Received: 6 November 2018 | Accepted: 20 October 2019

DOI: 10.1111/jpet.12412

ORIGINAL ARTICLE

JOURNAL OF PUBLIC ECONOMIC THEORY WILEY

Optimal interregional redistribution and local borrowing rules under migration and asymmetric information

Darong Dai[1] | **Liqun Liu**[2] | **Guoqiang Tian**[3,4]

[1]Institute for Advanced Research (IAR), Shanghai University of Finance and Economics, Shanghai, China

[2]Private Enterprise Research Center, Texas A&M University, College Station, Texas

[3]Department of Economics, Texas A&M University, College Station, Texas

[4]Institute for Advanced Studies in Finance and Economics, Hubei University of Economics, Wuhan, China

Correspondence
Darong Dai, Institute for Advanced Research (IAR), Shanghai University of Finance and Economics, 200433 Shanghai, China.
Email: dai.darong@mail.shufe.edu.cn; darongdai@hotmail.com

Abstract

Assuming two types of regions that differ only in the discount rate, Huber and Runkel show that optimal federal redistribution is from impatient to patient regions, and optimal local public debt is higher in impatient regions than that in patient regions. This paper extends their analysis by allowing for interregional migrations and by considering two alternative regional goals. When the regional governments maximize their respective residents' welfare, considering the interregional migrations does not change Huber and Runkel's analysis. When the regional governments maximize their respective natives' welfare, incorporating migrations would reverse Huber and Runkel's conclusions when migration intensity is sufficiently high and the regional difference is sufficiently large.

1 | INTRODUCTION

Interregional redistribution is implemented in developing countries, such as China, as an important policy tool for fiscal equalization (e.g., Li, 2018) that helps achieve balanced development (e.g., The Economist, 2016), and is also of policy relevance in developed economies, such as Canada, France, the UK, and the US (see, Mélitz & Zumer, 2002). The early literature either focuses on the cross-region spillover effects due to interregional migrations (e.g., Breuillé & Gary-Bobo, 2007; Brown & Oates, 1987; Manasse & Schultz, 1999; Wildasin, 1991) or focuses on the information asymmetries between the federal government and local governments (e.g., Bordignon, Manasse, & Tabellini, 2001; Cornes & Silva, 2002; Cremer, Marchand, & Pestieau, 1996; Cremer & Pestieau, 1997; Dai, Liu, & Tian, 2019; Huber & Runkel, 2008; Oates, 1999), but not much is known about their *joint effect* on the optimal mechanism to redistribute resources among regional governments. This paper represents an attempt to fill this gap and focuses on pure redistribution among two regional governments in the presence of both asymmetric information between the center and regions and the interregional migrations.

To this end, we extend Huber and Runkel's (2008) two-period model of a federation consisting of a benevolent federal government and two regions that differ in the rate of time preference—which is the private information of each region—by allowing individuals of each region to migrate to the other region with a given probability and by entertaining each of the two plausible regional goals in the presence of migration: Maximizing the welfare of the natives of a region or maximizing the welfare of the residents of a region. We find that considering the interregional migrations does not change Huber and Runkel's analysis when the regional governments maximize their respective residents' welfare. When the regional governments maximize their respective natives' welfare, on the other hand, incorporating migrations could reverse Huber and Runkel's conclusions as summarized below.

First, in the first-best optimum under full information, the prediction of Huber and Runkel (2008), namely that the impatient region should borrow more than the patient region and the federal government should redistribute from the impatient region to the patient region, holds only if the intensity of mutual migration is low; otherwise, the patient region should borrow more than the impatient region and the redistribution should be from the patient region to the impatient region.

Second, in the second-best optimum under asymmetric information, their prediction that the impatient region should borrow more than the patient region and the redistribution should be from the former to the latter holds only if either migration intensity is low or migration intensity is high and the regional difference in discounting is small; otherwise, the patient region should borrow more than the impatient region and the redistribution should be from the patient region to the impatient region.

Third, although it is more likely to be the case that the second-best optimum achieves less interregional redistribution than does the first-best optimum, we also identify conditions under which the reverse holds true. That is, on the one hand, we extend the set of circumstances in which the prediction of Huber and Runkel, namely that information asymmetry limits the ability of the federal government to adopt a tax-transfer system to redistribute resources across heterogeneous regions, carries through. On the other hand, we identify reasonable circumstances in which this prediction is overturned.

As such, the optimal interregional redistribution policy obtained by Huber and Runkel remains optimal only in special cases of our more general model with interregional migrations. Importantly, we establish under reasonable circumstances an optimal redistribution policy that is exactly the opposite of theirs.

We choose the discount factor as the source of heterogeneity among regions due to these two considerations. First, as argued by Huber and Runkel and empirically demonstrated by Evans and Sezer (2004), the discount factor is indeed difficult to observe and is more likely to be the private information of local regions. Second, it is a key factor affecting intertemporal resource allocation, and hence is relevant in causing individual welfare disparity among regions.

The rest of the paper is organized as follows. Section 2 describes the model of the economy. In Section 3, we derive optimal interregional redistribution and local borrowing policies under mutual migration. Section 4 establishes the budget institutions arranged for local governments so that the full-information optimum can be implemented by decentralized debt decisions. Section 5 concludes. Proofs are relegated to the appendix.

2 | THE MODEL

We consider a federation consisting of a federal government (also referred to as the center) and two regions (also referred to as jurisdictions), denoted A and B, respectively. They have the same initial population size that is normalized to one for notational simplicity. These individuals live for two periods with identical income[1] y_1 in Period 1, identical income y_2 in Period 2, and a lifetime utility function

$$u_1(c_1) + g_1(G_1) + \delta^R[u_2(c_2) + g_2(G_2)], \tag{1}$$

in which $0 < \delta^R < 1$ denotes the discount factor of individuals born in region R ($R = A$ or B), c_1 and c_2 are, respectively, private consumptions in Periods 1 and 2, G_1 and G_2 are, respectively, the public goods[2] in Periods 1 and 2, and all four functions are strictly increasing and strictly concave, and also satisfy the usual Inada conditions.

Throughout, we impose the following assumption without loss of generality:

Assumption 2.1. $\delta^A < \delta^B$.

That is, individuals born in jurisdiction A are less patient than those born in jurisdiction B. We follow Huber and Runkel (2008) to use the discount factor as the only source of heterogeneity among regions for the following two reasons. First, as empirically demonstrated by Evans and Sezer (2004), the discount factor is difficult to observe and is likely to be the private information of local regions. Second, it is a key factor affecting intertemporal resource allocation, and hence is highly relevant for interregional welfare disparities. In what follows, we may simply call region A the impatient region and region B the patient region.

For expositional convenience, we focus on the decisions of the representative individual and the local government in region A, because those of the representative individual and the local government in region B are symmetric.

The local government in region A collects lump sum taxes τ_1^A and τ_2^A to finance local public good provision. In Period 1, the local government receives a transfer z^A from the center and issues debt b^A. Debt plus interest has to be repaid in Period 2, taking as given the common interest rate $r > 0$.[3] The fiscal budget constraints of local government A in Periods 1 and 2 can be written as $G_1^A = \tau_1^A + b^A + z^A$ and $G_2^A = \tau_2^A - (1 + r)b^A$, respectively. If the transfer from the center is negative, then it means that the local government has to pay a tax to the center. Under pure redistribution, the budget constraint of the federal government is

$$z^A + z^B = 0, \tag{2}$$

which means that the center collects resources from one region to finance transfers to the other region.

[1]Since the optimal redistribution issue driven by regional income inequality has been well addressed in the literature, the current setting that assumes away interregional income disparity is helpful for us to focus on the primary concern. Notwithstanding, in the case of asymmetric information we admit that the multidimensional screening problem induced by discounting heterogeneity and income disparity would be of independent interest. We leave it for our future research.

[2]The formal results in Sections 3 and 4 hold regardless of whether they are interpreted as publicly provided public goods or publicly provided private goods.

[3]It seems reasonable to assume that there is a common capital market within a federation. So, there is a single price level of capital to eliminate arbitrage opportunities.

