ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Hur, Jinwook

Article

Political Economy of Immigration and Fiscal Sustainability

KDI Journal of Economic Policy

Provided in Cooperation with:
Korea Development Institute (KDI), Sejong

Suggested Citation: Hur, Jinwook (2022) : Political Economy of Immigration and Fiscal Sustainability, KDI Journal of Economic Policy, ISSN 2586-4130, Korea Development Institute (KDI), Sejong, Vol. 44, Iss. 1, pp. 1-47,
https://doi.org/10.23895/kdijep.2022.44.1.1

This Version is available at:
https://hdl.handle.net/10419/251901

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

https://creativecommons.org/licenses/by-sa/4.0/

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

KDI Journal of Economic Policy 2022, 44(1):1–47
http://dx.doi.org/10.23895/kdijep.2022.44.1.1

Political Economy of Immigration and Fiscal Sustainability†

By JINWOOK HUR*

This paper introduces a politico-economic model with a welfare state and immigration. In this model, policies on taxes and immigration are determined through a plurality voting system. While many studies of fiscal implications of immigration argue that relaxing immigration policies can substitute for tax reforms in an aging economy, I show that the democratic voting procedure can dampen the effect of relaxing immigration policies as desired policy reforms are not always implemented by the winner of an election. This political economy results in three types of social welfare losses. First, the skill composition is not balanced at a socially efficient level because workers are motivated to maximize their wages. Second, older retirees implement excessive taxes to maximize the size of the welfare state. Third, the volume of immigration is lower than the optimal level given the incentive by young workers to regain political power in the future.

Key Word: Immigration, Political Economy, Fiscal Sustainability
JEL Code: E60, F22, H20, J61, D72

I. Introduction

In many developed economies, sustainability of the welfare state is an important issue. The aging of the baby-boomer generation and falling fertility rates make it more difficult to maintain welfare programs without reforms of fiscal policies. Tax reform is one option, but increasing the tax burden causes other controversies given the potential negative effects of this strategy on productivity and social welfare.

In this sense, some economists argue that relaxing immigration policies can be an effective alternative to tax reforms. According to these studies, a more liberal immigration policy for young and skilled workers can increase tax revenues and solve fiscal problems in many economies with a pay-as-you-go (PAYG) system of welfare. For example, Storesletten (2000) argues that admitting 1.6 million high-

* Associate Fellow, Korea Development Institute (E-mail: jhur@kdi.re.kr)
* Received: 2021. 11. 28
* Referee Process Started: 2021. 12. 1
* Referee Reports Completed: 2021. 12. 29
† This paper is written based on Chapter 1 of Hur, 2016, *Essays on Fiscal Policy*, University of California, Los Angeles.

skilled immigrants aged 40-44 years old annually can replace tax reforms and resolve the fiscal problems arising in the U.S. economy as the U.S. population ages.

In most studies concentrating on the fiscal implications of immigration policies, the policy variables are assumed to be exogenous. However, during the actual immigration policy decision-making process, there are many factors, especially those that are political, that can keep desired outcomes from being implemented. For instance, young citizens of working age may worry that their wages would decrease due to competition from immigrants, causing them to vote for parties against generous immigration policies. Therefore, the extent to which the desired reforms of immigration policies can be implemented through a democratic decision-making process is important.

In this paper, I analyze political equilibria in a simple overlapping-generation model with two skill levels in labor, proportional taxes on labor income, and a welfare state in the form of simple lump-sum transfers. The native population, or the electorate, is divided into three different political parties depending on their age and skill level. The labor income tax rate, volume of immigration, and skill composition of the immigrants are then determined by the winning candidate after plurality voting. As there are three different political parties, the possibilities of strategic voting and political coalitions are also considered.

Using this model economy, I demonstrate the pattern of different preferences on fiscal and immigration policies by different cohorts. Older retirees want to maximize tax revenue per capita because they do not have labor income; hence, they prefer the Laffer rate of labor income tax and a balanced skill composition between skilled and unskilled workers to maximize production efficiency. On the other hand, the preference of younger workers depends on their skill level. Specifically, skilled workers want immigrants to be unskilled, and unskilled workers want them to be skilled. Moreover, depending on the state of the economy, young workers may not want to allow the maximum number of immigrants in, although their instantaneous utility is maximized by allowing in as many as possible. If young workers allow too many immigrants in, the probability that they lose the election in the next period increases given that the demographic composition of the electorate will be biased toward the younger generation in the next period due to the high fertility of the immigrants, which may remove the current young cohort's opportunity to enjoy the benefit of the welfare state after they retire. Thus, older retirees want the maximum amount of immigration and the strongest welfare state as pure beneficiaries of the welfare program, while skilled young workers want zero taxes with no welfare program in most of the parameter regions. Unskilled young workers are mostly in the middle of the two groups in the sense that they do not mind taxes as much as skilled workers do. This emphasizes the role of unskilled workers as a "swing" cohort in elections.

It is also shown that the skill composition of the economy exhibits "cycling" dynamics. That is, when the skill composition is biased toward unskilled labor at the beginning of a period, unskilled workers are highly likely to sway the election, thereby approving of a high volume of skilled immigrants and balancing the skill composition. Similarly, when the skill composition leans too much toward skilled labor, skilled workers are more likely to sway the election, and the winning candidate will allow a high volume of unskilled immigrants. Since the optimal skill composition of workers is to have more skilled workers than unskilled workers in an

economy with the simple Cobb-Douglas technology, the democratic procedure of plurality voting causes welfare losses due to the gap between the optimal skill composition of immigrants and the realized skill composition of immigrants, which tends to be lower than the optimal level.

I also show other factors causing welfare losses. First, the distortionary labor income tax in the model is a source of inefficiency. Because older retirees implement the Laffer rate whenever they win an election, this distortion is most severe when these retirees become the decisive political party. Second, the volume of immigration may be lower than the optimal level. Given that younger workers worry that they may not regain political power when they are older and lose their chance to obtain future social security benefits, young workers have an incentive to adjust the immigration volume to guarantee that they can win not only the election today but also those in the future.

This paper is in line with the literature on the macroeconomic implications of the political economy. Specifically, it builds on several studies on the political economy of immigration, such as Benhabib (1996), Dolmas and Huffman (2004) and Ortega (2005), and on studies of the political economy of social security systems, such as Boldrin and Rustichini (2000) and Galasso and Profeta (2002). Specifically, the main goal of this paper is to examine whether policy implications in the immigration literature[1] are feasible in a positive theory setup.

This study technically follows the analyses of Ortega (2005) and Suwankiri *et al.* (2016). Ortega (2005) introduces a dynamic politico-economic model with immigration and skill upgrading. In his model, the voting decision of the electorate affects the skill composition of the economy, thereby influencing not only the skill complementarity but also the skill composition of the future electorate. Therefore, voters confront the tradeoff between their skill premium and their future political power. On the other hand, the work by Suwankiri *et al.* (2016) abstracts from skill complementarity by assuming that the production function is linear in skilled and unskilled labor. Instead, their work concentrates on a more complicated political situation in which there exists a redistributive welfare state and three political parties. The three parties, formed according to skill levels and ages, make voting decisions in a plurality voting system and, therefore, potentially vote strategically. This paper builds on the techniques introduced in both of these papers. In a state of political equilibrium, each young worker considers not only the effect of the election result on his labor income but also how it will affect his political power when he becomes older and retires.

This paper proceeds as follows. The model economy is described in section II. The political procedure of policy implementation is explained in section III. Section IV describes the political equilibria of the model. In section V, the dynamics of the model and the welfare analysis are discussed. Section VI concludes the paper.

II. Model Economy

The model builds on the dynamic overlapping generation model as in Suwankiri

[1]As explained above, many studies report that relaxing immigration policies can reduce fiscal problems in developed economies. Among others, see Auerbach and Oreopoulos (1999) and Storesletten (2000).

et al. (2016) with the application of the democracy model with strategic voting behaviors in Besley and Coate (1997; 1998). In this model, there are two generations (the young and the old) and two skill levels (skilled and unskilled). Each individual works only in the first period of his life and retires in the second and last period. The government levies proportional labor income taxes from young workers and spends lump-sum transfers via a pay-as-you-go (PAYG) system. Furthermore, the government sets the tax rate, the volume of immigration, and the skill composition of immigrants by democratic voting for each period.

In this model, each period is divided into three sub-periods. In the first sub-period, native individuals - both young and old - hold a plurality election, and the winner of the election decides on the labor income tax rate (τ_t), the volume of immigration (μ_t), and the skill composition of immigrants (σ_t).[2] In the second sub-period, immigrants enter the economy as determined by the winning candidate. In the third and final sub-period, production, taxation, transfers, and consumption take place.

A. *Preference*

The lifetime utility function of a young worker with skill level $i \in \{s,u\}$ is assumed to be quasi-linear in terms of consumption, as in

$$U^y\left(c_t^{y,i}, l_t^i, c_{t+1}^{o,i}\right) = c_t^{y,i} - \frac{(l_t^i)^{1+1/\nu}}{1+1/\nu} + \beta c_{t+1}^{o,i},$$

where l_t^i is the labor supply of skill level $i \in \{s,u\}$, β is the discount factor for the next period, and ν is the Frisch elasticity of the labor supply.

The preference of an old retiree simply consists of the utility from today's consumption because he does not have a future continuation value. Thus, the utility function is given by

$$U^o\left(c_t^{o,i}\right) = c_t^{o,i}.$$

The budget constraints of the young and the old are given by

$$c_t^{y,i} + d_t^i \le (1-\tau_t) w_t^i l_t^i + b_t,$$

$$c_t^{o,i} \le (1+r_t) d_{t-1}^i + b_t,$$

where τ_t is the proportional labor income tax rate at period t, and d_t and b_t are savings and lump-sum transfers from the government, respectively. w_t represents the wage rate.

Given that the utility function is quasi-linear in terms of consumption, the Euler

[2]The process of voting is described in detail in section III.

equation pins down the equilibrium interest rate at a constant, i.e.,

$$r_t = \frac{1}{\beta} - 1 \text{ for all } t.$$

At this r_t, young workers are indifferent with respect to the amount of savings, as the marginal benefit of savings offsets exactly the marginal cost of savings such that any amount of savings can be an equilibrium solution. Throughout this paper, we will only consider the case of $d_t^i = 0$ for all $i \in \{s,u\}$ for simplicity. This makes old retirees simply consume the lump-sum transfer regardless of their skill level, allowing this setup to reduce the two groups of old retirees (skilled & unskilled) into a single group.

The labor supply function of a young worker with skill level $i \in \{s,u\}$ is derived as follows:

$$l_t^i = [(1-\tau_t)w_t^i]^\nu$$

Therefore, the indirect lifetime utility of each cohort is expressed as shown below.

$$\begin{aligned} v_t^i &= c_t^{y,i} - \frac{(l_t^i)^{1+\frac{1}{\nu}}}{1+\frac{1}{\nu}} + \beta c_{t+1}^{o,i} \\ &= \frac{[(1-\tau_t)w_t^i]^{1+\nu}}{1+\nu} + b_t + \beta b_{t+1} \end{aligned} \tag{1}$$

$$V_t^o = b_t \tag{2}$$

B. *Production Technology*

This model is abstracted from capital accumulation, meaning that skilled and unskilled labor will become the only factors of production.[3] The output is produced according to the Cobb-Douglas technology,

$$Y_t = (L_t^s)^\alpha (L_t^u)^{1-\alpha},$$

where L_t^i is the aggregate labor input of skill level $i \in \{s,u\}$ and α is assumed to be greater than 0.5. The labor demand equations for the two skill groups

[3]An alternative setup for this is to assume that capital stock per capita is fixed and that immigrants bring the same amount of capital stock when they are approved to immigrate.

corresponding to this production function are thus expressed as shown below.

$$w_t^s = \alpha Y_t (L_t^s)^{-1}$$

$$w_t^u = (1-\alpha) Y_t (L_t^u)^{-1}$$

C. *Immigration and Demographics*

Before production takes place in a period, the government sets the volume and the skill composition of immigrants. For simplicity, it is assumed that only young people are allowed to immigrate; thus, all immigrants supply labor immediately upon their entrance into the economy. Then, the aggregate labor supply can be characterized by the following formulae:

$$L_t^s = [s_t + \sigma_t \mu_t] N_t l_t^s,$$

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

in which N_t is the total size of the native young group of individuals and $s_t \in [0,1]$ is the fraction of native-born skilled workers in the labor force in the beginning of period t. $\mu_t \in [0, \bar{\mu}]$ the ratio of approved immigrants to the native-born young population, thereby governing the volume of immigration.[4] Finally, $\sigma_t \in [0,1]$ stands for the fraction of skilled immigrants in the group of new immigrants.

Also, there are certain additional assumptions for convenience. First, the skill levels of offspring follow the parent. Second, the offspring of immigrants are considered native, while the immigrants themselves are not.[5] Then, the demographic dynamics can be described by the following equations:

$$N_{t+1} = [1 + n + (1+m)\mu_t] N_t \quad \text{(level of population)}$$

$$s_{t+1} N_{t+1} = [(1+n)s_t + (1+m)\sigma_t \mu_t] N_t \quad \text{(skill composition)}$$

where n and m represent the population growth rates for natives and migrants, respectively.[6] Combining these two equations, we obtain the following law of motion for the skill composition of the natives:

[4]The volume of immigration has a finite upper bound because otherwise the choice of μt can be infinity in a number of political equilibria, as the production function is a constant-return-to-scale function. In an economy with DRS technology, μt has a finite value even without an exogenous upper bound. In this DRS case, however, the state of the economy depends not only on the relative population size of the voters but also the absolute number of young workers, thereby increasing the dimension of the state vector by one.

[5]In other words, immigrants cannot vote for their lifetimes, while their descendants are legal voters.

[6]In many developed countries, it is consistently observed in census data that the fertility rate is higher for immigrants than for natives, implying $m > n$. One recent example is the report of U.S. Census Bureau by Monte and Ellis (2014).

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

Equation (3) demonstrates the role of immigration policies. For instance, $\sigma_t < s_t$ implies that the fraction of skilled immigrants in the entire group of immigrants exceeds the fraction of skilled workers in the group of native young workers. Then, the fraction of skilled workers in the native young workers will increase in the next period given the assumption that the descendants succeed their parents' skill levels. Also, if $\mu_t = 0$ (no immigrants accepted), then s_t obviously remains the same in the next period. [7] As $\mu_t \to \infty$ (infinitely many immigrants are approved), $s_{t+1} \to \sigma_t$, implying that the skill composition of the natives will converge to the skill composition of accepted immigrants as more immigration is accepted.

D. *Government and the Welfare State*

The government levies proportional labor income taxes at the rate determined by the winner of the vote. The tax revenue is transferred equally regardless of age (young and old), skill level (skilled and unskilled), and/or residence status (natives and immigrants). The government is assumed to balance the budget every period; thus, there is no government debt. The lump-sum transfer is then expressed as

$$b_t = \frac{\tau_t \left[(s_t + \sigma_t \mu_t) w_t^s N_t l_t^s + (1 - s_t + (1-\sigma_t)\mu_t) w_t^u N_t l_t^u \right]}{(1+\mu_t)N_t + (1+\mu_{t-1})N_{t-1}}$$

in which the numerator is the total tax revenue and the denominator stands for the total population, including both natives and immigrants.

E. *Equilibrium-Characterizing Equations*

The general equilibrium of this economy is characterized by the following nine equations.

