

Online Appendix

A Details and Proofs for the Simple Model

A.1 Definition of Equilibrium

Definition 1. *Given an initial capital stock K_0 , an exogenous population $\{N_t(0, s, i), N_t(1, s, i)\}$ and government policy $\{\rho_t\}$ an equilibrium is a sequence of allocations and prices such that*

1. *Given wages $w_t(s, i)$, interest rates R_{t+1} and policies $\tau_t, b_{t+1}(s, i)$ for each t and each type (s, i) the allocation $c_t(0, s, i), c_{t+1}(1, s, i), a_{t+1}(s, i)$ maximizes lifetime utility (6) subject to the budget constraints (5).*
2. *Interest rates and wages (R_t, w_t) satisfy the marginal product pricing equations (11) and (11), and type-specific wages are given by (13).*
3. *Government policies satisfy the budget constraint (10).*
4. *Markets clear:*

(a) *Labor Markets*

$$L_t(hi) = N_t(0, hi, na) \tag{49}$$

$$L_t(lo, i) = \epsilon(lo, i)N_t(0, lo, i) \quad \text{for } i \in \{na, fo\} \tag{50}$$

(b) *Capital Market*

$$K_{t+1} = s_t w_t L_t \tag{51}$$

(c) *Goods Market*

$$C_t + K_{t+1} = K_t^\alpha L_t^{1-\alpha} \tag{52}$$

Equilibrium in the small open economy is defined in a similar fashion, but the capital market clearing condition is replaced by the condition that the real interest rate $R_t = R$ is fixed by the world capital market, which then from the firm's optimality conditions pins down the constant wage $w_t = w(R)$ and capital-labor ratio $k_t = k(R)$.

A.2 Relative Wages as Functions of Demographics

We summarize wages as functions of demographic variables as:

$$\begin{aligned}
\frac{L_t}{L_t(hi)} &= \frac{\left(L_t(lo)^{1-\frac{1}{\sigma_{lh}}} + L_t(hi)^{1-\frac{1}{\sigma_{lh}}}\right)^{\frac{1}{1-\frac{1}{\sigma_{lh}}}}}{L_t(hi)} = \left(\left(\frac{L_t(lo)}{L_t(hi)}\right)^{1-\frac{1}{\sigma_{lh}}} + 1\right)^{\frac{1}{1-\frac{1}{\sigma_{lh}}}} \\
\frac{L_t}{L_t(lo)} &= \left(\left(\frac{L_t(hi)}{L_t(lo)}\right)^{1-\frac{1}{\sigma_{lh}}} + 1\right)^{\frac{1}{1-\frac{1}{\sigma_{lh}}}} \\
\frac{L_t(lo)}{L_t(hi)} &= \frac{L_t(lo)}{L_t(lo, fo)} \cdot \frac{L_t(lo, fo)}{L_t(hi)} = \left(\left(\frac{\epsilon(lo, na)(1-\omega)\gamma_t^n}{\epsilon(lo, fo)\mu_t}\right)^{1-\frac{1}{\sigma_{nf}}} + 1\right)^{\frac{1}{1-\frac{1}{\sigma_{nf}}}} \cdot \frac{\epsilon(lo, fo)\mu_t}{\omega\gamma_t^n} \\
\frac{L_t(hi)}{L_t(lo)} &= \frac{L_t(hi)}{L_t(lo, fo)} \cdot \frac{L_t(lo, fo)}{L_t(lo)} = \left(\left(\frac{\epsilon(lo, na)(1-\omega)\gamma_t^n}{\epsilon(lo, fo)\mu_t}\right)^{1-\frac{1}{\sigma_{nf}}} + 1\right)^{\frac{-1}{1-\frac{1}{\sigma_{nf}}}} \cdot \frac{\omega\gamma_t^n}{\epsilon(lo, fo)\mu_t} \\
\frac{L_t(lo)}{L_t(lo, fo)} &= \frac{\left(L_t(lo, na)^{1-\frac{1}{\sigma_{nf}}} + L_t(lo, fo)^{1-\frac{1}{\sigma_{nf}}}\right)^{\frac{1}{1-\frac{1}{\sigma_{nf}}}}}{L_t(lo, na)} \\
&= \left(\left(\frac{\epsilon(lo, na)(1-\omega)\gamma_t^n}{\epsilon(lo, fo)\mu_t}\right)^{1-\frac{1}{\sigma_{nf}}} + 1\right)^{\frac{1}{1-\frac{1}{\sigma_{nf}}}} \\
\frac{L_t(lo)}{L_t(lo, na)} &= \frac{\left(L_t(lo, na)^{1-\frac{1}{\sigma_{nf}}} + L_t(lo, fo)^{1-\frac{1}{\sigma_{nf}}}\right)^{\frac{1}{1-\frac{1}{\sigma_{nf}}}}}{L_t(lo, na)} \\
&= \left(\left(\frac{\epsilon(lo, fo)\mu_t}{\epsilon(lo, na)(1-\omega)\gamma_t^n}\right)^{1-\frac{1}{\sigma_{nf}}} + 1\right)^{\frac{1}{1-\frac{1}{\sigma_{nf}}}}
\end{aligned}$$

$$\begin{aligned}
w_t(hi) &= w_t \cdot \left(\frac{L_t}{L_t(hi)} \right)^{\frac{1}{\sigma_{lh}}} = w_t \cdot \left(\left(\frac{L_t(lo)}{L_t(hi)} \right)^{1-\frac{1}{\sigma_{lh}}} + 1 \right)^{\frac{\frac{1}{\sigma_{lh}}}{1-\frac{1}{\sigma_{lh}}}} \\
&= w_t \cdot \left(\left(\left(\left(\frac{\epsilon(lo, na)(1-\omega)\gamma_t^n}{\epsilon(lo, fo)\mu_t} \right)^{1-\frac{1}{\sigma_{nf}}} + 1 \right)^{\frac{1}{1-\frac{1}{\sigma_{nf}}}} \cdot \frac{\epsilon(lo, fo)\mu_t}{\omega\gamma_t^n} \right)^{1-\frac{1}{\sigma_{lh}}} + 1 \right)^{\frac{1}{\sigma_{lh}-1}} \\
&= w_t \cdot \left(\left(\left(\left(\frac{\epsilon(lo, na)(1-\omega)}{\omega} \right)^{1-\frac{1}{\sigma_{nf}}} + \left(\frac{\epsilon(lo, fo)\mu_t}{\omega\gamma_t^n} \right)^{1-\frac{1}{\sigma_{nf}}} \right)^{\frac{1}{1-\frac{1}{\sigma_{nf}}}} \right)^{1-\frac{1}{\sigma_{lh}}} + 1 \right)^{\frac{1}{\sigma_{lh}-1}} \\
&= w_t \cdot \mathcal{W}_{hi}(\mu_t/\gamma_t^n) \\
w_t(lo, na) &= w_t \cdot \epsilon(lo, na) \cdot \left(\frac{L_t}{L_t(lo)} \right)^{\frac{1}{\sigma_{lh}}} \cdot \left(\frac{L_t(lo)}{L_t(lo, na)} \right)^{\frac{1}{\sigma_{nf}}} \\
&= w_t \cdot \epsilon(lo, na) \\
&\quad \cdot \left(\left(\left(\left(\frac{\epsilon(lo, na)(1-\omega)}{\omega} \right)^{1-\frac{1}{\sigma_{nf}}} + \left(\frac{\epsilon(lo, fo)\mu_t}{\omega\gamma_t^n} \right)^{1-\frac{1}{\sigma_{nf}}} \right)^{\frac{-1}{1-\frac{1}{\sigma_{nf}}}} \right)^{1-\frac{1}{\sigma_{lh}}} + 1 \right)^{\frac{1}{\sigma_{lh}-1}} \\
&\quad \cdot \left(\left(\frac{\epsilon(lo, fo)\mu_t}{\epsilon(lo, na)(1-\omega)\gamma_t^n} \right)^{1-\frac{1}{\sigma_{nf}}} + 1 \right)^{\frac{1}{\sigma_{nf}-1}} \\
&= w_t \mathcal{W}_{lo}(\mu_t/\gamma_t^n) \cdot \mathcal{W}_{na}(\mu_t/\gamma_t^n)
\end{aligned}$$

It follows from direct inspection that $\mathcal{W}_{hi}(\mu_t/\gamma_t^n)$, $\mathcal{W}_{na}(\mu_t/\gamma_t^n)$ are strictly increasing in μ_t/γ_t^n and $\mathcal{W}_{lo}(\mu_t/\gamma_t^n)$ is strictly decreasing in μ_t/γ_t^n .

A.3 Proof of Lemma 1 and Theorem 1

For lemma 1, we want to arrive at an expression for $\gamma_{t+1}^L = \frac{L_{t+1}}{L_t}$. Recall from (3) and (4) that

$$\begin{aligned}
L_t &= \left(L_t(lo)^{1-\frac{1}{\sigma_{lh}}} + L_t(hi)^{1-\frac{1}{\sigma_{lh}}} \right)^{\frac{1}{1-\frac{1}{\sigma_{lh}}}} \\
L_t(lo) &= \left(L_t(lo, na)^{1-\frac{1}{\sigma_{nf}}} + L_t(lo, fo)^{1-\frac{1}{\sigma_{nf}}} \right)^{\frac{1}{1-\frac{1}{\sigma_{nf}}}}
\end{aligned}$$

Work on (4):

$$\begin{aligned}
L_t(lo) &= \left((\epsilon(lo, na)N_t(0, lo, na))^{1-\frac{1}{\sigma_{nf}}} + (\epsilon(lo, fo)N_t(0, lo, fo))^{1-\frac{1}{\sigma_{nf}}} \right)^{\frac{1}{1-\frac{1}{\sigma_{nf}}}} \\
&= \left((\epsilon(lo, na)(1-\omega)\gamma_t^n N_{t-1}(0))^{1-\frac{1}{\sigma_{nf}}} + (\epsilon(lo, fo)\mu_t\gamma_t^n N_{t-1}(0))^{1-\frac{1}{\sigma_{nf}}} \right)^{\frac{1}{1-\frac{1}{\sigma_{nf}}}} \\
&= \left((\epsilon(lo, na)(1-\omega)\gamma_t^n)^{1-\frac{1}{\sigma_{nf}}} + (\epsilon(lo, fo)\mu_t\gamma_t^n)^{1-\frac{1}{\sigma_{nf}}} \right)^{\frac{1}{1-\frac{1}{\sigma_{nf}}}} N_{t-1}(0) \\
&= \Lambda(\gamma_t^n, \mu_t)N_{t-1}(0) = \Lambda_t N_{t-1}(0)
\end{aligned}$$

