Relativity, Inequality, and Optimal Taxation of Internationally Mobile Top Incomes

Darong Dai* Guoqiang Tian†

March 18, 2023

Abstract

We use the Mirrlees income tax model with migration between two countries to investigate the optimal labor income taxation of top earners, in which individuals differ in skill levels and migration costs and we apply intra-country comparisons of private goods consumption and inter-country comparisons of public goods consumption. We consider three scenarios of relative concerns ("relativity") for private goods consumption, including a comparison with the average consumption in the population, an upward comparison, and a comparison with the median-skill type. We first derive the tax formula for the optimal asymptotic marginal tax rate (AMTR) in closed form. We then characterize how inequality-attenuating and externality-correcting functions of the optimal AMTR are respectively shaped by the migration response of top incomes, which may work through both migration elasticity and the equilibrium size of cross-border labor flows. In addition to departures in qualitative features across the three scenarios, we find large quantitative differences in these optimal AMTRs under reasonable circumstances.

Keywords: Nonlinear taxation; top tax rate; relative consumption; corrective tax; tax competition; maxi-min social welfare function.

JEL classification codes: H21, H23, H41.

1 Introduction

Since the seminal works by Veblen (1899), Duesenberry (1949), and Festinger (1954), a growing number of empirical studies find substantial and robust evidence that people have strong preferences for social status and interpersonal comparisons, and feel envy toward others with higher social positions, which defines an important characteristic of human social life.[1] There is some evidence that positional concerns in social networks might even worsen economic inequality.[2] Individuals' concern regarding consumption (or income) relative to others ("relativity") will

*Corresponding author. E-mails: dai.darong@sufe.edu.cn; darongdai@hotmail.com. Institute for Advanced Research, Key Laboratory of Mathematical Economics (SUFE), Ministry of Education of China, Shanghai University of Finance and Economics, Shanghai 200433, People's Republic of China. Financial support from the National Natural Science Foundation of China (NSFC-72003115, NSFC-72192832) is gratefully acknowledged.

†E-mail: gtian@tamu.edu. Department of Economics, Texas A&M University, College Station, TX 77843, USA. Institute for Advanced Studies in Finance and Economics, Hubei University of Economics, Wuhan 430205, People's Republic of China.

[1]See Solnick and Hemenway (1998), Fong (2001), Blanchflower and Oswald (2004), Alpizar, Carlsson, and Johansson-Stenman (2005), Ferrer-i Carbonell (2005), Luttmer (2005), Fliessbach, Weber, Trautner, Dohmen, Sunde, Elger, and Falk (2007), Clark, Frijters, and Shields (2008), Grolleau and Saïd (2008), Dohmen, Falk, Fliessbach, Sunde, and Weber (2011), Heffetz and Frank (2011), and Corazzini, Esposito, and Majorano (2012).

[2]For example, Nishi, Shirado, Rand, and Christakis (2015) conduct experiments to examine the impact of wealth visibility on inequality in experimental social networks, finding that wealth visibility — a sort of revelation

lead to negative externalities.[3] Economists such as Layard (1980) and Frank (2008) highlight the implications of this type of negative positional externality for informing public policy design.

Meanwhile, income inequality has risen in the majority of countries in the past decades of economic globalization,[4] and those earning top incomes often have high degrees of international mobility, which is verified by the observation that a significant proportion of tax evasion and tax avoidance is perpetrated by the wealthy at the upper end of skills and earnings distribution.[5] Economic globalization and income inequality seem to be correlated, and both intensify individuals' relativity concerns through such channels as popular social media networks and international travel,[6] which further increases demand for income redistribution.[7] Therefore, it could be argued that a well-designed income tax policy for an open economy should have both inequality-attenuating and externality-correcting effects, subject to an internationally mobile tax base.

The goal of this article is to design optimal income tax schemes to account for the equity-efficiency tradeoff and the demand for correcting negative positional externalities in an open economy setting with two competing countries connected by mutual migration. The governments in these two countries are benevolent, and adopt the maxi-min social welfare function in the spirit of Rawls (1971).[8] Because the structure of optimal nonlinear income taxation à la Mirrlees (1971) is technically complex, we focus on characterizing optimal asymptotic marginal tax rates (AMTRs) levied on top labor incomes.[9] We seek to identify the channels through which migration might shape the optimal AMTR in equilibrium, and qualitatively analyze how the effect of inequality (respectively, relativity) on the optimal AMTR is shaped by the top income's migration elasticity. Using realistic parameter estimates, we numerically demonstrate how large the difference in the optimal AMTR can be made by migration, inequality, or relativity, and further investigate the problem of how these quantitative differences might vary across

of social positions — has adverse welfare consequences. For instance, in initially more unequal situations, wealth visibility leads to greater inequality than when wealth is not visible. Perhaps surprisingly, the authors find that high initial levels of economic inequality *alone* have relatively few deleterious welfare effects.

[3]In this paper, aligning with the related literature, the following terms are interchangeable: relativity concern, consumption/social comparison, comparison with a reference consumption, relative consumption concern, positional concern, envy, and social status concern.

[4]See Atkinson, Piketty, and Saez (2011), Piketty (2014), Xie and Zhou (2014), Chetty, Grusky, Hell, Hendren, Manduca, and Narang (2017), Stiglitz (2012, 2018), Schwab and Vanham (2021), and Blanchet, Chancel, and Gethin (2022).

[5]See Piketty, Saez, and Stantcheva (2014), Seim (2017), Alstadsæter, Johannesen, and Zucman (2019), and Brülhart, Gruber, Krapf, and Schmidheiny (2022).

[6]There is some evidence to demonstrate the significance of these two channels in the era of globalization. Since 2021, over 4.5 billion people around the world — about 57% of the global population — have used some form of social media (https://www.visualcapitalist.com/ranked-social-networks-worldwide-by-users/). With social media, users can see what others are doing around the world. One of the negative effects is that people tend to develop a negative self-image and begin to devalue their own ways of life when they see others having the wonderful lives represented on social media (www.uopeople.edu/blog/how-social-media-affected-communication/). Additionally, since 1980, the number of international tourist arrivals skyrocketed from 277 million to nearly 1.5 billion in 2019 before the COVID-19 pandemic hit (https://www.statista.com/chart/21793/international-tourist-arrivals-worldwide/).

[7]See Senik (2009), Clark and Senik (2010), and Sznycer, Seal, Sell, Lim, Porat, Shalvi, Halperin, Cosmides, and Tooby (2017).

[8]In a sense, this normative criterion is particularly relevant for developed countries. The demand for unskilled labor in advanced countries declines with trade opening, and with that, potentially higher wages, as multinational corporations may threaten to relocate if individuals with low inter-jurisdictional mobility do not accept lower wages (Stiglitz 2018). As documented by Gereffi and Luo (2014), many jobs held by individuals with the lowest skills are at the bottom of the global value chain and are characterized as low-paid, insecure, and dangerous.

[9]The approach to establish AMTRs is standard in the optimal taxation literature. We first derive a tax formula of second-best marginal tax rates (MTRs) for all skill types, and then the corresponding AMTR is a limit of the tax formula as the typical skill level approaches infinity, under such conditions as an unbounded Pareto distribution of skill types and quasilinear-in-consumption preferences.

the three definitions of relativity concern discussed below.

The benchmark relativity concern of individuals with any given skill type is defined as their utility decreasing from the average consumption in the entire population. Note that emulation-driven people make upward social comparisons; thus, a reasonable variation in the benchmark case is that the utility of a given skill type decreases in the quasi-average consumption of individuals with higher skill types than the given skill type.[10] In addition, in circumstances in which people compare themselves with the middle-income class, we define the relativity concern as the utility of a given skill type decreasing to the consumption (i.e., after-tax income) level of the median-skill type.[11] As such, in the first and third relativity scenarios, all skill types share the same reference consumption, whereas type-dependent reference consumptions are used in the second scenario.

When individual utility is additive in private goods consumption, labor supply, and public goods consumption, and is quasi-linear in private goods consumption, so that income effects are suppressed, we can derive the tax formula for an optimal AMTR in closed form. The tax formulas under alternative relativity scenarios make the qualitative analysis analytically manageable. We next report the major findings.

The first prediction is about how migration affects the optimal AMTR, and also about how this effect might vary across alternative definitions of relativity concern. We show that the optimal AMTR is determined by migration through two channels under a comparison with the average consumption or upward comparison, including the conventional channel of *migration elasticity* and the novel channel of equilibrium cross-border *net* labor flow of the top type,[12] termed *migration magnitude* and defined as the ratio of ex-ante density to ex post (after migration occurs) density of the top skill type. If this ratio is larger than 1, then the country under consideration faces net labor outflow of the top type in equilibrium; otherwise, it faces net labor inflow of the top type. However, under the comparison with the median-skill type, only the conventional channel matters.

This novel channel arises from accounting for relativity concerns in designing income tax policy. It arises in the first two relativity scenarios rather than in the third scenario because only in the former two cases will the top type generate a negative positional/consumption externality and face a positive Pigouvian tax. The Pigouvian tax is composed of two parts. The standard part is an aggregate measure of negative consumption externalities based on the ex-ante mass of top-type individuals, but this measure must be adjusted considering migration via being divided by the ex post mass of top-type individuals because an equilibrium net inflow (outflow) of such individuals will, ceteris paribus, enlarge (shrink) the tax base and hence reduce (increase) the Pigouvian tax facing each top-type individual. As a result, relativity concern and migration magnitude work jointly to enhance each other in determining the Pigouvian tax. The novel channel of migration magnitude emerges only in the first two relativity scenarios because optimal AMTRs under these two circumstances endow the function of externality correction.

Considered separately, these two channels exert influence on the optimal AMTR in the following manner. Under all three relativity scenarios, optimal AMTR decreases in migration elasticity. Intuitively, as top-type individuals become more sensitive in their migration response to a change in the income tax rate, migration-induced tax competition will cause efficiency considerations to have a more significant role than equity considerations in choosing the proper

[10]The formal definition is given in Section 5.

[11]To our knowledge, these sorts of consumption comparisons are comparatively analyzed by Dai (2020) for the first time; however, rather than conducting a normative analysis of income taxation, the research follows a political economy approach.

[12]In our model top-type individuals born in both countries are allowed to differ in migration costs, and subsequently may make different migration decisions. After they have decided on location choices, the equilibrium sizes of top-type labor inflow and labor outflow may differ for each country; thus, only the "net" labor flow matters in equilibrium.

tax rate. Under the first two scenarios, optimal AMTR rises in the term of migration magnitude. Intuitively, in the case of net labor inflow of top type in equilibrium (i.e., the term of migration magnitude is smaller than 1), then the equilibrium net labor inflow decreases if and only if the optimal AMTR rises. Similarly, in the case of facing net labor outflow of top type in equilibrium (i.e., the term of migration magnitude is larger than 1), then the equilibrium net labor outflow decreases if and only if the optimal AMTR goes down.

Given the separate effects of these two channels on *optimally* taxing the top incomes, we then examine their *joint effect* on the optimal AMTR under the first two forms of relativity concerns. We arrive at the second major prediction: the joint effect is positive (negative) if the term of migration magnitude is greater (smaller) than a certain threshold and migration elasticity is below (above) 1. The interpretation is twofold.

First, if both migration elasticity and equilibrium net inflow of top-type individuals have been small, then the channel of migration magnitude dominates the channel of migration elasticity, in the sense that an increase in migration elasticity leads, ceteris paribus, to a greater positive effect of migration magnitude imposed on the optimal AMTR, and an increase in the term of migration magnitude yields a smaller negative effect of migration elasticity imposed on the optimal AMTR. Given that relativity concern and migration magnitude work in tandem and reinforce one another in affecting the optimal AMTR, an increase in the term of migration magnitude generates a larger amount of negative positional externalities for a given intensity of relativity concerns. Therefore, the intuition of this result is: when the amount of negative positional externalities has been large and migration elasticity has been small, then the optimal AMTR should have a more significant role in correcting externalities than in keeping the top type at home.

Second, if migration elasticity has been large while the equilibrium net outflow of top-type individuals has been small, then the channel of migration magnitude is dominated by migration elasticity. This means that an increase in migration elasticity reduces the positive effect of migration magnitude imposed on the optimal AMTR, and an increase in the term of migration magnitude increases the negative effect of migration elasticity imposed on the optimal AMTR, all else remaining constant. The intuition of this result is that when the amount of negative positional externalities has been small, while migration elasticity has been large, then the optimal AMTR should have a more significant role in keeping the top type at home than correcting externalities.

We then characterize how the effect of inequality (relativity) on the optimal AMTR is modified by the top skill type's migration response. The main findings are summarized in three aspects.

First, in a reasonable range of migration elasticity,[13] we show that the larger the migration elasticity is, the smaller the effect of worsening inequality on increasing the optimal AMTR. Conversely, the higher the degree of inequality is, the greater the effect of an increase in migration elasticity on reducing the optimal AMTR. Intensifying migration-induced tax competition will weaken the upward pressure of worsening inequality for taxing top labor incomes. In contrast, increasing inequality strengthens the downward pressure of intensifying tax competition for taxing top labor incomes. This prediction holds true under a comparison with average consumption and with the median-skill type.

The implication is that although a government may adopt the most redistributive social welfare function (maxi-min), the pressure of the threat of migration from the top type dominates the demand for inequality attenuation (or income redistribution) in the current economic environment; thus, the optimal AMTR should feature a more important function for keeping the most skilled resources at home rather than taxing them heavily to reduce income inequality.

[13] Specifically, let the top type's migration elasticity be smaller than 3.

Second, if top-type individuals exhibit a strong relativity concern, and are minimally mobile with a migration elasticity smaller than 1, then an increase in migration elasticity will amplify the positive effect of relativity on the optimal AMTR. Otherwise, an increase in migration elasticity will reduce the positive effect of relativity concern on the optimal tax rate when they have a weak relativity concern and are highly mobile. When considered separately, an increase in international mobility tends to reduce the optimal tax rate, whereas a stronger relativity concern increases the optimal tax rate because of the increased demand for externality correction. The joint effect of these two factors will depend on the top type's status quo in these two dimensions. If they have already exhibited a strong relativity concern and a low degree of international mobility, then the upward pressure of relativity concern will dominate the downward pressure of migration placed on the optimal tax rate, so that migration complements relativity for optimally taxing top earners.

The implication of this result is that as migration-induced tax competition among national policymakers increases from a low level to a high level, migration first complements relativity, in causing the optimal tax rate to have a significant role in correcting externalities. Then, they seem to act as substitutes in shaping the optimal taxation of top labor incomes to capture the tradeoff between correcting the negative positional externality, which calls for a higher tax rate and keeping the highest skilled at home, which subsequently calls for a lower tax rate.

Third, when an increase in the competing country's public goods provision leads to a reduction in the current country's optimal public goods provision, the joint effect of migration elasticity and the competing country's public goods provision on the optimal AMTR is the same as that stated in the second finding. Elaborating further, if inter-country comparison of public goods consumption and intra-country comparison of private goods consumption generate similar effects on the optimal AMTR, then these two kinds of social comparison are interchangeable for *qualitatively* analyzing the joint effect of the top type's international mobility and relativity concern on optimally taxing their labor incomes.

Therefore, the broad implication is that a benevolent government facing an internationally mobile tax base must carefully address the tradeoffs among efficiency, equity, and externality correction, which is especially poignant for taxing top-income earners because the most skilled individuals often exhibit strong social status concerns and have the highest degree of international mobility in the course of economic globalization.[14]

We also provide a coarse numerical illustration to help quantitatively visualize how large the differences in the optimal AMTRs can be made by considering alternative definitions of relativity concern. We assume that preferences are quasi-linear in consumption, the disutility of labor takes an isoelastic functional form, and that the skill distribution obeys an untruncated Pareto distribution.[15] The realistic estimates of relevant parameters are borrowed from existing empirical studies.[16] We consider different plausible scenarios for the equilibrium size of net labor inflow or outflow of the highest skilled individuals. We then conduct controlled numerical exercises and obtain several findings that complement the qualitative characterizations above. In summary, our findings reveal that the distinction of various forms of consumption comparison does lead to numerically significant implications for optimally taxing top-income earners.

The remainder of the paper proceeds as follows. Following a brief literature review in the next section, Section 3 introduces the economic environment. Section 4 derives the main results when reference consumption is defined as the average consumption in the population. Section 5 extends the previous section's analysis to circumstances in which people make an upward consumption comparison, or when they compare themselves with the median-skill type. Section

[14] See, e.g., Docquier and Marfouk (2006), OECD (2008), Boeri, Brucker, Doquier, and Rapoport (2012), and Papademetriou and Sumption (2013).

[15] See, e.g., Diamond and Saez (2011), Boadway and Jacquet (2008), Saez (2001), and Diamond (1998).

[16] See, e.g., Piketty and Saez (2013), Saez, Slemrod, and Giertz (2012), and Clark, Frijters, and Shields (2008).

6 provides a numerical illustration. Section 7 concludes. Proofs of the results are presented in the Appendix.

2 Related Literature

To investigate optimal nonlinear income taxation in Mirrleesian economies with relative consumption concerns and international migrations, this paper contributes to two branches of existing studies.

The first branch of literature analyzes the influence of consumption/income comparison on the design of optimal nonlinear income tax schedules, such as Oswald (1983), Micheletto (2011), Alvarez-Cuadrado and Long (2012), Kanbur and Tuomala (2013), Mujcic and Frijters (2015), Aronsson and Johansson-Stenman (2010, 2013, 2014b, 2015, 2018), and Dai, Gao, and Tian (2020). The articles of Kanbur and Tuomala (2013) and Dai, Gao, and Tian (2020) are closely related to this paper in terms of the framework used and the major questions addressed. To our knowledge, Kanbur and Tuomala (2013) are the first to explicitly address how income inequality and relative consumption together determine the optimal taxation of top labor incomes. Extending to an open economy with international migration à la Lehmann, Simula, and Trannoy (2014), Dai, Gao, and Tian (2020) study how the joint effect of these two factors is shaped by the top income's migration elasticity and by alternative forms of strategic tax competition between two countries. While these two articles provide important insights, our study departs from, and therefore enriches, their analyses in three aspects.

First, they consider only the reference consumption defined as the average consumption in the population. In addition to this benchmark reference consumption, we also consider two reasonable variations — individuals making an upward social comparison or comparing themselves with the middle-income class — and investigate how the main results might vary across the three definitions of relativity concern.

Second, instead of considering social comparisons only in terms of private goods consumption, as in these two papers, a joint design of income taxation and public goods provision in an open economy enables us to incorporate inter-country social comparisons of public goods consumption into the optimum taxation model.[17] This is driven by the observation that the deepening globalization process and the decrease in transportation costs in the past decades have made international travel and communication easier, such that an individual's subjective wellbeing is influenced by both intra-country and inter-country relativity concerns.[18]

Indeed, our analysis reveals that the case of inter-country comparison in public goods consumption calls for an additional condition that is not needed in the case of intra-country comparison in private goods consumption to sharpen the qualitative prediction regarding the joint effect of relativity and migration on optimally taxing top labor incomes. The essential rationale is that public goods provision will also affect the endogenous location choices of individuals in

[17] Aronsson and Johansson-Stenman (2014a) also consider inter-country comparison in public goods consumption, but focus on the issue of the optimal provision of national and global public goods rather than on the design of optimal income tax policy.

[18] For example, using survey data for countries in Western Europe, Becchetti, Castriota, Corrado, and Ricca (2013) find that the contribution of inter-country comparisons to wellbeing increased from the early 1970s to 2002. Piketty (2014) argues that inter-country social comparisons constitute an important part of the motivation behind Thatcher's and Reagan's drastic income tax reductions in the early 1980s. Lehmann, Simula, and Trannoy (2014) argue that the globalization process has facilitated the easy transmission of ideas, meanings, and values across national borders, along with decreases in transportation costs, thereby reducing the barriers to international labor mobility. Consequently, inter-country social comparisons have become more relevant than when international communications and migration were highly restricted by informational, institutional, cultural, and political obstacles. The references cited in Aronsson and Johansson-Stenman (2014a) provide more indicative evidence.

both countries, which will then affect the fiscal revenue collected to support public goods supplies, and so between-country strategic interaction induced by migration must be considered by governments in the very first place.

Third, given that they investigate how inequality and relativity together determine the optimal MTR at the upper end of skill distribution, we focus on a complementary perspective, primarily addressing how the impact of inequality (relativity) on the optimal tax rate is modified by the top type's elasticity and equilibrium magnitude of international migration on the extensive margin. In other words, we investigate how the inequality-attenuating and externality-correcting roles of the optimal MTR at the top will be shaped by their optimal migration response.[19]

The second branch of literature investigates optimal nonlinear income tax models with migration-induced tax competition, such as Mirrlees (1982), Simula and Trannoy (2010, 2012), Morelli, Yang, and Ye (2012), Piketty and Saez (2013), Bierbrauer, Brett, and Weymark (2013), Lehmann, Simula, and Trannoy (2014), Lipatov and Weichenrieder (2015), Blumkin, Sadka, and Shem-Tov (2015), and Dai, Gao, and Tian (2020).[20] Among them, Lehmann, Simula, and Trannoy (2014) and Dai, Gao, and Tian (2020) are closely related to our analysis in the sense of addressing how migration and inequality together determine the optimal MTRs of top labor incomes.[21] This study extends and enriches previous analyses by investigating the joint design of redistributive tax policy and public goods provision, as well as examining alternative specifications of reference consumption, as emphasized previously.