(4) $$l_t^s = [(1-\tau_t) w_t^s]^{\nu}$$

(5) $$l_t^u = [(1-\tau_t) w_t^u]^{\nu}$$

(6) $$Y_t = (L_t^s)^{\alpha} (L_t^u)^{1-\alpha}$$

[7] In other words, immigration is the only method of changing the aggregate skill composition in this model, and there is no skill acquisition or loss. For studies focusing on the transition of the aggregate skill level or task composition through immigration, readers can refer to Hunt and Gauthier-Loiselle (2010) and Peri (2012), among others.

(7) $$w_t^s = \alpha Y_t (L_t^s)^{-1}$$

(8) $$w_t^u = (1-\alpha) Y_t (L_t^u)^{-1}$$

(9) $$L_t^s = [s_t + \sigma_t \mu_t] N_t l_t^s$$

(10) $$L_t^u = [1 - s_t + (1-\sigma_t)\mu_t] N_t l_t^u$$

(11) $$b_t = \frac{\tau_t \left[(s_t + \sigma_t \mu_t) w_t^s N_t l_t^s + (1 - s_t + (1-\sigma_t)\mu_t) w_t^u N_t l_t^u \right]}{(1+\mu_t) N_t + (1+\mu_{t-1}) N_{t-1}}$$

(12) $$N_t = [1 + n + (1+m)\mu_{t-1}] N_{t-1}$$

In this system of equations, the following variables exist:

- Nine endogenous variables: l_t^s, l_t^u, w_t^s, w_t^u, Y_t, L_t^s, L_t^u, b_t, *and* $\frac{N_t}{N_{t-1}}$
- Two state variables: μ_{t-1} *and* s_t, *and*
- Three policy variables: τ_t, μ_t, *and* σ_t

Note that all nine endogenous variables are uniquely solvable given the state variables and the policy variables for period t. It is obvious that the decisions of output produced and labor input are purely static because there is no inter-temporal component in production, such as capital stock in a neoclassical model. However, the demographic structure of this economy is dynamic as the ratio of the native young to the native old at time $t+1$ (N_{t+1} / N_t) and the skill composition of the natives at time $t+1$ depend on today's decisions with regard to immigration policy (μ_t and σ_t). Therefore, the only dynamic component of this model will be the voting decision, which will be discussed in detail in section III.

The equations are further normalized by dividing by N_t, which represents the population of native young workers. The equations above can then be rewritten as follows:

(13) $$l_t^s = [(1-\tau_t) w_t^s]^{\nu}$$

(14) $$l_t^u = [(1-\tau_t) w_t^u]^{\nu}$$

(15) $$\tilde{Y}_t = (\tilde{L}_t^s)^{\alpha} (\tilde{L}_t^u)^{1-\alpha}$$

(16) $$w_t^s = \alpha \tilde{Y}_t (\tilde{L}_t^s)^{-1}$$

(17) $$w_t^u = (1-\alpha) \tilde{Y}_t (\tilde{L}_t^u)^{-1}$$

(18) $$\tilde{L}_t^s = [s_t + \sigma_t \mu_t] N_t l_t^s$$

(19) $$\tilde{L}_t^u = [1 - s_t + (1-\sigma_t)\mu_t] N_t l_t^u$$

(20) $$b_t = \frac{\tau_t \left[w_t^s \tilde{L}_t^s + w_t^u \tilde{L}_t^u \right]}{(1+\mu_t) + (1+\mu_{t-1}) \dfrac{N_{t-1}}{N_t}}$$

(21) $$\frac{N_t}{N_{t-1}} = 1 + n + (1+m)\mu_{t-1}$$

where the variables with tildes are the original variables divided by N_t. For example, $\tilde{L}_t^s$ is the aggregate input of skilled labor per native worker.

III. Political Decision

As discussed in Section II, the policies pertaining to the income tax rate (τ_t), the volume of immigration (μ_t), and the skill composition of immigrants (σ_t) are decided at the beginning of each period by the winner of plurality voting. In this section, the political decision-making process is explained in detail. This modeling of the voting process is basically identical to the application of Besley and Coate (1997) by Suwankiri *et al.* (2016), while there are slight differences in the electorate composition.[8] In modeling the implementation procedure of the winning candidate, the idea of Markov-perfect equilibria is applied.

A. *Political Decision Process*

The political decision-making process of each period consists of the following three steps:

- Step 1: One individual is selected as a candidate in each of three political

[8]For example, Suwankiri *et al.* (2016) simplifies the political equilibria in two aspects. First, they assume that only the smallest party out of the three can choose to vote strategically, thereby eliminating the possibility that the smallest party wins the election. Second, they assume that immigrants have the right to vote when they become old. This simplifies the equilibrium because the young-to-old ratio in the electorate is consistent with that of the whole pre-immigration population. Both simplifications are relaxed in this paper; the smallest party can win an election if the members of the second-largest party have an incentive to vote strategically and the smallest party's members do not have incentive to vote for the second-largest party. Also, immigrants do not have the right to vote for their lifetimes in this paper.

parties (the "Skilled Young," the "Unskilled Young," and the "Old Retirees").

- Step 2: One and only one of the three candidates is elected by one single plurality voting event.[9]
- Step 3: The winning candidate chooses the set of policy variables (τ_t, μ_t, σ_t).

For convenience, these steps are described in reverse order in the following sub-sections.

1. Policy Implementation of the Winning Candidate

Once one of the three candidates is elected by plurality voting, the third and final step of the political process is decision making by the winning candidate. The winning candidate implements his preferred policies on taxation and immigration. As there is no mechanism to resolve commitment problems, all voters recognize *ex ante* that the winning candidate will pursue his optimum after being elected, and all other pledges are considered as 'cheap talk' under any voting equilibrium. Let the optimal policy of the decisive candidate at time t be denoted by a triplet $(\tau_t^d, \mu_t^d, \sigma_t^d) \equiv \Phi_t^d(\mu_{t-1}, s_t)$ for the winning candidate's party $d \in \{s,u,o\}$. Then, Φ_t^d solves the following maximization problem:

$$\Phi_t^d(\mu_{t-1}, s_t) = \text{argmax}_{\tau_t, \mu_t, \sigma_t} V^d(\tau_t, \mu_t, \sigma_t, \Phi_{t+1}),$$

where Φ_{t+1} is the perception of today's winning candidate of the policy variables in the next period. Obviously there can be multiple different rules of perception Φ_{t+1} that satisfy sub-game perfection; thus, the complexity of this optimization problem depends on the multiplicity of the expectation of the policy rule of the winning candidate in the next period. If there are multiple reasonable perceptions on future policies, this period's winning candidate must then consider multiple possibilities stemming from his decision today.

In order to refine this type of complexity and reduce the number of perceived future policy rules, this paper relies on the Markov-perfect property.[10] That is, the perception function Φ_{t+1} is assumed to be a function of the state variable for the period $t+1$ only without depending on anything else, such as signals to eliminate any types of "sunspot" equilibria. With the introduction of this Markov-perfect property, the winning candidate expects that the next period's winning candidate will make a decision according to a rule identical to those of individuals living today. Formally, the winning candidate's problem is expressed as follows

[9]In other words, there is no second voting such as, for example, the run-off primary in the U.S. presidential election.

[10]The Markov-perfection property in this paper is taken directly from Suwankiri *et al.* (2016), which builds on the Markov-perfection setup used in previous works on short-term government and political economy, such as Krusell and Rios-Rull (1996), Grossman and Helpman (1998), and Hassler *et al.* (2003).

$$\Phi^d(\mu_{t-1}, s_t) = \text{argmax}_{\tau_t, \mu_t, \sigma_t} V^d(\tau_t, \mu_t, \sigma_t, \Phi^d(\mu_t, s_{t+1}))$$
$$\text{subject to } s_{t+1} = \frac{(1+n)s_t + (1+m)\sigma_t \mu_t}{1+n+(1+m)\mu_t}$$

if the winning candidate belongs to the cohort $d \in \{s, u\}$. Intuitively, this means that today's winning candidate decides on policy variables with the belief that the winning candidate of the next period will decide on policy variables in exactly the same manner used by this period's candidates, and, in any equilibrium, this belief is precisely realized.

If the winning candidate is an old retiree, his decision rule is simple given that his utility consists only of the consumption of the lump-sum payment this period. Specifically, the decision when winning the election is given by the solution of the following problem:

$$\Phi^o(\mu_{t-1}, s_t) = \text{argmax}_{\tau_t, \mu_t, \sigma_t} V^o(\tau_t, \mu_t, \sigma_t) = b_t,$$

where b_t is represented as a function of only the policy variables and the states by solving (4)~(12).

2. Voting Process

One step backward from the policy implementation procedure in section III-A-1, now assume that one candidate is selected from each of the three political parties. The winning candidate will be determined by single plurality voting, meaning that the result of voting mainly depends on the sizes of the political parties. However, when analyzing the voting process, the possibility of strategic voting must be considered, as there are three parties. That is, if a voter believes that the candidate from his cohort has no possibility of winning the election, he may have an incentive to vote for the next-best candidate in order to prevent the worst candidate from winning the election.

Also, in this model all of the voters in the same cohort have identical preferences, implying that voters in the same political party vote identically. This setup means we must concentrate on the behavior of a representative voter instead of analyzing the complicated voting equilibrium within a cohort.

Formally, let $e_t = (p_t^{is}, p_t^{iu}, p_t^{io}) \in \Delta^2$ denote the voting decision of a representative voter in cohort $i \in \{s, u, o\}$, where p_t^{ij} stands for the probability that the voter in party i votes for the candidate from party j, and Δ^2 is a 2-simplex. Then, voting profile $\mathfrak{e}_t = (\mathfrak{e}_t^s, \mathfrak{e}_t^u, \mathfrak{e}_t^o)$ represents voting equilibrium if it solves the following problem: [11]

[11] It is implicitly assumed that there is no additional cost for the electorate to participate in voting, and only equilibria in which none of the voters abstains are considered.

$$\mathfrak{e}_t^i = \text{argmax}_{e_t^i} \left\{ \sum_{j \in \{s,u,o\}} \mathcal{P}^j \left(e_t^i, \mathfrak{e}_t^{-i} \right) V^i \left(\Phi_t^j, \Phi_{t+1}, \mathfrak{e}_{t+1} \right) \right\}$$

for all $i \in \{s,u,o\}$, where $\mathcal{P}^j(e_t)$ is the probability that the candidate from party j wins the election given voting profile e_t. That is, under a state of voting equilibrium, each voter chooses for whom to vote so that the expected value of himself is maximized given the other voter's decision and the perceived future policy of the winning candidate for the next period.

In order to reduce the possibility of multiple equilibria for a given set of state variables, I only consider the equilibria in which none of the voting decisions is a weakly dominated strategy. In addition, I assume the following tie-breaking rules for the election:

i) If there are two parties with the equally the largest number of votes, the candidate whose votes are from the fewest number of cohorts then becomes the winner.

ii) If two groups are willing to form a coalition by voting for either candidate of the two, the coalition follows the larger party of the two. If their sizes are equal, the tie is broken with equal probability.

iii) If the sizes of all three parties are identical, the tie is broken with equal probability.

These tie-breaking assumptions, together with the assumption that equilibria with weakly dominated strategies are excluded, guarantee that there exists a unique pure-strategy voting equilibrium for each state vector of (μ_{t-1}, s_t). In all of these equilibria, the members of the largest (not necessarily majority) cohort always vote for the candidate from their own party regardless of the formation of a coalition; thus, a political coalition is potentially formed only by the two smallest parties.[12]

B. *Sizes of Political Groups*

Because the size of each of the political groups (groups of skilled young workers, unskilled young workers, and old retirees) is a critical component of the political decision,[13] I start by looking at which group becomes the "largest" or the "majority" for all possible combinations of the state variables (μ_{t-1}, s_t).

The group of native skilled young workers becomes the largest group under the following conditions:

(i) $s_t > \frac{1}{2}$, and

(ii) $[1 + n + (1+m)\mu_{t-1}]s_t > 1$

[12]See Lemma 4 in Appendix C for proof.

[13]As emphasized before, note that the size of the electorate is not the only component, as whether a coalition is formed or not depends not only on the relative size of the two smallest groups but also on their policy profile preferences with regard to the others parties.

Specifically, the group of skilled young workers becomes the largest if it is larger than the group of unskilled young workers and the group of native old retirees. It is clear that the first condition dominates if $n+(1+m)\mu_{t-1}>1$ while the second condition dominates otherwise.

Similarly, the group of unskilled young workers is the largest group under the following two conditions:

(i) $1-s_t>\frac{1}{2}$, and

(ii) $[1+n+(1+m)\mu_{t-1}](1-s_t)>1$

and the first condition dominates if $n+(1+m)\mu_{t-1}>1,$ with the second condition dominating otherwise.

The political party of the native old retirees becomes the largest under the following condition:

$$\max\{[1+n+(1+m)\mu_{t-1}]s_t, [1+n+(1+m)\mu_{t-1}](1-s_t)\}<1$$

Therefore, the relative sizes of the political groups depend on the combination of the state (μ_{t-1}, s_t) and the parameters m (fertility rate of the immigrants) and n (fertility rate of the natives). Specifically, the relative size of the electorate groups depends on the value of $n+(1+m)\mu_{t-1}$, which governs the ratio between the native young and the native old. The following four cases categorize the relative size of each political group at different skill compositions s_t. The important fact here is that the largest group does not always win owing to the possibility that the two smaller parties form a coalition. Accordingly, this section does not describe who the winning candidate of the election is.

Case I. $\frac{N_t}{N_{t-1}}>3$

In this case, the size of the old retirees' party is less than a third of the size of the population of native young workers, meaning that this group can never be the largest regardless of the skill composition of the young generation. Figure 1 displays the relative size of each political group depending on the value of $s_t \in [0,1]$ in this case of $\frac{N_t}{N_{t-1}}>3$. The old retirees' party constitutes the smallest group unless the skill composition is extremely biased toward either skilled or unskilled labor, and either skilled or unskilled workers always become the largest group.

s_t	0	$\frac{1}{1+n+(1+m)\mu_{t-1}}$	$\frac{1}{2}-\frac{1}{2}\frac{1}{1+n+(1+m)\mu_{t-1}}$	$\frac{1}{2}$	$\frac{1}{2}+\frac{1}{2}\frac{1}{1+n+(1+m)\mu_{t-1}}$	$1-\frac{1}{1+n+(1+m))\mu_{t-1}}$	1
Largest	Unskilled*	Unskilled*	Unskilled	Skilled	Skilled*	Skilled*	
2nd Largest	Old	Skilled	Skilled	Unskilled	Unskilled	Old	
Smallest	Skilled	Old	Old	Old	Old	Unskilled	

FIGURE 1. RELATIVE SIZES OF POLITICAL GROUPS FOR DIFFERENT SKILL COMPOSITIONS - CASE I

Note: * More than 50%.

