+r}(\mathbb{P}'_{t+r}) \mid h_{t+r}, s_{t+r}\right) \in K\left(A_{t+r}(\mathbb{P}'_{t+r}) \mid h_{t+r}, s_{t+r}\right)$ only if $x^i_{t+r} \in I\left(\theta^p_{t+r} \mid h_{t+r}, s_{t+r}\right)$.

Case 2. $x^i_{t+r} \ngeq x^p_{t+r}$ and $x^i_{t+r} \nleq x^p_{t+r}$. Case 2.a $\theta^i_{t+r} > \theta^p_{t+r}$. Because X' is a complete lattice, $(x^i_{t+r} \vee x^p_{t+r}), (x^i_{t+r} \vee x^p_{t+r}) \in X'$ (see Milgrom and Shannon, 1994). Optimality implies $v_{t+r}\left(x^i_{t+r}; \theta^i_{t+r}, \varphi \mid h_{t+r}, s_{t+r}\right) \geq v_{t+r}\left(x^i_{t+r} \vee x^p_{t+r}; \theta^i_{t+r}, \varphi \mid h_{t+r}, s_{t+r}\right)$. *QSM* implies $v_{t+r}\left(x^i_{t+r} \wedge x^p_{t+r}; \theta^i_{t+r}, \varphi \mid h_{t+r}, s_{t+r}\right) \geq v_{t+r}\left(x^p_{t+r}; \theta^i_{t+r}, \varphi \mid h_{t+r}, s_{t+r}\right)$. *SSC* and $x^i_{t+r} \neq x^p_{t+r}$ imply $v_{t+r}\left(x^i_{t+r} \wedge x^p_{t+r}; \theta^p_{t+r}, \varphi \mid h_{t+r}, s_{t+r}\right) > v_{t+r}\left(x^p_{t+r}; \theta^p_{t+r}, \varphi \mid h_{t+r}, s_{t+r}\right)$, which implies in turn $x^p_{t+r} \notin I\left(v_{t+r}, \theta^p_{t+r} \mid h_{t+r}, s_{t+r}\right)$, which leads to a contradiction. Case 2.b $\theta^i_{t+r} < \theta^p_{t+r}$. Similarly to 2.a, optimality implies $v_{t+r}\left(x^i_{t+r}; \theta^i_{t+r}, \varphi \mid h_{t+r}, s_{t+r}\right) \geq v_{t+r}\left(x^i_{t+r} \wedge x^p_{t+r}; \theta^i_{t+r}, \varphi \mid h_{t+r}, s_{t+r}\right)$. *QSM* implies $v_{t+r}\left(x^i_{t+r} \vee x^p_{t+r}; \theta^i_{t+r}, \varphi \mid h_{t+r}, s_{t+r}\right) \geq v_{t+r}\left(x^p_{t+r}; \theta^i_{t+r}, \varphi \mid h_{t+r}, s_{t+r}\right)$. *SSC* and $x^i_{t+r} \neq x^p_{t+r}$ imply $v_{t+r}\left(x^i_{t+r} \vee x^p_{t+r}; \theta^p_{t+r}, \varphi \mid h_{t+r}, s_{t+r}\right) > v_{t+r}\left(x^p_{t+r}; \theta^p_{t+r}, \varphi \mid h_{t+r}, s_{t+r}\right)$, which implies in turn $x^p_{t+r} \notin I\left(v_{t+r}, \theta^p_{t+r} \mid h_{t+r}, s_{t+r}\right)$, which leads to a contradiction.

Thus, in each period $t+r$ for $r = 0, 1, 2, ...$ the tuple $\left\langle \left\{P^1_{t+r}\right\}, \left\{x^p_{t+r}\right\}), x^p_{t+r}\right\rangle$ satisfies the four conditions for *SPO* in Definition A.1 given v_t, h_t. Moreover, the collective strategy in each period $t+r$ is Markovian. In turn, this implies that a collective strategy s_t such that a tuple in the form $\left\langle \left\{P^1_{t+r}\right\}, \left\{x^p_{t+r}\right\}), x^p_{t+r}\right\rangle$ is played after any history h_t is a *MPPE*. This represents a contradiction. For stationarity, it is sufficient to notice that for $T \to \infty$ under the Markov assumption the dynamic problem is identical in each period t at given state g_t. Thus, if a strategy s_t is a *MPPE* in period t given state g_t, then $s_{t+r} = s_t$ is a *MPPE* in period $t+r$ given state $g_{t+r} = g_t$.

Part (ii) Suppose there is a *MPPE* s^*_t such that $x^*_{t+r} \notin I\left(\theta^p_{t+r} \mid h_{t+r}, s_{t+r}\right)$ for some $r = 0, 1, 2, ...$ This implies that $v_{t+r}\left(x^p_{t+r}; \theta^p_{t+r}, \varphi \mid h_{t+r}, s_{t+r}\right) > v_{t+r}\left(x^*_{t+r}; \theta^p_{t+r}, \varphi \mid h_{t+r}, s_{t+r}\right)$. Condition *(CS)* in Definition A.1 implies that any partition $\mathbb{P}'_{t+r}$ such that $\tilde{P}^d_{t+r} \in \mathbb{P}'_{t+r}$ with $\tilde{P}^d_{t+r} = \{\theta^p_{t+r}\}$ and platform profile $A_{t+r}(\mathbb{P}'_{t+r})$ with $\tilde{a}^d_{t+r} = x^p_{t+r}$ is such that either $x^p_{t+r} \notin K\left(A_{t+r}(\mathbb{P}'_{t+r}) \mid h_{t+r}, s_{t+r}\right)$ (including the possible case of an empty core of the voting game) or $x^j_{t+r} \in K\left(A_{t+r}(\mathbb{P}'_{t+r}) \mid h_{t+r}, s_{t+r}\right)$

for some $x^j_{t+r} \neq x^p_{t+r}$ in $A_{t+r}(\mathbb{P}'_{t+r})$, else all the citizens of type θ^p_{t+r} can be made strictly better off independently of other types' actions by coalescing in party $\tilde{P}^d_{t+r}$ and setting $\tilde{a}^d_{t+r} = x^p_{t+r}$. But in the proof to part (i) (Case 1-2) I have shown that $x^p_{t+r} \in K\left(A_{t+r}(\mathbb{P}'_{t+r}) \mid h_{t+r}, s_{t+r}\right)$ for any party system featuring $\tilde{P}^d_{t+r} \in \mathbb{P}'_{t+r}$ and $\tilde{a}^d_{t+r} = x^p_{t+r}$, and that there is no citizen with $\theta^i_{t+r} \neq \theta^p_{t+r}$ that possesses in his/her set of ideal policies an element $x^i_{t+r} \neq x^p_{t+r}$ that can defeat x^p_{t+r} under the majority rule, unless $x^i_{t+r} \in I\left(\theta^p_{t+r} \mid h_{t+r}, s_{t+r}\right)$. Because the core is non-empty, this implies that in both cases the deviation leads to a policy outcome $x'_t \in I\left(\theta^p_{t+r} \mid h_{t+r}, s_{t+r}\right)$. Thus, a deviation with $\tilde{P}^d_{t+r} = \{\theta^p_{t+r}\}$ and $a^d_{t+r} = x^p_{t+r}$ is strictly profitable for all members of party $\tilde{P}^d_{t+r}$ for all $\mathbb{P}'_{t+r}$ such that $\{\theta^p_{t+r}\} \in \mathbb{P}'_{t+r}$ and all corresponding $A_{t+s}(\mathbb{P}'_{t+r})$ with $\tilde{a}^j_{t+s} \in A_{t+s}(\mathbb{P}'_{t+r})$. Thus, the tuple $\left\langle \mathbb{P}_{t+s}, A_{t+s}(\mathbb{P}_{t+s}), x^*_{t+s}\right\rangle$ violates condition *(CS)*, which implies that it is not a *SPO*, and in turn that s^*_t is not a *MPPE*, leading to a contradiction.

Part (iii). The definition of g_{t+s} implies $g_{t+s} = \frac{\lambda}{\bar{\sigma}_{t+s-1}}$. The pivotal voter θ^p_{t+s} solves $F_t\left(\theta^p_{t+s}; \rho \mid h_t\right) n_{t+s} + \lambda(n_{t+s-1} + m_{t+s-1}) = \left[1 - F_t\left(\theta^p_{t+s}; \rho \mid h_t\right)\right] n_{t+s}$. Rearranging this equation I get $F_t\left(\theta^p_{t+s}; \rho \mid h_t\right) = \frac{1-g_{t+s}}{2}$ and $\theta^p_{t+s}(g_{t+s}) = F_t^{-1}\left(\frac{1-g_{t+s}}{2}; \rho \mid h_t\right)$. Thus, $\frac{\partial \theta^p_{t+s}(g_{t+s})}{\partial g_{t+s}} = -\frac{1}{2f_t\left(\theta^p_{t+s};\rho|h_t\right)} < 0$. Q.E.D.

Part (iv). The proofs requires the following Lemma..

Lemma 2. *There exists $\tilde{\Delta} > 0$ such that if $\Delta \in [0, \tilde{\Delta})$, then (i) the function $v_{t+r}\left((M_{t+r}, L_{t+r}); \theta^i_{t+r}, \varphi \mid h_{t+r}, s_{t+r}\right)$ is jointly continuous in $(M_{t+r}, L_{t+r}), g_{t+r}, \theta^i_{t+r}, \varphi$, and strictly concave in (M_{t+r}, L_{t+r}) for all $r = 1, 2, ..., T - t$, and (ii) the equilibrium policy $x^*_{t+r}(g_{t+r})$ is a continuous function of g_{t+r}, φ for all $r = 1, 2, ..., T - t$.*