Use this in (3) to get

$$\begin{aligned}
L_t &= \left((\Lambda(\cdot)N_{t-1}(0))^{1-\frac{1}{\sigma_{lh}}} + (\omega\gamma_t^n N_{t-1}(0))^{1-\frac{1}{\sigma_{lh}}} \right)^{\frac{1}{1-\frac{1}{\sigma_{lh}}}} \\
&= \left((\Lambda(\cdot))^{1-\frac{1}{\sigma_{lh}}} + (\omega\gamma_t^n)^{1-\frac{1}{\sigma_{lh}}} \right)^{\frac{1}{1-\frac{1}{\sigma_{lh}}}} N_{t-1}(0) \\
&= \Omega(\Lambda(\gamma_t^n, \mu_t), \gamma_t^n)N_{t-1}(0) = \Omega_t(\Lambda_t, \gamma_t^n)N_{t-1}(0).
\end{aligned}$$

Thus we get

$$\begin{aligned}
\gamma_{t+1}^L &= \frac{\Omega(\Lambda(\gamma_{t+1}^n, \mu_{t+1}), \gamma_{t+1}^n)N_t(0)}{\Omega(\Lambda(\gamma_t^n, \mu_t), \gamma_t^n)N_{t-1}(0)} \\
&= \frac{\Omega(\Lambda(\gamma_{t+1}^n, \mu_{t+1}), \gamma_{t+1}^n)}{\Omega(\Lambda(\gamma_t^n, \mu_t), \gamma_t^n)} \gamma_t \\
&= \frac{\Omega(\Lambda(\gamma_{t+1}^n, \mu_{t+1}), \gamma_{t+1}^n)}{\Omega(\Lambda(\gamma_t^n, \mu_t), \gamma_t^n)} (\gamma_t^n + \mu_t) \\
&= \frac{\Omega_{t+1}}{\Omega_t} (\gamma_t^n + \mu_t) \\
&= \gamma_t^n + \mu_t \text{ if } \gamma_{t+1}^n = \gamma_t^n, \text{ and } \mu_{t+1} = \mu_t
\end{aligned}$$

We make the following:

- Observation 1.**
1. Fix γ^n and consider a permanent change of μ_t from $\mu^l > 0$ to $\mu^h > \mu^l$ in period t . Since $\Lambda_{t+1} = \Lambda_t$ we have $\Omega_{t+1} = \Omega_t$ and thus γ_{t+1}^L jumps to $\gamma^n + \mu^l$.
 2. Fix μ and consider a permanent change of γ^n from $\gamma^{nl} > 0$ to $\gamma^{nh} > \gamma^{nl}$ in period t . Since $\Lambda_{t+1} = \Lambda_t$ and $\Omega_{t+1}(\Lambda_{t+1}, \gamma^{nh}) = \Omega_t(\Lambda_t, \gamma^{nh})$ we have that γ_{t+1}^L jumps to $\gamma^{nh} + \mu$.

The proof of theorem 1 then follows directly from lemma 1 as well as propositions 2 and 3. The only non-trivial part is to sign the general equilibrium effect. For this note that

$$\begin{aligned}
(1 + \beta) \ln(w_t) + \beta \ln(R_{t+1}) &= (1 + \beta) \ln((1 - \alpha)k_t^\alpha) + \beta \ln(\alpha k_{t+1}^{\alpha-1}) \\
&= v + (1 + \beta)\alpha \ln(k_t) - (1 - \alpha)\beta \ln(k_{t+1}) \\
&= v + (1 + \beta)\alpha \ln(k_t) - (1 - \alpha)\beta [\ln(s_t) + \alpha \ln(k_t) - \ln(\gamma_{t+1}^L)] \\
&= v + \alpha(1 + \alpha\beta) \ln(k_t) - (1 - \alpha)\beta [\ln(s_t) - \ln(\gamma_{t+1}^L)] \\
&= v + \alpha(1 + \alpha\beta) \ln(K_t) - (1 - \alpha)\beta \ln(s_t) \\
&\quad - [\alpha(1 + \alpha\beta) + (1 - \alpha)\beta] \ln(L_t) + (1 - \alpha)\beta \ln(L_{t+1}).
\end{aligned}$$

where v is a constant. The period t capital stock K_t is pre-determined. The saving rate s_t increases with the per-capita immigration cost κ_{t+1} which in turn rises as more migrants come in, increasing the capital-labor ratio in period $t + 1$ and thus reducing the real interest rate. This is the first negative general equilibrium effect (which would be absent if there are no resource costs for the newly arriving migrants, i.e. if $\kappa_{t+1} = 0$). Second, both L_t as well as $\gamma_{t+1}^L = L_{t+1}/L_t$ increase when μ_t increases permanently. As long as α is sufficiently large (trivially, if $\alpha = 1$), or as long as $\frac{\partial \ln(L_t)}{\partial \mu} \approx \frac{\partial \ln(L_{t+1})}{\partial \mu}$ (both of these terms only depend on model-exogenous variables) the general equilibrium effect of a permanent increase in migration flows is negative.

Remark 1. *Also note that in the absence of a resource cost ($\kappa_{t+1} = 0$) the saving rate is invariant to demographics, and an increase in migration triggers a decline in the current capital-labor ratio k_t and a further decline in future capita-labor ratios k_{t+s} through the permanent increase in the growth rate of labor γ_{t+s}^L . In the long-run the GE effect of these declines is negative as long as the economy remains dynamically efficient. To see this, observe that for all $t \geq 1$ the welfare difference along the transition is*

$$\begin{aligned}
\Delta [(1 + \beta) \ln(w_t) + \beta \ln(R_{t+1})] &= (1 + \beta) (\ln(w_t) - w_0) + \beta (\ln(R_{t+1}) - R_0) \\
&= (1 + \beta)\alpha (\ln(k_t) - \ln(k_0)) - (1 - \alpha)\beta (\ln(k_{t+1}) - \ln(k_0)).
\end{aligned}$$

For $t \rightarrow \infty$ this term is negative if

$$\frac{\alpha}{1 - \alpha} > \frac{\beta}{1 + \beta}.$$

It is straightforward to verify from the corresponding social planner's problem that this is the condition for dynamic efficiency of the economy. In the short run, for the period $t = 1$ old generation the effect is positive because the wage effect is absent. For all newborn generations

along the transition, the effect is negative if the capital share α is sufficiently large—notice that dynamic efficiency is thus only a necessary condition for the effect to be negative for all newborns along the transition—, because for all $t \geq 1$ the welfare change is negative if

$$(1 + \beta)\alpha |(\ln(k_t) - \ln(k_0))| - (1 - \alpha)\beta |(\ln(k_{t+1}) - \ln(k_0))| > 0$$

$$\Leftrightarrow \frac{\alpha}{1 - \alpha} \frac{|(\ln(k_t) - \ln(k_0))|}{|(\ln(k_{t+1}) - \ln(k_0))|} > \frac{\beta}{1 + \beta}.$$

and by the monotonic decline of the capital stock we know that $\frac{|(\ln(k_t) - \ln(k_0))|}{|(\ln(k_{t+1}) - \ln(k_0))|} < 1$.

B Quantitative Model Appendix

B.1 Assimilation Flows

We construct net migration numbers at the net addition to the population stock from migration flows in the next period, $M_{t+1}(j + 1, i, g)$, from which we then compute the migration rates $\mu_t(j, as, g) = \frac{M_{t+1}(j+1, i, g)}{N_t(j, i, g)}$. Denoting by $M_{t+1}^f(j + 1, as, g)$ the inflow from foreign countries to the group of asylum seekers, the net immigration flow to group as is

$$M_{t+1}(j + 1, as, g) = M_{t+1}^f(j + 1, as, g) - (\pi^l + (1 - \pi^l)\pi^{ar}) N_t(j, as, g)\psi_t(j, as, g)$$

and therefore

$$\mu_t(j, as, g) = \mu_t^f(j, as, g) - (\pi^l + (1 - \pi^l)\pi^{ar}) \psi_t(j, as, g).$$

Denoting by $M_{t+1}^f(j + 1, rw, g)$ the inflow from foreign countries to population group rw , the net inflow to the population group rw is

$$M_{t+1}(j + 1, rw, g) = M_{t+1}^f(j + 1, rw, g) + (1 - \pi^l)\pi^{ar}\psi_t(j, as, g)N_t(j, as, g) - \pi^{rh}\psi_t(j, rw, g)N_t(j, rw, g)$$

and thus

$$\mu_t(j, rw, g) = \mu_t^f(j, rw, g) + (1 - \pi^l)\pi^{ar}\psi_t(j, as, g)\frac{N_t(j, as, g)}{N_t(j, rw, g)} - \pi^{rh}\psi_t(j, rw, g).$$

Correspondingly, denoting by $M_{t+1}^f(j+1, ho, g)$ the inflow from foreign countries to population group ho , the total inflow to population group ho is

$$M_{t+1}(j+1, ho, g) = M_{t+1}^f(j+1, ho, g) + \pi^{rh}\psi_t(j, rw, g)N_t(j, rw, g)$$

and thus

$$\mu_t(j, ho, g) = \mu_t^f(j, ho, g) + \pi^{rh}\psi_t(j, rw, g)\frac{N_t(j, rw, g)}{N_t(j, ho, g)}.$$

B.2 First-Order Conditions of Firm Problem

Denote by $k_t = \frac{K_t}{A_t L_t}$ the ‘‘capital intensity’’, respectively the capital stock per efficiency unit of labor. Then, the first-order conditions of the static firm problem are given by

$$r_t = \alpha k_t^{-\frac{1}{\vartheta}} \left(\alpha k_t^{1-\frac{1}{\vartheta}} + (1-\alpha) \right)^{\frac{1}{1-\frac{1}{\vartheta}}} - \delta \quad (53a)$$

$$w_t = A_t(1-\alpha) \left(\alpha k_t^{1-\frac{1}{\vartheta}} + (1-\alpha) \right)^{\frac{1}{1-\frac{1}{\vartheta}}} \quad (53b)$$

$$w_t(s) = w_t \left(\frac{L_t}{L_t(s)} \right)^{\frac{1}{\sigma_{lmh}}} \quad (53c)$$

$$w_t(\bar{j}, s) = w_t(s) \quad (53d)$$

$$w_t(\bar{j}, s, na) = w_t(\bar{j}, s) \left(\frac{L_t(\bar{j}, s)}{L_t(\bar{j}, s, na)} \right)^{\frac{1}{\sigma_{nf}}}, \quad (53e)$$

$$\tilde{w}_t(\bar{j}, s, fo) = w_t(\bar{j}, s) \left(\frac{L_t(\bar{j}, s)}{\tilde{L}_t(\bar{j}, s, fo)} \right)^{\frac{1}{\sigma_{nf}}}, \quad (53f)$$

$$w_t(\bar{j}, s, ho) = \tilde{w}_t(\bar{j}, s, fo) \left(\frac{\tilde{L}_t(\bar{j}, s, fo)}{L_t(\bar{j}, s, ho)} \right)^{\frac{1}{\sigma_{hr}}} \quad (53g)$$

$$w_t(\bar{j}, s, o) = \tilde{w}_t(\bar{j}, s, fo) \left(\frac{\tilde{L}_t(\bar{j}, s, fo)}{\sum_{o \in \{rw, as\}} L_t(s, o)} \right)^{\frac{1}{\sigma_{hr}}}, \text{ for } o \in \{rw, as\}. \quad (53h)$$

We then get the age j , skill s , nationality i -specific aggregate wage component $w_t(j, s, i) = w_t(\bar{j}, s, i)$ if $j \in [j_l(\bar{j}), \dots, j_h(\bar{j})]$.