The representative individual in region A has private consumption $c_1^A = y_1 - \tau_1^A$ in Period 1. His private consumption in Period 2 depends on whether he stays in region A or migrates to region B. At the beginning of Period 2, the representative individual in region A migrates to region B with the probability $p \in [0, 1)$ due to exogenous reasons, such as schooling, marriage, social network, or geographic preference. By symmetry, the number of individuals in each region will not change after the process of mutual migration.[4] If he stays in region A in Period 2, he will pay the average tax in region A (τ_2^A) and have Period-2 private consumption $c_2^A = y_2 - \tau_2^A$. If he migrates to region B, on the other hand, he will pay the average tax in region B (τ_2^B) and have Period-2 private consumption $c_2^B = y_2 - \tau_2^B$. The problem of the representative individual in region A is choosing τ_1^A and τ_2^A (or equivalently choosing c_1^A and c_2^A) to maximize his expected utility, which, after incorporating both regional governments' budget constraints, can be written as

$$\max_{\{c_1^A, c_2^A\}} u_1\left(c_1^A\right) + g_1\left(y_1 + b^A + z^A - c_1^A\right) + \delta^A\Big[(1-p)u_2\left(c_2^A\right) \\ + (1-p)g_2\left(y_2 - b^A(1+r) - c_2^A\right) + pu_2\left(c_2^B\right) + pg_2\left(y_2 - b^B(1+r) - c_2^B\right)\Big],$$

taking as given b^A, z^A, b^B, and τ_2^B (or equivalently c_2^B). A higher value of p, namely a higher migration probability, implies that the representative individual imposes a larger weight on the Period-2 utility he obtains from migrating to the other region and a smaller weight on that he obtains from staying.

The first-order conditions are thus written as

$$u_1'\left(c_1^A\right) = g_1'\left(G_1^A\right) \quad \text{and} \quad u_2'\left(c_2^A\right) = g_2'\left(G_2^A\right). \tag{3}$$

By symmetry, $u_1'(c_1^B) = g_1'(G_1^B)$ and $u_2'(c_2^B) = g_2'(G_2^B)$ are the first-order conditions from solving the corresponding maximization problem of the representative individual in region B. We denote by $\hat{c}_1^A$, $\hat{c}_2^A$, $\hat{c}_1^B$, and $\hat{c}_2^B$ the corresponding solutions.

Now we are ready to give the value function of region A. As well documented by Cremer and Pestieau (2004), measuring social welfare has always been a controversial issue when labor is mobile. We shall consider two cases. First, if we assume that only the welfare of natives (individuals born in this region) matters,[5] then region A's value function is exactly the maximum expected utility of its representative native resident, that is,

$$V(b^A, z^A, \delta^A; b^B) \equiv u_1\left(\hat{c}_1^A\right) + g_1\left(y_1 + b^A + z^A - \hat{c}_1^A\right) \\ + \delta^A\Big[(1-p)u_2\left(\hat{c}_2^A\right) + (1-p)g_2\left(y_2 - b^A(1+r) - \hat{c}_2^A\right) \\ + pu_2\left(\hat{c}_2^B\right) + pg_2\left(y_2 - b^B(1+r) - \hat{c}_2^B\right)\Big]. \tag{4}$$

[4]Another interesting problem worth considering would be one in which the migration probabilities between the two regions are not symmetric—say, more individuals would move from region A to region B than from B to A in the second period. Given that this would affect also the local government's second-period budget constraints, how would the conclusions of the present analysis change? We thank an anonymous referee for encouraging us to offer this question up for potential future research.

[5]There are existing studies which assumed that only the citizens (or the natives) are taken into account in the social welfare function or in the electoral process. For example, Leite-Monteiro (1997) and the benchmark model of Cremer and Pestieau (2004) adopt a mobility-free criterion such that social welfare function involves only the natives regardless of location.

This specification can be interpreted from the following political economy perspective. Assume voting occurs at the beginning of Period 1. As native residents are assumed to be homogeneous, the representative individual is actually the median voter who wins a majority voting contest under direct democracy. A selfish median voter would just maximize his/her own expected lifetime utility when making policy decisions. In this sense, this perspective can be elaborated as following the selfishly optimal policy design suggested by Röell (2012), Bohn and Stuart (2013), Brett and Weymark (2016, 2017), and Dai and Tian (2019) in the study of majority voting over income tax schedules.

Second, if we assume that local government A is benevolent and maximizes the aggregate welfare of its residents, then region A's value function is written as[6]

$$
\begin{aligned}
V(b^A, z^A, \delta^A; b^B) \equiv\; & u_1\left(\hat{c}_1^A\right) + g_1\left(y_1 + b^A + z^A - \hat{c}_1^A\right) \\
& + \delta^A \Bigg[(1-p)u_2\left(\hat{c}_2^A\right) + (1-p)g_2\left(y_2 - b^A(1+r) - \hat{c}_2^A\right) \\
& + \underbrace{pu_2\left(\hat{c}_2^A\right) + pg_2\left(y_2 - b^A(1+r) - \hat{c}_2^A\right)}_{\text{welfare of immigrants}} \Bigg].
\end{aligned}
$$

Note that by simple rearrangement this value function no longer contains the migration probability p. As such, incorporating interregional migrations will not change the results derived by Huber and Runkel (2008) under this specification of regional governments' goals.[7] The rest of the paper shall focus on the potentially more interesting situation in which the regional governments maximize the welfare of their respective natives.

3 | OPTIMAL INTERREGIONAL REDISTRIBUTION WITH MUTUAL MIGRATION

We assume that the federal government cannot observe the degree of patience of each region. Applying the revelation principle, the center offers each local government a contract that stipulates the federal transfer and a region's debt in each of the two possible situations depending on the region's reported type. More specifically, we assume

- These two local governments privately observe their own δ^R.
- The federal government offers the contract (b^R, z^R) for any $R \in \{A, B\}$ (or, equivalently, the contract $(b(\delta), z(\delta))$ for any $\delta \in \{\delta^A, \delta^B\}$).
- These two local governments simultaneously pick a contract (or equivalently report their types), and the game ends.

[6]We thank a referee for pointing out such a connection between our model and the model of Huber and Runkel (2008).

[7]Note that we use the discount factor δ^A (instead of δ^B) for the Period-2 utilities of the residents migrated from region B because in the asymmetric-information world considered here, it is reasonable to assume that regional government A, consisting of natives of region A, uses δ^A as the discount factor for all Period-2 utilities, regardless of whom they are derived for.

Formally, the center solves the maximization problem[8]

$$\max_{\{b^A, z^A, b^B, z^B\}} V(b^A, z^A, \delta^A; b^B) + V(b^B, z^B, \delta^B; b^A)$$

subject to budget constraint (2) and these incentive-compatibility constraints:

$$V(b^A, z^A, \delta^A; b^B) \geq V(b^B, z^B, \delta^A; b^B) \ (\mathrm{IC}_A);$$
$$V(b^B, z^B, \delta^B; b^A) \geq V(b^A, z^A, \delta^B; b^A) \ (\mathrm{IC}_B).$$

By using (3) and (4), it is easy to show that the corresponding single-crossing property is satisfied. The Lagrangian can thus be written as

$$\begin{aligned}\mathcal{L}(b^A, z^A, b^B, z^B; \mu_A, \mu_B, \lambda) = (1 + \mu_A)V(b^A, z^A, \delta^A; b^B) - \mu_A V(b^B, z^B, \delta^A; b^B)\\ +(1 + \mu_B)V(b^B, z^B, \delta^B; b^A)\\ -\mu_B V(b^A, z^A, \delta^B; b^A) - \lambda(z^A + z^B),\end{aligned} \tag{5}$$

in which μ_A, μ_B, and λ are Lagrangian multipliers. Without loss of generality, we let the federal budget constraint (2) be binding so that $\lambda > 0$.

Let the regional value function, V, be defined by (4). As a standard benchmark, we consider first the case with symmetric information between the center and local governments, and hence $\mu_A = \mu_B = 0$. We index the first-best allocation by the superscript $^{\mathrm{FB}}$.