Proof. Part (i). Let $R = T - t$. Because the pivotal voter is unique in each period $t + r$ given the state g_{t+r} and collective continuation strategy s_{t+r} from Proposition 1 (ii), I can define a function $\theta^p_{t+R}(g_{t+R})$ that maps the state in period $t+r$ to the corresponding pivotal voter's type. Suppose v_{t+r} is not jointly continuous in $(M_{t+r}, L_{t+r}), g_{t+r}, \theta^i_{t+r}, \varphi$ and/or not strictly concave in (M_{t+r}, L_{t+r}) for some $r = 1, 2, ..., R$ for all values of Δ such that $\Delta > 0$. For old individuals $v_{t+r}\left((M_{t+r}, L_{t+r}); -1, \varphi \mid h_{t+r}, s_{t+r}\right) = d(\overline{G} - L_{t+r}) + c(M_{t+r})$, thus all these conditions are trivially satisfied given the assumptions on functions d, c. For a young citizen, start from $r = R$. In such period $g_{t+R+1} = \lambda/\sigma = \bar{g}$ which is invariant in x_{t+R}. Thus,

$$v_{t+R}\left((M_{t+R},L_{t+R});\theta^i_{t+R},\varphi \mid h_{t+R},s_{t+R}\right) = \begin{array}{l} A((M_{t+R},L_{t+R});\theta^i_{t+R},\varphi,g_{t+R})+ \\ +\beta\lambda B_{t+R+1}\left(M_{t+R},\varphi,g_{t+R}(h_{t+R})\right) \end{array} \tag{22}$$

where A is a jointly continuous function of $(M_{t+R},L_{t+R}),\theta^i_{t+R},\varphi,g_{t+R}$ and strictly concave in $x_{t+R}=(M_{t+R},L_{t+R})$, and B_{t+R+1} is constant in M_{t+R}. Thus, v_{t+R} is a jointly continuous function of $(M_{t+R},L_{t+R}),\theta^i_{t+R},\varphi,g_{t+R}$ and strictly concave in $x_{t+R}=(M_{t+R},L_{t+R})$. Strict concavity over a compact set implies that the pivotal citizen in period $t+R$ has a unique ideal point, i.e. $I\left(\theta^p_{t+R}\mid h_{t+R},s_{t+R}\right)=\{x^p_{t+R}\}$, which by Proposition 1 (ii) is also the unique equilibrium policy in all equilibria, i.e. $x^*_{t+R}(g_{t+R})=x^p_{t+R}$. Moreover, because $v_{t+R}\left((M_{t+R},L_{t+R});\theta^p_{t+R}(g_{t+R}),\varphi\mid h_{t+R},s_{t+R}\right)$ is jointly continuous in $(M_{t+R},L_{t+R}),\varphi,g_{t+R}$ and strictly concave in (M_{t+R},L_{t+R}), and X' is a convex set, the maximum theorem implies that $x^*_{t+R}(g_{t+R})=x^p_{t+R}$ is a jointly continuous function of $\theta^i_{t+R},\varphi,g_{t+R}$. In turn, this implies that $B_{t+R}\left(M_{t+R-1},\varphi,g_{t+R-1}(h_{t+R-1})\right)=d\left(L^*_{t+R}\right)+c\left(M^*_{t+R}\right)$ is jointly continuous in $\theta^i_{t+R},\varphi,g_{t+R}$. Thus, $v_{t+R-1}\left((M_{t+R-1},L_{t+R-1});\theta^i_{t+R-1},\varphi\mid h_{t+R-1},s_{t+R-1}\right)=$ $A((M_{t+R-1},L_{t+R-1});\theta^i_{t+R-1},\varphi,g_{t+R-1})+\beta\lambda B_{t+R}\left(M_{t+R-1},\varphi,g_{t+R-1}(h_{t+R-1})\right)$ is jointly continuous in $(M_{t+R-1},L_{t+R-1}),\theta^i_{t+R-1},\varphi,g_{t+R-1}$, and that $v_{t+R-1}\left((M_{t+R-1},L_{t+R-1});\theta^p_{t+R-1}(g_{t+R-1}),\varphi\mid h_{t+R-1},s_{t+R-1}\right)$ is jointly continuous in $(M_{t+R-1},L_{t+R-1}),\varphi,g_{t+R-1}$. Lastly, notice that $\lim_{\Delta\to 0}v_{t+R-1}\left((M_{t+R-1},L_{t+R-1});\theta^p_{t+R-1}(g_{t+R-1}),\varphi\mid h_{t+R-1},s_{t+R-1}\right)=$ $A((M_{t+R-1},L_{t+R-1});\theta^p_{t+R-1}(g_{t+R-1}),\varphi,g_{t+R-1})+\beta\lambda B_{t+R}\left(x^*_{t+R}(\bar g),x^*_{t+R+1}(\bar g);\theta^p_{t+R-1}(g_{t+R-1}),\varphi,\bar g\right)$, where B_{t+R} is constant in each element of $x_{t+R}=(M_{t+R-1},L_{t+R-1})$ and $A((M_{t+R-1},L_{t+R-1});\theta^p_{t+R-1}(g_{t+R-1}),\varphi,g_{t+R-1})$ is jointly continuous and strictly concave in (M_{t+R-1},L_{t+R-1}). Strict concavity implies $\alpha v_{t+R-1}\left(x';\theta^p_{t+R-1}(g_{t+R-1}),\varphi\mid h_{t+R-1},s_{t+R-1}\right)+(1-\alpha)v_{t+R-1}\left(x'';\theta^p_{t+R-1}(g_{t+R-1}),\varphi\mid h_{t+R-1},s_{t+R-1}\right)$ $-v_{t+R-1}\left(\alpha x'+(1-\alpha)x'';\theta^p_{t+R-1}(g_{t+R-1}),\varphi\mid h_{t+R-1},s_{t+R-1}\right)>0$ for all $x',x''\in X$ (condition A). Because v_{t+R-1} is jointly continuous in x,Δ, this implies that either (a.) condition (A) is satisfied for all $\Delta\geq 0$ and all $x',x''\in X'$, or (b.) there exists $\tilde\Delta_{t+R-1}>0$ such that if $\Delta<\tilde\Delta_{t+R-1}$ (B,$t+R-1$) then v_{t+R-1} is strictly concave in x. Set Δ such that condition (B,$t+R-1$) is satisfied. Then the pivotal voter in period $t+R-1$ has a unique ideal point, i.e. $I\left(\theta^p_{t+R-1}\mid h_{t+R-1},s_{t+R-1}\right)=\{x^p_{t+R-1}\}$, which is also the unique equilibrium policy in all equilibria given state g_{t+R-1}, i.e. $x^*_{t+R-1}(g_{t+R-1})=x^p_{t+R-1}$. Moreover, because $v_{t+R-1}\left(x_{t+R-1};\theta^p_{t+R-1}(g_{t+R-1}),\varphi\mid h_{t+R-1},s_{t+R-1}\right)$ is jointly continuous in $(M_{t+R-1},L_{t+R-1}),\varphi,g_{t+R-1}$ and strictly concave in $x_{t+R-1}=(M_{t+R-1},L_{t+R-1})$, and X' is a convex set, the maximum theorem implies that $x^*_{t+R-1}(g_{t+R-1}) = x^p_{t+R-1}$ is jointly continuous in $\theta^i_{t+R-1}, \varphi, g_{t+R-1}$. In turn, this implies that $B_{t+R-1}\left(M_{t+R-2}, \varphi, g_{t+R-2}(h_{t+R-2})\right) = d\left(L^*_{t+R-1}\right) + c\left(M^*_{t+R-1}\right)$ is jointly continuous in $\theta^i_{t+R-1}, \varphi, g_{t+R-1}$. Thus, $v_{t+R-2}\left(x_{t+R-2}; \theta^i_{t+R-2}, \varphi \mid h_{t+R-2}, s_{t+R-2}\right) = A((M_{t+R-2}, L_{t+R-2}); \theta^i_{t+R-2}, \varphi, g_{t+R-2}) + \beta\lambda B_{t+R-1}\left(M_{t+R-2}, \varphi, g_{t+R-2}(h_{t+S-2})\right)$ is jointly continuous in $(M_{t+R-2}, L_{t+R-2}), \theta^i_{t+R-2}, \varphi, g_{t+R-2}$, and that $v_{t+R-2}\left(x_{t+R-2}; \theta^p_{t+R-2}(g_{t+R-2}), \varphi \mid h_{t+R-2}, s_{t+R-2}\right)$ is jointly continuous in $(M_{t+R-2}, L_{t+R-2}), \varphi, g_{t+R-2}$. Iterate this procedure for each period $t+R-k$ and for $k = 3, 4, ..., R-r$, and assume that in each period the condition $\Delta < \tilde{\Delta}_{t+R-r}$ (B,$t+R-r$) is satisfied. Lastly, set $\tilde{\Delta} = \min\left\{\tilde{\Delta}_{t+R-r}\right\}_{r=2}^{R}$. As a result, if $\Delta < \tilde{\Delta}$, then the function v_{t+r} is jointly continuous in $(M_{t+r}, L_{t+r}), g_{t+r}, \theta^i_{t+r}, \varphi$ and strictly concave in (M_{t+r}, L_{t+r}) for each $r = 1, 2, ..., R$. This leads to a contradiction.

Part (ii). Suppose $x^*_{t+r}(g_{t+r})$ is not a continuous function of g_{t+r} for some $r = 1, 2, ..., R$. From part (i) we know that for $\Delta \in [0, \tilde{\Delta})$, $v_{t+r}\left(x_{t+r}; \theta^p_{t+r}(g_{t+r}), \varphi \mid h_{t+r}, s_{t+r}\right)$ is continuous in $x_{t+r}, g_{t+r}, \varphi$ and strictly concave in $x_{t+r} = (M_{t+r}, L_{t+r})$ for each $r = 1, 2, ..., R$, and X' is a convex set. Thus, Proposition 1 (ii) implies that $x^*_{t+r}(g_{t+r}) = x^p_{t+r}$ is the unique policy implemented in any equilibrium in each period $t+r$. Moreover, the maximum theorem implies that $x^*_{t+r}(g_{t+r}) = x^p_{t+r}$ is a continuous function of g_{t+r}, φ. This leads to a contradiction. Q.E.D.

Proposition 1. *Part (iv). $\tilde{\Delta} > 0$ such that if $\Delta \in [0, \tilde{\Delta})$ the equilibrium policy outcome x^*_t in each period t is unique given state g_t.*

Proof. Straightforward from Lemma 2 (ii).

Lemma 3. *I f $\Delta \in [0, \tilde{\Delta})$, then the function $v_{t+r}\left(x_{t+r}; \theta^i_{t+r}, \varphi \mid h_{t+r}, s_{t+r}\right)$ is infinitely jointly differentiable in $(x_{t+r}, g_{t+r}, \Delta)$ for all $r = 0, 1, 2, ..., T-1$ within the interval $\Delta \in [0, \tilde{\Delta})$.*

Proof. Assume $\Delta \in [0, \tilde{\Delta})$ and suppose $v_{t+r}\left(x_{t+r}; \theta^i_{t+r}, \varphi \mid g_{t+r}, s_{t+r}\right)$ is not infinitely jointly differentiable in x_{t+r}, g_{t+r}, Δ for some $r = 0, 1, 2, ..., R$. Recall that from period $T = t+R$ the birth rate of immigrants is such that $\Delta_T = 0$. Thus, in period T we have $g_{t+R+1} = \lambda/\sigma = \bar{g}$ which is constant in (M_{t+R}, L_{t+R}). Thus, the objective function of the pivotal voter is as in (22), in which B_{t+R+1} is constant in M_{t+R} and Δ. Thus, v_{t+R} is infinitely jointly differentiable in $(M_{t+R}, L_{t+R}), g_{t+R}, \Delta$.

