NBER WORKING PAPER SERIES

MIGRATION AND THE WELFARE STATE: DYNAMIC POLITICAL-ECONOMY THEORY

Assaf Razin
Efraim Sadka
Benjarong Suwankiri

Working Paper 14784
http://www.nber.org/papers/w14784

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2009

Migration and the welfare state: Dynamic Political-Economy Theory
Assaf Razin, Efraim Sadka, and Benjarong Suwankiri
NBER Working Paper No. 14784
March 2009
JEL No. E0,F2,H11

ABSTRACT

Milton Friedman, the Nobel-prize laureate economist, had it right: "It's just obvious that you can't have free immigration and a welfare state." That is, national welfare states can almost never coexist with the free movement of labor. This fact underscores the relevance of the analysis in this paper, which is a part of a forthcoming book on migration and the welfare state. It focuses on the demographic, and economic, fundamentals behind policy-restricted migration, and the policy-restricted generosity of the welfare state.

Assaf Razin
Department of Economics
Cornell University
Uris 422
Ithaca, NY 14853
and NBER
ar256@cornell.edu

Benjarong Suwankiri
Department of Economics
Cornell University
Ithaca, NY
14853
bs246@cornell.edu

Efraim Sadka
Tel Aviv University
Eitan Berglas School of Economics
P.O.B. 39040
Ramat Aviv, Tel Aviv, 69978, ISRAEL
sadka@post.tau.ac.il

Migration and the Welfare State: Dynamic Political-Economy Theory

Assaf Razin, Cornell University and Tel-Aviv University
Efraim Sadka, Tel-Aviv University
Benjarong Suwankiri, Cornell University

March, 2009

Abstract

Milton Friedman, the Nobel-prize laureate economist, had it right: "It's just obvious that you can't have free immigration and a welfare state." That is, national welfare states can almost never coexist with the free movement of labor. This fact underscores the relevance of the analysis in this paper, which is a part of a forthcoming book on migration and the welfare state. It focuses on the demographic, and economic, fundamentals behind policy-restricted (political-economy based) migration, and the policy-restricted (political-economy based) generosity of the welfare state.

1 Introduction

All over the world, the combination of declining population growth rates and rising life expectancy presents a major fiscal challenge to social security systems. From an economic perspective, a rise in the dependency ratio (i.e., the proportion of retirees per worker) increases the number of people drawing from the system; while it decreases the number of contributors. From a political perspective, the older is the decisive voter, the more relevant is the pension spending in the political agenda. One of the policy tools that are considered for mitigating these politico-economic forces which result in higher demand for, and lower supply of, social security benefits is migration policy.

The view that increased migration may come to the rescue of PAYG social security systems reflects the fact that the flow of migrants can alleviate the current demographic imbalance, by influencing the age structure of the host economy. A few empirical studies address this point by calibrating the equilibrium impact of a less restrictive policy towards migration according to U.S. data. Storesletten (2000) finds in a general equilibrium model that selective migration policies, involving increased inflow of working-age high and medium-skilled migrants, can remove the need for a future fiscal reform. By emphasizing the demographic side and abstracting from the migrants' factor prices effects, Lee and Miller (2000) conclude in a similar analysis that a higher number of migrants admitted into the economy can ease temporarily the projected fiscal burden of retiring baby boomers.

This paper combines two fields of the existing political economy literature, which have not been examined jointly, to our knowledge: the political economy of the PAYG social security systems (Cooley and Soares (1999), Bohn (2005), Boldrin and Rustichini (2000), Galasso (1999)) and the political economy of migration (Benhabib (1997)). There are also a few studies which deal with the effect of migrants on the PAYG social security system (Razin and Sadka (1999) and Scholten and Thum (1996)). This paper addresses the joint political economy decisions regarding both migration policy and social security policy in a dynamic set-up.

The paper, a part of a forthcoming book, develops a dynamic politico-economic model, in which both migration and taxes interact, focusing on inter- and intra-generational aspect of social security. The model is based on key demographic characteristics: that migrants are younger and have higher birth rates than the native born population. To isolate the inter-generational aspects, we abstract in this chapter from intra-generational income transfers considerations. (These considerations are brought up in subsequent chapters.) A standard dynamic equilibrium concept is employed in which migration policy and pay-as-you-go (PAYG) social security system are jointly determined through a majority voting process.

2 Background:Migration and Intergenerational Distribution Policy

We briefly describe the model of inter-generational distribution policy and migration is developed in Sand and Razin (2008). A perishable consumption good is produced using only labor as input; transfers from young to old (paid by flat tax rate on labor income) are an important supplement for private savings guaranteeing old-age consumption. Each generational cohort lives two periods, supplying labor elastically when young, and deriving utility from consumption in both periods of life.

If there were not to be migration, it is a standard outcome in this framework that if the population growth rate is positive, the young always outnumber the old. Therefore, a pay-as-you-go social security system cannot be sustained under majority voting. If, however, population growth is negative, so that the old outnumber the young, then the pay-as-you-go system can be sustained with a constant tax rate the maximizes the social security benefits (the preferred point of old cohort at each period). Now, introduce migration into the standard framework. Migrants arrive young but cannot vote until they are old. Their children, who are identical to the young native-born, can vote when young. Moreover, migrants (though not their offspring) have a birth rate that is larger than the native-born rate. Migration policy is described by an endogenously determined quota variable. The central tension faced by today's young in thinking about migration policy is that both the ratio of young to old in the next period, and the ratio of taxpayers to old dependents in the next period increase when the present period migration quota rises. A higher value of the latter this period will raise the number of young taxpayers per old dependent next period, but will also increase the voting power of the young next period, perhaps putting them in the majority. If the native born and the migrants' population growth rates are positive (while by assumption the latter rate exceeds the former), then young voters always outnumber old voters and the pay-as-you-go social security system will not be sustainable as a Markov equilibrium. So migration is of no help in this case. On the other hand, if the native-born population growth rate is negative, then the social security system is sustainable in the absence of migration. In this case, the quest is not whether migration helps sustain social security, but whether it threatens its sustainability. Assuming that the population growth rate of the native-born is negative, the sort of equilibrium that arises depends on the sum of native-born and migrants' population growth rates. If this sum is negative, admitting no migrants today guarantees an old majority tomorrow. Even if the current young chooses the maximum allowable migration so as to maximize next period's benefits, there will still be a majority of the old in the next period. Both the current old and the current young agree on letting in the maximal number of migrants, an except perhaps for the initial period, the majority of voters will always be old. Therefore, the tax rate is set at the "Laffer" rate. Migration does not yet add (nor subtract) much to the survival of the social security system in this case.

But when the sum of the native-born and the migrants population growth rates is positive and the native-born population growth rate is negative, migration adds an interesting twist. In essence, it poses a threat to social security that in the absence of migration will be assured. In this case, the numbers of old and young next period are equal and by assumption, ties are decided in favor of the old. Then current young's desire for higher migration, to maximize their old-age benefits is constrained by their desire to maintain an old majority next period. If the young are currently in the majority, they set the current tax rate equal to zero (implying no benefits for the current old), and set migration quota at an intermediate level that barely preserves the old majority in the next period. In the next period, the old median voter sets the tax rate at the "Laffer" rate and the migration quota at the maximum level. The latter guarantees that the young will be in majority in subsequent period; and the cycle repeats itself.

From this benchmark model, Sand and Razin (2008) develops a model which also includes capital accumulation and endogenous factor prices. The extended model has an additional demographic-steady equilibrium, where the young is steadily the median voter. Most importantly, the young does set the social security tax to a positive level, and thus sustains the social security system. As in Forni (2005), in the case of a positive native-born population growth rate, when the young are always in the majority, a pay-as-you-go social security system is sustained by a tax rate on labor income which varies with the level of the capital stock (a second state variable). Specifically, the tax rate on labor income is decreasing in the capital stock. In the case in which the population growth rates of the native-born and the migrants' are positive ($n, m > 0$), the number of next period young voters exceeds the number of next period old voters, which means that the decisive voter is always young. Still, if the capital per the native-born workforce is in some range, then the optimal strategy of the young is always to vote for a positive tax rate, and maximum migration quota, thus sustaining both migration and the social security system. The size of the social security system depends on the capital per native-born worker, and on the exogenously given ceiling on migration quota. Thus the polico-economic sustainable migration boosts up the tax base for financing the social security.