Proposition 3.1. *Let regional optimality be defined over the welfare of natives; then the following statements are true under Assumption 2.1 and symmetric information:*

(i) $G_1^{A,\mathrm{FB}} = G_1^{B,\mathrm{FB}}$ *and* $c_1^{A,\mathrm{FB}} = c_1^{B,\mathrm{FB}}$ *for* $\forall p \in [0, 1]$.
(ii) If $p < 1/2$, *then* $G_2^{A,\mathrm{FB}} < G_2^{B,\mathrm{FB}}$, $c_2^{A,\mathrm{FB}} < c_2^{B,\mathrm{FB}}$, $b^{A,\mathrm{FB}} > b^{B,\mathrm{FB}}$, $z^{A,\mathrm{FB}} < 0 < z^{B,\mathrm{FB}}$, IC_A *is violated, and* IC_B *is satisfied.*
(iii) If $p > 1/2$, *then* $G_2^{A,\mathrm{FB}} > G_2^{B,\mathrm{FB}}$, $c_2^{A,\mathrm{FB}} > c_2^{B,\mathrm{FB}}$, $b^{A,\mathrm{FB}} < b^{B,\mathrm{FB}}$, $z^{B,\mathrm{FB}} < 0 < z^{A,\mathrm{FB}}$, IC_A *is satisfied, and* IC_B *is violated.*
(iv) If $p = 1/2$, *then* $G_2^{A,\mathrm{FB}} = G_2^{B,\mathrm{FB}}$, $c_2^{A,\mathrm{FB}} = c_2^{B,\mathrm{FB}}$, $b^{A,\mathrm{FB}} = b^{B,\mathrm{FB}}$, $z^{A,\mathrm{FB}} = z^{B,\mathrm{FB}} = 0$, *and both* IC_A *and* IC_B *are satisfied.*

Proof. See the appendix. □

In the first-best allocation, Period-1 private and public consumptions are the same for both regions. If the intensity of mutual migration is low, namely $p < 1/2$, Period-2 private and public consumptions are higher in the patient region than those in the impatient region, the impatient region borrows more than does the patient region, and interregional redistribution is from the impatient region to the patient region. These characteristics are exactly the same as those found by Huber and Runkel (2008) in a federation without interregional migrations. As a result, the first-best allocation as well as the first-best interregional redistribution policy obtained by

[8]Following the common practice in the related literature, participation constraints are ignored. In practice, it may be politically and/or economically costly for a region to leave the federation, and hence it is not quite restrictive to ignore these constraints. In fact, exploring the breakup possibility of a nation is of independent interest (e.g., Bolton & Roland, 1997).

Huber and Runkel could be robust with respect to the introduction of interjurisdictional migrations, given that the intensity of mutual migration is not too high. We explain in more detail the driving force of this result as follows.

Under the symmetric-information case considered by Huber and Runkel, the benevolent center just offers the contract such that the intertemporal rate of substitution between Period-1 public good consumption and Period-2 public good consumption in each region is equal to their respective intertemporal rates of transformation, denoted by $\delta^R(1+r)$ for any $R \in \{A, B\}$, which are jointly determined the common interest rate and their respective discount rates. So, $\delta^A(1+r) < \delta^B(1+r)$ under Assumption 2.1 implies that the intertemporal rate of substitution is smaller in the impatient region than in the patient region, by which their first-best allocation as well as first-best interregional redistribution policy follows. Under the symmetric-information case considered here, a difference arises from the fact that the benevolent center needs to internalize the cross-region externality caused by interregional migrations. Now the intertemporal rates of substitution between Period-1 public good consumption and Period-2 public good consumption concerning the impatient region A and the patient region B are, respectively, given by $[\delta^A(1-p) + \delta^B p](1+r)$ and $[\delta^B(1-p) + \delta^A p](1+r)$. That is, the optimal intertemporal allocation must take into account the given probability p of between-region mutual migrations at the end of Period 1 (or at the beginning of Period 2). If $p < 1/2$, then we have $\delta^A(1-p) + \delta^B p < \delta^B(1-p) + \delta^A p$ under Assumption 2.1, yielding that the intertemporal rate of substitution is smaller in the impatient region than in the patient region, the same as what has been obtained by Huber and Runkel in terms of cross-region comparison.

However, as shown in (iii), if the intensity of mutual migration is high, namely $p > 1/2$, Period-2 private and public consumptions are lower in the patient region than those in the impatient region, the patient region borrows more than does the impatient region, and interregional redistribution is from the patient region to the impatient region. This novel redistribution policy hence *overturns* that obtained by Huber and Runkel. The reason is that we now have $\delta^A(1-p) + \delta^B p > \delta^B(1-p) + \delta^A p$ under Assumption 2.1, yielding that the intertemporal rate of substitution is larger in the impatient region than in the patient region, hence reversing what has been obtained by Huber and Runkel in terms of cross-region comparison. Finally, if $p = 1/2$, then $\delta^A(1-p) + \delta^B p = \delta^B(1-p) + \delta^A p$, namely the intertemporal rates of substitution of the patient region and the impatient region are the same in the first-best allocation, and hence the policy of interregional redistribution should no longer be used.

Under asymmetric information, the discount factor is private information so that the local government of a given region can mimic the local government of the other region to obtain transfers. We now index the second-best allocation by the superscript *.

Proposition 3.2. *Let regional optimality be defined over the welfare of natives; then the following statements are true under Assumption 2.1 and asymmetric information:*

(i) If $\mu_A > \mu_B \geq 0$, then (i-a) truth-telling requires that either $p \leq p^*_{A/B}$, or $p > p^*_{A/B}$ and $\delta^A/\delta^B > \delta^*$, for thresholds $p^*_{A/B} \in (1/2, 1)$ and $\delta^* \in (0, 1)$, (i-b) the second-best optimum satisfies $G_1^{A*} > G_1^{B*}$, $c_1^{A*} > c_1^{B*}$, $G_2^{A*} < G_2^{B*}$, $c_2^{A*} < c_2^{B*}$, and $b^{A*} > b^{B*}$, and (i-c) optimal interregional redistribution exhibits $z^{A*} < 0 < z^{B*}$ for $p \geq \frac{\mu_B(\delta^B - \delta^A)}{\delta^B + \mu_B(\delta^B - \delta^A)}$.

(ii) If $\mu_B > \mu_A \geq 0$, then (ii-a) truth-telling requires that either $p \geq p^*_{A/B}$, or $p < p^*_{A/B}$ and $\delta^A/\delta^B > \delta^*$, and (ii-b) the second-best optimum satisfies $G_1^{A*} < G_1^{B*}$, $c_1^{A*} < c_1^{B*}$, $G_2^{A*} > G_2^{B*}$, $c_2^{A*} > c_2^{B*}$, $b^{A*} < b^{B*}$, and $z^{B*} < 0 < z^{A*}$.

(iii) *If* $\mu_A = \mu_B > 0$, *then (iii-a) truth-telling requires that* $p = p^*_{A/A}$ *for threshold* $p^*_{A/A} \in (1/2, 1)$, *and (iii-b) the second-best optimum satisfies* $G_1^{A*} = G_1^{B*}$, $c_1^{A*} = c_1^{B*}$, $G_2^{A*} = G_2^{B*}$, $c_2^{A*} = c_2^{B*}$, $b^{A*} = b^{B*}$, *and* $z^{A*} = z^{B*} = 0$.

Proof. See the appendix. □

The second-best optimum derived by Huber and Runkel (2008) is just a special case of our Part (i) via imposing that $\mu_B = 0$ and $p = 0$. In particular, $\mu_B = 0$ implies that only the incentive-compatibility constraint of the impatient region A is binding, which turns out to be true only when $p = 0$. In other words, the introduction of a positive probability of between-region migrations prevents this clear-cut result from happening, and hence we need to consider three possible cases as shown in Proposition 3.2.

Part (i) of Proposition 3.2 studies the case in which the shadow price (μ_A) of incentive-compatibility constraint concerning the impatient region A is higher than that (μ_B) concerning the patient region B. As such, the impatient region obtains an information rent as their Period-1 private and public consumptions are higher than those of the patient region. Although the impatient region borrows more than does the patient region, the center redistributes from the former to the latter. The self-selection problem is resolved when either the intensity of mutual migration is not too high, or if it is high, the difference of the degree of patience between these two regions is not too large. In fact, all these conditions help guarantee that the intertemporal rate of substitution between Period-1 public consumption and Period-2 public consumption be smaller in the impatient region than that in the patient region under asymmetric information, a requirement follows from our assumption of $\mu_A > \mu_B \geq 0$. Roughly, through distorting along the channel of intertemporal rate of substitution, these conditions help separate the incentives facing these two types of regions under asymmetric information.

Part (ii) of Proposition 3.2 studies the case in which the shadow price (μ_A) of incentive-compatibility constraint concerning the impatient region is lower than that (μ_B) concerning the patient region. As such, the patient region obtains an information rent as their Period-1 private and public consumptions are higher than those of the impatient region. Although the patient region borrows more than does the impatient region, the center redistributes from the former to the latter. The self-selection problem is resolved when either the intensity of mutual migration is high, or if it is not high, the difference of the degree of patience between these two regions is not too large.