Because v_{t+R} is strictly concave in x_{t+R} by Lemma 2 (i), differentiable in x_{t+R}, and X' is compact, the conditions $c'(\underline{M}) = 0$, $c'(\overline{M}) = -\infty$, $b'(0) = +\infty$, and $c'(\overline{G}) = 0$, are sufficient for the equilibrium policy to be interior, solving the CFOs:

$$\begin{gathered} \psi\xi\theta^p_{t+R}(g_{t+R}) + c'(M^*_{t+R}) + \beta\tilde{p}_{t+R}\left(\xi\theta^p_{t+R}(g_{t+R})\right)\mathit{\Delta} = 0 \\ -\xi\theta^p_{t+R}(g_{t+R}) + b'(\overline{G} - L_{t+R}) = 0 \end{gathered} \tag{23}$$

Thus, I can solve the CFOs and define functions M^*_{t+R} and L^*_{t+R} representing the equilibrium levels of M_{t+R} and L_{t+R}, respectively, Notice that the optimal solution $x^*_{t+R}(g_{t+R}) = (M^*_{t+R}, L^*_{t+R})$ is infinitely jointly differentiable in g_{t+R} because the function $\theta^p_{t+R}(g_{t+R})$ is infinitely differentiable in g_{t+R} (see proof to Proposition 1 (ii)) and the functions c', b' are C^∞ by assumption. Moreover, $x^*_{t+R}(g_{t+R}) = (M^*_{t+R}, L^*_{t+R})$ is also infinitely jointly differentiable in $\mathit{\Delta}$ because all its components are infinitely differentiable in $\mathit{\Delta}$ within the interval $\mathit{\Delta} \in [0, \tilde{\mathit{\Delta}})$. Thus, in period $t + R - 1$ the objective function of the pivotal voter becomes:

$$\begin{aligned} & v_{t+R-1}\left(x_{t+R-1}; \theta^p_{t+R-1}(g_{t+R-1}), \varphi \mid h_{t+R-1}, s_{t+R-1}\right) = \\ & = \gamma\xi\theta_{t+R-1} + \psi\xi M_{t+R-1}\theta_{t+R-1} + \xi\theta_{t+R-1}L_{t+R-1} + b(\overline{G} - L_{t+R-1}) + \\ & + c(M_{t+R-1}) + \beta\tilde{p}_{t+R}\left(\xi\theta_{t+R-1}\right)\bar{\sigma}_{t+R-1} + \beta\lambda\left[d(\overline{G} - L^*_{t+R}) + c(M^*_{t+R})\right] \end{aligned} \tag{24}$$

which is infinitely jointly differentiable in $(M_{t+R-1}, L_{t+R-1}), g_{t+R-1}, \mathit{\Delta}$ because $x^*_{t+R}(g_{t+R}) = (M^*_{t+R}, L^*_{t+R})$ is infinitely differentiable in g_{t+R}, $\mathit{\Delta}$; $\theta^p_{t+R-1}(g_{t+R-1})$, $b(\overline{G} - L_{t+R-1})$, $c(M_{t+R-1})$ are infinitely jointly differentiable in $(M_{t+R-1}, L_{t+R-1}), g_{t+R-1}$, $\mathit{\Delta}$ and $\tilde{p}_{t+R-1}\left(\xi\theta_{t+R-1}\right)\mathit{\Delta}$ is C^∞ in $\xi\theta^p_{t+R-1}, \mathit{\Delta}$. Moreover, Proposition 1 (ii) implies that for $\mathit{\Delta} \in [0, \tilde{\mathit{\Delta}})$, v_{t+R-1} is strictly concave in x_{t+R-1}, X' is a compact set, and the conditions on functions b and c ensure that the solution is interior. Thus, $x^*_{t+R-1}(g_{t+R-1}) = (M^*_{t+S-1}, L^*_{t+S-1})$ solves the CFOs:

$$\begin{gathered} \psi\xi\theta^p_{t+R-1}(g_{t+R-1}) + c'(M_{t+R-1}) + \beta\tilde{p}_{t+R-1}\left(\xi\theta^p_{t+R-1}(g_{t+R-1})\right)\mathit{\Delta} + \\ -\beta\left[d'(\overline{G} - L^*_{t+R})\frac{\partial L^*_{t+S}(g_{t+R})}{\partial g_{t+R}} + c'(M^*_{t+R})\frac{\partial M^*_{t+R}(g_{t+R})}{\partial g_{t+R}}\right]\mathit{\Delta} g^2_{t+R} = 0 \\ \\ -\xi\theta^p_{t+R-1}(g_{t+R-1}) + b'(\overline{G} - L_{t+R-1}) = 0 \end{gathered} \tag{25}$$

Again, for $\mathit{\Delta} \in [0, \tilde{\mathit{\Delta}})$, the optimal solution $(M^*_{t+R-1}, L^*_{t+R-1})$ is infinitely jointly differentiable in $g_{t+R-1}, \mathit{\Delta}$, because $x^*_{t+R}(g_{t+R}) = (M^*_{t+R}, L^*_{t+R})$ is infinitely jointly differentiable in $g_{t+R-1}, \mathit{\Delta}$ at each state g_{t+R}. In turn, this implies that the objective function of the pivotal voter in period $t + R - 2$ is infinitely jointly differentiable in $(M_{t+R-2}, L_{t+R-2}), g_{t+R-2}, \Delta$. By Iterating this procedure for each period $t+R-k$ and for $k = 3, 4, ..., R-r$ I get that $v_{t+r}\left(x_{t+r}; \theta^i_{t+r}, \varphi \mid h_{t+r}, s_{t+r}\right)$ is infinitely jointly differentiable in $(M_{t+r}, L_{t+r}), g_{t+r}, \Delta$ for all $r = 1, 2, ..., R$ and solves the CFOs:

$$\begin{gathered}
\psi\xi\theta^p_{t+r}(g_{t+r}) + c'(M_{t+r}) + \beta\tilde{p}_{t+r}\left(\xi\theta^p_{t+r}(g_{t+r})\right)\xi\Delta + \\
-\beta\left[d'(\overline{G} - L^*_{t+r+1})\frac{\partial L^*_{t+r+1}(g_{t+r+1})}{\partial g_{t+r+1}} + c'(M^*_{t+r+1})\frac{\partial M^*_{t+r+1}(g_{t+r+1})}{\partial g_{t+r+1}}\right]\Delta g^2_{t+r+1} = 0 \\
\\
-\xi\theta^p_{t+r}(g_{t+r}) + b'(\overline{G} - L_{t+r}) = 0
\end{gathered} \tag{26}$$

This leads to a contradiction. Q.E.D.

Lemma 4. *If $\Delta \in [0, \tilde{\Delta})$, then equilibrium policy $x^*_{t+r}(g_{t+r}) = (M^*_{t+r}, L^*_{t+r})$ is such that the derivatives $\left\{\frac{\partial^k M^*_{t+r}(g_{t+r})}{\partial g^k_{t+r}}, \frac{\partial^k L^*_{t+r}(g_{t+r})}{\partial g^k_{t+r}}\right\}_{k=1}^{+\infty}$ exist and are (i) jointly continuous in (g_{t+r}, Δ). (ii) There exists $\acute{\Delta} > 0$ such that if $\Delta \in [0, \acute{\Delta})$ and $g_{t+r} \in [0, 1)$ for all $r = 1, 2, ..., R$, then the derivative $\frac{\partial M^*_{t+r}(g_{t+r})}{\partial g_{t+r}}$ possesses weakly negative finite values for all $r = 0, 1, 2, ..., R$.*

Proof. Part (i) Suppose $\frac{\partial^k M^*_{t+r}(g_{t+r})}{\partial g^k_{t+r}}$ ($\frac{\partial^k L^*_{t+r}(g_{t+r})}{\partial g^k_{t+r}}$) does not exist or it is not jointly continuous in (g_{t+r}, Δ). Because for $\Delta \in [0, \tilde{\Delta})$ the solution is always interior, v_{t+r} is infinitely jointly differentiable in $(x_{t+r}, g_{t+r}, \Delta)$ by Lemma 3, then all its derivatives (of any order) with respect to (g_{t+r}, Δ) are jointly continuous in (g_{t+r}, Δ). In turn, the CFOs in (26) are also jointly continuous in (g_{t+r}, Δ) in each period $t+r$ with $r = 1, 2, ..., R$, which implies that the derivatives $\left\{\frac{\partial^k L^*_{t+r+1}(g_{t+r+1})}{\partial g^k_{t+r+1}}, \frac{\partial^k M^*_{t+r+1}(g_{t+r+1})}{\partial g^k_{t+r+1}}\right\}_{k=1}^{\infty}$ exist and are jointly continuous in (g_{t+r}, Δ). This leads to a contradiction.

Part (ii) Set $\Delta = 0$. In each period $r = 1, 2, ..., R$, because the state g_{t+r+1} is constant in (M_{t+r}, L_{t+r}) and the Markov assumption, it must be true that $\frac{dM^*_{t+r+1}(g_{t+r+1})}{dM_{t+r}} = 0$. In turn, this implies $\frac{dM^*_{t+r+}(g_{t+r})}{dg_{t+r}} = -\frac{\psi\xi\theta'^p_{t+R}(g_{t+R})}{c''(M_{t+R})} < 0$ and $\frac{dL^*_{t+r}(g_{t+r})}{dg_{t+r}} = -\frac{\theta'^p_{t+r}(g_{t+r})}{b''\left(\overline{G} - L^*_{t+r}(g_{t+r})\right)} < 0$, which both possess strictly negative finite values for any $g_{t+r} \in [0, 1)$. Because $\frac{dM^*_{t+r+}(g_{t+r})}{dg_{t+r}}$ ($\frac{dL^*_{t+r}(g_{t+r})}{dg_{t+r}}$) is jointly continuous in (g_{t+r}, Δ) by part (i), this implies either (a) $\frac{dM^*_{t+r+}(g_{t+r})}{dg_{t+r}}$ ($\frac{dL^*_{t+r}(g_{t+r})}{dg_{t+r}}$) possesses weakly negative finite values for all $\Delta \geq 0$, or (b) by the intermediate value theorem, there exists a threshold $\acute{\Delta}_{M,t+r} > 0$ ($\acute{\Delta}_{L,t+r}$) such that if $\Delta \leq \acute{\Delta}_{M,t+r}$ ($\Delta \leq \acute{\Delta}_{M,t+r}$), then $\frac{dM^*_{t+r+}(g_{t+r})}{dg_{t+r}}$ ($\frac{dL^*_{t+r}(g_{t+r})}{dg_{t+r}}$) possesses weakly negative finite values. Set $\acute{\Delta} = \min\left\{\acute{\Delta}_{M,t+r}, \acute{\Delta}_{L,t+r}, \tilde{\Delta}\right\}_{r=0}^{R}$. Then $\frac{dM^*_{t+r+}(g_{t+r})}{dg_{t+r}}$ and $\frac{dL^*_{t+r}(g_{t+r})}{dg_{t+r}}$ possesses weakly negative finite values for all $r = 1, 2, ..., R$, which leads to a contradiction. Q.E.D.