The tax formulas for optimal AMTR obtained by Piketty and Saez (2013) and Lehmann, Simula, and Trannoy (2014) are simpler than ours, and can hence be referenced as special cases for our model. Because of this difference, some insights can only be yielded from our formula. For example, we find that migration will affect the optimal AMTR through both the conventional channel of migration elasticity and the equilibrium size of cross-border labor flows under relativity scenarios in which people compare with the average consumption or make type-dependent upward comparisons. Notably, we demonstrate that the presence of a novel channel of migration magnitude in the formula for optimal AMTR is precisely due to the joint consideration of relativity concerns.

3 The Model

3.1 Set-up

We consider an open economy with two countries, indexed by $i \in \{A, B\}$. The measure of individuals in country i is normalized to 1, while that of the competing country $-i$ is denoted by n_{-i}, where $0 < n_{-i} \leq 1$. Individuals born in country i differ in two dimensions: their skill levels $w \in [\underline{w}, \overline{w}]$, with $0 < \underline{w} < \overline{w} \leq \infty$, and the migration costs, $m \in \mathbb{R}^+$, they support if deciding to live abroad. If an individual faces an infinitely large migration cost, then she is immobile. Skills are distributed according to the function $F_i(w)$, and the density function, $f_i(w) = F_i'(w) > 0$, is differentiable for all $w \in [\underline{w}, \overline{w}]$, with a median at w_m.

[19] In this study, we consider endogenous migration decisions — individuals' optimal location choices must consider cross-country differentials in both income tax policy and public goods provision, in addition to the migration costs incurred.

[20] Wildasin (2006) and Wilson (2009) provide nontechnical discussions on the fiscal policy implications of global competition for mobile resources (labor and capital) for equity and efficiency.

[21] Recently, a growing number of articles are examining the welfare- or revenue-maximizing income taxation of top earners by using dynamic macro-models or considering other relevant frictions (e.g., increasing returns, human capital spillover, and general externalities) associated to tax design, such as Jones (2022), Kindermann and Krueger (2022), Brüggemann (2021), Uribe-Terán (2021), Badel, Huggett, and Luo (2020), Scheuer and Werning (2017), Ales and Sleet (2016), and the related references cited therein. Our study certainly departs from these articles with regard to the model used as well as to the source of frictions considered, and hence it complements these studies in providing additional insights for designing a welfare-maximizing tax policy for top labor incomes.

For each skill type w, we let $\psi_i(m|w)$ and $\Psi_i(m|w) = \int_0^m \psi_i(x|w)dx$ denote the conditional density function and conditional cumulative distribution function of migration costs, respectively. The initial joint density of (m, w) is thus $\psi_i(m|w)f_i(w)$, while $\Psi_i(m|w)f_i(w)$ is the mass of individuals with skill w whose migration costs are lower than m. Following common practice in the literature,[22] we assume that governments cannot observe individuals' *types*, (w, m), and can only condition transfers on earnings y via a nonlinear and twice differentiable tax function, $T_i(\cdot)$, for $i = A, B$. Further, we assume that taxes are levied according to the residence principle, and that international labor mobility induces Nash tax competition between these two governments.

Assume that all individuals have the same utility function:

$$u(c_i, l_i, G_i; \mu_i, G_{-i}) = v(c_i, G_i; \mu_i, G_{-i}) - h(l_i), \tag{1}$$

where c_i is consumption, l_i is labor,[23] G_i is public goods, μ_i is reference consumption, and G_{-i} is country $-i$'s public goods, with $v_c > 0$, $v_{cc} < 0$, $v_{G_i} > 0$, $v_{G_iG_i} < 0$, $v_{cG_i} \geq 0$, $h' > 0$, and $h'' > 0$. We assume that individuals have relativity concerns and exhibit jealousy regarding both private and public goods consumption, and hence we have $v_\mu < 0$, $v_{G_{-i}} < 0$, $v_{c\mu} < 0$, and $v_{G_iG_{-i}} < 0$.[24] That is, an increase in the reference consumption yields, ceteris paribus, a reduction in relative consumption, thereby leading to a lower level of satisfaction/happiness. By the same token, absolute consumption, c_i (or G_i), and reference consumption, μ_i (or G_{-i}), must act as substitutes (i.e., $v_{c\mu} < 0$ and $v_{G_iG_{-i}} < 0$) in determining an individual's subjective wellbeing.

A native individual of country i obtains her income from wages, with the before-tax income denoted by $y_i \equiv wl_i$ under a perfectly competitive wage-setting mechanism. Let $c_i(w)$, $l_i(w)$ and $y_i(w)$ be consumption (or after-tax income), labor supply, and income for an individual with skill type w. Her budget constraint is thus:

$$c_i(w) = y_i(w) - T_i(y_i(w)). \tag{2}$$

We also impose the following reference consumption constraint:

$$\begin{aligned} \mu_i \geq \frac{\int_{\underline{w}}^{\overline{w}} c_i(w) f_i(w) dw}{\int_{\underline{w}}^{\overline{w}} f_i(w) dw} &\equiv \int_{\underline{w}}^{\overline{w}} c_i(w) f_i(w) dw \\ &= \underbrace{\int_{\underline{w}}^{w} c_i(t) f_i(t) dt}_{\text{downward comparison for type } w} + \underbrace{\int_{w}^{\overline{w}} c_i(t) f_i(t) dt}_{\text{upward comparison for type } w} \end{aligned} \tag{3}$$

for any $w \in [\underline{w}, \overline{w}]$. The reference consumption μ_i is no smaller than the average consumption in the population, which can be roughly interpreted as the circumstance in which downward comparison (namely, a comparison with the lower incomes) and upward comparison (namely, a comparison with the higher incomes) are taken into account simultaneously. Taking μ_i, G_i and G_{-i} as given, type-w individuals maximize (1) subject to (2), yielding the first-order condition (FOC):

$$\frac{h'(l_i(w))}{wv_c(c_i(w), G_i; \mu_i, G_{-i})} = 1 - T_i'(y_i(w)). \tag{4}$$

[22]See Mirrlees (1971), Saez (2001), Lehmann, Simula, and Trannoy (2014), Dai, Gao, and Tian (2020), and Dai and Tian (2023).

[23]One may assume that all types of individuals have the same amount of time — e.g., $l_i \in [0, \bar{l}]$ — for allocation between leisure and labor. Given that we focus on interior solutions, this assumption is technically unnecessary.

[24]We assume that individuals care about status only with regard to the consumption goods because the empirical evidence of Carlsson, Johansson-Stenman, and Martinsson (2007) and Solnick and Hemenway (2005) shows that leisure time is the least positional good.

The individual optimality condition (4) means that the retention rate equals the marginal rate of substitution between income and consumption. We denote by $U_i(w)$ her indirect (or gross) utility.

We now proceed to her migration decision. As in Lehmann, Simula, and Trannoy (2014), we assume that migration occurs if and only if $m < U_{-i}(w) - U_i(w)$. After combining the migration decisions made by individuals born in both countries, the ex post density of residents with skill type w in country i can be written as:

$$\tilde{f}_i(w) \equiv \begin{cases} f_i(w) + \Psi_{-i}(\Delta_i(w)|w) f_{-i}(w) n_{-i} & \text{for } \Delta_i(w) \geq 0, \\ (1 - \Psi_i(-\Delta_i(w)|w)) f_i(w) & \text{for } \Delta_i(w) \leq 0 \end{cases} \tag{5}$$

where $\Delta_i(w) \equiv U_i(w) - U_{-i}(w)$. To ensure that $\tilde{f}_i(w)$ is differentiable with respect to Δ_i, we impose the technical restriction that $\psi_i(0|w) f_i(w) = \psi_{-i}(0|w) f_{-i}(w) n_{-i}$, which is verified when the two countries are symmetric, or when there is a fixed cost of migration. We can then define the semi-elasticity of migration and the elasticity of migration, respectively, as:

$$\eta_i(w) \equiv \frac{\partial \tilde{f}_i(w)}{\partial \Delta_i} \frac{1}{\tilde{f}_i(w)} \quad \text{and} \quad \theta_i(w) \equiv c_i(w) \eta_i(w). \tag{6}$$

3.2 Governments

We assume that the government in each country chooses the tax schedule, $T_i(\cdot)$ for $i = A, B$, to maximize the following social welfare function (SWF):

$$\text{SWF}_i = U_i(\underline{w}). \tag{7}$$

It is easy to show that the least-skilled individuals in the bottom of skill distribution are the worst-off ones, namely, $U_i(\underline{w}) = \min\{U_i(w) : w \in [\underline{w}, \overline{w}]\}$. Thus, the benevolent governments adopt the most redistributive objective by maximizing the welfare of the worst-off individuals.

Furthermore, each government faces two sets of constraints. The first is the fiscal budget constraint:

$$\int_{\underline{w}}^{\overline{w}} T_i(y_i(w)) \tilde{f}_i(w) dw \geq G_i, \tag{8}$$

where the per-unit costs of funding public goods are normalized to 1, which is mainly for the simplicity of notation and is nonessential for the following formal analysis. As $\tilde{f}_i$ is a twice differentiable function of the difference in gross utilities $(U_i(w) - U_{-i}(w))$, which depends on the tax policies and supplies of public goods in both countries, the strategic competition of these two countries has a nontrivial role affecting fiscal budget constraints. It is common knowledge that public policies will affect individuals' mobility decisions, which then determine the tax base of each government. In a sense, government budget constraint can be interpreted as accounting for the participation constraints of all skill types. Therefore, as a first priority, and as rational mechanism designers, the two governments must take such effects into consideration when designing public policies regarding income taxes, transfers, and public goods provision.

The second is the set of incentive-compatibility constraints:

$$v(c_i(w), G_i; \mu_i, G_{-i}) - h\left(\frac{y_i(w)}{w}\right) \geq v(c_i(w'), G_i; \mu_i, G_{-i}) - h\left(\frac{y_i(w')}{w}\right) \quad \forall w, w' \in [\underline{w}, \overline{w}]. \tag{9}$$

The necessary conditions for (9) to be satisfied are:

$$\dot{U}_i(w) = h'(l_i(w)) \frac{l_i(w)}{w} \quad \forall w \in [\underline{w}, \overline{w}], \tag{10}$$

which gives the first-order incentive compatibility (FOIC) conditions.

We will focus on the first-order approach of incentive mechanism design. Therefore, taking country $-i$'s tax schedule and public goods provision as given and exploiting the Taxation Principle (Hammond 1979, Guesnerie 1998, Bierbrauer 2011), country i solves the following maximization problem:

$$\max_{\{U_i(w),l_i(w),G_i,\mu_i\}_{w\in[\underline{w},\overline{w}]}} \mathrm{SWF}_i \tag{11}$$

subject to (3), (7), (8) and (10).

4 Optimal AMTR under Comparison with Average Consumption

4.1 Deriving the optimal tax formula

Solving Problem (11), we state the solution in the following lemma.

Lemma 4.1 *Given the public policy choice of country $-i$, we have the following statements for country i.*

(i) The second-best MTRs, $T_i'^*(y_i(\cdot))$, verify:

$$\frac{T_i'(y_i(w))}{1-T_i'(y_i(w))} = \overbrace{\frac{\gamma_i}{\lambda_i}\frac{f_i(w)}{\tilde{f}_i(w)}}^{\textit{Pigouvian-type tax}} + \overbrace{\mathcal{A}_i(w)\mathcal{B}_i(w)\mathcal{C}_i(w)}^{\textit{Mirrleesian-type tax}} \tag{12}$$

where: $\mathcal{A}_i(w)\equiv 1+[l_i(w)h''(l_i(w))/h'(l_i(w))]$, $\mathcal{B}_i(w)\equiv\left[\tilde{F}_i(\overline{w})-\tilde{F}_i(w)\right]/w\tilde{f}_i(w)$,

$$\mathcal{C}_i(w)\equiv\frac{v_c(w)\int_w^{\overline{w}}\left\{\frac{1}{v_c(t)}\left[1+\frac{\gamma_i}{\lambda_i}\frac{f_i(t)}{\tilde{f}_i(t)}\right]-T_i(y_i(t))\eta_i(t)\right\}\tilde{f}_i(t)dt}{\tilde{F}_i(\overline{w})-\tilde{F}_i(w)}, \tag{13}$$

$v_c(w)\equiv v_c(c_i(w),G_i;\mu_i,G_{-i})$ *denotes marginal utility of private goods consumption,* $\tilde{F}_i(w)\equiv\int_{\underline{w}}^{w}\tilde{f}_i(t)dt$ *denotes ex post skill distribution, and*

$$\frac{\gamma_i}{\lambda_i}=-\frac{\int_{\underline{w}}^{\overline{w}}\frac{v_\mu(c_i(w),G_i;\mu_i,G_{-i})}{v_c(c_i(w),G_i;\mu_i,G_{-i})}\tilde{f}_i(w)dw}{1+\int_{\underline{w}}^{\overline{w}}\frac{v_\mu(c_i(w),G_i;\mu_i,G_{-i})}{v_c(c_i(w),G_i;\mu_i,G_{-i})}f_i(w)dw}, \tag{14}$$

with Lagrange multipliers $\gamma_i>0$ and $\lambda_i>0$ associated with the reference consumption constraint and the government budget constraint, respectively.

(ii) Optimal public goods provision, G_i^*, satisfies:

$$\int_{\underline{w}}^{\overline{w}}\frac{v_{G_i}(c_i(w),G_i;\mu_i,G_{-i})}{v_c(c_i(w),G_i;\mu_i,G_{-i})}\tilde{f}_i(w)\left[1+\frac{\gamma_i}{\lambda_i}\frac{f_i(w)}{\tilde{f}_i(w)}\right]dw=1, \tag{15}$$

with γ_i/λ_i given by (14).

Proof. See the Appendix. ■

Using the optimal tax formula (12), we next derive the optimal asymptotic marginal tax rate (AMTR), denoted by $T_i'^*(y_i(\infty))$. We first need to impose the following two assumptions:

Assumption 4.1 $v(c_i(w),G_i;\mu_i,G_{-i})=c_i(w)-\sigma_i(w)\mu_i+v^p(G_i,G_{-i})$, *where* $\sigma_i(w)\in(0,1)$, $\sigma_i'(w)\geq 0$ *for all* $w\in[\underline{w},\overline{w}]$, $v^p_{G_i}>0, v^p_{G_iG_i}<0, v^p_{G_iG_{-i}}<0$, $v^p_{G_{-i}}<0$, *and* $v^p_{G_{-i}G_{-i}}>0$.

We thus focus on a quasilinear-in-consumption utility function. $\sigma_i(w) < 1$ means that absolute consumption (c_i) is more important than is reference consumption (μ_i), which is consistent with the empirical evidence reported by Clark, Frijters, and Shields (2008). $\sigma_i'(w) \geq 0$ reflects the empirical finding that high-income individuals often exhibit stronger relativity concerns than do low-income individuals. And bottom-income individuals are basically pursuing the satisfaction generated entirely by absolute consumption. The restrictions placed on $v^p(\cdot)$ are standard, and they are satisfied by such functions as $(G_i/G_{-i})^\alpha$, for a fixed parameter $\alpha \in (0, 1)$.

Assumption 4.2 *The disutility and distribution functions satisfy:*

(i) F_i obeys a Pareto distribution with the density:

$$f_i(w) = \frac{a_i}{w^{a_i+1}} \frac{1 - F_i(\hat{w})}{\hat{w}^{-a_i} - \overline{w}^{-a_i}} \text{ for } \hat{w} \geq w_{modal} \text{ and } \overline{w} \leq \infty,$$

where w_{modal} is the modal skill level, $\hat{w}$ is the truncation point to the left, and $a_i > 0$ is the Pareto index.

(ii) The disutility function of labor supply is isoelastic with a constant elasticity coefficient: $lh''(l)/h'(l) = \varepsilon > 0$.

(iii) $\lim_{w\uparrow\infty} \frac{f_i(w)}{\tilde{f}_i(w)}$ exists and is denoted by $q_i(\infty)$.

(iv) The limit and derivative operations of the Pigouvian-type tax formula can be exchanged as follows:

$$\begin{aligned} \lim_{w\uparrow\infty} \left[\frac{\mathrm{d}}{\mathrm{d}G_{-i}} \frac{\gamma_i}{\lambda_i} \frac{f_i(w)}{\tilde{f}_i(w)} \right] &= \frac{\mathrm{d}}{\mathrm{d}G_{-i}} \left[\lim_{w\uparrow\infty} \frac{\gamma_i}{\lambda_i} \frac{f_i(w)}{\tilde{f}_i(w)} \right], \\ \lim_{w\uparrow\infty} \left[\frac{\mathrm{d}}{\mathrm{d}\sigma_i} \frac{\gamma_i}{\lambda_i} \frac{f_i(w)}{\tilde{f}_i(w)} \right] &= \frac{\mathrm{d}}{\mathrm{d}\sigma_i} \left[\lim_{w\uparrow\infty} \frac{\gamma_i}{\lambda_i} \frac{f_i(w)}{\tilde{f}_i(w)} \right], \end{aligned}$$

where we assume that $\sigma_i(w) \equiv \sigma_i$ for all $w \in [\underline{w}, \overline{w}]$.

As argued by Saez (2001) and Boadway and Jacquet (2008), Pareto distribution is a good approximation for skills at the upper end, and hence Assumption 4.2(i) is reasonable for deriving the optimal AMTR. $\mathcal{A}_i(w)$ in (12) is a measure of efficiency effect, and is often considered to be constant through imposing Assumption 4.2(ii) (see Diamond 1998, and Boadway and Jacquet 2008). Assumption 4.2(iii)-(iv) are purely technical restrictions.

The following lemma is obtained by sending w to infinity on both sides of equation (12).

Lemma 4.2 *Suppose Assumptions 4.1 and 4.2 hold. Then, the formula for the optimal AMTR under a comparison with the average consumption level is given by*

$$\frac{T_i'^*(y_i(\infty))}{1 - T_i'^*(y_i(\infty))} = \frac{\left(1 + \frac{1+\varepsilon}{a_i}\right) \frac{\gamma_i}{\lambda_i}(\infty) \cdot q_i(\infty) + \frac{1+\varepsilon}{a_i}}{1 + \frac{1+\varepsilon}{a_i} \cdot \theta_i(\infty)}, \tag{16}$$

where $\frac{\gamma_i}{\lambda_i}(\infty) \equiv \frac{\gamma_i}{\lambda_i}\Big|_{\overline{w}=\infty}$ and $\theta_i(\infty) \equiv \lim_{w\uparrow\infty} \theta_i(w)$. Moreover, (16) can be equivalently expressed as

$$T_i'^*(y_i(\infty)) = \frac{\left(1 + \frac{1+\varepsilon}{a_i}\right) \frac{\gamma_i}{\lambda_i}(\infty) \cdot q_i(\infty) + \frac{1+\varepsilon}{a_i}}{\left(1 + \frac{1+\varepsilon}{a_i}\right) \left[1 + \frac{\gamma_i}{\lambda_i}(\infty) \cdot q_i(\infty)\right] + \frac{1+\varepsilon}{a_i} \cdot \theta_i(\infty)}. \tag{17}$$

Proof. See Proposition 3.2 in Dai, Gao, and Tian (2020). ■

From (16) or (17), we find that migration will affect the optimal AMTR through two channels. The first channel is migration elasticity ($\theta_i(\infty)$), which is conventional in the related literature. The second is migration magnitude ($q_i(\infty)$), which can be regarded as measuring the equilibrium cross-border flow of the most skilled individuals, regardless of whether it is net inflow or net outflow. Note that the emergence of this novel channel is due to the joint consideration of relativity concerns. Specifically, $q_i(\infty)$ will disappear from the formula whenever the term associated to relativity concerns is equal to 0; namely, $\frac{\gamma_i}{\lambda_i}(\infty) = 0$.

4.2 Qualitative characterization

To conduct a qualitative analysis of the optimal AMTR established in Lemma 4.2, we obtain the following two propositions. The first proposition provides a comparative static analysis of the optimal AMTR with respect to relevant factors.

Proposition 4.1 *Suppose Assumptions 4.1 and 4.2 hold. Then, we have comparative statics:*

$$\frac{\partial T_i'^*(y_i(\infty))}{\partial \theta_i(\infty)} < 0, \ \frac{\partial T_i'^*(y_i(\infty))}{\partial q_i(\infty)} > 0,$$

and

$$\frac{\partial T_i'^*(y_i(\infty))}{\partial \left(\frac{1+\varepsilon}{a_i}\right)} \begin{cases} > 0 & \textit{iff } \theta_i(\infty) < 1 + \frac{1}{\frac{\gamma_i}{\lambda_i}(\infty) \cdot q_i(\infty)}, \\ < 0 & \textit{iff } \theta_i(\infty) > 1 + \frac{1}{\frac{\gamma_i}{\lambda_i}(\infty) \cdot q_i(\infty)} \end{cases}$$

with

$$\frac{\partial T_i'^*(y_i(\infty))}{\partial \left(\frac{1+\varepsilon}{a_i}\right)} > 0 \quad \textit{if} \quad \theta_i(\infty) \leq 1.$$

Proof. See the Appendix. ■

The main message conveyed by Proposition 4.1 is the following. The optimal AMTR decreases in top type's migration elasticity, which yields that a larger migration elasticity leads, ceteris paribus, to a higher intensity of tax competition between the two countries, thereby imposing a downward pressure on the equilibrium tax rate. Note that $q_i(\infty) = \lim_{w \uparrow \infty} f_i(w)/\tilde{f}_i(w) < 1$ ($q_i(\infty) > 1$) implies that country i faces a net inflow (outflow) of the most skilled individuals. The interpretation of comparative statics of the optimal AMTR with respect to migration magnitude is that whenever country i faces a net inflow of the top type in equilibrium, then the equilibrium net inflow decreases if and only if ("iff") the optimal AMTR rises, all else remaining constant. In contrast, when country i faces a net outflow of the top type in equilibrium, then the equilibrium net outflow decreases iff the optimal AMTR decreases.