B.3 Annuity Income Stream of Leavers

Total wealth of a leaver includes the value of assets at the end of period t at age j net of fraction π^c of confiscated assets by the government of the country the leaver remigrates to,

a one time lump-sum payment by the German government b_t^l , and the discounted value of future labor income. We assume that in the country a household remigrates to it works full-time, $l = l_n$, does not pay or receive any transfers from a social insurance institution, and retires exogenously at age j_r . Accidental bequests are taxed at a confiscatory rate. We compute the continuation value in a partial equilibrium, taking the current period wage w_t and the interest rate in the period of leaving r_t as given.

Denote by a_t^l the savings of a leaver during the leaving period, i.e, in the last period in Germany. Initial assets after confiscation at the beginning of period $t + 1$ in the country the leaver migrates to are $a_{t+1} = (1 - \pi^c)a_t^l$. Total wealth of a leaver with education s and gender g in period $t + 1$, age $j + 1$, is accordingly given by

$$W_{t+1}(j + 1, s, g) = a_{t+1}(1 + r_t) + b_t^l + \eta \cdot \sum_{p=j+1}^{j_r-1} \left(\frac{1}{1 + r_t} \right)^{p-(j+1)} \epsilon(p, s, i, g) w_t(p, s, g) l_n$$

where $\eta \in (0, 1)$ is a productivity scaling parameter, reflecting lower productivity in the respective country as well as linear labor income taxes. The according annuity stream is

$$y_{t+1}^a(s, g) = \frac{r_t}{1 + r_t} \frac{(1 + r_t)^{J-j}}{(1 + r_t)^{J-j} - 1} W_{t+1}(j + 1, s, g). \quad (54)$$

B.4 Gains and Loss Term

For consumption equivalent variation of a cohort born in period $t - j$ for period t state variables age j , education s , nationality i , gender g , asset holdings a , denoted by $g_{t-j}^c(j, s, i, g, a)$ and corresponding cross-sectional $\Phi_t(j, s, i, g, a)$ in the baseline demographic scenario, we compute the average consumption equivalent variation

$$g_{t-j}^c(j, s, i, g) = \int g_{t-j}^c(j, s, i, g, a) \Phi_t(j, s, i, g, da). \quad (55)$$

For period 2013 we compute the above CEV for all cohorts $t - j$, ages $j = 0, \dots, J$, and consider the actual asset position and employment state in period 2013. For cohorts born after 2013 we evaluate the CEV at $j = 0, a = 0$, only. The period t consumption of the respective group given the consumption policy function in the baseline demographic scenario $c_t(j, s, i, g, a)$ in turn is

$$c_t(j, s, i, g) = \int c_t(j, s, i, g, a) \Phi_t(j, s, i, g, da). \quad (56)$$

These objects form the basis of the computation of the net gain term in equation (47).

B.5 Recursive Household Problem

State Variables. We collect state variables as follows, also see Table 1: age $j \in \{j_a, \dots, J\}$, education $s \in \{lo, me, hi\}$, economic nationality $i \in \{na, ho, rw, as\}$, gender $g \in \{fe, ma\}$, employment status $e \in \{em, re\}$, and assets $a \in \mathcal{A}$.

For asylum seekers the problem is slightly more complex because of the leaving shock and the assimilation shock. Also, immigrants from the rest of the world face an assimilation shock. We therefore first describe the problems of groups $i \in \{na, ho\}$ and then turn to relevant extensions for the remaining two population groups.

Dynamic Problem of Retired Households, $j \in \{j_r, \dots, J\}, i \in \{na, ho\}, e = re$. Retired agents solve³⁷

$$V_t(j, s, i, g, e, a) = \max_{c, a'} \left\{ U \left(\frac{c}{1 + \zeta n}, 1 \right) + \beta \psi_t(j, i, g) V_{t+1}(j + 1, s, i, g, e, a') \right\}$$

subject to

$$\begin{aligned} a' &= (a + tr_t)(1 + r_t(1 - \tau_t^k)) + y_t^p - (c(1 + \tau_t^c) + T_t(y_t^p)) \geq 0 \\ y_t^p &= (1 - \tau_t^h) b_t^p(s, i, g). \end{aligned}$$

Dynamic Problem of Working Households in Last Working Period, $j = j_r - 1, i \in \{na, ho\}, e = em$. In the last period of work, $j_r - 1$, the value function is the expected value of the maximized value functions of the discrete choice specific value functions $J(\cdot, l_i)$ from working $l_i \in \{l_1, \dots, l_n\}$ hours, which is, given the Gumbel distributed taste shocks ε with scale parameter ς ,

$$V_t(j, s, i, g, e = em, a) = \varsigma \log \left[\sum_{k=1}^n \exp \left\{ \frac{J_t(\cdot, l_k)}{\varsigma} \right\} \right]$$

with according choice probabilities for alternative k

$$\pi_t(j, s, i, g, e = em, a, l = l_k) = \frac{\exp \left(\frac{J_t(\cdot, l_k)}{\varsigma} \right)}{\sum_{m=1}^n \exp \left(\frac{J_t(\cdot, l_m)}{\varsigma} \right)}.$$

³⁷Recall that $\psi_t J, i, g = 0$ so that terminal (and trivial) decision problem of singles and couples at age J are nested in this description.

where $J_t(\cdot, l_k)$ is the choice specific value function for working $l_k \in \{l_1, \dots, l_n\}$ hours

$$J_t(j, s, i, g, e = em, a, l = l_k) = \max_{c, a'} \{U(c, 1 - l_k) + \beta \psi_t(j, i, g) V_{t+1}(j + 1, s, i, g, e = re, a')\}$$

subject to

$$\begin{aligned} a' &= (a + tr_t)(1 + r_t(1 - \tau_t^k)) + y_t - c(1 + \tau_t^c) - T_t(y_t) \geq 0 \\ y_t &= (1 - \tau_t^p - \tau_t^h) w_t(j, s, i) \epsilon(j, s, i, g) l_k \end{aligned}$$

Dynamic Problem of Working Households in Working Period $j \in \{j_s, \dots, j_r - 2\}, i \in \{n, h, r\}, e = em$. The structure is the same as previously, where continuation values at t, j are the value functions $V_{t+1}(j + 1, s, i, g, e = em, a')$.

Dynamic Problem of Households $i \in \{na, ho\}, j \in \{j_a, \dots, j_s\}, e = em$. The dynamic problem is the same as described above, but the current period labor productivity is further shifted by factor $\varrho(s) \in (0, 1)$.

Modifications for Asylees, $i = as$. Due to differences in access to the social insurance system and transfer payments to asylees as well as labor market restrictions, the problem of asylees in the first year of entry is different from other years, which we store in indicator $\mathbb{1}_a$. At the end of each period conditional on surviving asylum seekers face the risk of having to leave with respective probability π^l and, conditional on not leaving, they may assimilate to population group rw with probability π^{ar} , thus the unconditional probability of assimilating to group rw is $(1 - \pi^l)\pi^{ar}$ and the unconditional probability of staying in population group as is $(1 - \pi^l)(1 - \pi^{ar})$. For the remainder of the description we focus on asylum seekers during the working period and spell out later the adjustments needed for other stages of the life-cycle.

First, we compute the continuation value in case of leaving. An asylee being forced to leave at age $j + 1$ receives in each period a permanent income stream $y^a(s, g)$, which we compute for both partners in a couple according to equation (54). The household enjoys flow utility from consumption of the annuity in each period and is assumed to work full-time, $U(y^a(s, g), 1 - l_n)$, and thus the value function in case of being forced to leave can be

computed recursively as

$$V^l(j+1, s, i = as, a') = U\left(\frac{y^a(s, g)}{1 + \zeta n + \xi}, 1 - l_n\right) + \beta \psi_{t+1}(j+1, as, g) V^l(j+2, s, i = as, a'')$$

subject to

$$a'' = a'(1 + r) + \mathbb{1}_{j \leq j_r - 1} \cdot \eta \cdot w_t(j+1, s, as, g) l_n - y^a(s, g)$$

where indicator $\mathbb{1}_{j \leq j_r - 1}$ is equal to one if the household is of working age $j \leq j_r - 1$, r is the return on assets in the period of leaving the country, and the initial asset position is $(1 - \pi^c) \bar{a}'_t$, where \bar{a}'_t are savings during the leaving period.

Problem of Asylum Seekers at Age $j \in \{j_s, \dots, j_r - 2\}$. The problem of an asylum seeker in the working period is

$$V_t(j, s, as, e = em, a) = \max_{c, a'} \left\{ U\left(\frac{c}{1 + \zeta n + \xi}, 1 - \mathbb{1}_a \underline{l}^a - (1 - \mathbb{1}_a) \bar{l}^a\right) \beta \psi_t(j, i, g) \left(\pi^l V_{t+1}^l(j+1, s, as, a') + (1 - \pi^l) \left(\pi^{ar} V_{t+1}(j+1, s, i = rw, e' = em, a') + (1 - \pi^{ar}) V_{t+1}(j+1, s, i = as, e' = em, a') \right) \right) \right\}$$

subject to

$$\begin{aligned} a' &= (a + tr_t)(1 + r_t(1 - \tau_t^k)) + y_t + b_t^a(n) - (c(1 + \tau_t^c) + T_t(y_t)) \geq 0 \\ y_t &= w_t(s, j, as, g) \epsilon(s, j, as, g) (\mathbb{1}_a \underline{l}^a + (1 - \mathbb{1}_a) \bar{l}^a). \end{aligned}$$

Immigrants from Other Population Groups. Unlike asylum seekers groups rw, ho have full access to the labor market and to the German social insurance system in the first period after arrival. Immigrants from group rw face in each period the probability π^{rh} to assimilate to group ho , which they take into account in their continuation values.

