The Economic Geography of Global Warming*

José-Luis Cruz
Princeton University

Esteban Rossi-Hansberg
Princeton University

February 9, 2021

Online Appendix

A Proofs

In this section, we discuss the solution of the model, as well as the existence and uniqueness of a balanced growth path. Additionally, we outline the solution of the model backwards in time and the system of equations that solve for fundamental amenities, productivities and migration costs.

A.1 Forward Solution

In this subsection, we compute the forward solution of the model under the presence of proportional carbon taxes, $\tau_t(r)$, and clean energy subsidies, $s_t(r)$. Those taxes are charged to the firm and uniformly rebated to the households residing in region r itself through a lump-sum transfer, $\Phi_t(r)$.

The firm's cost minimization problem, in terms of fossil fuels and clean energy, is given by:

$$\begin{aligned} w_t(r)\mathcal{Q}_t(r)e_t^{\omega}(r) = \min_{e^f, e^c} \quad & w_t(r)(1+\tau_t(r))\mathcal{Q}_t^f(r)e_t^{f,\omega}(r) + w_t(r)(1-s_t(r))\mathcal{Q}_t^c(r)e_t^{c,\omega}(r) \\ \text{st} \quad & \left(\kappa e_t^{f,\omega}(r)^{\frac{\epsilon-1}{\epsilon}} + (1-\kappa)e_t^{c,\omega}(r)^{\frac{\epsilon-1}{\epsilon}}\right)^{\frac{\epsilon}{\epsilon-1}} = e_t^{\omega}(r) \end{aligned}$$

The First Order Conditions with respect to $e_t^{f,\omega}(r)$ and $e_t^{c,\omega}(r)$ imply the following relations, where $\mathcal{Q}_t(r)$ is the ideal energy price index.

$$\frac{e_t^{c,\omega}(r)}{e_t^{f,\omega}(r)} = \left(\frac{1-\kappa}{\kappa}\frac{1+\tau_t(r)}{1-s_t(r)}\frac{\mathcal{Q}_t^f(r)}{\mathcal{Q}_t^c(r)}\right)^{\epsilon} \tag{32}$$

$$\mathcal{Q}_t(r) = \left(\kappa^{\epsilon}(1+\tau_t(r))^{1-\epsilon}\mathcal{Q}_t^f(r)^{1-\epsilon} + (1-\kappa)^{\epsilon}(1-s_t(r))^{1-\epsilon}\mathcal{Q}_t^c(r)^{1-\epsilon}\right)^{\frac{1}{1-\epsilon}} \tag{33}$$

We define the lump-sum transfer, per unit of land, as shown in equation (34).

$$\begin{aligned} \Phi_t(r) &= \int_0^1 \Phi_t^{\omega}(r)d\omega \\ \Phi_t^{\omega}(r) &= \left(\tau_t(r)\mathcal{Q}_t^f(r)e_t^{f,\omega}(r) - s_t(r)\mathcal{Q}_t^c(r)e_t^{c,\omega}(r)\right) \\ &= \mathcal{Q}_t(r)e_t^{\omega}(r) - \left(\mathcal{Q}_t^f(r)e_t^{f,\omega}(r) + \mathcal{Q}_t^c(r)e_t^{f,\omega}(r)\right) \\ &= \mathcal{Q}_t(r)e_t^{\omega}(r) - \tilde{\mathcal{Q}}_t(r)^{1-\epsilon}\mathcal{Q}_t(r)^{\epsilon}e_t^{\omega}(r) \end{aligned} \tag{34}$$

$$\tilde{\mathcal{Q}}_t(r) = \left(\kappa^{\epsilon}(1+\tau_t(r))^{-\epsilon}\mathcal{Q}_t^f(r)^{1-\epsilon} + (1-\kappa)^{\epsilon}(1-s_t(r))^{-\epsilon}\mathcal{Q}_t^c(r)^{1-\epsilon}\right)^{\frac{1}{1-\epsilon}} \tag{35}$$

Equation (33) reduces the firm's problem to:

$$\begin{aligned} \max_{q,L,\phi,e} \quad & p_t^{\omega}(r,r)\phi_t^{\omega}(r)^{\gamma_1}z_t^{\omega}(r)\left(L_t^{\omega}(r)^{\chi}e_t^{\omega}(r)^{1-\chi}\right)^{\mu} - w_t(r)L_t^{\omega}(r) \\ & - w_t(r)\nu\phi_t^{\omega}(r)^{\xi} - w_t(r)\mathcal{Q}_t(r)e_t^{\omega}(r) - R_t(r) \end{aligned}$$

The First Order Conditions with respect to $e_t^\omega(r)$ and $L_t^\omega(r)$ imply:

$$\mathcal{Q}_t(r)e_t^\omega(r) = \left(\frac{1-\chi}{\chi}\right) L_t^\omega(r) \tag{36}$$

Equation (36) collapses the firm's problem to:

$$\max_{L,\phi} \quad p_t^\omega(r,r)\left(\frac{1-\chi}{\chi}\frac{1}{\mathcal{Q}_t(r)}\right)^{(1-\chi)\mu} \phi_t^\omega(r)^{\gamma_1} z_t^\omega(r) L_t^\omega(r)^\mu - \frac{w_t(r)L_t^\omega(r)}{\chi} - w_t(r)\nu\phi_t^\omega(r)^\xi - R_t(r)$$

The First Order Conditions with respect to $\phi_t^\omega(r)$ and $L_t^\omega(r)$ imply:

$$w_t(r)L_t^\omega(r) = \mu\chi\left(p_t^\omega(r,r)q_t^\omega(r)\right) \tag{37}$$

$$\chi\mu\nu\phi_t^\omega(r)^\xi = (\gamma_1/\xi)L_t^\omega(r) \tag{38}$$

We define total labor demand, $\bar{L}_t^\omega(r)$, as the demand from production, $L_t^\omega(r)$; innovation, $L_t^{\phi,\omega}(r)$; fossil fuel energy, $L_t^{f,\omega}(r)$; and clean energy, $L_t^{c,\omega}(r)$. Then, we insert equations (35), (36) and (38).

$$\begin{aligned}
\bar{L}_t^\omega(r) &= L_t^\omega(r) + L_t^{\phi,\omega}(r) + L_t^{f,\omega}(r) + L_t^{c,\omega}(r) \\
&= L_t^\omega(r) + \nu\phi_t^\omega(r)^\xi + \mathcal{Q}_t^f(r)e_t^{f,\omega}(r) + \mathcal{Q}_t^c(r)e_t^{c,\omega}(r) \\
&= L_t^\omega(r) + \frac{\gamma_1/\xi}{\mu\chi}L_t^\omega(r) + \left(\frac{\tilde{\mathcal{Q}}_t(r)}{\mathcal{Q}_t(r)}\right)^{1-\epsilon}\mathcal{Q}_t(r)e_t(r) \\
&= L_t^\omega(r) + \frac{\gamma_1/\xi}{\mu\chi}L_t^\omega(r) + \left(\frac{\tilde{\mathcal{Q}}_t(r)}{\mathcal{Q}_t(r)}\right)^{1-\epsilon}\left(\frac{1-\chi}{\chi}\right)L_t^\omega(r) \\
&= \frac{\mu\chi + \gamma_1/\xi + \mu(1-\chi)(\tilde{\mathcal{Q}}_t(r)/\mathcal{Q}_t(r))^{1-\epsilon}}{\mu\chi}L_t^\omega(r)
\end{aligned}$$

To ease notation, define the term $\varphi_t(r)$ as shown in (39). Observe that when taxes and subsidies are zero, $\varphi_t(r) = 1$.

$$\varphi_t(r) = \frac{\mu\chi + \gamma_1/\xi + \mu(1-\chi)(\tilde{\mathcal{Q}}_t(r)/\mathcal{Q}_t(r))^{1-\epsilon}}{\mu + \gamma_1/\xi} \tag{39}$$

$$\bar{L}_t^\omega(r) = \left(\frac{\mu + \gamma_1/\xi}{\mu\chi}\right)\varphi_t(r)L_t^\omega(r) \tag{40}$$

Due to the Cobb Douglas formulation of the production function, rents $R_t(r)$ can be expressed as a constant fraction of revenue, by inserting equations (37) and (38).

$$\begin{aligned}
R_t(r) &= p_t^\omega(r,r)q_t^\omega(r) - \frac{w_t(r)L_t^\omega(r)}{\chi} - w_t(r)\nu\phi_t^\omega(r)^\xi \\
&= (1 - \mu - \gamma_1/\xi)p_t^\omega(r,r)q_t^\omega(r)
\end{aligned} \tag{41}$$

Since in equilibrium $R_t(r)$ is taken as given by firms producing at r, the decisions of how much to innovate, $\phi_t^{\omega}(r)$, how many workers to hire per unit of land, $L_t^{\omega}(r)$, and how much energy to use from fossil fuels, $e_t^{f,\omega}(r)$, and clean energy, $e_t^{c,\omega}(r)$, are independent of the local idiosyncratic productivity shocks, $z_t^{\omega}(r)$, and so are identical across varieties ω. Hereinafter, we drop the superscript ω, except from production labor, $L_t^{\omega}(r)$, to differentiate it from total labor demand, $L_t(r)$.

We arrange firm sales by inserting equations (37) and (38).

$$\begin{aligned}
p_t(r,r)q_t(r) &= p_t(r,r)\left(\frac{1-\chi}{\chi}\frac{1}{\mathcal{Q}_t(r)}\right)^{(1-\chi)\mu}\phi_t(r)^{\gamma_1}z_t(r)L_t^{\omega}(r)^{\mu} \\
&= p_t(r,r)z_t(r)\left(\frac{1-\chi}{\chi}\frac{1}{\mathcal{Q}_t(r)}\right)^{(1-\chi)\mu}\left(\frac{\gamma_1\chi}{\xi\nu}\frac{p_t(r,r)q_t(r)}{w_t(r)}\right)^{\frac{\gamma_1}{\xi}}\left(\mu\chi\frac{p_t(r,r)q_t(r)}{w_t(r)}\right)^{\mu} \\
(p_t(r,r)q_t(r))^{1-\mu-\gamma_1/\xi)} &= p_t(r,r)z_t(r)\left(\frac{1-\chi}{\chi}\frac{1}{\mathcal{Q}_t(r)}\right)^{(1-\chi)\mu}\left(\frac{\gamma_1\chi}{\xi\nu}\right)^{\frac{\gamma_1}{\xi}}(\mu\chi)^{\mu}\,w_t(r)^{-\mu-\gamma_1/\xi} \qquad (42)
\end{aligned}$$

We define marginal cost, $mc_t(r)$, as shown below and insert equations (37) and (40) in equation (42).

$$\begin{aligned}
mc_t(r) &= p_t(r,r)z_t(r) \\
&= \left(\frac{1-\chi}{\chi}\frac{1}{\mathcal{Q}_t(r)}\right)^{-(1-\chi)\mu}\left(\frac{\gamma_1\chi}{\xi\nu}\right)^{-\frac{\gamma_1}{\xi}}(\mu\chi)^{-\mu}\,w_t(r)^{\mu+\gamma_1/\xi}\,(p_t(r,r)q_t(r))^{1-\mu-\gamma_1/\xi} \\
&= \left(\frac{1-\chi}{\chi}\frac{1}{\mathcal{Q}_t(r)}\right)^{-(1-\chi)\mu}\left(\frac{\gamma_1\chi}{\xi\nu}\right)^{-\frac{\gamma_1}{\xi}}(\mu\chi)^{\gamma_1/\xi-1}\,w_t(r)L_t^{\omega}(r)^{1-\mu-\gamma_1/\xi} \\
&= \left(\frac{1-\chi}{\chi}\right)^{-(1-\chi)\mu}\left(\frac{\gamma_1\chi}{\xi\nu}\right)^{-\frac{\gamma_1}{\xi}}\mu^{-\mu}\chi^{-(\mu+\gamma_1/\xi)} \\
&\times \varphi_t(r)^{-(1-\mu-\gamma_1/\xi)}\mathcal{Q}_t(r)^{(1-\chi)\mu}w_t(r)L_t(r)^{1-\mu-\gamma_1/\xi} \qquad (43)
\end{aligned}$$

We define total revenue per unit of land, $\mathcal{R}_t(r)$, as the sum of labor revenue, rents and lump sum transfer.

$$\begin{aligned}
\mathcal{R}_t(r) &= w_t(r)L_t(r) + R_t(r) + w_t(r)\Phi_t(r) \\
&= w_t(r)L_t(r) + (1-\mu-\gamma_1/\xi)y_t(r) + w_t(r)\left[1-\left(\frac{\tilde{\mathcal{Q}}_t(r)}{\mathcal{Q}_t(r)}\right)^{1-\epsilon}\right]\mathcal{Q}_t(r)e_t(r) \\
&= w_t(r)L_t(r) + \left(\frac{1-\mu-\gamma_1/\xi}{\mu\chi}\right)w_t(r)L_t(r) + w_t(r)\left[1-\left(\frac{\tilde{\mathcal{Q}}_t(r)}{\mathcal{Q}_t(r)}\right)^{1-\epsilon}\right]\left(\frac{1-\chi}{\chi}\right)L_t^{\omega}(r) \\
&= \left(\frac{1-\mu-\gamma_1/\xi+\mu\chi}{\mu\chi}\right)w_t(r)L_t(r) + w_t(r)\left[1-\left(\frac{\tilde{\mathcal{Q}}_t(r)}{\mathcal{Q}_t(r)}\right)^{1-\epsilon}\right]\left(\frac{\mu(1-\chi)}{\mu+\gamma_1/\xi}\right)\frac{1}{\varphi_t(r)}L_t(r) \\
&= \left(\frac{1}{\mu+\gamma_1/\xi}\right)\left(\frac{w_t(r)L_t(r)}{\varphi_t(r)}\right) \qquad (44)
\end{aligned}$$

Now, we combine equations (4) and (44), where $y_t(r) = \mathcal{R}_t(r)/L_t(r)$.

$$\begin{aligned} u_t(r) &= \bar{b}_t(r) L_t(r)^{-\lambda} \frac{\mathcal{R}_t(r)}{L_t(r) P_t(r)} \\ &= \bar{b}_t(r) L_t(r)^{-\lambda} \left(\frac{1}{\mu + \gamma_1/\xi}\right) \left(\frac{w_t(r)}{\varphi_t(r)}\right) \left(\frac{1}{P_t(r)}\right) \end{aligned} \tag{45}$$

We solve for $P_t(r)$ and employ equations (16) and (43).

$$P_t(r) = \frac{\bar{b}_t(r)}{u_t(r)} L_t(r)^{-\lambda} \left(\frac{1}{\mu + \gamma_1/\xi}\right) \left(\frac{w_t(r)}{\varphi_t(r)}\right) \tag{46}$$

$$\begin{aligned} &= \Gamma \left(\frac{-\rho}{(1-\rho)\theta + 1}\right)^{-\frac{1-\rho}{\rho}} \left(\int_S a_t(v) [mc_t(v) \varsigma(r,v)]^{-\theta} dv\right)^{-\frac{1}{\theta}} \\ &= (\bar{p}\bar{\kappa}) \left(\int_S \bar{a}_t(v) L_t(v)^{\alpha - (1-\mu-\gamma_1/\xi)\theta} \varphi_t(v)^{(1-\mu-\gamma_1/\xi)\theta} \mathcal{Q}_t(v)^{-(1-\chi)\mu\theta} w_t(v)^{-\theta} \varsigma(r,v)^{-\theta} dv\right)^{-\frac{1}{\theta}} \end{aligned} \tag{47}$$

And we define $\bar{p}$ and $\bar{\kappa}$ as shown below.

$$\begin{aligned} \bar{p} &= \Gamma \left(\frac{-\rho}{(1-\rho)\theta + 1}\right)^{-\frac{1-\rho}{\rho}} \\ \bar{\kappa} &= \left(\frac{1-\chi}{\chi}\right)^{-(1-\chi)\mu} \left(\frac{\gamma_1 \chi}{\xi \nu}\right)^{-\frac{\gamma_1}{\xi}} \mu^{-\mu} \chi^{-(\mu + \gamma_1/\xi)} \end{aligned}$$

We manipulate the equation (47) to obtain (48).

$$\begin{aligned} &\left(\frac{\bar{b}_t(r)}{u_t(r)}\right)^{-\theta} L_t(r)^{\lambda\theta} \left(\frac{w_t(r)}{\varphi_t(r)}\right)^{-\theta} = \kappa_1 \\ &\times \left(\int_S \bar{a}_t(v) L_t(v)^{\alpha - (1-\mu-\gamma_1/\xi)\theta} \varphi_t(v)^{(1-\mu-\gamma_1/\xi)\theta} \mathcal{Q}_t(v)^{-(1-\chi)\mu\theta} w_t(v)^{-\theta} \varsigma(r,v)^{-\theta} dv\right) \end{aligned} \tag{48}$$

Where $\kappa_1 = \left((\mu + \gamma_1/\xi)\, \bar{p}\bar{\kappa}\right)^{-\theta}$. Now, we use the condition that trade is balanced cell by cell, given by equation (17), considering the expression of revenue per unit of land found in equation (44), and insert the expression for trade shares, given by equation (15).

$$\begin{aligned} \frac{w_t(r) H(r) L_t(r)}{\varphi_t(r))} &= \int_S \frac{\pi_t(v,r) w_t(v) H(v) L_t(v)}{\varphi_t(v)} dv \\ &= \bar{p}^{-\theta} \int_S \frac{a_t(r) [mc_t(r) \varsigma(r,v)]^{-\theta} P_t(v)^{\theta} w_t(v) H(v) L_t(v)}{\varphi_t(v)} dv \end{aligned} \tag{49}$$

We insert equation (43) in the left-hand side of equation (49).

$$\begin{aligned} &\frac{w_t(r) H(r) L_t(r)}{\varphi_t(r) a_t(r)} mc_t(r)^{\theta} \\ &= \frac{w_t(r) H(r) L_t(r)^{1-\alpha}}{\varphi_t(r) \bar{a}_t(r)} \left(\bar{\kappa} \varphi_t(r)^{-(1-\mu-\gamma_1/\xi)} \mathcal{Q}_t(r)^{(1-\chi)\mu} w_t(r) L_t(r)^{1-\mu-\gamma_1/\xi}\right)^{\theta} \end{aligned} \tag{50}$$

We insert equation (46) in the right-hand side of equation (49).

$$\begin{aligned}
&\int_S P_t(v)^\theta \left(\frac{w_t(v)H(v)L_t(v)}{\varphi_t(v)}\right)\varsigma(r,v)^{-\theta}dv \\
&= \int_S \left(\frac{\bar{b}_t(v)}{u_t(v)}L_t(v)^{-\lambda}\left(\frac{1}{\mu+\gamma_1/\xi}\right)\left(\frac{w_t(v)}{\varphi_t(v)}\right)\right)^\theta \left(\frac{w_t(v)H(v)L_t(v)}{\varphi_t(v)}\right)\varsigma(r,v)^{-\theta}dv
\end{aligned} \tag{51}$$