Case II. $2 < \frac{N_t}{N_{t-1}} \leq 3$

This case is similar to Case I in that the old retirees never have a chance to be the largest group given that the population of the native old retirees is still less than half of the population of the native young workers.

s_t	0	$\frac{1}{2}-\frac{1}{2}\frac{1}{1+n+(1+m)\mu_{t-1}}$	$\frac{1}{1+n+(1+m)\mu_{t-1}}$	$\frac{1}{2}$	$1-\frac{1}{1+n+(1+m))\mu_{t-1}}$	$\frac{1}{2}+\frac{1}{2}\frac{1}{1+n+(1+m)\mu_{t-1}}$	1
Largest	Unskilled*	Unskilled	Unskilled	Skilled	Skilled	Skilled*	
2nd Largest	Old	Skilled	Skilled	Unskilled	Unskilled	Old	
Smallest	Skilled	Old	Old	Old	Old	Unskilled	

FIGURE 2. RELATIVE SIZES OF POLITICAL GROUPS FOR DIFFERENT SKILL COMPOSITIONS - CASE II

Note: * More than 50%.

Case III. $1 < \frac{N_t}{N_{t-1}} \leq 2$

This is likely to be the most interesting case. In this case, the native old population is still smaller than the population of native young workers but larger than half of it. Therefore, the group of old retirees has a chance to be the largest group if the number

s_t	0	$\frac{1}{2}-\frac{1}{2}\frac{1}{1+n+(1+m)\mu_{t-1}}$	$1-\frac{1}{1+n+(1+m)\mu_{t-1}}$	$\frac{1}{2}$	$\frac{1}{1+n+(1+m))\mu_{t-1}}$	$\frac{1}{2}+\frac{1}{2}\frac{1}{1+n+(1+m)\mu_{t-1}}$	1
Largest	Unskilled*	Unskilled	Old	Old	Skilled	Skilled*	
2nd Largest	Old	Old	Unskilled	Skilled	Old	Old	
Smallest	Skilled	Skilled	Skilled	Unskilled	Unskilled	Unskilled	

FIGURE 3. RELATIVE SIZES OF POLITICAL GROUPS FOR DIFFERENT SKILL COMPOSITIONS - CASE III

Note: * More than 50%.

of skilled young workers is close enough to the number of unskilled workers. However, because the population of old retirees is still less than half of the total native population as a whole, the old retirees' group cannot constitute the majority for any value of s_t.

Case IV. $\frac{N_t}{N_{t-1}} \leq 1$

In this case, the native old population exceeds that of the native young, meaning that the old retirees' party always constitutes the majority of the electorate.

s_t	0	$\frac{1}{2}$	1
Largest	Old*	Old*	
2nd Largest	Unskilled	Skilled	
Smallest	Skilled	Unskilled	

FIGURE 4. RELATIVE SIZES OF POLITICAL GROUPS FOR DIFFERENT SKILL COMPOSITIONS - CASE IV

Note: * More than 50%.

IV. Political Equilibria

As mentioned in section II, while the political decision of the policy-makers is forward-looking, the equilibrium allocation and prices are purely static in this model because there are no inter-temporal components linking different periods, such as the bond market or capital accumulation. Therefore, the equilibrium allocation can be solved analytically with only a few simplifying conditions. For instance, the equilibrium allocation is solved in a closed form if the following hold true:

(i) The marginal rate of substitution between consumption and leisure is multiplicative in labor input.

(ii) The production function is multiplicative in both types of labor (skilled and unskilled).

Therefore, a Cobb-Douglas production function and our functional form of the utility function enable the equilibrium to be solvable in closed form for a given state variable (μ_{t-1}, s_t) and a set of policy variables $(\tau_t, \mu_t, \sigma_t)$.[14] Specifically, the lump-sum transfer per capita and the wage of each type of labor are represented with respect to only (μ_{t-1}, s_t) and $(\tau_t, \mu_t, \sigma_t)$, as follows:

[14]For more details about the procedure of computing the closed-form equilibrium, see Appendix A.

(22) $$b_t = \frac{\tau_t(1-\tau_t)^{\nu}\alpha^{\alpha\nu}(1-\alpha)^{(1-\alpha)\nu}(s_t+\sigma_t\mu_t)^{a}[1-s_t+(1-\sigma_t)\mu_t]^{1-\alpha}}{1+\mu_t+(1+\mu_{t-1})[1+n+(1+m)\mu_{t-1}]}$$

(23) $$\omega_t^s = \alpha^{\frac{1+\alpha\nu}{1+\nu}}(1-\alpha)^{\frac{(1-\alpha)\nu}{1+\nu}}(s_t+\sigma_t\mu_t)^{-\frac{1-\alpha}{1+\nu}}[1-s_t+(1-\sigma_t)\mu_t]^{\frac{1-\alpha}{1+\nu}}$$

(24) $$\omega_t^u = \alpha^{\frac{\alpha\nu}{1+\nu}}(1-\alpha)^{\frac{1+(1-\alpha)\nu}{1+\nu}}(s_t+\sigma_t\mu_t)^{\frac{\alpha}{1+\nu}}[1-s_t+(1-\sigma_t)\mu_t]^{-\frac{\alpha}{1+\nu}}$$

Therefore, the lifetime utility of individuals in each cohort is also represented only with the state variables and the policy variables by (1) and (2).

A. *Equilibria without Forward-looking*

Before analyzing the equilibria of the full model, I start by demonstrating the equilibria without forward-looking behavior by individuals, which is simply equivalent to assuming $\beta = 0$. This simplified version is considered because

i) All equilibria associated with allocation and voting are solvable in a closed form, meaning that we can analyze all details of the equilibria, and

ii) More importantly, this is a good benchmark for analyzing how the forward-looking motive affects the decisions of the winning candidate and thus the voters' choice after all.

By assuming $\beta = 0$, this model becomes abstract from the inter-temporal components, and the winning candidate's policy decisions become a simple optimization problem of maximizing his instantaneous utility. First, we analyze the preference of each cohort with regard to policy variables, after which we examine the political equilibrium.

1. Preferences of Candidates on Policy Variables

After substituting the equilibrium allocation with closed-form solutions, the utility of an individual in each cohort can be represented by the following indirect utility functions with respect to policy variables and states:

$$V_t^i = b_t + \frac{\left[(1-\tau_t)w_t^i\right]^{1+\nu}}{1+\nu},$$

$$V_t^o = b_t,$$

where b_t and w_t^i take the form in (22)~(24) for $i \in \{s,u\}$. Note that the continuation value is zero for young workers as $\beta = 0$ is assumed. The ideal policies of each political party, therefore, can be analyzed with these indirect utility functions. The following proposition displays the old retiree's preferred policies.

Proposition 1. *If an old retiree wins the election, the implemented policy, as a result of his optimum, is characterized as follows:*[15]

$$(\tau_t, \mu_t, \sigma_t) = \begin{cases} \left(\frac{1}{1+\nu}, \bar{\mu}, 1\right) & \text{if } 0 \le s_t < \alpha - (1-\alpha)\bar{\mu} \\ \left(\frac{1}{1+\nu}, \bar{\mu}, \alpha + \frac{\alpha - s_t}{\bar{\mu}}\right) & \text{if } \alpha - (1-\alpha)\bar{\mu} \le s_t < \alpha + \alpha\bar{\mu} \\ \left(\frac{1}{1+\nu}, \bar{\mu}, 0\right) & \text{if } \alpha + \alpha\bar{\mu} \le s_t \le 1 \end{cases}$$

Proof. See Appendix B.1.

The implication of this proposition is clear: as old retirees do not have any labor endowment, they are pure beneficiaries of the welfare state. Therefore, if the candidate from the old retirees wins the election, he will choose policy variables such that the tax revenue of the government per capita is maximized. Figure 5 displays the relationship between the tax rate and the lifetime value of the old retirees. The tax rate is decided at $1/(1+\nu)$, which is the Laffer rate of the economy. Regarding the skill composition of immigrants (σ_t), the old retiree chooses $\sigma_t = 1$ if s_t is small enough and $\sigma_t = 0$ if s_t large enough, , as production and thus tax revenue is maximized when the proportion of skilled workers in the labor force equals α, the exponent in the Cobb-Douglas production function. To gain a better understanding of this implication, we assume that $\bar{\mu} \to \infty$. Then, the decision of σ_t converges to α, which means that the old retirees want the skill composition of the economy's workers to be consistent with the Cobb-Douglas exponent and the wage rates of skilled and unskilled labor groups to be equalized. Figure 5 shows that

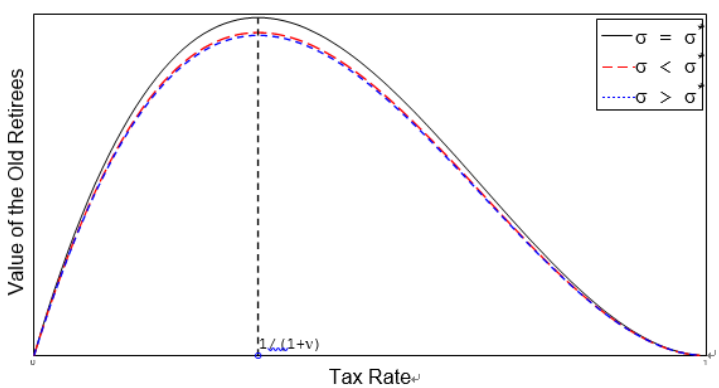

FIGURE 5. TAX RATES AND THE INDIRECT UTILITY OF THE OLD RETIREES

[15]Obviously, some of the ranges of s_t can be an empty set if $\bar{\mu}$ is large enough. For instance, if $\bar{\mu}$ is larger than $\alpha/(1-\alpha)$, the optimal σ_t always equals $\alpha + \frac{\alpha - s_t}{\bar{\mu}}$ because the first and the third intervals become empty sets.

the value of the old retirees is maximized when σ_t is optimal, as in Proposition 1. Finally, the immigration volume is decided at $\mu_t = \bar{\mu}$. This result relies on the assumption of constant-returns-to-scale technology. In other words, as long as the skill composition satisfies the optimality condition, the tax revenue is monotone increasing in the volume of immigration because the production efficiency is robust with regard to the scale of factor inputs.

The next two propositions show the optimality of the young workers with each skill level:

Proposition 2. *If a skilled young worker wins the election, the implemented policy, as a result of his optimum, is characterized as follows:*

$$(\tau_t, \mu_t, \sigma_t) = \begin{cases} (0, \bar{\mu}, 0) \ \ if \ \ 0 \le s_t < \alpha\left[1 + \bar{\mu} + (1+\mu_{t-1})\dfrac{N_{t-1}}{N_t}\right] \\ (\tau_t^s, \bar{\mu}, 0) \ \ if \ \ \alpha\left[1 + \bar{\mu} + (1+\mu_{t-1})\dfrac{N_{t-1}}{N_t}\right] < s_t \le 1 \end{cases}$$

where

$$\tau_t^s = \frac{s_t - \alpha\left[1 + \bar{\mu} + (1+\mu_{t-1})\dfrac{N_{t-1}}{N_t}\right]}{s_t(1+\nu) - \alpha\left[1 + \bar{\mu} + (1+\mu_{t-1})\dfrac{N_{t-1}}{N_t}\right]} \in \left(0, \frac{1}{1+\nu}\right)$$

Proof. See Appendix B.2.

Proposition 3. *If an unskilled young worker wins the election, the implemented policy, as a result of his optimum, is characterized as follows:*

$$(\tau_t, \mu_t, \sigma_t) = \begin{cases} (\tau_t^u, \bar{\mu}, 1) \ \ if \ \ 0 \le s_t < 1 - (1-\alpha)\left[1 + \bar{\mu} + (1+\mu_{t-1})\dfrac{N_{t-1}}{N_t}\right] \\ (0, \bar{\mu}, 1) \ \ if \ \ 1 - (1-\alpha)\left[1 + \bar{\mu} + (1+\mu_{t-1})\dfrac{N_{t-1}}{N_t}\right] < s_t \le 1 \end{cases}$$

where

$$\tau_t^u = \frac{1-s_t-(1-\alpha)\left[1+\bar{\mu}+(1+\mu_{t-1})\frac{N_{t-1}}{N_t}\right]}{(1-s_t)(1+\nu)-(1-\alpha)\left[1+\bar{\mu}+(1+\mu_{t-1})\frac{N_{t-1}}{N_t}\right]} \in \left(0, \frac{1}{1+\nu}\right)$$

Proof. The proof is omitted as it is symmetric to the proof of Proposition 2. Proof is provided simply by substituting s_t with $1-s_t$, σ_t with $1-\sigma_t$, and α with $1-\alpha$ in proposition 2.

The main implication of Propositions 2 and 3 is that young workers consider not only the lump-sum transfer but also their labor income. For example, when s_t is sufficiently small such that there are only a small number of skilled young workers in the economy, the wage rate of the skilled labor is so high that skilled young workers become net contributors to the welfare state, meaning that they prefer a zero tax rate. The upper panel in Figure 6 shows the utility of the skilled young workers for different tax rates and σ_t, the skill composition of the immigrants, when their optimal tax rate equals zero. Given a zero tax rate, the transfers also equal zero; thus, the utility of the skilled young workers depends only on their labor income. Because the wage rate for skilled labor is increasing in the input of unskilled labor due to skill complementarity, skilled workers prefer as many unskilled immigrants as possible.

The lower panel in Figure 6 displays the opposite case, i.e., where s_t is large enough. In this case, skilled workers now become net beneficiaries of the welfare state because there are already excessively many skilled workers such that the equilibrium wage for skilled labor is too small, meaning that skilled workers prefer a strictly positive tax rate τ_t^s. Nevertheless, their ideal tax rate is still smaller than the Laffer rate because they are in any case taxpayers, while the old retirees are not. In this case, how the skilled young workers select σ_t depends on two factors. First, their wage rate increases as more unskilled labor enters the economy, as in the previous case. Second, the direction of the changes of the lump-sum transfer depends

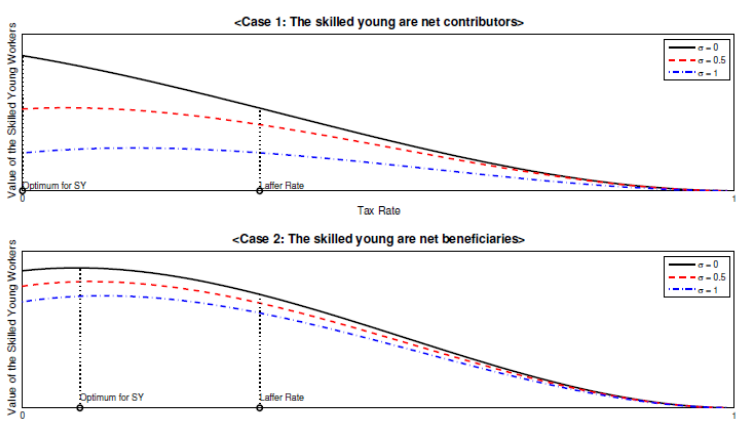

FIGURE 6. TAX RATES AND THE INDIRECT UTILITY OF THE SKILLED YOUNG WORKERS

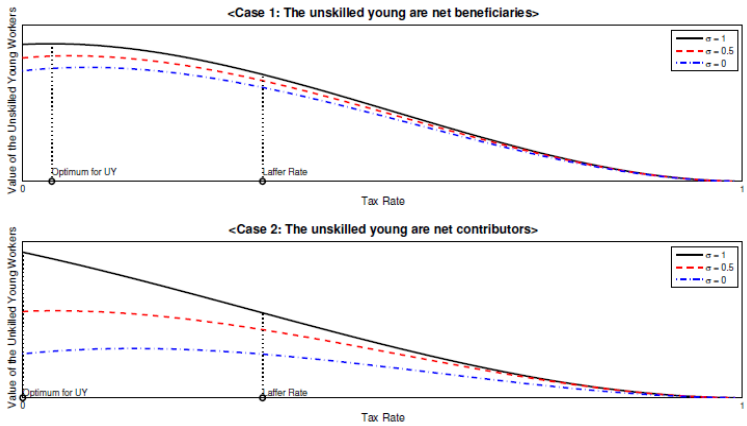

FIGURE 7. TAX RATES AND THE INDIRECT UTILITY OF THE UNSKILLED YOUNG WORKERS

on the pre-immigration skill composition of the economy. Specifically, the transfer increases in σ_t if s_t is relatively small and decreases if s_t is large. Regardless of the size of s_t, however, it can be shown that the wage effect dominates the transfer effect. Thus, skilled workers choose to approve of unskilled immigrants only.