Furthermore, the optimal AMTR increases in the degree of inequality ($1/a_i$), iff migration elasticity is smaller than a certain threshold; otherwise, it decreases in the degree of inequality. In particular, if migration elasticity is smaller than 1 (i.e., the top type can be seen as inelastic in terms of their migration response to the income tax rate), then the optimal AMTR always increases in the degree of inequality; that is, the optimal tax policy always features inequality attenuation for equity considerations when the top type somehow lacks mobility.

This comparative static result demonstrates the equity-efficiency tradeoff that the optimal AMTR must consider in an open economy with international labor mobility. If the out-migration threat from the highest skilled individuals is mild, then an increase in the tax rate represents an optimal response to worsening inequality because the equity gains brought by the tax increase outweigh efficiency losses; otherwise, the equity gains will be outweighed by the corresponding efficiency losses, and so the optimal AMTR no longer exhibits the function of reducing inequality.

In addition, since we have established the range of migration elasticity as the necessary and sufficient condition for the comparative static result to be true, the prediction can be interpreted from a reverse direction. If the optimal AMTR does (does not) feature inequality attenuation, in the sense of responding to a higher degree of inequality with a higher (lower) tax rate, then the top type's endogenous migration elasticity must be lower (higher) than the threshold.

The second proposition characterizes the joint effects of relevant factors on the optimal AMTR.

Proposition 4.2 *Suppose Assumptions 4.1 and 4.2 hold. Then, we have the following predictions.*

(i) Concerning the joint effect of migration elasticity and migration magnitude on the optimal AMTR, we have:

$$\frac{\partial^2 T_i'^*(y_i(\infty))}{\partial\theta_i(\infty)\partial q_i(\infty)} > 0 \quad \text{if } \frac{\gamma_i}{\lambda_i}(\infty)q_i(\infty) > \frac{a_i}{1+\varepsilon+a_i} \text{ and } \theta_i(\infty) \leq 1,$$

$$\frac{\partial^2 T_i'^*(y_i(\infty))}{\partial\theta_i(\infty)\partial q_i(\infty)} < 0 \quad \text{if } \frac{\gamma_i}{\lambda_i}(\infty)q_i(\infty) < \frac{a_i}{1+\varepsilon+a_i} \text{ and } \theta_i(\infty) \geq 1.$$

(ii) Concerning the joint effect of migration elasticity and the degree of inequality, we have:

$$\frac{\partial^2 T_i'^*(y_i(\infty))}{\partial\theta_i(\infty)\partial\left(\frac{1}{a_i}\right)} < 0 \quad \text{for } \theta_i(\infty) \leq 3.$$

(iii) Further assume that $F_i(w) = F_{-i}(w)$ and $\int_{\underline{w}}^{\infty} \tilde{f}_i(w)dw = 1$, then we have:

(iii-a) *Concerning the joint effect of migration elasticity and the intensity of relativity concern:*

$$\frac{\partial^2 T_i'^*(y_i(\infty))}{\partial\theta_i(\infty)\partial\sigma_i} > 0 \quad \text{if } \frac{\gamma_i}{\lambda_i}(\infty)q_i(\infty) > \frac{a_i}{1+\varepsilon+a_i} \text{ and } \theta_i(\infty) \leq 1,$$

$$\frac{\partial^2 T_i'^*(y_i(\infty))}{\partial\theta_i(\infty)\partial\sigma_i} < 0 \quad \text{if } \frac{\gamma_i}{\lambda_i}(\infty)q_i(\infty) < \frac{a_i}{1+\varepsilon+a_i} \text{ and } \theta_i(\infty) \geq 1,$$

where $\frac{\gamma_i}{\lambda_i}(\infty)q_i(\infty)$ is strictly increasing in σ_i.

(iii-b) *Concerning the joint effect of migration elasticity and the competing country's public goods provision: with $\mathrm{d}G_i^*/\mathrm{d}G_{-i} < 0$,*

$$\frac{\partial^2 T_i'^*(y_i(\infty))}{\partial\theta_i(\infty)\partial G_{-i}} > 0 \quad \text{if } \frac{\gamma_i}{\lambda_i}(\infty)q_i(\infty) > \frac{a_i}{1+\varepsilon+a_i} \text{ and } \theta_i(\infty) \leq 1,$$

$$\frac{\partial^2 T_i'^*(y_i(\infty))}{\partial\theta_i(\infty)\partial G_{-i}} < 0 \quad \text{if } \frac{\gamma_i}{\lambda_i}(\infty)q_i(\infty) < \frac{a_i}{1+\varepsilon+a_i} \text{ and } \theta_i(\infty) \geq 1,$$

where $\frac{\gamma_i}{\lambda_i}(\infty)q_i(\infty)$ is strictly increasing in G_{-i}. In particular, we have:

$$\frac{\mathrm{d}G_i^*}{\mathrm{d}G_{-i}} < 0 \text{ iff } \epsilon_{Q_i(w),G_i} < -\epsilon_{v_{G_i}^p,G_i},$$

where $\epsilon_{Q_i(w),G_i} > 0$ and $\epsilon_{v_{G_i}^p,G_i} < 0$ denote the elasticities of $Q_i(w)$ and $v_{G_i}^p$ with respect to G_i, respectively, with

$$Q_i(w) \equiv \tilde{f}_i(w) + \frac{\gamma_i}{\lambda_i}f_i(w) = \tilde{f}_i(w)\left[1 + \frac{\gamma_i}{\lambda_i}\frac{f_i(w)}{\tilde{f}_i(w)}\right]. \tag{18}$$

Proof. See the Appendix. ■

The interpretation of Proposition 4.2(i) is twofold. First, if migration elasticity is smaller than 1 and equilibrium net inflow of the most skilled individuals is small (i.e., a small migration elasticity implies that country i is more likely to face a net inflow of top-type individuals, and $q_i(\infty)$ being greater than a threshold implies that $\tilde{f}_i(\infty)$ must be smaller than a certain threshold), then using Proposition 4.1 yields:

$$\theta_i(\infty)\uparrow \Longrightarrow \frac{\partial T_i'^*(y_i(\infty))}{\partial q_i(\infty)}\uparrow \Longrightarrow \left|\frac{\partial T_i'^*(y_i(\infty))}{\partial q_i(\infty)}\right|\uparrow$$

and

$$q_i(\infty)\uparrow \Longrightarrow \frac{\partial T_i'^*(y_i(\infty))}{\partial \theta_i(\infty)}\uparrow \Longrightarrow \left|\frac{\partial T_i'^*(y_i(\infty))}{\partial \theta_i(\infty)}\right|\downarrow .$$

As such, if both migration elasticity and equilibrium net inflow of top-type individuals have been small, then the channel of migration magnitude dominates the channel of migration elasticity, in the sense that an increase in migration elasticity leads, ceteris paribus, to a greater positive effect of migration magnitude imposed on the optimal AMTR.

Second, if migration elasticity is larger than 1 and equilibrium net outflow of the most skilled individuals is small (i.e., a large migration elasticity implies that country i is more likely to face a net outflow of top-type individuals, and $q_i(\infty)$ being smaller than a threshold implies that $\tilde{f}_i(\infty)$ must be greater than a certain threshold), then using Proposition 4.1 again yields:

$$\theta_i(\infty)\uparrow \Longrightarrow \frac{\partial T_i'^*(y_i(\infty))}{\partial q_i(\infty)}\downarrow \Longrightarrow \left|\frac{\partial T_i'^*(y_i(\infty))}{\partial q_i(\infty)}\right|\downarrow$$

and

$$q_i(\infty)\uparrow \Longrightarrow \frac{\partial T_i'^*(y_i(\infty))}{\partial \theta_i(\infty)}\downarrow \Longrightarrow \left|\frac{\partial T_i'^*(y_i(\infty))}{\partial \theta_i(\infty)}\right|\uparrow .$$

Hence, if migration elasticity has been large and the equilibrium net outflow of top-type individuals has been small, then the channel of migration magnitude is dominated by the channel of migration elasticity, in the sense that an increase in migration elasticity reduces the positive effect of migration magnitude imposed on the optimal AMTR.

Proposition 4.2(ii) shows that migration elasticity and the degree of inequality act as substitutes in shaping the optimal AMTR, if migration elasticity is below 3. All else being equal, an increase in the degree of inequality will produce equity gains by increasing the tax rate, whereas an increase in migration elasticity will produce efficiency gains by reducing the tax rate. Therefore, these two factors represent opposite forces in shaping the optimal tax rate, leading to a negative joint effect. To further interpret the joint effect, note that

$$\frac{\partial^2 T_i'^*(y_i(\infty))}{\partial \theta_i(\infty)\partial\left(\frac{1}{a_i}\right)} < 0 \implies \frac{\partial\left(\frac{\partial T_i'^*(y_i(\infty))}{\partial \theta_i(\infty)}\right)}{\partial\left(\frac{1}{a_i}\right)} < 0 \text{ for } \frac{\partial T_i'^*(y_i(\infty))}{\partial \theta_i(\infty)} < 0,$$

which implies that the higher the degree of inequality is, the greater the effect of an increase in migration elasticity on reducing the optimal tax rate. As graphically shown in Figure 1, regardless of whether country i faces net labor inflow or outflow of the highest skilled individuals, the higher the degree of inequality (i.e., the smaller the value that a_i takes), the steeper $T_i'^*(y_i(\infty))$ as a decreasing function of migration elasticity is. Therefore, an increase in the degree of inequality strengthens the downward pressure of intensifying migration-induced competition imposed on the optimal tax rate.

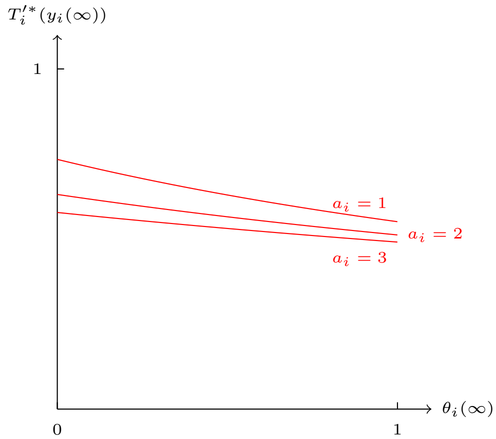

Figure 1: How $T_i'^*(y_i(\infty))$, as a function of $\theta_i(\infty)$, is shaped by inequality
Note: The other relevant parameters in the formula for the optimal AMTR take the following values: $\varepsilon = 0.25$, $\sigma_i = 0.5$, $q_i(\infty) = 0.67$ (net labor inflow) for the left panel, and $q_i(\infty) = 1.55$ (net labor outflow) for the right one.

Similarly, by the Young's Theorem,

$$\frac{\partial^2 T_i'^*(y_i(\infty))}{\partial \theta_i(\infty) \partial \left(\frac{1}{a_i}\right)} < 0 \implies \frac{\partial \left(\frac{\partial T_i'^*(y_i(\infty))}{\partial \left(\frac{1}{a_i}\right)} \right)}{\partial \theta_i(\infty)} < 0 \text{ for } \frac{\partial T_i'^*(y_i(\infty))}{\partial \left(\frac{1}{a_i}\right)} > 0,$$

which means that the larger the migration elasticity is, the smaller the effect of an increase in the degree of inequality on increasing the optimal tax rate. As graphically shown in Figure 2, regardless of whether country i faces net labor inflow or outflow of the highest skilled individuals, the larger the migration elasticity, the flatter $T_i'^*(y_i(\infty))$ as an increasing function of the degree of inequality is. Consequently, intensifying migration-induced competition weakens the upward pressure of worsening inequality imposed on the optimal tax rate.

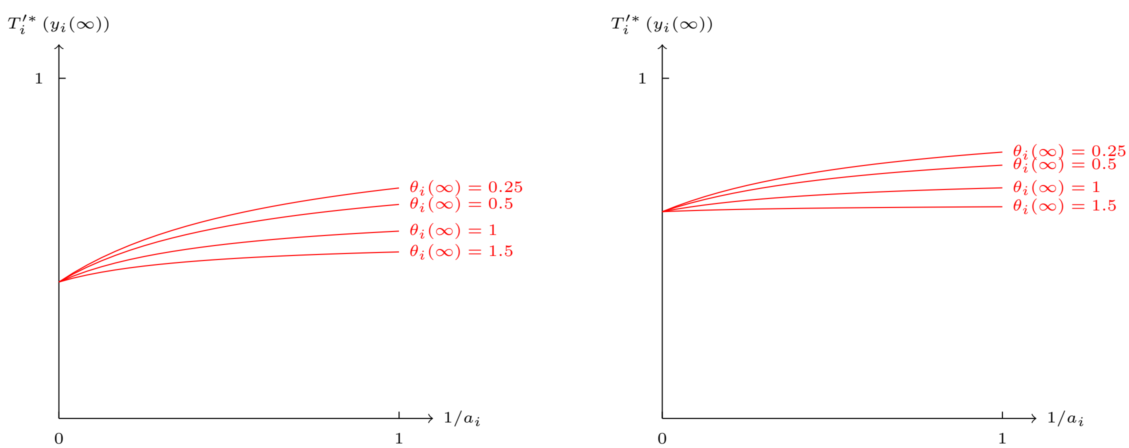

Figure 2: How $T_i'^*(y_i(\infty))$, as a function of $1/a_i$, is shaped by migration
Note: $\varepsilon = 0.25$, $\sigma_i = 0.5$, $a_i \geq 1$, $q_i(\infty) = 0.67$ (net labor inflow) for the left panel, and $q_i(\infty) = 1.55$ (net labor outflow) for the right one.

Proposition 4.2(iii-a) shows that if the intensity of relativity concern is greater than a threshold and migration elasticity is smaller than 1, then migration elasticity and relativity intensity

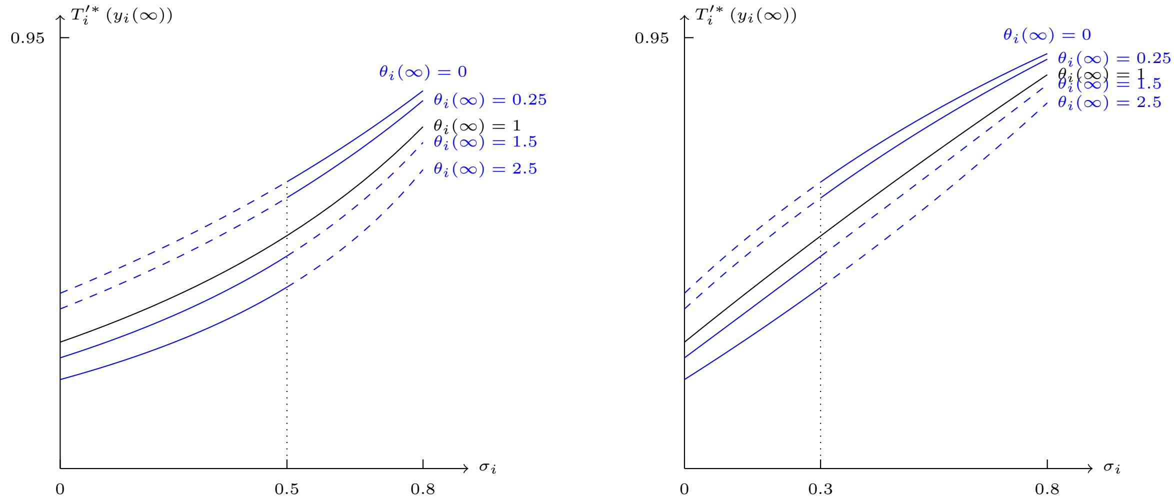

Figure 3: How $T_i'^*(y_i(\infty))$, as a function of σ_i, is shaped by migration
Note: $\varepsilon = 0.25$, $a_i = 2$, $q_i(\infty) = 0.67$ (net labor inflow) for the left panel, and $q_i(\infty) = 1.55$ (net labor outflow) for the right one.

act as complements in shaping the optimal AMTR. In contrast, if relativity intensity is smaller than a threshold and migration elasticity is larger than 1, then migration elasticity and relativity intensity act as substitutes in shaping the optimal AMTR.[25] As graphically shown in Figure 3, in an equilibrium with country i facing a net labor inflow, we have two main observations. (1) For the case in which relativity intensity is below 0.5 and migration elasticity is no smaller than 1, then the larger the migration elasticity is, the *flatter* $T_i'^*(y_i(\infty))$ as an increasing function of relativity intensity is. (2) For the case in which relativity intensity is beyond 0.5 and migration elasticity is no greater than 1, then the larger the migration elasticity is, the *steeper* $T_i'^*(y_i(\infty))$ as an increasing function of relativity intensity is. For an equilibrium with country i facing a net labor outflow of the highest skilled individuals, the corresponding observations can be explained in a similar way.

Considered separately, migration imposes a downward pressure on the tax rate, whereas relativity concern imposes an upward pressure. These two factors tend to act in opposite directions in shaping the optimal tax rate, leading to a negative joint effect when they are considered in tandem. However, Proposition 4.2(iii-a) shows that this is true only when relativity intensity has been low and migration elasticity has been large. If relativity intensity has been high, and dominates the force of migration, these two factors enhance each other in determining the optimal AMTR.

If relativity concern has been strong and migration-induced competition has been mild, then relativity will weaken the downward pressure of migration imposed on the tax rate, while migration will strengthen the upward pressure of relativity imposed on the tax rate. Consequently, relativity seems to dominate migration in terms of signing the joint effect, tending to result in a high tax rate. However, if the relativity concern has been weak and migration-induced competition has been intense, then relativity will strengthen the downward pressure of migration imposed on the tax rate, whereas migration will weaken the upward pressure of relativity imposed on the tax rate. Therefore, migration seems to dominate relativity in terms of signing the joint effect, tending to result in a low tax rate.

[25] The empirical evidence of Kleven, Landais, and Saez (2013), Kleven, Landais, Saez, and Schultz (2014), and Akcigit, Baslandze, and Stantcheva (2016) demonstrates that top earners' migration elasticity with respect to the tax rate can be larger than 1. For a recent review of such empirical evidence, please see Kleven, Landais, Muñoz, and Stantcheva (2020).

Note that $\frac{\gamma_i}{\lambda_i}(\infty)q_i(\infty)$ is strictly increasing in G_{-i} given $\mathrm{d}G_i^*/\mathrm{d}G_{-i} < 0$, and hence

$$\frac{\partial T_i'^*(y_i(\infty))}{\partial G_{-i}} > 0 \text{ for } \mathrm{d}G_i^*/\mathrm{d}G_{-i} < 0.$$

That is, an increase in the provision of public goods in the competing country $-i$ gives rise to a higher tax rate in country i, generating an effect similar to the effect of strengthening intra-country consumption comparisons, so the interpretation of Proposition 4.2(iii-b) is analogous to that of Proposition 4.2(iii-a).

Finally, we interpret the necessary and sufficient condition established in Proposition 4.2(iii-b) such that $\mathrm{d}G_i^*/\mathrm{d}G_{-i} < 0$. Using (15) and (18), we rewrite the rule for optimal public goods provision as

$$\int_{\underline{w}}^{\overline{w}} \mathrm{MRS}_{G_i c_i}(w)Q_i(w)dw = 1 = \mathrm{MRT}_{G_i c_i}, \tag{19}$$

where MRS and MRT denote marginal rate of substitution and marginal rate of transformation, respectively. This rule departs from Samuelson (1954)'s rule because the density weight assigned to skill type w, $Q_i(w)$, differs from the usual one, $f_i(w)$. Thanks to international labor mobility, $\tilde{f}_i(w)$ is generally different from $f_i(w)$; meanwhile, there is a positive shadow price of reference consumption, namely $\gamma_i > 0$, and hence the second component of $Q_i(w)$ is strictly positive. Under quasilinear-in-consumption preferences, (19) can be simplified as

$$\int_{\underline{w}}^{\overline{w}} \mathrm{MU}_{G_i}(w)Q_i(w)dw = 1, \tag{20}$$

where MU_{G_i} denotes the marginal utility of public goods consumption.

We now exploit (20) to interpret the sign of $\mathrm{d}G_i^*/\mathrm{d}G_{-i}$, which characterizes country i's strategic reaction of optimal public goods provision to country $-i$'s unilateral deviation on its public goods provision. $\mathrm{d}G_i^*/\mathrm{d}G_{-i} < 0$ means that an increase in country $-i$'s public goods provision will, ceteris paribus, induce a reduction in country i's optimal public goods provision. This statement is true iff the elasticity of Q_i with respect to G_i is smaller than the absolute value of the elasticity of MU_{G_i} with respect to G_i. For a unit increase in G_i, while the former elasticity measures a positive effect on the *extensive margin*, the latter elasticity measures a negative effect on the *intensive margin*.[26] Therefore, country i's optimal response to an increase in country $-i$'s public goods provision is providing a smaller amount of public goods, iff $\epsilon_{Q_i(w),G_i} < -\epsilon_{v^p_{G_i},G_i}$ — namely, for a unit decrease in G_i, the resulting negative effect on the extensive margin, measured by $\epsilon_{Q_i(w),G_i} > 0$, is outweighed by the corresponding positive effect on the intensive margin, measured by $-\epsilon_{v^p_{G_i},G_i} > 0$.