We manipulate equations (50) and (51) to obtain equation (52).

$$\begin{aligned}
&\bar{a}_t(r)^{-1}w_t(r)^{1+\theta}L_t(r)^{1-\alpha+(1-\mu-\gamma_1/\xi)\theta}\mathcal{Q}_t(r)^{(1-\chi)\mu\theta}\varphi_t(r)^{-(1+(1-\mu-\gamma_1/\xi)\theta)}H(r) \\
&= \kappa_1 \int_S \left(\frac{\bar{b}_t(v)}{u_t(v)}\right)^\theta L_t(v)^{1-\lambda\theta}\varphi_t(v)^{-(1+\theta)}w_t(v)^{1+\theta}H(v)\varsigma(r,v)^{-\theta}dv
\end{aligned} \tag{52}$$

Equations (3), (4), (9), (11), (12), (48) and (52) define a system of equations that solves for $u_t(r), L_t(r), w_t(r)$ and $\mathcal{Q}_t(r)$. Assuming trade costs are symmetric, we use equations (48) and (52) and introduce the function $f_1(\cdot)$, which is the ratio of the left-hand sides of (52) and (48).

$$\begin{aligned}
f_1(r) &= \frac{\bar{a}_t(r)^{-1}w_t(r)^{1+\theta}L_t(r)^{1-\alpha+(1-\mu-\gamma_1/\xi)\theta}\mathcal{Q}_t(r)^{(1-\chi)\mu\theta}\varphi_t(r)^{-(1+(1-\mu-\gamma_1/\xi)\theta)}H(r)}{\left(\frac{\bar{b}_t(r)}{u_t(r)}\right)^{-\theta}L_t(r)^{\lambda\theta}w_t(r)^{-\theta}\varphi_t(r)^\theta} \\
&= \bar{a}_t(r)^{-1}w_t(r)^{1+2\theta}L_t(r)^{1-\alpha-\theta(\lambda+\gamma_1/\xi-(1-\mu)} \\
&\times \varphi_t(r)^{-(1+\theta(2-\mu-\gamma_1/\xi)}\mathcal{Q}_t(r)^{(1-\chi)\mu\theta}H(r)\left(\frac{\bar{b}_t(r)}{u_t(r)}\right)^\theta
\end{aligned} \tag{53}$$

Obviously, $f_1(r)$ equals the right-hand sides.

$$f_1(r) = \frac{\int_S \left(\frac{\bar{b}_t(v)}{u_t(v)}\right)^\theta L_t(v)^{1-\lambda\theta}\varphi_t(v)^{-(1+\theta)}w_t(v)^{1+\theta}H(v)\varsigma(r,v)^{-\theta}dv}{\int_S \bar{a}_t(v)L_t(v)^{\alpha-(1-\mu-\gamma_1/\xi)\theta}\varphi_t(v)^{(1-\mu-\gamma_1/\xi)\theta}\mathcal{Q}_t(v)^{-(1-\chi)\mu\theta}w_t(v)^{-\theta}\varsigma(r,v)^{-\theta}dv}$$

Under the assumption that $\varsigma(r,v) = \varsigma(v,r)$, we can express $f_1(\cdot)$ as shown below.

$$\begin{aligned}
f_1(r) &= \frac{\int_S f_1(v)^{-\lambda}f_2(v,r)dv}{\int_S f_1(v)^{-(1+\lambda)}f_2(v,r)dv} \\
f_2(v,r) &= \left(\frac{\bar{b}_t(v)}{u_t(v)}\right)^{\theta(1+\lambda)}\bar{a}_t(v)^{-\lambda}L_t(v)^{1-\lambda\theta-\lambda[\alpha-1+\theta(\lambda+\gamma_1/\xi-(1-\mu))]}\mathcal{Q}_t(v)^{(1-\chi)\mu\theta\lambda} \\
&\times \varphi_t(v)^{-(1+\lambda)(1+\theta)+\lambda\theta(\gamma_1/\xi-(1-\mu)}w_t(v)^{(1+\lambda)(1+\theta)+\theta}H(v)^{1+\lambda}\varsigma(v,r)^{-\theta}
\end{aligned}$$

We follow the procedure of Desmet et al. (2018) and Theorem 2.1 in Zabreyko et al. (1975) to argue that

$f_1(r) = f_1$. Then, we solve for $w_t(r)/\varphi_t(r)$ using equation (53).

$$\left(\frac{w_t(r)}{\varphi_t(r)}\right) = f_1^{\frac{1}{1+2\theta}} \bar{a}_t(r)^{\frac{1}{1+2\theta}} L_t(r)^{-\frac{1-\alpha-\theta(\lambda+\gamma_1/\xi-(1-\mu)}{1+2\theta}} \times \varphi_t(r)^{-\frac{\mu+\gamma_1/\xi}{1+2\theta}} \mathcal{Q}_t(r)^{-\frac{(1-\chi)\mu\theta}{1+2\theta}} H(r)^{-\frac{1}{1+2\theta}} \left(\frac{\bar{b}_t(r)}{u_t(r)}\right)^{-\frac{\theta}{1+2\theta}}$$

We insert the expression for $w_t(r)/\varphi_t(r)$ in equation (48) and manipulate terms to obtain equation (54).

$$\begin{aligned}
&\left(\frac{\bar{b}_t(r)}{u_t(r)}\right)^{-\frac{\theta(1+\theta)}{1+2\theta}} \bar{a}_t(r)^{-\frac{\theta}{1+2\theta}} L_t(r)^{\flat_L} \\
&\times \mathcal{Q}_t(r)^{\frac{(1-\chi)\mu\theta^2}{1+2\theta}} H(r)^{\frac{\theta}{1+2\theta}} \varphi_t(r)^{-\frac{(\mu+\gamma_1/\xi)\theta^2}{1+2\theta}} \\
&= \kappa_1 \int_S \left(\frac{\bar{b}_t(v)}{u_t(v)}\right)^{\frac{\theta^2}{1+2\theta}} \bar{a}_t(v)^{\frac{1+\theta}{1+2\theta}} L_t(v)^{\flat_R} \mathcal{Q}_t(v)^{-(1-\chi)\mu\frac{\theta(1+\theta)}{1+2\theta}} \\
&\times H(v)^{\frac{\theta}{1+2\theta}} \varphi_t(v)^{\frac{(\mu+\gamma_1/\xi)\theta(1+\theta)}{1+2\theta}} \varsigma(r,v)^{-\theta} dv
\end{aligned} \tag{54}$$

Where the terms $\flat_L$ and $\flat_R$ are defined as follow.

$$\begin{aligned}
\flat_L &= \lambda\theta - \frac{\theta}{1+2\theta}[\alpha - 1 + \theta(\lambda + \gamma_1/\xi - (1-\mu))] \\
\flat_R &= 1 - \lambda\theta + \frac{1+\theta}{1+2\theta}[\alpha - 1 + \theta(\lambda + \gamma_1/\xi - (1-\mu))]
\end{aligned}$$

We utilize the expression for migration shares, given by equation (3), and solve for $L_t(r)$.

$$L_t(r) = H(r)^{-1} u_t(r)^{1/\Omega} m_2(r)^{-1/\Omega} \left(\frac{L_t}{\int_S u_t(v)^{1/\Omega} m_2(v)^{-1/\Omega} dv}\right) \tag{55}$$

We employ this expression and insert it in equation (54).

$$\begin{aligned}
&\bar{b}_t(r)^{-\frac{\theta(1+\theta)}{1+2\theta}} \bar{a}_t(r)^{-\frac{\theta}{1+2\theta}} \mathcal{Q}_t(r)^{\frac{(1-\chi)\mu\theta^2}{1+2\theta}} H(r)^{\frac{\theta}{1+2\theta}-\flat_L} \\
&\times \varphi_t(r)^{-\frac{(\mu+\gamma_1/\xi)\theta^2}{1+2\theta}} m_2(r)^{-\frac{\flat_L}{\Omega}} u_t(r)^{\frac{\flat_L}{\Omega}+\frac{\theta(1+\theta)}{1+2\theta}} \\
&= \kappa_1 \left(\frac{L_t}{\int_S u_t(v)^{1/\Omega} m_2(v)^{-1/\Omega} dv}\right)^{\flat_R-\flat_L} \\
&\times \int_S \bar{b}_t(v)^{\frac{\theta^2}{1+2\theta}} \bar{a}_t(v)^{\frac{1+\theta}{1+2\theta}} \mathcal{Q}_t(v)^{-(1-\chi)\mu\frac{\theta(1+\theta)}{1+2\theta}} \\
&\times H(v)^{\frac{\theta}{1+2\theta}-\flat_R} \varphi_t(v)^{\frac{(\mu+\gamma_1/\xi)\theta(1+\theta)}{1+2\theta}} m_2(v)^{-\frac{\flat_R}{\Omega}} u_t(r)^{\frac{\flat_R}{\Omega}-\frac{\theta^2}{1+2\theta}} \varsigma(r,v)^{-\theta} dv
\end{aligned} \tag{56}$$

We combine equations (9), (11) and (12) to find an expression for the price of energy.

$$\begin{aligned}\mathcal{Q}_t(r) &= \left(\kappa^{\epsilon}(1+\tau_t(r))^{1-\epsilon}\mathcal{Q}_t^f(r)^{1-\epsilon} + (1-\kappa)^{\epsilon}(1-s_t(r))^{1-\epsilon}\mathcal{Q}_t^c(r)^{1-\epsilon}\right)^{\frac{1}{1-\epsilon}} \\ &= \left(\kappa^{\epsilon}(1+\tau_t(r))^{1-\epsilon}\left(\frac{f(CumCO2_t)}{\zeta_t^f(r)}\right)^{1-\epsilon} + (1-\kappa)^{\epsilon}(1-s_t(r))^{1-\epsilon}\left(\frac{1}{\zeta_t^c(r)}\right)^{1-\epsilon}\right)^{\frac{1}{1-\epsilon}} \\ &= \left(\kappa^{\epsilon}(1+\tau_t(r))^{1-\epsilon}\left(\frac{f(CumCO2_t)}{\left(y_t^w/y_{t-1}^w\right)^{\upsilon^j}}\right)^{1-\epsilon} + (1-\kappa)^{\epsilon}(1-s_t(r))^{1-\epsilon}\left(\frac{1}{\left(y_t^w/y_{t-1}^w\right)^{\upsilon^j}}\right)^{1-\epsilon}\right)^{\frac{1}{1-\epsilon}} \end{aligned} \tag{57}$$

We rewrite global average real GDP, y_t^w, making use of equations (3) and (4).

$$\begin{aligned} y_t^w &= \int_S \left(\frac{L_t(v)H(v)}{L_t}\right) y_t(v)dv \\ &= \left(\frac{L_t}{\int_S u_t(v)^{1/\Omega} m_2(v)^{-1/\Omega} dv}\right)^{1+\lambda} \int_S L_t^{-1} H(v)^{-\lambda} u_t(v)^{1+\frac{1}{\Omega}[1+\lambda]} m_2(v)^{-\frac{1}{\Omega}[1+\lambda]} \bar{b}_t(v)^{-1} dv \end{aligned} \tag{58}$$

Therefore, equations (56), (57) and (58) define a system of equations that solves for $u_t(r)$ and $\mathcal{Q}_t(r)$. In order to guarantee existence and uniqueness of the solution, we can consider (i) $\epsilon = 1$ or (ii) $\upsilon^f = \upsilon^c = \upsilon$.

In the first case, the energy price $\mathcal{Q}_t(r)$ collapses to:

$$\mathcal{Q}_t(r) = \left(\frac{y_t^w}{y_{t-1}^w}\right)^{-(\kappa\upsilon^f + (1-\kappa)\upsilon^c)} \left(\frac{(1+\tau_t(r))f(CumCO2_t)}{\kappa}\right)^{\kappa} \left(\frac{(1-s_t(r))}{(1-\kappa)}\right)^{1-\kappa}$$

In the second case, the energy price $\mathcal{Q}_t(r)$ collapses to:

$$\mathcal{Q}_t(r) = \left(\frac{y_t^w}{y_{t-1}^w}\right)^{-\upsilon} \left(\kappa^{\epsilon}(1+\tau_t(r))^{1-\epsilon} f(CumCO2_t)^{1-\epsilon} + (1-\kappa)^{\epsilon}(1-s_t(r))^{1-\epsilon}\right)^{\frac{1}{1-\epsilon}}$$

Consequently, in both cases, we can represent $\mathcal{Q}_t(r)$ as the product a location-invariant term and an exogenous term $\hat{\mathcal{Q}}_t(r)$.

$$\mathcal{Q}_t(r) = \left(\frac{y_t^w}{y_{t-1}^w}\right)^{-\hat{\upsilon}} \hat{\mathcal{Q}}_t(r) \tag{59}$$

We insert equation (59) in equation (56).

$$\begin{aligned} B_{1,t}(r)\hat{u}_t(r)^{\frac{\flat_L}{\Omega} + \frac{\theta(1+\theta)}{1+2\theta})} &= \kappa_1 \int_S B_{2,t}(v)\hat{u}_t(v)^{\frac{\flat_R}{\Omega} - \frac{\theta^2}{1+2\theta}} \varsigma(r,v)^{-\theta} dv \\ B_{1,t}(r) &= \bar{b}_t(r)^{-\frac{\theta(1+\theta)}{1+2\theta}} \bar{a}_t(r)^{-\frac{\theta}{1+2\theta}} \hat{\mathcal{Q}}_t(r)^{\frac{(1-\chi)\mu\theta^2}{1+2\theta}} \\ &\quad \times H(r)^{\frac{\theta}{1+2\theta} - \flat_L} \varphi_t(r)^{-\frac{(\mu+\gamma_1/\xi)\theta^2}{1+2\theta}} m_2(r)^{-\frac{\flat_L}{\Omega}} \\ B_{2,t}(v) &= \bar{b}_t(v)^{\frac{\theta^2}{1+2\theta}} \bar{a}_t(v)^{\frac{1+\theta}{1+2\theta}} \hat{\mathcal{Q}}_t(v)^{-(1-\chi)\mu\frac{\theta(1+\theta)}{1+2\theta}} \\ &\quad \times H(v)^{\frac{\theta}{1+2\theta} - \flat_R} \varphi_t(v)^{\frac{(\mu+\gamma_1/\xi)\theta(1+\theta)}{1+2\theta}} m_2(v)^{-\frac{\flat_R}{\Omega}} \end{aligned} \tag{60}$$

Where $B_{1,t}(\cdot)$ and $B_{2,t}(\cdot)$ are exogenous functions and $\hat{u}_t(\cdot)$ is given by:

$$\hat{u}_t(r) = u_t(r)\left(y^w_{t-1}\right)^{-\hat{\upsilon}(1-\chi)\mu\theta}\left(\frac{L_t}{\int_S u_t(v)^{1/\Omega}m_2(v)^{-1/\Omega}dv}\right)^{\flat_R-\flat_L+\hat{\upsilon}(1-\chi)\mu\theta(1+\lambda)}$$
$$\times\left(\int_S L_t^{-1}H(v)^{-\lambda}u_t(v)^{1+\frac{1}{\Omega}[1+\lambda]}m_2(v)^{-\frac{1}{\Omega}[1+\lambda]}\bar{b}_t(v)^{-1}dv\right)^{\hat{\upsilon}(1-\chi)\mu\theta}$$

It follows from theorem 2.19 in Zabreyko et al. (1975) that the solution to equation (60) exists and is unique if:

$$\frac{\flat_R}{\Omega} - \frac{\theta^2}{1+2\theta} \leq \frac{\flat_L}{\Omega} + \frac{\theta(1+\theta)}{1+2\theta}$$
$$\frac{\alpha}{\theta} + \frac{\gamma_1}{\xi} \leq \lambda + \Omega + (1-\mu) \tag{61}$$

And $u_t(r)$ can be retrieved as:

$$u_t(r) = \hat{u}_t(r)\left(y^w_{t-1}\right)^{\frac{\hat{\upsilon}(1-\chi)\mu\theta}{\hat{\upsilon}(1-\chi)\mu\theta-\theta}}\left(\frac{L_t}{\int_S \hat{u}_t(v)^{1/\Omega}m_2(v)^{-1/\Omega}dv}\right)^{-\frac{\flat_R-\flat_L+\hat{\upsilon}(1-\chi)\mu\theta(1+\lambda)}{\hat{\upsilon}(1-\chi)\mu\theta-\theta}}$$
$$\times\left(\int_S L_t^{-1}H(v)^{-\lambda}\hat{u}_t(v)^{1+\frac{1}{\Omega}[1+\lambda]}m_2(v)^{-\frac{1}{\Omega}[1+\lambda]}\bar{b}_t(v)^{-1}dv\right)^{-\frac{\hat{\upsilon}(1-\chi)\mu\theta}{\hat{\upsilon}(1-\chi)\mu\theta-\theta}}$$

A.2 Fundamental Amenities and Productivities

To solve for the fundamentals $\bar{b}_t(r)/u_t(r)$ and $\bar{a}_t(r)$, employ the system of equations defined by equations (48) and (53) of Appendix A.1.

A.3 Balanced Growth Path

To prove the existence and uniqueness of a Balanced Growth Path, recall the evolution of amenities and technology, given by:

$$\bar{b}_t(r) = \left(1+\delta^b(T_t(r))\cdot\Delta T_t(r)\right)\bar{b}_{t-1}(r)$$
$$\bar{a}_t(r) = \left(1+\delta^a(T_t(r))\cdot\Delta T_t(r)\right)\left(\phi_{t-1}(r)^{\theta\gamma_1}\left[\int_S D(v,r)\bar{a}_{t-1}(v)dv\right]^{1-\gamma_2}\bar{a}_{t-1}(r)^{\gamma_2}\right)$$

For fundamental amenities to be constant over time, we require that local temperature reaches a Steady State for every cell. To show this behavior, remember that as the cost of extracting fossil fuels rises sharply, its use declines towards zero. Hence, in the long-run, carbon dioxide emissions converge to zero.

By assuming that emissions from deforestation also converge to zero in the long-run and the share of CO_2 emissions remaining in the atmosphere ℓ periods ahead, $(1-\delta_\ell)$, converge to a constant value in the long-run, then the carbon stock in the atmosphere also converges to a constant value.

More specifically, employing the characterization of the carbon cycle outlined in equations (86)-(88), $S_{i,t+1}$ converges to zero, $i \in \{1,2,3\}$, and $S_{0,t+1}$ converges to $S^* = S_{0,2000} + a_0(maxCumCO2^{f,x})$, where $maxCumCO2^{f,x}$ represents the cumulative and finite flow of emissions from fuel combustion and deforestation.