3 Elements of Strategic Voting with Multiple Groups

The initial motivation for our politico-economic setup is the class of models with citizen-candidate structure. Before the introduction of the citizen-candidate structure, earlier models in the fields of public choice and political economics utilize heavily the Downsian candidate setup that leads to the result of platform convergence of the candidates (Downs (1957)). The model assumes purely office-motivated candidates competing for a single office post. The competition to win the election will drive the policy platforms of all the candidates to the bliss point of the median voters, trying to attract as many votes as possible.[1] Thus the campaign among the candidates boils down to pursuing what drives the preference of the median voter and what may shift the distribution of voters. Moreover, the complete convergence in platforms does not seem to be observed in practice in most elections. Furthermore, candidates must arise from the citizen body and citizens are presumed to have some preferences for the policy chosen, regardless of the number of voters. Hence, assuming that candidates are only office-motivated misses out key policy determinants of voting models. The citizen-candidate model stands on the other end of the spectrum. First studied by Osborne and Slivinski (1996) and Besley and Coate (1997), the citizen-candidate model seeks to

[1]The politico-economic models we employed in the preceding chapters were in this spirit too.

endogenize the candidates' selection from within the body of the citizens, and how the policy is ultimately determined.

However, due to the richness of strategic choices in the model, the citizen-candidate model is not easily applicable for applied research. In particular, the model suffers from massive multiplicity of equilibria, even in a static setting. For those seeking a dynamic politico-economic framework, the citizen-candidate proves formidable. In a subsequent work, Besley and Coate (1998) have extended the static model to a two-period setting. Anything beyond two-period must face exponentiated complexity. All in all, the citizen-candidate model is appropriate for an analysis focussing on a small-scale election, and possibly static. Therefore, it remains just a motivation for our exposition in this chapter, as we have adapted the model into an easily applicable version.

3.1 Many candidates

Consider an economy with a continuum of citizens, normalizing the population size to a unit. The citizens are divided into N groups, indexed by $i \in \{1, 2, \ldots, N\}$, and each has a mass of $\omega_i \geq 0$, where $\sum_{i=1}^{N} \omega_i = 1$. We imagine N to be relatively small. This means that, with a large population, people with similar interests often get grouped together. This setup abstracts from the possibility that one individual may belong to more than one group, sharing many interests.[2]

To highlight the mechanics of the model, suppose that the voters must collectively choose a one-dimensional policy (that is, $p \in P = \mathbb{R}$).[3] We assume that any two citizens belonging to the same group will have identical

[2]This shortfall, nonetheless, is common even in literature concerning itself primarily with interest groups' influence.

[3]Besley and Coate (1997) studies a more general environment with possible multi-dimensional policy space.

preference over the policy. The representative citizen from group i has a preference defined over the policy space, represented by the utility function $v^i(p)$. These preferences are "singled-peaked" and we let p_i^* denotes group i's preferred policy.

We assume that there are N candidates running for office representing directly the interest of the group they belong to. We denote with $j \in \{1, \ldots, N\}$ the identity of the candidates. This is fully known to all voters. Only one candidate is present from each group. We assume that, if the candidate representing group j wins the election, the implemented policy will be p_j^*. Under plurality rule, candidates who receive the most votes win.

Each citizen has a single vote that can be cast for a candidate. In particular, because voters from the same group have identical preference, they will vote identically.[4]

Let $e^i \in \{1, \ldots, N\}$ denote the vote casted by voters of group i. How each chooses to vote depends on her preference and what we allow them to consider while voting. We consider two canonical voting behaviors: *sincere* and *strategic*.

3.2 Sincere Voting

Voting sincerely is the simpler of the two. Under sincere voting behavior, voters will vote for candidates $j \in \{1, \ldots, N\}$ whose policy platform maximizes their utility, that is

$$\widetilde{e}^{i*} = \arg\max \left\{ v^i \left(p_j^* \right) \mid e^i \in \{1, \ldots, N\} \right\}.$$

We can denote the voting vector as $\widetilde{\mathbf{e}}^* = \left(\widetilde{e}^{1*}, \ldots, \widetilde{e}^{N*} \right)$. Under this voting behavior, voters belonging to group i will vote for candidate representing their

[4]We allow no abstentions within the model. Abstention can be built directly into voting choices. Depending on the context, however, it may appear unrealistic because, if one voter from a group abstains, all members of the same group must accordingly abstain.

group. That is $\widetilde{e}^{i*} = i$. The winner of the election will be decided purely by the size of the groups. Under plurality rule, the winning candidate will come from the group with the largest size, as reflected by ω_i. In the special case with two groups ($N = 2$), then the winning candidate will be represent the median voter of the economy. However, as N gets larger, it is no longer the case that the winning candidate will represent the preference of the median voter. When there are more fractions in the economy, and no collusion is allowed (that is, assuming everyone votes sincerely), the preference of the largest group in the economy will dictate the implemented policy.

3.3 Strategic Voting

Strategic voting relaxes the assumption of sincere voting. People are no longer required to vote for the candidate they like most, but rather they take into account the probability of that candidate winning the election. A voter is said to be voting *strategically* if she votes for the candidate with a policy platform that maximizes her.expected utility, where the expectation is taken over all the candidates and their probability of winning the election. Moreover, the votes must be consistent with the induced probability of winning of each candidate. Formally, voting decisions $\mathbf{e}^* = (e^{1*}, \ldots, e^{N*})$ form a *voting equilibrium*[5] if

$$e^{i*} = \arg\max \left\{ \sum_{j=1}^{N} \mathcal{P}^j(e^i, \mathbf{e}^*_{-i}) v^i \left(p^*_j\right) \mid e^i \in \{1, \ldots, N\} \right\}$$

for $i \in \{1, \ldots, N\}$, where $\mathcal{P}^j(e^i, \mathbf{e}^*_{-i})$ denotes the probability that candidate $j \in \{1, \ldots, N\}$ will win given the voting decisions, and $\mathbf{e}^*_{-i}$ is the optimal voting decisions of other groups that is not i. Thus we also require that each vote cast by each group is a best-response to the votes by the other groups. In addition, this also means that the representative voter of each group must

[5]The original definition of this voting equilibrium is due to Besley and Coate (1997).

take into the account the *pivotal* power of her vote, because the entire group will also vote accordingly. After the election, the votes are tallied by adding up the size of each group that have chosen to vote for the candidate. The candidate with the most votes wins the election and gets to implement her ideal set of policies. The winning probability quantity, $\mathcal{P}^j(e^i, \mathbf{e}^*_{-i})$, must be determined endogenously from the voting vector and the groups' weight. Lastly, we define a *political equilibrium* to consists of two vectors, $\mathbf{e}^*$ and $\mathbf{p}^*$, where the latter is the vector listing the policies preferred by every candidate.

It is important to contrast the strategic voting scenario with the sincere counterpart. We do this by a couple of examples, which will also demonstrate how the probability a candidate would win is determined, $\mathcal{P}^j(\mathbf{e}^*)$. Under sincere voting, voters assume that the policy of their most-preferred candidate will be implemented with probability one, while under strategic voting, the probability depends on how other groups vote. A special case arises when a certain group form more than 50% of the population. In this case, the winning candidate, who will also represent the preference of the median, will belong to this group, irrespective of the voting profiles of the other groups. Therefore, the probability that its candidate will win is 1. One can easily construct other examples with different conclusions. For example, let $N = 3$, and $\omega_i = \frac{1}{4}, \frac{1}{3}, \frac{5}{12}$ for $i = 1, 2, 3$ respectively. No one group consists of more than 50% of the population; group 3 is the largest. However, if group 1 and 2 both dislike the policy preferred by group 3, they could collude to surpass 50% and win the election. The implemented policies will be decided by the voting equilibrium. If collusion means voters from group 1 and group 2 both vote from group 2's candidate, the ideal policy of group 2 will be implemented in equilibrium. The probability of winning for candidates representing group 1 and 2 are $\mathcal{P}^1(\mathbf{e}^*) = 0$ and $\mathcal{P}^2(\mathbf{e}^*) = 1$. Likewise, group 1 and 2 could both vote for group 1's representative candidate, hence resulting in policy preferred by group 1 in equilibrium. In this case, the probability of winning for candidates representing group 1 and 2 are reversed $\mathcal{P}^1(\mathbf{e}^*) = 1$ and $\mathcal{P}^2(\mathbf{e}^*) = 0$. By either collusions, the preferred policy of the largest group, group 3, will be blocked in equilibrium. These two voting equilibrium will generate $\mathcal{P}^3(\mathbf{e}^*) = 0$.