5 Alternative Reference Consumption Constraints

In the previous section, we assume that all skill types face the same reference consumption, and their relativity concern is defined with respect to the average (mean) consumption in the population. This specification is widely adopted in the related literature; hence, it is reasonable to serve as a benchmark.[27] In this section, we consider two variations of this specification. The first is type-dependent, varying with skill type, and is motivated by the fact that individuals often compare themselves with those earning higher incomes. The second is specifying the consumption (after-tax income) of median-skill individuals as the common reference consumption for all skill types, which is motivated by the observation that people may choose to target the

[26] If $v^p(G_i, G_{-i}) = (G_i/G_{-i})^\alpha$ for $\alpha \in (0,1)$, then $-\epsilon_{v^p_{G_i},G_i} = 1-\alpha$.

[27] See, e.g., Oswald (1983), Kanbur and Tuomala (2013), and Aronsson and Johansson-Stenman (2015).

median income rather than the mean income because the former measure provides an accurate and unbiased representation of the average income in a society.

5.1 Upward consumption comparison

An emulation-driven individual with type w becomes worse off when the average consumption of the higher skill levels is larger, which suggests the following constraint on the reference consumption:

$$\mu_i(w) \;\geq\; \frac{\int_w^{\overline{w}} c_i(t) f_i(t) dt}{\int_{\underline{w}}^{\overline{w}} f_i(t) dt} \equiv \int_w^{\overline{w}} c_i(t) f_i(t) dt \quad \text{for } w \in [\underline{w}, \overline{w}] \text{ and } i \in \{A, B\}. \tag{21}$$

We will call the term on the right hand side of inequality (21) the quasi-average consumption of the higher skill levels, for any given skill type w under consideration. In particular, for the top type $\overline{w}$, constraint (21) degenerates to $\mu_i(\overline{w}) \geq 0$, and hence is trivially binding at an optimum, whereas for the bottom type $\underline{w}$, it is the same as the previous relativity scenario featuring a comparison with the average consumption in the entire population.

Replacing constraint (3) with (21) in Problem (11) and then solving this problem, we arrive at:

Lemma 5.1 *Given the public policy choice of country $-i$, we have the following statements for country i.*

(i) The second-best MTRs, $T_i'^{\star}(y_i(\cdot))$, verify:

$$\frac{T_i'(y_i(w))}{1 - T_i'(y_i(w))} \;=\; \overbrace{\frac{\Gamma_i(w)}{\lambda_i} \frac{f_i(w)}{\tilde{f}_i(w)}}^{\textit{Pigouvian-type tax}} \;+\; \overbrace{\mathcal{A}_i(w) \mathcal{B}_i(w) \mathcal{C}_i(w)}^{\textit{Mirrleesian-type tax}} \tag{22}$$

where: $\mathcal{A}_i(w) \equiv 1 + [l_i(w) h''(l_i(w))/h'(l_i(w))]$, $\mathcal{B}_i(w) \equiv \left[\tilde{F}_i(\overline{w}) - \tilde{F}_i(w)\right] / w \tilde{f}_i(w)$,

$$\mathcal{C}_i(w) \equiv \frac{v_c(w) \int_w^{\overline{w}} \left\{ \frac{1}{v_c(t)} \left[1 + \frac{\Gamma_i(t)}{\lambda_i} \frac{f_i(t)}{\tilde{f}_i(t)}\right] - T_i(y_i(t)) \eta_i(t) \right\} \tilde{f}_i(t) dt}{\tilde{F}_i(\overline{w}) - \tilde{F}_i(w)}, \tag{23}$$

$v_c(w) \equiv v_c(c_i(w), G_i; \mu_i, G_{-i})$, $\tilde{F}_i(w) \equiv \int_{\underline{w}}^{w} \tilde{f}_i(t) dt$ *denotes the ex post skill distribution,* $\Gamma_i(w) \equiv \int_{\underline{w}}^{w} \gamma_i(t) dt$, *and*

$$\frac{\gamma_i(w)}{\lambda_i} = -\frac{v_\mu(c_i(w), G_i; \mu_i, G_{-i})}{v_c(c_i(w), G_i; \mu_i, G_{-i})} \tilde{f}_i(w) \left[1 + \frac{\Gamma_i(w)}{\lambda_i} \frac{f_i(w)}{\tilde{f}_i(w)}\right] \quad \forall w, \tag{24}$$

with Lagrange multipliers $\gamma_i(w) > 0$ and $\lambda_i > 0$ associated with the type-dependent reference consumption constraint and the government budget constraint, respectively.

(ii) Optimal public goods provision, $G_i^{\star}$, satisfies:

$$\int_{\underline{w}}^{\overline{w}} \frac{v_{G_i}(c_i(w), G_i; \mu_i, G_{-i})}{v_c(c_i(w), G_i; \mu_i, G_{-i})} \tilde{f}_i(w) \left[1 + \frac{\Gamma_i(w)}{\lambda_i} \frac{f_i(w)}{\tilde{f}_i(w)}\right] dw = 1, \tag{25}$$

with $\Gamma_i(w)/\lambda_i \equiv \int_{\underline{w}}^{w} \frac{\gamma_i(t)}{\lambda_i} dt$ *determined by (24).*

Proof. See the Appendix. ■

Comparing Lemma 5.1 with Lemma 4.1, the differences in the optimal tax formula and in the rule for optimal public goods provision are precisely due to the distinction in the specification of reference consumption.

Now, we are ready to derive the optimal AMTR:

Lemma 5.2 *Suppose Assumptions 4.1 and 4.2 hold. Then, the formula for the optimal AMTR under upward consumption comparison is given by*

$$\frac{T_i'^{\star}(y_i(\infty))}{1-T_i'^{\star}(y_i(\infty))} = \frac{\left(1+\frac{1+\varepsilon}{a_i}\right)\frac{\Gamma_i(\infty)}{\lambda_i}\cdot q_i(\infty)+\frac{1+\varepsilon}{a_i}}{1+\frac{1+\varepsilon}{a_i}\cdot\theta_i(\infty)}, \tag{26}$$

where $\frac{\Gamma_i(\infty)}{\lambda_i} \equiv \lim_{w\uparrow\infty}\frac{\Gamma_i(w)}{\lambda_i}$ *and* $\theta_i(\infty) \equiv \lim_{w\uparrow\infty}\theta_i(w)$. *Moreover, (26) can be equivalently expressed as*

$$T_i'^{\star}(y_i(\infty)) = \frac{\left(1+\frac{1+\varepsilon}{a_i}\right)\frac{\Gamma_i(\infty)}{\lambda_i}\cdot q_i(\infty)+\frac{1+\varepsilon}{a_i}}{\left(1+\frac{1+\varepsilon}{a_i}\right)\left[1+\frac{\Gamma_i(\infty)}{\lambda_i}\cdot q_i(\infty)\right]+\frac{1+\varepsilon}{a_i}\cdot\theta_i(\infty)}. \tag{27}$$

Proof. See the Appendix. ■

Comparing tax formula (27) with (17), the difference and similarity of these two optimal AMTRs are straightforward. Specifically, $\frac{\gamma_i}{\lambda_i}(\infty)$ in (17) is replaced by $\Gamma_i(\infty)/\lambda_i$ in (27), which is due to the distinction in reference consumption constraint. Given the functional similarity between tax formulas (27) and (17), some of the comparative statics stated in Proposition 4.1 remain valid in the present case, and hence we give the following proposition without providing a formal proof.

Proposition 5.1 *Suppose Assumptions 4.1 and 4.2 hold. Then, we have comparative statics:*

$$\frac{\partial T_i'^{\star}(y_i(\infty))}{\partial\theta_i(\infty)} < 0 \quad \text{and} \quad \frac{\partial T_i'^{\star}(y_i(\infty))}{\partial q_i(\infty)} > 0.$$

Analogous to Proposition 4.2, we can use tax formula (27) to establish the following proposition, which characterizes the joint effects of relevant factors on the optimal AMTR.

Proposition 5.2 *Suppose Assumptions 4.1 and 4.2 hold. Then, we have the following predictions.*

(i) Concerning the joint effect of migration elasticity and migration magnitude on the optimal AMTR, we have:

$$\frac{\partial^2 T_i'^{\star}(y_i(\infty))}{\partial\theta_i(\infty)\partial q_i(\infty)} > 0 \quad \text{if } \frac{\Gamma_i(\infty)}{\lambda_i}q_i(\infty) > \frac{a_i}{1+\varepsilon+a_i} \text{ and } \theta_i(\infty) \le 1,$$

$$\frac{\partial^2 T_i'^{\star}(y_i(\infty))}{\partial\theta_i(\infty)\partial q_i(\infty)} < 0 \quad \text{if } \frac{\Gamma_i(\infty)}{\lambda_i}q_i(\infty) < \frac{a_i}{1+\varepsilon+a_i} \text{ and } \theta_i(\infty) \ge 1.$$

(ii) If $-\epsilon_{\tilde{f}_i(w),\sigma_i} \le 1$ *holds for almost all* $w \in [\underline{w}, \overline{w}]$, *then the joint effect of migration elasticity and the intensity of relativity concern is characterized as follows:*

$$\frac{\partial^2 T_i'^{\star}(y_i(\infty))}{\partial\theta_i(\infty)\partial\sigma_i} > 0 \quad \text{for } \frac{\Gamma_i(\infty)}{\lambda_i}q_i(\infty) > \frac{a_i}{1+\varepsilon+a_i} \text{ and } \theta_i(\infty) \le 1,$$

$$\frac{\partial^2 T_i'^{\star}(y_i(\infty))}{\partial\theta_i(\infty)\partial\sigma_i} < 0 \quad \text{for } \frac{\Gamma_i(\infty)}{\lambda_i}q_i(\infty) < \frac{a_i}{1+\varepsilon+a_i} \text{ and } \theta_i(\infty) \ge 1,$$

where $\frac{\Gamma_i(\infty)}{\lambda_i}q_i(\infty)$ *is strictly increasing in* σ_i, *and* $\epsilon_{\tilde{f}_i(w),\sigma_i}$ *represents the elasticity of the ex post skill density with respect to the relativity intensity* σ_i.

(iii) We have

$$\frac{\mathrm{d}G_i^{\star}}{\mathrm{d}G_{-i}} < 0 \text{ iff } \epsilon_{S_i,G_i} < -\epsilon_{v^p_{G_i},G_i},$$

where $\epsilon_{S_i,G_i} > 0$ *and* $\epsilon_{v^p_{G_i},G_i} < 0$ *denote the elasticities of* S_i *and* $v^p_{G_i}$ *with respect to* G_i*, respectively, with*

$$S_i \equiv \int_{\underline{w}}^{\overline{w}} \tilde{f}_i(w) \left[1 + \frac{\Gamma_i(w)}{\lambda_i} \frac{f_i(w)}{\tilde{f}_i(w)}\right] dw. \tag{28}$$

Then, with $\mathrm{d}G_i^\star/\mathrm{d}G_{-i} < 0$ *the joint effect of migration elasticity and the competing country's public goods provision on the optimal AMTR is characterized as follows:*

(iii-a) *If* $\epsilon_{\Gamma_i(\infty)/\lambda_i,G_{-i}} > \epsilon_{\tilde{f}_i(\infty),G_{-i}}$*, then*

$$\frac{\partial^2 T_i'^\star(y_i(\infty))}{\partial\theta_i(\infty)\partial G_{-i}} > 0 \quad \textit{for } \frac{\Gamma_i(\infty)}{\lambda_i} q_i(\infty) > \frac{a_i}{1+\varepsilon+a_i} \textit{ and } \theta_i(\infty) \leq 1,$$

$$\frac{\partial^2 T_i'^\star(y_i(\infty))}{\partial\theta_i(\infty)\partial G_{-i}} < 0 \quad \textit{for } \frac{\Gamma_i(\infty)}{\lambda_i} q_i(\infty) < \frac{a_i}{1+\varepsilon+a_i} \textit{ and } \theta_i(\infty) \geq 1,$$

where $\frac{\Gamma_i(\infty)}{\lambda_i} q_i(\infty)$ *is strictly increasing in* G_{-i}*, and* $\epsilon_{\Gamma_i(\infty)/\lambda_i,G_{-i}}$ *and* $\epsilon_{\tilde{f}_i(\infty),G_{-i}}$ *represent elasticities of* $\Gamma_i(\infty)/\lambda_i$ *and* $\tilde{f}_i(\infty)$ *with respect to* G_{-i}*, respectively.*

(iii-b) *If* $\epsilon_{\Gamma_i(\infty)/\lambda_i,G_{-i}} < \epsilon_{\tilde{f}_i(\infty),G_{-i}}$*, then*

$$\frac{\partial^2 T_i'^\star(y_i(\infty))}{\partial\theta_i(\infty)\partial G_{-i}} < 0 \quad \textit{for } \frac{\Gamma_i(\infty)}{\lambda_i} q_i(\infty) > \frac{a_i}{1+\varepsilon+a_i} \textit{ and } \theta_i(\infty) \leq 1,$$

$$\frac{\partial^2 T_i'^\star(y_i(\infty))}{\partial\theta_i(\infty)\partial G_{-i}} > 0 \quad \textit{for } \frac{\Gamma_i(\infty)}{\lambda_i} q_i(\infty) < \frac{a_i}{1+\varepsilon+a_i} \textit{ and } \theta_i(\infty) \geq 1,$$

where $\frac{\Gamma_i(\infty)}{\lambda_i} q_i(\infty)$ *is strictly decreasing in* G_{-i}*.*

Proof. See the Appendix. ■

Given the functional similarity between tax formulas (17) and (27), the interpretation of the joint effects of relevant factors on the optimal AMTR is basically the same under both circumstances. The minor departure lies in the (sufficient) conditions required. For instance, concerning the joint impact of migration elasticity and relativity intensity on the optimal AMTR, Proposition 5.2 requires that the negative effect of stronger relativity concerns on the extensive margin, measured by $-\epsilon_{\tilde{f}_i(w),\sigma_i} > 0$, be smaller than a constant threshold, whereas Proposition 4.2 relies on a different set of assumptions. Roughly, the key insights conveyed by Proposition 4.2 continue to be true for Proposition 5.2.

Moreover, Proposition 5.2(iii) provides the necessary and sufficient condition — for an increase in public goods provision, the resulting negative effect on the intensive margin, measured by $-\epsilon_{v^p_{G_i},G_i} > 0$, outweighs the resulting positive effect on the extensive margin, measured by $\epsilon_{S_i,G_i} > 0$ — under which the optimal amount of public goods provision should decrease, in response to an increase in the competing country's public goods provision.

5.2 Comparison with the median-skill type

If median-skill individuals represent the middle class of the society, then it is reasonable to consider the following reference consumption constraint:

$$\mu_i \;\geq\; c_i(w_m) \quad \text{for } i \in \{A, B\}, \tag{29}$$

in which $c_i(w_m)$ denotes the consumption level of median-skill individuals.

Replacing constraint (3) with (29) in Problem (11) and then solving this problem, we have:

Lemma 5.3 *Given the public policy choice of country $-i$, we have the following statements for country i.*

(i) The second-best MTRs, $\hat{T}_i'(y_i(\cdot))$, verify:

$$\frac{T_i'(y_i(w))}{1-T_i'(y_i(w))} = \overbrace{\frac{\gamma_i}{\lambda_i}\frac{f_i(w)}{\tilde{f}_i(w)}\cdot \mathbb{I}_{w=w_m}}^{\text{Pigouvian-type tax}} + \overbrace{\mathcal{A}_i(w)\mathcal{B}_i(w)\mathcal{C}_i(w)}^{\text{Mirrleesian-type tax}} \tag{30}$$

where: $\mathcal{A}_i(w) \equiv 1 + [l_i(w)h''(l_i(w))/h'(l_i(w))]$, $\mathcal{B}_i(w) \equiv \left[\tilde{F}_i(\overline{w}) - \tilde{F}_i(w)\right]/w\tilde{f}_i(w)$,

$$\mathcal{C}_i(w) \equiv \frac{v_c(w)\int_w^{\overline{w}}\left\{\frac{1}{v_c(t)}\left[1+\frac{\gamma_i}{\lambda_i}\frac{f_i(t)}{\tilde{f}_i(t)}\mathbb{I}_{t=w_m}\right] - T_i(y_i(t))\eta_i(t)\right\}\tilde{f}_i(t)dt}{\tilde{F}_i(\overline{w}) - \tilde{F}_i(w)}, \tag{31}$$

$v_c(w) \equiv v_c(c_i(w), G_i; \mu_i, G_{-i})$, $\tilde{F}_i(w) \equiv \int_{\underline{w}}^{w}\tilde{f}_i(t)dt$ *denotes the ex post skill distribution,* $\mathbb{I}_{t=w_m}$ *is equal to 1 if* $t = w_m$ *and to 0 otherwise, and*

$$\frac{\gamma_i}{\lambda_i} = -\frac{\int_{\underline{w}}^{\overline{w}}\frac{v_\mu(c_i(w),G_i;\mu_i,G_{-i})}{v_c(c_i(w),G_i;\mu_i,G_{-i})}\tilde{f}_i(w)dw}{\left[1+\frac{v_\mu(c_i(w_m),G_i;\mu_i,G_{-i})}{v_c(c_i(w_m),G_i;\mu_i,G_{-i})}\right]f_i(w_m)}, \tag{32}$$

with Lagrange multipliers $\gamma_i > 0$ *and* $\lambda_i > 0$ *associated with the reference consumption constraint and the government budget constraint, respectively.*

(ii) Optimal public goods provision satisfies:

$$\int_{\underline{w}}^{\overline{w}}\frac{v_{G_i}(c_i(w),G_i;\mu_i,G_{-i})}{v_c(c_i(w),G_i;\mu_i,G_{-i})}\tilde{f}_i(w)dw + \frac{\gamma_i}{\lambda_i}\cdot\frac{v_{G_i}(c_i(w_m),G_i;\mu_i,G_{-i})}{v_c(c_i(w_m),G_i;\mu_i,G_{-i})}f_i(w_m) = 1,$$

with γ_i/λ_i *given by (32).*

Proof. See the Appendix. ■

A significant difference arises from comparing tax formula (30) with tax formula (12), revealing that only median-skill individuals face a positive Pigouvian-type tax when reference consumption is the median-skill's consumption, whereas all skill types face a positive Pigouvian-type tax when reference consumption is the average consumption in the population. The rationale is logical, given that Pigouvian tax is used to correct the negative positional externality generated by consumption comparisons. Under reference consumption constraint (29), everyone targets the median-skill type; hence, the negative positional externality is produced by median-type individuals. Therefore, levying positive Pigouvian taxes on them would drive down their consumption to a socially desirable level and internalize the negative externality to restore (constrained) efficiency. The same logic applies to the optimal Pigouvian tax that appears in Lemmas 4.1 and 5.1.

Using (30)-(32) yields the following:

Proposition 5.3 *Suppose Assumptions 4.1 and 4.2 hold. Then, the optimal AMTR under comparison with the median-skill type amounts to*

$$\hat{T}_i'(y_i(\infty)) = \frac{\frac{1+\varepsilon}{a_i}}{1+\frac{1+\varepsilon}{a_i}+\frac{1+\varepsilon}{a_i}\cdot\theta_i(\infty)}, \tag{33}$$

where $\theta_i(\infty) \equiv \lim_{w\uparrow\infty}\theta_i(w)$. *Moreover, using (33) yields*

$$\frac{\partial \hat{T}_i'(y_i(\infty))}{\partial\left(\frac{1+\varepsilon}{a_i}\right)} > 0, \quad \frac{\partial \hat{T}_i'(y_i(\infty))}{\partial \theta_i(\infty)} < 0, \quad \text{and} \quad \frac{\partial^2 \hat{T}_i'(y_i(\infty))}{\partial\left(\frac{1}{a_i}\right)\partial\theta_i(\infty)} < 0.$$

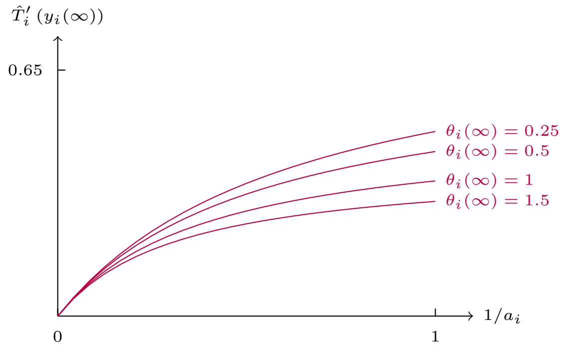

Figure 4: How $\hat{T}_i'(y_i(\infty))$, as a function of $1/a_i$, is shaped by migration
Note: the elasticity of labor supply takes $\varepsilon = 0.25$ and the Pareto index takes $a_i \geq 1$.