Part (iii) of Proposition 3.2 studies the case in which the shadow price (μ_A) of incentive-compatibility constraint concerning the impatient region is the same as that (μ_B) concerning the patient region. The self-selection problem is resolved when the migration probability is equal to a specific value larger than 1/2. We find in this special context that these two heterogeneous regions should be treated in the same way, and hence the federal policy of interregional redistribution should not be used at all.

To identify the effect of the asymmetric information between the center and local governments on optimal debt and interregional redistribution policies, we give the following:

Proposition 3.3. *Let* λ^{FB} *and* λ^* *denote the shadow values of federal funds in the first-best and the asymmetric-information optimum, respectively. Under Assumption 2.1, the following statements are true when regional optimality is defined over the welfare of natives:*

(i) *Suppose* $\mu_A > \mu_B$. *(i-a) If* $\delta^A/\delta^B > \mu_B/\mu_A$ *and* $\lambda^{\mathrm{FB}} = \lambda^*$, *then* $b^{B,\mathrm{FB}} < b^{B*} < b^{A*} < b^{A,\mathrm{FB}}$ *and* $z^{A,\mathrm{FB}} < z^{A*} < 0 < z^{B*} < z^{B,\mathrm{FB}}$. *(i-b) If* $p < 1/2$, $\delta^A/\delta^B \leq \mu_B/\mu_A$, *and either*

$\lambda^*/\lambda^{\mathrm{FB}} \geq 1 + \mu_A - \mu_B$ *or* $\lambda^*/\lambda^{\mathrm{FB}} \leq 1 + \mu_B - \mu_A$ *is fulfilled, then* $b^{B*} < b^{B,\mathrm{FB}} < b^{A,\mathrm{FB}} < b^{A*}$ *and* $z^{A*} < z^{A,\mathrm{FB}} < 0 < z^{B,\mathrm{FB}} < z^{B*}$.

(ii) *Suppose* $\mu_B > \mu_A$. *If either* $\lambda^{\mathrm{FB}} = \lambda^*$, *or* $\lambda^*/\lambda^{\mathrm{FB}} \geq 1 + \mu_B - \mu_A$, *or* $\lambda^*/\lambda^{\mathrm{FB}} \leq 1 + \mu_A - \mu_B$ *is fulfilled, then* $b^{A,\mathrm{FB}} < b^{A*} < b^{B*} < b^{B,\mathrm{FB}}$ *and* $z^{B,\mathrm{FB}} < z^{B*} < 0 < z^{A*} < z^{A,\mathrm{FB}}$.

Proof. See the appendix. □

Parts (i-a) and (ii) show that the second-best allocation achieves less interregional redistribution than does the first-best allocation. That is, under these conditions given in Parts (i-a) and (ii), information asymmetry does limit the ability of the center to redistribute resources across different types of regions (or local governments). In terms of the optimal debt policy for these local governments, the contributor region borrows less in the second-best allocation than it does in the first-best allocation, whereas the recipient region borrows more in the second-best allocation than it does in the first-best allocation. The reason is that the contributor pays less tax to and the recipient receives less transfer from the center under asymmetric information than they do under symmetric information.

Part (i-b), however, confirms the possibility that the second-best allocation achieves more interregional redistribution than that of the first-best allocation. The sufficient conditions that guarantee this prediction are the following: (a) The shadow price of truth-telling constraint of the impatient region A is greater than that of the patient region B, that is, the impatient region is more likely to misreport its type; (b) mutual migration probability is below 50%, that is, regions face a small migration intensity; (c) the degree of impatience (or patience) of region A (or B) is sufficiently large, that is, regional difference in discounting is large; and (d) the shadow value of federal funds in the asymmetric-information optimum is sufficiently either smaller or larger than that in the first-best, which can be interpreted as informational asymmetry either weakening or strengthening the redistributive motive of the center.[9] Although this prediction follows from the joint effect of these four conditions, we conjecture condition (d) plays a determinant role. Also, in terms of the optimal debt policy for these local governments, the contributor region borrows more in the second-best allocation than it does in the first-best allocation, whereas the recipient region borrows less in the second-best allocation than it does in the first-best allocation. The reason is that the contributor pays more tax to and the recipient receives more transfer from the center under asymmetric information than they do under symmetric information.

4 | JUSTIFYING DIFFERENTIATED BUDGET INSTITUTIONS

A main goal of the theoretical model of Huber and Runkel (2008) is to justify differentiated budget institutions in a federation, and they show that recipient regions of a federal transfer/redistribution scheme face more stringent borrowing constraints than contributing regions. We now consider the case with regionally decentralized debt decisions. That is, both regions choose a level of public debt to maximize their regional welfare, taking as given the

[9]We thank a referee for pointing out that the shadow value of federal funds in the full-information optimum is not necessarily the same as that in the asymmetric-information optimum.

redistribution scheme of the federal government. Formally, the maximization problem of region A is

$$\max_{b^A} V\left(b^A, z, \delta^A; b^B\right)$$

for any given z. If regional optimality is defined over the welfare of natives, then the first-order condition reads as

$$g_1'\left(y_1 + b^A + z - c_1^A\right) = \delta^A(1-p)(1+r)g_2'\left(y_2 - (1+r)b^A - c_2^A\right), \tag{6}$$

which characterizes the debt level in region A the representative native in this region would like to choose without any restrictions on debt issuance.

In contrast, the full-information optimal debt level in region A is characterized by Equation (A1) in the proof of Proposition 3.1:

$$g_1'\left(G_1^A\right) = [\delta^A(1-p) + \delta^B p](1+r)g_2'\left(G_2^A\right).$$

Comparing Equation (6) with this equation, it is straightforward to see that, due to the $\delta^B p$ term in Equation (A1), the full-information optimal debt level in region A is smaller than the debt level in region A that is chosen, absent any restrictions on debt issuance, under decentralized debt decisions. We can therefore establish the following result regarding the implementation of the full-information optimum through decentralized regional debt decisions.[10]

Proposition 4.1. *Suppose that regional welfare is specified to be the welfare of natives. The full-information optimum is attained through decentralized regional debt decisions by setting $z^A = z^{A,\text{FB}}$, $z^B = z^{B,\text{FB}}$, an upper bound $b^{A,\text{FB}}$ on the public debt in region A, and an upper bound $b^{B,\text{FB}}$ on the public debt in region B.*

Proof. See the appendix. □

In the model of Huber and Runkel (2008), the full-information optimum is attained through decentralized regional debt decisions by simply setting $z^A = z^{A,\text{FB}}$ and $z^B = z^{B,\text{FB}}$. In our more general model with interregional migrations, however, an upper bound must be imposed on each region's debt issuance to implement the full-information optimum.[11] Without such restrictions, the representative native citizen in each region ignores the fact that debt financing in the region will also place a burden on the migrants from the other region, thereby having a tendency to borrow too much from the point of view of the entire federation. In addition, Proposition 4.1 justifies under complete information the existence of differentiated budget institutions with different debt ceilings.

[10]We thank a referee for pointing out an error in the original version of this proposition.

[11]In a different setting with interjurisdictional migrations, Dai, Jansen, and Liu (2019) also present a theoretical argument in favor of imposing debt limits on local governments. Indeed, the use of debt limits on subnational government borrowing is of practical relevance, for example, such a local borrowing rule is enforced in China (see, Huang, Ning, & Tian, 2018).

5 | CONCLUDING SUMMARY

We enrich the two-region, two-period model of Huber and Runkel (2008) by allowing for between-jurisdiction mutual migrations to study optimal interregional redistribution and local borrowing policies under a benevolent federal government. Following Huber and Runkel, these two regions are assumed to differ only in the discount factor for Period-2 utility (or utility when old). The federal government cannot observe the value of each region's discount factor, and hence the second-best interregional redistribution scheme is established by solving a mechanism-design problem.

In the presence of migration, the goal of a regional government could be maximizing the welfare of either the region's residents or the region's natives. When the regional governments maximize their respective residents' welfare, considering the interregional migrations does not change Huber and Runkel's analysis and conclusions. Hence our analysis focuses on the more interesting situation where the regional governments maximize their respective natives' welfare.

The main results are summarized as follows. First, in the benchmark of the first-best optimum, Huber and Runkel's results regarding the relative optimal debt level in the two regions and the direction of interregional redistribution continue to hold when migration intensity is low, but are reversed when migration intensity is high. The intuition behind this reversal is that for sufficiently high migration intensity, the two regions' populations are essentially swapped over time, but the legal persons inheriting any accumulated debt (i.e., regional governments) still remain the same.[12] From a practical perspective, this can also be interpreted as the two polities exchanging their levels of indebtedness, in so far as migration changes the identity of the people who are de facto responsible for repayment. A welfare-maximizing federal government takes this debt-swapping phenomenon into account when deciding on the optimal level of interregional taxes, subsidies, and debt ceilings.