Note that a rule for a tie breaker should be defined. That is, if two candidates receive the same amount of votes, how will this be resolved. Besley and Coate (1997) proposes equal probability across all leading candidates. Alternatively, one can also assign some other arbitrary rules, such as the candidate belonging the larger group always win or the candidate with a smaller group index wins. Whichever rule one chooses, it should complement the analysis underlying the usage of the model.

4 Migration, *Inter-* and *Intra*-generational Redistribution

We employ a two-period, overlapping-generations model. The old cohort retires, while the young cohort works. There are two skill levels: skilled and unskilled. The welfare-state is modeled simply as in Part I of the book, by a proportional tax on labor income to finance a demogrant in a balanced-budget manner.[6] Therefore, some (the unskilled workers and old retirees) are net beneficiaries from the welfare state and others (the skilled workers) are net contributors to it. Migration policies are set to determine the total migration volume and its skill composition. As in Chapter 5, we characterize subgame-perfect Markov politico-economic equilibria consisting of the tax rate (which determines the demogrant), skill composition and the the total number of migrants. We distinguish between two voting behaviors: sincere and strategic voting (see Chapter 6). As illustrated in that chapter, when

[6] We draw heavily on Suwankiri (2009).

participating in political decisions, as we indeed have, sincere voting is too simplistic. We therefore study also the case of strategic voting among the native-born in order to enable the formation of strategic political coalitions.

4.1 Analytical Framework

Consider an economy consisting of overlapping generations. Each individual lives for two periods, working in the first period when young, and retiring in the second period when old. The population is divided into two groups according to their exogenously given skills: skilled (s) and unskilled (u).

4.2 Preferences and Technology

The utility of each individual in period t, for young and old, is given, respectively, by

$$U^y(c_t^y, l_t^i, c_{t+1}^o) = c_t^y - \frac{\varepsilon (l_t^i)^{\frac{1+\varepsilon}{\varepsilon}}}{1+\varepsilon} + \beta c_{t+1}^o, \; i = s, u \tag{1}$$

$$U^o(c_t^o) = c_t^o. \tag{2}$$

where, as in Part I, s and u denote skilled and unskilled labor. Here, y and o denote to young and old, l^i is labor, ε is the elasticity of the labor supply, and $\beta \in (0,1)$ is the discount factor.[7] Note that c_t^o is the consumption of an old individual at period t (who was born in period $t-1$). Agents in the economy maximize the above utility functions subject to their respective budget constraints. Given the linearity of U in c_t and c_{t+1}, a non-corner solution can be attained on ly when $1 = \beta(1+r)$, where r is the interest rate. We indeed assume that the interest rate r equal $\frac{1}{\beta} - 1$ and individuals have no incentive to either save or dissave. Fore simplicity, we set saving at zero.[8]

[7]This functional form of U^y is similar to the one used in Part I.

[8]In fact, any saving level is an optimal choice. Assuming no saving is for pure convenience. With saving, since old individuals do not work the last period of their life, they

This essentially reduces the two groups of old retirees (skilled and unskilled) to just one because they have identical preference irrespective of their skill level. In addition to consumption, the young also decide on how much labor to supply. Individual's labor supply is given by

$$l_t^i = \left(A_t w^i (1-\tau)\right)^{\varepsilon}, \ i = s, u \tag{3}$$

where w^i is the wage rate of a worker of skill level $i = s, u$.

There is just one good, which is produced by using the two types of labor as perfect substitute.[9] The production function is given by

$$Y_t = w^s L_t^s + w^u L_t^u \tag{4}$$

where L_t^i is the aggregate labor supply of skill $i = s, u$. Labor markets are competitive, ensuring the wages going to the skilled and unskilled workers are indeed equal to their marginal products, w^s and w^u, respectively. We naturally assume that $w^s > w^u$.

As before, we denote the demogrant by b_t and the tax rate by τ_t. The agents in the economy take these policy variables as given when maximizing their utilities. Because the old generation has no income, its only source of income comes from the demogrant. The model yields the following indirect utility function (recall that saving is zero):

$$\begin{aligned} V^{y,i} &= \frac{\left((1-\tau_t)w^i\right)^{1+\varepsilon}}{1+\varepsilon} + b_t + \beta b_{t+1} \\ V^o &= b_t, \end{aligned}$$

will consume savings plus any transfer. Through both these channels, the old individuals benefit from migration. To keep the analysis short, we will just focus on the costs and benefits in terms of the welfare state.

[9] This simplification, nonetheless, allows us to focus solely on the linkages between the welfare state and migration, leaving aside any labor market consideration. In Appendix 7A.1, we consider the case where the two inputs are not perfect substitute.

for $i \in \{s, u\}$. For brevity, we will use V^i to denote $V^{y,i}$ because only the young workers need to be distinguished by their skill level.

In addition to the parameters of the welfare state (τ_t and, consequently, b_t), the political process also determines migration policy. This policy consists of two parts: one determining the volume of migration, and the other its skill composition. We denote by μ_t the ratio of allowed migrants to the native-born young population and denote by σ_t the fraction of skilled migrants in the the total number of migrant entering the country in period t.

Migrants are assumed to have identical preference to the native-born. As before, we assume all migrants come young and they are naturalized one period after their entrance. Hence, they gain voting rights when they are old, as in the intergenerational model of chapter 5.

As in chapters 2 and 3, let s_t denote the fraction of native-born skilled workers in the labor force in period t (where $s_0 > 0$). The aggregate labor supply in the economy of each type of labor is given by

$$L_t^s = [s_t + \sigma_t \mu_t] \, N_t l_t^s \tag{5}$$

and

$$L_t^u = [1 - s_t + (1 - \sigma_t)\mu_t] \, N_t l_t^u, \tag{6}$$

where N_t is the number of native-born young individuals in period t.

4.3 Dynamics

The dynamics of the economy are given by two dynamic equations: one governs the *aggregate* population, while the other governs the *skill* composition dynamics. Because skills are not endogeneous within the model, we assume for simplicity that the offspring replicate exactly the skill level of their parents.[10] That is,

$$N_{t+1} = [1 + n + (1+m)\mu_t]\, N_t \tag{7}$$
$$s_{t+1}N_{t+1} = [(1+n)s_t + (1+m)\sigma_t\mu_t]\, N_t,$$

where n and m are the population growth rates of the native-born population and the migrants, respectively. As in chapter 5, we plausibly assume that $n < m$, and we allow the population growth rates to be negative. Combining the two equations in (7) together, we get the dynamics of the labor supply of skilled native-born as follows:

$$s_{t+1} = \frac{(1+n)s_t + (1+m)\sigma_t\mu_t}{1 + n + (1+m)\mu_t}. \tag{8}$$

Equation (8) implies that the fraction of the native-born skilled in the native-born labor force will be higher in period $t+1$ than in period t if the proportion of skilled migrants in period t is higher than that of the native-born, that is, if $\sigma_t > s_t$. Naturally, when there is no migration the share of skilled workers out of (native-born) young population does not change over time, by assumption. When migration is allowed and its share of skilled labor is larger than that of the native-born, the share of skilled labor in the population will grow over time.

