Online Appendix for "Nonlinear Taxation and International Mobility in General Equilibrium"*

Eckhard Janeba[†] **Karl Schulz**[‡]

University of Mannheim University of Mannheim

May 26, 2021

Abstract

This is the Online Appendix for "Nonlinear Taxation and International Mobility in General Equilibrium" by Janeba and Schulz (2021). In Section 1, we generalize our results to an arbitrary number of N types. We demonstrate that our main insights regarding the bottom and top tax rates continue to hold. Moreover, we show that the moderation effect of migration in general equilibrium on the taxation of middle incomes depends on the local progressivity of the tax scheme absent of migration. In Section 2, we study a continuous-type setting allowing for the overlap of different types' income distributions. We provide a rigorous derivation of the sufficient statistics relevant for studying the nonlinear tax incidence and the optimal taxation in general equilibrium with migration.

*Eckhard Janeba gratefully acknowledges the support from the Collaborative Research Center (SFB) 884 "Political Economy of Reforms", funded by the German Research Foundation (DFG). Karl Schulz gratefully acknowledges the support by the University of Mannheim's Graduate School of Economic and Social Sciences, funded by the German Research Foundation (DFG). For helpful comments and discussions, we thank Hans Peter Grüner, Duk Gyoo Kim, Mathilde Muñoz, Nicolas Werquin, and numerous seminar and conference participants.

†University of Mannheim, Department of Economics, L7 3-5, 68131 Mannheim, Germany, janeba@uni-mannheim.de.

‡University of Mannheim, Center for Doctoral Studies in Economics, L7 3-5, 68131 Mannheim, Germany, karl.schulz@gess.uni-mannheim.de.

1 Extension to N Types

We go beyond the two-type setting by generalizing our results to an arbitrary set of types $\theta \in \Theta = \{1,...,N\}$. Most importantly, this allows us to speak to the effects of migration and general equilibrium on the optimal taxation of the 'middle class'. Without loss of generality, we order types such that their equilibrium wages are increasing in types $w_{i,\theta} > w_{i,\theta-1}$. Therefore, one can interpret a household's type as its position (e.g., percentile) in the equilibrium wage distribution. Moreover, consider a more general class of utility functions, $u(c,l)$, that satisfies the Spence-Mirrlees single-crossing property $\frac{d}{dw}\frac{-u_l(c,y/w)/w}{u_c(c,y/w)} < 0$. Suppose that labor and consumption are separable (e.g., $u(c,l) = h(c) - v(l)$).

The set of incentive constraints reads as $u\left(c_{i,\theta}, l_{i,\theta}\right) \geq u\left(c_{i,\theta'}, \frac{l_{i,\theta'}w_{i,\theta'}}{w_{i,\theta}}\right)$, $\forall \theta, \theta' \in \Theta$. Observe that one can omit non-local incentive constraints where $\theta' < \theta - 1$ and $\theta' > \theta + 1$ (see Milgrom and Shannon (1994)). As in our baseline setup, we focus on solutions where only the local downward incentive constraints, $u\left(c_{i,\theta}, l_{i,\theta}\right) \geq u\left(c_{i,\theta-1}, \frac{l_{i,\theta-1}w_{i,\theta-1}}{w_{i,\theta}}\right)$, bind, and income is redistributed, $T_i\left(y_{i,\theta}\right) \geq T_i\left(y_{i,\theta-1}\right)$. For a set of Pareto weights $\left\{\psi_{i,\theta}\right\}_{\theta\in\Theta}$ with $\psi_{i,\theta} \geq \psi_{i,\theta+1}$, the planner problem of country i's government is given by

$$\max_{\left\{c_{i,\theta}, l_{i,\theta}, N_{i,\theta}\right\}_{\theta\in\Theta}} \sum_{\theta\in\Theta} \psi_{i,\theta} u\left(c_{i,\theta}, l_{i,\theta}\right) N_{i,\theta} \tag{1}$$

subject to

$$u\left(c_{i,\theta}, l_{i,\theta}\right) \geq u\left(c_{i,\theta-1}, \frac{l_{i,\theta-1}w_{i,\theta-1}}{w_{i,\theta}}\right) \; for \; \theta \in \{2,...,N\}, \tag{2}$$

$$\sum_{\theta\in\Theta} N_{i,\theta}c_{i,\theta} \leq F_i\left(\left\{N_{i,\theta}l_{i,\theta}\right\}_{\theta\in\Theta}\right), \tag{3}$$

$$w_{i,\theta} = \frac{\partial F_i\left(\left\{N_{i,\theta}l_{i,\theta}\right\}_{\theta\in\Theta}\right)}{\partial\left(N_{i,\theta}l_{i,\theta}\right)} \; for \; \theta \in \Theta, \tag{4}$$

and

$$N_{i,\theta} = \rho_i\left(\Delta_i; \theta\right) \; for \; \theta \in \Theta. \tag{5}$$

Observe that this planner problem is identical to the one in Ales, Kurnaz, and Sleet (2015) except for the endogeneity of the equilibrium population (Equation (5)).

Inner problem. Optimizing the Lagrangian

$$\mathscr{L}_i\left(\{N_{i,\theta}\}_{\theta\in\Theta}\right)=\sum_{\theta\in\Theta}\psi_{i,\theta}u\left(c_{i,\theta},l_{i,\theta}\right)N_{i,\theta}+\sum_{\theta\in\{2,\ldots,N\}}\mu_{i,\theta}\left[u\left(c_{i,\theta},l_{i,\theta}\right)-u\left(c_{i,\theta-1},\frac{l_{i,\theta-1}w_{i,\theta-1}}{w_{i,\theta}}\right)\right]$$
$$+\xi_i\left[F_i\left(\{N_{i,\theta}l_{i,\theta}\}_{\theta\in\Theta}\right)-\sum_{\theta\in\Theta}N_{i,\theta}c_{i,\theta}\right]+\sum_{\theta\in\Theta}\lambda_{i,\theta}\left[N_{i,\theta}-\rho_i\left(\Delta_i;\theta\right)\right],$$

with respect to $\left\{c_{i,\theta},l_{i,\theta}\right\}_{\theta\in\Theta}$, the first-order conditions of the inner problem can be written as