By assuming that radiative forcing from non-CO_2 GHG converges to zero in the long-run, then radiative forcing, F_{t+1}, converges to a stable value given by $F^* = \varphi \log_2(S^*/S_{\text{pre-ind}})$, according to equation (19). Finally, under the assumption that the temperature response to an increase in radiative force ℓ periods ago, ζ_ℓ, converges to a constant value, global temperature reaches a Steady State. More precisely, employing equations (89) and (90), global temperature reaches a Steady State given by $T^* = \left(\frac{c_1/d_1}{1-e^{-1/d_1}} + \frac{c_2/d_2}{1-e^{-1/d_2}}\right) F^*$. Since global temperature achieves a stable level in the long-run, equation (21) implies the same conclusion for local temperature.

By the aforementioned arguments, fundamental amenities are constant in the Balanced Growth Path. When additionally imposing that total natality rates, $1 + n_t(r)$, converge to one as income grows, global population achieves a stable value, L^*. Consequently, productivity growth rates follow:

$$\frac{\bar{a}_{t+1}(r)}{\bar{a}_t(r)} = \phi_t(r)^{\theta\gamma_1} \left[\int_S D \frac{\bar{a}_t(v)}{\bar{a}_t(r)} dv\right]^{1-\gamma_2} \tag{62}$$

In a BGP in which technology growth rates are constant, so $\bar{a}_{t+1}(r)/\bar{a}_t(r)$ is constant over time and space and $\bar{a}_t(s)/\bar{a}_t(r)$ is constant over time, the investment decision will also be constant over time but potentially different across locations. We divide equation (62) evaluated at region r with respect to that evaluated at region s to obtain:

$$\frac{\bar{a}_t(s)}{\bar{a}_t(r)} = \left(\frac{\phi(s)}{\phi(r)}\right)^{\frac{\theta\gamma_1}{1-\gamma_2}} = \left(\frac{L(s)/\varphi_t(s)}{L(r)/\varphi_t(r)}\right)^{\frac{\theta\gamma_1}{\xi(1-\gamma_2)}} \tag{63}$$

Where the last equality comes from equation (40). We arrange equation (63), integrate over s and solve for $\bar{a}_t(r)$.

$$\begin{aligned} L^* &= \int_S L(s)H(s)ds = \int_S \left(\frac{\bar{a}_t(s)}{\bar{a}_t(r)}\right)^{\frac{\xi(1-\gamma_2)}{\theta\gamma_1}} \left(\frac{\varphi_t(s)}{\varphi_t(r)}\right) L(r)H(s)ds \\ \bar{a}_t(r) &= \kappa_{2,t} \left(\frac{L(r)}{\varphi_t(r)}\right)^{\frac{\theta\gamma_1}{\xi(1-\gamma_2)}} \\ \kappa_{2,t} &= \left(\int_S \bar{a}_t(s)^{\frac{\xi(1-\gamma_2)}{\theta\gamma_1}} \varphi_t(s)H(s)ds\right)^{\frac{\theta\gamma_1}{\xi(1-\gamma_2)}} (L^*)^{-\frac{\theta\gamma_1}{\xi(1-\gamma_2)}} \end{aligned} \tag{64}$$

We insert equation (64) in equation (54) and define $\sharp_L$ and $\sharp_R$ as shown below.

$$\begin{aligned}
&\left(\frac{\bar{b}(r)}{u_t(r)}\right)^{-\frac{\theta(1+\theta)}{1+2\theta}} L(r)^{\sharp_L}\mathcal{Q}_t(r)^{\frac{(1-\chi)\mu\theta^2}{1+2\theta}} H(r)^{\frac{\theta}{1+2\theta}}\varphi_t(r)^{-\frac{\theta^2}{1+2\theta}\left(\mu+\frac{\gamma_1}{\xi}-\frac{\gamma_1}{\xi(1-\gamma_2)}\right)} \\
&= \kappa_1\kappa_{2,t}\int_S \left(\frac{\bar{b}(v)}{u_t(v)}\right)^{\frac{\theta^2}{1+2\theta}} L(v)^{\sharp_R}\mathcal{Q}_t(v)^{-(1-\chi)\mu\frac{\theta(1+\theta)}{1+2\theta}} H(v)^{\frac{\theta}{1+2\theta}}\varphi_t(v)^{\frac{\theta(1+\theta)}{1+2\theta}\left(\mu+\frac{\gamma_1}{\xi}-\frac{\gamma_1}{\xi(1-\gamma_2)}\right)}\varsigma(r,v)^{-\theta}dv \quad (65) \\
&\sharp_L = \lambda\theta - \frac{\theta}{1+2\theta}\left[\alpha - 1 + \theta\left(\lambda + \frac{\gamma_1}{\xi} + \frac{\gamma_1}{\xi(1-\gamma_2)} - (1-\mu)\right)\right] \\
&\sharp_R = 1 - \lambda\theta + \frac{1+\theta}{1+2\theta}\left[\alpha - 1 + \theta\left(\lambda + \frac{\gamma_1}{\xi} + \frac{\gamma_1}{\xi(1-\gamma_2)} - (1-\mu)\right)\right]
\end{aligned}$$

We substitute the migration shares, given by equation (55), in equation (65) and arrange terms.

$$\begin{aligned}
&\bar{b}(r)^{-\frac{\theta(1+\theta)}{1+2\theta}}\mathcal{Q}_t(r)^{\frac{(1-\chi)\mu\theta^2}{1+2\theta}} H(r)^{\frac{\theta}{1+2\theta}-\sharp_L} \\
&\times \varphi_t(r)^{-\frac{\theta^2}{1+2\theta}\left(\mu+\frac{\gamma_1}{\xi}-\frac{\gamma_1}{\xi(1-\gamma_2)}\right)} m_2(r)^{-\frac{\sharp_L}{\Omega}} u_t(r)^{\frac{\sharp_L}{\Omega}+\frac{\theta(1+\theta)}{1+2\theta}} \\
&= \kappa_1\kappa_{2,t}\left(\frac{L^*}{\int_S u_t(v)^{1/\Omega} m_2(v)^{-1/\Omega}dv}\right)^{\sharp_R-\sharp_L} \\
&\times \int_S \bar{b}(v)^{\frac{\theta^2}{1+2\theta}}\mathcal{Q}_t(v)^{-(1-\chi)\mu\frac{\theta(1+\theta)}{1+2\theta}} H(v)^{\frac{\theta}{1+2\theta}-\sharp_R} \\
&\times \varphi_t(v)^{\frac{\theta(1+\theta)}{1+2\theta}\left(\mu+\frac{\gamma_1}{\xi}-\frac{\gamma_1}{\xi(1-\gamma_2)}\right)} m_2(v)^{-\frac{\sharp_R}{\Omega}} u_t(r)^{\frac{\sharp_R}{\Omega}-\frac{\theta^2}{1+2\theta}}\varsigma(r,v)^{-\theta}dv \quad (66)
\end{aligned}$$

For the price of energy to be the product of a location-invariant term and a location-specific time-invariant term, $\mathcal{Q}_t(r) = \left(\frac{y_t^w}{y_{t-1}^w}\right)^{-\hat{\upsilon}}\mathcal{Q}(r)$, we assume (i) $\epsilon = 1$ or (ii) $\upsilon^f = \upsilon^c = \upsilon$.

When considering $\epsilon = 1$, $\varphi_t(r)$ collapses to:

$$\begin{aligned}
\varphi_t(r) &= \frac{\mu\chi + \gamma_1/\xi + \mu(1-\chi)\tilde{\kappa}(r)}{\mu + \gamma_1/\xi} \\
\lim_{\epsilon\to 1}\left(\frac{\tilde{\mathcal{Q}}_t(r)}{\mathcal{Q}_t(r)}\right)^{1-\epsilon} &= \lim_{\epsilon\to 1}\left(\frac{\kappa^\epsilon(1+\tau(r))^{-\epsilon}\mathcal{Q}_t^f(r)^{1-\epsilon} + (1-\kappa)^\epsilon(1-s(r))^{-\epsilon}\mathcal{Q}_t^c(r)^{1-\epsilon}}{\kappa^\epsilon(1+\tau(r))^{1-\epsilon}\mathcal{Q}_t^f(r)^{1-\epsilon} + (1-\kappa)^\epsilon(1-s(r))^{1-\epsilon}\mathcal{Q}_t^c(r)^{1-\epsilon}}\right) \\
&= \left(\frac{\kappa}{1+\tau(r)} + \frac{1-\kappa}{1-s(r)}\right) = \tilde{\kappa}(r)
\end{aligned}$$

And when considering $\upsilon^f = \upsilon^c = \upsilon$, $\varphi_t(r)$ collapses to:

$$\begin{aligned}
\varphi_t(r) &= \frac{\mu\chi + \gamma_1/\xi + \mu(1-\chi)\breve{\kappa}(r)}{\mu + \gamma_1/\xi} \\
\left(\frac{\tilde{\mathcal{Q}}_t(r)}{\mathcal{Q}_t(r)}\right)^{1-\epsilon} &= \left(\frac{\kappa^\epsilon(1+\tau(r))^{-\epsilon}(\tilde{f})^{1-\epsilon}(y_t^w/y_{t-1}^w)^{-\upsilon(1-\epsilon)} + (1-\kappa)^\epsilon(1-s(r))^{-\epsilon}(y_t^w/y_{t-1}^w)^{-\upsilon(1-\epsilon)}}{\kappa^\epsilon(1+\tau(r))^{1-\epsilon}(\tilde{f})^{1-\epsilon}(y_t^w/y_{t-1}^w)^{-\upsilon(1-\epsilon)} + (1-\kappa)^\epsilon(1-s(r))^{1-\epsilon}(y_t^w/y_{t-1}^w)^{-\upsilon(1-\epsilon)}}\right) \\
&= \left(\frac{\kappa^\epsilon(1+\tau(r))^{-\epsilon}(\tilde{f})^{1-\epsilon} + (1-\kappa)^\epsilon(1-s(r))^{-\epsilon}}{\kappa^\epsilon(1+\tau(r))^{1-\epsilon}(\tilde{f})^{1-\epsilon} + (1-\kappa)^\epsilon(1-s(r))^{1-\epsilon}}\right) = \breve{\kappa}(r)
\end{aligned}$$

Consequently, $\varphi_t(r)$ might vary across locations, but is constant over time. Now, we combine equation

(66) and the definition of the energy price and manipulate terms.

$$
\begin{aligned}
B_1(r)\hat{u}_t(r)^{\frac{\sharp_L}{\Omega}+\frac{\theta(1+\theta)}{1+2\theta}} &= \kappa_1\kappa_{2,t}\int_S B_2(v)\hat{u}_t(v)^{\frac{\sharp_R}{\Omega}-\frac{\theta^2}{1+2\theta}}\varsigma(r,v)^{-\theta}dv \qquad (67)\\
B_1(r) &= \bar{b}(r)^{-\frac{\theta(1+\theta)}{1+2\theta}}\mathcal{Q}(r)^{\frac{(1-\chi)\mu\theta^2}{1+2\theta}}H(r)^{\frac{\theta}{1+2\theta}-\sharp_L}\varphi(r)^{-\frac{\theta^2}{1+2\theta}\left(\mu+\frac{\gamma_1}{\xi}-\frac{\gamma_1}{\xi(1-\gamma_2)}\right)}m_2(r)^{-\frac{\sharp_L}{\Omega}}\\
B_2(r) &= \bar{b}(r)^{\frac{\theta^2}{1+2\theta}}\mathcal{Q}(r)^{-\frac{(1-\chi)\mu\theta(1+\theta)}{1+2\theta}}H(r)^{\frac{\theta}{1+2\theta}-\sharp_R}\varphi(r)^{\frac{\theta(1+\theta)}{1+2\theta}\left(\mu+\frac{\gamma_1}{\xi}-\frac{\gamma_1}{\xi(1-\gamma_2)}\right)}m_2(r)^{-\frac{\sharp_R}{\Omega}}
\end{aligned}
$$

Where $B_1(\cdot)$ and $B_2(\cdot)$ are exogenous functions and $\hat{u}_t(\cdot)$ is given by:

$$
\begin{aligned}
\hat{u}_t(r) = u_t(r)\left(y^w_{t-1}\right)^{-\hat{\upsilon}(1-\chi)\mu\theta}&\left(\frac{L^*}{\int_S u_t(v)^{1/\Omega}m_2(v)^{-1/\Omega}dv}\right)^{\sharp_R-\sharp_L+\hat{\upsilon}(1-\chi)\mu\theta(1+\lambda)}\\
\times&\left(\int_S L^{*-1}H(v)^{-\lambda}u_t(v)^{1+\frac{1}{\Omega}[1+\lambda]}m_2(v)^{-\frac{1}{\Omega}[1+\lambda]}\bar{b}(v)^{-1}dv\right)^{\hat{\upsilon}(1-\chi)\mu\theta}
\end{aligned}
$$

It follows from theorem 2.19 in Zabreyko et al. (1975) that the solution to equation (67) exists and is unique if:

$$
\begin{aligned}
\frac{\sharp_R}{\Omega}-\frac{\theta^2}{1+2\theta} &\leq \frac{\sharp_L}{\Omega}+\frac{\theta(1+\theta)}{1+2\theta}\\
\frac{\alpha}{\theta}+\frac{\gamma_1}{\xi}+\frac{\gamma_1}{\xi(1-\gamma_2)} &\leq \lambda+\Omega+(1-\mu) \qquad (68)
\end{aligned}
$$

And $u_t(\cdot)$ can be retrieved as:

$$
\begin{aligned}
u_t(r) = \hat{u}_t(r)\left(y^w_{t-1}\right)^{\frac{\hat{\upsilon}(1-\chi)\mu\theta}{\hat{\upsilon}(1-\chi)\mu\theta-\theta}}&\left(\frac{L^*}{\int_S \hat{u}_t(v)^{1/\Omega}m_2(v)^{-1/\Omega}dv}\right)^{-\frac{\sharp_R-\sharp_L+\hat{\upsilon}(1-\chi)\mu\theta(1+\lambda)}{\hat{\upsilon}(1-\chi)\mu\theta-\theta}}\\
\times&\left(\int_S L_t^{-1}H(v)^{-\lambda}\hat{u}_t(v)^{1+\frac{1}{\Omega}[1+\lambda]}m_2(v)^{-\frac{1}{\Omega}[1+\lambda]}\bar{b}(v)^{-1}dv\right)^{-\frac{\hat{\upsilon}(1-\chi)\mu\theta}{\hat{\upsilon}(1-\chi)\mu\theta-\theta}}
\end{aligned}
$$

A.4 Backward Solution

Rewrite equations (2) and (7) as shown below and define the terms $\natural_L$, $\natural_R$ and κ_2.

$$\begin{aligned}
\bar{b}_t(r) &= \frac{\bar{b}_{t+1}(r)}{\Lambda^b_{t+1}(r)} \\
\bar{a}_t(r) &= \left(\frac{\mu+\gamma_1/\xi}{\gamma_1/\xi}\nu\right)^{\frac{\theta\gamma_1}{\xi\gamma_2}} \left(\int_S D(r,v)\bar{a}_t(v)\right)^{\frac{\gamma_2-1}{\gamma_2}} \left(\frac{\bar{a}_{t+1}(r)}{\Lambda^a_{t+1}(r)}\right)^{\frac{1}{\gamma_2}} L_t(r)^{-\frac{\theta\gamma_1}{\xi\gamma_2}} \\
\Lambda^x_{t+1}(r) &= \Lambda^x(T_{t+1}(r), \Delta T_t(r)), \quad x \in \{a,b\} \\
\natural_L &= \lambda\theta - \frac{\theta}{1+2\theta}[\alpha - 1 + \theta(\lambda + \gamma_1/\xi - \gamma_1/(\xi\gamma_2) - (1-\mu))] \\
\natural_R &= 1 - \lambda\theta + \frac{1+\theta}{1+2\theta}[\alpha - 1 + \theta(\lambda + \gamma_1/\xi - \gamma_1/(\xi\gamma_2) - (1-\mu))] \\
\kappa_2 &= \kappa_1 \left(\frac{\mu+\gamma_1/\xi}{\gamma_1/\xi}\nu\right)^{\frac{\theta\gamma_1}{\xi\gamma_2}}
\end{aligned}$$

Then, insert them in equation (54).

$$\begin{aligned}
&\left(\frac{\bar{b}_{t+1}(r)}{\Lambda^b_{t+1}(r)}\right)^{-\frac{\theta(1+\theta)}{1+2\theta}} \left(\left(\int_S D(r,s)\bar{a}_t(s)ds\right)^{\frac{\gamma_2-1}{\gamma_2}} \left(\frac{\bar{a}_{t+1}(r)}{\Lambda^a_{t+1}(r)}\right)^{\frac{1}{\gamma_2}}\right)^{-\frac{\theta}{1+2\theta}} \\
&\times L_t(r)^{\natural_L} \mathcal{Q}_t(r)^{\frac{(1-\chi)\mu\theta^2}{1+2\theta}} H(r)^{\frac{\theta}{1+2\theta}} \varphi_t(r)^{\frac{-(\mu+\gamma_1/\xi)\theta^2}{1+2\theta}} u_t(r)^{\frac{\theta(1+\theta)}{1+2\theta}} \\
&= \kappa_2 \int_S \left(\frac{\bar{b}_{t+1}(v)}{\Lambda^b_{t+1}(v)}\right)^{\frac{\theta^2}{1+2\theta}} \left(\left(\int_S D(v,s)\bar{a}_t(s)ds\right)^{\frac{\gamma_2-1}{\gamma_2}} \left(\frac{\bar{a}_{t+1}(v)}{\Lambda^a_{t+1}(v)}\right)^{\frac{1}{\gamma_2}}\right)^{\frac{1+\theta}{1+2\theta}} \\
&\times L_t(v)^{\natural_R} \mathcal{Q}_t(v)^{-(1-\chi)\mu\frac{\theta(1+\theta)}{1+2\theta}} H(v)^{\frac{\theta}{1+2\theta}} \varphi_t(v)^{(\mu+\gamma_1/\xi)\frac{\theta(1+\theta)}{1+2\theta}} u_t(v)^{-\frac{\theta^2}{1+2\theta}} \varsigma(r,v)^{-\theta} dv
\end{aligned} \tag{69}$$

And substitute the migration shares, given by equation (4), in equation (69), assuming $D(r,s) = D$.