C Calibration Appendix

Table 5 summarizes the exogenously calibrated and Table 6 the endogenously calibrated parameters of the model.

Table 5: Exogenous Calibration Parameters

Parameter	Interpretation	Value
<i>Population</i>		
$N_t(j, i)$	Population Stock Data	Section 6.2
j_a	Age at labor market entrance	17
j_l	Age of education completion low-skilled	16
j_m	Age of education completion middle-skilled	20
j_h	Age of education completion high-skilled	24
j_f	Fertility Age	15
j_c	Age of completing fertility	50
j_r	Statutory Retirement Age	66
J	Max. Lifetime	100
$\{\psi_t(j, i, g)\}$	Survival rates	Section 6.2
ϕ	Fraction of baby girls	0.48
$\{\chi_t(j, i)\}$	Fertility rates	Section 6.2
π^l	Leaving probability	0.06
π^{ar}	Assimilation probability $as \Rightarrow rw$	0.008
π^{rh}	Assimilation probability $rw \Rightarrow hi$	0.060
$\phi(s, i)$	Fraction of skill s among population i	Table 3
<i>Endowments</i>		
$\epsilon(j, s, i)$	Age Profile	Figure 16
$\epsilon(g)$	Productivity Shifter by Gender	[0.8074, 1]
$\{l_1, \dots, l_n\}$	Discrete labor supply levels	{0.036, 0.18, 0.36}
$\{\underline{l}^a, \bar{l}^a\}$	Fraction of full-time work of group as	{0.109, 0.369}
η	Relative productivity of leavers	0.45
$\varrho(s)$	Productivity loss in education	0.5
π^c	Confiscation rate of assets for leaving asylum seekers	1
<i>Preferences</i>		
θ	Relative risk aversion parameter	1
σ_ϵ	Scale parameter of taste shocks	0.1

<i>Production</i>		
α	Capital share	0.33
δ	Depreciation rate	0.05
λ	Rate of technological progress	0.015
ϑ	Elast. of substitution b/w capital and labor	1
σ_{lmh}	Elast. of substitution b/w labor of different skill levels	3.05
σ_{nf}	Elast. of substitution b/w fo and na	13.22
σ_{hr}	Elast. of substitution b/w ho and rw	22.61
<i>Government</i>		
α^p	Sensitivity parameter in the pension formula	0 (0.25)
τ^p	Pension contribution rate	Figure 17
τ^h	Health system contribution rate	Figure 17
ι	Private contribution factor	0 (0.04)
b^a	Transfer payments to asylum seekers	Section 6.6.1
$\{b^h(j)\}$	Health insurance payments	Figure 18
τ^c	Consumption tax rate (in steady-state)	19%
τ^k	Capital income tax rate	25%
$\{G/Y_t\}$	Government consumption to GDP ratio	Section 6.6.2

Notes: Exogenous calibration parameters from various source described in Section 6.

Table 6: Endogenous Calibration Parameters

Parameter	Interpretation	Value	Target	Target Value
β	Discount Factor	0.9827	K/Y	3.66
$\phi(g)$	Util. Wgths.	[2.846,1.227]	Uncond. Hours/Full-Time Hours	[0.475,0.815]
ω_0	Level Param. Tax Func.	0.8416	G/Y	19.12%

Notes: Parameters calibrated endogenously by matching 2010 auxiliary steady state moments.

C.1 Population Model

Population data is from the German Federal Statistical Office (Statistisches Bundesamt /Destatis; HMD) and from the Central Foreign Population Registry (Ausländerzentralstatistik, AZR). In this data foreigners include all persons who do not have German citizenship, and we have explicit information on the stock of first and second generation foreigners. A first generation foreigner is a person that was born outside Germany, whereas a second generation foreigner in the data is born in Germany but holds foreign nationality. By our economic perspective we consider second generation foreigners as natives irrespective of their legal citizenship, cf. equation (30), and accordingly assign them to population group with “nationality” na . With this assumption, we first construct the age-specific population stock $N_t(j, i, g)$ for groups $i \in \{na, ho, rw, as\}$ for the years 2008 – 2019.

Next, we impute from this data the implied net addition to the population stock $M_t(j, i, g)$ from the law of motion of the population in equation (29), taking into account the adjustments of the dynamics that are implied by the assimilation probability π^{ar} and the leave probability π^l . We refer to the net change of the stock also as the migration flow.³⁸

To compute this net flow from (29) we also need data on age, group, and time specific mortality rates. We take those from the Human Mortality Database (HMD) for years 1960-2017 and, since we lack data on group specific mortality rates, we assume that all immigrants immediately after entry have the same mortality process as the average German population and thus set $\psi_t(j, i, g) = \psi_t(j, g) \forall i$.

For the predictions of the population beyond 2019 we make the following assumptions:

1. For all groups $\{na, rw, ho, as\}$ we compute the average age distribution of constructed net migration numbers $\bar{M}(j, i, g)$ over the years 2007-2018. We assume that aggregate migration in each group reverts to a long-run average until 2022. This reversal takes place according to the timing assumptions for each scenario described in Section 6.1. To compute long-run average migration in each group we assume—consistent with conventional assumptions by the German Federal Statistical Office (Statistisches Bundesamt)—that total migration over all groups is 200,000 annually and then distribute this total migration to the three groups ho, rw, as according to the relative shares during the years 2008-2012.

³⁸The advantage of constructing the flow data from the information on the population stock is that we can meaningfully measure the net addition to the stock caused by migration. Also, direct information on flows features statistical inaccuracies because of double counting of multiple within year migration. The disadvantage is that we do have to make assumptions on mortality and survival rates for all population groups. However, mortality is relevant only at higher ages at which migration numbers are close to zero.

2. Age and group specific fertility distributions are constant at their respective age specific averages taken over the years 2007-2018 until year 2100. Thereafter, fertility rates adjust such that the number of newborns is constant in each period. With this assumption (and with the assumption of constant survival rates and constant migration numbers) the population will reach a stationary distribution with constant population growth by about year 2200.
3. Survival rates increase according to the predictions from the Lee-Carter model until year 2100 and are constant thereafter.

During the phase-in period from 1960 to 2012 we have the exact data on the population stocks only from 2008 onwards. Leading towards 2008 we forward shoot on the population dynamics using data on the annual flow of migration and distribute those across the four groups such that we minimize the distance between the model implied population stocks in the four groups in 2008 and the respective actual population stock.

The next subsections contain a more detailed description of the construction of fertility rates, mortality rates and migration numbers.

C.1.1 Fertility Rates

In the data the number of newborns is

$$N_{t+1}(0, i) = \sum_{j=j_f}^{j_c} f_t(j, i) N_t(j, i, fe)$$

where $f_t(j, i)$ is the group i age j and time t specific fertility rate. Since we lack information on $f_t(j, i)$ and on the number of newborns for all population groups, we construct fertility rate as follows. We take time and age specific fertility rates of the average German population from the Federal Statistical Office and on the number of birth from the Human Mortality Database, separately for East and West Germany. Based on the stock of the population in both regions, we next adjust the age- and time-specific fertility rates such that the fertility distribution is consistent with the number of newborns. We then take the population weighted average of the East and West German constructed data.

C.1.2 Mortality Rates

We take a time series of gender specific mortality rates for 1950 to 2017 from the Human Mortality Database, computed as the weighted average of East and West German mortality

rates, and decompose mortality rates as

$$\ln(1 - \psi_t(j, g)) = a(j, g) + b(j, g)d_t(g)$$

where $\psi_t(j, g)$ is the survival rate applying the Lee-Carter procedure (Lee and Carter 1992). Next, we assume that the estimated time specific factor $\hat{d}_t(g)$ obeys a unit root process

$$\hat{d}_{t+1}(g) = \alpha(g) + \hat{d}_t(g) + \epsilon_{t+1}(g).$$

Based on the estimates $\left\{ \hat{a}(j, g), \hat{b}(j, g) \right\}_{j=0}^J, \hat{d}_t(g), \hat{\alpha}(g)$ we predict (future) survival rates by setting to zero the innovation terms $\hat{\epsilon}_t(g)$ and initialize the process assuming that $\hat{d}_0(g) = \hat{d}_0(g)$.

After construction of the population numbers (and the migration flows, see next subsection) we take population weighted average survival rates and recompute the population dynamics.

C.1.3 Migration Numbers

We construct the net addition to the respective population stock in group i by backing out the net flow from equation (29):³⁹ Since we lack data on group specific mortality rates, we assume that all immigrants immediately after entry have the same mortality process as the average German population and thus set $\psi_t(j, i, g) = \psi_t(j, g) \forall i$.