$$[c_{i,\theta}]:0=\left(\psi_{i,\theta}u_c\left(c_{i,\theta},l_{i,\theta}\right)-\xi_i\right)N_{i,\theta}+\mathbb{1}\left[\theta>1\right]\mu_{i,\theta}u_c\left(c_{i,\theta},l_{i,\theta}\right)-\mathbb{1}\left[\theta<N\right]\mu_{i,\theta+1}u_c\left(c_{i,\theta},\frac{l_{i,\theta}w_{i,\theta}}{w_{i,\theta+1}}\right)$$
$$-\lambda_{i,\theta}\eta_{i,\theta}N_{i,\theta}u_c\left(c_{i,\theta},l_{i,\theta}\right)\tag{6}$$
$$[l_{i,\theta}]:0=\psi_{i,\theta}u_l\left(c_{i,\theta},l_{i,\theta}\right)N_{i,\theta}+\mathbb{1}\left[\theta>1\right]\mu_{i,\theta}u_l\left(c_{i,\theta},l_{i,\theta}\right)-\mathbb{1}\left[\theta<N\right]\mu_{i,\theta+1}u_l\left(c_{i,\theta},\frac{l_{i,\theta}w_{i,\theta}}{w_{i,\theta+1}}\right)\frac{w_{i,\theta}}{w_{i,\theta+1}}$$
$$+\xi_iw_{i,\theta}N_{i,\theta}-\lambda_{i,\theta}\eta_{i,\theta}N_{i,\theta}u_l\left(c_{i,\theta},l_{i,\theta}\right)-\sum_{k\in\{2,\ldots,N\}}\mu_{i,k}u_l\left(c_{i,k-1},\frac{l_{i,k-1}w_{i,k-1}}{w_{i,k}}\right)\frac{l_{i,k-1}w_{i,k-1}}{l_{i,\theta}w_{i,k}}\left(\gamma_{i,k-1,\theta}-\gamma_{i,k,\theta}\right),\tag{7}$$

for $\theta\in\Theta$, where $\mathbb{1}\left[\cdot\right]$ is the indicator function. Plug (6) into (7) and use the individual worker's first-order condition

$$u_c\left(c_{i,\theta},l_{i,\theta}\right)w_{i,\theta}\left(1-T_i'\left(y_{i,\theta}\right)\right)=-u_l\left(c_{i,\theta},l_{i,\theta}\right)$$

to obtain an expression for the optimal tax scheme

$$T_i'\left(y_{i,\theta}\right)=\frac{\mathbb{1}\left[\theta<N\right]\frac{\mu_{i,\theta+1}}{\xi_i}\frac{u_c\left(c_{i,\theta},\frac{l_{i,\theta}w_{i,\theta}}{w_{i,\theta+1}}\right)}{N_{i,\theta}}\left[1+\frac{u_l\left(c_{i,\theta},\frac{l_{i,\theta}w_{i,\theta}}{w_{i,\theta+1}}\right)}{w_{i,\theta+1}u_c\left(c_{i,\theta},l_{i,\theta}\right)}\right]}{1+\mathbb{1}\left[\theta<N\right]\frac{\mu_{i,\theta+1}}{\xi_i}\frac{u_c\left(c_{i,\theta},\frac{l_{i,\theta}w_{i,\theta}}{w_{i,\theta+1}}\right)}{N_{i,\theta}}}$$
$$+\frac{\frac{1}{y_{i,\theta}N_{i,\theta}}\sum_{k\in\{2,\ldots,N\}}\frac{\mu_{i,k}}{\xi_i}u_l\left(c_{i,k-1},\frac{l_{i,k-1}w_{i,k-1}}{w_{i,k}}\right)\frac{l_{i,k-1}w_{i,k-1}}{w_{i,k}}\left(\gamma_{i,k-1,\theta}-\gamma_{i,k,\theta}\right)}{1+\mathbb{1}\left[\theta<N\right]\frac{\mu_{i,\theta+1}}{\xi_i}\frac{u_c\left(c_{i,\theta},\frac{l_{i,\theta}w_{i,\theta}}{w_{i,\theta+1}}\right)}{N_{i,\theta}}}.$$

In their Proposition (1), Ales et al. (2015) show that one can decompose the formula for the optimal tax rate in the N-type Stiglitz (1982) setting without migration ($\eta_{i,\theta}=0$, $\forall\theta\in\Theta$) into a Mirrleesian and a wage compression term. A similar decomposition applies here with the difference that one needs to account for migration responses ($\eta_{i,\theta}>0$).

On the one hand, the Mirrleesian term is augmented by the direct partial equilibrium impact of migration on the objective function, which Lehmann, Simula, and Trannoy (2014) label as the "migration threat." On the other hand, migration interacts with the wage compression term. Thus, this setting nests the well-known partial equilibrium effect of migration on the optimal taxation and adds general equilibrium moderation effects. In the following, we derive conditions under which the classical partial equilibrium downward force of migration on taxes is offset by our novel general equilibrium moderation effects.

From now on, we assume that the production function has constant returns to scale

$$F_i\left(\{N_{i,\theta}l_{i,\theta}\}_{\theta\in\Theta}\right) = \left[\sum_{\theta\in\Theta} a_{i,\theta}\left(N_{i,\theta}l_{i,\theta}\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$$

such that the wage rate reads as $w_{i,\theta} = a_{i.\theta}\left[L_{i,\theta}/F_i\left(\{N_{i,\theta}l_{i,\theta}\}_{\theta\in\Theta}\right)\right]^{-\frac{1}{\sigma}}$. Then, type θ's tax rate simplifies to

$$T_i'\left(y_{i,\theta}\right) = \frac{\mathbb{1}\left[\theta<N\right]\frac{\mu_{i,\theta+1}}{\xi_i}\frac{u_c\left(c_{i,\theta},\frac{l_{i,\theta}w_{i,\theta}}{w_{i,\theta+1}}\right)}{N_{i,\theta}}\left[1+\frac{u_l\left(c_{i,\theta},\frac{l_{i,\theta}w_{i,\theta}}{w_{i,\theta+1}}\right)}{w_{i,\theta+1}u_c\left(c_{i,\theta},l_{i,\theta}\right)}\frac{\sigma-1}{\sigma}\right]+\mathbb{1}\left[\theta>1\right]\frac{\mu_{i,\theta}}{\xi_i}\frac{y_{i,\theta-1}}{y_{i,\theta}N_{i,\theta}}\frac{1}{\sigma}\frac{u_l\left(c_{i,\theta-1},\frac{l_{i,\theta-1}w_{i,\theta-1}}{w_{i,\theta}}\right)}{w_{i,\theta}}}{1+\mathbb{1}\left[\theta<N\right]\frac{\mu_{i,\theta+1}}{\xi_i}\frac{u_c\left(c_{i,\theta},\frac{l_{i,\theta}w_{i,\theta}}{w_{i,\theta+1}}\right)}{N_{i,\theta}}}.$$