Second, in the second-best optimum, their results continue to hold when either migration intensity is low or migration intensity is high and the regional difference in discounting is small; otherwise, their results are reversed. Third, although we extend the set of circumstances in which their result that the optimal level of redistribution is smaller in the second-best optimum than that in the first-best optimum continues to hold in the presence of mutual migration, we also identify reasonable circumstances in which such a prediction is reversed. Fourth, in the course of implementing the full-information optimum via decentralized leadership of local borrowing, we justify the existence of differentiated budget institutions with different debt ceilings.

ACKNOWLEDGMENTS

We are grateful to the editor, Helmuth Cremer, the anonymous Associate Editor, and two anonymous referees for excellent comments and suggestions that lead to substantial improvement. All remaining errors are our own.

ORCID

Darong Dai http://orcid.org/0000-0002-3997-2574
Liqun Liu http://orcid.org/0000-0003-4506-1272

[12]We thank a referee for pointing out this intuition.

REFERENCES

Bohn, H., & Stuart, C. (2013). *Revenue extraction by median voters* (Unpublished Manuscript). Department of Economics, University of California, Santa Barbara, CA.

Bolton, P., & Roland, G. (1997). The breakup of nations: A political economy analysis. *Quarterly Journal of Economics*, *112*(4), 1057–1090.

Bordignon, M., Manasse, P., & Tabellini, G. (2001). Optimal regional redistribution under asymmetric information. *American Economic Review*, *91*(3), 709–723.

Brett, C., & Weymark, J. A. (2016). Voting over selfishly optimal nonlinear income tax schedules with a minimum-utility constraint. *Journal of Mathematical Economics*, *67*, 18–31.

Brett, C., & Weymark, J. A. (2017). Voting over selfishly optimal nonlinear income tax schedules. *Games and Economic Behavior*, *101*, 172–188.

Breuillé, M.-L., & Gary-Bobo, R. J. (2007). Sharing budgetary austerity under free mobility and asymmetric information: An optimal regulation approach to fiscal federalism. *Journal of Public Economics*, *91*(5–6), 1177–1196.

Brown, C. C., & Oates, W. E. (1987). Assistance to the poor in a federal system. *Journal of Public Economics*, *32*(3), 307–330.

Cornes, R. C., & Silva, E. C. D. (2002). Local public goods, interregional transfers and private information. *European Economic Review*, *46*(2), 329–356.

Cremer, H., & Pestieau, P. (2004). Factor mobility and redistribution. In J. V. Henderson, & J. F. Thisse (Eds.), *Handbook of regional and urban economics* (Vol. 4, pp. 2529–2560).

Cremer, H., & Pestieau, P. (1997). Income redistribution in an economic union: The trade off between inter- and intra-national redistribution. *International Tax and Public Finance*, *4*(3), 325–335.

Cremer, H., Marchand, M., & Pestieau, P. (1996). Interregional redistribution through tax surcharge. *International Tax and Public Finance*, *3*(2), 157–173.

Dai, D., Jansen, D. W., & Liu, L. (2019). *Inter-jurisdiction migration and the fiscal policies of local governments* (Working Paper). SUFE and TAMU.

Dai, D., Liu, L., & Tian, G. (2019). Interregional redistribution and budget institutions with private information on intergenerational externality. *Review of Economic Design*, *23*(3), 127–154.

Dai, D., & Tian, G. (2019). *Voting over selfishly optimal income tax schedules with tax-driven migrations* (Working Paper). SUFE and TAMU.

Evans, D. J., & Sezer, H. (2004). Social discount rates for six major countries. *Applied Economics Letters*, *11*(9), 557–560.

Huang, K. X. D., Ning, L., & Tian, G. (2018). Conquering China's unbalanced and inadequate development: Macroeconomic outlook, policy simulations, and reform implementation—A summary of annual SUFE macroeconomic report (2017–2018). *Frontiers of Economics in China*, *13*(2), 147–170.

Huber, B., & Runkel, M. (2008). Interregional redistribution and budget institutions under asymmetric information. *Journal of Public Economics*, *92*(12), 2350–2361.

Leite-Monteiro, M. (1997). Redistributive policy with labour mobility across countries. *Journal of Public Economics*, *65*(2), 229–244.

Li, Y. (2018). Asymmetric decentralization, intergovernmental transfers, and expenditure policies of local governments. *Frontiers of Economics in China*, *13*(2), 223–248.

Manasse, P., & Schultz, C. (1999). *Regional redistribution and migration* (IGIER Working Paper No. 146). Bocconi University.

Mélitz, J., & Zumer, F. (2002). Regional redistribution and stabilization by the center in Canada, France, the UK and the US: A reassessment and new tests. *Journal of Public Economics*, *86*(2), 263–286.

Oates, W. E. (1999). An essay on fiscal federalism. *Journal of Economic Literature*, *37*(3), 1120–1149.

Röell, A. A. (2012). *Voting over nonlinear income tax schedules* (Unpublished Manuscript). School of International and Public Affairs, Columbia University.

The Economist. (2016). Regional development—Rich province, poor province: The government is struggling to spread wealth more evenly. *The Economist*, October 1, 2016, Beijing.

Wildasin, D. E. (1991). Income redistribution in a common labor market. *American Economic Review*, *81*(4), 757–774.

How to cite this article: Dai D, Liu L, Tian G. Optimal interregional redistribution and local borrowing rules under migration and asymmetric information. *J Public Econ Theory*. 2019;21:1266–1285. https://doi.org/10.1111/jpet.12412

APPENDIX A: PROOFS

Proof of Proposition 3.1. Using (4), (5), and Envelope Theorem, we have these first-order conditions (FOCs):

$$\begin{aligned} g_1'\left(G_1^A\right) &= [\delta^A(1-p)+\delta^B p](1+r)g_2'\left(G_2^A\right) \\ g_1'\left(G_1^A\right) &= \lambda. \end{aligned} \tag{A1}$$

By symmetry, we have another two FOCs:

$$\begin{aligned} g_1'\left(G_1^B\right) &= [\delta^B(1-p)+\delta^A p](1+r)g_2'\left(G_2^B\right) \\ g_1'\left(G_1^B\right) &= \lambda. \end{aligned} \tag{A2}$$

It is immediate by (A1) and (A2) that $G_1^{A,\mathrm{FB}} = G_1^{B,\mathrm{FB}}$. This combined with (3) yields the desired assertion (i). Note that

$$[\delta^A(1-p)+\delta^B p] - [\delta^B(1-p)+\delta^A p] = (1-2p)(\delta^A-\delta^B),$$

then $p < 1/2$ combined with Assumption 2.1, (A1), and (A2) implies that $G_2^{A,\mathrm{FB}} < G_2^{B,\mathrm{FB}}$. This combined with (3) yields $c_2^{A,\mathrm{FB}} < c_2^{B,\mathrm{FB}}$. It follows from Period-2 public budget constraint that $0 > G_2^A - G_2^B = (b^B - b^A)(1+r) + (c_2^B - c_2^A)$, which yields $(b^A - b^B)(1+r) > (c_2^B - c_2^A) > 0$, and hence $b^{A,\mathrm{FB}} > b^{B,\mathrm{FB}}$. This combined with Period-1 public budget constraint, assertion (i), and federal budget constraint (2) implies that $z^{A,\mathrm{FB}} < 0 < z^{B,\mathrm{FB}}$. Therefore, we get from (4) and assertion (i) that

$$\begin{aligned} & V(b^{A,\mathrm{FB}}, z^{A,\mathrm{FB}}, \delta^A; b^{B,\mathrm{FB}}) - V(b^{B,\mathrm{FB}}, z^{B,\mathrm{FB}}, \delta^A; b^{B,\mathrm{FB}}) \\ &= \delta^A(1-p)\left[u_2\left(c_2^{A,\mathrm{FB}}\right) + g_2\left(G_2^{A,\mathrm{FB}}\right) - u_2\left(c_2^{B,\mathrm{FB}}\right) - g_2\left(G_2^{B,\mathrm{FB}}\right)\right] \\ &< 0 \end{aligned}$$

and by symmetry,

$$\begin{aligned} & V(b^{B,\mathrm{FB}}, z^{B,\mathrm{FB}}, \delta^B; b^{A,\mathrm{FB}}) - V(b^{A,\mathrm{FB}}, z^{A,\mathrm{FB}}, \delta^B; b^{A,\mathrm{FB}}) \\ &= \delta^B(1-p)\left[u_2\left(c_2^{B,\mathrm{FB}}\right) + g_2\left(G_2^{B,\mathrm{FB}}\right) - u_2\left(c_2^{A,\mathrm{FB}}\right) - g_2\left(G_2^{A,\mathrm{FB}}\right)\right] \\ &> 0. \end{aligned}$$

The proof of assertion (ii) is thus complete. Since assertions (iii) and (iv) can be similarly proved, we thus omit them to economize on the space. □

Proof of Proposition 3.2. We shall complete the proof in six steps.