B.2 Steady-State

Proposition 2. *There exists $\check{\Delta} > 0$ such that if (i) $\Delta \in [0, \check{\Delta})$, then a steady-state exists, which is unique and globally stable. If $\Delta = 0$, then (ii) any MPPE at time t is history-independent and the economy converges immediately to the steady state after any history h_t.*

Proof. Part (i). Suppose a steady state does not exists for any $\Delta > 0$. Consider $\Delta \in [0, \tilde{\Delta})$. By Lemma 2 (i) and Lemma 3 the pivotal voter's objective function each period $t + r$ is strictly concave and differentiable in (M_{t+r}, L_{t+r}), thus the equilibrium policy satisfies the CFOs. Thus, to prove the existence of a steady-state it is sufficient to show that there exists $M_{SS} \in [\underline{M}, \overline{M}]$ that solves

$$\psi\xi\theta^p(\tilde{g}(M_{SS}; \Delta)) + c'(M_{SS}) + \beta\tilde{p}\left(\theta^p(\tilde{g}(M_{SS}; \Delta))\right)\Delta - \beta D(\tilde{g}(M_{SS}; \Delta))\Delta\tilde{g}(M_{SS}; \Delta)^2 = 0 \qquad (27)$$

where $\tilde{g}(M_{SS}; \Delta) = \lambda/[\sigma^m - \Delta(1 - M_{SS})]$ and $D(\tilde{g}(M_{SS}; \Delta)) = d'(\overline{G} - L^*_{t+r+1}(g))\frac{\partial L^*_{t+r+1}(g)}{\partial g} + c'(M^*_{t+r+1}(g))\frac{\partial M^*_{t+r+1}(g)}{\partial g}\Big|_{g=\tilde{g}(M_{SS};\Delta)}$. If such M_{SS}, then both the state $g_t = \tilde{g}(M_{SS}; \Delta) = g_{SS}$ and the equilibrium L^*_{t+r} do not vary with time, i.e. $-\theta^p(g(M_{SS}; \Delta)) + b'(G - L_{SS}) = 0$. Firstly, notice that at $M = \underline{M}$ by assumption $c'(\underline{M}) = 0$. Secondly, at $M = \overline{M}$ by assumption $c'(\overline{M}) = -\infty$. Secondly, Lemma 4 (ii) implies that if $\Delta \in [0, \tilde{\Delta})$, then $D(g)$ has finite value for all $g \in [0, 1)$. This implies that at $\Delta = 0$ the LHS of equation (27) is strictly positive for $M = \underline{M}$ *(3.A)* and strictly negative for $M = \overline{M}$ *(3.B)*. Thus, either conditions *(3.A)* and *(3.B)* hold true for all $\Delta \in [0, \tilde{\Delta})$, or by the intermediate value theorem there exists $\tilde{\Delta} \geq \check{\Delta}_1 > 0$ such that if $\Delta \in [0, \check{\Delta}_1)$ then *(3.A)* and *(3.B)* hold true. Consider $\Delta \in [0, \check{\Delta}_1)$. Lemma 2 (i) implies that the function in (27) is continuous in M. Thus, the intermediate value theorem implies that there exists $M_{SS} \in [\underline{M}, \overline{M}]$ such that equation (27) is satisfied. Thus, a steady state exists for any $\Delta \in [0, \check{\Delta}_1)$, which leads to a contradiction. Now suppose the steady-state is not unique or globally stable for some $\Delta \in [0, \check{\Delta}_1)$. Notice that (M^*_{t+r}, L^*_{t+r}) is a continuous function of g_{t+r} by Lemma 3 and g_{t+r} is itself a continuous function of M_{t+r-1} for each $r = 0, 1, 2, ..., R$. Thus, for uniqueness and global stability of the steady-state it is sufficient to show that $-1 < \frac{dM^*_{t+r}}{dg_{t+r}}\frac{\partial g_{t+r}}{\partial M_{t+r-1}} < 1$ for all $M_{t+r-1} \in [\underline{M}, \overline{M}]$ in each period $t + r$ for $r = 0, 1, 2, ..., R$ *(3.C)*. Notice that $\frac{dM^*_{t+r}}{dg_{t+r}}\frac{\partial g_{t+r}}{\partial M_{t+r-1}} = -\frac{dM^*_{t+r}}{dg_{t+r}}\frac{\Delta}{\lambda}g^2_{t+r}$, where $\frac{dM^*_{t+r}}{dg_{t+r}}$ has finite value by Lemma 4 (ii) and $g_{t+r} \in [0, 1)$. Thus, at $\Delta = 0$ I get $\frac{dM^*_{t+r}}{dg_{t+r}}\frac{\partial g_{t+r}}{\partial M_{t+r-1}} = 0$ and the condition is satisfied. Lastly, continuity of $\frac{dM^*_{t+r}}{dg_{t+r}}$ by Lemma 3 implies that either (a) the condition *(3.C)* is satisfied for all $\Delta \in [0, \check{\Delta}_1)$, or (b) there exists $\check{\Delta}_2 > 0$ such that if $\Delta \in [0, \check{\Delta}_2)$, then $0 < \frac{dM^*_{t+r}}{dg_{t+r}} \frac{\partial g_{t+r}}{\partial M_{t+r-1}} < 1$ for all $M_{t+r-1} \in [\underline{M}, \overline{M}]$ in each period $t+r$ for $r = 0, 1, 2, ..., R$. Set $\check{\Delta} = \min\left\{\check{\Delta}_1, \check{\Delta}_2\right\}$. In turn, this implies that if $\Delta \in [0, \check{\Delta})$, then the steady-state exists, it is unique and globally stable. This leads to a contradiction.

Part (ii). Suppose the equilibrium is history-dependent. If $\Delta = 0$, then $\frac{\partial g_{t+r}}{\partial M_{t+r-1}} = 0$ in each period $t+r$ for $r = 0, 1, 2, ..., R$. Moreover, at $\Delta = 0$, $\frac{dM^*_{t+r}}{dg_{t+r}}$ has finite value by Lemma 4 (ii). Then $\frac{dM^*_{t+r}}{dg_{t+r}} \frac{\partial g_{t+r}}{\partial M_{t+r-1}} = 0$ which implies that the condition $-1 < \frac{dM^*_{t+r}}{dg_{t+r}} \frac{\partial g_{t+r}}{\partial M_{t+r-1}} < 1$ for all $M_{t+r-1} \in [0, \overline{M}]$ in each period $t+r$ for $r = 0, 1, 2, ..., R$ is trivially satisfied. Moreover, $\frac{dM^*_{t+r}}{dg_{t+r}} \frac{\partial g_{t+r}}{\partial M_{t+r-1}} = 0$ also implies that $M^*_{t+r} = M^*_t = M_{SS}$, $L^*_{t+r} = L^*_t = L_{SS}$, and $g_{t+r+1} = g_{t+1} = g_{SS}$ for any state g_t, and for all $r = 1, 2, ..., R$, i.e. the equilibrium policy is history-independent and the economy converges immediately to the steady-state after any history h_t. This leads to a contradiction. Q.E.D.

B.3 Comparative Statics

Proposition 3. *(Effect of population ageing, increasing inequality, and economic depression). There exists $\hat{\Delta} > 0$ such that if $\Delta \in [0, \hat{\Delta})$, then (i) an increase in longevity and/or (ii) an increase in income inequality and/or (iii) a critical decrease in fertility, and/or (iv) an economic depression translate to (1) a less open immigration policy M_{SS}, a less liberal economic policy L_{SS}, (3) a larger size of government τ_{SS}, and (4) a higher old-age dependency ratio g_{SS}.*

Proof. Part (i)-(1), -(2). Suppose (i)-(1) or -(2) does not hold true (or both). Consider any $\lambda', \lambda'' \in [\underline{\lambda}, 1]$ such that $\lambda' > \lambda''$. I define the set $\Phi_\lambda(\varphi) := \{\hat{\varphi} \in \Phi \mid \hat{\varphi}_j = \varphi_j \; \forall j \neq 3\}$ and the ordering $\leq_\lambda$ over $\Phi_\lambda(\varphi)$ such that $\varphi' \leq \varphi''$ if and only if $\lambda' \geq \lambda''$. Consider any two elements $\varphi' = (\beta, \gamma, \lambda'', \Delta, \sigma^m, \xi, l, \rho)$ and $\varphi'' = (\beta, \gamma, \lambda', \Delta, \sigma^m, \xi, l, \rho)$ of $\Phi_\lambda(\varphi)$ such that $\varphi' \leq \varphi''$. Lastly, let $g'_t(h_t) = \lambda' / [\sigma^m - \Delta(1 - M_{t-1})]$ and t $g''_t(h_t) = \lambda'' / [\sigma^m - \Delta(1 - M_{t-1})]$. Consider any two policies $(M''_t, L''_t), (M'_t, L'_t) \in X'$ such that $(M''_t, L''_t) \geq (M'_t, L'_t)$. Then v^p_t satisfies the single crossing property *(SC)* in (M_t, L_t, φ) over $\Phi_\lambda(\varphi)$ if:

$$\begin{aligned} v_t\left((M''_t, L''_t); \theta^p_t\left(g''_t(h_t)\right), \varphi'' \mid h_t, s_t\right) - v_t\left((M'_t, L'_t); \theta^p_t\left(g''_t(h_t)\right), \varphi'' \mid h_t, s_t\right) \geq \\ v_t\left((M''_t, L''_t); \theta^p_t\left(g'_t(h_t)\right), \varphi' \mid h_t, s_t\right) - v_t\left((M'_t, L'_t); \theta^p_t\left(g'_t(h_t)\right), \varphi' \mid h'_t, s_t\right) \end{aligned} \tag{28}$$

Recall $g_t \in [0, 1)$. implies $\theta^p_t > 0$. Using (14) condition (28) rewrites:

$$
\begin{aligned}
&\xi\left[\psi\left(M_t''-M_t'\right)+\left(L_t''-L_t'\right)\right]\left[\theta_t^p\left(g_t''(h_t)\right)-\theta_t^p\left(g_t'(h_t)\right)\right]+\\
&+\beta\left[\tilde{p}\left(\xi\theta_t^p\left(g_t''(h_t)\right)\right)-\tilde{p}\left(\xi\theta_t^p\left(g_t'(h_t)\right)\right)\right]\Delta\left(M_t''-M_t'\right)+\\
&+\beta\lambda''\left[B_{t+1}\left(M_t'',\varphi'',g_t''(h_t)\right)-B_{t+1}\left(M_t',\varphi'',g_t''(h_t)\right)\right]+\\
&-\beta\lambda'\left[B_{t+1}\left(M_t'',\varphi'',g_t''(h_t)\right)-B_{t+1}\left(M_t',\varphi'',g_t''(h_t)\right)\right]\geq 0
\end{aligned}
\tag{29}
$$