The optimum level of unskilled young workers is symmetric to that of skilled workers, as shown in Figure 7. They set a zero tax rate when s_t is so large that the unskilled workers are net contributors to the welfare state, and they prefer a positive tax rate when s_t is small and they are net beneficiaries of the welfare state. In either case, unskilled workers prefer the maximum number of skilled immigrants.

In summary, old retirees want to maximize the government's tax revenue by levying a high tax rate (at the Laffer rate) and balancing the skill composition of workers for efficient production. On the other hand, young workers prefer a lower (even zero) tax rate and want all immigrants to have opposite skills to maximize their disposable labor income.

2. Political Equilibria

Essentially, the result of the election depends on the relative sizes of the three political parties, as described in Section III-B, while there is no guarantee that the largest party always wins the election given that the two smallest parties potentially form a coalition when the largest party's preferred policy is believed to be the worst for either of the smaller groups. Therefore, not only the relative size of the electorate but also the preference of each party with regard to the preferred policies of one another will determine the result of the election.

The pattern of political equilibria depends on the relative values of the parameters. For instance, a simple case is when $\frac{N_t}{N_{t-1}} > 1$ (the size of the native old population exceeds that of the native young population), in which the candidate of the old retirees' party always wins the election by a majority. While equilibria can be represented in a closed form in this non-forward-looking case, I will describe only

one form of equilibrium with interesting implications rather than considering every possible case of parameter values and states. Specifically, the analysis concentrates on the case in which

i) $1 < \frac{N_t}{N_{t-1}} \leq 2$

ii) α is sufficiently large and $\bar{\mu}$ is sufficiently small such that $\frac{1}{2} < \alpha - (1-\alpha)\bar{\mu} < \alpha(1+\bar{\mu})$

iii) μ_{t-1} is sufficiently large such that $\alpha - (1-\alpha)\left[\bar{\mu} + (1+\mu_{t-1})\frac{N_{t-1}}{N_t}\right] < \frac{1}{2}$

In this case, the patterns of political equilibrium can be described by the range of s_t, as follows:

Case I. $0 \leq s_t \leq \frac{1}{2} - \frac{1}{2} \cdot \frac{N_{t-1}}{N_t}$

In this case, the population of native unskilled young workers constitutes the majority, meaning that the winning candidate's policy decision maximizes the lifetime value for the unskilled young. Because s_t is so small in this case that the unskilled young workers become net beneficiaries of the welfare state, the choice of the winner's policy variables is $\Phi^u(\mu_{t-1}, s_t) = (\tau_t^u, \bar{\mu}, 1)$, where τ_t^u is defined as in Proposition 3. Whether a political coalition arises is immaterial as it cannot take the majority.

Case II. $\frac{1}{2} - \frac{1}{2} \cdot \frac{N_{t-1}}{N_t} \leq s_t \leq 1 - \frac{N_{t-1}}{N_t}$

In this case, the party of the unskilled young workers is still the largest but does not secure a majority. However, the smallest parties (the skilled young workers and the old retirees) do not have any incentive to form a coalition because they perceive that the policies of the other party are the worst. For the old retirees, the preferred policy of the skilled young includes a zero tax rate, which is the worst for the old retirees as it leads to a zero welfare state. For the skilled young workers, the policies of σ_t for both the unskilled young workers and the old retirees are equally the worst because both prefer $\sigma_t = 1$, while the old retirees even prefer a higher tax rate. Therefore, for both the old retirees and the skilled young workers, it is the next-best choice to let the unskilled workers win the election, thereby forgoing the possibility of a political coalition. As a result, the unskilled young workers' party wins the election and the winner implements $(\tau_t, \mu_t, \sigma_t) = (\tau_t^u, \bar{\mu}, 1)$ according to Proposition 3.

Case III. $1-\frac{N_{t-1}}{N_t} < s_t \le \alpha-(1-\alpha)\left[\bar{\mu}+(1+\mu_{t-1})\frac{N_{t-1}}{N_t}\right]$

In this case, the population of old retirees is the largest, the unskilled workers are the second, and the skilled workers form the smallest group. Therefore, it is important as to whether the skilled and unskilled workers have an incentive to form a coalition with each other. It is clear that the skilled workers have an incentive to vote strategically for the unskilled candidate. The preferred tax rate of the unskilled workers is $\tau^u \in (0, 1/(1+\nu))$, which is advantageous for the skilled workers as it levies fewer taxes than the old retirees would levy. Both the old retirees and the unskilled workers prefer the maximum number of skilled immigrants. Therefore, the skilled workers are willing to vote for the political party of the unskilled workers. Because the population of unskilled workers exceeds that of skilled workers, the coalition is formed such that the unskilled candidate represents the overall coalition. Therefore, the unskilled workers win the election.[16] The policy implementation becomes $(\tau_t, \mu_t, \sigma_t) = (\tau^u, \bar{\mu}, 1)$ according to Proposition 3.

Case IV. $\alpha-(1-\alpha)\left[\bar{\mu}+(1+\mu_{t-1})\frac{N_{t-1}}{N_t}\right] < s_t < \frac{1}{2}$

In this case, the population size is still the order of the old, the unskilled, and the skilled. The only difference from the previous case is that now the unskilled workers are net contributors and thus prefer zero taxes, as also wanted by the skilled workers. Therefore, the coalition is formed so that the unskilled party wins the election, and the policy of $(\tau_t, \mu_t, \sigma_t) = (0, \bar{\mu}, 1)$ is implemented according to Proposition 3.

Case V. $\frac{1}{2} \le s_t < \frac{N_{t-1}}{N_t}$

In this case, the old retirees are the largest group, the skilled are the second-largest, and the unskilled are the smallest in size. Also, because both skilled and unskilled workers are net contributors to the welfare state, both groups prefer a zero tax rate; thus, the old retirees equally dislike both young candidates and want neither to win. The skilled workers obviously prefer the unskilled workers' policy profile to that of the old retirees. Therefore, the important problem is whether the unskilled prefer the skilled workers' ideal policy to the old retirees'. If the unskilled workers prefer skilled workers, the skilled candidate would then represent the coalition. Otherwise, the unskilled workers would have no reason to vote for the skilled workers, meaning

[16] I do not consider whether the unskilled workers have incentive to vote strategically for the skilled candidate because the unskilled workers' party is a larger group than that of the skilled workers and the skilled group does have an incentive to vote for the unskilled candidate. In this case, regardless whether the unskilled party's second-best choice is the skilled workers' candidate or not, the coalition is formed such that the unskilled workers' candidate represents the overall coalition according to the equilibrium refinement procedure introduced in Section III-A-2. Obviously, the case in which the skilled party wins the election by forming a coalition with the unskilled workers can be a form of Nash equilibrium if the unskilled workers have incentive to do so.

that the skilled workers strategically vote for the unskilled workers' candidate, and the unskilled candidate will represent the coalition.

For the unskilled workers, the preferred policies of the other two parties have the following trade-offs: Though the skilled workers' preferred policy, denoted by $(\tau_t, \mu_t, \sigma_t) = (0, \bar{\mu}, 0)$, is preferable because it levies zero taxes, it will allow many unskilled immigrants in and make the competition in the unskilled labor market tighter. Though the old retirees' preferred policy, denoted by $(\tau_t, \mu_t, \sigma_t) = (1/(1+\nu), \bar{\mu}, 0)$, is beneficial because it will allow in a maximum number of skilled immigrants, it levies high taxes on labor income. Therefore, the problem is whether the unskilled workers can bear the high tax rate (at the Laffer rate) preferred by the old retirees in order to allow as many skilled immigrants to enter as possible. Formally, this trade-off is characterized by the threshold tax rate, denoted by $\tilde{\tau}(\mu_{t-1}, s_t)$ and determined implicitly by solving for the following equation:

$V^u(\mu_{t-1}, s_t; \tau_t = \tilde{\tau},\ \mu_t = \bar{\mu},\ \sigma_t = 1) = V^u(\mu_{t-1}, s_t; \tau_t = 0,\ \mu_t = \bar{\mu},\ \sigma_t = 0)$, where the left-hand side is the utility of the unskilled workers when the tax rate is at $\tilde{\tau}$ and the maximum number of skilled immigrants are approved, and the right-hand side is the utility of the unskilled workers when the skilled young workers' candidate wins the election.[17] In other words, $\tilde{\tau}$ is the maximum tax rate that the unskilled young workers can bear in return for approving the maximum number of skilled immigrants. If $\tilde{\tau} < 1/(1+\nu)$, then the tax rate to be levied by the old retirees is excessively burdensome to the unskilled, meaning that the unskilled will strategically vote for the skilled workers' party. Otherwise, if $\tilde{\tau} \geq 1/(1+\nu)$, the unskilled workers then prefer the old retirees' policy at least as much as the skilled workers'; thus, a coalition is formed such that the unskilled candidate represents the overall coalition and wins the election.

As a result, the skilled young workers' candidate wins the election with the support of the unskilled party if $\tilde{\tau} < 1/(1+\nu)$, which implements $(\tau_t, \mu_t, \sigma_t) = (0, \bar{\mu}, 0)$ according to Proposition 2. If $\tilde{\tau} \geq 1/(1+\nu)$, then the unskilled workers win the election with the support of the skilled workers and the policy implementation is $(\tau_t, \mu_t, \sigma_t) = (0, \bar{\mu}, 1)$ via Proposition 3.

Case VI. $\dfrac{N_{t-1}}{N_t} \leq s_t < \alpha - (1-\alpha)\bar{\mu}$

In this case, the order of electorate size is as follows: skilled, the old retirees, and then the unskilled. As in the previous case, the old retirees do not have an incentive for strategic voting because all of the young workers prefer a zero tax rate. Therefore, the problem is whether the unskilled are willing to vote strategically for the old candidate or not, and this decision depends on the level of the unskilled workers' threshold tax rate $\tilde{\tau}$. If $\tilde{\tau} \leq 1/(1+\nu)$, he unskilled do not participate in the coalition with the old retirees and the skilled party wins; thus , the policy implementation will be $(\tau_t, \mu_t, \sigma_t) = (0, \bar{\mu}, 0)$ according to Proposition 2. If $\tilde{\tau} > 1/(1+\nu)$, then the old

[17]The existence of this form of $\tilde{\tau}$ is guaranteed by the Intermediate Value Theorem.

retirees win the election with the support of the unskilled workers, and the resulting policy implementation is given by $(\tau_t, \mu_t, \sigma_t) = (1/(1+\nu), \bar{\mu}, 1)$ according to Proposition 1.

Case VII. $\alpha - (1-\alpha)\bar{\mu} < s_t \le \alpha + \alpha\bar{\mu}$

In this case, the order of the electorate size is identical to that in the previous case (the skilled, the old retirees, and then the unskilled). The difference is that now s_t is so large that the old retirees prefer a "balanced" skill composition of immigrants at $\sigma_t = \alpha + \frac{\alpha - s_t}{\bar{\mu}}$ instead of allowing skilled immigrants only. This makes the unskilled workers less favorable to the old retirees, thereby decreasing the threshold tax rate for strategic voting. Formally, the new threshold tax, denoted by $\hat{\tau}(\mu_{t-1}, s_t)$, is determined implicitly by the following equation:

$$V^u\left(\mu_{t-1}, s_t; \hat{\tau}, \bar{\mu}, \alpha + \frac{\alpha - s_t}{\bar{\mu}}\right) = V^u(\mu_{t-1}, s_t; 0, \bar{\mu}, 0)$$

If $\hat{\tau} \le 1/(1+\nu)$, the unskilled do not participate in the coalition with the old retirees and the skilled party wins; thus, the policy implementation will be $(\tau_t, \mu_t, \sigma_t) = (0, \bar{\mu}, 0)$ according to Proposition 2. If $\hat{\tau} > 1/(1+\nu)$, then the old retirees win the election with the support of the unskilled workers; thus, $(\tau_t, \mu_t, \sigma_t) = \left(1/(1+\nu), \bar{\mu}, \alpha + \frac{\alpha - s_t}{\bar{\mu}}\right)$ is implemented via Proposition 1.

Case VIII. $\alpha + \alpha\bar{\mu} < s_t < \frac{1}{2} + \frac{1}{2} \cdot \frac{N_{t-1}}{N_t}$

In this case, the electorate size is identical to that in the previous case (the skilled, the old retirees, and then the unskilled). However, now s_t is so high that the old retirees want to implement $\sigma_t = 0$. Therefore, neither the old retirees nor the unskilled workers have an incentive to support each other. As a result, no coalition is formed and the skilled candidate wins the election as the largest party. The policy of $(\tau_t, \mu_t, \sigma_t) = (0, \bar{\mu}, 0)$ is implemented according to Proposition 2.

Case IX. $\frac{1}{2} + \frac{1}{2} \cdot \frac{N_{t-1}}{N_t} < s_t \le 1$

In this case, the skilled young workers' party constitutes the majority of the electorate. The skilled young candidate wins the election and implements $(\tau_t, \mu_t, \sigma_t) = (0, \bar{\mu}, 0)$ via Proposition 2.

The nine cases above describe the patterns of political equilibria depending on the skill composition and the age composition of the economy. In short, a highly unbalanced skill composition makes the larger skill group more likely win the election, and the old retirees are likely to win when the young-to-old ratio is small enough and the skill composition of the young voters is relatively balanced. The forward-looking case will be shown to exhibit a similar pattern of equilibria in the next section.

B. *Equilibria with Forward-Looking Characteristics*

In this section, the young workers now care not only the instantaneous utility but also for their own future value (alternatively, $\beta > 0$). This results in a trade-off for the young workers. For example, for an unskilled young worker, it is still advantageous to approve of many skilled immigrants because the skill complementarity will increase his wage during this period, as discussed in the previous section, without forward-looking behavior. However, forward-looking behavior now imposes a drawback on this strategy. If too many skilled young immigrants enter the economy during this period, the descendants of those immigrants will be native skilled young workers in the next period (meaning that they have the right to vote in the next period as members of the skilled young party). Therefore, it is highly likely that the skilled young workers' party, which is most hostile to the welfare state, will win the election during the next period. This will deprive the present unskilled workers of the future welfare state after they retire. The unskilled young workers of this period can predict this and thus may choose to approve of only a small number of skilled workers so that the skilled young workers will not constitute a majority in the next period.

Formally, if a young candidate with skill level d wins the election, his policy rule is determined by the following optimization problem:

$$\Phi^d(\mu_{t-1}, s_t) = \operatorname{argmax}_{\tau_t, \mu_t, \sigma_t} \frac{(1-\tau_t)\left[w^j(\tau_t, \mu_t, \sigma_t; \mu_{t-1}, s_t)\right]^{1+\nu}}{1+\nu} + b\left(\tau_t, \mu_t, \sigma_t; \mu_{t-1}, s_t\right) + \beta \cdot b\left(\Phi^d(\mu_t, s_{t+1})\right) \tag{25}$$

subject to

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

An old retiree's decision when he wins is identical to that in the previous section when $\beta = 0$ because his utility does not depend on the future.