Proof. See the Appendix. ■

First, it is intuitive that the optimal tax rate decreases in migration elasticity and increases in the degree of inequality. Second, as previously, the joint effect of these two factors on the optimal tax rate can be interpreted as follows. On the one hand, using Young's Theorem yields:

$$\frac{\partial \left(\frac{\partial \hat{T}_i'(y_i(\infty))}{\partial \left(\frac{1}{a_i}\right)} \right)}{\partial \theta_i(\infty)} = \frac{\partial^2 \hat{T}_i'(y_i(\infty))}{\partial \theta_i(\infty) \partial \left(\frac{1}{a_i}\right)} < 0.$$

This result implies that the larger the migration elasticity is, the smaller the effect of an increase in the degree of inequality on increasing the optimal tax rate. As graphically shown in Figure 4, the larger the migration elasticity, the flatter $\hat{T}_i'(y_i(\infty))$ as an increasing function of the degree of inequality is. As such, intensifying migration-induced tax competition will, ceteris paribus, weaken the upward pressure of worsening inequality on the optimal tax rate.

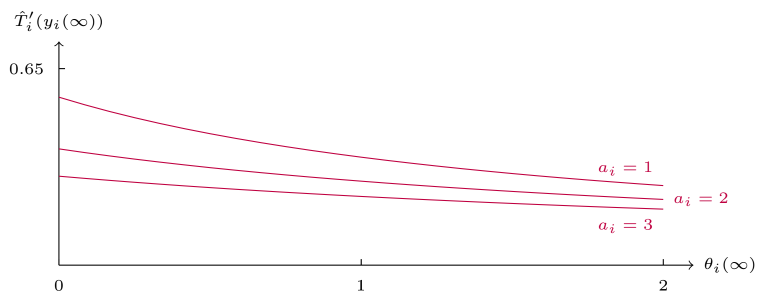

Figure 5: How $\hat{T}_i'(y_i(\infty))$, as a function of $\theta_i(\infty)$, is shaped by inequality
Note: the elasticity of labor supply takes $\varepsilon = 0.25$.

On the other hand, we arrive at

$$\frac{\partial \left(\frac{\partial \hat{T}_i'(y_i(\infty))}{\partial \theta_i(\infty)} \right)}{\partial \left(\frac{1}{a_i}\right)} = \frac{\partial^2 \hat{T}_i'(y_i(\infty))}{\partial \theta_i(\infty) \partial \left(\frac{1}{a_i}\right)} < 0.$$

This result implies that the higher the degree of inequality is, the greater the effect of an increase in migration elasticity on reducing the optimal tax rate. As graphically shown in Figure 5, the higher the degree of inequality (i.e., the smaller the value that a_i takes), the steeper $\hat{T}_i'(y_i(\infty))$ as

Table 1: Parameter values

	Value	Description	Source/Target
$\theta_i(\infty)$	0.25	migration elasticity	Piketty and Saez (2013)
ε	(0.1,0.4)	labor-supply elasticity	Saez, Slemrod, and Giertz (2012)
a_i	2	Pareto index	Diamond and Saez (2011)
σ_i	0.5	relativity intensity	Clark, Frijters, and Shields (2008)

a decreasing function of migration elasticity is. Consequently, worsening inequality will, ceteris paribus, strengthen the downward pressure of intensifying tax competition placed on the optimal tax rate.

6 A Comparative Analysis with Numerical Examples

In this section, we conduct a comparative analysis of the optimal AMTRs given by tax formulas (17), (27), and (33) under alternative reference consumption constraints using numerical examples. The numerical illustration helps us to quantitatively determine the difference in these optimal AMTRs from various specifications of the reference consumption constraint, and demonstrate how the difference varies with the degree of inequality, elasticity of labor supply, and intensity of relativity concerns.

We consider three elasticity scenarios regarding labor supply. $\varepsilon = 0.25$ is realistic (Chetty 2012), $\varepsilon = 0.33$ is a midrange estimate (Chetty 2012, and Saez, Slemrod, and Giertz 2012), and $\varepsilon = 0.5$ is slightly larger than the average estimate for an experimental deviation. We consider two inequality scenarios: $a_i = 2$ is based on the 2005 US empirical income distribution (see Diamond and Saez 2011), while $a_i = 3$ represents an experimental deviation toward a more equal economy.

We consider two relativity scenarios, in which $\sigma_i = 0.5$ is consistent with the empirical finding of Alpizar, Carlsson, and Johansson-Stenman (2005), given the preferences specified in Assumption 4.1, whereas $\sigma_i = 0.25$ represents an experimental deviation enabling us to determine how large the difference in optimal AMTR can be made by the reduction of relativity intensity, other things being equal. Following Piketty and Saez (2013), we set migration elasticity to be 0.25. We summarize all realistic estimates of relevant parameters in Table 1.

Tables 2-3 report the numerical values of optimal AMTRs calculated using the tax formulas given by Lemmas 4.2 and 5.2, and Proposition 5.3. We use black, red, and brown numbers to denote the optimal AMTRs under comparison with the average consumption, upward comparison, and comparison with the median-skill type, respectively. In addition, in view of Assumption 4.2 we see that if $q_i(\infty) = \lim_{w\uparrow\infty} f_i(w)/\tilde{f}_i(w) < 1$, then country i faces a net inflow of top-type individuals under tax competition; otherwise, it faces a net outflow of top-type individuals when $q_i(\infty) > 1$. We will calculate the optimal AMTRs in both cases. And we consider four scenarios for the value that $q_i(\infty)$ takes, which can be interpreted in the following manner.

Imposing symmetry of the underlying environments of the two countries and using (5), we can show that $q_i(\infty) = 0.95 \Leftrightarrow Pr(m < U_i(\infty) - U_{-i}(\infty)|\overline{w}) \approx 5.3\%$ and $q_i(\infty) = 0.67 \Leftrightarrow Pr(m < U_i(\infty) - U_{-i}(\infty)|\overline{w}) \approx 49.3\%$, namely, the conditional out-migration probabilities of top-type individuals in country $-i$ are approximately 5.3% and 49.3%, respectively, under these two values of $q_i(\infty)$. Likewise, we have $q_i(\infty) = 1.05 \Leftrightarrow Pr(m < U_{-i}(\infty) - U_i(\infty)|\overline{w}) \approx 4.8\%$ and $q_i(\infty) = 1.55 \Leftrightarrow Pr(m < U_{-i}(\infty) - U_i(\infty)|\overline{w}) \approx 35.5\%$, namely, the conditional out-migration probabilities of top-type individuals in country i are approximately 4.8% and 35.5%, respectively, under these two values of $q_i(\infty)$. The knife-edge case with $q_i(\infty) = 1$ will serve as a benchmark for the equilibrium size of cross-country labor flows.

To calculate the optimal AMTRs with formula (27), we put $\Gamma_i(\infty)/\lambda_i = 0.15$ for $\sigma_i = 0.25$, and $\Gamma_i(\infty)/\lambda_i = 0.50$ for $\sigma_i = 0.50$. It follows from (24) that $\Gamma_i(\infty)/\lambda_i$ cannot be explicitly solved for under Assumptions 4.1-4.2, but we can show that $\Gamma_i(\infty)/\lambda_i < \frac{\gamma_i}{\lambda_i}(\infty) = \sigma_i/(1-\sigma_i)$ must hold true. Note that $\sigma_i/(1-\sigma_i) = 0.33$ for $\sigma_i = 0.25$ and $\sigma_i/(1-\sigma_i) = 1$ for $\sigma_i = 0.5$, we thus set $\Gamma_i(\infty)/\lambda_i$ to be almost the half of the corresponding value that $\sigma_i/(1-\sigma_i)$ takes to make significant the quantitative differences between these two kinds of optimal AMTR. This is reasonable because one can easily show that the quantitative differences decrease as $\Gamma_i(\infty)/\lambda_i$ approaches $\frac{\gamma_i}{\lambda_i}(\infty) = \sigma_i/(1-\sigma_i)$ from below.

Table 2: AMTRs (%) under alternative sizes for net *inflow* of the most skilled individuals

	$\varepsilon = 0.25$	$\varepsilon = 0.25$	$\varepsilon = 0.33$	$\varepsilon = 0.33$	$\varepsilon = 0.5$	$\varepsilon = 0.5$
	$a_i = 2$	$a_i = 3$	$a_i = 2$	$a_i = 3$	$a_i = 2$	$a_i = 3$
benchmark $q_i(\infty) = 1$:						
$\sigma_i = 0.25$	50.1, 42.9, 35.1	44.5, 36.3, 27.4	51.0, 44.0, 36.3	45.3, 37.2, 28.5	52.8, 46.0, 38.7	46.9, 39.2, 30.8
$\sigma_i = 0.50$	66.1, 55.4, 35.1	62.4, 50.5, 27.4	66.6, 56.2, 36.3	62.9, 51.2, 28.5	67.8, 57.8, 38.7	64.0, 52.6, 30.8
tax increase (%) as $\sigma_i \uparrow$	31.9, 29.1, 0	40.2, 39.1, 0	30.6, 27.7, 0	38.9, 37.6, 0	28.4, 25.7, 0	36.5, 34.2, 0
$q_i(\infty) = 0.95$:						
$\sigma_i = 0.25$	49.4, 42.4, 35.1	43.7, 35.8, 27.4	50.3, 43.5, 36.3	44.5, 40.0, 28.5	52.1, 45.6, 38.7	46.2, 38.7, 30.8
$\sigma_i = 0.50$	65.2, 54.9, 35.1	61.5, 49.8, 27.4	65.8, 55.7, 36.3	62.0, 50.6, 28.5	67.0, 57.2, 38.7	63.1, 52.0, 30.8
tax increase (%) as $\sigma_i \uparrow$	32.0, 29.5, 0	40.7, 39.1, 0	30.8, 28.0, 0	39.3, 26.5, 0	28.6, 25.4, 0	36.6, 34.4, 0
$q_i(\infty) = 0.67$:						
$\sigma_i = 0.25$	45.9, 40.5, 35.1	39.7, 33.6, 27.4	46.9, 41.6, 36.3	40.7, 34.6, 28.5	48.9, 43.8, 38.7	42.5, 36.6, 30.8
$\sigma_i = 0.50$	59.7, 50.5, 35.1	55.3, 44.9, 27.4	60.4, 51.4, 36.3	55.9, 45.7, 28.5	61.8, 53.1, 38.7	57.2, 47.3, 30.8
tax increase (%) as $\sigma_i \uparrow$	30.1, 24.7, 0	39.3, 33.6, 0	28.8, 23.6, 0	37.3, 32.1, 0	26.4, 21.2, 0	34.6, 29.2, 0
tax cut (%) as 0.95 $\downarrow$ 0.67:						
$\sigma_i = 0.25$	7.1, 4.5, 0	9.2, 6.1, 0	6.8, 4.4, 0	8.5, 13.5, 0	6.1, 3.9, 0	8.0, 5.4, 0
$\sigma_i = 0.50$	8.4, 8.0, 0	10.1, 9.8, 0	8.2, 7.7, 0	9.8, 9.7, 0	7.8, 7.2, 0	9.4, 9.0, 0

Table 3: AMTRs (%) under alternative sizes for net *outflow* of the most skilled individuals

	$\varepsilon = 0.25$	$\varepsilon = 0.25$	$\varepsilon = 0.33$	$\varepsilon = 0.33$	$\varepsilon = 0.5$	$\varepsilon = 0.5$
	$a_i = 2$	$a_i = 3$	$a_i = 2$	$a_i = 3$	$a_i = 2$	$a_i = 3$
benchmark $q_i(\infty) = 1$:						
$\sigma_i = 0.25$	50.1, 42.9, 35.1	44.5, 36.3, 27.4	51.0, 44.0, 36.3	45.3, 37.2, 28.5	52.8, 46.0, 38.7	46.9, 39.2, 30.8
$\sigma_i = 0.50$	66.1, 55.4, 35.1	62.4, 50.5, 27.4	66.6, 56.2, 36.3	62.9, 51.2, 28.5	67.8, 57.8, 38.7	64.0, 52.6, 30.8
tax increase (%) as $\sigma_i \uparrow$	31.9, 29.1, 0	40.2, 39.1, 0	30.6, 27.7, 0	38.9, 37.6, 0	28.4, 25.7, 0	36.5, 34.2, 0
$q_i(\infty) = 1.05$:						
$\sigma_i = 0.25$	50.8, 43.4, 35.1	45.2, 36.8, 27.4	51.7, 44.4, 36.3	46.1, 37.8, 28.5	53.4, 46.4, 38.7	47.7, 39.7, 30.8
$\sigma_i = 0.50$	66.8, 56.2, 35.1	63.3, 51.4, 27.4	67.4, 57.0, 36.3	63.8, 52.1, 28.5	68.5, 58.6, 38.7	64.8, 53.5, 30.8
tax increase (%) as $\sigma_i \uparrow$	31.5, 29.5, 0	40.0, 39.7, 0	30.4, 28.4, 0	38.4, 37.8, 0	28.3, 26.3, 0	35.8, 34.8, 0
$q_i(\infty) = 1.55$:						
$\sigma_i = 0.25$	55.7, 46.3, 35.1	50.8, 40.2, 27.4	56.5, 47.3, 36.3	51.5, 41.1, 28.5	58.0, 49.3, 38.7	52.9, 42.9, 30.8
$\sigma_i = 0.50$	73.1, 62.1, 35.1	70.3, 57.9, 27.4	73.6, 62.7, 36.3	70.7, 58.6, 28.5	74.5, 64.0, 38.7	71.5, 59.7, 30.8
tax increase (%) as $\sigma_i \uparrow$	31.2, 34.1, 0	38.4, 44.0, 0	30.3, 32.6, 0	37.3, 42.6, 0	28.4, 29.8, 0	35.2, 39.2, 0
tax increase (%) as 1.05 $\uparrow$ 1.55:						
$\sigma_i = 0.25$	9.6, 6.7, 0	12.4, 9.2, 0	9.3, 6.5, 0	11.7, 8.7, 0	8.6, 6.3, 0	10.9, 8.1, 0
$\sigma_i = 0.50$	9.4, 10.5, 0	11.1, 12.6, 0	9.2, 10.0, 0	10.8, 12.5, 0	8.8, 9.2, 0	10.3, 11.6, 0

Comparing Tables 2-3, we obtain four main findings from the perspective of country i, and the analysis for the competing country can be conducted by symmetry. First, for each combination of the parameter values that we consider, the optimal AMTR is highest when people's relativity concerns target the average consumption in the population, lowest when targeting the median-skill's consumption, and in between when making an upward comparison. This finding is consistent with our intuition that the optimal corrective tax should increase in the population size that generates a negative positional externality.

Second, regardless of whether people target the average or make an upward comparison, the optimal tax rates in an equilibrium with net labor inflow are smaller than the optimal tax rates in an equilibrium with net labor outflow, which verifies the assumption that international migration of the most skilled individuals generally intensifies tax competition; hence, a country should impose a lower tax rate to attract more migrants of such skill type. Under a net labor inflow, optimal tax rates decrease as the size of labor inflow increases, and the magnitude of the tax cut seems to be larger when people target the average level than when they make an upward comparison. Under a net labor outflow, optimal tax rates increase as the size of labor outflow increases, and if the intensity of relativity is low, the magnitude of tax increase tends to be larger when people target the average level than when they make an upward comparison. Otherwise, if the intensity of relativity is high, the magnitude of the tax increase may be smaller when people target the average level than when they make an upward comparison.

Specifically, for all the combinations under a net labor inflow, as the asymptotic ratio of ex-ante skill density to ex post skill density ($\lim_{w\uparrow\infty} f_i(w)/\tilde{f}_i(w)$) decreases from 0.95 to 0.67, registering a 29.5% reduction in skill density, the *tax cut* is over 6% when people target the average consumption level, and is over 3.5% when people make an upward comparison. Correspondingly, for all the combinations under a net labor outflow, as the asymptotic ratio of ex-ante skill density to ex post skill density increases from 1.05 to 1.55, registering a 47.6% increase in skill density, the *tax increase* is over 8.5% when people target the average level, and is over 6% when people make an upward comparison.

Table 4: Tax increase (%) as the degree of inequality ($1/a_i$) increases from 1/3 to 1/2

	$\varepsilon = 0.25$	$\varepsilon = 0.33$	$\varepsilon = 0.5$
(1) net labor inflow			
$q_i(\infty) = 0.95$:			
$\sigma_i = 0.25$	13.0, 18.4, 28.1	13.0, 8.8, 27.4	12.8, 17.8, 25.6
$\sigma_i = 0.50$	6.0, 10.2, 28.1	6.1, 10.1, 27.4	6.2, 10.0, 25.6
$q_i(\infty) = 0.67$:			
$\sigma_i = 0.25$	15.6, 20.5, 28.1	15.2, 20.2, 27.4	15.1, 19.7, 25.6
$\sigma_i = 0.50$	8.0, 12.5, 28.1	8.1, 12.5, 27.4	8.0, 12.3, 25.6
(2) net labor outflow			
$q_i(\infty) = 1.05$:			
$\sigma_i = 0.25$	12.4, 17.9, 28.1	12.1, 17.5, 27.4	11.9, 16.9, 25.6
$\sigma_i = 0.50$	5.5, 9.3, 28.1	5.6, 9.4, 27.4	5.7, 9.5, 25.6
$q_i(\infty) = 1.55$:			
$\sigma_i = 0.25$	9.6, 15.2, 28.1	9.7, 15.1, 27.4	9.6, 14.9, 25.6
$\sigma_i = 0.50$	4.0, 7.3, 28.1	4.1, 7.0, 27.4	4.2, 7.2, 25.6

Third, under average level and upward comparisons, an increase in relativity intensity (σ_i) contributes to significant tax increases of over 20% for all the cases that we consider, although the specific magnitude of tax increase varies in reference to consumption constraint and differs from the circumstance with a net labor inflow to that with a net labor outflow. In an equilibrium with a net labor inflow, as relativity intensity increases from 0.25 to 0.5, the increase in the optimal tax rate is larger under average level comparison than upward comparison for the two sizes of net labor inflow that we consider. If the size of net labor outflow is small, this prediction carries over to an equilibrium with a net labor outflow. In contrast, the increase in optimal tax rate is rather smaller under the average level comparison than under an upward comparison.

Fourth, all else being equal, as the degree of inequality ($1/a_i$) increases from 1/3 to 1/2, optimal AMTRs will increase in all the cases that we consider. In particular, Table 4 shows that the tax increase is highest when people target the median-skill's consumption (over 25%), lowest when targeting the average level (over 4%), and in between when making an upward comparison (over 7%).

Note that we set migration elasticity to be $\theta_i(\infty) = 0.25$ in Tables 2-4. It would be worthwhile

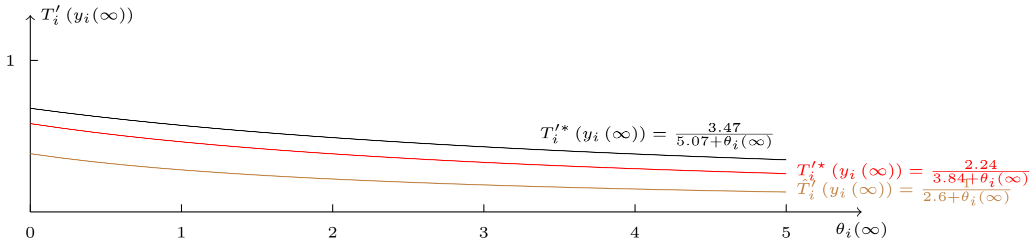

Figure 6: Optimal AMTRs change in migration elasticity under net labor inflow
Note: $\varepsilon = 0.25$, $a_i = 2$, $\Gamma_i(\infty)/\lambda_i = 0.5$, and $q_i(\infty) = 0.95$.

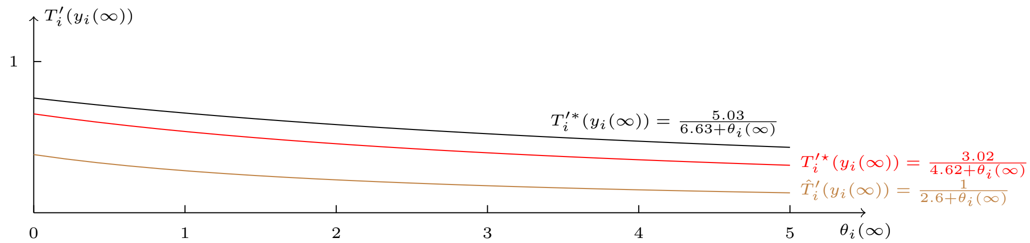

Figure 7: Optimal AMTRs change in migration elasticity under net labor outflow
Note: $\varepsilon = 0.25$, $a_i = 2$, $\Gamma_i(\infty)/\lambda_i = 0.5$, and $q_i(\infty) = 1.55$.

investigating how the optimal AMTR varies with migration elasticity and how this feature may differ under alternative reference consumption constraints. With realistic estimates of relevant parameters, Figures 6-7 show that the optimal AMTR is strictly decreasing and convex in migration elasticity under the three reference consumption constraints. Moreover, on the interval of migration elasticity that we consider, the optimal tax rate is highest when comparing with the average consumption level, lowest when comparing with the median-skill's consumption, and in between when making an upward comparison.