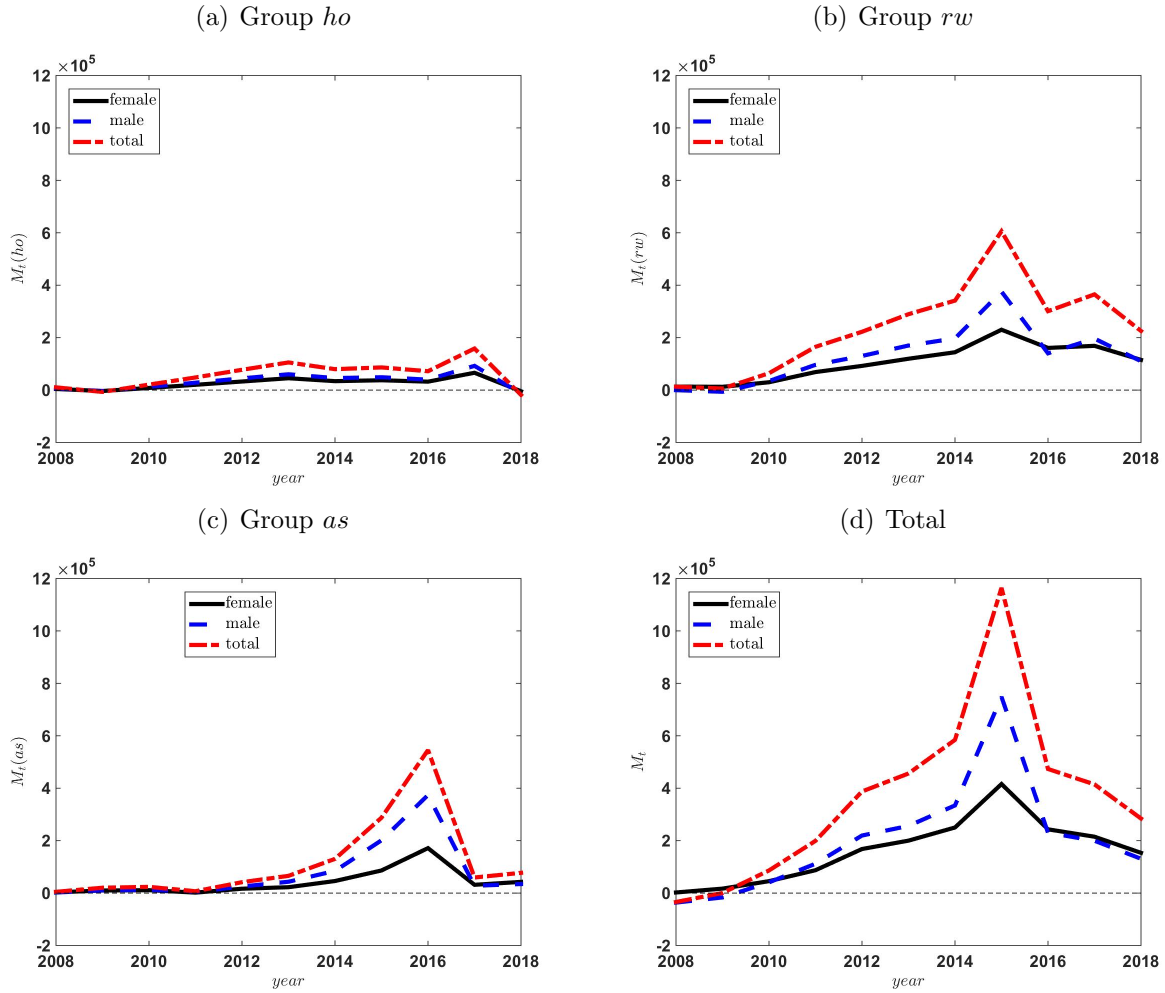
Figure 13 summarizes the constructed migration flows in the three groups of the foreign population $\{ho, rw, as\}$, Figure 14 contains the according age distribution of the migration flow, and Figure 15 shows the resulting age distribution of the population in groups $i \in \{na, ho, rw, as\}$.

C.2 Age Wage Profiles

Figure 16 displays the age wage profiles for natives and for foreigners from group rw . Age wage profiles for foreigners from groups ho are similar to those of natives.

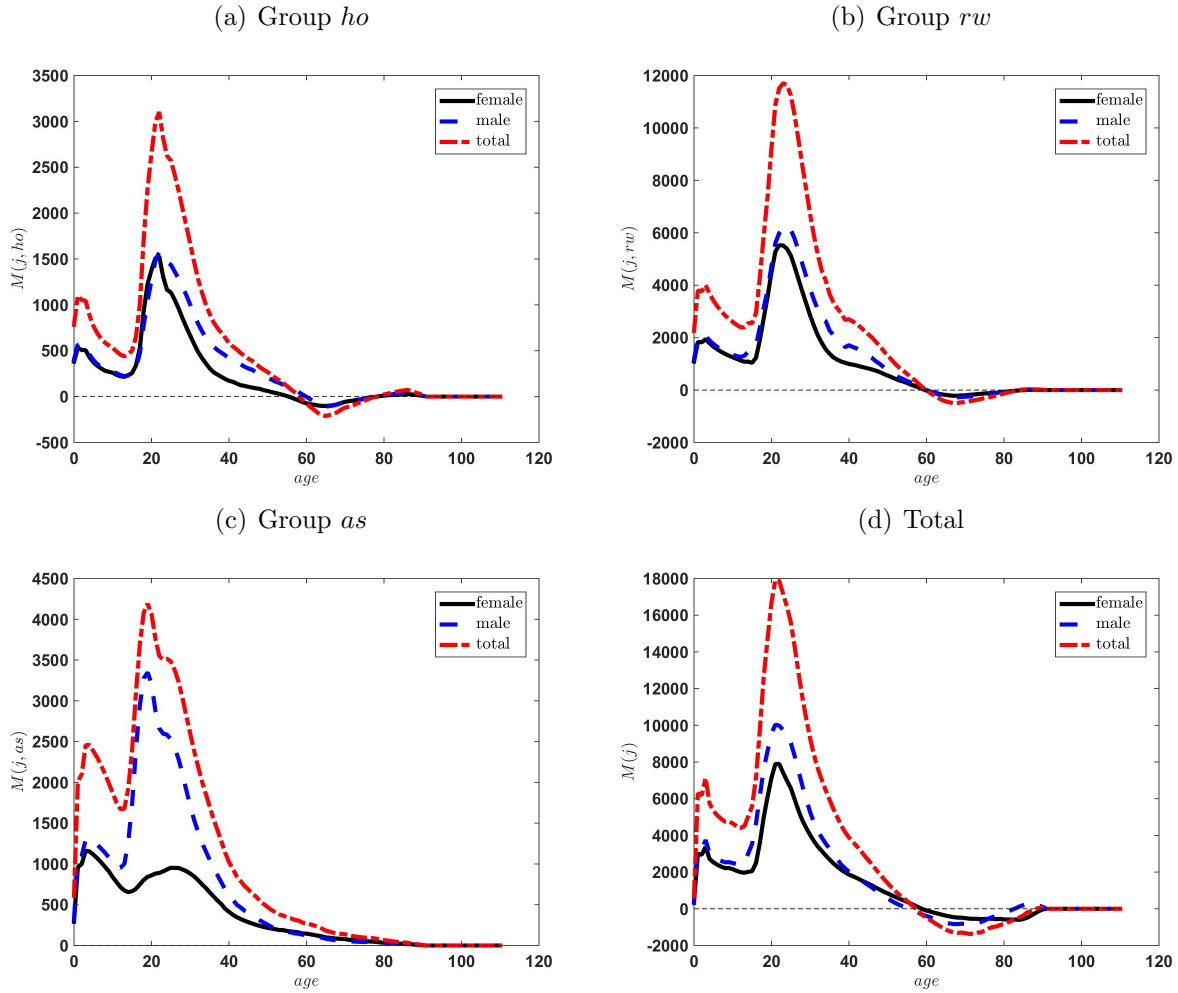
³⁹In the data, the population stock is reported at the end of a given calendar year which we accordingly interpret as the beginning of the next calendar year. Thus the population stock reported in the data at the end of calendar year 2007 is taken to be the population stock at the beginning of year 2008.

Figure 13: Net Migration Flows



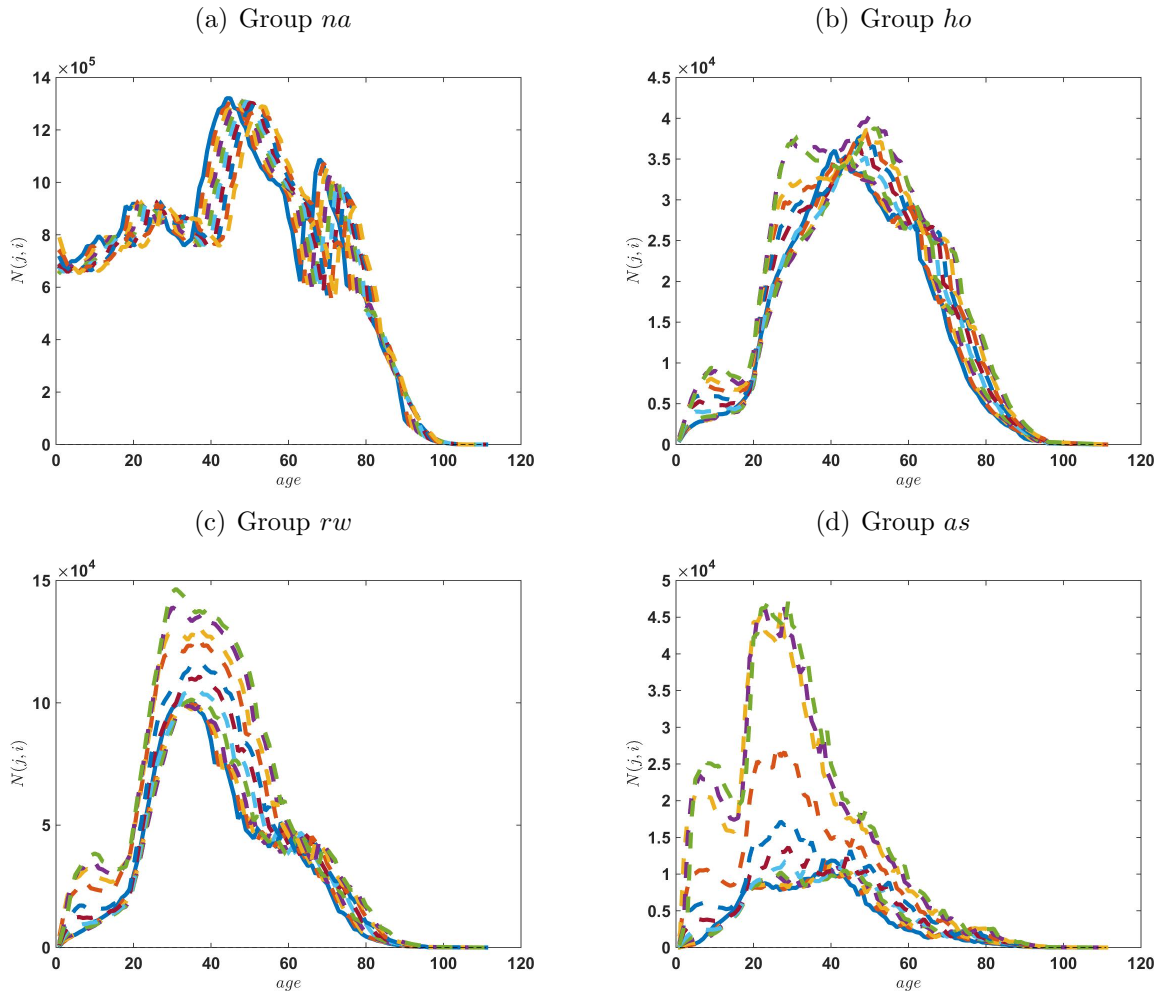
Notes: Annual aggregate net migration numbers from 2008 to 2018 by nationality group. Panel (a): HIOECD, Panel (b): RW, Panel (c): AS, Panel (d): total. Source: Own calculations based on Central Foreign Population Registry (Ausländerzentralstatistik, AZR).