Due to the complexity of this model, I rely on a numerical solution for this forward-looking case. Essentially, the solution numerically searches for the

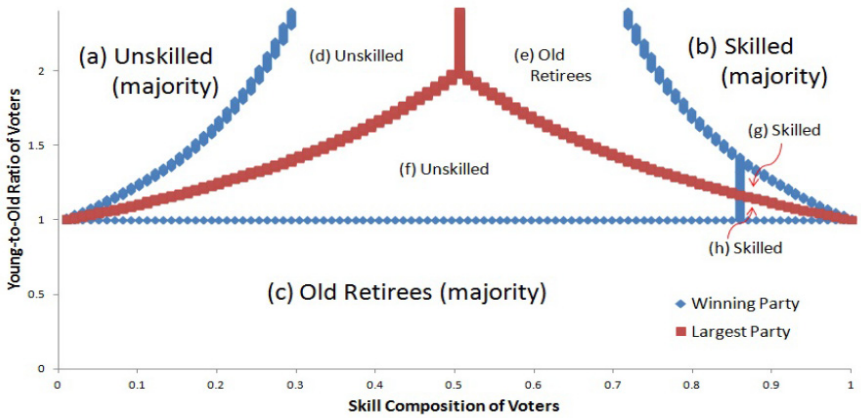

FIGURE 8. WINNING CANDIDATE FOR EACH STATE ($\alpha = 0.8$)

functional form of $\Phi^d(\mu_{t-1}, s_t)$ which solves the functional equation (25).[18] In other words, the solution describes what the winning candidate's optimal choice of the policy profile is given the state vector.

Figure 8 displays which cohort actually wins the election given the state vector (skill composition of the electorate on the horizontal axis and the young-to-old ratio of the electorate on the vertical axis). First, the red line divides the state space into three regions according to in which the largest party is. For instance, regions (a) and (d) are where the unskilled workers' party forms the largest group, regions (b), (e), and (g) are where the skilled workers' party is the largest, and the old retirees' party is the largest in (c), (f), and (h). The blue line further divides the regions according to which group becomes the winner and how this group wins. The following details can explain this:

- In regions (a), (b), and (c), either one of the groups constitutes the majority (more than 50%) of the electorate. Therefore, the political party with the majority wins the election.
- In region (d), the unskilled party is the largest. It wins the election because the two smaller parties (the skilled and the old retirees) are expected to implement a policy profile which is the worst to the other; thus, no coalition is organized.
- In region (e), the old retirees' party wins while the skilled workers group is the largest because the unskilled young workers prefer the old retirees' policy profile to that of the skilled. In other words, the unskilled workers can bear the high tax rates that the old retirees are expected to levy as the old retirees will approve of many skilled immigrants and increase the wages of the unskilled labor.

[18]For the solution to (25), a grid for the state vector (μ, s') is constructed and fixed points are found by iteration. Though the solution is checked for several initial points for robustness, uniqueness is not mathematically proved in this study. It is hoped that this issue will be solved in future studies.

- In region (f), the unskilled workers' party wins the election though the old retirees are the largest because the skilled workers strategically vote for the unskilled candidate to keep the old retirees from winning the election.
- In region (g), the skilled workers' party wins the election as the largest party. The other two parties do not form a coalition.
- In region (h), the skilled workers' party wins the election by forming a coalition with the unskilled. The old retirees' party is the largest but fails to win because their preferred tax rate (the Laffer rate) is excessively burdensome to the unskilled workers.

Therefore, a political coalition is formed in (e), (f), and (h), while the smaller parties let the largest to win without the formation of a coalition in (d) and (g).

Figure 9 shows how the winning party and the equilibrium volume of immigration differ according to the skill composition of the voters for a fixed young-to-old ratio of the electorate. While the old retirees always approve of the maximum number of immigrants, young workers can choose a value smaller than the maximum for certain values of s_t. Note that this occurs solely due to forward-looking behavior, as the desired immigration volume is always the maximum level without the forward-looking component, as shown in Propositions 1~3.

Figures 10~12 describe why unskilled workers prefer $\mu_t < \bar{\mu}$ in some states and $\mu_t = \bar{\mu}$ in others. In Figure 10, the current state is given by point A. Path (i) shows how the economy will move to a new state in the next period if the winning candidate decides to approve of the maximum number of immigrants. As the diagram demonstrates, allowing the maximum number of skilled immigrants will make skilled workers the majority group in the next period, thereby removing the possibility of receiving transfers after they retire. To prevent this, an alternative option for the unskilled workers is to approve of a slightly smaller number of immigrants instead (path ii), which will guarantee that today's unskilled young workers will win the election again in the next period (and thus receive the maximum lump-sum transfer), though it will reduce today's labor income of the unskilled workers relative to path (i). Figure 11 shows an example in which it is too costly to reduce the immigration volume so as to win in the next period if the economy is currently at point B. In this case, the unskilled workers will abandon the possibility of winning in the next period and attempt to maximize today's utility by allowing the maximum number of immigrants, which explains why the immigration volume increases to its maximum at around $s = 0.6$ in Figure 9.

Finally, Figure 12 depicts a case in which it is not feasible for unskilled workers to regain political power in the next period by adjusting the immigration volume. If the skill composition is sufficiently biased toward unskilled labor, unskilled workers cannot make the economy move toward a state in which old retirees will win the election in the next period even if the maximum number of skilled immigrants is accepted. In this case, the unskilled workers' optimal choice is simply to approve of the maximum number of skilled immigrants to maximize the instantaneous piece of their utility.

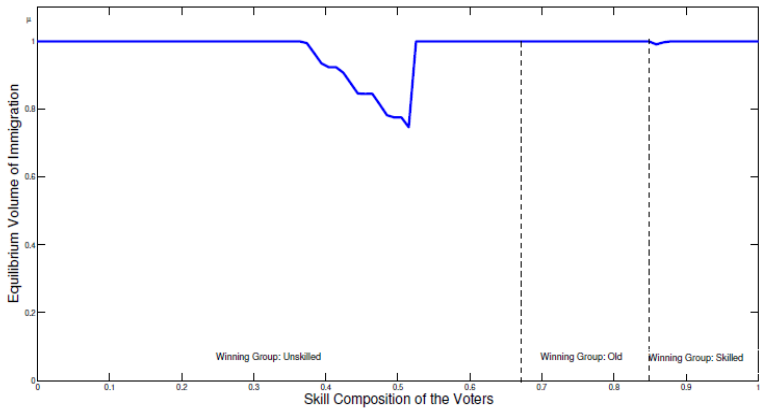

FIGURE 9. DECISION OF IMMIGRATION VOLUME BY SKILL COMPOSITION

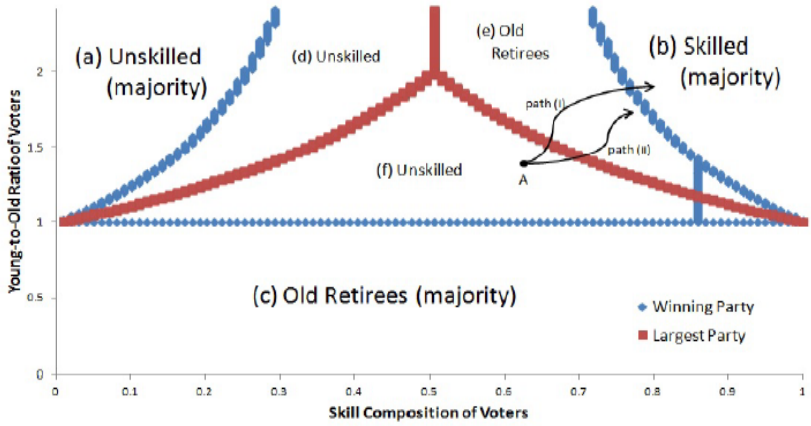

FIGURE 10. CASE I: LESS THAN MAXIMUM VOLUME OF IMMIGRATION IS OPTIMAL

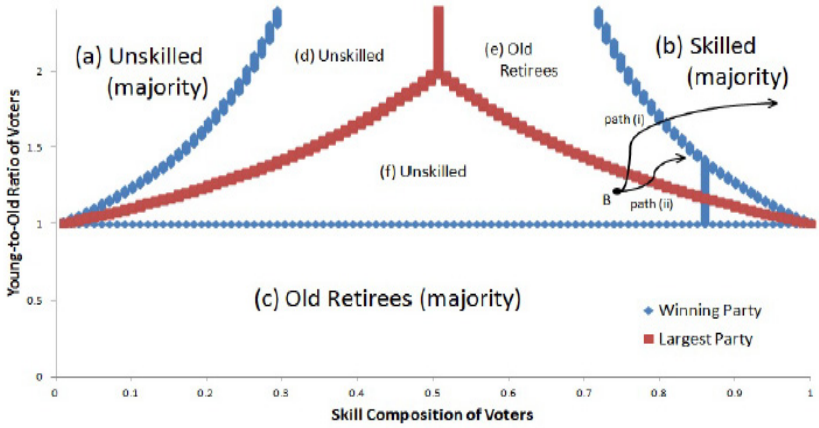

FIGURE 11. CASE II: LESS THAN MAXIMUM VOLUME OF IMMIGRATION IS TOO COSTLY

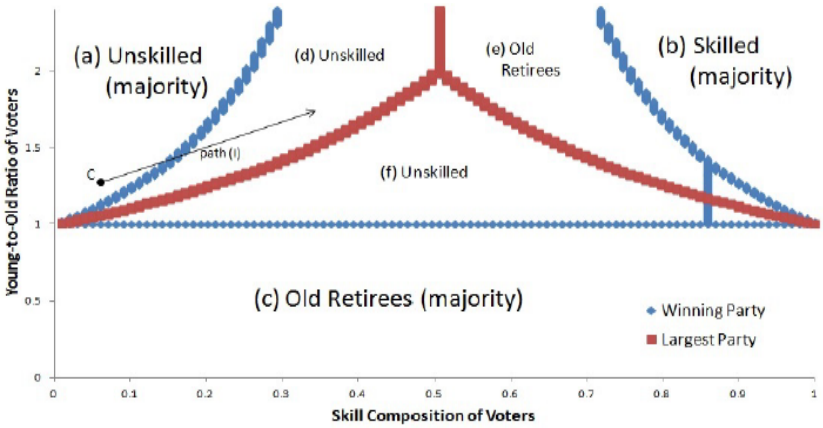

FIGURE 12. CASE III: LESS THAN MAXIMUM VOLUME OF IMMIGRATION OFFERS NO GAIN

V. Dynamics and Social Welfare

In this section, the dynamics of the key macroeconomic variables and their implications are analyzed. According to the analysis of political equilibria shown in the previous section, the skill composition of the economy is the key component which determines who wins the election and how the policies are implemented. Moreover, based on the results, it is expected that the skill composition exhibits cyclical movement; if the skill level of the young natives (i.e., young voters) is biased toward unskilled labor, the skill composition will move toward skilled labor in the next period because the unskilled workers' party is highly likely to win the election and approve of the maximum number of skilled immigrants, whose descendants will be skilled workers in the next period. Similarly, if the proportion of skilled workers is very large, the skill composition will move toward unskilled labor in the next period by the same logic.

This cyclical behavior of the skill composition leads to welfare loss in the economy. Because the production function of this model is the Cobb-Douglas aggregation of skilled and unskilled labor with exponent α for skilled labor, the socially optimal skill composition is obviously such that α is the proportion of skilled labor and $1-\alpha$ is that of unskilled labor. In this sense, the skill composition selected by the old retirees is socially optimal, while the young workers' ideal decision may be suboptimal.

In this model, immigration per se does not cause any inefficiency due to the constant-returns-to-scale technology. In other words, it is socially optimal to approve of as many immigrants as possible as long as the post-immigration skill composition becomes more balanced toward α. Therefore, the welfare loss in this economy is three-fold. First, the skill composition is not always balanced because young workers prefer maximal skill complementarity to maximize their own wages. Second, the proportional labor income tax is distortionary. Third, young workers may not choose to allow the maximum number of immigrants in to regain their political power in the following period.

A. *Welfare Loss due to a Skill Imbalance*

In order to check whether the skill composition exhibits a suboptimal path, the dynamics of the skill composition is examined in this section. Figure 13 displays the dynamics of the post-immigration skill composition of workers when $\alpha = 0.8$, and the initial skill composition equals 0.5. As surmised in the previous paragraphs, the skill composition fluctuates over time. For instance, skilled workers win the election at $t = 2$, decreasing the proportion of skilled workers by allowing the maximum number of unskilled immigrants. This causes, at $t = 3$, the unskilled workers to win the election and pushes the skill composition upward. The old retirees win the election at $t = 4$ because the unskilled workers adjust the immigration volume at $t = 3$ so that they can be the decisive party again when they retire. Given that the old retirees want the skill composition to be balanced at α, the skill composition of immigrants is adjusted such that the skill composition of workers is precisely at α after the immigrants enter the economy. This causes the skilled workers to win the election at $t = 5$, and this pattern is repeated later on.

It is notable that the value of s_t always stays below α. Because α is assumed to be larger than 0.5 for the skill premium, the skilled workers' party is likely to win the election when s_t is close to α. Therefore, the welfare loss from the skill imbalance is mainly caused by the skilled workers in this model. This skill imbalance not only causes a welfare loss but also dampens the effect of immigration in sustaining the welfare state. Given that the welfare state is most efficiently financed when $s_t = \alpha$, $\forall t$, the gap between α and s_t caused by the skilled workers' ideal policy deteriorates the inefficiency of the welfare state.

Furthermore, the fluctuation of the skill composition converges to the same pattern of cycles regardless of the initial skill composition. Figure 14 shows how different initial skill compositions in the end exhibit the same cyclical behavior. The skill composition converges to identical cycles because the value of s_t is pulled into α whenever the old retirees win an election. In other words, whenever the old retirees' party wins an election, the economy is absorbed into the state of $s_t = \alpha$ in the next period.