Table 5: AMTRs (%) when $q_i(\infty) = 1$ and migration elasticity varies

	$\varepsilon = 0.25$	$\varepsilon = 0.25$	$\varepsilon = 0.33$	$\varepsilon = 0.33$	$\varepsilon = 0.5$	$\varepsilon = 0.5$
	$a_i = 2$	$a_i = 3$	$a_i = 2$	$a_i = 3$	$a_i = 2$	$a_i = 3$
$\theta_i(\infty) = 0.25$:						
$\sigma_i = 0.25$	50.1, 42.9, 35.1	44.5, 36.3, 27.4	51.0, 44.0, 36.3	45.3, 37.2, 28.5	52.8, 46.0, 38.7	46.9, 39.2, 30.8
$\sigma_i = 0.50$	66.1, 55.4, 35.1	62.4, 50.5, 27.4	66.6, 56.2, 36.3	62.9, 51.2, 28.5	67.8, 57.8, 38.7	64.0, 52.6, 30.8
tax increase (%) as $\sigma_i \uparrow$	31.9, 29.1, 0	40.2, 39.1, 0	30.6, 27.7, 0	38.9, 37.6, 0	28.4, 25.7, 0	36.5, 34.2, 0
$\theta_i(\infty) = 0$:						
$\sigma_i = 0.25$	53.7, 46.4, 38.5	46.9, 38.6, 29.4	54.8, 47.7, 39.9	47.9, 39.7, 30.7	57.0, 50.3, 42.9	49.9, 42.0, 33.3
$\sigma_i = 0.50$	69.2, 58.9, 38.5	64.7, 52.9, 29.4	70.0, 59.9, 39.9	65.3, 53.8, 30.7	71.4, 61.9, 42.9	66.7, 55.5, 33.3
tax increase (%) as $\sigma_i \uparrow$	28.9, 26.9, 0	38.0, 37.0, 0	27.7, 25.6, 0	36.3, 35.5, 0	25.3, 23.1, 0	33.7, 32.1, 0
tax cut (%) as $0 \uparrow 0.25$:						
$\sigma_i = 0.25$	6.7, 7.5, 8.8	5.1, 6.0, 6.8	6.9, 7.8, 9.0	5.4, 6.3, 7.2	7.4, 8.5, 9.8	6.0, 6.7, 7.5
$\sigma_i = 0.50$	4.5, 5.9, 8.8	3.6, 4.5, 6.8	4.9, 6.2, 9.0	3.7, 4.8, 7.2	5.0, 6.6, 9.8	4.0, 5.2, 7.5

Finally, to identify the effect of migration on AMTRs only through the elasticity channel, we set $q_i(\infty) = 1$ to exclude (or control for) the level effect of migration; namely, there is a zero amount of *net* labor flow, implying that even though labor flows between the two countries do occur, there is neither a positive amount of net inflow nor a positive amount of net outflow. We compare the case with $\theta_i(\infty) = 0$, which can be interpreted as an autarky economy circumstance without mobile labor across countries, to the case with $\theta_i(\infty) = 0.25$. From Table 5, we obtain three main findings.

First, the equilibrium tax cut induced by the increase in migration elasticity from 0 to 0.25

is highest under a comparison with the median-skill type, lowest under a comparison with the average consumption level, and in between under an upward comparison. Roughly, all things being equal, the intensity of migration-induced tax competition will decrease for the population that serves as the social comparison target in a society. In other words, as the externality-correcting role of the income tax policy becomes more relevant, the downward pressure of tax competition placed on the equilibrium tax rate will be weakened.

Second, while the size of the tax cut induced by the increase in migration elasticity from 0 to 0.25 is independent of relativity intensity when people make comparisons toward the median-skill type, it actually decreases as relativity intensity increases from 0.25 to 0.5 for the other two comparison scenarios. The relevance of the form of social comparison is in the definitions of reference consumptions, and only in the latter two scenarios will the highest skilled be involved in the creation of positional externalities; hence, the tax policy facing them must exhibit an externality-corrective function. Consequently, the demand for correcting for negative externalities will weaken the influence of tax competition in reducing the AMTRs levied on the most skilled individuals.

Third, concerning the tax increase induced by an increase of relativity intensity, which, as discussed in the preceding point, is relevant only for the two cases in which people compare themselves with average consumption or make an upward comparison, we find that the corresponding tax increases are larger for $\theta_i(\infty) = 0.25$ than for $\theta_i(\infty) = 0$. For instance, let $\varepsilon = 0.25$ and $a_i = 2$, and the tax increases are 31.9% and 29.1%, respectively, for the two comparison scenarios under $\theta_i(\infty) = 0.25$, whereas they are 28.9% and 26.9% under $\theta_i(\infty) = 0$. Moreover, the difference in the magnitude of tax increase is totally due to the variation of $\theta_i(\infty)$. An intuitively appealing explanation is the following. In the autarky economy, the positional externality is generated entirely by the social comparison of private goods consumption, while in the open economy, besides this type of externality, there is an additional positional externality caused by the between-country comparison of public goods consumption.

7 Concluding Summary

The main purpose of this study was to examine how migration would modify the effects of inequality and relativity on the optimal MTRs levied on top labor incomes in an open Mirrleesian economy with two competing countries. We consider three definitions of reference consumption for any given skill type — the average consumption of all skill types, the quasi-average consumption of the higher skill types, and the consumption of the median-skill type. The first definition is the often used in the related literature, and can be regarded as a trivial combination of downward and upward social comparisons for any given skill type between the lower and upper ends of skill distribution, serving as a benchmark for a comparative analysis. The second scenario captures the effect of upward social comparisons, and the last scenario can be interpreted as a comparison with the middle-income class. Moreover, to capture inter-country social comparisons arising from the deepening globalization process, we also let individuals derive utility from the relative consumption of national public goods; for example, defined as the ratio of domestic public goods provision to the competing country's public goods provision.

Thus, taking the income tax policy and public goods provision of the competing country as a given, the government in each country adopts maxi-min normative criterion for designing an optimal nonlinear tax profile that is incentive compatible — truth-telling in terms of privately observable skill types — and based on the residence principle, as well as for determining an optimal level of public goods provision that balances government budget constraints. In addition to the response on the intensive margin by adjusting the allocation of a given amount of time between labor and leisure, individuals can move between the two countries after paying a certain amount of migration costs; hence, the governments face an internationally mobile tax base, which must be considered in the first place toward a joint design of the optimal income tax policy and the optimal level of public goods provision.

We obtain different optimal MTRs for the highest skilled individuals under the three forms of reference consumption, and our numerical examples reveal that the optimal tax rate is highest under a comparison with the average consumption level, lowest when targeting the median-skill's consumption, and in between under an upward consumption comparison. Consequently, it is practically relevant to design surveys to identify which of the three forms might represent a better approximation for a given population.

Concerning the interplay of relativity and migration for optimally taxing top-income earners, we provide a set of sufficient conditions under which, if the intensity of relativity concerns has been high, while migration elasticity has been small, then an increase in migration elasticity will, ceteris paribus, reinforce the positive effect of relativity placed on the optimal tax rate. Otherwise, if the intensity of relativity concerns has been low, while migration elasticity has been large, then a further increase in migration elasticity will, ceteris paribus, weaken the positive effect of relativity concerns. Thus, the joint effect of relativity and migration on the optimal tax rate is dependent on the relative strength of these two opposite forces in the status quo under consideration.

Additionally, to examine the interplay of inequality and migration for optimally taxing top labor incomes, we show that an increase in migration elasticity over a reasonable range of migration elasticity will, ceteris paribus, weaken the upward pressure of worsening inequality placed on the optimal tax rate. Imposing a high tax rate on top earners helps to correct the negative positional externality and simultaneously produce equity gains, but doing so would also drive out some individuals with the highest skill level, resulting in efficiency losses. Therefore, even when governments adopt the most redistributive social objective (maxi-min), for the proper design of MTRs levied on top labor incomes, the concern of retaining and attracting top talent is more determinant than the demand for income redistribution.

Statements and Declarations

We declare that we have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Akcigit, U., S. Baslandze, and S. Stantcheva (2016): "Taxation and the international mobility of inventors," *American Economic Review*, 106(10), 2930–81.

Ales, L., and C. Sleet (2016): "Taxing top CEO incomes," *American Economic Review*, 106(11), 3331–3366.

Alpizar, F., F. Carlsson, and O. Johansson-Stenman (2005): "How much do we care about absolute versus relative income and consumption?," *Journal of Economic Behavior & Organization*, 56(3), 405–421.

Alstadsæter, A., N. Johannesen, and G. Zucman (2019): "Tax evasion and inequality," *American Economic Review*, 109(6), 2073–2103.

Alvarez-Cuadrado, F., and N. V. Long (2012): "Envy and inequality," *The Scandinavian Journal of Economics*, 114(3), 949–973.

Aronsson, T., and O. Johansson-Stenman (2010): "Positional concerns in an OLG model: optimal labor and capital income taxation," *International Economic Review*, 51(4), 1071–1095.

——— (2013): "Veblen's theory of the leisure class revisited: Implications for optimal income taxation," *Social Choice and Welfare*, 41(3), 551–578.

——— (2014a): "When Samuelson met Veblen abroad: National and global public good provision when social comparisons matter," *Economica*, 81(322), 224–243.

——— (2014b): "Positional preferences in time and space: Optimal income taxation with dynamic social comparisons," *Journal of Economic Behavior & Organization*, 101, 1–23.

——— (2015): "Keeping up with the Joneses, the Smiths and the Tanakas: On international tax coordination and social comparisons," *Journal of Public Economics*, 131, 71–86.

——— (2018): "Paternalism against Veblen: Optimal taxation and non-respected preferences for social comparisons," *American Economic Journal: Economic Policy*, 10(1), 39–76.

Atkinson, A. B., T. Piketty, and E. Saez (2011): "Top incomes in the long run of history," *Journal of Economic Literature*, 49(1), 3–71.

Badel, A., M. Huggett, and W. Luo (2020): "Taxing top earners: A human capital perspective," *The Economic Journal*, 130(629), 1200–1225.

Becchetti, L., S. Castriota, L. Corrado, and E. G. Ricca (2013): "Beyond the Joneses: Inter-country income comparisons and happiness," *Journal of Socio-Economics*, 45, 187–195.

Bierbrauer, F., C. Brett, and J. A. Weymark (2013): "Strategic nonlinear income tax competition with perfect labor mobility," *Games and Economic Behavior*, 82, 292–311.

Bierbrauer, F. J. (2011): "On the optimality of optimal income taxation," *Journal of Economic Theory*, 146(5), 2105–2116.

Blanchet, T., L. Chancel, and A. Gethin (2022): "Why is Europe more equal than the United States?," *American Economic Journal: Applied Economics*, 14(4), 480–518.

Blanchflower, D. G., and A. J. Oswald (2004): "Well-being over time in Britain and the USA," *Journal of Public Economics*, 88(7-8), 1359–1386.

Blumkin, T., E. Sadka, and Y. Shem-Tov (2015): "International tax competition: Zero tax rate at the top re-established," *International Tax and Public Finance*, 22(5), 760–776.

Boadway, R., and L. Jacquet (2008): "Optimal marginal and average income taxation under maximin," *Journal of Economic Theory*, 143(1), 425–441.

Boeri, T., H. Brucker, F. Doquier, and H. Rapoport (2012): *Brain Drain and Brain Gain: The Global Competition to Attract High-Skilled Migrants*. Oxford University Press.

Brüggemann, B. (2021): "Higher taxes at the top: The role of entrepreneurs," *American Economic Journal: Macroeconomics*, 13(3), 1–36.

Brülhart, M., J. Gruber, M. Krapf, and K. Schmidheiny (2022): "Behavioral responses to wealth taxes: Evidence from Switzerland," *American Economic Journal: Economic Policy*, 14(4), 111–150.

Carlsson, F., O. Johansson-Stenman, and P. Martinsson (2007): "Do you enjoy having more than others? Survey evidence of positional goods," *Economica*, 74(296), 586–598.

Chetty, R. (2012): "Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply," *Econometrica*, 80(3), 969–1018.

Chetty, R., D. Grusky, M. Hell, N. Hendren, R. Manduca, and J. Narang (2017): "The fading American dream: Trends in absolute income mobility since 1940," *Science*, 356(6336), 398–406.

Clark, A. E., P. Frijters, and M. A. Shields (2008): "Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles," *Journal of Economic Literature*, 46(1), 95–144.

Clark, A. E., and C. Senik (2010): "Who compares to whom? The anatomy of income comparisons in Europe," *The Economic Journal*, 120(544), 573–594.

Corazzini, L., L. Esposito, and F. Majorano (2012): "Reign in hell or serve in heaven? A cross-country journey into the relative vs absolute perceptions of wellbeing," *Journal of Economic Behavior and Organization*, 81(3), 715–730.

Dai, D. (2020): "Voting over selfishly optimal tax schedules: Can Pigouvian tax redistribute income?," *Journal of Public Economic Theory*, 22(5), 1660–1686.

Dai, D., W. Gao, and G. Tian (2020): "Relativity, mobility, and optimal nonlinear income taxation in an open economy," *Journal of Economic Behavior & Organization*, 172(C), 57–82.

Dai, D., and G. Tian (2023): "Voting over selfishly optimal income tax schedules with tax-driven migrations," *Social Choice and Welfare*, 60, 183–235.

Diamond, P. (1998): "Optimal income taxation: An example with a U-shaped pattern of optimal marginal tax rates," *American Economic Review*, 88(1), 83–95.

Diamond, P., and E. Saez (2011): "The case for a progressive tax: From basic research to policy recommendations," *Journal of Economic Perspectives*, 25(4), 165–90.

Docquier, F., and A. Marfouk (2006): "International migration by education attainment (1990–2000)," *International Migration, Remittances and the Brain Drain*, pp. 151–199.

Dohmen, T., A. Falk, K. Fliessbach, U. Sunde, and B. Weber (2011): "Relative versus absolute income, joy of winning, and gender: Brain imaging evidence," *Journal of Public Economics*, 95(3-4), 279–285.

Duesenberry, J. S. (1949): *Income, Saving, and the Theory of Consumer Behavior*. Harvard University Press.

Ferrer-i Carbonell, A. (2005): "Income and well-being: An empirical analysis of the comparison income effect," *Journal of Public Economics*, 89(5-6), 997–1019.

Festinger, L. (1954): "A theory of social comparison processes," *Human Relations*, 7(2), 117–140.

Fliessbach, K., B. Weber, P. Trautner, T. Dohmen, U. Sunde, C. E. Elger, and A. Falk (2007): "Social comparison affects reward-related brain activity in the human ventral striatum," *Science*, 318(5854), 1305–1308.

Fong, C. (2001): "Social preferences, self-interest, and the demand for redistribution," *Journal of Public Economics*, 82(2), 225–246.

Frank, R. H. (2008): "Should public policy respond to positional externalities?," *Journal of Public Economics*, 92(8-9), 1777–1786.

Gereffi, G., and X. Luo (2014): "Risks and opportunities of participation in global value chains," Discussion paper, World Bank.

Grolleau, G., and S. Saïd (2008): "Do you prefer having more or more than others? Survey evidence on positional concerns in France," *Journal of Economic Issues*, 42(4), 1145–1158.

Guesnerie, R. (1998): *A Contribution to the Pure Theory of Taxation*. Cambridge University Press.

Hammond, P. J. (1979): "Straightforward individual incentive compatibility in large economies," *The Review of Economic Studies*, 46(2), 263–282.

Heffetz, O., and R. H. Frank (2011): "Preferences for status: Evidence and economic implications," in *Handbook of Social Economics*, vol. 1, pp. 69–91. Elsevier.

Jones, C. I. (2022): "Taxing top incomes in a world of ideas," *Journal of Political Economy*, 130(9), 2227–2274.

Kanbur, R., and M. Tuomala (2013): "Relativity, inequality, and optimal nonlinear income taxation," *International Economic Review*, 54(4), 1199–1217.

Kindermann, F., and D. Krueger (2022): "High marginal tax rates on the top 1 percent? Lessons from a life-cycle model with idiosyncratic income risk," *American Economic Journal: Macroeconomics*, 14(2), 319–366.

Kleven, H., C. Landais, M. Muñoz, and S. Stantcheva (2020): "Taxation and migration: Evidence and policy implications," *Journal of Economic Perspectives*, 34(2), 119–142.

Kleven, H. J., C. Landais, and E. Saez (2013): "Taxation and international migration of superstars: Evidence from the European football market," *American Economic Review*, 103(5), 1892–1924.

Kleven, H. J., C. Landais, E. Saez, and E. Schultz (2014): "Migration and wage effects of taxing top earners: Evidence from the foreigners tax scheme in Denmark," *Quarterly Journal of Economics*, 129(1), 333–378.

Layard, R. (1980): "Human satisfactions and public policy," *The Economic Journal*, 90(360), 737–750.

Lehmann, E., L. Simula, and A. Trannoy (2014): "Tax me if you can! Optimal nonlinear income tax between competing governments," *Quarterly Journal of Economics*, 129(4), 1995–2030.

Lipatov, V., and A. Weichenrieder (2015): "Welfare and labor supply implications of tax competition for mobile labor," *Social Choice and Welfare*, 45(2), 457–477.

Luttmer, E. F. (2005): "Neighbors as negatives: Relative earnings and well-being," *Quarterly Journal of Economics*, 120(3), 963–1002.

Micheletto, L. (2011): "Optimal nonlinear redistributive taxation and public good provision in an economy with Veblen effects," *Journal of Public Economic Theory*, 13(1), 71–96.

Mirrlees, J. A. (1971): "An exploration in the theory of optimum income taxation," *Review of Economic Studies*, 38(2), 175–208.

——— (1982): "Migration and optimal income taxes," *Journal of Public Economics*, 18(3), 319–341.

Morelli, M., H. Yang, and L. Ye (2012): "Competitive nonlinear taxation and constitutional choice," *American Economic Journal: Microeconomics*, 4(1), 142–175.

Mujcic, R., and P. Frijters (2015): "Conspicuous consumption, conspicuous health, and optimal taxation," *Journal of Economic Behavior & Organization*, 111, 59–70.

Nishi, A., H. Shirado, D. G. Rand, and N. A. Christakis (2015): "Inequality and visibility of wealth in experimental social networks," *Nature*, 526(7573), 426–429.

OECD (2008): "The global competition for talent: Mobility of the highly skilled," Discussion paper, Paris: OECD.

Oswald, A. J. (1983): "Altruism, jealousy and the theory of optimal non-linear taxation," *Journal of Public Economics*, 20(1), 77–87.

Papademetriou, D. G., and M. Sumption (2013): "Attracting and selecting from the global talent pool — Policy challenges," Discussion paper, Migration Policy Institute, Washington DC.

Piketty, T. (2014): *Capital in the Twenty-First Century*. Harvard University Press.

Piketty, T., and E. Saez (2013): "Optimal labor income taxation," in *Handbook of Public Economics*, vol. 5, pp. 391–474. Elsevier.

Piketty, T., E. Saez, and S. Stantcheva (2014): "Optimal taxation of top labor incomes: A tale of three elasticities," *American Economic Journal: Economic Policy*, 6(1), 230–271.

Rawls, J. (1971): *A Theory of Justice*. Harvard University Press.

Saez, E. (2001): "Using elasticities to derive optimal income tax rates," *Review of Economic Studies*, 68(1), 205–229.

Saez, E., J. Slemrod, and S. H. Giertz (2012): "The elasticity of taxable income with respect to marginal tax rates: A critical review," *Journal of Economic Literature*, 50(1), 3–50.

Samuelson, P. A. (1954): "The pure theory of public expenditure," *The Review of Economics and Statistics*, pp. 387–389.

Scheuer, F., and I. Werning (2017): "The taxation of superstars," *The Quarterly Journal of Economics*, 132(1), 211–270.

Schwab, K., and P. Vanham (2021): *Stakeholder Capitalism: A Global Economy that Works for Progress, People and Planet*. Wiley.

Seim, D. (2017): "Behavioral responses to wealth taxes: Evidence from Sweden," *American Economic Journal: Economic Policy*, 9(4), 395–421.

Senik, C. (2009): "Direct evidence on income comparisons and their welfare effects," *Journal of Economic Behavior & Organization*, 72(1), 408–424.

Simula, L., and A. Trannoy (2010): "Optimal income tax under the threat of migration by top-income earners," *Journal of Public Economics*, 94(1-2), 163–173.

——— (2012): "Shall we keep the highly skilled at home? The optimal income tax perspective," *Social Choice and Welfare*, 39(4), 751–782.

Solnick, S. J., and D. Hemenway (1998): "Is more always better?: A survey on positional concerns," *Journal of Economic Behavior & Organization*, 37(3), 373–383.

Solnick, S. J., and D. Hemenway (2005): "Are positional concerns stronger in some domains than in others?," *American Economic Review*, 95(2), 147–151.

Stiglitz, J. E. (2012): *The Price of Inequality*. W.W. Norton & Company.

Stiglitz, J. E. (2018): "Rethinking globalization in the Trump era: US-China relations," *Frontiers of Economics in China*, 13(2), 133–146.

Sznycer, D., M. F. L. Seal, A. Sell, J. Lim, R. Porat, S. Shalvi, E. Halperin, L. Cosmides, and J. Tooby (2017): "Support for redistribution is shaped by compassion, envy, and self-interest, but not a taste for fairness," *The Proceedings of the National Academy of Sciences*, 114(31), 8420–8425.

Uribe-Terán, C. (2021): "Higher taxes at the top? The role of tax avoidance," *Journal of Economic Dynamics and Control*, 129(104187).

Veblen, T. (1899): *The Theory of the Leisure Class: An Economic Study in the Evolution of Institutions*. Macmillan.