Step 1 Using (4), (5), and Envelope Theorem, we have these FOCs:

$$\begin{aligned} \frac{g_1'(G_1^A)}{(1+r)g_2'(G_2^A)} &= \frac{(1+\mu_A)(1-p)\delta^A + [(1+\mu_B)p - \mu_B]\delta^B}{1+\mu_A-\mu_B} \\ g_1'(G_1^A) &= \frac{\lambda}{1+\mu_A-\mu_B}. \end{aligned} \tag{A3}$$

By symmetry, we have another two FOCs:

$$\begin{aligned} \frac{g_1'(G_1^B)}{(1+r)g_2'(G_2^B)} &= \frac{(1+\mu_B)(1-p)\delta^B + [(1+\mu_A)p - \mu_A]\delta^A}{1+\mu_B-\mu_A} \\ g_1'(G_1^B) &= \frac{\lambda}{1+\mu_B-\mu_A}. \end{aligned} \tag{A4}$$

It follows from Proposition 3.1 that truth-telling generally requires that μ_A and μ_B cannot be zero at the same time. We thus just need to consider these three cases shown in Proposition 3.2.

Step 2 Consider first the case with $\mu_A > \mu_B \geq 0$, then it is immediate by (A3), (A4) and $\lambda > 0$ that $\mu_A - \mu_B < 1$ and $G_1^{A*} > G_1^{B*}$. Then, using (3) produces $c_1^{A*} > c_1^{B*}$. These results combined with (4) reveal that

$$\begin{aligned} 0 &= V(b^{A*}, z^{A*}, \delta^A; b^{B*}) - V(b^{B*}, z^{B*}, \delta^A; b^{B*}) \\ &= \underbrace{\left[u_1\left(c_1^{A*}\right) - u_1\left(c_1^{B*}\right)\right] + \left[g_1\left(G_1^{A*}\right) - g_1\left(G_1^{B*}\right)\right]}_{>0} \\ &\quad + \delta^A(1-p)\left[u_2\left(c_2^{A*}\right) + g_2\left(G_2^{A*}\right) - u_2\left(c_2^{B*}\right) - g_2\left(G_2^{B*}\right)\right]. \end{aligned} \tag{A5}$$

By symmetry, we have

$$\begin{aligned} 0 &\leq V(b^{B*}, z^{B*}, \delta^B; b^{A*}) - V(b^{A*}, z^{A*}, \delta^B; b^{A*}) \\ &= \underbrace{\left[u_1\left(c_1^{B*}\right) - u_1\left(c_1^{A*}\right)\right] + \left[g_1\left(G_1^{B*}\right) - g_1\left(G_1^{A*}\right)\right]}_{<0} \\ &\quad + \delta^B(1-p)\left[u_2\left(c_2^{B*}\right) + g_2\left(G_2^{B*}\right) - u_2\left(c_2^{A*}\right) - g_2\left(G_2^{A*}\right)\right]. \end{aligned} \tag{A6}$$

We thus get from (A5) and (A6) that $G_2^{A*} < G_2^{B*}$. By (3), we have $c_2^{A*} < c_2^{B*}$.

Step 3 In addition, we must have

$$\begin{aligned} &\frac{(1+\mu_A)(1-p)\delta^A + [(1+\mu_B)p - \mu_B]\delta^B}{1+\mu_A-\mu_B} \\ &\quad < \frac{(1+\mu_B)(1-p)\delta^B + [(1+\mu_A)p - \mu_A]\delta^A}{1+\mu_B-\mu_A} \end{aligned}$$

by (A3) and (A4). This inequality can be equivalently simplified as

$$[(1-2p)(1+\mu_A)+\mu_B]\delta^A < [(1-2p)(1+\mu_B)+\mu_A]\delta^B. \tag{A7}$$

We just need to figure out the conditions such that (A7) holds. Note that

$$(1-2p)(1+\mu_B)+\mu_A-[(1-2p)(1+\mu_A)+\mu_B] = 2p(\mu_A-\mu_B)>0$$

and

$$(1-2p)(1+\mu_A)+\mu_B \geq 0 \Leftrightarrow p \leq \frac{1}{2}+\frac{\mu_B}{2(1+\mu_A)} \equiv p^*_{B/A}, \tag{A8}$$

thus (A7) automatically holds true under (A8) and Assumption 2.1. If, otherwise, $p > p^*_{B/A}$, then using (A7) yields

$$\frac{\delta^A}{\delta^B} > \frac{1+\mu_A+\mu_B-2p(1+\mu_B)}{1+\mu_A+\mu_B-2p(1+\mu_A)} \equiv \delta^*.$$

Note that

$$1+\mu_A+\mu_B-2p(1+\mu_B) \geq 0 \Leftrightarrow p \leq \frac{1}{2}+\frac{\mu_A}{2(1+\mu_B)} \equiv p^*_{A/B}$$

and

$$p^*_{A/B} > p^*_{B/A},$$

so $\delta^* \leq 0$ and hence (A7) automatically holds true whenever $p \leq p^*_{A/B}$. If $p > p^*_{A/B}$, then $\delta^* \in (0,1)$ and (A7) holds true under another condition, namely $\delta^A/\delta^B > \delta^*$, as desired in Part (i) of this proposition. Also, it is easy to get $b^{A*} > b^{B*}$ from $0 > G_2^{A*} - G_2^{B*} = (b^{B*}-b^{A*})(1+r)+(c_2^{B*}-c_2^{A*})$.

Step 4 To complete the proof of Part (i), we need to characterize the second-best interregional redistribution scheme. First, using Envelope Theorem and (4), we get

$$\left.\frac{dz^A}{db^A}\right|_{dV=0} = \frac{\delta^A(1-p)(1+r)g_2'(G_2^A)-g_1'(G_1^A)}{g_1'(G_1^A)}.$$

And also, using (3) and Implicit Function Theorem gives that

$$\frac{\partial c_1^A}{\partial b^A} = \frac{\partial c_1^A}{\partial z^A} = \frac{g_1''(G_1^A)}{u_1''(c_1^A)+g_1''(G_1^A)} \quad \text{and} \\ \frac{\partial c_2^A}{\partial b^A} = -\frac{(1+r)g_2''(G_2^A)}{u_2''(c_2^A)+g_2''(G_2^A)}. \tag{A9}$$

We thus obtain by simplifying the algebra that

$$\left.\frac{d^2z^A}{d(b^A)^2}\right|_{dV=0} = -\frac{\delta^A(1-p)(1+r)^2 g_2''(G_2^A)u_2''(c_2^A)}{g_1'(G_1^A)[u_2''(c_2^A)+g_2''(G_2^A)]} - \frac{[\delta^A(1-p)(1+r)g_2'(G_2^A)]^2 g_1''(G_1^A)u_1''(c_1^A)}{[g_1'(G_1^A)]^3[u_1''(c_1^A)+g_1''(G_1^A)]} > 0, \tag{A10}$$

so the welfare indifference curve of region A is strictly convex and U-shaped in the (b, z)-space, and the minimum is achieved at $\delta^A(1-p)(1+r)g'(G_2^A) = g'(G_1^A)$. Second, it follows from (A5) that (z^{A*}, b^{A*}) and (z^{B*}, b^{B*}) lie on the same welfare indifference curve of region A. Third, noting from (A3) that

$$\begin{aligned}
&\frac{(1+\mu_A)(1-p)\delta^A + [(1+\mu_B)p - \mu_B]\delta^B}{1+\mu_A-\mu_B} - (1-p)\delta^A \\
&= \frac{\mu_B(1-p)(\delta^A-\delta^B) + p\delta^B}{1+\mu_A-\mu_B} \geq 0 \\
&\Leftrightarrow p \geq \frac{\mu_B(\delta^B-\delta^A)}{\delta^B+\mu_B(\delta^B-\delta^A)},
\end{aligned} \tag{A11}$$

so (A10) implies that both (z^{A*}, b^{A*}) and (z^{B*}, b^{B*}) lie on the decreasing part of the welfare indifference curve of region A whenever $p \geq \frac{\mu_B(\delta^B-\delta^A)}{\delta^B+\mu_B(\delta^B-\delta^A)}$. We, therefore, obtain $z^{A*} < 0 < z^{B*}$ by using $b^{A*} > b^{B*}$.