4.4 The Welfare-State System

As before, we model the welfare-state system as balanced period-by-period. In essence, it operates like a pay-as-you-go system. The proceeds from the labor tax of rate τ_t in period t serve entirely to finance the demogrant b_t in

[10]Razin, Sadka, and Swagel (2002a, 2002b) and Casarico and Devillanova (2003) provide a synthesis with endogeneous skill analysis. The first work focuses on the shift in skill distribution of current population, while the latter studies skill-upgrading of future population.

the same period. Therefore, the equation for the demogrant, b_t, is given by

$$b_t = \frac{\tau_t\left((s_t + \sigma_t\mu_t)w^s N_t l_t^s + (1 - s_t + (1-\sigma_t)\mu_t)\, w^u N_t l_t^u\right)}{(1+\mu_t)\, N_t + (1+\mu_{t-1})\, N_{t-1}}, \tag{9}$$

which upon some manipulation reduces to

$$b_t = \frac{\tau_t\left((s_t + \sigma_t\mu_t)w^s l_t^s + (1 - s_t + (1-\sigma_t)\mu_t)\, w^u l_t^u\right)}{1 + \mu_t + \frac{1+\mu_{t-1}}{1+n+\mu_{t-1}(1+m)}}, \tag{10}$$

where the individual's labor supplies are given above in equation (3). It is straightforward to see that a larger σ_t increases the demogrant (recall that $w^s l_t^s > w^u l_t^u$). That is, a higher skill composition of migrants brings about higher tax revenues, and, consequently, enables more generous welfare state, other things being equal. Similarly, upon differentiation of b_t with respect to μ_t, we can conclude that a higher volume of migration enables a more generous welfare system if the share of the skilled among the migrants exceeds the share of the skilled among the native-born workers ($\sigma_t > s_t$).

4.5 Political Economy Equilibrium: Sincere Voting

In this section, we study the politico-economic equilibrium in the model. We imagine the economy with three candidates representing each group of voters. In the text, we discuss only the equilibrium with sincere voting. In appendix 7A, we consider the equilibrium with strategic voting.

We focus on "sincere voting," where individuals vote according to their *sincere* preference irrespective of what the final outcome of the political process will be; see chapter 6. In this case, the outcome of the voting is determined by the largest voting group.[11] Therefore, it is important to see who forms the largest voting group in the economy and under what conditions. Note that there are only three voting groups: the skilled native-born young,

[11]Evidently, this assumption amounts to majority voting when there are only two voting groups.

the unskilled native-born young, and the old (recall that there is no saving, so that all the old care only about the size of the demogrant and thus have identical interest.

1. The group of skilled native-born workers is the largest group ("the skilled group") under two conditions. First, its size must dominates the unskilled young, and, second, it must also dominate the old cohort. Algebraically, these are
$$s_t > \frac{1}{2} \tag{11}$$
and
$$s_t > \frac{1 + \mu_{t-1}}{1 + n + \mu_{t-1}(1 + m)} \tag{12}$$
, respectively. It can be shown that, because $n < m \leq 1$, only the second of the two conditions is sufficient.

2. The group of unskilled native-born workers is the largest group ("the unskilled group") under two similar conditions; that are reduced to just one:
$$1 - s_t > \frac{1 + \mu_{t-1}}{1 + n + \mu_{t-1}(1 + m)}. \tag{13}$$

3. The group of old retirees is the largest group ("the old group"), when its size is larger than each one of the former groups, that is,
$$\frac{1 + \mu_{t-1}}{1 + n + \mu_{t-1}(1 + m)} \geq \max\{s_t, 1 - s_t\}. \tag{14}$$

4.6 Equilibrium

We first describe what are the variables relevant for each of the three types of voters when casting the vote in period t. First, s_t is the variable which describes the state of the economy. Also, each voter takes into account how her choice of the policy variables in period t will affect the chosen policy variables in period $t+1$ which depends on s_{t+1} (recall that the benefit she will get in period $t+1$, b_{t+1}, depends on τ_{t+1}, σ_{t+1}, and μ_{t+1}). Therefore each voter will cast her vote on the set of policy variables τ_t, σ_t, and μ_t which maximizes her utility given the values of s_t, taking also into account how this will affect s_{t+1}. Thus, there is a link between the policy chosen in period t to the one chosen in period $t+1$. The outcome of the voting is the triplet of the policy variables most preferred by the largest voting group.

The mechanism (policy rule or function) that characterizes the choice of the policy variables (τ_t, σ_t, and μ_t) is invariant over time. This mechanism relates the choice in any period to the choice of the preceding period (τ_{t-1}, σ_{t-1}, and μ_{t-1}). This choice depend also on the current state of the economy, s_t. Thus, we are looking for a triplet policy function $(\tau_t, \sigma_t, \mu_t) = \Phi(s_t, \tau_{t-1}, \sigma_{t-1}, \mu_{t-1})$, which is a solution to the following functional equation

$$\Phi(s_t, \tau_{t-1}, \sigma_{t-1}, \mu_{t-1}) = \underset{\tau_t, \sigma_t, \mu_t}{\arg\max} V^d \left\{s_t, \tau_t, \sigma_t, \mu_t, \Phi(s_{t+1}, \tau_t, \sigma_t, \mu_t)\right\} \tag{15}$$
$$\text{s.t. } s_{t+1} = \frac{(1+n)s_t + (1+m)\sigma_t\mu_t}{1+n+(1+m)\mu_t},$$

where V^d is defined in equations (7.5) and (7.11), and $d \in \{s, u, o\}$ is the identity of the largest voting group in the economy.

This equation states that the decisive (largest) group in period t chooses, given the state of the economy s_t, the most preferred policy variables τ_t, σ_t, and μ_t. In doing so, this group realizes that her utility is affected not only by these (current) variables, but also the policy variables of the next period $(\tau_{t+1}, \sigma_{t+1}, \mu_{t+1})$. This group further realizes that the future policy variables are affected by the current variables according to the policy function $\Phi(s_{t+1}, \tau_t, \sigma_t, \mu_t)$. Furthermore, this intertemporal functional relationship between the policy variables in periods $t+1$ and t is the same as the one existed between period t and $t-1$. Put differently, what the decisive group in period t chooses is related to $s_t, \tau_{t-1}, \sigma_{t-1}$, and μ_{t-1} in exactly the same way (through $\Phi(\cdot)$) as what the decisive group in period $t+1$ is expected to be related to $s_{t+1}, \tau_t, \sigma_t$, and μ_t.

Denoting the policy function, $\Phi(s_t, \tau_{t-1}, \sigma_{t-1}, \mu_{t-1})$, by $(\tau_t, \sigma_t, \mu_t)$, we can show that the outcomes of the policy rule are:

$$\tau_t = \begin{cases} 0 & \text{, if the skilled group is the largest} \\ \frac{1-\frac{1}{J}}{1+\varepsilon-\frac{1}{J}} & \text{, if the unskilled group is the largest} \\ \frac{1}{1+\varepsilon} & \text{, if the old group is the largest} \end{cases}$$

$$\sigma_t = \begin{cases} 1 & \text{, if either the skilled or unskilled group is the largest and } s_t < \frac{1}{1+n} \\ \widehat{\sigma} < \frac{1}{2} & \text{, if the skilled group is the largest and } s_t \geq \frac{1}{1+n} \\ 1 & \text{, if the old group is the largest.} \end{cases} \tag{16}$$

$$\mu_t = \begin{cases} \frac{1-(1+n)s_t}{m} & \text{, if the unskilled group is the largest and } \Psi > 0 \text{ or if the skilled group is the largest and } s_t < \frac{1}{1+n} \\ \widehat{\mu} < 1 & \text{, if the skilled group is the largest and } s_t \geq \frac{1}{1+n} \\ 1 & \text{, if the unskilled group is the largest and } \Psi \leq 0 \text{ or if the old group is the largest.} \end{cases}$$

where

$$J = \frac{(s_t + \sigma_t \mu_t)\left(\frac{w_t^s}{w_t^u}\right)^{1+\varepsilon} + 1 - s_t + (1-\sigma_t)\mu_t}{1 + \mu_t + \frac{1+\mu_{t-1}}{1+n+\mu_{t-1}(1+m)}} \tag{17}$$

$$\Psi = b_t^u + \beta b_{t+1}^o - \widehat{b}_t, \tag{18}$$

where we denote by $\widehat{b}_t$ the demogrant period t with $\mu_t = 1 = \sigma_t$, and b_t^u the demogrant in period t with $\sigma_t = 1$ and $\mu_t = \frac{1-(1+n)s_t}{m}$ (both demogrants are associated with the tax rate preferred by the unskilled group). Similarly, b_{t+1}^o is the demogrant in period $t+1$ associated with the set of policy variables preferred by the old group.