$$\begin{aligned}
&\left(\frac{\bar{b}_{t+1}(r)}{\Lambda^b_{t+1}(r)}\right)^{-\frac{\theta(1+\theta)}{1+2\theta}} \left(\frac{\bar{a}_{t+1}(r)}{\Lambda^a_{t+1}(r)}\right)^{-\frac{\theta}{1+2\theta}\frac{1}{\gamma_2}} \\
&\times \mathcal{Q}_t(r)^{\frac{(1-\chi)\mu\theta^2}{1+2\theta}} H(r)^{-\natural_L+\frac{\theta}{1+2\theta}} \varphi_t(r)^{-\frac{(\mu+\gamma_1/\xi)\theta^2}{1+2\theta}} m_2(r)^{-\frac{\natural_L}{\Omega}} u_t(r)^{\frac{\natural_L}{\Omega}+\frac{\theta(1+\theta)}{1+2\theta}} \\
&= \kappa_2 \left(\int_S D\bar{a}_t(s)ds\right)^{\frac{\gamma_2-1}{\gamma_2}} \left(\frac{L_t}{\int_S u_t(v)^{1/\Omega} m_2(v)^{-1/\Omega} dv}\right)^{\natural_R-\natural_L} \\
&\int_S \left(\frac{\bar{b}_{t+1}(v)}{\Lambda^b_{t+1}(v)}\right)^{\frac{\theta^2}{1+2\theta}} \left(\frac{\bar{a}_{t+1}(v)}{\Lambda^a_{t+1}(v)}\right)^{\frac{1+\theta}{1+2\theta}\frac{1}{\gamma_2}} \\
&\times \mathcal{Q}_t(v)^{-(1-\chi)\mu\frac{\theta(1+\theta)}{1+2\theta}} H(v)^{-\natural_R+\frac{\theta}{1+2\theta}} \varphi_t(v)^{\frac{(\mu+\gamma_1/\xi)\theta(1+\theta)}{1+2\theta}} m_2(v)^{-\frac{\natural_R}{\Omega}} u_t(v)^{\frac{\natural_R}{\Omega}-\frac{\theta^2}{1+2\theta}} \varsigma(r,v)^{-\theta} dv
\end{aligned} \tag{70}$$

A.5 Migration Costs

To solve for the migration costs, $m_2(r)$, recall the expression for migration shares.

$$u_t(r) = m_2(r)H(r)^{\Omega}L_t(r)^{\Omega}\left(\frac{L_t}{\int_S u_t(v)^{1/\Omega}m_2(v)^{-1/\Omega}dv}\right)^{-\Omega}$$

And substitute it in equation (54).

$$\begin{aligned}
&\bar{b}_t(r)^{-\frac{\theta(1+\theta)}{1+2\theta}}\bar{a}_t(r)^{-\frac{\theta}{1+2\theta}}L_t(r)^{\flat_L+\frac{\Omega\theta(1+\theta)}{1+2\theta}}\mathcal{Q}_t(r)^{\frac{(1-\chi)\mu\theta^2}{1+2\theta}}\\
&\times H(r)^{\frac{\theta[1+\Omega(1+\theta)]}{1+2\theta}}m_2(r)^{\frac{\theta(1+\theta)}{1+2\theta}}\varphi_t(r)^{\frac{-(\mu+\gamma_1/\xi)\theta^2}{1+2\theta}}\\
&=\kappa_1\left(\frac{L_t}{\int_S u_t(s)^{1/\Omega}m_2(s)^{-1/\Omega}}\right)^{\frac{\Omega\theta}{1+2\theta}}\\
&\times\int_S \bar{b}_t(v)^{\frac{\theta^2}{1+2\theta}}\bar{a}_t(v)^{\frac{1+\theta}{1+2\theta}}L_t(v)^{\flat_R-\frac{\Omega\theta^2}{1+2\theta}}\mathcal{Q}_t(v)^{-(1-\chi)\mu\frac{\theta(1+\theta)}{1+2\theta}}\\
&\times H(v)^{\frac{\theta[1-\Omega\theta]}{1+2\theta}}m_2(v)^{-\frac{\theta^2}{1+2\theta}}\varphi_t(v)^{(\mu+\gamma_1/\xi)\frac{\theta(1+\theta)}{1+2\theta}}\varsigma(r,v)^{-\theta}dv
\end{aligned} \tag{71}$$

For simplicity, consider the normalization $\min_s m_2(s) = 1$.

B Data

Population and GDP at $1° \times 1°$. Data on population and GDP, in Power Purchasing Parities, for $1° \times 1°$ cells across the entire world come from the G-Econ 4.0 research project. For the estimation of fundamental amenities and productivities, we use data for the years 1990, 1995, 2000, 2005. We consider the same 17,048 cells that in 2000 have positive population, GDP and land. If some of these cells display missing values for 1990, 1995 or 2005, we linearly extrapolate the missing data, and, in each period, we cap GDP per capita at the percentile 97.13.

Human Development Index at $1° \times 1°$. Kummu et al. (2018) provide data on the Human Development Index at a yearly frequency, from 1990 to 2015, at a subnational level, considering around 700 geographic units. This data is presented at a resolution of 5 arc-min, so we aggregate it at a resolution of 60 minutes by considering the mode across cells.

Geographical attributes at $1° \times 1°$. Elevation data, measured as meters above the sea level at a resolution of $1° \times 1°$ is taken from the University of Washington, Join Institute for the Study of the Atmosphere and the Ocean, http://research.jisao.washington.edu/data_sets/elevation/. In order to construct the standard deviation and the mean absolute error, also known as *roughness*, within each $1° \times 1°$ cell, we use the aforementioned dataset at a resolution of $0.25° \times 0.25°$ and compute these statistics over the cells with positive land.

Distance to the coast is taken from the NASA Ocean Biology Processing Group https://oceancolor.

gsfc.nasa.gov/docs/distfromcoast/. This data is provided at a resolution of $0.1° \times 0.1°$, so we compute the average distance within each $1° \times 1°$ cell. Distance to non-frozen oceans is taken from G-Econ database and displayed at a resolution of $1° \times 1°$. Distance to nearest water body either inland or sealand (ice-covered land areas are not considered as water bodies) is taken from Carrea et al. (2015), which presents data at $1° \times 1°$.

The following data are obtained from the NASA Earth Observations, https://neo.sci.gsfc.nasa.gov at a geographical resolution of $1° \times 1°$. Vegetation density is taken as the average over the period 1951-1980.[55] Share of ice-covered land is taken over April of 2010. Albedo, which is the ratio of light that a surface reflects compared to the total sunlight that falls, is taken over April of 2010.[56] Land cover classification considers the year 2010.[57]

Temperature at $1° \times 1°$. Gridded temperature data comes from the Berkeley Earth Surface Temperature (http://berkeleyearth.org/data-new/). This dataset provides information as far as 1750, as frequent as daily maximum, minimum and average temperature and as fine as $0.25° \times 0.25°$. We employ the database that provides annual temperature at a resolution of $1° \times 1°$.

For the cells with missing temperature, we take the simple average temperature across the nine surrounding cells, that is, we create a block of cells of size 3×3 centered at the cell with missing data. If there are still cells with missing temperature (which occurs in the case of small islands), we create a block of cells of size 5×5 centered at the cell with missing data and take the simple average temperature. We continue with this procedure until the cell is filled with temperature data.

Historical CO_2 emissions and clean energy at country-level. Crippa et al. (2019) (https://edgar.jrc.ec.europa.eu/overview.php?v=booklet2020) provide CO_2 emissions for all countries considered in this analysis, except for Greenland. We supplement this observation with data from the World Bank. Additionally, Crippa et al. (2019) consider international marine and international aviation. We split those emissions across countries according to the distribution provided in IEA (2020).

Since IEA (2020) provides information for aggregate regions that comprise several countries, like Former Soviet Union or Other Africa, we partition the emissions of those aggregate regions across countries according to the pattern displayed in Crippa et al. (2019) for total emissions.

As for the use of clean energy, expressed in tons of oil equivalent (toe), we use information from BP (2019) and define clean energy as the sum of nuclear, hydroelectricity and renewables (wind, solar, among others). Since this database provides information for some aggregate regions, like Other South America and

[55]The vegetation index ranges from -0.1 to 0.9 and have no unit. Rather, they are index values in which higher values (0.4 to 0.9) show lands covered by green, leafy vegetation and lower values (0 to 0.4) show lands where there is little or no vegetation.

[56]Surfaces that reflect a large share of the light falling on them are bright, and have high albedo, like snow. Surfaces that do not reflect much light are dark, and have low albedo, like forests.

[57]This dataset partitions land based on characteristics of the surface that satellites can detect, such as water, soil, and vegetation types. There are 17 categories of land cover: 9 classes of natural vegetation, 3 classes of developed lands, 2 classes of mosaic lands, and 3 classes of non-vegetated lands (snow/ice, bare soil/rocks, water).

Other Middle East, we partition the energy use of those aggregate regions across countries according to the pattern for CO_2 emissions presented in Crippa et al. (2019). In order to make comparable CO_2 emissions and use of clean energy, we take the ratio of tons of CO_2 per ton of oil equivalent to be 2.8466.

CO_2 emissions and clean energy at $1° \times 1°$. The Emission Database for Global Atmospheric Research (EDGAR) v4.0 https://jeodpp.jrc.ec.europa.eu/ftp/jrc-opendata/EDGAR/datasets/v40/ contains global emission inventories for greenhouse gases and air pollutants. These emissions are calculated as totals by country from 1970 to 2008, and distributed at a resolution of $0.1° \times 0.1°$ using proxy data. We aggregate thi data at $1° \times 1°$ by considering the sum across cells.

We employ the CO_2 distribution from residential emissions. We prefer this specification over total emissions or emissions from production, because the former considers emissions that occur over cells with no land and the latter displays high level of CO_2 in cells scarcely populated and with low income levels (due to presence of plants producing steel or cement, for instance), we consider that such pattern does not represent the long-run trend of CO_2 emissions.

In order to define carbon dioxide emissions at the cell-level, we disaggregate the country-level emissions according to the spatial pattern displayed in the EDGAR database. As no gridded data for clean energy exists, we split the country-level clean energy use using the spatial pattern of the EDGAR database. Figure 25 displays the spatial distribution of CO_2 emissions and clean energy in 2000.

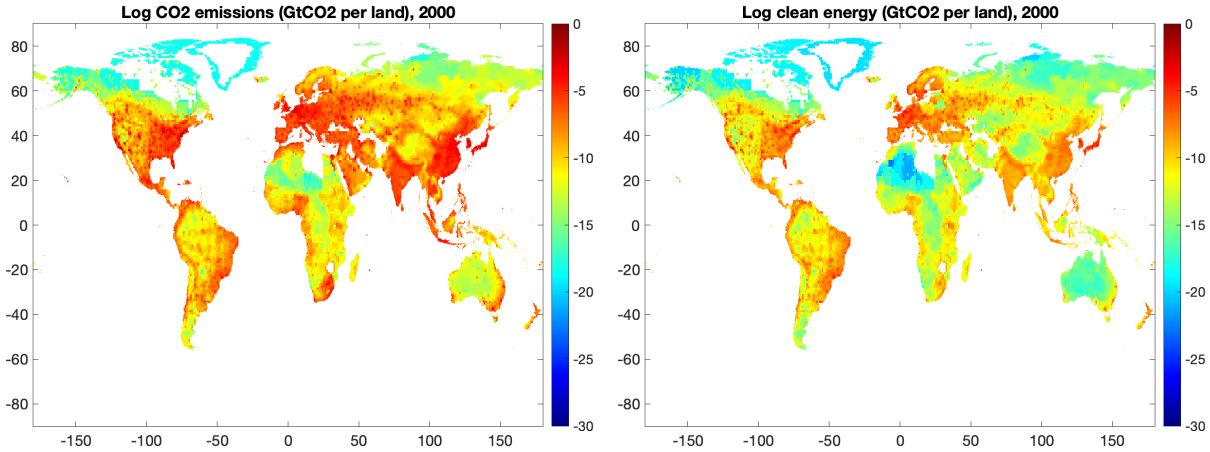

Figure 25: Spatial distribution of CO_2 emissions and clean energy in 2000.

Historical net natality at country-level. Crude birth rates and crude death rates at the country-level since 1950 at a yearly frequency are taken from the World Population Prospects of the United Nations (https://population.un.org/wpp/Download/Standard/Population/).

Projections of non-fuel combustion CO_2 emissions and non-CO_2 forcing. Forecasts up to 2500 for CO_2 flow and forcing for RCP 8.5, 6.0, 4.5 and 2.6 are taken from the RCP Database version 2.0 (http://www.iiasa.ac.at/web-apps/tnt/RcpDb). Carbon dioxide from deforestation is considered as *OtherCO2* and we consider that 1 GtC equals 44/12 $GtCO_2$. Non-CO_2 forcing is considered as *Total anthropogenic and natural*

radiative forcing minus *CO2 forcing*.

Projections of global population. Global population levels at a quinquenial frequency for the medium scenario, as well as 80% and 95% confidence intervals are taken from the World Population Prospects of the United Nations (https://population.un.org/wpp/Download/Standard/Population/). In order to make consistent total population from United Nations and G-Econ in the year 2000, from the former dataset I substract the total population of the initial period and add the total population of the year 2000 displayed in the G-Econ database.[58]

Cost of extracting fossil fuel. Bauer et al. (2017) estimate the cost of extracting fossil fuels and present estimates for different Socio Economic Share Pathways (SSP), which consider alternative assumption regarding the evolution of the world economy. We choose the scenario SSP5 (development based on fossil fuels), which is the one that closest resembles the RCP 8.5. Then, we aggregate the costs of hard coal and lignite into a single fossil fuel in terms of tCO_2 per usd, considering the conversion factors of 0.0946 and 0.1012 $GtCO_2$ per EJ, respectively. Finally, we rank costs from the least to the most expensive.

C Damage Functions on Amenities and Productivities

In this section, we detail how we employ the Human Development Index (HDI) as our measure of utility, how we estimate our main empirical specification for the damage function on amenities and productivities, and perform some robustness exercises of these functions.

C.1 HDI as a Measure of Utility

The HDI is constructed as a geometric mean of indices evaluating three dimensions of human development: a long and healthy life, being knowledgeable and have a decent standard of living. The health dimension is assessed by life expectancy at birth; the education dimension by mean of years of schooling for adults aged 25 years and more and expected years of schooling for children of school entering age; and the standard of living dimension by Gross National Income (GNI) per capita. The HDI uses the logarithm of income, to reflect the diminishing importance of income with increasing GNI.

$$HDI_t(r) = \left(I_t^{\text{Health}}(r) \cdot I_t^{\text{Educ}}(r) \cdot I_t^{\text{Income}}(r) \right)^{1/3} \tag{72}$$

$$I_t^{\text{Income}}(r) = \begin{cases} 1 & \text{if} \quad GNI_t(r) > 75,000 \\ \frac{\log(GNI_t(r)) - \log(100)}{\log(75,000) - \log(100)} & \text{if} \quad 100 \leq GNI_t(r) \leq 75,000 \\ 0 & \text{if} \quad GNI_t(r) < 100 \end{cases} \tag{73}$$

[58] This adjustment is performed since the G-Econ database does not display information for some regions of the world, like Libya and some parts of Africa and Asia.

We manipulate equations (72) and (73), in order to obtain a measure of well-being that is linear in log-real income.

$$HDI_t(r)^3 = \psi_{0,t}(r) + \psi_{1,t}(r) \log\left(GNI_t(r)\right) + \log\left(\varepsilon_t^i(r)\right). \tag{74}$$

Consequently, there is a direct relationship between equation (74) and the definition of utility in our model, given by equation (1). Ignoring migration costs, the flow utility of an individual i residing in location r in the model is given by:

$$\log\left(u_t^i(r)\right) = \log\left(b_t(r)\right) + \log\left(y_t(r)\right) + \log\left(\varepsilon_t^i(r)\right)$$

Accordingly, we estimate the following regression.

$$HDI_t(r)^3 = \psi_0 + \psi \log\left(y_t(r)\right) + \epsilon_t^i(r) \tag{75}$$

Where the HDI is obtained from Kummu et al. (2018), which present this index at a yearly frequency, from 1990 to 2015 at a subnational level, considering around 700 geographic units. Real-income, $y_t(r) = L_t(r)u_t(r)/\bar{b}_t(r)$, is computed using data on population and the ratio of utility in terms of amenities, derived from the model inversion. We estimate equation (75) aggregating cells at the subnational level and weighting observations by population size.

Table 5 shows the results of the estimation. Column (1) denotes the specification of equation (75). Column (2) allows the intercept to be time-dependent, $\psi_{0,t}$. Column (3) allows the intercept and the slope to be time-dependent, $\psi_{0,t}$ and ψ_t. As can be observed, the estimates of ψ_t are relatively stable over time. Columns (4)-(6) repeat the aforementioned exercises weighting each subnational unit by land size rather than population size as in Columns (1)-(3). We take the central estimate of $\psi = 0.045$, given by Column (2).

Finally, we compare the spatial pattern of the HDI with that of the Cantril ladder. This measure of subjective well-being is employed in Desmet et al. (2018). Figure 26 evidences that both measures of utility display a similar spatial configuration for the year 2000. However, we prefer the HDI as it provides more cross-sectional and temporal variation, which allows to better identify the effect of temperature on fundamentals.

C.2 Estimation of Damage Functions

In order to estimate the effect of temperature on amenities and productivities, we employ equation (26) as our main empirical specification. The variable $T_t(r)$ denotes the average temperature over the last decade for January in the Northern Hemisphere and July in the Southern Hemisphere, and $\mathbb{1}\{T_t(r) \in \mathcal{T}_j\}$ is an indicator function for temperature $T_t(r)$ being in interval $\mathcal{T}_j$. We partition the distribution of temperature into $J = 20$ bins, each comprising 5% of the observed temperature values.

The variable $Z(r)$ is a vector of cell-level geographic attributes. Namely, mean, standard deviation and

	(1)	(2)	(3)	(4)	(5)	(6)
logrealgdp	0.107***	0.0450***		0.0974***	0.0447***	
	(0.00657)	(0.00813)		(0.00657)	(0.00813)	
1990×logrealgdp			0.0338***			0.0360***
			(0.00658)			(0.00193)
1995×logrealgdp			0.0412***			0.0407***
			(0.00574)			(0.00188)
2000×logrealgdp			0.0459***			0.0424***
			(0.00529)			(0.00185)
2005×logrealgdp			0.0510***			0.0427***
			(0.00516)			(0.00184)
subcountry fe	X	X	X	X	X	X
year fe		X	X		X	X
weight pop	X	X	X			
weight land size				X	X	X
N	2,952	2,952	2,952	2,952	2,952	2,952
R^2	0.9822	0.9880	0.9910	0.9863	0.9927	0.9933
RMSE	0.0297	0.0245	0.0211	0.0300	0.0219	0.0211

Standard errors in parentheses, clustered by country

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Estimation of utility measure from HDI.