Recall $g_t \in [0,1)$. Notice that for $\mathit{\Delta} \in [0,\tilde{\mathit{\Delta}})$ the LHS of (29) is continuous in $\mathit{\Delta}$ by Lemma 2 (i) and that $\lim_{\mathit{\Delta}\to 0}\left[B_{t+1}\left(M_t'',\varphi,g_t(h_t)\right)-B_{t+1}\left(M_t',\varphi,g_t(h_t)\right)\right]=0$ for all $g_t(h_t)$. Thus, either the inequality above is satisfied for all values of $\mathit{\Delta} \in [0,\tilde{\mathit{\Delta}})$ for any two $\varphi',\varphi'' \in \Phi_\lambda(\varphi)$ and for all $(M_t,L_t)\in X'$, or the intermediate value theorem implies that there exists $\mathring{\mathit{\Delta}}_1 > 0$ such that if $\mathit{\Delta} \in [0,\mathring{\mathit{\Delta}}_1)$, then the inequality above is satisfied for any two $\varphi',\varphi'' \in \Phi_\lambda(\varphi)$ and for all $(M_t,L_t)\in X'$. Thus, there exists threshold $\mathring{\mathit{\Delta}}_1 > 0$ such that for $\mathit{\Delta} \in [0,\mathring{\mathit{\Delta}}_1)$ the equilibrium policy (M_t,L_t) is weakly increasing in φ over $\Phi_\lambda(\varphi)$, and therefore weakly decreasing in λ. Lastly, notice that a decrease in (M_t,L_t) implies $g_{t+1} \geq g_t$, which for $\mathit{\Delta} \leq \acute{\mathit{\Delta}}$ and $g_{t+r} \in [0,1)$ by Lemma 4 (ii) implies $(M_{t+1},L_{t+1}) \leq (M_t,L_t)$. Set $\hat{\mathit{\Delta}}_1 \leq \min\left\{\mathring{\mathit{\Delta}}_1,\acute{\mathit{\Delta}}\right\}$. Iterating this procedure for each period $t+r$ and for $r=1,2,...$, this implies $(M_{t+r+1},L_{t+r+1}) \leq (M_{t+r},L_{t+r})$ for all $r=0,1,2,...,R$. Thus, the new steady-state must be such that $(M_{SS},L_{SS}) \leq (M_{t-1},L_{t-1})$. This leads to a contradiction. Part (i)-(3), -(4) are straightforward from (i)-(1), -(2) given that $\tau_{SS}=\tau\left(M_{SS},L_{SS}\right)$, which by formula (5) is decreasing in both M_{SS} and L_{SS} and constant in λ, and that g_{SS} is decreasing in M_{SS}, constant in L_{SS} and increasing in λ.

Part (ii)-(1), -(2). Suppose (ii)-(1) or -(2) does not hold true (or both). First, I prove that the type of the pivotal voter is decreasing in ρ. Recall $g_t \in [0,1)$. Using (7) I get $\theta_t^p = Q^{-1}\left(0.5(1-g_t);\rho\right)$ with $Q(\theta_t;\rho)=\rho Q_2(\theta_t)+(1-\rho)Q_1(\theta_t)$. Assumption $Q_1(\check{\theta}) \geq 0.5$ implies $Q(\theta_t^p;\rho) < 0.5$ and $Q_2(\theta_t^p)-Q_1(\theta_t^p) < 0$. Thus, I get

$$
\frac{\partial\theta_t^p}{\partial\rho}=\frac{Q_1(\theta_t^p)-Q_2(\theta_t^p)}{\rho q_2(\theta_t^p)+(1-\rho)q_1(\theta_t^p)}<0 \tag{30}
$$

i.e., an increase in income inequality ρ simply means that θ_t^p decreases at constant mean income.

Consider any two values of $\rho \in [0,1]$ such that $\rho'' < \rho'$. I define the set $\Phi_\rho(\varphi) := \{\hat{\varphi} \in \Phi \mid \hat{\varphi}_j = \varphi_j \ \forall j \neq 8\}$ and the ordering $\leq_\rho$ over $\Phi_\rho(\varphi)$ such that $\varphi' \leq \varphi''$ if and only if $\rho'' < \rho'$. Consider any two elements $\varphi'=(\beta,\gamma,\lambda,\mathit{\Delta},\sigma^m,\xi,l,\rho')$ and $\varphi''=(\beta,\gamma,\lambda,\mathit{\Delta},\sigma^m,\xi,l,\rho'')$ of $\Phi_\rho(\varphi)$ such that $\varphi' \leq \varphi''$. Lastly, let $(\theta_t^p)'$ and $(\theta_t^p)''$ denote the type of the pivotal voter under ρ' and ρ'', respectively, and note that (30) implies $(\theta_t^p)'' > (\theta_t^p)'$. Consider any two policies $(M_t'',L_t''),(M_t',L_t') \in X'$ such that

$(M_t'', L_t'') \geq (M_t', L_t')$. Then v_t^p satisfies *(SC)* in (M_t, L_t, φ) over $\Phi_\rho(\varphi)$ if:

$$\begin{gathered} v_t\left((M_t'', L_t''); (\theta_t^p)'', \varphi'' \mid h_t, s_t\right) - v_t\left((M_t', L_t'); (\theta_t^p)'', \varphi_t'' \mid h_t, s_t\right) \geq \\ v_t\left((M_t', L_t'); (\theta_t^p)', \varphi_t' \mid h_t, s_t\right) - v_t\left((M_t', L_t'); (\theta_t^p)', \varphi_t' \mid h_t, s_t\right) \end{gathered} \tag{31}$$

Using (14) condition (28) rewrites:

$$\begin{gathered} \xi\left[\psi\left(M_t'' - M_t'\right) + \left(L_t'' - L_t'\right)\right]\left[(\theta_t^p)'' - (\theta_t^p)'\right] + \\ +\beta\left[\tilde{p}\left(\xi(\theta_t^p)''\right) - \tilde{p}\left(\xi(\theta_t^p)'\right)\right]\Delta\left(M_t'' - M_t'\right) \\ +\beta\lambda\left[B_{t+1}\left(M_t'', \varphi'', g_t(h_t)\right) - B_{t+1}\left(M_t', \varphi'', g_t(h_t)\right)\right] + \\ -\beta\lambda\left[B_{t+1}\left(M_t'', \varphi', g_t(h_t)\right) - B_{t+1}\left(M_t', \varphi', g_t(h_t)\right)\right] \geq 0 \end{gathered} \tag{32}$$

Recall $g_t \in [0, 1)$. Notice that for $\mathit{\Delta} \in [0, \tilde{\mathit{\Delta}})$ the LHS of (32) is continuous in $\mathit{\Delta}$ by Lemma 2 (i) and that $\lim_{\mathit{\Delta} \to 0}\left[B_{t+1}\left(M_t'', \varphi, g_t(h_t)\right) - B_{t+1}\left(M_t', \varphi, g_t(h_t)\right)\right] = 0$ for all $g_t(h_t)$. Thus, either the inequality above is satisfied for all values of $\mathit{\Delta} \in [0, \tilde{\mathit{\Delta}})$ for any two $\varphi', \varphi'' \in \Phi_\rho(\varphi)$ and for all $(M_t, L_t) \in X'$, or the intermediate value theorem implies that there exists $\mathring{\mathit{\Delta}}_2 > 0$ such that if $\mathit{\Delta} \in [0, \mathring{\mathit{\Delta}}_2)$, then the inequality above is satisfied for any two $\varphi', \varphi'' \in \Phi_\rho(\varphi)$ and for all $(M_t, L_t) \in X'$. Thus, there exists threshold $\mathring{\mathit{\Delta}}_2 > 0$ such that for $\mathit{\Delta} \in [0, \mathring{\mathit{\Delta}}_2)$ the equilibrium policy (M_t, L_t) is weakly increasing in φ over $\Phi_\rho(\varphi)$, and therefore weakly decreasing in ρ. Lastly, notice that a decrease in (M_t, L_t) implies $g_{+1} \geq g_t$, which for $\mathit{\Delta} \leq \acute{\mathit{\Delta}}$ and $g_{t+r} \in [0, 1)$ by Lemma 4 (ii) implies $(M_{t+1}, L_{t+1}) \leq (M_t, L_t)$. Set $\hat{\mathit{\Delta}}_2 \leq \left\{\mathring{\mathit{\Delta}}_2, \acute{\mathit{\Delta}}\right\}$. Iterating this procedure for each period $t + r$ with $r = 1, 2, ...$, this implies $(M_{t+r+1}, L_{t+r+1}) \leq (M_{t+r}, L_{t+r})$ for all $r = 0, 1, 2, ..., R$. Thus, the new steady-state must be such that $(M_{SS}, L_{SS}) \leq (M_{t-1}, L_{t-1})$. This leads to a contradiction. Part (ii)-(3), -(4) are straightforward from (ii)-(1), -(2) given that $\tau_{SS} = \tau\left(M_{SS}, L_{SS}\right)$, which by formula (5) is decreasing in both M_{SS} and L_{SS} and constant in ρ, and that g_{SS} is decreasing in M_{SS}, constant in L_{SS} and and constant in ρ,

Part (iii)-(1), -(2). Suppose (iii)-(1) or -(2) does not hold true (or both). Consider any $\mathit{\Delta}', \mathit{\Delta}'' \in [0, \hat{\mathit{\Delta}}_3]$ for some threshold $\hat{\mathit{\Delta}}_3 \leq \tilde{\mathit{\Delta}}$ such that $\mathit{\Delta}' > \mathit{\Delta}''$. I define the set $\Phi_\Delta(\varphi) := \{\hat{\varphi} \in \Phi \mid \hat{\varphi}_j = \varphi_j \ \forall j \neq 4, \hat{\varphi}_4 \leq \hat{\mathit{\Delta}}_3\}$ and the ordering $\leq_\Delta$ over $\Phi_\Delta(\varphi)$ such that $\varphi' \leq \varphi''$ if and only if $\mathit{\Delta}' \geq \mathit{\Delta}''$. Consider any two elements $\varphi' = (\beta, \gamma, \lambda, \Delta'', \sigma^m, \xi, l)$ and $\varphi'' = (\beta, \gamma, \lambda, \Delta', \sigma^m, \xi, l)$ of $\Phi_\lambda(\varphi)$ such that $\varphi' \leq \varphi''$. Lastly, let $g_t'(h_t) = \lambda / \left[\sigma^m - \Delta'(1 - M_{t-1})\right]$ and t $g_t''(h_t) = \lambda / \left[\sigma^m - \Delta''(1 - M_{t-1})\right]$. Consider any two policies $(M_t'', L_t''), (M_t', L_t') \in X'$ such that $(M_t'', L_t'') \geq (M_t', L_t')$. Then v_t^p satisfies *(SC)* in (M_t, L_t, φ)

over $\Phi_{\Delta}(\varphi)$ if:

$$\begin{aligned} v_t\left((M_t'', L_t''); \theta_t^p\left(g_t''(h_t)\right), \varphi'' \mid h_t, s_t\right) - v_t\left((M_t', L_t'); \theta_t^p\left(g_t''(h_t)\right), \varphi'' \mid h_t, s_t\right) \geq \\ v_t\left((M_t'', L_t''); \theta_t^p\left(g_t'(h_t)\right), \varphi' \mid h_t, s_t\right) - v_t\left((M_t', L_t'); \theta_t^p\left(g_t'(h_t)\right), \varphi' \mid h_t, s_t\right) \end{aligned} \tag{33}$$