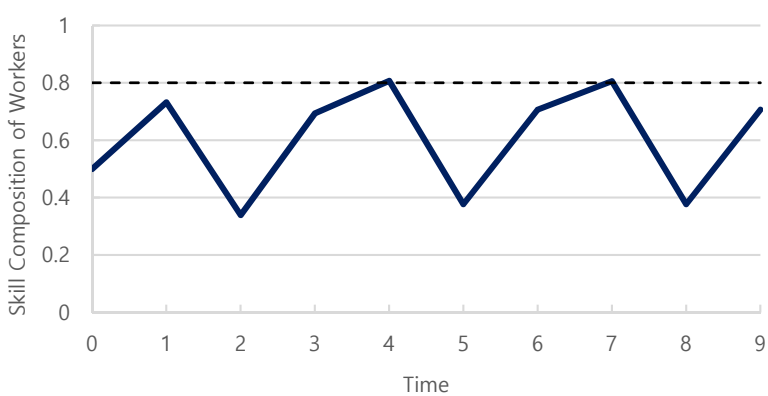

FIGURE 13. DYNAMICS OF THE SKILL COMPOSITION OF WORKERS ($\alpha = 0.8$)

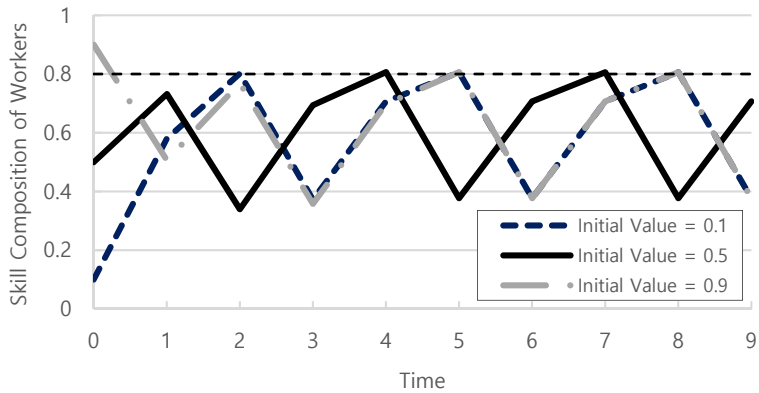

FIGURE 14. DYNAMICS OF THE SKILL COMPOSITION OF WORKERS FOR DIFFERENT INITIAL VALUES

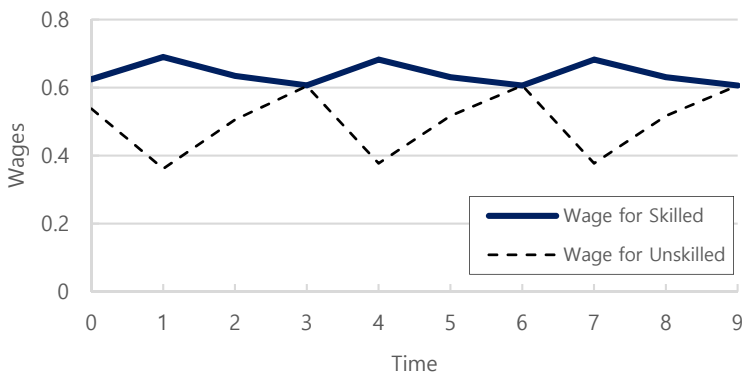

FIGURE 15. DYNAMICS OF PRE-TAX WAGES

Figure 15 displays the dynamics of wage rates for skilled and unskilled labor, respectively. It is clear that the skilled workers implement a policy profile to maximize their own wage (e.g., $t=1$) and that unskilled workers maximize their own wage (e.g., $t=2$). The wage rates are equalized at $t=3, t=6,$ and $t=9$, denoting the period when the old retirees win an election. The implication is simple; as discussed above, the Cobb-Douglas exponent and the actual skill composition will be equalized by the old retirees. At the same time, the income share of the skilled workers is always α according to the property of the Cobb-Douglas technology. Therefore, the wage rates for both skills are equalized when the old retirees win an election and implement the "skill-balancing" immigration policy.

B. *Welfare Loss by Distortionary Taxation*

It is obvious that the labor income tax in this model distorts the labor supply of both skills. This type of welfare loss is mainly caused by the old retirees because they prefer the Laffer rate to maximize the welfare state. As proved in Proposition 2 and Proposition 3, the tax distortion is either zero or is smaller than that caused by

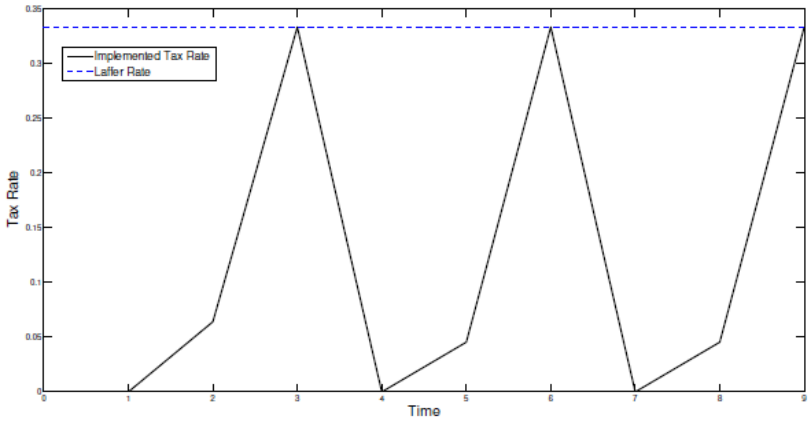

FIGURE 16. DYNAMICS OF TAX RATES

the old retirees if any cohort of the young workers becomes the decisive party. Therefore, the principal contributors of the welfare loss by taxes are the opposite of those in the case of the skill imbalance: the old retirees are the main contributor, and the skilled workers' preferred tax policy is the least distortionary in terms of taxation.[19]

Figure 16 shows the implementation of the tax rate over time. It is straightforward to observe that the old retirees levy $\tau_t = \frac{1}{1+\nu}$ at the Laffer rate, while the skilled workers levy zero taxes. The unskilled young workers implement a small non-zero tax rate when they win, as proved in Proposition 3.

C. *Welfare Loss by an Immigration Volume Adjustment*

In this model, the only reason for the immigration volume to be smaller than $\bar{\mu}$ is the young workers' motive to regain political power after they retire, as discussed in section VI-B. Figure 17 shows the implemented volume of immigration over time. Note that the volume does not reach its maximum at $t = 5$ and $t = 7$, both of which are the periods when the unskilled workers win an election. However, in this numerical example, the gap between $\bar{\mu}$ and the implemented μ_t is very small; moreover, such a value of μ_t that is smaller than $\bar{\mu}$ is still very large (nearly 1, which means that an identical number of immigrants equal to that of the existing native young workers enters the economy).

[19]Although skilled workers potentially prefer a non-zero tax rate, as proved in Proposition 2, it is mostly not realized because these workers prefer a non-zero rate only if $\alpha\left[1+\bar{\mu}+(1+\mu_{t-1})\frac{N_{t-1}}{N_t}\right] < s_t \leq 1$, which is an empty set for most reasonable values of $\frac{N_{t-1}}{N_t}$, $\bar{\mu}$, and α. For instance, the interval becomes empty for the following conservative choice of parameters: $\alpha = 0.8$, $\bar{\mu} = 0$, $\frac{N_{t-1}}{N_t} = 0.25$.

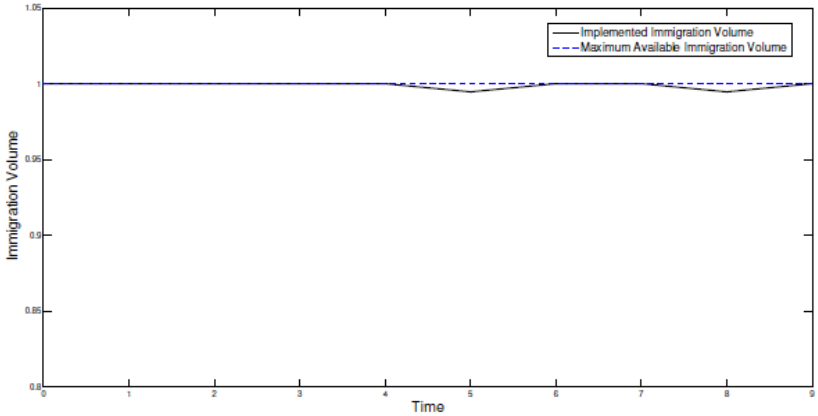

FIGURE 17. DYNAMICS OF THE IMMIGRATION VOLUME

VI. Concluding Remarks

This paper analyzes a politico-economic model of immigration and the welfare state to demonstrate the patterns of the political equilibria and the theoretical implications of sustainability of the welfare state and efficiency of the economy. In this two-period overlapping- generation model, the size of the welfare state, the volume of immigration, and the skill composition of immigrants are determined by the winner of a plurality voting system for each period. Because there are more than two political parties, the possibilities of strategic voting and consequent political coalitions among small political parties are also considered.

In the political processes, there are two key conflicts among the cohorts: the old vs. the young (maximizing the welfare state vs. low taxes) and skilled workers vs. unskilled workers (more unskilled immigrants vs. more skilled immigrants). In equilibrium, the skill composition of immigrants is biased toward unskilled labor relative to a socially efficient level because, at the socially efficient level of skill composition, the skilled workers' party is more likely to win the election. Due to this inefficiency, the effect of the immigration policy on the sustainability of the welfare state is weaker than expected when considering the literature.

Different sources of welfare loss are also analyzed. Not only does the imbalance of the skill composition contribute to the welfare loss but so also do distortionary taxation and the suboptimal volume of immigration. Specifically, the ruling of the old retirees causes maximal tax distortion, as these retirees want to maximize tax revenues by the government. If a young cohort becomes the ruling party, the volume of immigration can be determined at a level lower than the optimal level due to the young workers' incentive to regain political power in the future.

Finally, this study does not consider the possibility that the winning candidate's implementation of immigration policy is not feasible due to, for example, a lack of foreign-born workers who want to migrate to the host country. Also, another interesting case not studied here is one in which there are masses of refugees potentially entering several countries, and the refugees' country choice depends on the policy of each country as well as the economic state, such as the skill composition, aggregate productivity, and demographic structure.

APPENDIX

A. *Full Equilibrium Solution of Allocation and Prices*

If the production function is a typical Cobb-Douglas function in skilled and unskilled labor, the equilibrium can be represented by closed-form expressions.[20] In other words, all instances of equilibrium allocation and prices are represented in closed-form functions of the policy variables (the labor-income tax rate, the volume of immigration, and the skill composition of immigrants), state variables (μ_{t-1} and s_t), and the parameters. For a Cobb-Douglas production function, the equilibrium is characterized by the following nine equations:

(26) $$l_t^s = [(1-\tau_t)w_t^s]^\nu$$

(27) $$l_t^u = [(1-\tau_t)w_t^u]^\nu$$

(28) $$\tilde{Y}_t = (\tilde{L}_t^s)^\alpha (\tilde{L}_t^u)^{1-\alpha}$$

(29) $$\omega_t^s = \alpha \tilde{Y}_t (\tilde{L}_t^s)^{-1}$$

(30) $$\omega_t^u = (1-\alpha)\tilde{Y}_t (\tilde{L}_t^u)^{-1}$$

(31) $$\tilde{L}_t^s = [s_t + \sigma_t \mu_t] l_t^s$$

(32) $$\tilde{L}_t^u = [1 - s_t + (1-\sigma_t)\mu_t] l_t^u$$

(33) $$b_t = \frac{\tau_t \left[\omega_t^s \tilde{L}_t^s + \omega_t^u \tilde{L}_t^u \right]}{(1+\mu_t) + (1+\mu_{t-1})\dfrac{N_{t-1}}{N_t}}$$

(34) $$\frac{N_t}{N_{t-1}} = 1 + n + (1+m)\mu_{t-1}$$

with nine unknowns $\left(l_t^s, l_t^u, \tilde{Y}_t, \tilde{L}_t^s, \tilde{L}_t^u, b_t, \dfrac{N_t}{N_{t-1}}, \omega_t^s, \omega_t^w \right)$, each of which is defined as in section 2.

[20]More rigorously, closed-form solutions are available when the marginal rate of substitution between the labor input and consumption is represented as a multiplicative function of the labor input, and the production function is multiplicative in both skills.

Plugging (29) and (30) into (26) and (27), respectively, leads to

$$(l_t^s)^{\frac{1}{\nu}} = (1-\tau_t)\alpha(\tilde{L}_t^s)^{\alpha-1}(\tilde{L}_t^u)^{1-\alpha}$$
$$(l_t^u)^{\frac{1}{\nu}} = (1-\tau_t)(1-\alpha)(\tilde{L}_t^s)^{\alpha}(\tilde{L}_t^u)^{-\alpha}.$$

Because at this point the equations are expressed in terms of variables for labor input only, this becomes a system of two equations and two unknowns (l_t^s and l_t^u) when substituting $\tilde{L}_t^s$ and $\tilde{L}_t^u$ with the (31) and (32) solutions. After some algebra, labor input per skilled and unskilled worker is represented only by the following policy variables and parameters:

$$l_t^s = (1-\tau_t)^{\nu}\alpha^{\frac{(1+\alpha\nu)\nu}{1+\nu}}(1-\alpha)^{\frac{(1-\alpha)\nu^2}{1+\nu}}(s_t+\sigma_t\mu_t)^{\frac{(\alpha-1)\nu}{1+\nu}}[1-s_t+(1-\sigma_t)\mu_t]^{\frac{(1-\alpha)\nu}{1+\nu}}$$
$$l_t^u = (1-\tau_t)^{\nu}\alpha^{\frac{\alpha\nu^2}{1+\nu}}(1-\alpha)^{\frac{[1+(1-\alpha)\nu]\nu}{1+\nu}}(s_t+\sigma_t\mu_t)^{\frac{\alpha\nu}{1+\nu}}[1-s_t+(1-\sigma_t)\mu_t]^{\frac{-\alpha\nu}{1+\nu}}$$

By plugging these back into (31) and (32), we obtain the following closed-form expressions of total skilled and unskilled labor divided by the native young population ($\tilde{L}_t^s$ and $\tilde{L}_t^u$, respectively):

$$\tilde{l}_t^s = (1-\tau_t)^{\nu}\alpha^{\frac{(1+\alpha\nu)\nu}{1+\nu}}(1-\alpha)^{\frac{(1-\alpha)\nu^2}{1+\nu}}(s_t+\sigma_t\mu_t)^{\frac{1+\alpha\nu}{1+\nu}}[1-s_t+(1-\sigma_t)\mu_t]^{\frac{(1-\alpha)\nu}{1+\nu}}$$
$$\tilde{l}_t^u = (1-\tau_t)^{\nu}\alpha^{\frac{\alpha\nu^2}{1+\nu}}(1-\alpha)^{\frac{[1+(1-\alpha)\nu]\nu}{1+\nu}}(s_t+\sigma_t\mu_t)^{\frac{\alpha\nu}{1+\nu}}[1-s_t+(1-\sigma_t)\mu_t]^{\frac{1+(1-\alpha)\nu}{1+\nu}}$$

By plugging these further back into (28), the output per native young worker ($\tilde{Y}_t$) is also represented in a closed form, as follows:

$$\tilde{Y}_t = (1-\tau_t)^{\nu}\alpha^{\alpha\nu}(1-\alpha)^{(1-\alpha)\nu}(s_t+\sigma_t\mu_t)^{\alpha}[1-s_t+(1-\sigma_t)\mu_t]^{1-\alpha}$$

Consequently the amount of the lump-sum transfer (b_t) is represented in the function of policy variables and states according to the results above.

$$b_t = \frac{\tau_t(1-\tau_t)^{\nu}\alpha^{\alpha\nu}(1-\alpha)^{(1-\alpha)\nu}(s_t+\sigma_t\mu_t)^{\alpha}[1-s_t+(1-\sigma_t)\mu_t]^{1-\alpha}}{(1+\mu_t)+(1+\mu_{t-1})\frac{N_{t-1}}{N_t}} \tag{35}$$

Finally, the wage rates of skilled and unskilled labor are calculated as shown below.