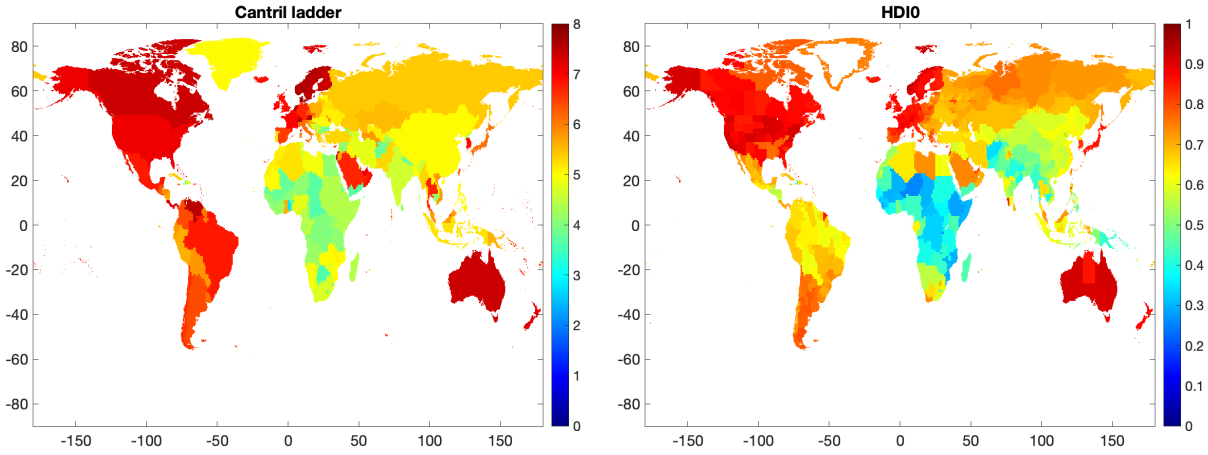

Figure 26: Cantril ladder and Human Development Index in 2000.

average deviation of elevation, distance to the coast, to a water body and to non-frozen oceans. Each of these six covariates, $Z_i(r)$, is transformed by means of a Chebyshev polynomial, $Z_i^j(r)$, $j \in \{1, \cdots, 5\}$. Accordingly, the vector $Z(r)$ comprises the 30 elements of $Z_i^j(r)$. Additionally, we consider an indicator

function for 16 different types of land.[59]

$$Z_i^j(r) = \cos\left(j \cdot \arccos\left(\frac{\tilde{Z}_i(r)}{\max_{s\in S}|\tilde{Z}_i(s)|}\right)\right)$$
$$\tilde{Z}_i(r) = \left(Z_i(r) - \frac{1}{2}\left(\min_{s\in S} Z_i(s) + \max_{s\in S} Z_i(s)\right)\right)$$

With respect to the time-invariant fixed effects, $\iota(b)$, we partition the $180° \times 360°$ gridded map into blocks of size $2° \times 2°$. As for the time-varying fixed effects at the subnational level, $\iota_t(s)$, we take as basis the subnational levels, as delimited in Kummu et al. (2018), and aggregate the subnational units in Europe at the (i) country-level, as defined in Desmet et al. (2018), and (ii) at the region level, considering North, South, West and East. Figure 27 displays the spatial demarcations of the subnational units.[60]

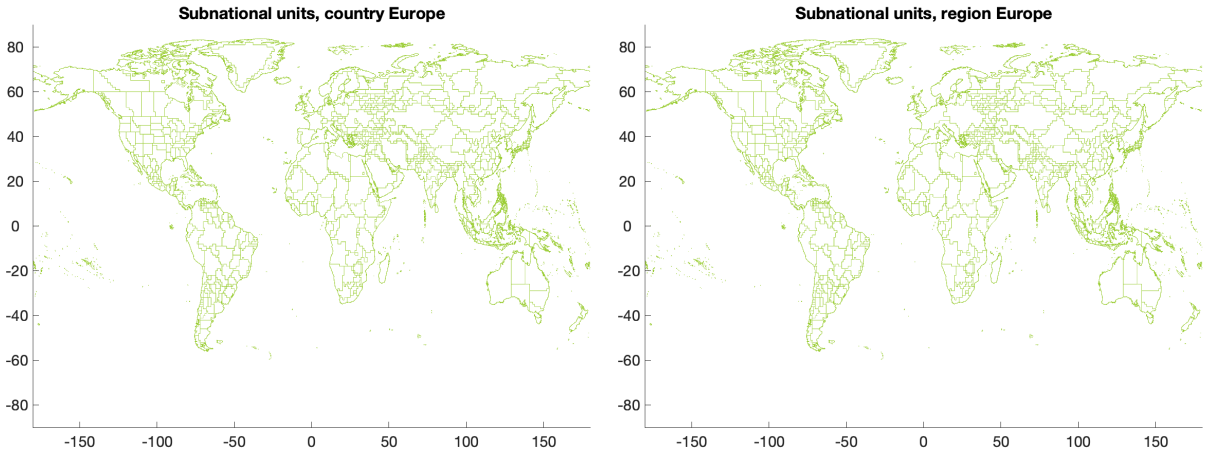

Figure 27: Time-varying fixed effects at the subnational level, $\iota_t(s)$.

We estimate equation (26), modeling the error term as in Conley (1999). That is, we consider that errors are spatially correlated, so that correlation linearly declines from one to zero as distance increases. When distance is greater or equal than 550 km (5 cells) correlation equals zero, as in Schlenker and Roberts (2009). We implement this error structure through the Stata package `acreg`, developed by Colella et al. (2019).

To smooth the behavior of the point estimates across temperature bins, $\hat{\delta}_j^x$, we fit the logistic curve $\delta^x(T)$, given by equation (77), across the point estimates of each bin, weighting them by the inverse of their standard errors, $se(\hat{\delta}_j^x)$, to provide a greater weight to the more accurate estimates. In other words, we

[59] $Z_i(r)$ denotes the variable in raw units (for instance, in meters when considering elevation) and $Z_i^j(r)$ the Chebyshev transformation for the j-th polynomial so that it lies between minus one and one.

[60] We aggregate the subnational units in Europe, because some of them have a size of one cell, precluding a proper identification.

estimate the coefficients $(\delta^x_{\min}, \delta^x_{\max}, \delta^x_{\text{center}})$ that solve (76).[61]

$$\min \quad \sum_{j=1}^{J} \frac{1}{se(\hat{\delta}_j)} \left(\hat{\delta}_j - \delta^x(T_j)\right)^2 \tag{76}$$

$$\text{st} \quad \delta^x(T) = \delta^x_{\min} + \frac{\delta^x_{\max} - \delta^x_{\min}}{1 + \mathrm{e}^{0.15*(T-\delta^x_{\text{center}})}} \tag{77}$$

Where T_j is the $(2j-1)/(2J)$-th percentile of the temperature observations, $j \in \{1, \cdots, J = 20\}$. To estimate the upper and lower α% confidence interval of the logistic function, we solve (76), but replace the point estimates of each bin j by their upper and lower α% confidence intervals, respectively.

C.3 Robustness Exercises

In this subsection, we present robustness exercises regarding the estimation of the damage functions $\Lambda^a(\cdot), \Lambda^b(\cdot)$. Figure 28 presents the damage function considering January temperature for all cells in the world, rather than January temperatures in the Northern Hemisphere and July temperatures in the Southern Hemisphere. The quantitative results are in line with those of the main specification.

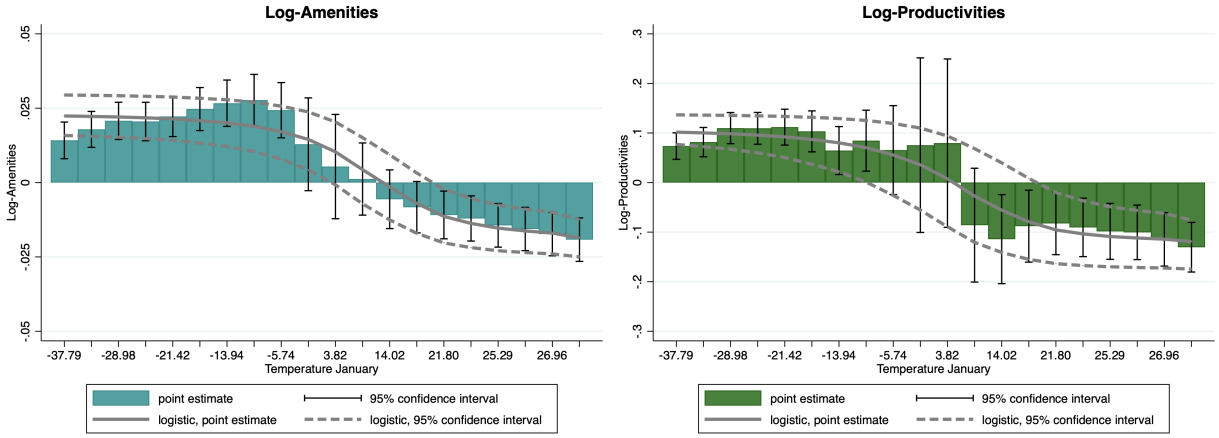

Figure 28: Effect of temperature on fundamental amenities and productivities, using January temperature for all cells.

Figure 29 plots the main specification using January-July temperatures for one year, rather than the ten-year average. Although the results are similar to the main specification, the use of only one year provides noisier results in terms of the standard errors, and the transition from benefits in cold regions to damages in warm places is less smooth.

Figure 30 implements a finer partition of the temperature observations, by considering 50 bins, each comprising 2% of the observed temperature values. Given the finer partition of the data, there are less observations in each bin and therefore the standard errors tend to be larger with respect to the baseline

[61] To achieve convergence of the non linear estimation, we exogenously set to 0.15 the slope coefficient of the logistic function.

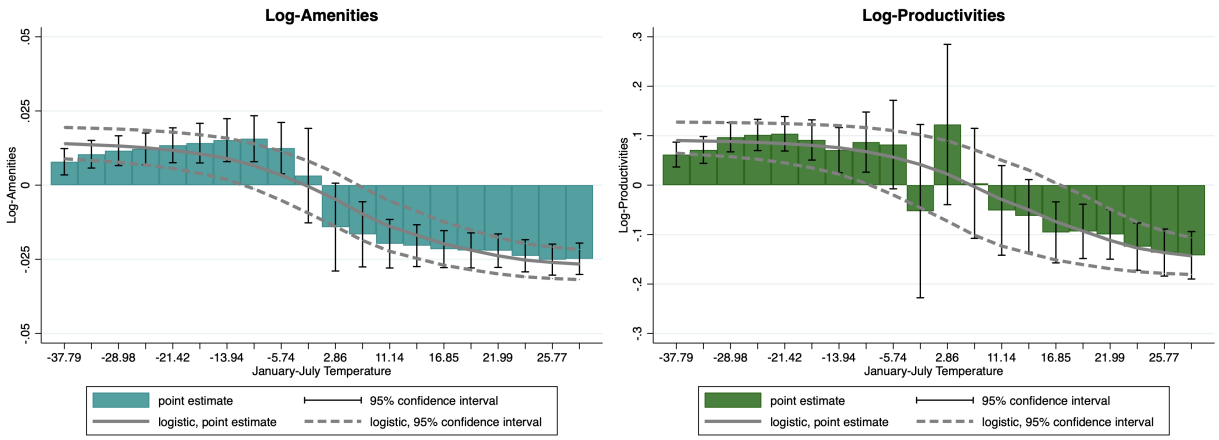

Figure 29: Effect of temperature on fundamental amenities and productivities, using one-year variation of January-July.

estimation. This behavior is more pronounced at the middle of the temperature distribution. However, the logistic smoothing is similar to that in the main specification.

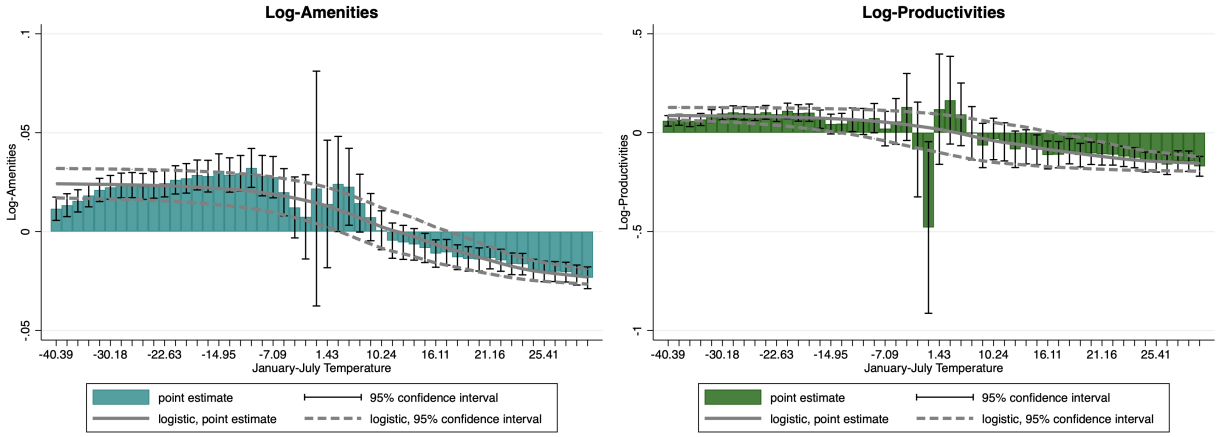

Figure 30: Effect of temperature on fundamental amenities and productivities, using 50 bins.

Tables 6 and 7 evaluate how the standard errors of the main specification of the damage functions for amenities and productivities vary under different assumptions of the errors. Column (1) considers that errors are spatially correlated with a cutoff of 550km, as in the main specification. Column (2) extends the degree of correlation to 1,100km. Column (3) considers that errors are homoskedastic and no correlation between observations exists. Column (4) replaces the homoskedastic assumption by heteroskedasticity. Column (5) clusters the errors within subnational units, as defined in Kummu et al. (2018). The last three columns are estimated with the stata package `reghdfe`, implemented by Correia (2016).

Damage Function on Amenities

Coefficient	Standard Errors				
	(1)	(2)	(3)	(4)	(5)
0.0138	(0.0031)***	(0.0039)***	(0.0013)***	(0.0014)***	(0.0076)*
0.0175	(0.0031)***	(0.0039)***	(0.0014)***	(0.0015)***	(0.0073)**
0.0203	(0.0032)***	(0.0040)***	(0.0014)***	(0.0016)***	(0.0072)***
0.0200	(0.0033)***	(0.0041)***	(0.0015)***	(0.0017)***	(0.0074)***
0.0216	(0.0034)***	(0.0042)***	(0.0016)***	(0.0018)***	(0.0078)***
0.0240	(0.0037)***	(0.0045)***	(0.0017)***	(0.0020)***	(0.0084)***
0.0258	(0.0039)***	(0.0047)***	(0.0019)***	(0.0023)***	(0.0092)***
0.0266	(0.0043)***	(0.0049)***	(0.0023)***	(0.0029)***	(0.0092)***
0.0228	(0.0046)***	(0.0051)***	(0.0032)***	(0.0038)***	(0.0088)***
0.0077	(0.0076)	(0.0083)	(0.0060)	(0.0069)	(0.0111)
0.0083	(0.0081)	(0.0081)	(0.0047)*	(0.0060)	(0.0095)
0.0047	(0.0054)	(0.0056)	(0.0028)*	(0.0033)	(0.0086)
-0.0057	(0.0037)	(0.0039)	(0.0023)**	(0.0026)**	(0.0059)
-0.011	(0.0033)***	(0.0035)***	(0.0021)***	(0.0023)***	(0.0055)**
-0.0142	(0.0029)***	(0.0031)***	(0.0019)***	(0.0021)***	(0.0048)***
-0.0168	(0.0028)***	(0.0029)***	(0.0018)***	(0.0019)***	(0.0045)***
-0.0167	(0.0026)***	(0.0027)***	(0.0017)***	(0.0018)***	(0.0042)***
-0.0192	(0.0026)***	(0.0026)***	(0.0016)***	(0.0017)***	(0.0043)***
-0.0213	(0.0025)***	(0.0025)***	(0.0016)***	(0.0016)***	(0.0042)***
-0.0227	(0.0025)***	(0.0026)***	(0.0016)***	(0.0016)***	(0.0044)***

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Standard errors of the damage function on amenities, when the error term is (1) spatially correlated up to 550 km, (2) spatially correlated up to 1100 km, (3) homoskedastic, (4) heteroskedastic and (5) clustered at the subnational level.

D Natality Rates, Energy Elasticities and Migration Costs

In this section, we outline the procedure to estimate the parameters of the natality rate function, b^ℓ, b^h, b^T, b^w, the elasticities of energy productivity growth to global real GDP growth, υ^f, υ^c and the migration cost function, $m_2(\cdot)$. To construct the natality rate function $\eta(\cdot)$, we set the parameter b_0^T to target the global natality rate observed in the year 2000, n_0^w.

$$n_0^w L_0 = \int_S \eta^y(y_0(v); b^\ell, b^h) L_0(v) H(v) dv + \int_S \eta^T(T_0(v), \log(y_0^w); b^T, b_w) L_0(v) H(v) dv$$

And we impose the functional form of equation (29).

$$\begin{aligned} n_0^w L_0 - \int_S \eta^y(y_0(v); b^\ell, b^h) L_0(v) H(v) dv &= \int_S \eta^T(T_0(v), \log(y_0^w); b^T, b_w) L_0(v) H(v) dv \\ &= 0.5 \int_S \left(b_0^T + b_2^T \mathrm{e}^{-b_1^T \left(T_0(v) - b^{*T}\right)^2} \right) L_0(v) H(v) dv \\ &= (0.5) b_0^T L_0 + (0.5) b_2^T \int_S \left(\mathrm{e}^{-b_1^T \left(T_0(v) - b^{*T}\right)^2} \right) L_0(v) H(v) dv \end{aligned}$$

Damage Function on Productivities

Coefficient	Standard Errors				
	(1)	(2)	(3)	(4)	(5)
0.0638	(0.0137)***	(0.0168)***	(0.0049)***	(0.0056)***	(0.0279)**
0.0718	(0.0151)***	(0.0183)***	(0.0054)***	(0.0063)***	(0.0313)**
0.0997	(0.0160)***	(0.0193)***	(0.0057)***	(0.0066)***	(0.0309)***
0.0991	(0.0163)***	(0.0191)***	(0.0060)***	(0.0071)***	(0.0315)***
0.1009	(0.0184)***	(0.0216)***	(0.0065)***	(0.0078)***	(0.0346)***
0.0912	(0.0209)***	(0.0236)***	(0.0076)***	(0.009)***	(0.0351)***
0.0516	(0.0243)**	(0.0273)*	(0.0090)***	(0.0105)***	(0.0409)
0.0683	(0.0308)**	(0.0349)*	(0.0115)***	(0.0134)***	(0.0436)
0.0484	(0.0448)	(0.0514)	(0.0178)***	(0.0206)**	(0.0677)
0.0578	(0.0871)	(0.0962)	(0.0418)	(0.0469)	(0.1001)
0.1216	(0.0905)	(0.0413)***	(0.0285)***	(0.0343)***	(0.1588)
-0.0482	(0.0579)	(0.0407)	(0.0146)***	(0.0183)***	(0.1400)
-0.0537	(0.0470)	(0.0378)	(0.0113)***	(0.0141)***	(0.1159)
-0.0897	(0.0392)**	(0.0353)**	(0.0098)***	(0.0120)***	(0.1057)
-0.1150	(0.0329)***	(0.0301)***	(0.0086)***	(0.0104)***	(0.0892)
-0.1104	(0.0289)***	(0.0256)***	(0.0078)***	(0.0093)***	(0.0780)
-0.1147	(0.0265)***	(0.0244)***	(0.0072)***	(0.0086)***	(0.0735)
-0.1396	(0.0253)***	(0.0244)***	(0.0068)***	(0.0082)***	(0.0708)**
-0.1544	(0.0251)***	(0.0250)***	(0.0069)***	(0.0084)***	(0.0701)**
-0.1574	(0.0253)***	(0.0269)***	(0.0070)***	(0.0086)***	(0.0696)**

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Standard errors of the damage function on productivities, when the error term is (1) spatially correlated up to 550 km, (2) spatially correlated up to 1100 km, (3) homoskedastic, (4) heteroskedastic and (5) clustered at the subnational level.