One can obtain a value for the shadow value of public funds by summing up Equation (6), which yields $\xi_i = \frac{\sum_{\theta\in\Theta}\psi_{i,\theta}N_{i,\theta}-\sum_{\theta\in\Theta}\lambda_{i,\theta}\eta_{i,\theta}N_{i,\theta}}{\sum_{\theta\in\Theta}N_{i,\theta}/u_c\left(c_{i,\theta},l_{i,\theta}\right)}$. Plugging this value back into (6), the normalized multiplier on the incentive constraint for $k\in\{2,...,N\}$ reads as

$$\frac{\mu_{i,k}}{\xi_i} = \frac{\sum_{l=k}^{N}N_{i,l}/u_c\left(c_{i,l},l_{i,l}\right)\sum_{l=1}^{k-1}\left(\psi_{i,l}N_{i,l}-\lambda_{i,l}\eta_{i,l}N_{i,l}\right)-\sum_{l=1}^{k-1}N_{i,l}/u_c\left(c_{i,l},l_{i,l}\right)\sum_{l=k}^{N}\left(\psi_{i,l}N_{i,l}-\lambda_{i,l}\eta_{i,l}N_{i,l}\right)}{\sum_{\theta\in\Theta}\psi_{i,\theta}N_{i,\theta}-\sum_{\theta\in\Theta}\lambda_{i,\theta}\eta_{i,\theta}N_{i,\theta}}.$$

Then, we define the normalized multiplier without migration responses and a scaling factor as

$$\overline{\frac{\mu_{i,k}}{\xi_i}} \equiv \frac{\sum_{l=k}^{N}N_{i,l}/u_c\left(c_{i,l},l_{i,l}\right)\sum_{l=1}^{k-1}\psi_{i,l}N_{i,l}-\sum_{l=1}^{k-1}N_{i,l}/u_c\left(c_{i,l},l_{i,l}\right)\sum_{l=k}^{N}\psi_{i,l}N_{i,l}}{\sum_{\theta\in\Theta}\psi_{i,\theta}N_{i,\theta}}$$

and $\Psi_{i,k}\left(\left\{\eta_{i,\theta}\right\}_{\theta\in\Theta}\right) \equiv \frac{\mu_{i,k}}{\xi_i}/\overline{\frac{\mu_{i,\theta}}{\xi_i}}$, respectively.

Thus, one can write the Lagrange multiplier with migration, $\frac{\mu_{i,\theta}}{\xi_i} = \overline{\frac{\mu_{i,\theta}}{\xi_i}}\cdot\Psi_{i,\theta}\left(\left\{\eta_{i,\theta}\right\}_{\theta\in\Theta}\right)$, as the product of the multiplier without migration, $\overline{\frac{\mu_{i,\theta}}{\xi_i}}$, and a scaling factor $\Psi_{i,\theta}\left(\left\{\eta_{i,\theta}\right\}_{\theta\in\Theta}\right)$. Below, we describe conditions under which $\Psi_{i,\theta}\left(\left\{\eta_{i,\theta}\right\}_{\theta\in\Theta}\right) < 1$. Similarly, the tax rate with migration is a function of the tax set by the exogenous technology planner $T_i'\left(y_{i,\theta}\right) = \overline{T}_i'\left(y_{i,\theta}\right)|_{T_i'}\cdot$

$f_{i,\theta}\left(\left\{\Psi_{i,\theta}\left(\{\eta_{i,\theta}\}_{\theta\in\Theta}\right)\right\}_{\theta\in\Theta}\right)$. In the following, we compare $T_i'\left(y_{i,\theta}\right)$ and $\overline{T}_i'\left(y_{i,\theta}\right)|_{T_i'}$ at different positions in the wage distribution.

Bottom and top tax rate. For a CES production technology, the top tax rate simplifies to

$$0 > T_i'\left(y_{i,N}\right) = \Psi_{i,N}\left(\{\eta_{i,\theta}\}_{\theta\in\Theta}\right)\frac{\overline{\mu_{i,N}}}{\xi_i}\frac{y_{i,N-1}}{y_{i,N}N_{i,N}}\frac{1}{\sigma}\frac{u_l\left(c_{i,N-1},\frac{l_{i,N-1}w_{i,N-1}}{w_{i,N}}\right)}{w_{i,N}} = \Psi_{i,N}\overline{T}_i'\left(y_{i,N}\right)|_{T_i'} > \overline{T}_i'\left(y_{i,N}\right)|_{T_i'}. \tag{8}$$

Therefore, as in our two-type setup, the optimal tax rate in the Nash equilibrium rises (becomes less negative) at the top. At the same time, the bottom tax declines

$$0 < T_i'\left(y_{i,1}\right) = \frac{\frac{\overline{\mu_{i,2}}}{\xi_i}\Psi_{i,2}\left(\{\eta_{i,\theta}\}_{\theta\in\Theta}\right)\frac{u_c\left(c_{i,1},\frac{l_{i,1}w_{i,1}}{w_{i,2}}\right)}{N_{i,1}}}{1+\frac{\overline{\mu_{i,2}}}{\xi_i}\Psi_{i,2}\left(\{\eta_{i,\theta}\}_{\theta\in\Theta}\right)\frac{u_c\left(c_{i,1},\frac{l_{i,1}w_{i,1}}{w_{i,2}}\right)}{N_{i,1}}}\frac{1+\frac{\overline{\mu_{i,2}}}{\xi_i}\frac{u_c\left(c_{i,1},\frac{l_{i,1}w_{i,1}}{w_{i,2}}\right)}{N_{i,1}}}{\frac{\overline{\mu_{i,2}}}{\xi_i}\frac{u_c\left(c_{i,1},\frac{l_{i,1}w_{i,1}}{w_{i,2}}\right)}{N_{i,1}}}\overline{T}_i'\left(y_{i,1}\right)|_{T_i'} < \overline{T}_i'\left(y_{i,1}\right)|_{T_i'}. \tag{9}$$

Thus, our main finding that migration raises optimal tax rates at the top and reduces them at the bottom continues to hold in this N-type economy.