Step 5 We now establish the thresholds required in Part (ii). As before, we need to figure out conditions under which the following inequality holds:

$$[(1-2p)(1+\mu_A)+\mu_B]\delta^A > [(1-2p)(1+\mu_B)+\mu_A]\delta^B. \tag{A12}$$

Noting that

$$(1-2p)(1+\mu_A)+\mu_B - [(1-2p)(1+\mu_B)+\mu_A] = 2p(\mu_B-\mu_A) > 0$$

provided that $\mu_B > \mu_A \geq 0$, as well as that $(1-2p)(1+\mu_B)+\mu_A \geq 0 \Leftrightarrow p \leq p^*_{A/B}$, so we get that (A12) automatically holds for either $p = p^*_{A/B}$, or $p < p^*_{A/B}$ and $\delta^A/\delta^B > \delta^*$ for $\delta^* \in (0,1)$. If, nevertheless, $p > p^*_{A/B}$, then (A12) holds whenever $(1-2p)(1+\mu_A)+\mu_B \geq 0$, that is, $p \leq p^*_{B/A}$. That is, (A12) holds for any $p \in \left[p^*_{A/B}, p^*_{B/A}\right]$ by noting that $p^*_{A/B} < p^*_{B/A}$. Moreover, if $p > p^*_{B/A}$, then we get by (A12) that

$$\frac{\delta^A}{\delta^B} < \frac{(1-2p)(1+\mu_B)+\mu_A}{(1-2p)(1+\mu_A)+\mu_B}. \tag{A13}$$

Since we now have $0 > (1-2p)(1+\mu_A)+\mu_B > (1-2p)(1+\mu_B)+\mu_A$, it must be that the right-hand side of inequality (A13) is strictly greater than one. In consequence, requirement (A13) is always met under Assumption 2.1, so (A12) holds true accordingly. To summarize, (A12) holds for any $p \in [p^*_{A/B}, 1)$, as desired.

Step 6 In particular, to obtain the characterization of optimal interregional redistribution shown in Part (ii), we just need to note from (A4) that

$$\begin{aligned} &\frac{(1+\mu_B)(1-p)\delta^B + [(1+\mu_A)p - \mu_A]\delta^A}{1+\mu_B-\mu_A} - (1-p)\delta^B \\ &= \frac{\mu_A(1-p)(\delta^B-\delta^A) + p\delta^A}{1+\mu_B-\mu_A} > 0 \end{aligned} \tag{A14}$$

under Assumption 2.1. That is, an application of the counterpart of (A10) yields that both (z^{A*}, b^{A*}) and (z^{B*}, b^{B*}) lie on the decreasing part of the welfare indifference curve of region B, which immediately produces the desired interregional redistribution in Part (ii). Finally, with the threshold of migration probability given by

$$p^*_{A/A} \equiv \frac{1}{2} + \frac{\mu_A}{2(1+\mu_A)},$$

assertions in Part (iii) can be analogously proved. □

Proof of Proposition 3.3. We shall complete the proof in two steps. In particular, as the proof of Part (ii) is similar to that shown below, we will omit it here to economize on the space.

Step 1 We first show Part (i-a) under the restriction that $\lambda^{\rm FB} = \lambda^* \equiv \lambda$. If $\mu_A > \mu_B$, then we get from (A1) and (A3) that

$$\begin{aligned} g_1'(G_1^{A*}) &= \frac{\lambda}{1+\mu_A-\mu_B} < \lambda = g_1'(G_1^{A,\rm FB}) \\ &\Rightarrow G_1^{A*} > G_1^{A,\rm FB}, \end{aligned}$$

which combined with (3) yields $c_1^{A*} > c_1^{A,\rm FB}$. Making use of (A1) and (A3) again reveals that

$$\begin{aligned} \frac{\lambda}{(1+r)g_2'(G_2^{A,\rm FB})} &= \delta^A(1-p) + \delta^B p, \\ \frac{\lambda}{(1+r)g_2'(G_2^{A*})} &= \delta^A(1-p) + \delta^B p + (1-p)(\mu_A\delta^A - \mu_B\delta^B). \end{aligned}$$

Letting $\mu_A\delta^A > \mu_B\delta^B$, then it is straightforward that $G_2^{A*} > G_2^{A,\rm FB}$ and $c_2^{A*} > c_2^{A,\rm FB}$. Note that $0 < G_2^{A*} - G_2^{A,\rm FB} = (b^{A,\rm FB} - b^{A*})(1+r) + (c_2^{A,\rm FB} - c_2^{A*})$, we thus have $(b^{A,\rm FB} - b^{A*})(1+r) > c_2^{A*} - c_2^{A,\rm FB} > 0$, as desired. Similarly, note that $0 < G_1^{A*} - G_1^{A,\rm FB} = (b^{A*} - b^{A,\rm FB}) + (z^{A*} - z^{A,\rm FB}) + (c_1^{A,\rm FB} - c_1^{A*})$, we thus immediately obtain $z^{A*} - z^{A,\rm FB} > b^{A,\rm FB} - b^{A*} + c_1^{A*} - c_1^{A,\rm FB} > 0$, which combined with Propositions 3.1 and 3.2 gives rise to the desired assertion. It follows from (A2) and (A4) that

$$\begin{aligned} g_1'(G_1^{B*}) &= \frac{\lambda}{1-(\mu_A-\mu_B)} > \lambda = g_1'(G_1^{B,\rm FB}) \\ &\Rightarrow G_1^{B*} < G_1^{B,\rm FB}, \end{aligned}$$

which combined with (3) yields $c_1^{B*} < c_1^{B,\mathrm{FB}}$. Making use of (A2) and (A4) again reveals that

$$\frac{\lambda}{(1+r)g_2'(G_2^{B,\mathrm{FB}})} = \delta^B(1-p) + \delta^A p,$$
$$\frac{\lambda}{(1+r)g_2'(G_2^{B*})} = \delta^B(1-p) + \delta^A p + (1-p)(\mu_B\delta^B - \mu_A\delta^A).$$

Under the restriction that $\mu_A\delta^A > \mu_B\delta^B$, we obtain $G_2^{B*} < G_2^{B,\mathrm{FB}}$ and $c_2^{B*} < c_2^{B,\mathrm{FB}}$. Note that $0 > G_2^{B*} - G_2^{B,\mathrm{FB}} = (b^{B,\mathrm{FB}} - b^{B*})(1+r) + (c_2^{B,\mathrm{FB}} - c_2^{B*})$, we thus have $(b^{B,\mathrm{FB}} - b^{B*})(1+r) < c_2^{B*} - c_2^{B,\mathrm{FB}} < 0$, as desired. Similarly, note that $0 > G_1^{B*} - G_1^{B,\mathrm{FB}} = (b^{B*} - b^{B,\mathrm{FB}}) + (z^{B*} - z^{B,\mathrm{FB}}) + (c_1^{B,\mathrm{FB}} - c_1^{B*})$, we thus immediately obtain $z^{B*} - z^{B,\mathrm{FB}} < b^{B,\mathrm{FB}} - b^{B*} + c_1^{B*} - c_1^{B,\mathrm{FB}} < 0$, which combined with Propositions 3.1 and 3.2 gives rise to the desired assertion.