Hence, we can define b_0^T as a function of the remaining parameters $\flat = (b^\ell, b^h, b_1^T, b_2^T, b^{*T})$ and data on the initial period $x_0 = (n_0^w, L_0(\cdot), y_0(\cdot), T_0(\cdot))$. Where local population is obtained from the G-Econ database, local real income is constructed as $y_0(r) = L_0(r)^\lambda u_0(r)/\bar{b}_0(r)$ so that the ratio of utility to amenities is computed as in Appendix A.2,[62] and local temperature comes from the Berkeley Earth Surface Temperature Database (BEST).

$$b_0^T(\flat, x_0) = 2n_0^w - \frac{1}{L_0}\int_S \left(2\eta^y(y_0(v); b^\ell, b^h) + b_2^T \mathrm{e}^{-b_1^T\left(T_0(v) - b^{*T}\right)^2}\right) L_0(v)H(v)dv \tag{78}$$

Analogously, we set the parameter b_w to target the global natality rate observed in the year 2020, n_{20}^w.

[62] To solve for this ratio we require data on local wages and energy prices, which come from G-Econ, EDGAR and BP databases. Further details of the data are described in Appendix B.

The natality rates n_0^w and n_{20}^w are taken from the World Population Prospects of the United Nations.

$$\begin{aligned} & n_{20}^w L_{20} - \int_S \eta^y(y_{20}(v); b^\ell, b^h) L_{20}(v) H(v) dv \\ &= \int_S \eta^T(T_{20}(v), \log(y_{20}^w); b^T, b_w) L_{20}(v) H(v) dv \\ &= \frac{1}{1 + \mathrm{e}^{b_w [\log(y_{20}^w / y_0^w)]}} \int_S \left(b_0^T(\mathfrak{b}, x_0) + b_2^T \mathrm{e}^{-b_1^T \left(T_{20}(v) - b^{*T}\right)^2} \right) L_{20}(v) H(v) dv \end{aligned} \tag{79}$$

Hence, we can define b_w as a function of the parameters $\mathfrak{b}$ and data on the initial and twentieth period.

$$b_w(\mathfrak{b}, x_0, x_{20}) = \frac{1}{\log(y_{20}^w / y_0^w)} \log \left(\frac{\int_S \left(b_0^T(\mathfrak{b}, x_0) + b_2^T \mathrm{e}^{-b_1^T \left(T_{20}(v) - b^{*T}\right)^2} \right) L_{20}(v) H(v) dv}{n_{20}^w L_{20} - \int_S \eta^y(y_{20}(v); b^\ell, b^h) L_{20}(v) H(v) dv} - 1 \right) \tag{80}$$

Now, we describe the algorithm employed to jointly solve for the parameters and functions of interest.

1. Guess $m_2(\cdot), \mathfrak{b}, \upsilon^f, \upsilon^c$ and x_{20}.

2. Run the model backwards for 50 periods using equation (70) of Appendix A.4, taking the behavior of local historical temperature as in the BEST dataset,[63] and retrieve $L_t(\cdot)$ and $y_t(\cdot)$.

3. Compute the natality rates at the country-level induced from the model.[64] If the difference between the model induced and the historical natality rates is small enough, go to the next step. Otherwise, use the solution of (81) to update $\mathfrak{b}$ and go back to step 2.

$$\min_{\mathfrak{b}} \quad SSR(\mathfrak{b}) = \sum_{c=1}^{168} \sum_{t=1950}^{1999} L_t^c \left(u_t^c\right)^2 \tag{81}$$

$$\text{st} \qquad u_t^c = \left(n_t^{c,data} - n_t^{c,model}(\mathfrak{b}) \right)^2 \tag{82}$$

$$n_t^{c,model}(\mathfrak{b}) = \frac{\int_{v \in c} \eta(y_t(v), T_t(v); \mathfrak{b}) L_t(v) H(v) dv}{\int_{v \in c} L_t(v) H(v) dv}$$

Where c denotes countries, t periods of time, L_t^c is the weight given to each observation, based on historical country-level population data, and u_t^c is the error between the model and the data for each observation. Let B be the number of elements in the vector $\mathfrak{b}$.

We solve (81) as follows: If $SSR(\cdot)$, evaluated at the guess, is small enough, the procedure ends. Otherwise, we update $\mathfrak{b}$ in the j-th iteration as shown below.

$$\mathfrak{b}_{j+1} = \mathfrak{b}_j + \varrho (X_j' D X_j)^{-1} (X_j' D u_j) \tag{83}$$

[63] The backward solution of the climatic model does not provide very accurate results with respect to the historical observations. Therefore, we prefer to employ observed past data.

[64] We consider 168 countries, following the classification of Desmet et al. (2018).

Where X_j is a matrix of size $(168 \cdot 50) \times B$ comprising the derivatives of $n_t^{c,model}(\cdot)$ evaluated at $\flat_j$, D is a matrix of size $(168 \cdot 50) \times (168 \cdot 50)$ comprising the population weights, and u_j is a matrix of size $B \times (168 \cdot 50)$ comprising the errors evaluated at $\flat_j$. Additionally, ϱ is *step size* scalar parameter chosen at each iteration to improve convergence.

At each iteration, ϱ is set to one, and a candidate $\flat_{j+1}^*$ is compute from (83). If $SSR(\flat_{j+1}^*) < SSR(\flat_j)$, then $\flat_{j+1} = \flat_{j+1}^*$ and the iteration is completed. Otherwise, ϱ is halved, a new $\flat_{j+1}^*$ is computed, and the process is repeated. The NLLS estimation concludes when $SSR(\flat_{j+1})$ is small enough.

4. Compute the migration costs that target population distribution in the year 2005 using equation (71) of Appendix A.5. If the difference between the guess and the targeted migration costs is small enough, run the model 20 periods in the future, update x_{20} and go to the next step. Otherwise, use the targeted migration costs to update $m_2(\cdot)$ and go back to step 2. Figure 31 displays the logarithm of migration costs of the last iteration of the algorithm.

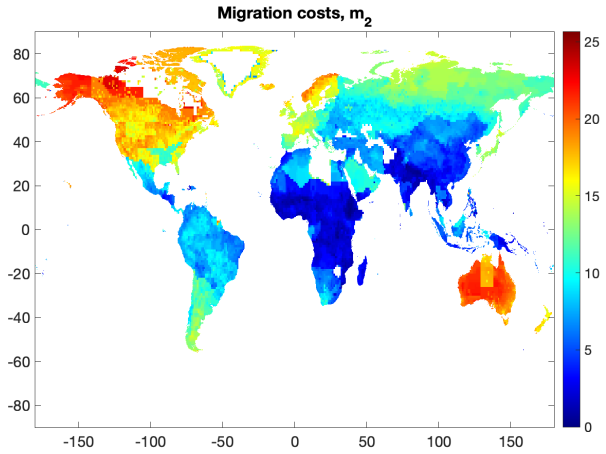

Figure 31: Log-Migration costs.

5. If the difference between the model induced, $E_t^{f,model}$, and historical data on global carbon dioxide emissions, $E_t^{f,data}$, from 1950 to 1999 is small enough, run the model 20 periods in the future, update x_{20} and go to the next step. Otherwise, update υ^f based on (84) and go back to step 2.

$$\upsilon^{f'} = \frac{E_{1999}^{f,data} - E_{1950}^{f,data}}{E_{1999}^{f,model} - E_{1950}^{f,model}} \tag{84}$$

6. If the difference between the model induced, $E_t^{c,model}$, and historical data on global clean energy use, $E_t^{c,data}$, from 1965 to 1999 is small enough, run the model 20 periods in the future, update x_{20} and the algorithm concludes.[65] Otherwise, update υ^c based on (85) and go back to step 2. Figure 32 compares

[65]The BP database does not contain information on clean energy use for the years 1950-1964.

the historical and model estimated flow of global CO_2 emissions and clean energy use.[66]

$$\upsilon^{c'} = \frac{E^{c,data}_{1999} - E^{c,data}_{1965}}{E^{c,model}_{1999} - E^{c,model}_{1965}} \tag{85}$$

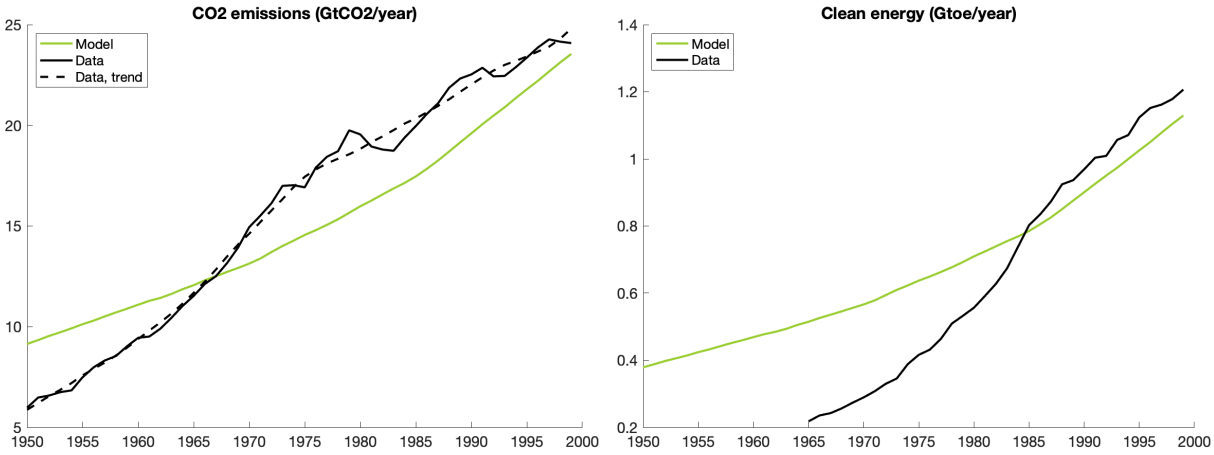

Figure 32: Historical global CO_2 emissions and clean energy use.

Figure 33 compares the evolution of global population and natality rates in the data and the model. Figure 34 plots the cross-section of the country-level natality rates in the data and the cell-level natality rates estimated by the algorithm for the year 2000. Figure 35 displays in solid black curves the natality rate functions on log real GDP and temperature, and in green bubbles the observed relationships between country-level natality rates, temperature and model induced log real GDP, where the size of the bubbles depends on population levels.[67]

These figures reveal that the model predicts lower dispersion in natality rates with respect to the data. Steeper natality rate functions might exacerbate the long-run effects of shocks to the economy. For instance, a rise in iceberg trade costs that reduces income at impact would generate a huge increase in the natality rates of the subsequent periods, generating greater levels of population in the long-run and therefore higher growth rates in the BGP. In other words, the direct effect of higher commercial frictions would be overshadowed by the evolution of global population.

E Carbon Cycle and Temperature Down-scaling

In this section, we describe the parametrization of the carbon cycle and the temperature down-scaling. Regarding the evolution of the stock of carbon dioxide, displayed in equation (18), the share of CO_2 remaining

[66] The dashed black curve in Figure 32 represents the 11-year moving average of global CO_2 emissions.

[67] We consider the model-induced log real GDP, because it is hard to construct a measure of real income from the data that captures the same elements of this model for every year of the period 1950-1999.

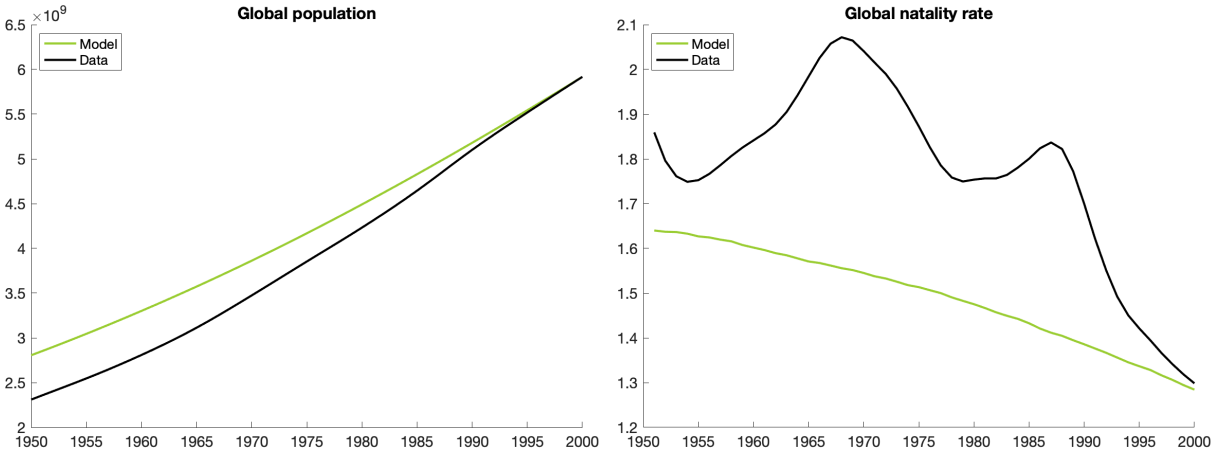

Figure 33: Historical global population and natality rates: model and data.

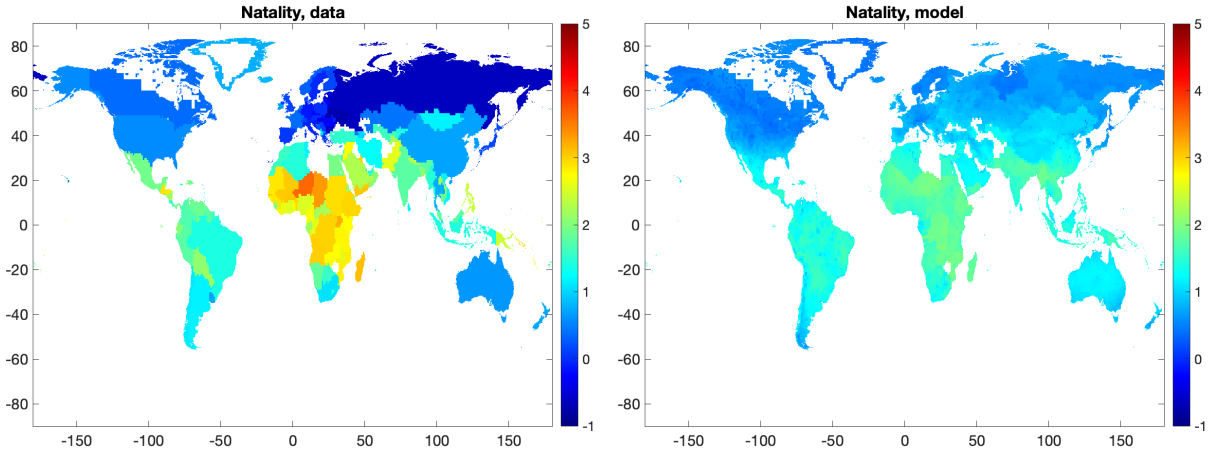

Figure 34: Natality rates in 2000: country-level data and cell-level model fit.

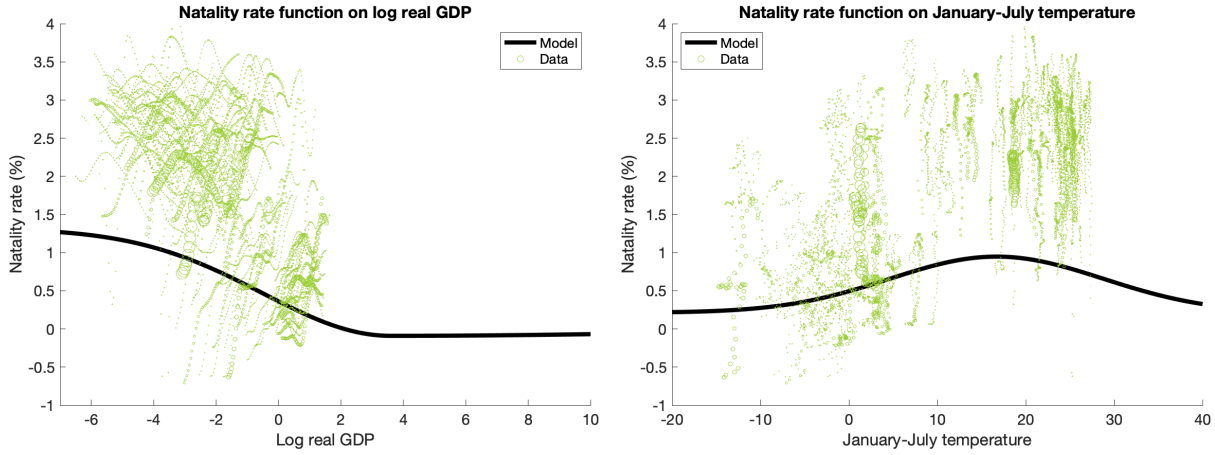

Figure 35: Natality rate function on real GDP and temperature: model and data.

in the atmosphere ℓ periods ahead, $(1-\delta_\ell)$, is approximated by a sum of exponentials, as in Forster et al. (2007) and Joos et al. (2013).

$$(1-\delta_\ell) = a_0 + \sum_{i=1}^{3}(a_i \cdot \mathrm{e}^{-\ell/b_i})$$

According to IPCC (2013), we set $a_0 = 0.2173, a_1 = 0.2240, a_2 = 0.2824, a_3 = 0.2763, b_1 = 394.4, b_2 = 36.54, b_3 = 4.304$. To simplify the evolution of carbon stock, we rewrite the law of motion as a recursive vector representation.

$$S_{t+1} = S_{0,t+1} + \sum_{i=1}^{3} S_{i,t+1} \tag{86}$$

$$S_{0,t+1} = S_{0,t} + a_0(E_t^f + E_t^x) \tag{87}$$

$$S_{i,t+1} = (\mathrm{e}^{-1/b_i})S_{i,t} + a_i(E_t^f + E_t^x), \quad i \in \{1,2,3\} \tag{88}$$

Thus, the model requires the initial values of the four layers, $S_{0,i}, i \in \{0,1,2,3\}$. Following Golosov et al. (2014), we consider $S_{0,2000} = 2,429$ GtCO$_2$ as the sum of the pre-industrial stock, $S_{\text{pre-ind}} = 2,200$ GtCO$_2$, plus a share a_0 of the historical cumulative carbon emissions, $\sum_{\ell=1945}^{1999}(E_\ell^f + E_\ell^x) = 1,054$ GtCO$_2$. The remaining three layers, $S_{1,2000} = 224, S_{2,2000} = 178, S_{3,2000} = 37$ GtCO$_2$, are computed as the discounted sum of past emissions.[68]

With respect to the effect of forcing on global temperature, the response to a unit forcing can be represented by a sum of two exponentials, as in Boucher and Reddy (2008).

$$\zeta_\ell = \sum_{j=1}^{2} \frac{c_j}{d_j} \cdot \mathrm{e}^{-\ell/d_j}$$

We take the forcing sensitivity to be $\varphi = 5.35$ and the climate parameters to be $c_1 = 0.631, c_2 = 0.429, d_1 = 8.4, d_2 = 409.5$.[69] Analogously to the carbon circulation, we rewrite the temperature module as a recursive vector representation.

$$T_{t+1} = T_{1,t+1} + T_{2,t+1} \tag{89}$$

$$T_{j,t+1} = (\mathrm{e}^{-1/d_j})T_{j,t} + \frac{c_j}{d_j}F_{t+1}, \quad j \in \{1,2\} \tag{90}$$

Where $T_{1,2000} = 1.01$°C is the discounted sum of past forcings (from 1825 to 2000) and $T_{2,2000} = 8.19$°C

[68] Historical data for CO$_2$ stock and projections for CO$_2$ emissions and forcing for every RCP are taken from http://www.iiasa.ac.at/web-apps/tnt/RcpDb.