Proposition 1. *In the Nash equilibrium of our N-type economy, the optimal marginal tax at the top is higher, $T_i'\left(y_{i,N}\right) > \overline{T}_i'\left(y_{i,N}\right)|_{T_i'}$, and the optimal marginal tax at the bottom is lower, $T_i'\left(y_{i,1}\right) < \overline{T}_i'\left(y_{i,1}\right)|_{T_i'}$, than in the closed economy.*

Middle tax rate. The effect of migration on the tax rate of middle incomes, $\theta \in \{2,...,N-1\}$, is case-specific. Firstly, considering the binding incentive constraint of θ-workers,

$$u\left(c_{i,\theta},l_{i,\theta}\right) = u\left(c_{i,\theta-1},\frac{l_{i,\theta-1}w_{i,\theta-1}}{w_{i,\theta}}\right),$$

observe that the government can raise θ-workers' utility by amplifying pre-tax wage inequality between $\theta-1$ and θ. Secondly, by the binding incentive constraint at $\theta+1$,

$$u\left(c_{i,\theta+1},l_{i,\theta+1}\right) = u\left(c_{i,\theta},\frac{l_{i,\theta}w_{i,\theta}}{w_{i,\theta+1}}\right),$$

a higher pre-tax wage inequality between θ and $\theta+1$ increases utility for $\theta+1$-workers.

To raise θ-workers' utility, the government has an incentive to tax type θ more to reduce their labor supply and, thereby, raise $\frac{w_{i,\theta}}{w_{i,\theta-1}}$. We call the described channel "local trickle-down" since a tax cut on the relatively richer θ-workers would trickle down to the poorer workers of type $\theta-1$. At the same time, the second incentive constraint calls for a higher $\frac{w_{i,\theta+1}}{w_{i,\theta}}$, which the government can achieve by taxing θ-workers less. We label this mechanism as "local trickle-up" because the tax cut on the poor (type θ) trickles up to the rich (type $\theta+1$). To simplify the exposition, we make the following assumption.

Assumption 1. *Let the interaction between the migration semi-elasticity and general equilibrium production complementarities be small:* $\frac{1}{\sigma}\eta_{i,\theta} \to 0$ *for all* $\theta \in \Theta$.

Under Assumption 1, one can easily relate the effect of migration in general equilibrium to the local progressivity of the exogenous technology planner's tax rate. The presence of migration raises θ-workers' optimal tax rate, i.e., $T_i\left(y_{i,\theta}\right) > \overline{T}'_i\left(y_{i,\theta}\right)|_{T'_i}$, if and only if

$$\overline{T}'_i\left(y_{i,\theta}\right)|_{T'_i} < \frac{\Psi_{i,\theta+1}\left(\{\eta_{i,\theta}\}_{\theta\in\Theta}\right) - \Psi_{i,\theta}\left(\{\eta_{i,\theta}\}_{\theta\in\Theta}\right)}{1-\Psi_{i,\theta+1}\left(\{\eta_{i,\theta}\}_{\theta\in\Theta}\right)} \frac{\overline{\mu_{i,\theta}}}{\xi_i} \frac{y_{i,\theta-1}}{y_{i,\theta}N_{i,\theta}} \frac{1}{\sigma} \frac{-u_l\left(c_{i,\theta-1}, \frac{l_{i,\theta-1}w_{i,\theta-1}}{w_{i,\theta}}\right)}{w_{i,\theta}}. \qquad (10)$$

In general, the right-hand side of this equation depends on the sign of $\Psi_{i,\theta+1}\left(\{\eta_{i,\theta}\}_{\theta\in\Theta}\right) - \Psi_{i,\theta}\left(\{\eta_{i,\theta}\}_{\theta\in\Theta}\right)$ for $\Psi_{i,\theta}\left(\{\eta_{i,\theta}\}_{\theta\in\Theta}\right) < 1$ (see below). However, under Assumption 1, the term tends to zero, and Proposition 2 follows.

Proposition 2. *Let Assumption 1 hold. Then, in the Nash equilibrium of our N-type economy, the optimal marginal tax of a worker with type* θ *is higher in the open economy,* $T_i\left(y_{i,\theta}\right) > \overline{T}'_i\left(y_{i,\theta}\right)|_{T'_i}$*, if and only if the tax set by the exogenous technology planner is locally regressive,* $\overline{T}'_i\left(y_{i,\theta}\right)|_{T'_i} < 0$*. The optimal marginal tax of a* θ*-type worker is lower in the open economy if and only if the exogenous technology planner's tax is locally progressive,* $\overline{T}'_i\left(y_{i,\theta}\right)|_{T'_i} > 0$*.*

Thus, if the exogenous technology planner's tax rate is locally regressive, the tax rate will rise due to the presence of migration. In this situation, the exogenous technology planner's incentive to decrease θ's tax rate to raise pre-tax wage inequality between $\theta-1$ and θ is sufficiently strong. The mechanism at the top that calls for a rise in wage inequality $\frac{w_{i,\theta}}{w_{i,\theta-1}}$ (local trickle-down) dominates the local trickle-up mechanism that calls for a rise in $\frac{w_{i,\theta+1}}{w_{i,\theta}}$. If the exogenous technology planner sets a locally progressive tax rate, then the presence of migration responses has the intuitive negative impact on the optimal marginal tax rate.

Condition for $\Psi_{i,\theta}\left(\{\eta_{i,\theta}\}_{\theta\in\Theta}\right) < 1$. To show that $\Psi_{i,\theta}\left(\{\eta_{i,\theta}\}_{\theta\in\Theta}\right) < 1$, $\forall\theta\in\Theta$, first, notice that

$$\Psi_{i,\theta}\left(\{\eta_{i,\theta}\}_{\theta\in\Theta}\right)=\frac{\sum_{l=\theta}^{N}N_{i,l}/u_c\left(c_{i,l},l_{i,l}\right)\sum_{l=1}^{\theta-1}\left(\psi_{i,l}N_{i,l}-\lambda_{i,l}\eta_{i,l}N_{i,l}\right)-\sum_{l=1}^{\theta-1}N_{i,l}/u_c\left(c_{i,l},l_{i,l}\right)\sum_{l=\theta}^{N}\left(\psi_{i,l}N_{i,l}-\lambda_{i,l}\eta_{i,l}N_{i,l}\right)}{\sum_{l=\theta}^{N}N_{i,l}/u_c\left(c_{i,l},l_{i,l}\right)\sum_{l=1}^{\theta-1}\psi_{i,l}N_{i,l}-\sum_{l=1}^{\theta-1}N_{i,l}/u_c\left(c_{i,l},l_{i,l}\right)\sum_{l=\theta}^{N}\psi_{i,l}N_{i,l}}$$
$$\cdot\frac{\sum_{\theta\in\Theta}\psi_{i,\theta}N_{i,\theta}}{\sum_{\theta\in\Theta}\psi_{i,\theta}N_{i,\theta}-\sum_{\theta\in\Theta}\lambda_{i,\theta}\eta_{i,\theta}N_{i,\theta}}<1$$

if and only if

$$\left(\sum_{l=\theta}^{N}\psi_{i,l}N_{i,l}\right)\sum_{l=1}^{\theta-1}\lambda_{i,l}\eta_{i,l}N_{i,l}-\left(\sum_{l=1}^{\theta-1}\psi_{i,l}N_{i,l}\right)\sum_{l=\theta}^{N}\lambda_{i,l}\eta_{i,l}N_{i,l}>0.$$

This condition is fulfilled for $\psi_{i,\theta'}\lambda_{i,\theta}\eta_{i,\theta}-\psi_{i,\theta}\lambda_{i,\theta'}\eta_{i,\theta'}>0$, $\forall\theta'>\theta$. Observe the similarity to the respective condition in the two-type setting ($\psi_{i,H}\lambda_{i,L}\eta_{i,L}-\psi_{i,L}\lambda_{i,H}\eta_{i,H}>0$).