Notice that the case $s_t > \frac{1}{1+n}$ cannot happen if the unskilled group is the largest (because $n < 1$). In this case, the special migration policy variables preferred by the skilled group, $\widehat{\sigma}$, and $\widehat{\mu}$, are given implicitly from the maximization exercise

$$\langle\widehat{\sigma},\widehat{\mu}\rangle = \arg\max_{\sigma_t,\mu_t} V_t^s = \frac{\left(A_t w_t^s\right)^{1+\varepsilon}}{1+\varepsilon} + \beta b_{t+1}^o \tag{19}$$
$$\text{s. t.} \qquad (1+n)s_t - 1 \leq \mu_t(1-(1+m)\sigma_t).$$

When the solution to the problem in (19) is interior, we can describe it by

$$\frac{\frac{\partial V^s}{\partial \sigma_t}}{\frac{\partial V^s}{\partial \mu_t}} = \frac{\widehat{\mu}(1+m)}{(1+m)\widehat{\sigma}-1}. \tag{20}$$

There are also two possible corner solutions: $\langle\widehat{\sigma},\widehat{\mu}\rangle = \langle 0,(1+n)s_t-1\rangle$ and $\langle\widehat{\sigma},\widehat{\mu}\rangle = \left\langle \frac{2-(1+n)s_t}{1+m}, 1\right\rangle$.

4.7 Interpretation: Migration and Tax Policies

The intuition for the aforementioned results is as follows. The skilled are the net contributor to the welfare state, while the other two groups are net beneficiaries. Preferences of the old retirees are simple. If the old cohort is the largest, it wants maximal social security benefits, which means taxing to the Laffer point $(\frac{1}{1+\varepsilon})$. They also allow the maximal number of skilled migrants in to the economy because of the tax contribution this generates to the welfare system.

It is interesting to note that, although the unskilled young are net beneficiaries in this welfare state, they are, nevertheless, still paying taxes. Hence the preferred tax policy of the unskilled voters is smaller than the Laffer point with a wedge $\frac{1}{J}$. (We will provide further discussions on this deviation factor below.) Clearly, the unskilled workers also prefer to let in more skilled immigrants due to their contribution to the welfare state. How many will they let in depends on the function Ψ, which weighs the future benefits against the cost at the present. Basically, if the unskilled workers are not forward-looking, it is in their best interest to let in as many skilled migrants as possible. However, this will lead to no redistribution in the next period because the skilled workers will be the largest. Hence, the function Ψ is the difference between the benefits they get by being, as they are, forward-looking and being myopic.

The skilled native-born prefer more skilled migrants for a different reason than the earlier two groups. They prefer to let in skilled migrants in this case because this will provide a higher number of skilled native workers in the *next* period. Thus, because the skilled are forward-looking, they too will prefer to have more skilled workers in their retirement period. However, they cannot let in too many of them because their high birth rate may render the skilled young in the next period as the largest group who will vote to abolish the welfare state altogether (similar to chapter 5).

A common feature among models with subgame-perfect Markov equilibrium is the idea that today's voters have the power to influence the identity of future policymakers. Such feature is also prominent in our analysis here (as well as in chapter 5). The migration policy of either young group reflects the fact that they may want to put themselves as the largest group in the next period. Thus, instead of letting in too many migrants, who will give birth to a large new skilled generation, they will want to let in as much as possible before the threshold is crossed. This threshold is $\frac{1-(1+n)s_t}{m}$. This strategic motive on migration quota is previously fleshed out in chapter 5. Letting $s_t = 1$ gets the result of the chapter. There are two differences between this threshold and the one in chapter 5. First, the equilibrium here has a bite even if the population growth rate is *positive*, which cannot be done when there are only young and old cohort, as in chapter 5, unless there is a negative population growth rate. Another fundamental is that, in order to have some transfer in the economy, the young decisive largest group has a choice of placing the next period's decisive power either in the hand of next period's unskilled or the old. So we need to verify an additional condition that it is better for this period's decisive young to choose the old generation next period, which is the case.

When $s_t \geq \frac{1}{1+n}$, we have a unique situation (which is only possible when $n > 0$). In this range of values, the number of skilled is growing too fast to be curbed by reducing migration volume alone. To ensure that the decisive power lands in the right hand (that is, the old), the skilled voters (who are the largest in this period) must make the unskilled cohort grow to weigh down the growth rate of the skilled workers. This is done by restricting both the skill composition as well as the size of total migration.[12]

The tax choice of the unskilled young deserves an independent discussion. In Razin, Sadka and Swagel (2002a, 2002b), it is maintained that the "fiscal leakage" to the native-born and to the migrants who are net beneficiaries may result in a lower tax rate chosen by the median voter. They assume that all migrants possess lower skill than the native-born. Because this increases the burden on the fiscal system, the median voter vote to reduce the size of the welfare state, instead of increasing it. To see such a resemblance the our result, we must first take the migration volume, μ_t, and the skill composition, σ_t, as given. Letting τ_t^u denote the tax rate preferred by the unskilled group, one can verify from equation (17) that $\frac{\partial \tau_t^u}{\partial \sigma_t} > 0$, and there exists $\overline{\sigma}$ such that, for any $\sigma_t < \overline{\sigma}$, we have $\frac{\partial \tau_t^u}{\partial \mu_t} < 0$. Conversely, for any $\sigma_t > \overline{\sigma}$, we would get an expansion of the welfare state, because $\frac{\partial \tau_t^u}{\partial \mu_t} > 0$.[13]

[12]Empirically, with the population growth rate of the major host countries for migration like the U.S. and Europe going below 1%, it is unlikely that this case should ever be of much concern. Barro and Lee (2000) provides an approximation of the size of the skilled. While Barro and Lee statistics capture those 25 years and above, they also cite OECD statistics which capture age group between 25 and 64. The percentage of this group who received tertiary education or higher in developed countries falls in the range of 15% to 47%.

[13]Recall that the tax rate preferred by the unskilled young workers is less than the level

The inequalities tell us that higher number of skilled migrants will prompt a higher demand for intra-generational redistribution. The fiscal leakage channel shows that unskilled migration creates more fiscal burden, such that the decisive "unskilled" voters would rather have the welfare state shrink. In addition, an increase in inequality in the economy, reflected in the skill premium ratio $\frac{w_t^s}{w_t^u}$, leads to a larger welfare state demanded by the unskilled.

5 Conclusion

In this paper, which is part of a forthcoming book, we built a dynamic politico-economic model featuring three groups of voters: skilled workers, unskilled workers, and retirees. The model features both *inter*- and *intra*-generational redistribution, resembling a welfare state. The skilled workers are net contributors to the welfare state whereas the unskilled workers and old retirees are net beneficiaries. When the skilled cohort grows rapidly, it may be necessary to bring in unskilled migrants to counter balance the expanding size of the skilled group.

As in chapter 5, the native-born young, whether skilled or unskilled, benefit from letting in migrants of all types, because their high birth rates can help increase the tax base in the next period. In this respect, skilled migrants help the welfare state more than unskilled migrants, to the extent that the offspring resemble their parents with respect to skill. On the other hand, more migrants in the present will strengthen the political power of the young in the next period who, relatively to the old, are less keen on the generosity of the welfare state. In this respect, unskilled migrants pose less of a threat

that is preferred by the old retirees. The tax rate preferred by the old retirees, $\tau_t^o = \frac{1}{1+\varepsilon}$ is the Laffer point that attains the maximum welfare size, given immigration policies. Therefore the size of the welfare state is monotonic in the tax rate when $\tau \in [0, \frac{1}{1+\varepsilon}]$. Thus, our use of "shrink" and "expand" is justified.

to the generosity of the welfare state then skilled migrants.