[69] Etminan et al. (2016) provide new calculations of the radiative force with respect to Myhre et al. (1998). They propose a GHG concentration-dependence for the forcing sensitivity, so that it can be expressed as $\varphi = 5.35 + \varphi_1|S_t - S_{\text{pre-ind}}| + \varphi_2(S_t - S_{\text{pre-ind}})^2 + \varphi_N(N_t + N_{\text{pre-ind}})/2$, where N_t denotes the stock of nitrous oxide.

is the discounted sum of past forcings plus the pre-industrial temperature $T_{\text{pre-ind}} = 8.1°\text{C}$. We interpret temperature as that over land, excluding that over water.

To construct the mapping from global to local temperature, we estimate equation (21), where the object of interest is the temperature scaler function, $g(\cdot)$. We parametrize this function as an additive separable Chebyshev polynomial of order 10 in the following arguments: latitude, longitude, product of latitude and longitude, mean elevation, distance to the coast, distance to the ocean, distance to a water body, vegetation density, albedo and share of land covered by ice.

Therefore, we can define the function $g(\cdot)$ as shown in equation (91), where $Z_i(\cdot)$ denotes each of the ten covarites mentioned in the previous paragraph, $Z_i^j(\cdot)$ is the Chebyshev polynomial of order $j \in \{1, \cdots, 10\}$ of covariate i and β_i^j is the set of coefficients to be estimated by OLS.[70]

$$
\begin{aligned}
g(r) &= \sum_{i=1}^{10} \sum_{j=1}^{10} \beta_i^j Z_i^j(r) \\
Z_i^j(r) &= \cos\left(j \cdot \arccos\left(\frac{\tilde{Z}_i(r)}{\max_{s \in S} |\tilde{Z}_i(s)|} \right) \right) \\
\tilde{Z}_i(r) &= \left(Z_i(r) - \frac{1}{2} \left(\min_{s \in S} Z_i(s) + \max_{s \in S} Z_i(s) \right) \right)
\end{aligned}
\tag{91}
$$

To estimate (21) and (91), we construct the temperature variables as follows: $T_{\text{base}}(r)$ is the average temperature from 1950 to 1979 in cell r, $T_t(r)$ is the yearly temperature from 1980 to 2017 in cell r, and T_{base} is the global average of $T_{\text{base}}(r)$, where each cell is weighted by land size. Finally, we provide more weight to the more recent observations, according to $\omega_t = (2018 - t)^{-1}$. The estimation procedure is able to capture 83% of the variance of the data.

F Sensitivity Results of the Baseline Scenario

In this section, we provide additional results for the baseline scenario and test the sensitivity of the numerical results to different discount factors, elasticities of energy substitution and size of fuel deposits.

F.1 Additional Results of the Baseline Scenario

Figure 36 shows the evolution of local temperature from the year 2000 to 2200 and Figure 37 the histogram of amenity and productivity losses attributed to global warming in the year 2200.

[70] When constructing the product of latitude and longitude, we first normalize latitude and longitude so that they lie between minus one and one, and then we multiply them. That is, we define this product as $Z_{\text{lat}}^j(\cdot) \times Z_{\text{lon}}^j(\cdot)$.

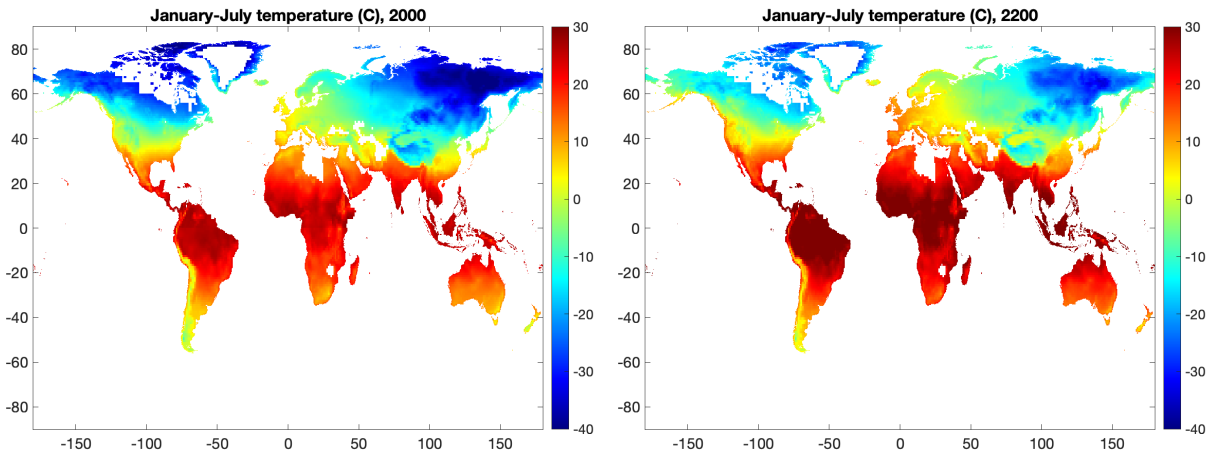

Figure 36: Local January-July temperature in 2000 and 2200.

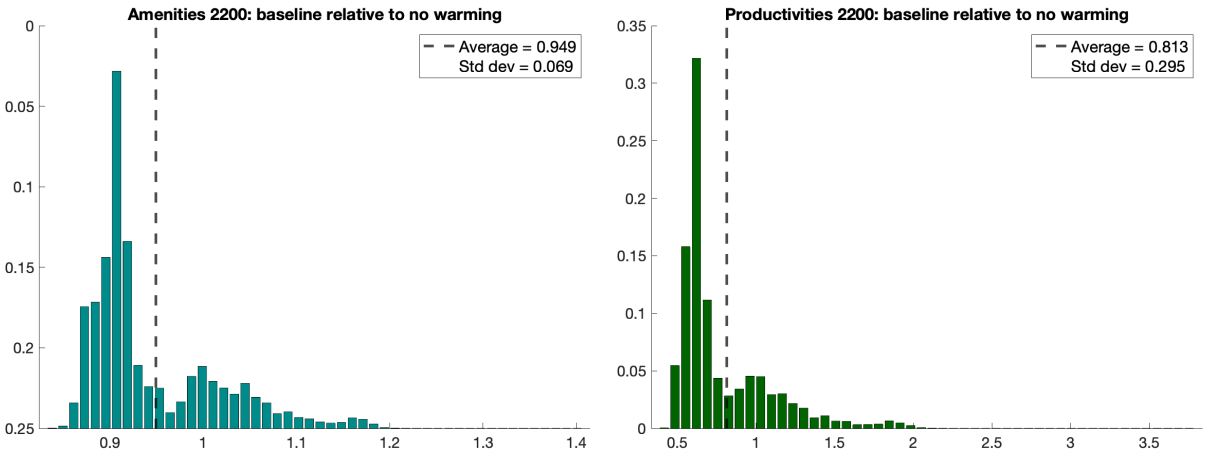

Figure 37: Histogram of losses in amenities and productivities from global warming in the year 2200.

F.2 Additional Results of the Environmental Policies

Figures 38 and 39 present the spatial distribution of real GDP gains from a 200% carbon tax and a 75% clean energy subsidy relative to the baseline scenario with no environmental policies, respectively. The spatial pattern is in line with that of welfare. However, North of Africa and Middle East tend to display larger distortions in real GDP, since those regions are more intensive in fossil fuels.

F.3 Baseline Results with a Discount Factor of $\beta = 0.969$

Figures 40 and 41 explore the spatial dimension of welfare and real GDP losses when considering a higher discount factor of $\beta = 0.969$, rather than $\beta = 0.965$ as in Section 4. The spatial patterns and the shape of the histograms resemble those of Figures 8 and 9. However, when considering this higher discount factor, losses in global average welfare and real GDP are roughly 1% larger.

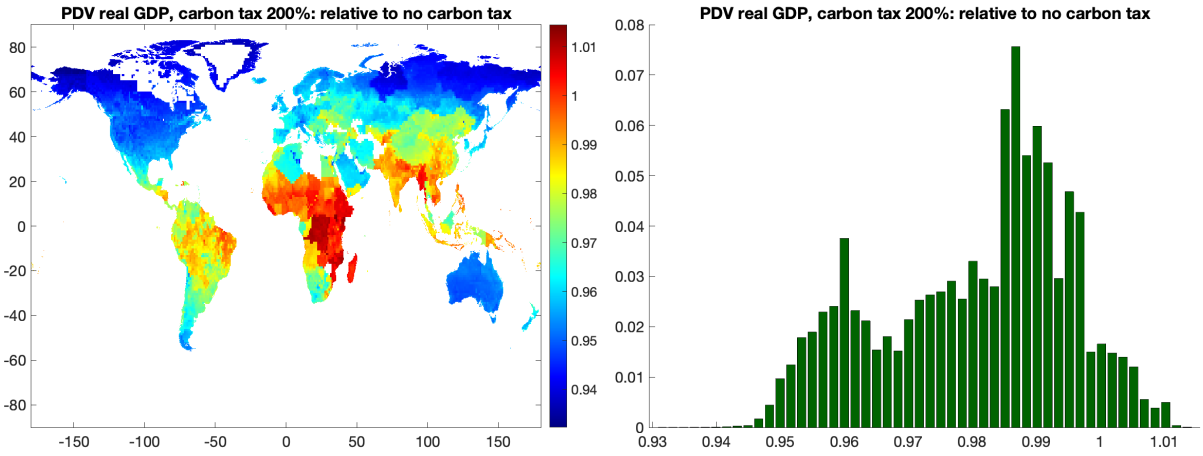

Figure 38: Local real GDP effects of a carbon tax of 200% with a discount factor of $\beta = 0.965$.

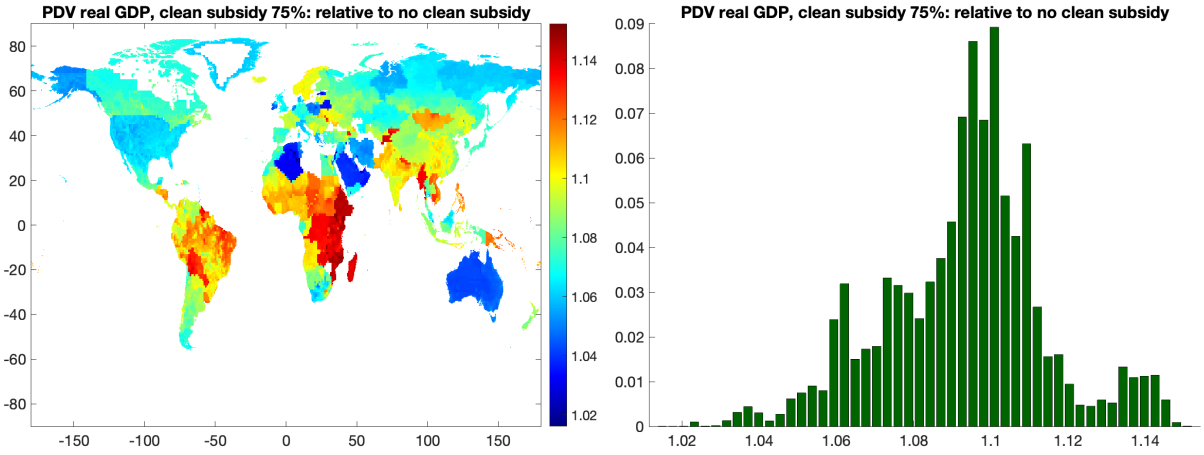

Figure 39: Local real GDP effects of a clean energy subsidy of 75% with a discount factor of $\beta = 0.965$.

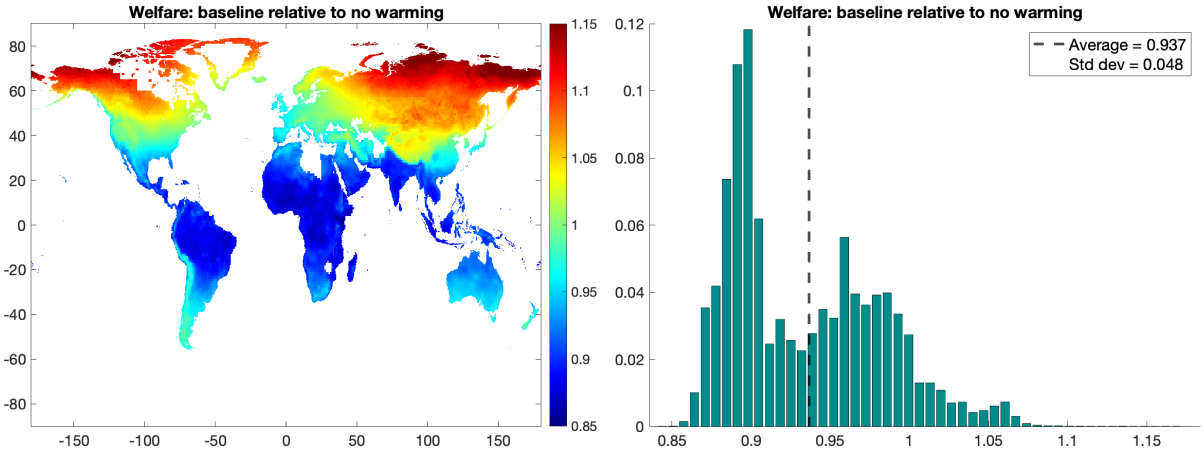

Figure 40: Welfare losses due to global warming with a discount factor of $\beta = 0.969$.

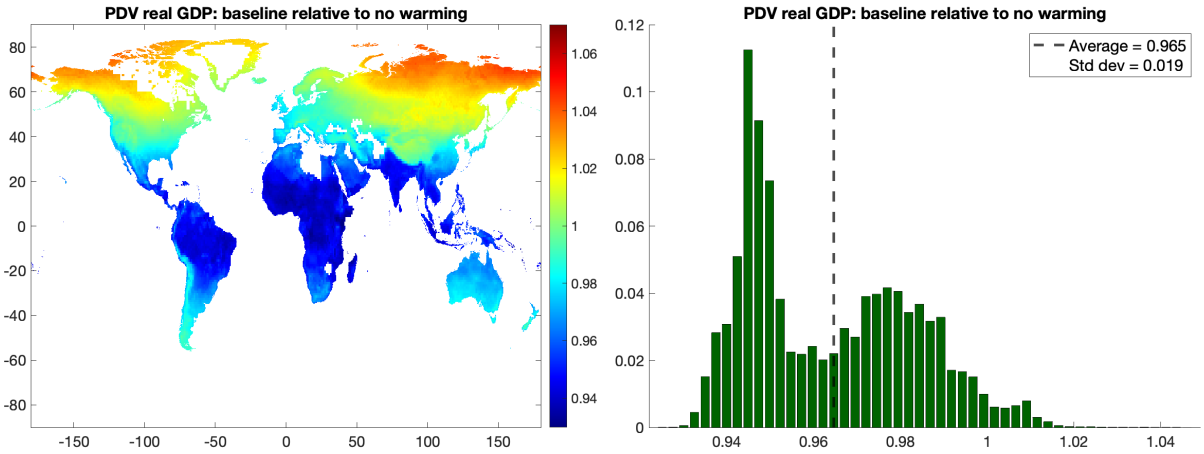

Figure 41: Real GDP losses due to global warming with a discount factor of $\beta = 0.969$.

Figure 42 displays the spatial distribution of real GDP and welfare for the baseline case, the worst- and best-scenario. The overall shape of the distributions is similar to that of Figure 11, but the greater discount factor rises the level of damages.

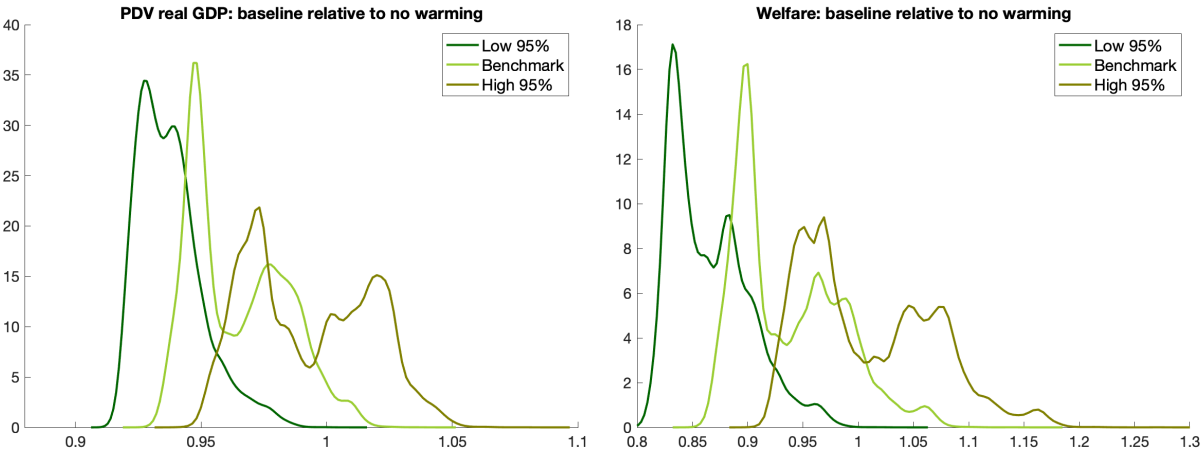

Figure 42: Real GDP and welfare distribution losses by uncertainty level with a discount factor of $\beta = 0.969$.