Figure 14: Age Distribution of Net Migration



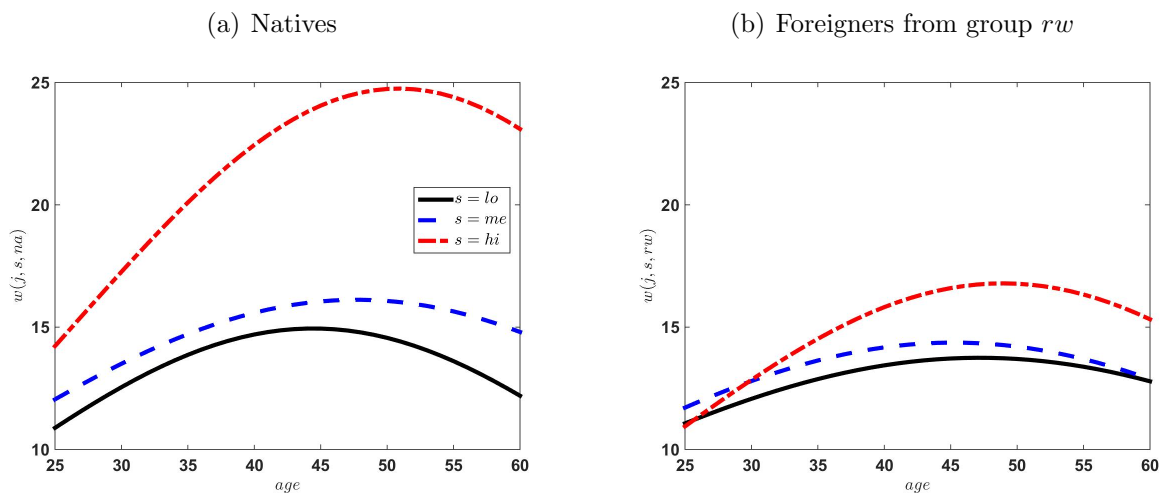
Notes: Age distribution of net migration, average of years 2008 to 2018. Panel (a): HIOECD, Panel (b): RW, Panel (c): AS, Panel (d): total. Source: Own calculations based on Central Foreign Population Registry (Ausländerzentralstatistik, AZR).

Figure 15: Age Distribution of Population Stock



Notes: Age distribution of population stock by years 2008 to 2018. Panel (a): Natives, Panel (b): HIOECD, Panel (c): RW, Panel (d): AS. Source: Own calculations based on Central Foreign Population Registry (Ausländerzentralstatistik, AZR) and German Federal Statistical Office (Statistisches Bundesamt).

Figure 16: Age Wage Profiles



Notes: Predicted age wage profiles for natives and foreigners from RW for the three skill categories $s \in \{lo, me, hi\}$. Source: Own calculations based on SOEP.

C.3 Technology

For the estimation of the substitution elasticities in production, we exploit the homogeneity of the production function in each nest and add productivity parameters $\tilde{\epsilon}(\cdot)$ at each nest, which are normalized to one. Thus, at the estimation, we write labor at each nest as

$$\begin{aligned}
L_t &= \left(\sum_{s \in \{lo, me, hi\}} \tilde{\epsilon}(s) \tilde{L}_t(s)^{1 - \frac{1}{\sigma_{lmh}}} \right)^{\frac{1}{1 - \frac{1}{\sigma_{lmh}}}} \\
\tilde{L}_t(s) &= \sum_{\bar{j}=1}^{n_{\bar{j}}} \tilde{\epsilon}(\bar{j}, s) \tilde{L}_t(\bar{j}, s) \\
\tilde{L}_t(\bar{j}, s) &= \left(\tilde{\epsilon}(\bar{j}, s, na) \tilde{L}_t(\bar{j}, s, na)^{1 - \frac{1}{\sigma_{nf}}} + \tilde{\epsilon}(\bar{j}, s, fo) \tilde{L}_t(\bar{j}, s, fo)^{1 - \frac{1}{\sigma_{nf}}} \right)^{\frac{1}{1 - \frac{1}{\sigma_{nf}}}} \\
\tilde{\tilde{L}}_t(\bar{j}, s, fo) &= \left(\tilde{\epsilon}(\bar{j}, s, ho) \tilde{L}_t(\bar{j}, s, ho)^{1 - \frac{1}{\sigma_{hr}}} + \left(\sum_{i \in \{rw, as\}} \tilde{\epsilon}(\bar{j}, s, i) \tilde{L}_t(\bar{j}, s, i) \right)^{1 - \frac{1}{\sigma_{hr}}} \right)^{\frac{1}{1 - \frac{1}{\sigma_{hr}}}} \\
\tilde{L}_t(\bar{j}, s, i) &= \sum_{g \in \{fe, ma\}} \tilde{\epsilon}(\bar{j}, s, i, g) L_t(\bar{j}, s, i, g),
\end{aligned}$$

and we assume that $\tilde{\epsilon}(\bar{j}, s, i, g) = \tilde{\epsilon}(\bar{j}, s, i) \tilde{\epsilon}(g)$. Given the homogeneity of the production function, it is straightforward to show that the productivity scaling parameters $\tilde{\epsilon}(\cdot)$ can be mapped into labor productivity $\epsilon(\bar{j}, s, i) \epsilon(g)$, where $\epsilon(g) = \tilde{\epsilon}(g)$, and $\epsilon(\bar{j}, s, i)$ is an average over $j \in [j_l(\bar{j}, s), \dots, j_h(\bar{j}, s)]$ of the (s, i) -specific productivity profile $\epsilon(j, s, i)$.

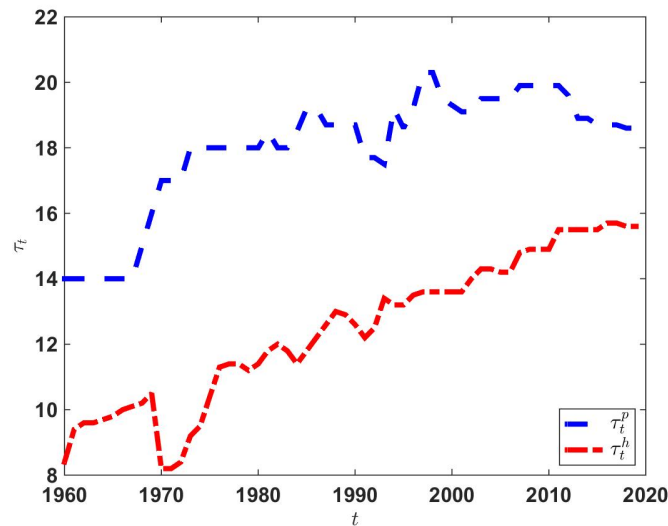
We estimate the elasticities of substitution at the different nests jointly with the productivity parameters following the standard approach in the literature (cf., e.g., Borjas 2003). For example, at the level of immigrant groups, we translate the first order conditions into estimation equations and identify the relative productivity parameters and the elasticity of substitution across immigrant groups using variation over time in the relative labor supply (hours worked) and the relative wages. We then use these estimates to obtain the CES aggregator of labor supply of immigrants, and the implied wage aggregate. We use this together with labor supply and wages of natives to estimate the next layer of the CES, and then move up nest by nest in the same fashion. At each nest, we use the population size of a given group as an instrument for the labor supply in order to address a potential endogeneity problem of the estimation equations. At the highest nest, we allow the education group specific productivity components to follow quadratic time trends in order to accommodate the possibility of skill-biased technological change in the estimation. We implement the estimator using SOEP data from 1984-2015. For each year, group specific hours worked are

aggregate hours worked by individuals up to age 60 of a given group, and the group specific wage rate is estimated using workers up to age 60 who work at least 520 hours.

C.4 Social Insurance

Figure 17 shows the contribution rates to the German PAYG pension system and to the public health insurance system (including long-term care insurance).

Figure 17: Contribution Rates to Social Security & Health Insurance



Notes: Data on contribution rates to social security and health insurance. Source: <http://www.sozialpolitik-aktuell.de>.

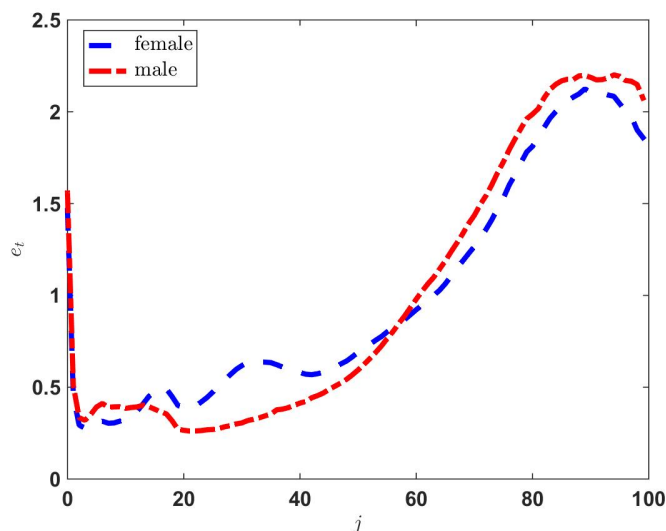
Our data on health expenditures cover ages 0-99 for years 2010-2017. We normalize these expenditures by nominal GDP data (which leads to almost identical profiles for all years pointing to strong time effects) and take the average across these years. Figure 18 shows the age profile for females and males.

D Appendix: Further Results

D.1 Population Shares by Groups

Figures 19 and 20 shows the population shares by nationality and their changes relative to the baseline scenario.

Figure 18: Health Expenditures over the Life-Cycle [Index, centralized data]



Notes: Data on age-specific health expenditures. *Source:* Federal Insurance Office (Bundesversicherungsamt).