6 Appendix 7A: Strategic Voting Equilibrium

Recall that we have only three groups: the skilled native-born, the unskilled native-born, and the old. Let the set of three candidates be $\{s, u, o\}$, denoting their identity. Then, as in Chapter 6, the decision to vote of any individual must be optimal under the correctly anticipated probability of winning and policy stance of each candidate. Because identical voters vote identically, we can focus on the decision of a representative voter from each group. Let $e_t^i \in \{s, u, o\}$ be the vote of individual of type $i \in \{s, u, o\}$ cast for a candidate. In the same spirit as in Chapter 6, voting decisions $\mathbf{e}_t^* = (e_t^{s*}, e_t^{u*}, e_t^{o*})$ form a *voting equilibrium* at time t if

$$e_t^{i*} = \arg\max \left\{ \sum_{j \in \{s,u,o\}} \mathcal{P}^j(e_t^i, \mathbf{e}_{-it}^*) V^i \left(\Phi_t^j, \Phi_{t+1}, \mathbf{e}_{t+1} \right) \mid e_t^i \in \{s, u, o\} \right\} \quad (21)$$

for $i \in \{s, u, o\}$, where $\mathcal{P}^j(e_t^i, \mathbf{e}_{-it}^*)$ denotes the probability that candidate $j \in \{s, u, o\}$ will win given the voting decisions, and $\mathbf{e}_{-it}^*$ is the optimal voting decision of other groups that is not i, and $\Phi_t^j = \left(\tau_t^j, \sigma_t^j, \mu_t^j\right)$ is the policy vector if candidate j wins. Thus we require that each vote cast by each group is a best-response to the votes by the other groups. In addition, the representative voter of each group must take into the account the *pivotal* power of their vote, because the entire group will also vote accordingly. The voting decision of the old voters is simple, because they have no concern for the future,

$$e_t^{o*} = \arg\max \left\{ \sum_{j \in \{s,u,o\}} \mathcal{P}^j(e_t^o, \mathbf{e}_{-ot}^*) V^i \left(\tau_t^j, \sigma_t^j, \mu_t^j \right) \mid e_{ot} \in \{s, u, o\} \right\}.$$

After the election, the votes are tallied by adding up the size of each group that have chosen to vote for the candidate. The candidate with the most votes wins the election and gets to implement his ideal set of policies.

Clearly, each individual prefers the ideal policies of their representative candidate. Strategic voting opens up the possibility of voting for someone else that is not the most preferred candidate in order to avoid the least favorable candidate. For the skilled young, they prefer the least amount of taxes and some migration for the future. Thus, they will prefer the policy choice of the unskilled over the old candidate. As for the old retirees, the higher the transfer benefits, the better. Clearly, the unskilled candidate promises some benefits whereas the skilled promises none, so they would choose the policies of the unskilled over the skilled.

As for the unskilled workers, both rankings are possible: either they prefer the policy choice of the skilled over the old, or vice versa. The parameters of the model will dictate the direction of their votes. The cut-off tax policy, $\widetilde{\tau}$, is the break-even point for the unskilled between getting taxed but receiving transfer (policies of the old candidate) or pay no tax at all (policies of the skilled candidate).Formally, this tax level, $\widetilde{\tau}$, is defined implicitly by the equation

$$\frac{(w^u)^{1+\varepsilon}}{1+\varepsilon} = \frac{((1-\widetilde{\tau})w^u)^{1+\varepsilon}}{1+\varepsilon} + \frac{\widetilde{\tau}(1-\widetilde{\tau})^{\varepsilon}\left((s_t+\sigma_t\mu_t)(w^s)^{1+\varepsilon} + (1-s_t+(1-\sigma_t)\mu_t)(w^u)^{1+\varepsilon}\right)}{1+\mu_t+\frac{1+\mu_{t-1}}{1+n+\mu_{t-1}(1+m)}}. \tag{22}$$

We know that such a tax policy exists, because, take next period's policy as given, the payoff in this period to the unskilled is maximized at its preferred policy and zero at $\tau = 1$. Therefore, at some $\widetilde{\tau}$, the equality will hold. This cut-off tax rate will play an important role for the unskilled young' voting decision.

The main problem with ranking the utility streams of the voters is due to the multiplicity of *future* equilibria once we extend our work to strategic voting. This makes it impossible for the voters to get a precise prediction of what will happen as a result of their action today. Even if we could pin down all the relative sizes of all possible payoffs in the next period, multiple voting equilibria do not allow a prediction of which equilibrium will be selected in the future. To deal with the problem, we restrict the voting equilibrium to satisfy the stationary Markov-perfect property, similarly to the policy choices in previous subsection. Now, we are ready to define the subgame-perfect Markov political equilibrium under strategic voting. We are looking for the a triplet policy function $(\tau_t, \sigma_t, \mu_t) = \Phi(s_t, \tau_{t-1}, \sigma_{t-1}, \mu_{t-1}, \mathbf{e}_t^*)$ with the voting vector $\mathbf{e}_t^*$ that solve the following two problems:

$$\Phi(s_t, \tau_{t-1}, \sigma_{t-1}, \mu_{t-1}, \mathbf{e}_t^*) = \underset{\tau_t, \sigma_t, \mu_t}{\arg\max} \quad V^d\left(s_t, , \tau_t, \sigma_t, \mu_t, \Phi(s_{t+1}, \tau_t, \sigma_t, \mu_t, \mathbf{e}_t^*)\right) \tag{23}$$

$$\text{s.t. } s_{t+1} = \frac{(1+n)s_t + (1+m)\sigma_t\mu_t}{1+n+\mu_t(1+m)},$$

where $d \in \{s, u, o\}$ is the identity of the the winning candidate, decided by the voting equilibrium $\mathbf{e}_t^*$ that satisfies the subgame-perfect Markov property and solves

$$\begin{aligned} e_t^{i*} &= \mathbf{e}^*\left(s_t, \tau_{t-1}, \sigma_{t-1}, \mu_{t-1}, \mathbf{e}_{t-1}^*\right) \qquad (24) \\ &= \underset{e_t^i \in \{s,u,o\}}{\arg\max} \sum_{j \in \{s,u,o\}} \mathcal{P}^j(e_t^i, \mathbf{e}_{-it}^*) V^i\left(\Phi_t^j, \Phi(s_{t+1}, \tau_t, \sigma_t, \mu_t, \mathbf{e}_t^*), \mathbf{e}^*\left(s_{t+1}, \tau_t, \sigma_t, \mu_t, \mathbf{e}_t^*\right)\right) \end{aligned}$$

where $\mathcal{P}^j(e_t^i, \mathbf{e}_{-it}^*)$ denotes the winning probability of the representative candidate $j \in \{s, u, o\}$ given the voting decisions, and $\mathbf{e}_{-it}^*$ is the optimal voting decision of other groups that is not i, and $\Phi_t^j = \left\langle \tau_t^j, \sigma_t^j, \mu_t^j \right\rangle$ is the vector of preferred policy of candidate from group j.

The stationary Markov-perfect equilibrium defined above introduces another functional equation exercise. The first exercise is to find a policy profile that satisfies the usual Markov-perfect definition, as discussed in the case of sincere voting in the text. The second exercise restricts the voting decision to be cast on the belief that individuals in the same situation next period will vote in exactly the same way. With this property, the voters in this period know exactly how future generations will vote and can evaluate the stream of payoffs accordingly.

Lastly, the keep the analysis simple, we focus on voting equilibria that are consistent with policies derived in the text for the case of sincerely voting. This will be the case if the policies are always coupled with a voting equilibrium featuring the largest group always voting for its representative candidate. In particular, if the group forms the absolute majority, all votes cast from this group will go to its representative candidate. The economy can go through different equilibrium paths depending on n, m, and s_0, as follows:

1. If $n+m \leq 0$, the old group is always the absolute majority. Tax rate is at the Laffer point and the economy is fully open to skilled migration.

2. If $n+m > 0$, then the dynamics depend on the initial state of the economy, s_0. If $s_0 \geq \frac{1+\frac{n}{2}}{1+n}$, then the skilled workers are the majority (controlling 50% of the population), and zero tax rate with limited skilled migration will be observed. If $\frac{n}{2(1+n)} \geq s_0$, the unskilled workers are the majority, then there will be a positive tax rate (less than at the Laffer point) and some skilled migration. If $n < 0$, then *initially* the old cohort is the majority; the tax rate will be at the Laffer point and the skilled migration will be maximal. Otherwise, the policies implemented are given in the equilibrium below.