Outer problem. To show that $\psi_{i,\theta'}\lambda_{i,\theta}\eta_{i,\theta}-\psi_{i,\theta}\lambda_{i,\theta'}\eta_{i,\theta'}>0$, $\forall\theta'>\theta$, we need to consider solution to the outer problem. That is, $\underset{\left\{N_{i,\theta}\right\}_{\theta\in\Theta}}{max}\ \mathscr{L}_i\left(\left\{N_{i,\theta}\right\}_{\theta\in\Theta}\right)$, which yields the first-order condition

$$\begin{aligned}[N_{i,\theta}]:0=\psi_{i,\theta}u\left(c_{i,\theta},l_{i,\theta}\right)-\sum_{k\in\{2,\dots,N\}}\mu_{i,k}u_l\left(c_{i,\theta-1},\frac{l_{i,k-1}w_{i,k-1}}{w_{i,k}}\right)\frac{l_{i,k-1}w_{i,k-1}}{N_{i,\theta}w_{i,k}}\left(\gamma_{i,k-1,\theta}-\gamma_{i,k,\theta}\right)\\+\xi_i\left(l_{i,\theta}w_{i,\theta}-c_{i,\theta}\right)+\lambda_{i,\theta},\end{aligned}\tag{11}$$

for any $\theta\in\Theta$.

Assuming a CES production function, this expression reads as

$$\begin{aligned}-\lambda_{i,\theta}\eta_{i,\theta}=&\,\psi_{i,\theta}u\left(c_{i,\theta},l_{i,\theta}\right)\eta_{i,\theta}+\xi_iT_i\left(y_{i,\theta}\right)\eta_{i,\theta}\\&+\frac{1}{\sigma}\eta_{i,\theta}\left[\mu_{i,\theta+1}u_l\left(c_{i,\theta},\frac{l_{i,\theta}w_{i,\theta}}{w_{i,\theta+1}}\right)\frac{l_{i,\theta}w_{i,\theta}}{N_{i,\theta}w_{i,\theta+1}}-\mu_{i,\theta}u_l\left(c_{i,\theta-1},\frac{l_{i,\theta-1}w_{i,\theta-1}}{w_{i,\theta}}\right)\frac{l_{i,\theta-1}w_{i,\theta-1}}{N_{i,\theta}w_{i,\theta}}\right]\end{aligned}$$

and, under Assumption 1, simplifies to $-\lambda_{i,\theta}\eta_{i,\theta}=\psi_{i,\theta}u\left(c_{i,\theta},l_{i,\theta}\right)\eta_{i,\theta}+\xi_iT_i\left(y_{i,\theta}\right)\eta_{i,\theta}$. Therefore, for any $\theta'>\theta$,

$$\begin{aligned}\psi_{i,\theta'}\lambda_{i,\theta}\eta_{i,\theta}-\psi_{i,\theta}\lambda_{i,\theta'}\eta_{i,\theta'}=&\,\psi_{i,\theta}\psi_{i,\theta'}\left[u\left(c_{i,\theta'},l_{i,\theta'}\right)\eta_{i,\theta'}-u\left(c_{i,\theta},l_{i,\theta}\right)\eta_{i,\theta}\right]\\&+\xi_i\left[\psi_{i,\theta}T_i\left(y_{i,\theta'}\right)\eta_{i,\theta'}-\psi_{i,\theta'}T_i\left(y_{i,\theta}\right)\eta_{i,\theta}\right]>0,\end{aligned}$$

if $\eta_{i,\theta'}\geq\eta_{i,\theta}$ and $T_i\left(y_{i,\theta'}\right)\geq T_i\left(y_{i,\theta}\right)$. Hence, we conclude that $\Psi_{i,\theta}\left(\left\{\eta_{i,\theta}\right\}_{\theta\in\Theta}\right)<1$.

2 Extension to Continuous Types

So far, we have dealt with households who differ in their migration cost and their skill. Now, we add an additional source of heterogeneity: within-skill heterogeneity. We return to the two-type setup (between-skill heterogeneity), as in Stiglitz (1982), and introduce, for each skill group, a continuum of individual productivities with measure one. Individuals are indexed by $j \in \mathscr{I}_i$, with $\mathscr{I}_i$ denoting country i's set of households ($i = A, B$) and $\mathscr{I}_{i,\theta}$ the mass of households with skill $\theta = L, H$. Hence, $\mathscr{I}_i \equiv \mathscr{I}_{i,L} \cup \mathscr{I}_{i,H}$ and $\mathscr{I}_{i,L} \cap \mathscr{I}_{i,H} = \emptyset$. Individual j's labor income reads as $y_{i,j} \equiv l_{i,j} w_{i,\theta_j} \omega_{i,j}$, where $l_{i,j}$ are the hours worked, w_{i,θ_j} is the endogenous return to the individual's skill $\theta_j \in \{L, H\}$, and $\omega_{i,j}$ denotes her individual labor productivity draw. This setup is known as the canonical model (Acemoglu and Autor (2011)). Define the aggregate effective labor of each skill group as $L_{i,\theta} \equiv \int_{\mathscr{I}_{i,\theta}} l_{i,j} N_{i,j} \omega_{i,j} dj$, where $N_{i,j}$ is the endogenous mass of workers with skill θ and productivity $\omega_{i,j}$. Then, each skill group's wage is given by $w_{i,\theta} = \frac{\partial F_i\left(L_{i,L}, L_{i,H}\right)}{\partial L_{i,\theta}}$, where $F_i\left(L_{i,L}, L_{i,H}\right)$ is the constant-returns-to-scale production function. As before, let $\gamma_{i,\theta,\theta} \equiv \frac{\partial w_{i,\theta}}{\partial L_{i,\theta}} \frac{L_{i,\theta}}{w_{i,\theta}} < 0$ and $\gamma_{i,\theta,\theta'} \equiv \frac{\partial w_{i,\theta}}{\partial L_{i,\theta'}} \frac{L_{i,\theta'}}{w_{i,\theta}} > 0$ be the own- and cross-wage elasticities. Denote the semi-elasticity of migration as $\eta_{i,j} \equiv \frac{\partial N_{i,\theta}}{\partial \Delta_i} \frac{1}{N_{i,\theta}}$, where Δ_i is the utility gain from being in country i, and similarly, the migration elasticity as $\nu_{i,j} \equiv c_{i,j} \eta_{i,j}$ (see Lehmann et al. (2014)). As in our baseline setup, we abstract from income effects. Each country's government taxes labor income non-linearly according to the tax scheme $T_i\left(y_{i,j}\right)$.