Using (14) condition (33) rewrites:

$$\begin{aligned} \xi\left[\psi\left(M_t'' - M_t'\right) + \left(L_t'' - L_t'\right)\right]\left[\theta_t^p\left(g_t''(h_t)\right) - \theta_t^p\left(g_t'(h_t)\right)\right] + \\ +\beta\left\{\tilde{p}\left(\xi\theta_t^p\left(g_t''(h_t)\right)\right)\Delta'' - \tilde{p}\left(\xi\theta_t^p\left(g_t'(h_t)\right)\right)\Delta'\right\}\left(M_t'' - M_t'\right) + \\ \beta\lambda\left[B_{t+1}\left(M_t'', \varphi', g_t'(h_t)\right) - B_{t+1}\left(M_t', \varphi', g_t'(h_t)\right)\right] + \\ -\beta\lambda\left[B_{t+1}\left(M_t'', \varphi'', g_t''(h_t)\right) - B_{t+1}\left(M_t', \varphi'', g_t''(h_t)\right)\right] \geq 0 \end{aligned} \tag{34}$$

Firstly, a sufficient condition for the first two lines of (34) to be positive is $\frac{1-M_{t-1}}{2q(\theta_t^p(g_t'(h_t));\rho)}\frac{g_t'(h_t)^2}{\lambda} > \beta\tilde{p}\left(\xi\theta_t^p\left(g_t''(h_t)\right)\right)$. Notice that under the assumptions $\lim_{\theta_t \to 0^+} q\left(\theta_t; \rho\right) = 0$ and $\overline{M} < 1$ such condition is always satisfied for $g_t \to 1$. Thus, either it is also satisfies for all $g_t \in [0, 1)$, or by the intermediate value theorem there exists a threshold $\hat{g}_t \in (0, 1)$ such that for any state $g_t \in [\hat{g}_t, 1)$ the the first two lines of (34) have positive value. Secondly, notice that for $\Delta \in [0, \tilde{\Delta})$ the LHS of (34) is continuous in Δ by Lemma 2 (i) and that $\lim_{\Delta \to 0}\left[B_{t+1}\left(M_t', \varphi, g_t(h_t)\right) - B_{t+1}\left(M_t'', \varphi, g_t(h_t)\right)\right] = 0$. Thus, for $g_t \in [\hat{g}_t, 1)$ either the inequality above is satisfied for all $\Delta \in [0, \tilde{\Delta})$, for any two $\varphi', \varphi'' \in \Phi_{\Delta}(\varphi)$ and for all $(M_t, L_t) \in X'$, or the intermediate value theorem implies that there exists $\mathring{\Delta}_3 > 0$ such that if $\Delta \in [0, \mathring{\Delta}_3)$, then the inequality above is satisfied for any two $\varphi', \varphi'' \in \Phi_{\Delta}(\varphi)$ and for all $(M_t, L_t) \in X'$. Thus, there exists threshold $\mathring{\Delta}_3 > 0$ such that for $\Delta \in [0, \mathring{\Delta}_3)$ the equilibrium policy (M_t, L_t) is weakly increasing in φ over $\Phi_{\Delta}(\varphi)$, and therefore weakly decreasing in Δ. Lastly, notice that a decrease in (M_t, L_t) implies $g_{t+1} \geq g_t$, which for $\Delta \in [0, \acute{\Delta})$ and $g_{t+r} \in [0, 1)$ by Lemma 4 (ii) implies $(M_{t+1}, L_{t+1}) \leq (M_t, L_t)$. Set $\hat{\Delta}_3 \leq \min\{\mathring{\Delta}_3, \acute{\Delta}\}$. Iterating for each period $t + r$ for $r = 1, 2, ...$ this implies $(M_{t+r+1}, L_{t+r+1}) \leq (M_{t+r}, L_{t+r})$ for all $r = 0, 1, 2, ..., R$. Thus, the new steady-state must be such that $(M_{SS}, L_{SS}) \leq (M_{t-1}, L_{t-1})$. This leads to a contradiction. Part (iii)-(3), -(4) are straightforward from (i)-(1), -(2) given that $\tau_{SS} = \tau\left(M_{SS}, L_{SS}\right)$, which by formula (5) is decreasing in both M_{SS} and L_{SS} and constant in Δ, and that g_{SS} is decreasing in M_{SS}, constant in L_{SS} and and increasing in Δ.

Part (iv)-(1), -(2). Suppose (iv)-(1) or -(2) does not hold (or both). Consider any $\xi', \xi'' \in (0, +\infty)$ such that $\xi'' > \xi'$. I define the following notation. $\Phi_{\xi}(\varphi) := \{\hat{\varphi} \in \Phi \mid \hat{\varphi}_j = \varphi_j \ \forall j \neq 6\}$ and the ordering $\leq_{\xi}$ over $\Phi_{\xi}(\varphi)$ such that $\varphi' \leq \varphi''$ if and only if $\xi'' > \xi'$. Consider any two elements $\varphi' = (\beta, \gamma, \lambda, \Delta, \sigma^m, \xi', l)$ and $\varphi'' = (\beta, \gamma, \lambda, \Delta, \sigma^m, \xi'', l)$ of $\Phi_{\xi}(\varphi)$ such that $\varphi' \leq \varphi''$ and any two policies $(M_t'', L_t''), (M_t', L_t') \in X'$ such that $(M_t'', L_t'') \geq (M_t', L_t')$. Then v_t^p satisfies *(SC)* in (M_t, L_t, φ) over $\Phi_\xi(\varphi)$ if:

$$\begin{gathered} v_t\left((M_t'', L_t''); \theta_t^p\left(g_t(h_t)\right), \varphi'' \mid h_t, s_t\right) - v_t\left((M_t', L_t'); \theta_t^p\left(g_t(h_t)\right), \varphi'' \mid h_t, s_t\right) \geq \\ v_t\left((M_t'', L_t''); \theta_t^p\left(g_t(h_t)\right), \varphi' \mid h_t, s_t\right) - v_t\left((M_t', L_t'); \theta_t^p\left(g_t(h_t)\right), \varphi' \mid h_t, s_t\right) \end{gathered} \tag{35}$$

Using (14) condition (35) rewrites:

$$\begin{gathered} \left[\psi\left(M_t'' - M_t'\right) + \left(L_t'' - L_t'\right)\right] \theta_t^p\left(g_t(h_t)\right)\left(\xi'' - \xi'\right) + \\ +\beta\left\{\tilde{p}\left(\xi''\theta_t^p\left(g_t(h_t)\right)\right) - \tilde{p}\left(\xi'\theta_t^p\left(g_t(h_t)\right)\right)\right\} \Delta\left(M_t'' - M_t'\right) + \\ \beta\lambda\left[B_{t+1}\left(M_t'', \varphi'', g_t(h_t)\right) - B_{t+1}\left(M_t', \varphi'', g_t(h_t)\right)\right] + \\ -\beta\lambda\left[B_{t+1}\left(M_t'', \varphi', g_t(h_t)\right) - B_{t+1}\left(M_t', \varphi'', g_t(h_t)\right)\right] \geq 0 \end{gathered} \tag{36}$$

Notice that g_{t+1} is constant in ξ'. This implies that $B_{t+1}\left(M_t'', \varphi'', g_t(h_t)\right) - B_{t+1}\left(M_t'', \varphi', g_t(h_t)\right) = d\left(\overline{G} - L_{t+1}^*(g_{t+1}(h_{t+1}''))\right) + c\left(M_{t+1}^*(g_{t+1}(h_{t+1}'')\right) - d\left(\overline{G} - L_{t+1}^*(g_{t+1}(h_{t+1}''))\right) - c\left(M_{t+1}^*(g_{t+1}(h_{t+1}'')\right) = 0$, where h_t'' denotes the history after policy choice M_t''. Thus, the inequality in (36) is always satisfied given the assumption that $\tilde{p}$ is a weakly increasing function. Lastly, notice that a decrease in (M_t, L_t) implies $g_{t+1} \geq g_t$, which for $\Delta \in [0, \acute{\Delta})$ and $g_{t+r} \in [0, 1)$ by Lemma 4 (ii) implies $(M_{t+1}, L_{t+1}) \leq (M_t, L_t)$. Set $\hat{\Delta}_4 = \acute{\Delta}$. Iterating for each period $t + r$ for $r = 1, 2, ...$ this implies $(M_{t+r+1}, L_{t+r+1}) \leq (M_{t+r}, L_{t+r})$ for all $r = 0, 1, 2, ..., R$. Thus, the new steady-state must be such that $(M_{SS}, L_{SS}) \leq (M_{t-1}, L_{t-1})$. This leads to a contradiction. Part (i)-(3), -(4) are straightforward from (i)-(1), -(2) given that $\tau_{SS} = \tau\left(M_{SS}, L_{SS}\right)$, which by formula (5) is decreasing in both M_{SS} and L_{SS} and constant in ξ, and that g_{SS} is decreasing in M_{SS}, constant in L_{SS} and constant in ξ. Lastly, define $\hat{\Delta} = \min\left\{\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \hat{\Delta}_4\right\}$ and note that $\hat{\Delta} > 0$. Then for$\Delta \in [0, \hat{\Delta})$ all the statements in parts (i)-(ii)-(iii)-(iv) hold true. Q.E.D.