(36)
$$\omega_t^s = \alpha^{\frac{1+\alpha\nu}{1+\nu}}(1-\alpha)^{\frac{(1-\alpha)\nu}{1+\nu}}(s_t+\sigma_t\mu_t)^{-\frac{1-\alpha}{1+\nu}}[1-s_t+(1-\sigma_t)\mu_t]^{\frac{1-\alpha}{1+\nu}}$$
$$\omega_t^u = \alpha^{\frac{\alpha\nu}{1+\nu}}(1-\alpha)^{\frac{1+(1-\alpha)\nu}{1+\nu}}(s_t+\sigma_t\mu_t)^{\frac{\alpha}{1+\nu}}[1-s_t+(1-\sigma_t)\mu_t]^{-\frac{\alpha}{1+\nu}}$$

B. *Proofs*

B.1. Proof of Proposition 1

The old retirees' objective function at time t is simply the lump-sum transfer per capita:

$$V_t^o = b_t$$

Thus, we can differentiate it by each of τ_t, μ_t, and σ_t using (35).

First, the derivative with respect to the tax rate is

(37)
$$\frac{\partial V_t^o}{\partial \tau_t} = \frac{\partial b_t}{\partial \tau_t},$$

(38)
$$= B_1 \cdot [(1-\tau_t)^\nu - \tau_t \nu (1-\tau_t)^{\nu-1}]$$

where

$$B_1 = \frac{\alpha^{\alpha\nu}(1-\alpha)^{(1-\alpha)\nu}(s_t+\sigma_t\mu_t)^\alpha[1-s_t+(1-\sigma_t)\mu_t]^{1-\alpha}}{1+\mu_t+(1+\mu_{t-1})\frac{N_{t-1}}{N_t}}$$
$$> 0$$

Obviously $\frac{\partial V_t^o}{\partial \tau_t} = 0$ at either $\tau_t = 1/(1+\nu)$ or $\tau_t = 1$. However, $\tau_t = 1$ implies no lump-sum payment by (35) and thus zero utility for the old retirees, meaning that $\tau_t = 1/(1+\nu)$ is the optimal tax rate when the old retiree group wins the election.

Second, the derivative of V_t^o with respect to the skill composition of immigrants is as follows:

(39)
$$\frac{\partial V_t^o}{\partial \sigma_t} = \frac{\partial b_t}{\partial \sigma_t},$$
$$= B_2 \cdot (\alpha + \alpha\mu_t - s_t - \sigma_t\mu_t)$$

where

$$B_2 = \frac{\tau_t(1-\tau_t)^{\nu}\alpha^{\alpha\nu}(1-\alpha)^{(1-\alpha)\nu}(s_t+\sigma_t\mu_t)^{\alpha-1}[1-s_t+(1-\sigma_t)\mu_t]^{-\alpha}\mu_t}{1+\mu_t+(1+\mu_{t-1})\dfrac{N_{t-1}}{N_t}}$$

$$>0 \quad (\because 0<\tau_t^* = 1/(1+\nu)<1 \text{ as shown above})$$

Note that $\frac{\partial V_t^o}{\partial \sigma_t}<0$ for any $\sigma_t \in (0,1)$ if $s_t > \alpha+\alpha\mu_t$, and $\frac{\partial V_t^o}{\partial \sigma_t}>0$ for any $\sigma_t \in (0,1)$ if $s_t < \alpha+\alpha\mu_t-\mu_t$. Therefore, the optimal choice of the skill composition of immigrants when the old retirees' party wins an election is expressed as shown below.

(40)
$$\sigma_t = \begin{cases} 1 & if \;\; 0 \le s_t < \alpha+\alpha\mu_t-\mu_t \\ \alpha+\dfrac{\alpha-s_t}{\mu_t} & if \;\; \alpha+\alpha\mu_t-\mu_t \le s_t \le \alpha+\alpha\mu_t \\ 0 & if \;\; \alpha+\alpha\mu_t < s_t \le 1 \end{cases}$$

Finally, the derivative of V_t^o with respect to the volume of immigration is

(41)
$$\frac{\partial V_t^o}{\partial \mu_t} = \frac{\partial b_t}{\partial \mu_t} = B_3\left[\begin{array}{l}\left\{1+\mu_t+(1+\mu_{t-1})\dfrac{N_{t-1}}{N_t}\right\}\{\alpha\sigma_t[1-s_t+(1-\sigma_t)\mu_t] \\ +(s_t+\sigma_t\mu_t)(1-\alpha)(1-\sigma_t)\}-(s_t+\sigma_t\mu_t)[1-s_t+(1-\sigma_t)\mu_t]\end{array}\right],$$

where

$$B_3 = \frac{\tau_t(1-\tau_t)^{\nu}\alpha^{\alpha\nu}(1-\alpha)^{(1-\alpha)\nu}(s_t+\sigma_t\mu_t)^{\alpha-1}[1-s_t+(1-\sigma_t)\mu_t]^{-\alpha}}{\left[1+\mu_t+(1+\mu_{t-1})\dfrac{N_{t-1}}{N_t}\right]^2}$$

$$>0 \quad (\because 0<\tau_t = 1/(1+\nu)<1 \text{ as shown above}).$$

There can be three cases according to the optimal decision for σ_t.

Case I. $s_t \in (\alpha + \alpha\mu_t, 1]$

In this case, $\sigma_t = 0$ as shown in (40). After substitution, (41) is rewritten as follows:

$$\begin{aligned}
\frac{\partial V_t^o}{\partial \mu_t} &= B_3 \cdot \left\{ \left[1 + \mu_t + (1+\mu_{t-1}) \frac{N_{t-1}}{N_t} \right] (1-\alpha) s_t - (1 + \mu_t - s_t) s_t \right\} \\
&= B_3 \cdot s_t \cdot \left\{ \left[1 + \mu_t + (1+\mu_{t-1}) \frac{N_{t-1}}{N_t} \right] (1-\alpha) - (1 + \mu_t - s_t) \right\} \\
&> B_3 \cdot s_t \cdot \left\{ \left[1 + \mu_t + (1+\mu_{t-1}) \frac{N_{t-1}}{N_t} \right] (1-\alpha) - (1 + \mu_t)(1-\alpha) \right\} \\
&\qquad (\because s_t > \alpha + \alpha\mu_t) \\
&> 0
\end{aligned}$$

Therefore, because $\frac{\partial V_t^o}{\partial \mu_t} > 0$ for any value of μ_t, it is optimal to choose $\mu_t = \bar{\mu}$.

Case II. $s_t \in [0, \alpha + \alpha\mu_t - \mu_t)$

In this case, $\sigma_t = 1$ as shown in (40). After substitution, (41) is rewritten as follows:

$$\begin{aligned}
\frac{\partial V_t^o}{\partial \mu_t} &= B_3 \cdot \left\{ \left[1 + \mu_t + (1+\mu_{t-1}) \frac{N_{t-1}}{N_t} \right] \alpha (1 - s_t) - (s_t + \mu_t)(1 - s_t) \right\} \\
&= B_3 \cdot (1 - s_t) \cdot \left\{ \left[1 + \mu_t + (1+\mu_{t-1}) \frac{N_{t-1}}{N_t} \right] \alpha - s_t - \mu_t \right\} \\
&> B_3 \cdot (1 - s_t) \cdot \left\{ \left[1 + \mu_t + (1+\mu_{t-1}) \frac{N_{t-1}}{N_t} \right] \alpha - (1 - \mu_t)\alpha \right\} \\
&\qquad (\because s_t < \alpha + \alpha\mu_t - \mu_t) \\
&> 0
\end{aligned}$$

Therefore, because $\frac{\partial V_t^o}{\partial \mu_t} > 0$ for any value of μ_t, it is optimal to choose $\mu_t = \bar{\mu}$.

Case III. $s_t \in [\alpha + \alpha\mu_t - \mu_t, \alpha + \alpha\mu_t]$

In this case, $\sigma_t = \alpha + \frac{\alpha - s_t}{\mu_t}$ as shown in (40). After reorganization, (41) is rewritten as follows:

$$\frac{\partial V_t^o}{\partial \mu_t} = B_3 \cdot \left\{ \begin{array}{l} \left[1 + \mu_t + (1+\mu_{t-1})\frac{N_{t-1}}{N_t}\right]\sigma_t\{\alpha[1 - s_t + (1-\sigma_t)\mu_t] - (1-\alpha)(s_t + \sigma_t\mu_t)\} \\ + \left[1 + \mu_t + (1+\mu_{t-1})\frac{N_{t-1}}{N_t}\right](s_t + \sigma_t\mu_t)(1-\alpha) - (s_t + \sigma_t\mu_t)[1 - s_t + (1-\sigma_t)\mu_t] \end{array} \right\}$$

After substituting $\sigma_t = \alpha + \frac{\alpha - s_t}{\mu_t}$, the first term equals zero, and the entire expression is simplified into the following form:

$$\frac{\partial V_t^o}{\partial \mu_t} = B_3 \cdot \left\{ \left[1 + \mu_t + (1+\mu_{t-1})\frac{N_{t-1}}{N_t}\right](1-\alpha) - (1+\mu_t)(1-\alpha) \right\} > 0 \tag{42}$$

Therefore the optimal choice is $\mu_t = \bar{\mu}$.

In conclusion, given that $\frac{\partial V_t^o}{\partial \mu_t}$ is strictly positive, the optimal choice of the old retiree is to choose $\mu_t = \bar{\mu}$ regardless of the state of the economy, which ends the proof.

B.2. Proof of Proposition 2

The skilled young worker's objective function at time t consists of two terms: the utility from the lump-sum transfer and the utility from the labor income net of the disutility of working:

$$V_t^s = b_t + \frac{[(1-\tau_t)\omega_t^s]^{1+\nu}}{1+\nu}$$

First, the derivative of V_t^s with respect to the tax rate is expressed as follows:

$$\frac{\partial V_t^s}{\partial \tau_t} = \frac{\partial b_t}{\partial \tau_t} + [(1-\tau_t)\omega_t^s]^{\nu}\left[(1-\tau_t)\frac{\partial \omega_t^s}{\partial \tau_t} - \omega_t^s\right]$$
$$= \frac{\partial b_t}{\partial \tau_t} - (1-\tau_t)^{\nu}(\omega_t^s)^{1+\nu}$$

Because $\frac{\partial \omega_t^s}{\partial \tau_t} = 0$ by (36). Plugging (36) and (38) into this expression yields

$$\frac{\partial V_t^s}{\partial \tau_t} = B_4 \cdot \left[\frac{s_t + \sigma_t \mu_t}{1+\mu_t + (1+\mu_{t-1})\frac{N_{t-1}}{N_t}}[1-(1+\nu)\tau_t] - (1-\tau_t)\alpha\right]$$
$$= B_4 \cdot \left\{\frac{s_t + \sigma_t \mu_t}{1+\mu_t + (1+\mu_{t-1})\frac{N_{t-1}}{N_t}} - \alpha - \left[\frac{s_t + \sigma_t \mu_t}{1+\mu_t + (1+\mu_{t-1})\frac{N_{t-1}}{N_t}}(1+\nu) - \alpha\right]\cdot \tau_t\right\},$$

where

$$B_4 = (1-\tau_t)^{\nu-1}\alpha^{\alpha\nu}(1-\alpha)^{(1-\alpha)\nu}(s_t + \sigma_t\mu_t)^{\alpha-1}[1-s_t + (1-\sigma_t)\mu_t]^{1-\alpha}$$
$$> 0.$$

Therefore, the optimal tax rate for a skilled young worker depends on the state of the economy (s_t, μ_{t-1}), and the following two cases are possible:

Case I. $s_t < \alpha\left[1+\mu_t + (1+\mu_{t-1})\frac{N_{t-1}}{N_t}\right] - \sigma_t\mu_t$

In this case we can observe the following two consequences:

i) $\frac{s_t + \sigma_t \mu_t}{1+\mu_t + (1+\mu_{t-1})\frac{N_{t-1}}{N_t}} - \alpha < 0$, and

ii) $\frac{s_t + \sigma_t \mu_t}{1+\mu_t + (1+\mu_{t-1})\frac{N_{t-1}}{N_t}}(1+\nu) - \alpha$ is either is either a positive number or a negative number smaller in terms of the absolute value compared to $\frac{s_t + \sigma_t \mu_t}{1+\mu_t + (1+\mu_{t-1})\frac{N_{t-1}}{N_t}} - \alpha < 0$.

These two observations guarantee that $\frac{\partial V_t^s}{\partial \tau_t} < 0$. Therefore, the optimal tax rate for a skilled young worker in this case is given by $\tau_t = 0$.

Case II. $s_t \geq \alpha\left[1+\mu_t+(1+\mu_{t-1})\frac{N_{t-1}}{N_t}\right]-\sigma_t\mu_t$

In this case, we have an interior solution at

$$\tau_t = \frac{\dfrac{s_t+\sigma_t\mu_t}{1+\mu_t+(1+\mu_{t-1})\dfrac{N_{t-1}}{N_t}}-\alpha}{\dfrac{s_t+\sigma_t\mu_t}{1+\mu_t+(1+\mu_{t-1})\dfrac{N_{t-1}}{N_t}}(1+\nu)-\alpha}$$

$$= \frac{s_t+\sigma_t\mu_t-\alpha\left[1+\mu_t+(1+\mu_{t-1})\dfrac{N_{t-1}}{N_t}\right]}{(s_t+\sigma_t\mu_t)(1+\nu)-\alpha\left[1+\mu_t+(1+\mu_{t-1})\dfrac{N_{t-1}}{N_t}\right]}$$

$$\in (0, 1)$$

Therefore, these two cases show that the optimal decision of τ_t for the skilled young worker is given by

$$\tau_t = \begin{cases} 0 & \text{if } s_t < \alpha\left[1+\mu_t+(1+\mu_{t-1})\dfrac{N_{t-1}}{N_t}\right]-\sigma_t\mu_t \\ \dfrac{s_t+\sigma_t\mu_t-\alpha\left[1+\mu_t+(1+\mu_{t-1})\dfrac{N_{t-1}}{N_t}\right]}{(s_t+\sigma_t\mu_t)(1+\nu)-\alpha\left[1+\mu_t+(1+\mu_{t-1})\dfrac{N_{t-1}}{N_t}\right]} & \text{if } s_t \geq \alpha\left[1+\mu_t+(1+\mu_{t-1})\dfrac{N_{t-1}}{N_t}\right]-\sigma_t\mu_t \end{cases} \tag{43}$$

The derivative of V_t^s with respect to the skill composition of immigrants is

$$\frac{\partial V_t^s}{\partial \sigma_t} = \frac{\partial b_t}{\partial \sigma_t} + (1-\tau_t)^{1+\nu}(\omega_t^s)^{\nu}\frac{\partial \omega_t^s}{\partial \sigma_t}.$$

Plugging (36) and (39) leads to

(44) $$\frac{\partial V_t^s}{\partial \sigma_t} = B_5 \cdot \left\{ \begin{array}{l} \dfrac{\tau_t(s_t + \sigma_t \mu_t)}{1 + \mu_t + (1+\mu_{t-1})\dfrac{N_{t-1}}{N_t}} (\alpha - s_t + \alpha\mu_t - \sigma_t \mu_t) \\ - \dfrac{(1-\tau_t)(1+\mu_t)\alpha(1-\alpha)}{1+\nu} \end{array} \right\}$$

where

$$B_5 = (1-\tau_t)^{\nu} \mu_t \alpha^{\alpha\nu} (1-\alpha)^{(1-\alpha)\nu} (s_t + \sigma_t \mu_t)^{\alpha-2} [1 - s_t + (1-\sigma_t)\mu_t]^{-\alpha}$$
$$> 0.$$

Again, the optimal decision of σ_t can be separated into two cases, as in (43).