This section aims to express the nonlinear incidence of tax reforms (and the resulting optimal tax formulas) in terms of empirically observable sufficient statistics. Thereby, we identify a set of novel interaction effects between wage and migration responses. It is more convenient to turn to the dual (instead of the primal) approach and directly maximize the objective function with respect to the tax schedule (see, for instance, Saez (2001) and Golosov, Tsyvinski, and Werquin (2014)). As before, the government's objective is to maximize welfare subject to a budget constraint. Defining each household's indirect utility as $\mathscr{U}_{i,j}$, the planner problem in country i reads as

$$\max_{T_i()} \mathscr{G}_i \equiv \max_{T_i()} \int_{\mathscr{I}_{i,L}} \psi_{i,j} \mathscr{U}_{i,j} N_{i,j} dj + \int_{\mathscr{I}_{i,H}} \psi_{i,j} \mathscr{U}_{i,j} N_{i,j} dj. \tag{12}$$

subject to

$$\mathscr{R}_i \equiv \int_{\mathscr{I}_{i,L}} T_i\left(y_{i,j}\right) N_{i,j} dj + \int_{\mathscr{I}_{i,H}} T_i\left(y_{i,j}\right) N_{i,j} dj \geq 0 \tag{13}$$

as well as subject to the households' optimal labor supply and migration decisions, the endogeneity of wages, and the other country's tax scheme. Denote the social marginal weights as $\Gamma_{i,j} \equiv \frac{1}{\xi_i}\psi_{i,j}$, where ξ_i is, again, the shadow value of public funds.

In the following, we derive the individual and aggregate effects of reforming an initial (potentially sub-optimal) tax scheme, e.g., the US tax code, T_i, to a new tax code, $T_i + \delta\hat{T}_i$. As a byproduct, by equalizing the aggregate marginal benefits to the aggregate marginal costs, this analysis also delivers the optimal tax code. To formalize this reform-approach, we define the Gateaux derivative of the functional $\mathscr{F} : \mathscr{C}(\mathbb{R}_+, \mathbb{R}) \to \mathbb{R}$ at T_i in the direction $\hat{T}_i$ by $\hat{\mathscr{F}} \equiv \frac{d}{d\delta}\mathscr{F}\left(T_i + \delta\hat{T}_i\right)|_{\delta=0}$ (see Golosov et al. (2014)).

Labor supply and migration responses. By perturbing the individuals' first-order condition, $v'\left(l_{i,j}\right) = \omega_{i,j} w_{i,\theta_j}\left(1 - T_i'\left(y_{i,j}\right)\right)$, the labor supply response to the tax reform reads as $\frac{\hat{l}_{i,j}}{l_{i,j}} = -\varepsilon_{i,j}^{l,\left(1-T_i'\right)} \frac{\hat{T}_i'\left(y_{i,j}\right)}{1-T_i'\left(y_{i,j}\right)} + \varepsilon_{i,j}^{l,w}\frac{\hat{w}_{i,\theta_j}}{w_{i,\theta_j}}$, where $\varepsilon_{i,j}^{l,\left(1-T_i'\right)}$ is the hours elasticity with respect to the retention rate of the labor income tax rate and $\varepsilon_{i,j}^{l,w}$ is the hours elasticity to a change in the wage rate, both evaluated along the nonlinear budget line. The main difference to a partial equilibrium setting is that, in general equilibrium, one needs to account for wage changes that affect labor supply which is the second part of the equation (see Sachs, Tsyvinski, and Werquin (2020)).

Using the Envelope theorem, one can decompose the migration response into two parts $\frac{\hat{N}_{i,j}}{N_{i,j}} = -\hat{T}_i\left(y_{i,j}\right)\eta_{i,j} + y_{i,j}\left(1 - T_i'\left(y_{i,j}\right)\right)\frac{\hat{w}_{i,\theta_j}}{w_{i,\theta_j}}\eta_{i,j}$. The first part is the migration response to a tax-induced reduction in the utility level in country i, as initially shown in Lehmann et al. (2014). The second part captures the migration response to a change in the individual's wage (*wage effect on migration*).

Wage responses. The latter effect captures the first interaction between general equilibrium responses and mobility. Now, we also describe the effect of labor migration on individual wages (*migration effect on wages*). Using the definitions of the own- and cross-wage elasticities one can write down worker type θ's wage response as

$$\frac{\hat{w}_{i,\theta}}{w_{i,\theta}} = \gamma_{i,\theta,\theta}\frac{\int_{\mathscr{I}_{i,\theta}} l_{i,j}N_{i,j}\omega_{i,j}\left(\frac{\hat{l}_{i,j}}{l_{i,j}} + \frac{\hat{N}_{i,j}}{N_{i,j}}\right)dj}{L_{i,\theta}} + \gamma_{i,\theta,\theta'}\frac{\int_{\mathscr{I}_{i,\theta'}} l_{i,j}N_{i,j}\omega_{i,j}\left(\frac{\hat{l}_{i,j}}{l_{i,j}} + \frac{\hat{N}_{i,j}}{N_{i,j}}\right)dj}{L_{i,\theta'}},$$