B.4 Equilibrium Party System

Proposition 4. *(Equilibrium party system). If $\Delta \in [0, \hat{\Delta})$, then (i) in any MPPE, the tuple $\langle \mathbb{P}_{t+r}, A_{t+r}\left(\mathbb{P}_{t+r}\right)\rangle$ is a populist vs. libertarian party system in each period $t + r$ for all $r = 1, 2, ...$; (ii) each member of any right-wing populist party is weakly older and weakly lower-income than any member of any strictly libertarian party; (iii) if a marginal shock of type (a), (b), (c), and/or (d) occurs in period t, then in any period $t + r$ with $r = 0, 1, 2, ...$ the winning party is right-wing populist; in particular, (iv) there is an MPPE such that in each period $t+r$ the tuple $\langle \mathbb{P}_{t+r}, A_{t+r}\left(\mathbb{P}_{t+r}\right), x_{t+r}\rangle$ is a stable two-party structure in which the winning party is right-wing populist and includes all citizens* *with $\theta^i_{t+r} \leq \theta^p_{t+r}$ (i.e., the old and the low-income citizens).*

Proof. Part (i) Suppose there is a *MPPE* such that $\langle \mathbb{P}_{t+r}, A_{t+r}(\mathbb{P}_{t+r})\rangle$ is not a populist vs. libertarian party system. Using Definition 3, this is true only if $A_{t+r}(\mathbb{P}_{t+r})$ is not totally ordered under $\leq$. Because for $\Delta \in [0, \tilde{\Delta})$ the objective function of each citizen's type is strictly concave by Lemma 2 (i) and satisfies *QSM* and *SSC* in $(x_t, \theta_t)$by Lemma 1, Theorem 4 in Milgrom and Shannon (1994) implies that the set of all citizen's ideal policies $UI(h_{t+r}, s_{t+r}) \equiv \cup_{\theta_{t+r}\in\Theta} I(\theta_{t+r} \mid h_{t+r}, s_{t+r})$ in each period $t+r$ is totally ordered under $\leq$. Any possible policy profile $A_{t+r}(\mathbb{P}_{t+r})$ in period $t+r$ that is part of a *MPPE* in period t is such that $A_{t+r}(\mathbb{P}_{t+r}) \subseteq UI(h_{t+r}, s_{t+r})$ because of the assumption *(CC)*. Thus, $A_{t+r}(\mathbb{P}_{t+r})$ is also totally ordered under $\leq$, which leads to a contradiction.

Part (ii). First, this statement trivially holds true for one-party systems. If $|\mathbb{P}_{t+r}| > 1$ consider two parties $P^j_{t+r}, P^k_{t+r} \in \mathbb{P}_{t+r}$ with $M^j_{t+r} \geq M^k_{t+r}$ and $L^j_{t+r} \geq L^k_{t+r}$ and $(M^j_{t+r}, L^j_{t+r}) \neq (M^k_{t+r}, L^k_{t+r})$. Suppose (ii) does not hold. Then there exists $\theta^i_{t+r} \in P^j_{t+r}$ and $\theta^l_{t+r} \in P^k_{t+r}$ with $\theta^i_{t+r} < \theta^l_{t+r}$. Part (i) implies that party platforms are totally ordered. *(PM)* implies $v_t\left(a^j_t; \theta^i_t, \varphi \mid h_t, s_t\right) \geq v_t\left(a^k_t; \theta^i_t, \varphi \mid h_t, s_t\right)$. The *SSC* and $a^j_t \geq a^k_t$ imply $v_t\left(a^j_t; \theta^l_t, \varphi \mid h_t, s_t\right) > v_t\left(a^k_t; \theta^l_t, \varphi \mid h_t, s_t\right)$. But then *(PM)* implies that $\theta^l_t \notin P^k_t$. This leads to a contradiction.

Part (iii). Straightforward from Definition 2.4 and Proposition 3.

Part (iv). Suppose a shock of type (a) occurs in period t such that $\lambda'' > \lambda'$, and that a *MPPE* that satisfies (iv) does not exists. Consider in each period $t+r$ a tuple $\left\langle \{P^1_{t+r}, P^2_{t+r}\}, A_{t+r}, x^p_{t+r}\right\rangle$ with $A_{t+r} = \left\{x^p_{t+r}, x^j_{t+r}\right\}$ where $x^p_{t+r} \in I\left(\theta^p_{t+r} \mid h_{t+r}, s_{t+r}\right)$ and $x^j_{t+r} \in I\left(\theta^j_{t+r} \mid h_{t+r}\right)$ for some $\theta^j_{t+r} > \theta^p_{t+r}$. The latter inequality implies $x^j_{t+s} \geq x^p_{t+s}$ because of monotonicity of the optimal policy (Milgrom ans Shannon 1994). Optimality implies $v_{t+r}\left(x^p_{t+r}; \theta^p_{t+r}, \varphi \mid h_{t+r}, s_{t+r}\right) \geq v_{t+s}\left(x^j_{t+r}; \theta^p_{t+r}, \varphi \mid h_{t+r}, s_{t+r}\right)$ and $v_{t+r}\left(x^p_{t+r}; \theta^j_{t+r}, \varphi \mid h_{t+r}, s_{t+r}\right) \leq v_{t+r}\left(x^j_{t+r}; \theta^j_{t+r}, \varphi \mid h_{t+r}, s_{t+r}\right)$. Because v_{t+r} is continuous in θ_{t+r} by Lemma 2 (i), by the intermediate value theorem there exists $\theta^k_{t+r} \in [\theta^p_{t+r}, \theta^j_{t+r}]$ such that $v_{t+r}\left(x^p_{t+r}; \theta^k_{t+r}, \varphi \mid h_{t+r}, s_{t+r}\right) = v_{t+r}\left(x^j_{t+r}; \theta^k_{t+r}, \varphi \mid h_{t+r}, s_{t+r}\right)$. Thus, construct the partition $\{P^1_{t+r}, P^2_{t+r}\}$ such that all types $\theta^i_{t+r} \leq \theta^k_{t+r}$ are members of the right-populist party P^1_{t+r}, i.e. $P^1_{t+r} = \left\{\theta_{t+r} \in \Theta \mid \theta_{t+r} \leq \theta^k_{t+r}\right\}$ and $P^2_{t+r} = \Theta \setminus P^1_{t+r}$. The tuple $\left\langle \{P^1_{t+r}, P^2_{t+r}\}, A_{t+r}, x^p_{t+r}\right\rangle$ is a *SPO* because it satisfy all the conditions in Definition A.1. Consider a collective strategy s_t that consists in collectively playing a tuple of such kind in each period $t+r$ after any history h_t, and that satisfies the Markov property and subgame perfection. Then, s_t is a *MPPE* such that $\langle \mathbb{P}_{t+r}, A_{t+r}(\mathbb{P}_{t+r})\rangle$ is a populist vs. libertarian party system in each period $t+r$. This leads to a contradiction. In a similar way one can show that the same is true for a shock of type (b), (c), and/or (d). Q.E.D.

B.5 Welfare Analysis

Proposition 5. *For any Social Welfare Function $SWF\left((M_t, L_t); \varphi \mid h_t, s_t\right)$ that assigns a strictly positive weight to each native individual of working age, there exist thresholds $\breve{g}_t \in [0,1)$ and $\check{\Delta} > 0$ such that if $g_t \in [\breve{g}_t, 1)$ and $\Delta \in [0, \check{\Delta})$, then a marginal loosening in the immigration policy is welfare-enhancing.*

Proof. Suppose a marginal increase in M_t evaluated at M_t^* is not welfare-enhancing for some SWF with $\mu(\theta_t) > 0$ for all θ_t with $f_t\left(\theta_t; \rho \mid h_t\right) > 0$. First, notice that for $\Delta \leq \tilde{\Delta}$ the function SWF is differentiable in M_t because is the integral over functions v_{t+r} which are differentiable by Lemma 3. Thus, I define the marginal social welfare function as follows:

$$MSW\left((M_t^*, L_t^*); \varphi \mid h_t, s_t\right) := \left.\frac{\partial SWF\left((M_t, L_t^*); \varphi \mid h_t, s_t\right)}{\partial M_t}\right|_{M_t = M_t^*} \tag{37}$$

which exists and it is continuous and differentiable as a direct consequence of Lemma 3. Set $\Delta = 0$. Consider the effect of an increase in M_t on the induces utility of a young individual:

$$\left.\frac{\partial v_t\left((M_t, L_t^*); \theta_t^i, \varphi \mid h_t, s_t\right)}{\partial M_t}\right|_{M_t = M_t^*} = \psi \xi \theta_t^i + c'(M_t^*) \tag{38}$$

and for an old individual:

$$\left.\frac{\partial v_t\left((M_t, L_t^*); -1, \varphi \mid h_t, s_t\right)}{\partial M_t}\right|_{M_t = M_t^*} = c'(M_t^*) \tag{39}$$

Lastly, consider an individual born in period $t+r$ for $r > 0$. given $\Delta = 0$ I get:

$$\begin{gathered}\frac{\partial}{\partial M_t} E_t\left[v_{t+r}\left((M_{t+r}, L_{t+r}); \theta_{t+r}^i, \varphi \mid h_{t+r}, s_{t+r}\right) \mid h_t, s_t, (M_t, L_t^*)\right] = \\ \frac{dv_{t+r}\left((M_{t+r}^*, L_{t+r}^*); \theta_{t+r}^i, \varphi \mid h_{t+r}, s_{t+r}\right)}{dg_{t+r}} \left(\prod_{l=1}^{r} \frac{dg_{t+l}}{dg_{t+l-1}}\right) \frac{dg_{t+1}}{dM_t} = 0\end{gathered} \tag{40}$$

i.e. if $\Delta = 0$ current policy choices do not affect future outcomes. Thus, for $\Delta = 0$ I get:

$$MSW\left((M_t^*, L_t^*); \varphi \mid h_t, s_t\right) = \psi \xi \int_\theta \mu(\theta_t)\theta_t q(\theta_t, \rho) d\theta_t + \left[\int_\theta \mu(\theta_t) q(\theta_t, \rho) d\theta_t + \mu_t(-1)\right] c'(M_t^*) \tag{41}$$

Lastly, notice that if $g_t \to 1$ then $\lim_{g_t \to 1} \theta_t^p(g_t) = \lim_{g_t \to 1} Q^{-1}\left(0.5 - 0.5 g_t; \rho\right) = Q^{-1}\left(0; \rho\right) = 0$ and therefore:

$$\lim_{g_t \to 1} \frac{\partial v_t\left((M_t, L_t^*); \theta_t^p, \varphi \mid h_t, s_t\right)}{\partial M_t} = c'(M_t)$$

which implies that the ideal policy of the pivotal voter is $\lim_{g_t \to 1} \left\{\arg\max_{(M_t, L_t) \in X'} v_t\left((M_t, L_t); 0, \varphi \mid h_t, s_t\right)\right\} = (\underline{M}, L_t^*)$ where $\underline{M}$ solves $c'(\underline{M}) = 0$. Set $M_t^* = \underline{M}$ into (41) and take the limit for $g_t \to 1$ to get

$$\lim_{g_t \to 1} MSW\left((M_t^*, L_t^*); \varphi \mid h_t, s_t\right) = \psi\xi \int_{\theta} \mu_t(\theta_t)\theta_t q(\theta_t, \rho) d\theta_t > 0 \tag{42}$$

which is strictly positive for any weight function that satisfies $\mu_t(\theta_t) > 0$ for all $\theta_t \in \Theta$ with $f_t\left(\theta_t; \rho \mid h_t\right) > 0$. Because $MSW\left((M_t^*, L_t^*); \varphi \mid h_t, s_t\right)$ is jointly continuous in $(g_t, \varDelta)$ by Lemma 2, then either $MSW\left((M_t^*, L_t^*); \varphi \mid h_t, s_t\right) > 0$ for all $(g_t, \varDelta) \in [0,1] \times [0, \tilde{\varDelta})$, or the intermediate value theorem implies that there exists threshold $(\check{g}_t, \check{\varDelta})$ with $\check{g}_t \in [0,1)$ and $\check{\varDelta} > 0$ such that if $g_t \in [\check{g}_t, 1)$ and $\varDelta \in [0, \check{\varDelta})$, then $MSW\left((M_t^*, L_t^*); \varphi \mid h_t, s_t\right) > 0$. In turn, $MSW\left((M_t^*, L_t^*); \varphi \mid h_t, s_t\right) > 0$ implies that a marginal increase in M_t evaluated at M_t^* is strictly welfare-enhancing for any SWF that satisfies $\mu(\theta_t) > 0$ for all $\theta_t \in \Theta$ with $f_t\left(\theta_t; \rho \mid h_t\right) > 0$. This leads to a contradiction. Q.E.D.