The first equilibrium we look at is dubbed "Intermediate" because it captures the essence that the preferred policies of the unskilled workers are a compromise from the extremity of the other two groups. We can show that the following strategy profile forms a subgame-perfect Markov Equilibrium with strategic voting

$$
\begin{aligned}
e_t^{s*} &= \begin{cases} s & \text{, if } s_t \geq \frac{1+\mu_{t-1}}{1+n+\mu_{t-1}(1+m)} \\ u & \text{, otherwise} \end{cases} \\
e_t^{u*} &= u \\
e_t^{o*} &= \begin{cases} o & \text{, if } \frac{1+\mu_{t-1}}{1+n+\mu_{t-1}(1+m)} \geq \max\{s_t, 1-s_t\} \\ u & \text{, otherwise} \end{cases}
\end{aligned}
\tag{25}
$$

and the policies implemented when no group is the absolute majority are

$$
\Phi_t = \left(\tau_t = \frac{1-\frac{1}{J}}{1+\varepsilon-\frac{1}{J}}, \sigma_t = 1, \mu_t = \frac{2+n-2(1+n)s_t}{m}\right) \tag{26}
$$

where $J = J(\mu_t, \sigma_t, s_t, \mu_{t-1})$ is as in equation (17).

The equilibrium features the unskilled voters always voting for their representative, whereas the other two groups vote for their respective candidate only if they are the largest group, or for the unskilled candidate otherwise. With these voting strategy, if no group captures 50% of the voting populations, the policy choice preferred by the unskilled candidate will prevail. One notable difference is the policy related to the immigration volume. In period $t+1$, as long as the skilled workers do not form 50% of the voting population, the policies preferred by the unskilled workers will be implemented. To make sure that this is the case, skilled migration is restricted to just the threshold that would have put the skilled voters as the absolute majority in period $t+1$. The volume of migration, $\mu_t^* = \frac{2+n-2(1+n)s_t}{m}$, reflects the fact that the threshold value for this variable has been pushed slightly farther. This level can be shown to be higher than the restricted volume in sincerely voting equilibrium.

In the preceding equilibrium, we let the preference of the skilled workers and the old retirees decide the fate of the the policies. In the following analysis, the unskilled workers consider who they want to vote for. This will depend on how extractive the tax policy preferred by old is. We call the next equilibrium "Left-wing", because it features a welfare state of the size greater-than-or-equal to that of the intermediate policy equilibrium. This may arise when the tax rate preferred by the old voters is not excessively to redistributive. When $\frac{1}{1+\varepsilon} \leq \widetilde{\tau}$, we can show that we have an equilibrium of the following form

$$
\begin{aligned}
e_t^{s*} &= \begin{cases} s & \text{, otherwise} \\ u & \text{, if } \frac{1+\mu_{t-1}}{1+n+\mu_{t-1}(1+m)} \geq s_t \geq \frac{1+\frac{n-m}{2}}{1+n} \end{cases} \\
e_t^{u*} &= \begin{cases} u & \begin{cases} \text{, if } 1-s_t \geq \frac{1+\mu_{t-1}}{1+n+\mu_{t-1}(1+m)}, \text{ or} \\ \frac{1+\mu_{t-1}}{1+n+\mu_{t-1}(1+m)} \geq s_t \geq \frac{1+\frac{n-m}{2}}{1+n} \end{cases} \\ o & \text{, otherwise} \end{cases} \\
e_t^{o*} &= o
\end{aligned} \tag{27}
$$

and the policies implemented when no group is the absolute majority are

$$
\Phi_t = \begin{cases} \left(\tau_t = \frac{1-\frac{1}{J}}{1+\varepsilon-\frac{1}{J}}, \sigma_t = 1, \mu_t = \frac{2+n-2(1+n)s_t}{m}\right) & \text{, if } \frac{1+\mu_{t-1}}{1+n+\mu_{t-1}(1+m)} \geq s_t \geq \frac{1+\frac{n-m}{2}}{1+n} \\ \left(\tau_t^* = \frac{1}{1+\varepsilon}, \sigma_t = 1, \mu_t = 1\right) & \text{, otherwise} \end{cases} \tag{28}
$$

where $J = J(\mu_t, \sigma_t, s_t, \mu_{t-1})$ is as in equation (17) and $\widetilde{\tau}$ is given implicitly in equation (22).

When the tax rate preferred by the old voters is not excessively redistributive in the eyes of the unskilled, we could have an equilibrium where the unskilled voters strategically vote for the old candidate to avoid the policies preferred by the skilled voters. This will be an equilibrium when the size of the skilled is not "too large." Recall that, voting to implement the policies selected by the old candidate leads to opening the economy fully to the skilled immigrants. If the size of the skilled group is currently too large, there is a risk of making the skilled voters the absolute majority in the next period and will result in no welfare state in the retirement of this period's workers.

The cutoff level before this happens is given by $\frac{1+\frac{n-m}{2}}{1+n}$. Therefore, voting for the old will only be compatible with the interest of the unskilled voters when the tax rate is not excessively high and when the size of the skilled is not too large.

We turn our attention to the next equilibrium. When $\frac{1}{1+\varepsilon} > \widetilde{\tau}$, we can show that there is an equilibrium with the following functions:

$$
\begin{aligned}
e_t^{s*} &= \begin{cases} s & \text{, otherwise} \\ u & \text{, if } 1 - s_t \geq \frac{1+\mu_{t-1}}{1+n+\mu_{t-1}(1+m)} \end{cases} \\
e_t^{u*} &= \begin{cases} u & \text{, otherwise} \\ s & \text{, if } \frac{1+\mu_{t-1}}{1+n+\mu_{t-1}(1+m)} \geq \max\{s_t, 1-s_t\}. \end{cases} \\
e_t^{o*} &= \begin{cases} o & \text{, otherwise} \\ u & \text{, if } s_t \geq \frac{1+\mu_{t-1}}{1+n+\mu_{t-1}(1+m)} \end{cases}
\end{aligned}
\tag{29}
$$

and the policies implemented when no group is the absolute majority are

$$
\Phi_t = \begin{cases} \left(\tau_t = 0, \sigma_t = 1, \mu_t = \frac{2+n-2(1+n)s_t}{m}\right) & \text{, if } \frac{1+\mu_{t-1}}{1+n+\mu_{t-1}(1+m)} \geq \max\{s_t, 1-s_t\} \\ \left(\tau_t = \frac{1-\frac{1}{J}}{1+\varepsilon-\frac{1}{J}}, \sigma_t = 1, \mu_t = \frac{2+n-2(1+n)s_t}{m}\right) & \text{, otherwise} \end{cases}
\tag{30}
$$

where $J = J(\mu_t, \sigma_t, s_t, \mu_{t-1})$ is as in equation (17) and $\widetilde{\tau}$ is given in equation (22).

When the Laffer point is higher than $\widetilde{\tau}$, the tax rate is read as excessive. In this case, the unskilled voters will instead choose to vote for the skilled over the old candidate. The resulting equilibrium as the size of the welfare state less-than-or-equal to that in the intermediate policy equilibrium, hence we refer to it as "Right-wing." When the tax preferred by the old is excessive from the perspective of the unskilled, the political process could implement the policies preferred by the skilled in order to avoid the worst possible outcome. This happens when the old voters constitute the largest group, and the unskilled voters vote strategically for the skilled candidate. In other cases, however, the policies preferred by the unskilled will be implemented, irrespective of the identity of the largest group in the economy.

For our results with multidimensional policies, it is important to note here that the ranking of candidates by individual voters allows us to escape the well-known agenda-setting cycle (the "Condorcet paradox"). Such a cycle, which arises when any candidate could be defeated in a pairwise majority voting competition, leads to massive indeterminacy and non-existence of a political equilibrium. The agenda-setting cycle will have a bite if the rankings of the candidates for all groups are unique: no group occupies the same ranked position more than once. However, this does not arise here, because, in all equilibria, some political groups have a *common* enemy. That is, because they will never vote for the least-preferred candidate (the "common" enemy), the voting cycle breaks down to determinate policies above, albeit their multiplicity. This occurs when voters agree on who is the least-preferred candidate and act together to block her from winning the election. The literature typically avoids the Condorcet paradox by restricting political preferences with some ad hoc assumptions. For our case, the preferences induced from economic assumption lead to the escape of the Condorcet paradox. For discussions on agenda-setting cycle, see Drazen (2000, page 71-72), and Persson and Tabellini (2000, page 29-31).