Case I. $s_t < \alpha \left[1 + \mu_t + (1+\mu_{t-1}) \dfrac{N_{t-1}}{N_t} \right] - \sigma_t \mu_t$

In this case, the optimal τ_t is zero according to (43). After plugging $\tau_t = 0$ into (44), obviously $\dfrac{\partial V_t^s}{\partial \sigma_t} < 0$, implying that the optimal σ_t equals zero.

Case II. $s_t \geq \alpha \left[1 + \mu_t + (1+\mu_{t-1}) \dfrac{N_{t-1}}{N_t} \right] - \sigma_t \mu_t$

This case is equivalent to $-s_t - \sigma_t \mu_t \leq -\alpha \left[1 + \mu_t + (1+\mu_{t-1}) \dfrac{N_{t-1}}{N_t} \right]$. Plugging this inequality into (44),

$$\frac{\partial V_t^s}{\partial \sigma_t} \leq B_5 \cdot \left\{ \begin{array}{l} \dfrac{\tau_t(s_t + \sigma_t \mu_t)}{1 + \mu_t + (1+\mu_{t-1})\dfrac{N_{t-1}}{N_t}} \left[\alpha(1+\mu_t) - \alpha \left[1 + \mu_t + (1+\mu_{t-1}) \dfrac{N_{t-1}}{N_t} \right] \right] \\ - \dfrac{(1-\tau_t)(1+\mu_t)\alpha(1-\alpha)}{1+\nu} \end{array} \right\}$$
$$< 0.$$

Because $\dfrac{\partial V_t^s}{\partial \sigma_t} < 0$ for both cases, the optimal decision of σ_t is always zero for the skilled young worker.

Finally, the derivative of the skilled young worker's objective function with respect to the volume of immigration is expressed as follows:

$$\frac{\partial V_t^s}{\partial \mu_t} = \frac{\partial b_t}{\partial \mu_t} + (1-\tau_t)^{1+\nu}(\omega_t^s)^{\nu}\frac{\partial \omega_t^s}{\partial \mu_t}$$

As it was already shown that $\sigma_t = 0$ in the optimum state, we can use (42) to derive the following:

$$\left.\frac{\partial b_t}{\partial \mu_t}\right|_{\sigma_t=0} = B_3 \cdot s_t \cdot \left\{(1-\alpha)\left[1+\mu_t+(1+\mu_{t-1})\frac{N_{t-1}}{N_t}\right]-(1-s_t+\mu_t)\right\}$$

and there can be two cases depending on the value of s_t:

Case I. $s_t < \alpha\left[1+\mu_t+(1+\mu_{t-1})\frac{N_{t-1}}{N_t}\right]$

In this case, $\tau_t = 0$ as shown above; therefore, $b_t = 0$ and thus $\frac{\partial b_t}{\partial \mu_t} = 0$.

Case II. $s_t \geq \alpha\left[1+\mu_t+(1+\mu_{t-1})\frac{N_{t-1}}{N_t}\right]$

In this case:

$$\begin{aligned}\left.\frac{\partial b_t}{\partial \mu_t}\right|_{\sigma_t=0} &= B_3 \cdot s_t \cdot \left\{(1-\alpha)\left[1+\mu_t+(1+\mu_{t-1})\frac{N_{t-1}}{N_t}\right]-(1-s_t+\mu_t)\right\} \\ &\geq B_3 \cdot s_t(1+\mu_{t-1})\frac{N_{t-1}}{N_t} \\ &\geq 0\end{aligned}$$

Therefore, $\frac{\partial b_t}{\partial \mu_t}$ is non-negative in either case. Now, the derivative of the wage rate is written in detail as follows:

$$\frac{\partial \omega_t^s}{\partial \mu_t} = \alpha^{\frac{1+\alpha\nu}{1+\nu}}(1-\alpha)^{\frac{(1-\alpha)\nu}{1+\nu}}\left[\begin{array}{l}-\frac{1-\alpha}{1+\nu}\sigma_t(s_t+\sigma_t\mu_t)^{-\frac{1-\alpha}{1+\nu}-1}\left[1-s_t+(1-\sigma_t)\mu_t\right]^{\frac{1-\alpha}{1+\nu}} \\ +\frac{1-\alpha}{1+\nu}(1-\sigma_t)(s_t+\sigma_t\mu_t)^{-\frac{1-\alpha}{1+\nu}}[1-s_t+(1-\sigma_t)\mu_t]^{\frac{1-\alpha}{1+\nu}-1}\end{array}\right]$$

Substituting $\sigma_t = 0$, it is readily apparent that $\frac{\partial \omega_t^s}{\partial \mu_t} > 0$, thereby implying that $\frac{\partial V_t^s}{\partial \mu_t} > 0$. Therefore, the optimal choice of the volume of immigration for the skilled

young worker is $\bar{\mu}$.

At this point, we return to the intervals of s_t in (43). Because now we know that $\mu_t = \bar{\mu}$ and $\sigma_t = 0$ in optimality, we can substitute μ_t and σ_t with $\bar{\mu}$ and 0, respectively, ending the proof.

C. *Political Coalition and Equilibria*

In this section, I discuss the details of the assumption of the determination process of political equilibria as discussed in section III-A-2. As in the model, there are three political parties, and all of the individuals in one party have identical preferences. Throughout this section, let $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ denote the three political parties in the order of their population sizes and $\#(\mathcal{A})$, $\#(\mathcal{B})$, and $\#(\mathcal{C})$ denote the relative sizes of the three parties, respectively. The sizes are normalized such that $\#(\mathcal{A}) + \#(\mathcal{B}) + \#(\mathcal{C}) = 1$ such that the measure of the size of the overall electorate equals 1.[21] We assume that all parties have at least one member; formally, this is $\#(\mathcal{X}) > 0$ for all $\mathcal{X} \in \{\mathcal{A}, \mathcal{B}, \mathcal{C}\}$.

First, it will be proved that the largest party's members always vote for their own party regardless of whether they constitute the majority or not. It is obvious that $\#(\mathcal{A}) > 0.5$, which is a trivial case. Hence, we assume that it does not exceed 50%. Then, Lemma 4 follows:

Lemma 4. A member of $\mathcal{A}$ votes only for $\mathcal{A}$ in any political equilibrium.

This implies that the members of the largest political party do not vote strategically for a candidate from another party in any of the equilibria after repeatedly eliminating weakly dominating strategies. The proof is shown below.

Proof. It is obvious that the members of $\mathcal{A}$ prefer $\mathcal{A}$'s policy profile most. Accordingly, we can assume, without a loss of generality, that the members of $\mathcal{A}$ prefer $\mathcal{B}$ to $\mathcal{C}$. Then, it is obvious that $\mathcal{C}$ is weakly dominated by the other two strategies. Therefore, we assume that there is a state of equilibrium in which $\mathcal{A}$'s members vote for $\mathcal{B}$. Then, the following possible cases exist:

Case I. $\mathcal{B}$'s members vote for $\mathcal{B}$, and $\mathcal{C}$'s members vote for $\mathcal{C}$.

In this case, for $\mathcal{C}$'s members, voting for $\mathcal{B}$ gives a weakly smaller payoff than voting for $\mathcal{A}$ because if $\mathcal{A}$'s members deviate to voting for $\mathcal{A}$, then $\mathcal{A}$ wins the election.

Case II. $\mathcal{B}$'s members vote for $\mathcal{C}$, and $\mathcal{C}$'s members vote for $\mathcal{B}$.

Similar to Case I.

[21] Therefore, $\#(\mathcal{A}) \geq \#(\mathcal{B}) \geq \#(\mathcal{C})$ holds because A, B, and C are labeled in the order of size.

Case III. $\mathcal{B}$'s members vote for $\mathcal{B}$, and $\mathcal{C}$'s members vote for $\mathcal{B}$.

Then, for the members of $\mathcal{A}$, voting for $\mathcal{A}$ and $\mathcal{B}$ results in an identical payoff because $\mathcal{C}$ wins after all.

Case IV. $\mathcal{B}$'s members vote for $\mathcal{C}$, and $\mathcal{C}$'s members vote for $\mathcal{C}$.

Similar to Case III.

Therefore, in any of the cases above, voting for $\mathcal{B}$ is at least weakly dominated by voting for $\mathcal{A}$, which contradicts the assumption that voting for $\mathcal{B}$ occurs in equilibrium and ends the proof.

Therefore, it is clear that the members of the largest cohort always vote for the candidate of their own cohort. Together with this lemma, the following assumptions refine the political equilibrium so that there is a unique form of pure-strategy equilibrium per state.

(i) Weakly dominated strategies are repeatedly eliminated.

(ii) If there are two parties with the equally the largest number of votes, the candidate whose votes are from the least number of cohorts then becomes the winner.

(iii) Voters strategically vote for the candidate of another cohort only if it is strictly preferred to not doing so.

(iv) Suppose there are two parties. If all members of both parties prefer to form a coalition with each other, the candidate of the larger party out of the two becomes the candidate of the overall coalition. If the sizes of the two parties are identical, the candidate is determined by equal probability.

(v) Any remaining possibility of ties is broken with equal probability.

These assumptions, together with Lemma 4, imply the following results of the election.

Case I. $\#(\mathcal{A}) \geq 0.5$

This is the simplest case to solve, where $\mathcal{A}$'s candidate wins the election by a majority. Note that $\#(\mathcal{A}) = 0.5$ also guarantees $\mathcal{A}$'s victory due to assumption (ii) above.

Case II. $0.5 > \#(\mathcal{A}) > \#(\mathcal{B}) \geq \#(\mathcal{C})$

This is the most interesting case. Because $\mathcal{A}$ does not take the majority, it is important as to whether $\mathcal{B}$ and $\mathcal{C}$ form a political coalition or not, which depends on the preference of $\mathcal{B}$'s and $\mathcal{C}$'s members. The following exclusive and exhaustive cases are possible:

- **Case II-1.** $\mathcal{B}$'s members at least weakly prefer $\mathcal{A}$ to $\mathcal{C}$, and $\mathcal{C}$'s members at least weakly prefer $\mathcal{A}$ to $\mathcal{B}$.

In this case, neither $\mathcal{B}$ nor $\mathcal{C}$ wants a coalition because they perceive the other party's ideal policy to be the worst. Therefore, $\mathcal{A}$ wins the election by receiving the largest number of votes.

- **Case II-2.** $\mathcal{B}$'s members at least weakly prefer $\mathcal{A}$ to $\mathcal{C}$, and $\mathcal{C}$'s members strongly prefer $\mathcal{B}$ to $\mathcal{A}$.

In this case, while $\mathcal{B}$'s members do not have any incentive to vote for $\mathcal{C}$ strategically, $\mathcal{C}$'s members have an incentive to vote for $\mathcal{B}$'s candidate. Therefore, $\mathcal{B}$'s candidate represents the political coalition by $\mathcal{B}$ and $\mathcal{C}$ and wins the election in the equilibrium.

- **Case II-3.** $\mathcal{B}$'s members strongly prefer $\mathcal{C}$ to $\mathcal{A}$, and $\mathcal{C}$'s members at least weakly prefer $\mathcal{A}$ to $\mathcal{B}$.

In this case, $\mathcal{C}$'s candidate represents the coalition by $\mathcal{B}$ and $\mathcal{C}$ wins the election via the same logic presented in Case II-2. Note that $\mathcal{C}$ wins the election despite the fact that it is the smallest cohort in the electorate.

- **Case II-4.** $\mathcal{B}$'s members strongly prefer $\mathcal{C}$ to $\mathcal{A}$, and $\mathcal{C}$'s members strongly prefer $\mathcal{B}$ to $\mathcal{A}$.

In this case, both of the two smaller parties want to form a coalition. If $\#(\mathcal{B}) > \#(\mathcal{C})$, then the candidate of $\mathcal{B}$ represents the entire coalition according to assumption (iv) and wins the election. If $\#(\mathcal{B}) = \#(\mathcal{C})$, then the candidate of the coalition is determined by a coin toss, and the resulting candidate wins the election.

Case III. $\#(\mathcal{A}) = \#(\mathcal{B}) > \#(\mathcal{C})$

In this case, $\mathcal{C}$ serves as the swing cohort. If $\mathcal{C}$'s members strongly prefer $\mathcal{A}$ to $\mathcal{B}$, then $\mathcal{A}$ wins the election by forming a coalition with $\mathcal{C}$. If' $\mathcal{C}$'s members strongly prefer $\mathcal{B}$ to $\mathcal{A}$, $\mathcal{B}$ wins the election similarly. If $\mathcal{C}$'s members are indifferent, then they do not vote strategically, and either $\mathcal{A}$ or $\mathcal{B}$ then wins the election with equal probability (a coin toss).

Case IV. $\#(\mathcal{A}) = \#(\mathcal{B}) = \#(\mathcal{C})$

In this case, each party wins the election with a probability of 1/3.

REFERENCES

Auerbach, Alan J. and Philip Oreopoulos. 1999. "Analyzing the Fiscal Impact of US Immigration," *The American Economic Review*, 89(2): 176-180.

Benhabib, Jess. 1996. "On the political economy of immigration." *European Economic Review*, 40(9): 1737-1743.

Besley, Timothy and Stephen Coate. 1997. "An economic model of representative democracy," *The Quarterly Journal of Economics*, 112(1): 85-114.

Besley, Timothy and Stephen Coate. 1998. "Sources of inefficiency in a representative democracy: a dynamic analysis," *American Economic Review*, 88(1): 139-156.

Boldrin, Michele and Aldo Rustichini. 2000. "Political equilibria with social security," *Review of Economic Dynamics,* 3(1): 41-78.

Dolmas, Jim and Gregory W. Huffman. 2004. "On the political economy of immigration and income redistribution," *International Economic Review,* 45(4): 1129-1168.

Galasso, Vincenzo and Paola Profeta. 2002. "The political economy of social security: a survey," *European Journal of Political Economy,* 18(1): 1-29.

Grossman, Gene and Elhanan Helpman. 1998. "Intergenerational redistribution with short-lived governments," *The Economic Journal*, 108(450): 1299-1329.

Hassler, John, *et al*. 2003. "The survival of the welfare state," *American Economic Review*, 93(1): 87-112.

Hunt, Jennifer and Marjolaine Gauthier-Loiselle. 2010. "How much does immigration boost innovation?" *American Economic Journal: Macroeconomics*, 2(2): 31-56.

Krusell, Per and Jose-Victor Rios-Rull. 1996. "Vested interests in a positive theory of stagnation and growth," *The Review of Economic Studies,* 63(2): 301-329.

Monte, Lindsay M. and Renee Ellis. 2014. *Fertility of women in the United States: 2012*, US Department of Commerce, Economics and Statistics Administration, US Census Bureau.

Ortega, Francesc. 2005. "Immigration quotas and skill upgrading," *Journal of Public Economics,* 89(9-10): 1841-1863.

Peri, Giovanni. 2012. "The effect of immigration on productivity: Evidence from US states," *Review of Economics and Statistics*, 94(1): 348-358.

Storesletten, Kjetil. 2000. "Sustaining Fiscal Policy through Immigration," *Journal of Political Economy*, 108(2): 300-323.

Suwankiri, Benjarong, Assaf Razin, and Efraim Sadka. 2016. "The welfare State and migration: a dynamic analysis of political coalitions," *Research in Economics*, 70(1): 122-142.