Wildasin, D. E. (2006): "Global competition for mobile resources: Implications for equity, efficiency and political economy," *CESifo Economic Studies*, 52(1), 61–110.

Wilson, J. D. (2009): "Income taxation and skilled migration: The analytical issues," in *Skilled Immigration Today: Prospects, Problems, and Policies*, chap. 10, pp. 285–314. Jagdish Bhagwati (ed.) and Gordon Hanson (ed.), Oxford University Press.

Xie, Y., and X. Zhou (2014): "Income inequality in today's China," *The Proceedings of the National Academy of Sciences*, 111(19), 6928–6933.

Appendix: Proofs

Proof of Lemma 4.1. Treating consumption $c_i(w)$ as an implicit function of $U_i(w)$, $l_i(w)$, μ_i, G_i and G_{-i}, and rewriting it as $\phi_i(\cdot)$, the indirect utility of a type-w individual in country $i \in \{A, B\}$ can be written as

$$U_i(w) = v(\phi_i(U_i(w), l_i(w), \mu_i, G_i, G_{-i}), G_i; \mu_i, G_{-i}) - h(l_i(w)). \tag{34}$$

Applying the Implicit Function Theorem, we get from (34) the following derivatives:

$$\frac{\partial \phi_i}{\partial l_i} = \frac{h'}{v_c} > 0, \frac{\partial \phi_i}{\partial U_i} = \frac{1}{v_c} > 0, \frac{\partial \phi_i}{\partial \mu_i} = -\frac{v_\mu}{v_c} > 0, \frac{\partial \phi_i}{\partial G_i} = -\frac{v_{G_i}}{v_c} < 0, \frac{\partial \phi_i}{\partial G_{-i}} = -\frac{v_{G_{-i}}}{v_c} > 0. \tag{35}$$

The Lagrangian of Problem (11) is given by:

$$\begin{aligned} & \mathcal{L}_i\left(\{U_i(w), l_i(w)\}_{w\in[\underline{w},\overline{w}]}, G_i, \mu_i; \lambda_i, \gamma_i, \{\varsigma_i(w)\}_{w\in[\underline{w},\overline{w}]}\right) \\ = \; & U_i(\underline{w}) + \lambda_i \int_{\underline{w}}^{\overline{w}} \left[w l_i(w) - \phi_i(U_i(w), l_i(w), \mu_i, G_i, G_{-i})\right] \tilde{f}_i(w) dw - \lambda_i G_i \\ & + \int_{\underline{w}}^{\overline{w}} \varsigma_i(w) \left[h'(l_i(w)) \frac{l_i(w)}{w} - \dot{U}_i(w)\right] dw \\ & + \gamma_i \left[\mu_i - \int_{\underline{w}}^{\overline{w}} \phi_i(U_i(w), l_i(w), \mu_i, G_i, G_{-i}) f_i(w) dw\right] \end{aligned} \tag{36}$$

where λ_i is the nonnegative multiplier associated with the fiscal budget constraint (8), $\varsigma_i(w)$ is the multiplier associated with the FOIC conditions (10), and γ_i is the nonnegative multiplier associated with the reference consumption constraint (3). It is easy to verify that both the government budget constraint and the reference consumption constraint are binding in an optimum, and so the standard complementary-slackness conditions yield $\lambda_i > 0$ and $\gamma_i > 0$.

Integrating by parts, we obtain

$$\int_{\underline{w}}^{\overline{w}} \varsigma_i(w) \dot{U}_i(w) dw = \varsigma_i(\overline{w}) U_i(\overline{w}) - \varsigma_i(\underline{w}) U_i(\underline{w}) - \int_{\underline{w}}^{\overline{w}} \dot{\varsigma}_i(w) U_i(w) dw. \tag{37}$$

Plugging (37) in (36) gives rise to

$$\begin{aligned} & \mathcal{L}_i\left(\{U_i(w), l_i(w)\}_{w\in[\underline{w},\overline{w}]}, G_i, \mu_i; \lambda_i, \gamma_i, \{\varsigma_i(w)\}_{w\in[\underline{w},\overline{w}]}\right) \\ = \; & U_i(\underline{w}) + \lambda_i \int_{\underline{w}}^{\overline{w}} \left[w l_i(w) - \phi_i(U_i(w), l_i(w), \mu_i, G_i, G_{-i})\right] \tilde{f}_i(w) dw - \lambda_i G_i \\ & + \varsigma_i(\underline{w}) U_i(\underline{w}) - \varsigma_i(\overline{w}) U_i(\overline{w}) + \int_{\underline{w}}^{\overline{w}} \left[\varsigma_i(w) h'(l_i(w)) \frac{l_i(w)}{w} + \dot{\varsigma}_i(w) U_i(w)\right] dw \\ & + \gamma_i \left[\mu_i - \int_{\underline{w}}^{\overline{w}} \phi_i(U_i(w), l_i(w), \mu_i, G_i, G_{-i}) f_i(w) dw\right]. \end{aligned}$$

Given the existence of an interior solution, the necessary conditions under (35) are:

$$\begin{aligned} \frac{\partial \mathcal{L}_i}{\partial l_i(w)} = \; & \lambda_i \left[w - \frac{h'(l_i(w))}{v_c(c_i(w), G_i; \mu_i, G_{-i})}\right] \tilde{f}_i(w) - \gamma_i \frac{h'(l_i(w))}{v_c(c_i(w), G_i; \mu_i, G_{-i})} f_i(w) \\ & + \frac{\varsigma_i(w) h'(l_i(w))}{w} \left[1 + \frac{l_i(w) h''(l_i(w))}{h'(l_i(w))}\right] = 0 \quad \forall w \in [\underline{w}, \overline{w}], \end{aligned} \tag{38}$$

$$\begin{aligned} \frac{\partial \mathcal{L}_i}{\partial U_i(w)} = & - \frac{\lambda_i \tilde{f}_i(w)}{v_c(c_i(w), G_i; \mu_i, G_{-i})} + \lambda_i T_i(y_i(w)) \eta_i(w) \tilde{f}_i(w) \\ & - \frac{\gamma_i f_i(w)}{v_c(c_i(w), G_i; \mu_i, G_{-i})} + \dot{\varsigma}_i(w) = 0 \quad \forall w \in (\underline{w}, \overline{w}), \end{aligned} \tag{39}$$

$$\begin{aligned} \frac{\partial \mathcal{L}_i}{\partial U_i(\underline{w})} &= 1 + \varsigma_i(\underline{w}) = 0, \\ \frac{\partial \mathcal{L}_i}{\partial U_i(\overline{w})} &= -\varsigma_i(\overline{w}) = 0, \end{aligned} \tag{40}$$

$$\begin{aligned} \frac{\partial \mathcal{L}_i}{\partial G_i} = \; & \lambda_i \int_{\underline{w}}^{\overline{w}} \frac{v_{G_i}(c_i(w), G_i; \mu_i, G_{-i})}{v_c(c_i(w), G_i; \mu_i, G_{-i})} \tilde{f}_i(w) dw \\ & + \gamma_i \int_{\underline{w}}^{\overline{w}} \frac{v_{G_i}(c_i(w), G_i; \mu_i, G_{-i})}{v_c(c_i(w), G_i; \mu_i, G_{-i})} f_i(w) dw - \lambda_i = 0, \end{aligned} \tag{41}$$

and

$$\begin{aligned}\frac{\partial \mathcal{L}_i}{\partial \mu_i} &= \lambda_i \int_{\underline{w}}^{\overline{w}} \frac{v_\mu(c_i(w), G_i; \mu_i, G_{-i})}{v_c(c_i(w), G_i; \mu_i, G_{-i})} \tilde{f}_i(w) dw \\ &\quad + \gamma_i \left[1 + \int_{\underline{w}}^{\overline{w}} \frac{v_\mu(c_i(w), G_i; \mu_i, G_{-i})}{v_c(c_i(w), G_i; \mu_i, G_{-i})} f_i(w) dw\right] = 0.\end{aligned} \tag{42}$$

Using (39), we get

$$\frac{\dot{\varsigma}_i(w)}{\lambda_i} = \frac{\gamma_i}{\lambda_i} \frac{f_i(w)}{v_c(c_i(w), G_i; \mu_i, G_{-i})} + \frac{\tilde{f}_i(w)}{v_c(c_i(w), G_i; \mu_i, G_{-i})} - T_i(y_i(w))\eta_i(w)\tilde{f}_i(w).$$

Integrating on both sides of this equation and using the transversality condition (40), we obtain

$$\begin{aligned}-\frac{\varsigma_i(w)}{\lambda_i} &= \frac{\gamma_i}{\lambda_i} \int_w^{\overline{w}} \frac{f_i(t)}{v_c(c_i(t), G_i; \mu_i, G_{-i})} dt \\ &\quad + \int_w^{\overline{w}} \left[\frac{1}{v_c(c_i(t), G_i; \mu_i, G_{-i})} - T_i(y_i(t))\eta_i(t)\right] \tilde{f}_i(t) dt.\end{aligned} \tag{43}$$

Rearranging (38) via using FOC (4), we have

$$\frac{T_i'(y_i(w))}{1 - T_i'(y_i(w))} = \frac{\gamma_i}{\lambda_i} \frac{f_i(w)}{\tilde{f}_i(w)} - \frac{\varsigma_i(w)}{\lambda_i} \frac{v_c(c_i(w), G_i; \mu_i, G_{-i})}{w\tilde{f}_i(w)} \left[1 + \frac{l_i(w)h''(l_i(w))}{h'(l_i(w))}\right] \quad \forall w \in [\underline{w}, \overline{w}]. \tag{44}$$

Substituting (43) into (44) gives the tax formula characterizing second-best MTRs, with γ_i/λ_i determined by solving (42). And the optimality condition for public goods provision is given by (41). ■

Proof of Proposition 4.1. Note that

$$\frac{\partial \frac{T_i'^*(y_i(\infty))}{1 - T_i'^*(y_i(\infty))}}{\partial \theta_i(\infty)} = \underbrace{\frac{\partial \frac{T_i'^*(y_i(\infty))}{1 - T_i'^*(y_i(\infty))}}{\partial T_i'^*(y_i(\infty))}}_{+} \cdot \frac{\partial T_i'^*(y_i(\infty))}{\partial \theta_i(\infty)}$$

the comparative static result with respect to $\theta_i(\infty)$ is immediate by using tax formula (16). Using (17) yields:

$$\frac{\partial T_i'^*(y_i(\infty))}{\partial q_i(\infty)} = \frac{\left(1 + \frac{1+\varepsilon}{a_i}\right) \frac{\gamma_i}{\lambda_i}(\infty) \left[1 + \frac{1+\varepsilon}{a_i} \cdot \theta_i(\infty)\right]}{\left\{\left(1 + \frac{1+\varepsilon}{a_i}\right) \left[1 + \frac{\gamma_i}{\lambda_i}(\infty) \cdot q_i(\infty)\right] + \frac{1+\varepsilon}{a_i} \cdot \theta_i(\infty)\right\}^2} > 0. \tag{45}$$

Additionally, it follows from (16) that

$$\frac{\partial \frac{T_i'^*(y_i(\infty))}{1 - T_i'^*(y_i(\infty))}}{\partial \left(\frac{1+\varepsilon}{a_i}\right)} = \frac{1 + \frac{\gamma_i}{\lambda_i}(\infty) q_i(\infty) \left[1 - \theta_i(\infty)\right]}{\left[1 + \frac{1+\varepsilon}{a_i} \cdot \theta_i(\infty)\right]^2},$$

and hence the proof is completed. ■

Proof of Proposition 4.2. To prove Proposition 4.2(i), we take partial derivative of (45) with respect to $\theta_i(\infty)$ and simplifying the algebra yields:

$$\begin{aligned}&\frac{\partial^2 T_i'^*(y_i(\infty))}{\partial q_i(\infty) \partial \theta_i(\infty)} \\ &= \frac{\frac{1+\varepsilon}{a_i} \left(1 + \frac{1+\varepsilon}{a_i}\right) \frac{\gamma_i}{\lambda_i}(\infty) \left[\left(1 + \frac{1+\varepsilon}{a_i}\right) \frac{\gamma_i}{\lambda_i}(\infty) q_i(\infty) - 1 + \frac{1+\varepsilon}{a_i} \left(1 - \theta_i(\infty)\right)\right]}{\left\{\left(1 + \frac{1+\varepsilon}{a_i}\right) \left[1 + \frac{\gamma_i}{\lambda_i}(\infty) \cdot q_i(\infty)\right] + \frac{1+\varepsilon}{a_i} \cdot \theta_i(\infty)\right\}^3}.\end{aligned}$$

Hence the comparative static result given in Proposition 4.2(i) is immediate.

To prove Proposition 4.2(ii), we use (17) and obtain

$$\frac{\partial T_i^{\prime *}(y_i(\infty))}{\partial \theta_i(\infty)} = -\frac{\frac{1+\varepsilon}{a_i}\left(1+\frac{1+\varepsilon}{a_i}\right)\frac{\gamma_i}{\lambda_i}(\infty)\cdot q_i(\infty)+\left(\frac{1+\varepsilon}{a_i}\right)^2}{\left\{\left(1+\frac{1+\varepsilon}{a_i}\right)\left[1+\frac{\gamma_i}{\lambda_i}(\infty)\cdot q_i(\infty)\right]+\frac{1+\varepsilon}{a_i}\cdot\theta_i(\infty)\right\}^2}. \tag{46}$$

Differentiating (46) with respect to $(1+\varepsilon)/a_i$ and simplifying the algebra yields:

$$\begin{aligned}&\frac{\partial^2 T_i^{\prime *}(y_i(\infty))}{\partial \theta_i(\infty)\partial\left(\frac{1+\varepsilon}{a_i}\right)}\\ =&\ -\frac{\left(1+\frac{1+\varepsilon}{a_i}\right)\left(\frac{\gamma_i}{\lambda_i}(\infty)\cdot q_i(\infty)\right)^2+2\left(\frac{1+\varepsilon}{a_i}\right)+\left[1+(3-\theta_i(\infty))\frac{1+\varepsilon}{a_i}\right]\frac{\gamma_i}{\lambda_i}(\infty)\cdot q_i(\infty)}{\left\{\left(1+\frac{1+\varepsilon}{a_i}\right)\left[1+\frac{\gamma_i}{\lambda_i}(\infty)\cdot q_i(\infty)\right]+\frac{1+\varepsilon}{a_i}\cdot\theta_i(\infty)\right\}^3},\end{aligned}$$

so the joint effect of $\theta_i(\infty)$ and $(1+\varepsilon)/a_i$ on $T_i^{\prime *}(y_i(\infty))$ is characterized as stated in Proposition 4.2(ii).

Similarly, we could have

$$\begin{aligned}&\frac{\partial^2 T_i^{\prime *}(y_i(\infty))}{\partial \theta_i(\infty)\partial\left(\frac{\gamma_i}{\lambda_i}(\infty)q_i(\infty)\right)}\\ =&\ \frac{\frac{1+\varepsilon}{a_i}\left(1+\frac{1+\varepsilon}{a_i}\right)\left[\left(1+\frac{1+\varepsilon}{a_i}\right)\frac{\gamma_i}{\lambda_i}(\infty)q_i(\infty)-1+\frac{1+\varepsilon}{a_i}(1-\theta_i(\infty))\right]}{\left\{\left(1+\frac{1+\varepsilon}{a_i}\right)\left[1+\frac{\gamma_i}{\lambda_i}(\infty)\cdot q_i(\infty)\right]+\frac{1+\varepsilon}{a_i}\cdot\theta_i(\infty)\right\}^3},\end{aligned}$$

which thus leads to the joint effect of $\theta_i(\infty)$ and $\frac{\gamma_i}{\lambda_i}(\infty)q_i(\infty)$ on $T_i^{\prime *}(y_i(\infty))$:

$$\frac{\partial^2 T_i^{\prime *}(y_i(\infty))}{\partial \theta_i(\infty)\partial\left(\frac{\gamma_i}{\lambda_i}(\infty)q_i(\infty)\right)} > 0 \text{ for } \begin{cases}\frac{\gamma_i}{\lambda_i}(\infty)q_i(\infty) > \frac{a_i}{1+\varepsilon+a_i} \text{ and } \theta_i(\infty)\le 1,\\ \text{or } \frac{\gamma_i}{\lambda_i}(\infty)q_i(\infty) \ge \frac{a_i}{1+\varepsilon+a_i} \text{ and } \theta_i(\infty) < 1;\end{cases} \tag{47}$$

and

$$\frac{\partial^2 T_i^{\prime *}(y_i(\infty))}{\partial \theta_i(\infty)\partial\left(\frac{\gamma_i}{\lambda_i}(\infty)q_i(\infty)\right)} < 0 \text{ for } \begin{cases}\frac{\gamma_i}{\lambda_i}(\infty)q_i(\infty) < \frac{a_i}{1+\varepsilon+a_i} \text{ and } \theta_i(\infty)\ge 1,\\ \text{or } \frac{\gamma_i}{\lambda_i}(\infty)q_i(\infty) \le \frac{a_i}{1+\varepsilon+a_i} \text{ and } \theta_i(\infty) > 1.\end{cases} \tag{48}$$

Using the assumption that $F_i(w) = F_{-i}(w)$ and (5) yields:

$$\frac{f_i(w)}{\tilde{f}_i(w)} = \begin{cases}\frac{1}{1+\Psi_{-i}(\Delta_i(w)|w)n_{-i}} & \text{for } \Delta_i(w)\ge 0,\\ \frac{1}{1-\Psi_i(-\Delta_i(w)|w)} & \text{for } \Delta_i(w)\le 0.\end{cases} \tag{49}$$

Using Assumption 4.1, (14) can be rewritten as

$$\frac{\gamma_i}{\lambda_i} \;=\; \frac{\int_{\underline{w}}^{\overline{w}}\sigma_i(w)\tilde{f}_i(w)dw}{1-\int_{\underline{w}}^{\overline{w}}\sigma_i(w)f_i(w)dw}. \tag{50}$$

Applying the assumption $\int_{\underline{w}}^{\infty}\tilde{f}_i(w)dw = 1$ and Assumption 4.2 to (50) produces:

$$\frac{\gamma_i}{\lambda_i}(\infty) \;=\; \frac{\sigma_i}{1-\sigma_i}. \tag{51}$$

Also, using Assumption 4.1, (5) and (50) yields:

$$\frac{\mathrm{d}\tilde{f}_i(w)}{\mathrm{d}\sigma_i(w)} = -\mu_i \cdot \frac{\partial \tilde{f}_i(w)}{\partial \Delta_i(w)} < 0. \tag{52}$$

It follows from (49), (51), (52) and Assumption 4.2 that

$$\frac{\partial\left(\frac{\gamma_i}{\lambda_i}(\infty)q_i(\infty)\right)}{\partial \sigma_i} = \left(\frac{1}{1-\sigma_i}\right)^2 \lim_{w\uparrow\infty}\frac{f_i(w)}{\tilde{f}_i(w)} + \frac{\sigma_i}{1-\sigma_i}\cdot \underbrace{\lim_{w\uparrow\infty}\left[\frac{f_i(w)}{\tilde{f}_i(w)^2}\left(-\frac{\mathrm{d}\tilde{f}_i(w)}{\mathrm{d}\sigma_i}\right)\right]}_{+}.$$

Applying this result to (47)-(48), the assertion stated in Proposition 4.2(iii-a) thus follows.

Treating G_i^* as a function of G_{-i}, we apply the Implicit Function Theorem and the Envelope Theorem to (5) and (1) and find that:

$$\begin{aligned}\frac{\mathrm{d}\tilde{f}_i(w)}{\mathrm{d}G_{-i}} &= \underbrace{\frac{\partial \tilde{f}_i(w)}{\partial \Delta_i(w)}\left(\frac{\partial U_i(w)}{\partial G_i^*} - \frac{\partial U_{-i}(w)}{\partial G_i^*}\right)}_{+}\cdot\frac{\mathrm{d}G_i^*}{\mathrm{d}G_{-i}} \\ &\quad + \underbrace{\frac{\partial \tilde{f}_i(w)}{\partial \Delta_i(w)}\left(\frac{\partial U_i(w)}{\partial G_{-i}} - \frac{\partial U_{-i}(w)}{\partial G_{-i}}\right)}_{-}.\end{aligned} \tag{53}$$

In light of (53), (51), and Assumption 4.2 we have

$$\frac{\partial\left(\frac{\gamma_i}{\lambda_i}(\infty)q_i(\infty)\right)}{\partial G_{-i}} = \frac{\sigma_i}{1-\sigma_i}\cdot \underbrace{\lim_{w\uparrow\infty}\left[\frac{f_i(w)}{\tilde{f}_i(w)^2}\left(-\frac{\mathrm{d}\tilde{f}_i(w)}{\mathrm{d}G_{-i}}\right)\right]}_{+},$$

for $\mathrm{d}G_i^*/\mathrm{d}G_{-i} < 0$.