References

[1] Acemoglu, D., Egorov, G., and Konstantin Sonin, K. (2013). "A political theory of populism," *Quarterly Journal of Economics*, 771–805.

[2] Bag P.K., Sabourian, H., and Winter, E. (2009). "Multi-stage voting, sequential elimination and Condorcet consistency," *Journal of Economic Theory,* 144, 1278–1299.

[3] Barone, G., D'Ignazio, A., de Blasio, G., and Naticchioni, P. (2016). "Mr. Rossi, Mr. Hu and politics. The role of immigration in shaping natives' voting behavior," *Journal of Public Economics*, 136, 1–13,

[4] Becker, S.O., and Fetzer, T. (2017). "Does Migration Cause Extreme Voting?," CAGE Online Working Paper Series 306, Competitive Advantage in the Global Economy (CAGE).

[5] Becker, S.O., Fetzer, T., and Novy, D. (2017). "Who voted for Brexit? A comprehensive district-level analysis," *Economic Policy*, 32(92), 601–650.

[6] Ben-Gad, M. (2012). "On deficit bias and immigration", City University London Department of Economics Discussion Paper Series No 12-09.

[7] Benhabib, J. (1996). "On the Political Economy of Immigration," *European Economic Review,* 40(9), 1737–1743.

[8] Besley, T., and Coate, S. (1997). "An Economic Model of Representative Democracy", *Quarterly Journal of Economics*, 108(1), 85–114.

[9] Black, D. (1948). "On the Rationale of Group Decision-Making," *Journal of Political Economy*, 56, 23–34.

[10] Boeri, T. and Brücker, H. (2005). "Why are Europeans so tough on migrants?," *Economic Policy*, 20(44), 629–703.

[11] Boeri, T. (2010). "Immigration to the land of redistribution," *Economica*, 77, 651–687.

[12] Bohn, H. and Lopez-Velasco, A.R. (2019). "Immigration and Demographics: Can High Immigrant Fertility Explain Voter support for Immigration?," *Macroeconomic Dynamics*, 23, 1815–1837.

[13] Boldrin, M. and Montes, A. (2005). "The Intergenerational State. Education and Pensions," *Review of Economic Studies*, 72, 651–664.

[14] Borge, L.-E. and Rattsø, J.J. (2004). "Income distribution and tax structure: Empirical test of the Meltzer-Richard hypothesis," *European Economic Review*, 48(4), 805–826.

[15] Borjas, G. (1999). 'Immigration and welfare magnets', *Journal of Labor Economics*, vol. 17(4), pp. 607–37.

[16] Card, D., Dustmann, C. and Preston, I. (2012). "Immigration, wages and compositional amenities", *Journal of the European Economic Association*, 10, 78–119.

[17] Curtice, J. (2012) "A Revolt on The Right? The social and political attitudes of UKIP supporters," *British Social Attitudes 32 – UKIP report*, 1–15.

[18] Dhillon, A. (2005). "Political Parties and Coalition Formation", Chapter in Group Formation in Economics; Networks, Clubs and Coalition, CUP, eds. G. Demange and M.Wooders, 2005.

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

[20] Dotti, V. (2019a). "Political Parties and Policy Outcomes: Do Parties Block Reforms?", Mimeo.

[21] Dotti, V. (2019b). "The Political Economy of Public Education," *Research in Economics*, 73, 35–52.

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

[23] Dustmann, C. and Preston, I. (2006). "Is immigration good or bad for the economy? Analysis of attitudinal responses", *Research in Labor Economics*, vol. 24, pp. 3-34.

[24] Dustmann, C. and Preston, I. (2007). "Racial and economic factors in attitudes to immigration", *Berkeley Electronic Journal of Economic Analysis and Policy*, vol. 7, Art. 62.

[25] Dustmann, C. and Frattini, T. (2014), "The Fiscal Effects of Immigration to the UK," *The Economic Journal*, 124; 593–643.

[26] Edo, A., Yvonne Giesing Y., Öztunc, J., and Poutvaara, P. "Immigration and electoral support for the far-left and the far-right," *European Economic Review,* 115, 99–143.

[27] Facchini, G. and Mayda, A.M., (2008). "From individual attitudes towards migrants to migration policy outcomes: Theory and evidence," *Economic Policy*,23, 651–713.

[28] Facchini, G. and Mayda, A.M., (2009). "Does the Welfare State Affect Individual Attitudes toward Immigrants? Evidence across Countries," *The Review of Economics and Statistics*, MIT Press, vol. 91(2), 295–314.

[29] Galasso, V. and Profeta, P. (2004). "Lessons for an ageing society: the political sustainability of social security systems," *Economic Policy*, CEPR;CES;MSH, 19(38), 63–115.

[30] Grandmont, J.-M. (1978). "Intermediate preferences and the majority rule", *Econometrica*, 46(2), 317–330.

[31] Guiso, L., Herrera, H., Morelli, M., and T. Sonno, T. (2019). "Populism: Demand and Supply," mimeo.

[32] Haupt, A. and Peters, W. (1998). "Public Pensions and Voting on Immigration," *Public Choice*, Springer, vol. 95(3-4), pages 403-13, June.

[33] Heckman, J. and Robb, R. (1985). "Using longitudinal data to estimate age, period and cohort effects in earnings equations." In Mason, W.M. & Fienberg, S.E. (eds.) Cohort Analysis in Social Research, New York, Springer, 137–150.

[34] HM Treasury. (2018). Public Expenditure Statistical Analysis. Available at: https://www.gov.uk/government/statistics/public-expenditure-statistical-analyses-2018. [Accessed: 12 February, 2020].

[35] Kemnitz, A. (2003). "Immigration, unemployment and pensions", *Scandinavian Journal of Economics*, vol. 105, pp. 31-47.

[36] Krieger, T. (2003). "Voting on low-skill immigration under different pension regimes", *Public Choice*, vol. 117, pp. 51-78.

[37] Krieger, T. (2005). Public Pensions and Immigration: A Public Choice Approach, Edward Elgar Publishing Ltd., Cheltenham and Northampton.

[38] Lee, R. and Miller, T. (2000). "Immigration, Social Security, and Broader Fiscal Impacts," *American Economic Review*, 90(2), 350–354.

[39] Leers, T., Meijdam, L. and Verbon, H.A.A. (2003). "Ageing, migration and endogenous public pensions", *Journal of Public Economics*, vol. 88, pp. 131-159.

[40] Levy, G. (2004). "A model of political parties, " *Journal of Economic Theory*, 115, 250-277.

[41] Levy, G. (2005). "The Politics of Public Provision of Education", *The Quarterly Journal of Economics*, Vol. 120, No. 4, pp. 1507-1534.

[42] Milgrom, P. and Shannon, C. (1994). "Monotone Comparative Statics", *Econometrica*, Vol. 62(1): 157-180.

[43] Milic, T. (2015). "For They Knew What They Did" – What Swiss Voters Did (Not) Know About The Mass Immigration Initiative. *Swiss Political Science Review*, 21, 48-62.

[44] Minkenberg, M. (2000). "The Renewal of the Radical Right: Between Modernity and Anti-modernity," *Government and Opposition*, 35(2), 170–188.

[45] Mudde, C. (2007). Populist Radical Right Parties in Europe. Cambridge University Press.

[46] Mudde, C. (2016). On Extremism and Democracy in Europe. London: Routledge.

[47] NatCen Social Research. (2019). British Social Attitudes Survey, 1995–2017. [Accessed: 12 February, 2020]. UK Data Service. SN: 8450, http://doi.org/10.5255/UKDA-SN-8450-1.

[48] Office for National Statistics (2019). Population estimates for the UK, England and Wales, Scotland and Northern Ireland: mid-2018. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/mid2018. [Accessed: 10 January, 2020].

[49] Orrenious, P. (2017). "New Findings on the Fiscal Impact of Immigration in the United States," Federal Reserve Bank of Dallas, Working Paper 17-13.

[50] Ortega, F. (2005). "Immigration quotas and skill upgrading," *Journal of Public Economics,* 89 (2005) 1841 – 1863.

[51] Ortega, F. and Peri, G. (2009). "The Causes and Effects of International Migrations: Evidence from OECD Countries 1980-2005," NBER Working Papers 14833, National Bureau of Economic Research, Inc.

[52] Osborne, M.J. and Slivinski, A. (1996). "A Model of Political Competition with Citizen-Candidates", *The Quarterly Journal of Economics*, Vol. 111(1), 65-96.

[53] Peri, G., and Sparber, C. (2009). 'Task Specialization, Immigration, and Wages', *American Economic Journal: Applied Economics*, 1(3), 135–69.

[54] Plott, C.R. (1967). "A notion of equilibrium and its possibility under majority rule", *American Economic Review*, 57 (Sept.): 787-806.

[55] Preston, I. (2014), "The Effect of Immigration on Public Finances," *The Economic Journal*, 124: F569–F592.

[56] Ray, D. and Vohra, R. (1997). "Equilibrium Binding Agreements", *Journal of Economic Theory*, 73, 30-78.

[57] Razin, A. and Sadka, E. (1999). "Migration and Pension with International Capital Mobility," *Journal of Public Economics* 74(1): 141-150.

[58] Razin, A. and Sadka, E. (2000). “Unskilled Migration: A Burden or a Boon for the Welfare State,” *Scandinavian Journal of Economics* 102(3): 463-479.

[59] Razin, A., Sadka, E. and Swagel, P. (2002). “The ageing Population and the Size of the Welfare State,” *Journal of Political Economy*, University of Chicago Press, vol. 110(4), pages 900-918, August.

[60] Razin, A., Sadka, E. and Suwankiri, B. (2014). “The Welfare State and Migration: a Dynamic Analysis of Political Coalitions”, NBER Working Paper Series, WP 20806.

[61] Scholten, U. and Thum, M. (1996). “Public pensions and immigration policy in a democracy”, *Public Choice*, vol. 87, pp. 347-361.

[62] United Nations (2000). Replacement migration: Is it a solution to declining and ageing populations? New York: UN Population.

[63] Van der Brug, W., Fennema, M., and Tillie, J. (2000). “Anti-Immigrant Parties in Europe: Ideological or Protest Vote?” *European Journal of Political Research*, 37 77–102.