which, using the previously derived individual responses, reads as

$$\frac{\hat{w}_{i,\theta}}{w_{i,\theta}} = \sum_{\theta''=\theta,\theta'} \gamma_{i,\theta,\theta''} \int_{\mathscr{I}_{i,\theta''}} \frac{l_{i,j} N_{i,j} \omega_{i,j}}{L_{i,\theta''}} \left[-\varepsilon_{i,j}^{l,(1-T')} \frac{\hat{T}_i'(y_{i,j})}{1-T_i'(y_{i,j})} - \hat{T}_i(y_{i,j}) \eta_{i,j} \right] dj$$
$$+ \sum_{\theta''=\theta,\theta'} \gamma_{i,\theta,\theta''} \int_{\mathscr{I}_{i,\theta''}} \frac{l_{i,j} N_{i,j} \omega_{i,j}}{L_{i,\theta''}} \left[\varepsilon_{i,j}^{l,w} + y_{i,j} \left(1 - T_i'(y_{i,j})\right) \eta_{i,j} \right] \frac{\hat{w}_{i,\theta''}}{w_{i,\theta''}} dj$$

for $\theta = L, H$. Noting that, by the constant returns to scale in production, $\gamma_{i,\theta,\theta} = -\gamma_{i,\theta',\theta} \frac{w_{i,\theta'} L_{i,\theta'}}{w_{i,\theta} L_{i,\theta}}$, the wage response reads as

$$\frac{\hat{w}_{i,\theta}}{w_{i,\theta}} = -\frac{\sum_{\theta''=\theta,\theta'} \gamma_{i,\theta,\theta''} \int_{\mathscr{I}_{i,\theta''}} \frac{l_{i,j} N_{i,j} \omega_{i,j}}{L_{i,\theta''}} \left[\varepsilon_{i,j}^{l,(1-T')} \frac{\hat{T}_i'(y_{i,j})}{1-T_i'(y_{i,j})} + \hat{T}_i(y_{i,j}) \eta_{i,j} \right] dj}{1 - \sum_{\theta''=\theta,\theta'} \gamma_{i,\theta'',\theta''} \int_{\mathscr{I}_{i,\theta''}} \frac{l_{i,j} N_{i,j} \omega_{i,j}}{L_{i,\theta''}} \left[\varepsilon_{i,j}^{l,w} + y_{i,j} \left(1 - T_i'(y_{i,j})\right) \eta_{i,j} \right] dj}. \tag{14}$$

There are two migration effects on each type's wages. In the denominator, a rise in the semi-elasticity leads ceteris paribus to a reduction in the strength of the wage response. The response in the numerator is case-specific. Consider a tax reform that raises everyone's tax liability ($\hat{T}_i(y_{i,j}) > 0$, $\forall y_{i,j}$). Then, a rise in the migration semi-elasticity of type θ', say low-skilled workers who become more mobile, reduces high-skilled (type θ) workers' wages. Put differently, in the presence of labor complementarities, there is an indirect (fiscal) benefit of low-skilled immigration on high-skilled workers' wages, as described in Colas and Sachs (2020). Similarly, a rise in the migration semi-elasticity of θ-type workers reduces the their equilibrium population in country i and, thereby, tends to raise their wages due to the decreasing marginal product of labor.

Revenue and welfare effects. Having dealt with the individual-level incidence, we, now, turn to the aggregate revenue and welfare effects of tax reforms in general equilibrium with migration. First, we describe the partial equilibrium effects without any migration responses, e.g., described by Saez (2001),

$$\hat{\mathscr{R}}_i^{Saez} \equiv \int_{\mathscr{I}_i} \hat{T}_i(y_{i,j}) N_{i,j} dj - \int_{\mathscr{I}_i} y_{i,j} T_i'(y_{i,j}) \varepsilon_{i,j}^{l,(1-T')} \frac{\hat{T}_i'(y_{i,j})}{1-T_i'(y_{i,j})} N_{i,j} dj \tag{15}$$

and

$$\hat{\mathscr{G}}_i^{Saez} \equiv -\xi_i \int_{\mathscr{I}_i} \Gamma_{i,j} \hat{T}_i(y_{i,j}) N_{i,j} dj \tag{16}$$

that lead, in the partial equilibrium optimum ($\frac{1}{\xi_i} \hat{\mathscr{G}}_i^{Saez} + \hat{\mathscr{R}}_i^{Saez} = 0$), to the usual inverse elasticity rule of optimal income taxation. The first term in Equation (15) and the term in (16) are the mechanical changes in revenues and welfare to a rise in everyone's tax liability. The second term in (15) describes the behavioral responses triggered by the perturbation of marginal tax rates. There are no first-order behavioral effects on welfare.

Lehmann et al. (2014) (LST) find that there is an additional effect of tax reforms in partial equilibrium with migration

$$\hat{\mathscr{R}}_i^{LST} \equiv -\int_{\mathscr{I}_i} \hat{T}_i\left(y_{i,j}\right) T_i\left(y_{i,j}\right) \eta_{i,j} N_{i,j} dj \tag{17}$$

and

$$\hat{\mathscr{G}}_i^{LST} \equiv -\xi_i \int_{\mathscr{I}_i} \Gamma_{i,j} \hat{T}_i\left(y_{i,j}\right) \mathscr{U}_{i,j} \eta_{i,j} N_{i,j} dj, \tag{18}$$

accounting for the tax-induced migration response of each individual j. This negative partial equilibrium response is proportional to the semi-elasticity of migration and leads to the well-known result of lower marginal tax rates in the presence of migration. As we show here, this "migration threat" channel is also present in our framework.

As shown by Sachs et al. (2020), in general equilibrium, each workers' wage responses augment the revenue and welfare effects of tax reforms. In our two-type Stiglitz (1982) setting, this effect reads as

$$\hat{\mathscr{R}}_i^{Stiglitz} \equiv \sum_{\theta=L,H} \int_{\mathscr{I}_{i,\theta}} y_{i,j} T_i'\left(y_{i,j}\right) \left(1+\varepsilon_{i,j}^{l,w}\right) \left(\frac{\hat{w}_{i,\theta}}{w_{i,\theta}}\right)^{Stiglitz} N_{i,j} dj \tag{19}$$

and

$$\hat{\mathscr{G}}_i^{Stiglitz} \equiv \sum_{\theta=L,H} \xi_i \int_{\mathscr{I}_{i,\theta}} \Gamma_{i,j} y_{i,j} \left(1-T_i'\left(y_{i,j}\right)\right) \left(\frac{\hat{w}_{i,\theta}}{w_{i,\theta}}\right)^{Stiglitz} N_{i,j} dj, \tag{20}$$