Step 2 We now show Part (i-b). Letting $\lambda^*/\lambda^{\mathrm{FB}} \geq 1 + \mu_A - \mu_B > 1$ under $\mu_A > \mu_B$, then we get from (A1) and (A3) that

$$g_1'(G_1^{A*}) = \frac{\lambda^*}{1+\mu_A-\mu_B} \geq \lambda^{\mathrm{FB}} = g_1'(G_1^{A,\mathrm{FB}}),$$

hence $G_1^{A*} \leq G_1^{A,\mathrm{FB}}$ and $c_1^{A*} \leq c_1^{A,\mathrm{FB}}$. Also, it follows from (A1) and (A3) that

$$\frac{\lambda^{\mathrm{FB}}}{(1+r)g_2'(G_2^{A,\mathrm{FB}})} = \delta^A(1-p) + \delta^B p,$$
$$\frac{\lambda^*}{(1+r)g_2'(G_2^{A*})} = \delta^A(1-p) + \delta^B p + (1-p)(\mu_A\delta^A - \mu_B\delta^B). \tag{A15}$$

Letting $\mu_A\delta^A \leq \mu_B\delta^B$, then we have by (A15) that $G_2^{A*} < G_2^{A,\mathrm{FB}}$ and $c_2^{A*} < c_2^{A,\mathrm{FB}}$ under $\lambda^* > \lambda^{\mathrm{FB}}$. Note that $0 > G_2^{A*} - G_2^{A,\mathrm{FB}} = (b^{A,\mathrm{FB}} - b^{A*})(1+r) + (c_2^{A,\mathrm{FB}} - c_2^{A*})$, we thus must have $b^{A,\mathrm{FB}} < b^{A*}$. Also, note that

$$\begin{aligned} 0 &\geq G_1^{A*} - G_1^{A,\mathrm{FB}} \\ &= \underbrace{(b^{A*} - b^{A,\mathrm{FB}})}_{>0} + (z^{A*} - z^{A,\mathrm{FB}}) + \underbrace{\left(c_1^{A,\mathrm{FB}} - c_1^{A*}\right)}_{\geq 0}, \end{aligned}$$

we thus immediately obtain $z^{A*} < z^{A,\mathrm{FB}}$, which combined with Part (ii) of Proposition 3.1 reveals that

$$z^{A*} < z^{A,\mathrm{FB}} < 0 < z^{B,\mathrm{FB}} < z^{B*} \tag{A16}$$

must hold under $p < 1/2$. By $\mu_A > \mu_B$, $\lambda^* > \lambda^{\mathrm{FB}}$, (A2) and (A4), we have

$$g_1'\left(G_1^{B*}\right) = \frac{\lambda^*}{1-(\mu_A-\mu_B)} > \lambda^{\mathrm{FB}} = g_1'\left(G_1^{B,\mathrm{FB}}\right),$$

so $G_1^{B*} < G_1^{B,\mathrm{FB}}$ and $c_1^{B*} < c_1^{B,\mathrm{FB}}$. In addition, it follows from (A2) and (A4) that

$$\begin{aligned}\frac{\lambda^{\mathrm{FB}}}{(1+r)g_2'(G_2^{B,\mathrm{FB}})} &= \delta^B(1-p)+\delta^A p,\\ \frac{\lambda^*}{(1+r)g_2'(G_2^{B*})} &= \delta^B(1-p)+\delta^A p+(1-p)(\mu_B\delta^B-\mu_A\delta^A).\end{aligned} \tag{A17}$$

We thus have $g_2'(G_2^{B,\mathrm{FB}})/g_2'(G_2^{B*}) \geq \lambda^{\mathrm{FB}}/\lambda^*$ under $\mu_B\delta^B \geq \mu_A\delta^A$. Since here we assume that $\lambda^* > \lambda^{\mathrm{FB}}$, it seems indeterminate whether $g_2'(G_2^{B,\mathrm{FB}})/g_2'(G_2^{B*})$ is greater than one or not. Suppose, for the sake of contradiction, $g_2'(G_2^{B,\mathrm{FB}})/g_2'(G_2^{B*}) \leq 1$; then $G_2^{B*} \leq G_2^{B,\mathrm{FB}}$ and $c_2^{B*} \leq c_2^{B,\mathrm{FB}}$. Note that $0 \geq G_2^{B*} - G_2^{B,\mathrm{FB}} = (b^{B,\mathrm{FB}} - b^{B*})(1+r) + (c_2^{B,\mathrm{FB}} - c_2^{B*})$, we thus have $b^{B,\mathrm{FB}} \leq b^{B*}$. Also, note that

$$\begin{aligned} 0 &> G_1^{B*} - G_1^{B,\mathrm{FB}}\\ &= \underbrace{(b^{B*} - b^{B,\mathrm{FB}})}_{\geq 0} + (z^{B*} - z^{B,\mathrm{FB}}) + \underbrace{\left(c_1^{B,\mathrm{FB}} - c_1^{B*}\right)}_{>0},\end{aligned}$$

we thus immediately obtain $z^{B*} < z^{B,\mathrm{FB}}$, which however contradicts with (A16). As such, we have $g_2'(G_2^{B,\mathrm{FB}})/g_2'(G_2^{B*}) > 1$, and hence $b^{B,\mathrm{FB}} > b^{B*}$ follows. Making use of Part (ii) of Proposition 3.1 again, the desired assertion in Part (i-b) follows. To complete the proof of Part (i-b), we now assume that $\lambda^*/\lambda^{\mathrm{FB}} \leq 1 + \mu_B - \mu_A < 1$ under $\mu_A > \mu_B$. By (A2) and (A4), it is immediate that $G_1^{B*} \geq G_1^{B,\mathrm{FB}}$ and $c_1^{B*} \geq c_1^{B,\mathrm{FB}}$. Under $\mu_B\delta^B \geq \mu_A\delta^A$, we get from (A17) that $G_2^{B*} > G_2^{B,\mathrm{FB}}$ and $c_2^{B*} > c_2^{B,\mathrm{FB}}$. Then, as before, we can easily show that $b^{B,\mathrm{FB}} > b^{B*}$ and $z^{B*} > z^{B,\mathrm{FB}}$. Also, by Part (ii) of Proposition 3.1 we get the above (A16). We now consider region A, and we get by (A1) and (A3) that $G_1^{A*} > G_1^{A,\mathrm{FB}}$ and $c_1^{A*} > c_1^{A,\mathrm{FB}}$ under $\lambda^*/\lambda^{\mathrm{FB}} < 1$ and $\mu_A > \mu_B$. Using (A15), we can also show that $g_2'(G_2^{A,FB}) < g_2'(G_2^{A*})$ by means of contradiction. As a result, the remaining proof of Part (i-b) can be mechanically done. □

Proof of Proposition 4.1. To prove this proposition we just need to show that region A will choose debt level $b^{A,\mathrm{FB}}$ when receiving transfers $z^{A,\mathrm{FB}}$, and region B will choose debt level $b^{B,\mathrm{FB}}$ when receiving transfers $z^{B,\mathrm{FB}}$. Under the budget institution stated in Proposition 4.1, the maximization problem facing region A reads as $\max_{b^A} V(b^A, z^{A,\mathrm{FB}}, \delta^A; b^B)$ subject to $b^A \leq b^{A,\mathrm{FB}}$. Evaluating $V_b(b^A, z^{A,\mathrm{FB}}, \delta^A; b^B)$ at $b^A = b^{A,\mathrm{FB}}$, we get by (3), (6) and (A1) that

$$\begin{aligned} &V_b(b^{A,\mathrm{FB}}, z^{A,\mathrm{FB}}, \delta^A; b^B)\\ &\quad= g_1'\left(y_1 + b^{A,\mathrm{FB}} + z^{A,\mathrm{FB}} - c_1^{A,\mathrm{FB}}\right) - \delta^A(1-p)(1+r)g_2'\left(y_2 - (1+r)b^{A,\mathrm{FB}} - c_2^{A,\mathrm{FB}}\right)\\ &\quad= (1+r)g_2'\left(G_2^{A,\mathrm{FB}}\right)\left[\frac{g_1'(G_1^{A,\mathrm{FB}})}{(1+r)g_2'(G_2^{A,\mathrm{FB}})} - \delta^A(1-p)\right]\\ &\quad= (1+r)g_2'\left(G_2^{A,\mathrm{FB}}\right)[\delta^A(1-p) + \delta^B p - \delta^A(1-p)] > 0.\end{aligned}$$

Making use of (6) and (A9), we can get that

$$V_{bb}(b^A, z^{A,\mathrm{FB}}, \delta^A; b^B) = \frac{u_1''g_1''}{u_1'' + g_1''} + \delta^A(1-p)(1+r)^2\left(\frac{u_2''g_2''}{u_2'' + g_2''}\right) < 0$$

for any feasible b^A. We thus get by applying the budget restriction $b^A \leq b^{A,\mathrm{FB}}$ that

$$V_b(b^A, z^{A,\mathrm{FB}}, \delta^A; b^B) \geq V_b(b^{A,\mathrm{FB}}, z^{A,\mathrm{FB}}, \delta^A; b^B) > 0$$

which implies that region A shall choose $b^A = b^{A,\mathrm{FB}}$ whenever receiving $z^{A,\mathrm{FB}}$, that is, the first-best allocation $(b^{A,\mathrm{FB}}, z^{A,\mathrm{FB}})$ is realized. The result corresponding to region B can be obtained by symmetry. □