References

[1] Auerbach, A. and P. Oreopoulos (1999), "Analyzing the Economic Impact of U.S. Immigration," *American Economic Review Papers and Proceedings*, 89(2), 176-180.

[2] Benhabib, J. (1996), "On the Political Economy of Immigration," *European Economic Review*, 40, 1737-43.

[3] Besley, Timothy and Stephen Coate (1997), "An Economic Model of Representative Democracy," *Quarterly Journal of Economics*, 112(1), February, 85-114.

[4] Besley, Timothy and Stephen Coate (1998), "Sources of Inefficiency in a Representative Democracy: A Dynamic Approach," *American Economic Review*, 88(1), 139-156.

[5] Blank, Rebecca M. (1988), "The Effect of Welfare and Wage Levels on the Location Decisions of Female-Headed Households." *Journal of Urban Economics*, 24, 186.

[6] Boeri, Tito, Gordon Howard Hanson and Barry McCormick (2002), *Immigration Policy and the Welfare System: A Report for the Fondazione Rodolfo Debenedetti*, Oxford University Press.

[7] Bohn, Henning (2005), " Will Social Security and Medicare Remain Viable as the U.S. Population Is Aging: An Update," In Robin Brooks and Assaf Razin (eds.), *The Politics and Finance of Social Security Reform*, Cambridge University Press, 44-72.

[8] Boldrin, M. and A. Rustichini (2000), "Political Equilibria with Social Security," *Review of Economic Dynamics*, 3, 41-78.

[9] Borjas, George J. (1999), *Heaven's Door: Immigration Policy and the American Economy*, Princeton University Press, Princeton, New Jersey.

[10] Brucker, Herbert, Gil Epstein, Barry McCormick, Gilles Saint-Paul, Alessandra Venturini, and Klaus Zimmerman (2001), "Managing Migration in the European Welfare State," mimeo, IZA Bonn, Germany.

[11] Casarico, Alessandra and Carlo Devillanova (2003), "Social Security and Migration with Endogenous Skill Upgrading," *Journal of Public Economics*, 87 (3-4), 773-797.

[12] Cohen, Alon, and Assaf Razin, "The Skill Composition of Immigrants and the Generosity of the Welfare State: Free versus Policy-controlled Migration," *NBER Working Paper No. 144459*, October.

[13] Cooley, T. F. and J. Soares (1999), "A Positive Theory of Social Security Based on Reputation," *Journal of Political Economy*, 107, 135-160.

[14] De Giorgi, Giacomo and Michele Pellizzari (2006), "Welfare Migration in Europe and the Cost of a Harmonized Social Assistance," *IZA Discussion Paper No. 2094.*

[15] Docquier, Frederic and Abdeslam Marfouk (2006), "International Migration by Educational Attainment 1990-2000," in Caglar Ozden and Maurice Schiff (eds.), *International Migration, Remittances ad the Brain Drain*, McMillan and Palgrave: New York.

[16] Docquier, Frederic, Oliver Lohest and Abdeslam Marfouk (2006), "What Determines Migrants' Destination Choice?," working paper.

[17] Dolmas, J. and G.W. Huffman (2004), "On the Political Economy of Immigration and Income Redistribution," *International Economic Review*, 45, 1129-68.

[18] Drazen, Alan (2000), *Political Economy in Macroeconomics*, Princeton University Press: New Jersey.

[19] Downs, Anthony (1957), *An Economic Thoery of Democracy*, Harper and Row.

[20] Enchautegui, Maria E. (1997), "Welfare Payments and Other Determinants of Female Migration," *Journal of Labor Economics*, 15, 529.

[21] Forni, L. (2005), "Social Security as Markov Equilibrium in OLG Models," *Review of Economic Dynamics*, 8, 178-194.

[22] Frankel, Jeffrey A. and David Romer (1999), "Does Trade Cause Growth?," *American Economic Review*, 89(2), 379-399.

[23] Galasso, V. and P. Profeta (2002), "The Political Economy of Social Security: A Survey," *European Journal of Political Economy*, 18, 1-29.

[24] Gelbach, Jonah B. (2000), "The Life-cycle Welfare Migration Hypothesis: Evidence from the 1980 and 1990 Censuses," working paper.

[25] Gramlich, Edward M. and Deborah S. Laren (1984), "Migration and Income Redistribution Resposibilities," *Journal of Human Resources*, 19(4), 489.

[26] Hanson, Gordon H. (2008), "The Economic Consequence of the International Migration of Labor," *NBER Working Paper No. 14490*, November.

[27] Krusell, P. and J.V. Rios-Rull (1996), "Vested Interests in a Positive Theory of Stagnation and Growth," *Review of Economic Studies*, 63, 301-29.

[28] Lee, R. and T. Miller (2000), "Immigration, Social Security, and Broader Fical Impacts," *American Economic Review Papers and Proceedings*, 90(2), 350-354.

[29] Levine, Phillip B. and David J. Zimmerman (1999), "An Empirical Analysis of the Welfare Magnet Debate Using the NLSY," *Journal of Population Economics*, 12(3), 391.

[30] McKinnish, Terra (2005), "Importing the Poor: Welfare Magnetism and Cross-Border Welfare Migration," *Journal of Human Resources,* 40(1), 57.

[31] Meyer, Bruce D. (2000), "Do the Poor Move to Receive Higher Welfare Benefits?," unpublished paper.

[32] Ortega, F. (2005), Immigration Quotas and Skill Upgrading," *Journal of Public Economics*, 89(9-10), 1841-1863.

[33] Persson, Torsten and Guido Tabellini (2000), " *Political Economics: Explaining Economic Policy*, MIT Press: Cambridge, MA.

[34] Peridy, Nicolas (2006), "The European Union and Its New Neighbors: An Estimation of Migration Potentials," *Economic Bulletin*, 6(2), 1.

[35] Razin, Assaf and Efraim Sadka (1999), "Migration and Pension with International Capital Mobility," *Journal of Public Economics*, 74, 141-150.

[36] Razin, Assaf and Efraim Sadka (2004), "Welfare Migration: Is the Net Fiscal Burden a Good Measure of Its Economic Impact on the Welfare of the Native-born Population?," *CESifo Economic Studies*, 50(4), 709-716.

[37] Razin, Assaf, Efraim Sadka and Phillips Swagel (2002a), "The Aging Population and the Size of the Welfare State," *Journal of Political Economy*, 110 900-918.

[38] Razin, Assaf, Efraim Sadka and Phillips Swagel (2002b), "Tax Burden and Migration: A Political Theory and Evidence," *Journal of Public Economic*, 85, 167-190.

[39] Sand, Edith, and Assaf Razin (2007), "The Political-Economy Positive Role of Social Security in Sustaining Migration (But Not Vice Versa)," *NBER Working Paper 13598.*

[40] Scholten, U., and M.P. Thum (1996), "Public Pensions and Immigration Policy in a Democracy," *Public Choice*, 87, 347-361.

[41] Sinn, Hans-Werner (2003), "EU Enlargement, Migration, and Lessons from German Unification," *German Economic Review*, 1(3), 299 - 314.

[42] Southwick, Lawrence Jr. (1981), "Public Welfare Programs and Recipient Migration," *Growth and Change*, 12(4), 22.

[43] Storesletten, K. (2000), "Sustaining Fiscal Policy Through Immigration," *Journal of Political Economy*, 108(2), 300-323.

[44] Suwankiri, Benjarong (2009), "Three Essays in Dynamic Political Economy: Migration, Welfare State, and Poverty," Ph.D. Dissertation, unpublished: Cornell University.

[45] Walker, James (1994), "Migration Among Low-income Households: Helping he Witch Doctors Reach Consensus," unpublised paper.