D.2 The Fiscal Side

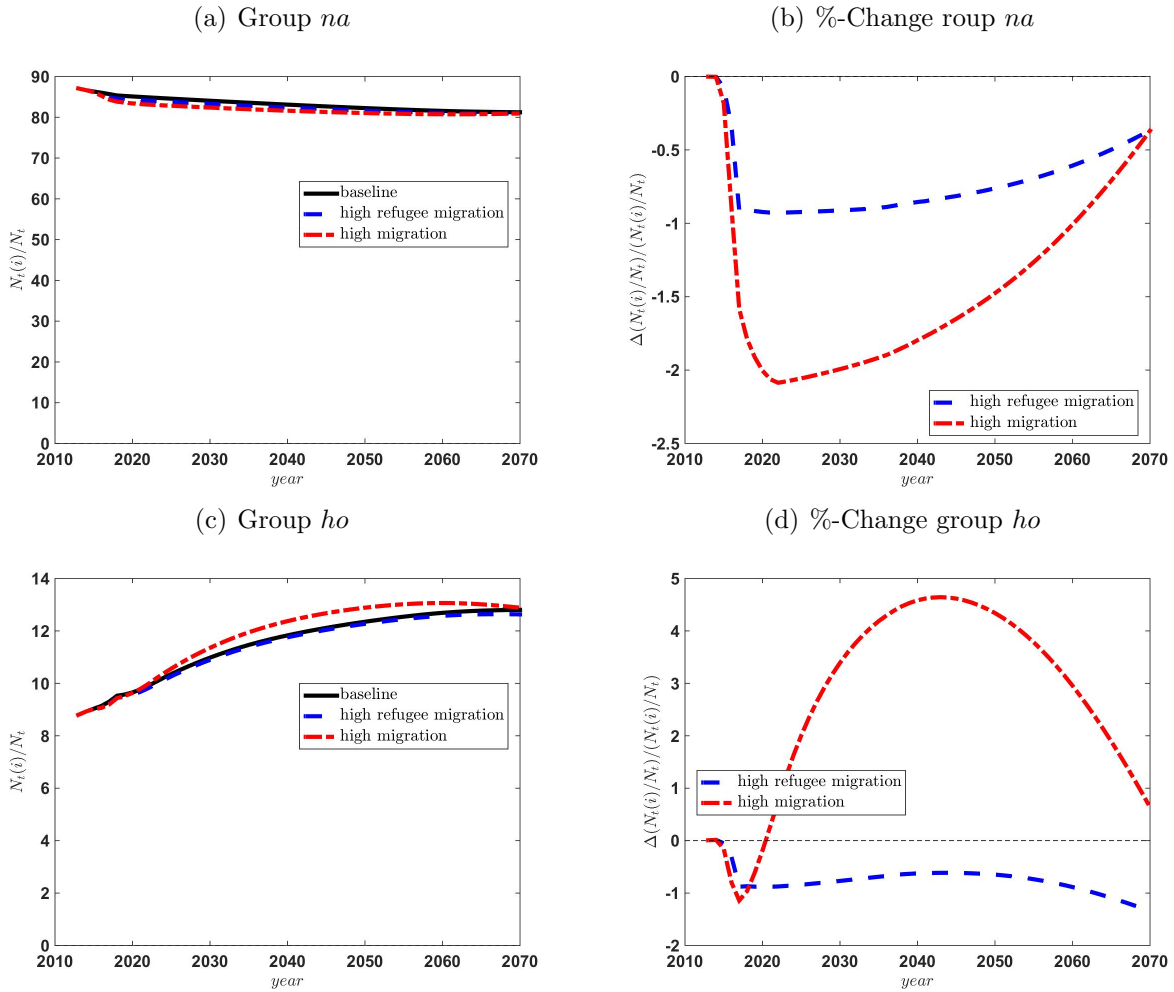
Figure 21 shows the health contribution rate (and its change). In the refugee migration scenario the health contribution rate increases slightly; refugee immigrants receive the same age contingent lump-sum payments but contribute little to the system. In the high migration scenario the contribution rate initially increases when the effect of young in-migration dominates.

Figure 22 shows total government expenditures as the sum of government consumption G_t and all outlays to finance incoming and leaving refugees E_t . In the migration scenarios we observe the initial blip due to the incoming wave of migrants, but overtime overall expenditures decrease slightly relative to GDP because GDP increases.

D.3 Per Capita GDP and Consumption

Figure 23 de-trended per capita GDP and consumption, where de-trending is by the technology level A_t .

Figure 19: Population Shares by Region of Origin I



Notes: Fractions as a share of total population and respective percentage changes. Panels (a)-(b): Natives, Panels (c)-(d): HIOECD.

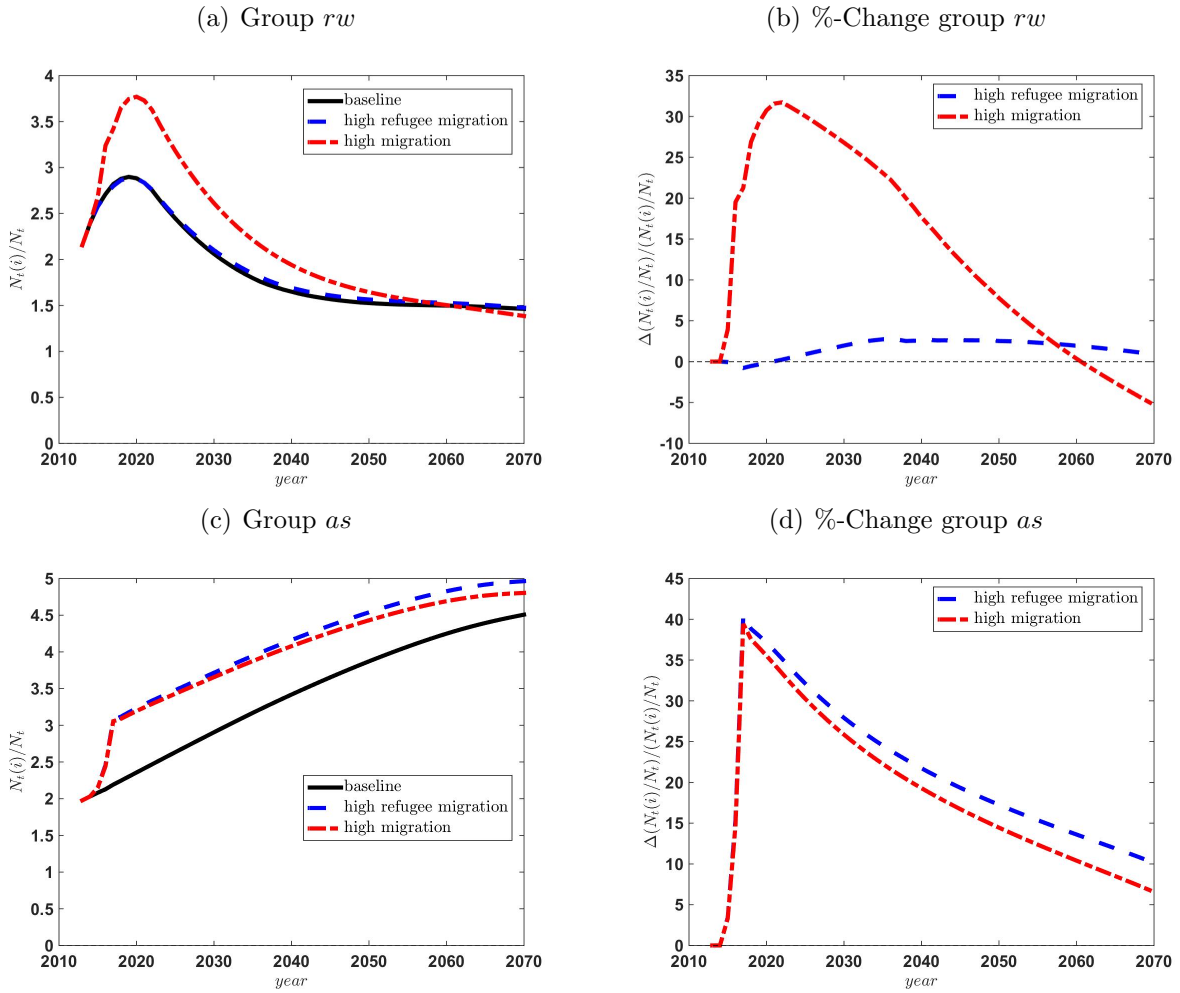
D.4 Rate of Return & Wages

Figure 24 shows the rate of return to capital and its change to the baseline demographic model. Figure 25 shows gross and net wages of low skilled natives as weighted averages of the three age groups.

D.5 Wage Changes

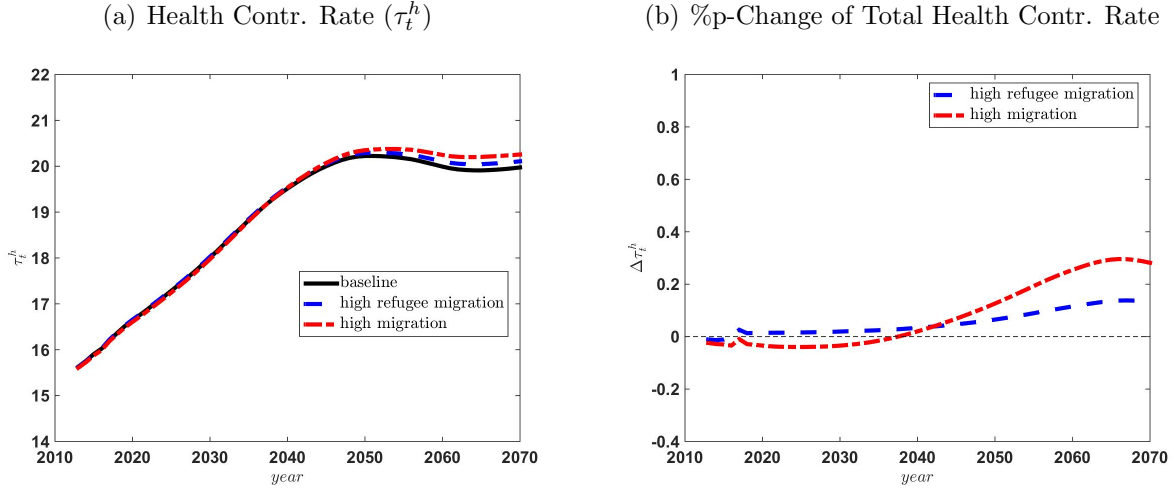
Figure 26 displays the change of the skill ratios for the low-skilled natives.

Figure 20: Population Shares by Region of Origin II



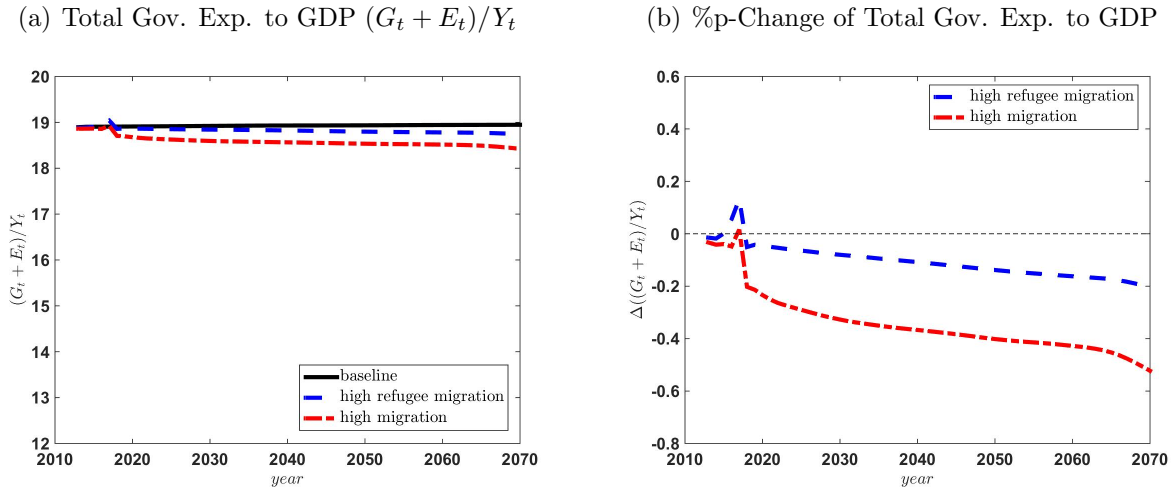
Notes: Fractions as a share of total population and respective percentage changes. Panels (a)-(b): RW, Panels (c)-(d): AS.