where

$$\left(\frac{\hat{w}_{i,\theta}}{w_{i,\theta}}\right)^{Stiglitz} \equiv -\frac{\sum_{\theta''=\theta,\theta'} \gamma_{i,\theta,\theta''} \int_{\mathscr{I}_{i,\theta''}} \frac{l_{i,j} N_{i,j} \omega_{i,j}}{L_{i,\theta''}} \varepsilon_{i,j}^{l,(1-T')} \frac{\hat{T}_i'\left(y_{i,j}\right)}{1-T_i'\left(y_{i,j}\right)} dj}{1-\sum_{\theta''=\theta,\theta'} \gamma_{i,\theta'',\theta''} \int_{\mathscr{I}_{i,\theta''}} \frac{l_{i,j} N_{i,j} \omega_{i,j}}{L_{i,\theta''}} \varepsilon_{i,j}^{l,w} dj}$$

describes workers' general equilibrium wage responses absent of migration. These result from labor supply externalities between worker types. On the one hand, a rise in the marginal tax rate of θ-workers (e.g., high-skilled workers) reduces their labor supply, which raises their wages. On the other hand, the tax-induced reduction in high-skilled workers' hours lowers low-skilled wages. Similarly, in response to a tax cut at the bottom, high-skilled workers experience a wage increase, whereas low-skilled workers' wages go down. Stiglitz (1982) discovered this trickle-down effect that leads, in the optimum, to a lower tax progressivity. More recently, Sachs et al. (2020) show that, in this setting, the aggregate effects (e.g., on tax revenues) of changing tax progressivity become less clear when the initial tax code is progressive.

The novelty of this paper is to consider the tax policy implications of migration in general equilibrium. In the presence of migration and general equilibrium wage responses, there are two novel effects on revenues

$$\hat{\mathscr{R}}_i^{new} = \hat{\mathscr{R}}_i^{new,W\to M} + \hat{\mathscr{R}}_i^{new,M\to W} \tag{21}$$

and welfare

$$\hat{\mathscr{G}}_i^{new} = \hat{\mathscr{G}}_i^{new,W\to M} + \hat{\mathscr{G}}_i^{new,M\to W}, \tag{22}$$

where

$$\hat{\mathscr{R}}_i^{new,W\to M} \equiv \sum_{\theta=L,H} \int_{\mathscr{I}_{i,\theta}} T_i\left(y_{i,j}\right) y_{i,j} \left(1-T_i'\left(y_{i,j}\right)\right) \frac{\hat{w}_{i,\theta}}{w_{i,\theta}} \eta_{i,j} N_{i,j} dj$$

and

$$\hat{\mathscr{G}}_i^{new,W\to M} \equiv \sum_{\theta=L,H} \xi_i \int_{\mathscr{I}_{i,\theta}} \Gamma_{i,j} \mathscr{U}_{i,j} y_{i,j} \left(1-T_i'\left(y_{i,j}\right)\right) \frac{\hat{w}_{i,\theta}}{w_{i,\theta}} \eta_{i,j} N_{i,j} dj$$

describe the aggregate impact of endogenous wages on labor mobility (*aggregate wage effect on migration*).

$$\hat{\mathscr{R}}_i^{new,M\to W} \equiv \sum_{\theta=L,H} \int_{\mathscr{I}_{i,\theta}} y_{i,j} T_i'\left(y_{i,j}\right) \left(1+\varepsilon_{i,j}^{l,w}\right) \left[\frac{\hat{w}_{i,\theta}}{w_{i,\theta}} - \left(\frac{\hat{w}_{i,\theta}}{w_{i,\theta}}\right)^{Stiglitz}\right] N_{i,j} dj$$

and

$$\hat{\mathscr{G}}_i^{new,M\to W} \equiv \sum_{\theta=L,H} \xi_i \int_{\mathscr{I}_{i,\theta}} \Gamma_{i,j} y_{i,j} \left(1-T_i'\left(y_{i,j}\right)\right) \left[\frac{\hat{w}_{i,\theta}}{w_{i,\theta}} - \left(\frac{\hat{w}_{i,\theta}}{w_{i,\theta}}\right)^{Stiglitz}\right] N_{i,j} dj$$

measure the effect of migration on equilibrium wages (*aggregate migration effect on wages*). These effects demonstrate the interaction between mobility and wage endogeneity. Observe that either without migration ($\eta_{i,j} = 0,\ \forall j \in \mathscr{I}_i$) or without general equilibrium responses ($\gamma_{i,\theta,\theta} = \gamma_{i,\theta,\theta'} = 0,\ \forall \theta, \theta' \in \{L,H\}$) these interactions vanish ($\hat{\mathscr{R}}_i^{new} = 0$ and $\hat{\mathscr{G}}_i^{new} = 0$).

Intuition for novel effects. The aggregate wage effect on migration is the aggregate effect of migration responses that come from wage adjustments. Consider tax revenues and suppose that a tax reform, say a rise in tax progressivity, raises low-skilled wages. This induces an inflow of low-skilled workers into the country and raises tax payments by these workers. At the same time, when the reform reduces high-skilled workers' wages, they emigrate out of the country, reducing tax revenues there.

The aggregate migration effect on wages collects the impact of all migration-induced wage responses. These responses depend on the nature of the tax reform. Consider a rise in the tax liability of high-skilled workers and a tax cut for low-skilled workers. This leads to an inflow of low-skilled workers and an outflow of high-skilled ones. As a result, wages adjust. By diminishing marginal products, the low-skilled workers' immigration tends to reduce low-skilled wages and, by the complementarity of labor, raises high-skilled workers' wages (also see Colas and Sachs (2020)). Simultaneously, the loss of high-skilled workers in the country increases the wages of high types and reduces those of low ones.

References

Acemoglu, D., and Autor, D. (2011). "Skills, tasks and technologies: Implications for employment and earnings." *Handbook of Labor Economics*, *4*, 1043–1171.

Ales, L., Kurnaz, M., and Sleet, C. (2015). "Technical change, wage inequality, and taxes." *American Economic Review*, *105*(10), 3061–3101.

Colas, M., and Sachs, D. (2020). "The indirect fiscal benefits of low-skilled immigration." *CESifo Working Paper No. 8604*.

Golosov, M., Tsyvinski, A., and Werquin, N. (2014). "A variational approach to the analysis of tax systems." *National Bureau of Economic Research, Working Paper No. 20780*.

Janeba, E., and Schulz, K. (2021). "Nonlinear taxation and international mobility in general equilibrium." *Working Paper*.

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

Milgrom, P., and Shannon, C. (1994). "Monotone comparative statics." *Econometrica*, 157–180.

Sachs, D., Tsyvinski, A., and Werquin, N. (2020). "Nonlinear tax incidence and optimal taxation in general equilibrium." *Econometrica*, *88*(2), 469–493.

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

Stiglitz, J. E. (1982). "Self-selection and pareto efficient taxation." *Journal of Public Economics*, *17*(2), 213–240.