To complete the proof, we need to identify the condition such that $\mathrm{d}G_i^*/\mathrm{d}G_{-i} < 0$ is satisfied. Applying (53) and Assumption 4.1 to (50), we arrive at

$$\frac{\mathrm{d}(\gamma_i/\lambda_i)}{\mathrm{d}G_{-i}} = \underbrace{\frac{\int_{\underline{w}}^{\overline{w}}\sigma_i(w)\xi_i(w)dw}{1-\int_{\underline{w}}^{\overline{w}}\sigma_i(w)f_i(w)dw}}_{+}\cdot\frac{\mathrm{d}G_i^*}{\mathrm{d}G_{-i}} + \underbrace{\frac{\int_{\underline{w}}^{\overline{w}}\sigma_i(w)\beta_i(w)dw}{1-\int_{\underline{w}}^{\overline{w}}\sigma_i(w)f_i(w)dw}}_{-}. \tag{54}$$

Applying (53) and (54) to (18) yields:

$$\begin{aligned}\frac{\mathrm{d}Q_i(w)}{\mathrm{d}G_{-i}} &= \left[\xi_i(w) + \frac{f_i(w)\int_{\underline{w}}^{\overline{w}}\sigma_i(w)\xi_i(w)dw}{1-\int_{\underline{w}}^{\overline{w}}\sigma_i(w)f_i(w)dw}\right]\frac{\mathrm{d}G_i^*}{\mathrm{d}G_{-i}} + \beta_i(w) + \frac{f_i(w)\int_{\underline{w}}^{\overline{w}}\sigma_i(w)\beta_i(w)dw}{1-\int_{\underline{w}}^{\overline{w}}\sigma_i(w)f_i(w)dw} \\ &\equiv \underbrace{\frac{\partial Q_i(w)}{\partial G_i^*}}_{+}\cdot\frac{\mathrm{d}G_i^*}{\mathrm{d}G_{-i}} + \underbrace{\frac{\partial Q_i(w)}{\partial G_{-i}}}_{-}.\end{aligned} \tag{55}$$

Using Assumption 4.1, (18) and (15) gives the following rule for optimal public goods provision:

$$v_{G_i}^p\left(G_i, G_{-i}\right)\int_{\underline{w}}^{\overline{w}} Q_i(w)dw = 1. \tag{56}$$

Differentiating both sides of equation (56) with respect to G_{-i} and applying the Implicit Function Theorem yields:

$$\left(v^p_{G_iG_i} \cdot \frac{\mathrm{d}G^*_i}{\mathrm{d}G_{-i}} + v^p_{G_iG_{-i}}\right) \int_{\underline{w}}^{\overline{w}} Q_i(w)dw + v^p_{G_i} \cdot \int_{\underline{w}}^{\overline{w}} \left(\frac{\partial Q_i(w)}{\partial G^*_i}\frac{\mathrm{d}G^*_i}{\mathrm{d}G_{-i}} + \frac{\partial Q_i(w)}{\partial G_{-i}}\right) dw = 0.$$

Collecting terms and using (55) and Assumption 4.1 yields:

$$\begin{aligned} &\int_{\underline{w}}^{\overline{w}} \left(\underbrace{v^p_{G_iG_i} \cdot Q_i(w)}_{-} + \underbrace{v^p_{G_i} \cdot \frac{\partial Q_i(w)}{\partial G^*_i}}_{+}\right) dw \cdot \frac{\mathrm{d}G^*_i}{\mathrm{d}G_{-i}} \\ = \; & -\int_{\underline{w}}^{\overline{w}} \left(\underbrace{v^p_{G_iG_{-i}} \cdot Q_i(w)}_{-} + \underbrace{v^p_{G_i} \cdot \frac{\partial Q_i(w)}{\partial G_{-i}}}_{-}\right) dw, \end{aligned}$$

Therefore, the proof of Proposition 4.2(iii-b) is completed. ∎

Proof of Lemma 5.1. Replacing constraint (3) with constraint (21), the Lagrangian of Problem (11) amounts to

$$\begin{aligned} &\mathcal{L}_i\left(\{U_i(w), l_i(w), \mu_i(w)\}_{w\in[\underline{w},\overline{w}]}, G_i; \lambda_i, \{\gamma_i(w), \varsigma_i(w)\}_{w\in[\underline{w},\overline{w}]}\right) \\ = \; & U_i(\underline{w}) + \lambda_i \int_{\underline{w}}^{\overline{w}} \left[wl_i(w) - \phi_i(U_i(w), l_i(w), \mu_i, G_i, G_{-i})\right] \tilde{f}_i(w)dw - \lambda_i G_i \\ & + \int_{\underline{w}}^{\overline{w}} \varsigma_i(w) \left[h'(l_i(w))\frac{l_i(w)}{w} - \dot{U}_i(w)\right] dw \\ & + \int_{\underline{w}}^{\overline{w}} \gamma_i(w) \left[\mu_i(w) - \int_{w}^{\overline{w}} \phi_i(U_i(t), l_i(t), \mu_i, G_i, G_{-i}) f_i(t)dt\right] dw. \end{aligned} \tag{57}$$

Reversing the order of integration gives rise to

$$\begin{aligned} & \int_{\underline{w}}^{\overline{w}} \gamma_i(w) \int_{w}^{\overline{w}} \phi_i(U_i(t), l_i(t), \mu_i, G_i, G_{-i}) f_i(t)dtdw \\ = \; & \int_{\underline{w}}^{\overline{w}} \phi_i(U_i(t), l_i(t), \mu_i, G_i, G_{-i}) f_i(t) \left[\int_{\underline{w}}^{t} \gamma_i(w)dw\right] dt \\ \equiv \; & \int_{\underline{w}}^{\overline{w}} \phi_i(U_i(t), l_i(t), \mu_i, G_i, G_{-i}) f_i(t)\Gamma_i(t)dt. \end{aligned} \tag{58}$$

Substituting (58) into (57) and integrating by parts yields:

$$\begin{aligned} &\mathcal{L}_i\left(\{U_i(w), l_i(w), \mu_i(w)\}_{w\in[\underline{w},\overline{w}]}, G_i; \lambda_i, \{\gamma_i(w), \varsigma_i(w)\}_{w\in[\underline{w},\overline{w}]}\right) \\ = \; & U_i(\underline{w}) + \lambda_i \int_{\underline{w}}^{\overline{w}} \left[wl_i(w) - \phi_i(U_i(w), l_i(w), \mu_i, G_i, G_{-i})\right] \tilde{f}_i(w)dw - \lambda_i G_i \\ & + \varsigma_i(\underline{w})U_i(\underline{w}) - \varsigma_i(\overline{w})U_i(\overline{w}) + \int_{\underline{w}}^{\overline{w}} \left[\varsigma_i(w)h'(l_i(w))\frac{l_i(w)}{w} + \dot{\varsigma}_i(w)U_i(w)\right] dw \\ & + \int_{\underline{w}}^{\overline{w}} \gamma_i(w)\mu_i(w)dw - \int_{\underline{w}}^{\overline{w}} \phi_i(U_i(w), l_i(w), \mu_i, G_i, G_{-i}) f_i(w)\Gamma_i(w)dw. \end{aligned}$$

The remaining details are similar to those appear in the proof of Lemma 4.1, and hence are omitted to economize on the space. ■

Proof of Lemma 5.2. Using Assumption 4.2, the definition of $\mathcal{B}_i(w)$ given in Lemma 5.1 and the L'Hôpital's rule, we have

$$\lim_{w\uparrow\infty}\mathcal{B}_i(w)=\lim_{w\uparrow\infty}\frac{1-F_i(w)}{wf_i(w)}\cdot\frac{\frac{\tilde{F}_i(\infty)-\tilde{F}_i(w)}{1-F_i(w)}}{\tilde{f}_i(w)/f_i(w)}=\frac{1}{a_i}\cdot\lim_{w\uparrow\infty}\frac{\frac{\tilde{F}_i(\infty)-\tilde{F}_i(w)}{1-F_i(w)}}{\tilde{f}_i(w)/f_i(w)}=\frac{1}{a_i}.$$

Similarly, using Assumptions 4.1-4.2, (6), (23), and the L'Hôpital's rule, we arrive at

$$\lim_{w\uparrow\infty}\mathcal{C}_i(w)=1+\frac{\Gamma_i(\infty)}{\lambda_i}q_i(\infty)-\frac{T_i'(y_i(\infty))}{1-T_i'(y_i(\infty))}\theta_i(\infty).$$

Therefore, applying these two results and Assumption 4.2 to tax formula (22) produces the desired result. ■

Proof of Proposition 5.2. Given the similarity of tax formulas (27) and (17), the current proof is roughly the same as that of Proposition 4.2, and hence we just need to show some additional technical details as follows.

Applying Assumption 4.1 to (24) gives

$$\begin{aligned}\frac{\gamma_i(w)}{\lambda_i}&=\sigma_i(w)\tilde{f}_i(w)\left[1+\frac{\Gamma_i(w)}{\lambda_i}\frac{f_i(w)}{\tilde{f}_i(w)}\right]\\&=\sigma_i(w)\left[\tilde{f}_i(w)+f_i(w)\int_{\underline{w}}^{w}\frac{\gamma_i(t)}{\lambda_i}dt\right]\quad\forall w.\end{aligned}\tag{59}$$

Integrating over $[\underline{w},\overline{w}]$ on both sides of (59) and reversing the order of integration yields:

$$\int_{\underline{w}}^{\overline{w}}\frac{\gamma_i(w)}{\lambda_i}dw=\int_{\underline{w}}^{\overline{w}}\sigma_i(w)\tilde{f}_i(w)dw+\int_{\underline{w}}^{\overline{w}}\frac{\gamma_i(w)}{\lambda_i}\left[\int_{w}^{\overline{w}}\sigma_i(t)f_i(t)dt\right]dw.\tag{60}$$

Firstly, setting $\overline{w}=\infty$ and applying Assumption 4.2 to (60) gives rise to

$$\int_{\underline{w}}^{\infty}\frac{\gamma_i(w)}{\lambda_i}dw=\sigma_i\int_{\underline{w}}^{\infty}\tilde{f}_i(w)dw+\sigma_i\int_{\underline{w}}^{\infty}\frac{\gamma_i(w)}{\lambda_i}\left[1-F_i(w)\right]dw.\tag{61}$$

Taking derivative with respect to σ_i on both sides of equation (61) and collecting terms shows that, for almost all $w\in[\underline{w},\overline{w}]$,

$$\frac{\mathrm{d}\left(\frac{\gamma_i(w)}{\lambda_i}\right)}{\mathrm{d}\sigma_i}=\underbrace{\frac{\sigma_i}{1-\sigma_i[1-F_i(w)]}}_{+}\cdot\underbrace{\frac{\mathrm{d}\tilde{f}_i(w)}{\mathrm{d}\sigma_i}}_{-}+\underbrace{\frac{\tilde{f}_i(w)+\frac{\gamma_i(w)}{\lambda_i}[1-F_i(w)]}{1-\sigma_i[1-F_i(w)]}}_{+}.\tag{62}$$

If

$$-\epsilon_{\tilde{f}_i(w),\sigma_i}\equiv-\frac{\mathrm{d}\tilde{f}_i(w)}{\mathrm{d}\sigma_i}\frac{\sigma_i}{\tilde{f}_i(w)}\leq 1$$

holds true for almost all $w\in[\underline{w},\overline{w}]$, then using (62) produces:

$$\frac{\mathrm{d}\left(\frac{\Gamma_i(\infty)}{\lambda_i}\right)}{\mathrm{d}\sigma_i}=\int_{\underline{w}}^{\infty}\frac{\mathrm{d}\left(\frac{\gamma_i(w)}{\lambda_i}\right)}{\mathrm{d}\sigma_i}dw>0.$$

Now, using Assumption 4.2 yields:

$$\frac{\mathrm{d}\left(\frac{\Gamma_i(\infty)}{\lambda_i}q_i(\infty)\right)}{\mathrm{d}\sigma_i} = \underbrace{q_i(\infty)\cdot\frac{\mathrm{d}\left(\frac{\Gamma_i(\infty)}{\lambda_i}\right)}{\mathrm{d}\sigma_i}}_{+} + \frac{\Gamma_i(\infty)}{\lambda_i}\lim_{w\uparrow\infty}\left[-\frac{f_i(w)}{\tilde{f}_i(w)^2}\underbrace{\frac{\mathrm{d}\tilde{f}_i(w)}{\mathrm{d}\sigma_i}}_{-}\right] > 0,$$

whenever $-\epsilon_{\tilde{f}_i(w),\sigma_i} \leq 1$ holds for almost all $w \in [\underline{w}, \overline{w}]$.

Secondly, we differentiate (61) with respect to G_{-i} and rearrange the algebra, then for almost all $w \in [\underline{w}, \overline{w}]$ we arrive at:

$$\frac{\mathrm{d}\left(\frac{\gamma_i(w)}{\lambda_i}\right)}{\mathrm{d}G_{-i}} = \underbrace{\frac{\sigma_i}{1-\sigma_i[1-F_i(w)]}}_{+}\cdot\underbrace{\left(\frac{\partial\tilde{f}_i(w)}{\partial G_i}\frac{\mathrm{d}G_i^\star}{\mathrm{d}G_{-i}}+\frac{\partial\tilde{f}_i(w)}{\partial G_{-i}}\right)}_{-}$$

when $\mathrm{d}G_i^\star/\mathrm{d}G_{-i} < 0$. We thus have:

$$\frac{\mathrm{d}\left(\frac{\Gamma_i(\infty)}{\lambda_i}\right)}{\mathrm{d}G_{-i}} = \int_{\underline{w}}^{\infty}\frac{\mathrm{d}\left(\frac{\gamma_i(w)}{\lambda_i}\right)}{\mathrm{d}G_{-i}}dw < 0 \quad \text{for } \frac{\mathrm{d}G_i^\star}{\mathrm{d}G_{-i}} < 0.$$

Consequently, we make use of Assumption 4.2 again and arrive at:

$$\frac{\mathrm{d}\left(\frac{\Gamma_i(\infty)}{\lambda_i}q_i(\infty)\right)}{\mathrm{d}G_{-i}} = \frac{\Gamma_i(\infty)}{\lambda_i}q_i(\infty)\left\{\frac{\lambda_i}{\Gamma_i(\infty)}\cdot\underbrace{\frac{\mathrm{d}\left(\frac{\Gamma_i(\infty)}{\lambda_i}\right)}{\mathrm{d}G_{-i}}}_{-} - \lim_{w\uparrow\infty}\left[\frac{1}{\tilde{f}_i(w)}\underbrace{\frac{\mathrm{d}\tilde{f}_i(w)}{\mathrm{d}G_{-i}}}_{-}\right]\right\}$$

for $\mathrm{d}G_i^\star/\mathrm{d}G_{-i} < 0$.

Finally, to prove the necessary and sufficient condition for $\mathrm{d}G_i^\star/\mathrm{d}G_{-i} < 0$, we take derivative of (60) with respect to G_{-i} and rearrange the algebra, yielding

$$\int_{\underline{w}}^{\overline{w}}\underbrace{\left[1-\int_{w}^{\overline{w}}\sigma_i(t)f_i(t)dt\right]}_{+}\frac{\mathrm{d}\left(\frac{\gamma_i(w)}{\lambda_i}\right)}{\mathrm{d}G_{-i}}dw = \int_{\underline{w}}^{\overline{w}}\sigma_i(w)\frac{\mathrm{d}\tilde{f}_i(w)}{\mathrm{d}G_{-i}}dw,$$

in which we have used Assumption 4.1 again. We thus arrive at:

$$\frac{\mathrm{d}\left(\frac{\gamma_i(w)}{\lambda_i}\right)}{\mathrm{d}G_{-i}} = \frac{\sigma_i(w)}{1-\int_w^{\overline{w}}\sigma_i(t)f_i(t)dt}\frac{\mathrm{d}\tilde{f}_i(w)}{\mathrm{d}G_{-i}} \tag{63}$$

for almost all $w \in [\underline{w}, \overline{w}]$. It follows from using (28) and reversing the order of integration that

$$\begin{aligned} S_i &= \int_{\underline{w}}^{\overline{w}}\left[\tilde{f}_i(w)+f_i(w)\int_{\underline{w}}^{w}\frac{\gamma_i(t)}{\lambda_i}dt\right]dw \\ &= \int_{\underline{w}}^{\overline{w}}\tilde{f}_i(w)dw + \int_{\underline{w}}^{\overline{w}}\frac{\gamma_i(t)}{\lambda_i}\left[\int_t^{\overline{w}}f_i(w)dw\right]dt \\ &= \int_{\underline{w}}^{\overline{w}}\tilde{f}_i(w)dw + \int_{\underline{w}}^{\overline{w}}\frac{\gamma_i(w)}{\lambda_i}\left[1-F_i(w)\right]dw. \end{aligned} \tag{64}$$

Using (64), (63) and (53) yields:

$$
\begin{aligned}
\frac{\mathrm{d}S_i}{\mathrm{d}G_{-i}} &= \int_{\underline{w}}^{\overline{w}} \left\{1+\frac{\sigma_i(w)[1-F_i(w)]}{1-\int_w^{\overline{w}}\sigma_i(t)f_i(t)dt}\right\}\left(\frac{\partial \tilde{f}_i(w)}{\partial G_i}\cdot\frac{\mathrm{d}G_i^\star}{\mathrm{d}G_{-i}}+\frac{\partial \tilde{f}_i(w)}{\partial G_{-i}}\right)dw \\
&= \underbrace{\int_{\underline{w}}^{\overline{w}} \left\{1+\frac{\sigma_i(w)[1-F_i(w)]}{1-\int_w^{\overline{w}}\sigma_i(t)f_i(t)dt}\right\}\frac{\partial \tilde{f}_i(w)}{\partial G_i}dw}_{+}\cdot\frac{\mathrm{d}G_i^\star}{\mathrm{d}G_{-i}} \\
&\quad + \underbrace{\int_{\underline{w}}^{\overline{w}} \left\{1+\frac{\sigma_i(w)[1-F_i(w)]}{1-\int_w^{\overline{w}}\sigma_i(t)f_i(t)dt}\right\}\frac{\partial \tilde{f}_i(w)}{\partial G_{-i}}dw}_{-} \\
&= \underbrace{\frac{\partial S_i}{\partial G_i}}_{+}\cdot\frac{\mathrm{d}G_i^\star}{\mathrm{d}G_{-i}}+\underbrace{\frac{\partial S_i}{\partial G_{-i}}}_{-}.
\end{aligned}
\tag{65}
$$

Applying Assumption 4.1 and (28) to (25) gives rise to

$$
v^p_{G_i}(G_i, G_{-i})\cdot S_i = 1. \tag{66}
$$

We differentiate both sides of (66) with respect to G_{-i} and make use of Assumption 4.1 and (65), and then we find that

$$
\left(\underbrace{v^p_{G_iG_i}\cdot S_i}_{-}+\underbrace{v^p_{G_i}\cdot\frac{\partial S_i}{\partial G_i}}_{+}\right)\frac{\mathrm{d}G_i^\star}{\mathrm{d}G_{-i}} = -\left(\underbrace{v^p_{G_iG_{-i}}\cdot S_i}_{-}+\underbrace{v^p_{G_i}\cdot\frac{\partial S_i}{\partial G_{-i}}}_{-}\right).
$$

Therefore, we arrive at

$$
\mathrm{sgn}\left(v^p_{G_iG_i}\cdot S_i+v^p_{G_i}\cdot\frac{\partial S_i}{\partial G_i}\right) = \mathrm{sgn}\left(\frac{\mathrm{d}G_i^\star}{\mathrm{d}G_{-i}}\right),
$$

by which the desired assertion follows. ■

Proof of Lemma 5.3. Since the proof is quite similar to that of Lemma 4.1, we just provide the Lagrange function of the maximization problem. Replacing constraint (3) with constraint (29), the Lagrangian of Problem (11) amounts to

$$
\begin{aligned}
&\mathcal{L}_i\left(\{U_i(w), l_i(w)\}_{w\in[\underline{w},\overline{w}]}, G_i, \mu_i; \lambda_i, \gamma_i, \{\varsigma_i(w)\}_{w\in[\underline{w},\overline{w}]}\right) \\
=\ & U_i(\underline{w})+\lambda_i\int_{\underline{w}}^{\overline{w}}[wl_i(w)-\phi_i(U_i(w), l_i(w), \mu_i, G_i, G_{-i})]\,\tilde{f}_i(w)dw-\lambda_iG_i \\
&+\int_{\underline{w}}^{\overline{w}}\varsigma_i(w)\left[h'(l_i(w))\frac{l_i(w)}{w}-\dot{U}_i(w)\right]dw \\
&+\gamma_i\left[\mu_i-\phi_i(U_i(w_m), l_i(w_m), \mu_i, G_i, G_{-i})\right]f_i(w_m).
\end{aligned}
$$

Performing differentiation with respect to the choice variables and making use of (35), the optimal tax formula and the rule for optimal public goods provision can be accordingly obtained. ■

Proof of Proposition 5.3. Similar to the proof of Lemma 5.2, using Assumptions 4.1-4.2, (6), (31), and the L'Hôpital's rule yields:

$$
\lim_{w\uparrow\infty}\mathcal{C}_i(w) = 1-\frac{T_i'(y_i(\infty))}{1-T_i'(y_i(\infty))}\theta_i(\infty).
$$

Therefore, applying Assumption 4.2 to tax formula (30) yields the optimal AMTR.

Using (33) and differentiating the optimal AMTR with respect to $(1+\varepsilon)/a_i$ yields:

$$\frac{\partial \hat{T}_i'(y_i(\infty))}{\partial \left(\frac{1+\varepsilon}{a_i}\right)} = \frac{1}{\left[1 + \frac{1+\varepsilon}{a_i} + \frac{1+\varepsilon}{a_i} \cdot \theta_i(\infty)\right]^2} > 0,$$

as desired. ■