Figure 21: Health Contribution Rate



Notes: Panel (a): contribution rate to health insurance system, Panel (b): percentage point change of contribution rate to health insurance system.

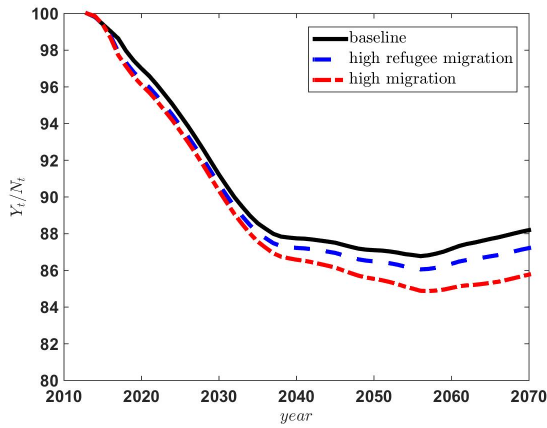
Figure 22: Total Government Expenditures



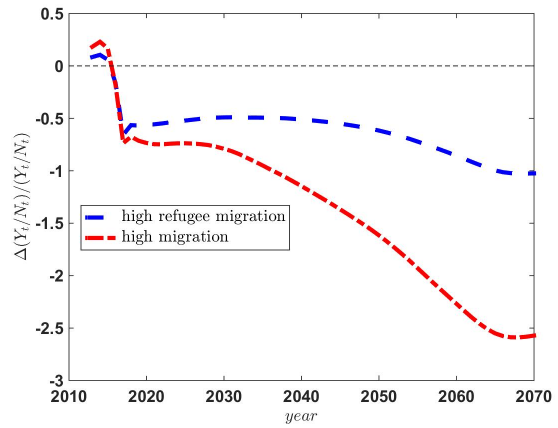
Notes: Panel (a): ratio of total government expenditures to GDP, Panel (b): percentage point change of ratio of total government expenditures to GDP.

Figure 23: Detrended Per Capita GDP & Consumption [Index]

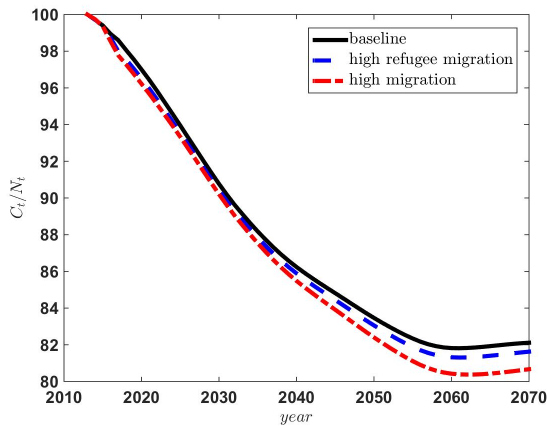
(a) Per Capita GDP



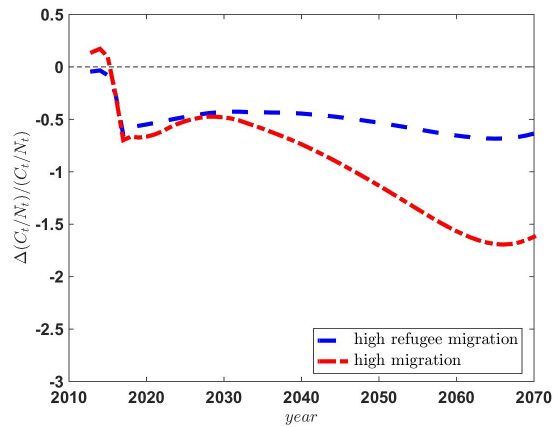
(b) %-Change of P.C.GDP



(c) Per Capita Consumption

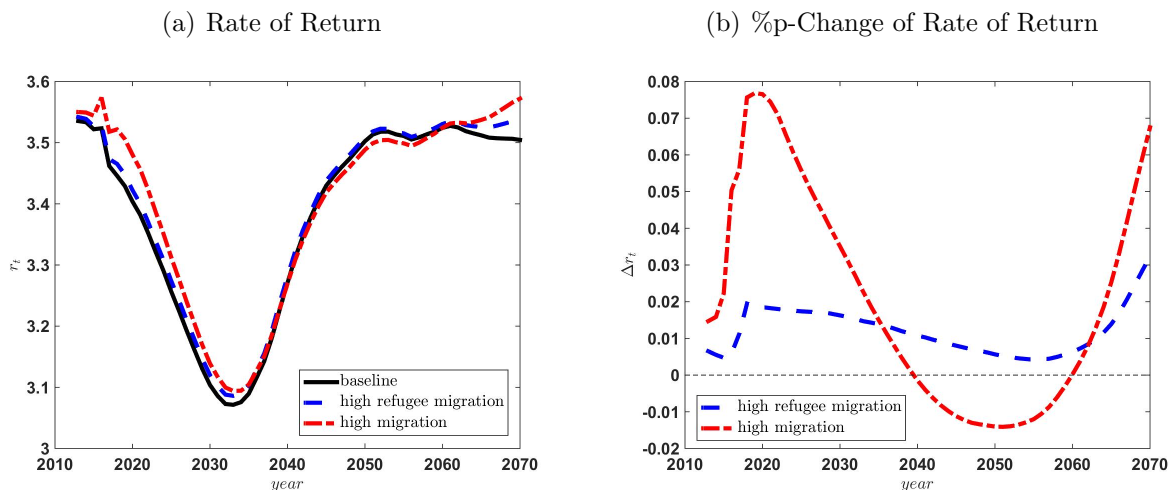


(d) %-Change of P.C.Consumption



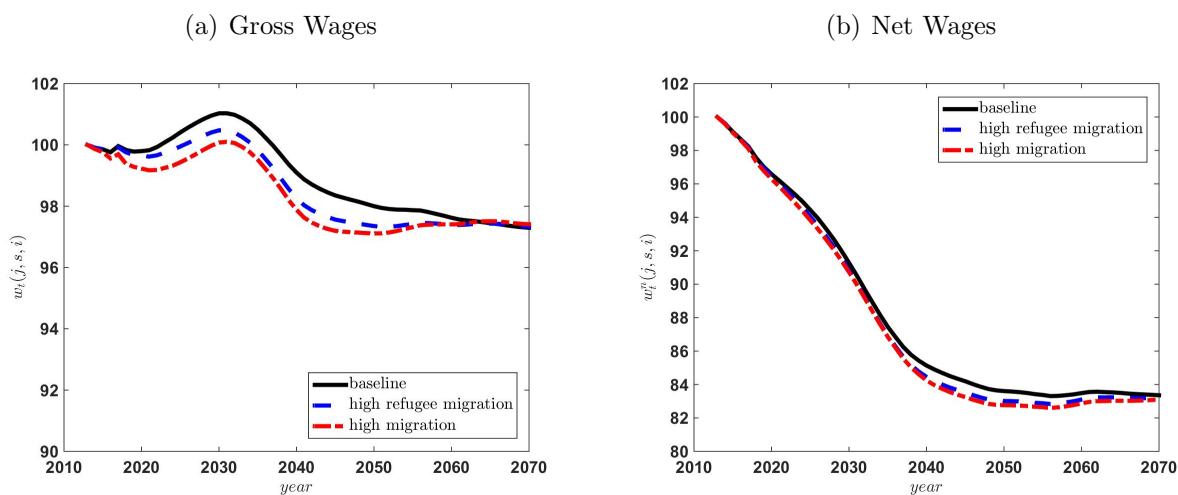
Notes: Panel (a): de-trended per capita GDP (Index, 2013=100), Panel (b): percent change of per capita GDP; Panel (c): de-trended per capita consumption (Index, 2013=100), Panel (d): percent change of per capita consumption.

Figure 24: Rate of Return



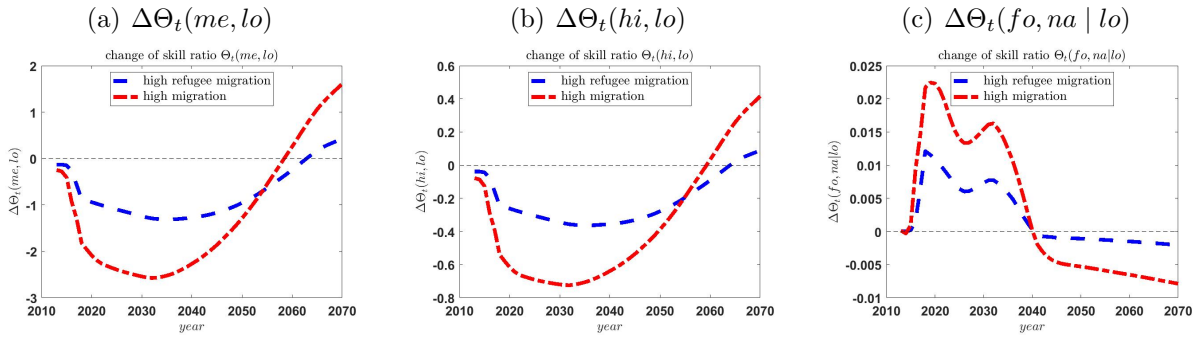
Notes: Panel (a): rate of return, Panel (b): percentage point change of rate of return.

Figure 25: Gross & Net Wages, Low-Skilled Natives, Age-Group $\bar{j} = 1$



Notes: Panel (a): gross wages, Panel (b): net wages of low skilled natives in age group $\bar{j} = 1$.

Figure 26: Change of Skill Ratios, Low-Skilled Natives, Age Group $\bar{j} = 1$



Notes: Change of skill ratios as defined in equation (46) for low skilled natives in age group $\bar{j} = 1$.