F.4 Sensitivity to Elasticity of Substitution between Energy Sources

In this subsection, we explore the robustness of the quantitative results when considering different degrees of substitution between fossil fuel and clean energy, $\varepsilon \in \{0.5, 1.6, 3, 6\}$. Figure 43 displays the evolution of global carbon dioxide emissions and temperature, across different energy substitution levels, where $\varepsilon = 1.6$ denotes the benchmark calibration.

Since the elasticity of energy productivity growth with respect to global real GDP growth is higher for clean than for dirty energy ($\upsilon^c = 1.05 > 0.95 = \upsilon^f$), then increases in worldwide income imply greater

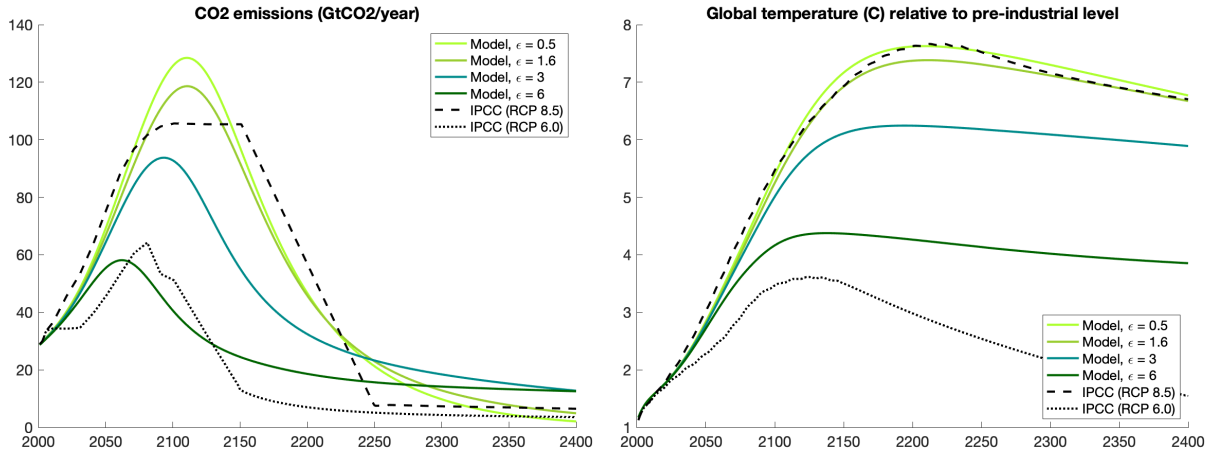

Figure 43: CO_2 emissions and global temperature across different elasticities of energy substitution.

reductions in the price of clean energy compared to that of fossil fuels.[71] Thus, with higher energy substitution, the economy consumes more of the relatively cheap source of energy, and less of the expensive one, leading to a flatter evolution of the CO_2 flow.

Figure 44 presents the transition of losses in worldwide real GDP and welfare. As global temperature rises at a slower speed with higher energy substitution, the adverse consequences of global warming attenuate.[72] In this sense, policies aiming at rising the substitutability between clean and dirty energy sources, like research to reduce the cost of clean energy storage, are projected to have a remarkable impact. Furthermore, the attenuation of warming damages is expected to be larger the higher (lower) the elasticity of clean (dirty) energy productivity growth to real income growth, ϑ^c (ϑ^f), or the steeper extraction cost functions, $f(\cdot)$, as they increase the relative price of dirty to clean energy.

When exploring the cross-section of warming damages, the coldest regions are relatively better-off under lower energy substitution, as in those cases, temperature reaches greater values, inducing amenity- and productivity-gains in such places. The converse occurs for tropical regions. Henceforth, greater energy substitution, diminishes both the gains in cold places and the losses in warm places, reducing the dispersion of warming damages on welfare. Figure 45 shows the ratio of welfare losses in the benchmark scenario ($\varepsilon = 1.6$) with respect to that in high energy substitution case ($\varepsilon = 6$) and the comparison of the spatial distribution of welfare losses under different values for ε.

Now, we evaluate the sensitivity of environmental policies to different degrees of energy substitution. First, we assess the effect of an spatial- and time-invariant carbon tax of 200%, keeping clean energy subsidies at 0%. Figure 46 displays in solid curves the evolution of global carbon dioxide emissions and tem-

[71] If the relationship between the elasticities of clean and dirty energy productivity growth to real income growth were such that $\upsilon^c < \upsilon^f$, then with more energy substitution, the evolution of the carbon dioxide flow would be steeper.

[72] To the extent that all the carbon available in the ground is depleted, the long-run global temperature is insensitive to the elasticity of substitution across energy sources.

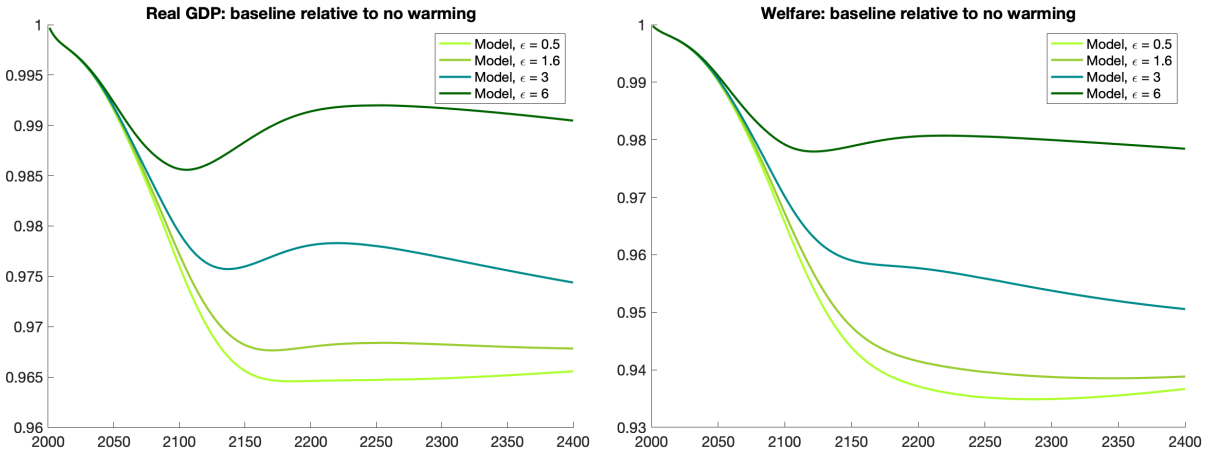

Figure 44: Real GDP and welfare losses over time across different elasticities of energy substitution.

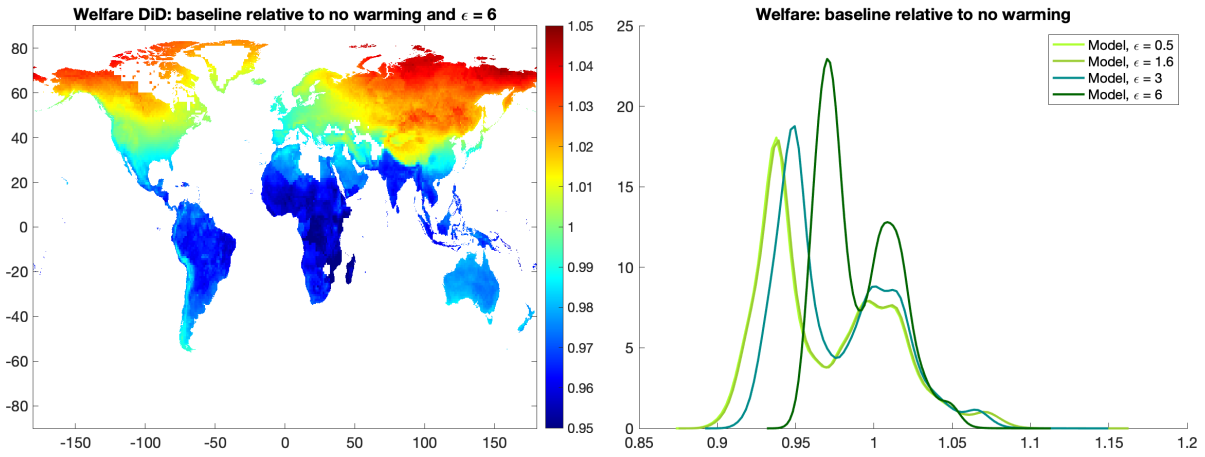

Figure 45: Welfare losses over space across different elasticities of energy substitution.

perature without abatement and in dashed curves those with the introduction of a free geoengineering technology in 2100.

Conditional on the same proportional carbon tax, its implementation leads to a larger decline in the use of fossil fuels at impact with higher energy substitution, as firms can more easily substitute energy consumption towards clean sources. Over time, the evolution of carbon dioxide emissions is more protracted with greater values of ε.

Since the cumulative CO_2 flow released to the atmosphere monotonically declines as the energy substitution increases, the introduction of an abatement yields stronger differences in long-run temperature with respect to the absence of such technology.

When examining the global real GDP and welfare losses, Figure 47, we observe that the distortion in output in the first periods due to the implementation of a carbon tax is smaller when energy types are more substitutable. In the transition, even though, temperature is more sensitive to environmental policies with

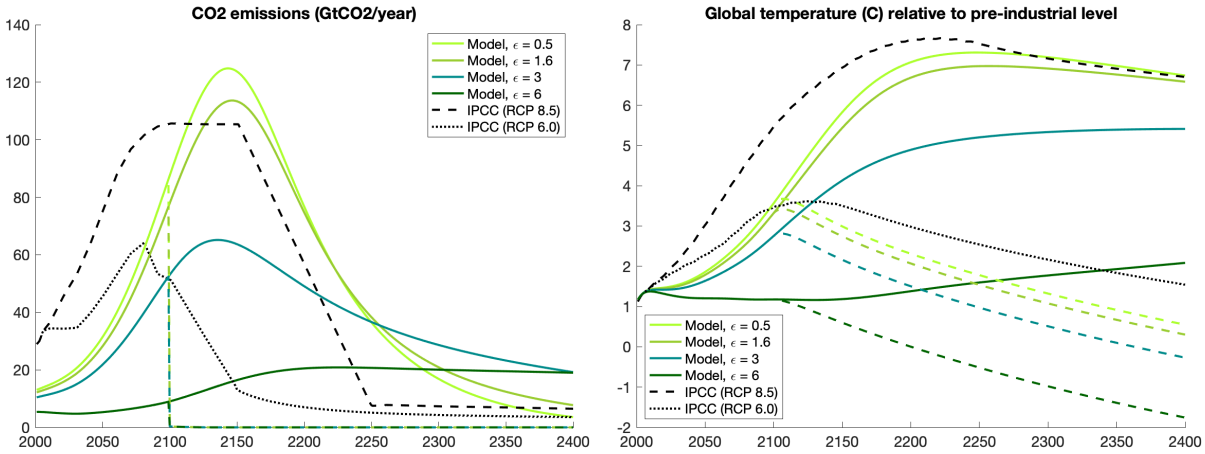

Figure 46: CO_2 emissions and global temperature with a carbon tax of 200% across different elasticities of energy substitution.

greater energy substitution, the evolution of economic benefits arising from the carbon tax is more modest. This effect is explained by the behavior of total population: the introduction of the CO_2 levy reduces income in the first periods, being this decline greater with lower energy substitution. Accordingly, the smaller ε, the larger the increase in natality rates and global population. Summarizing, the larger rise in productivity from higher population levels outweighs the temperature differences when ϵ is low. So, in those cases, the economic benefits of a carbon tax augment for the next centuries.[73]

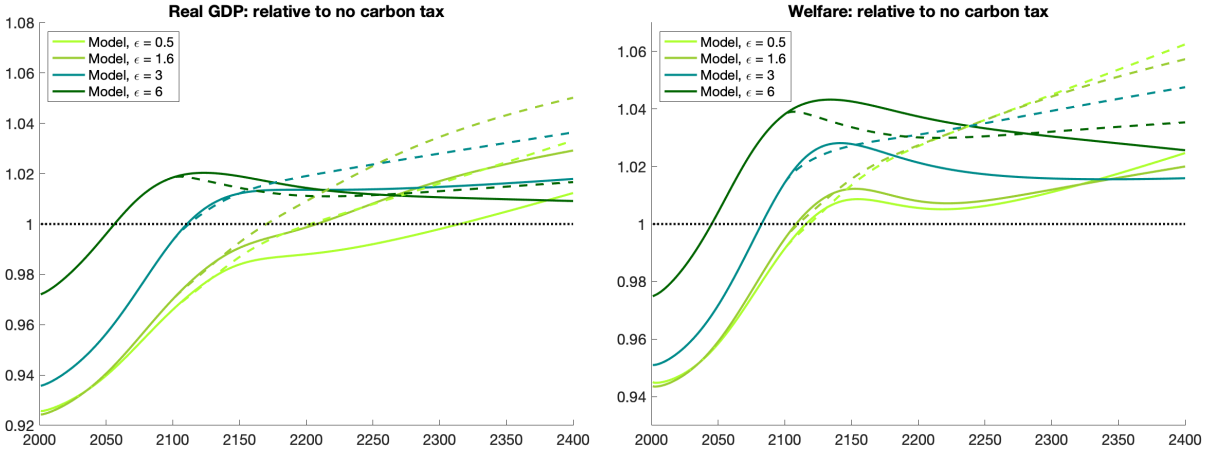

Figure 47: Real GDP and welfare with a carbon tax of 200% across different elasticities of energy substitution.

[73]Note that relative real GDP and welfare can be slightly lower in the abatement case for a couple of decades after the invention of the abatement technology. The reason is that the difference in temperatures between the benchmark scenario with and without abatement can be larger than the difference in temperature with and without abatement in the scenario with a carbon tax, depending on the second derivative of the temperature function at the time the abatement technology arrives. After a few decades, this effect is always dominated by the faster increases in temperature in the case without abatement.

Tables 8 and 9 summarize the global real GDP and welfare gains for a carbon tax of 200% under different discount factors and elasticities of substitution between fossil fuel and clean energy. The former compares the scenario with carbon taxes and no abatement with respect to the benchmark case, whereas the latter compares the scenario with carbon taxes and the introduction of a costless geoengineering technology in 2100 with respect to the benchmark case.

For low discount factors, the initial production distortions have a greater weight, so economic benefits rise with energy substitution. For high discount factors, real GDP and welfare display a concave shape in the energy elasticity of substitution.

	PDV of real GDP			Welfare		
	BGP gr	β=0.965	β=0.969	BGP gr	β=0.965	β=0.969
ε=0.5	2.888%	0.970	0.999	2.868%	0.990	1.017
ε=1.6	3.053%	0.981	1.042	3.032%	0.993	1.033
ε=3.0	3.048%	0.993	1.021	3.023%	1.002	1.016
ε=6.0	3.045%	1.005	1.005	3.017%	1.017	1.005

Table 8: PDV of real GDP and welfare gains with a carbon tax of 200%, with no abatement, under different elasticities of energy substitutions and discount factors.

	PDV of real GDP			Welfare		
	BGP gr	β=0.965	β=0.969	BGP gr	β=0.965	β=0.969
ε=0.5	2.879%	0.976	1.013	2.883%	1.001	1.042
ε=1.6	3.065%	0.989	1.074	3.051%	1.006	1.082
ε=3.0	3.064%	0.998	1.055	3.049%	1.012	1.068
ε=6.0	3.060%	1.006	1.028	3.046%	1.018	1.043

Table 9: PDV of real GDP and welfare gains with a carbon tax of 200%, with the introduction of an abatement technology in 2100, under different elasticities of energy substitutions and discount factors.

As for the cross-section of welfare gains under a carbon tax of 200% with no abatement, higher levels of energy substitution shift to the right and increase the dispersion of the welfare distribution. In other words, tropical regions gain more from carbon taxes with higher energy substitution, due to the larger decline in temperature and the smaller production losses in the first periods, as the rise in the price of the energy composite is less pronounced.[74]

Now, we examine the effect of a spatial- and time-invariant clean energy subsidy of 75%, keeping carbon taxes at 0%. As expected, higher energy substitution leads to greater declines in CO_2 emissions and a more prolonged path for global temperature, as illustrated in Figure 49.

At impact, the subsidy on clean energy leads to a reduction in the composite price of energy, so that

[74] Tropical regions tend to be more carbon intensive and a carbon tax generates larger production distortions.

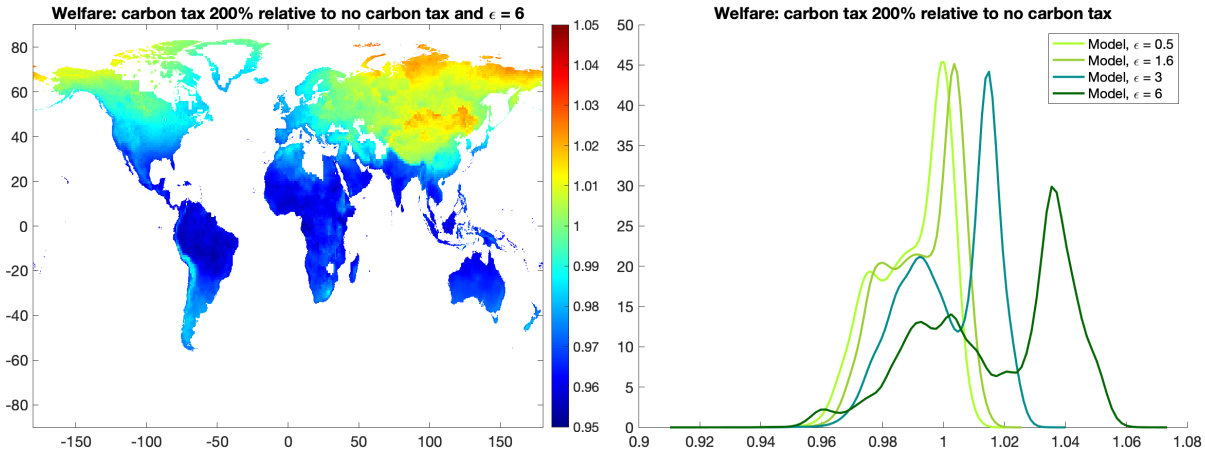

Figure 48: Welfare gains over space with a carbon tax of 200% across different elasticities of energy substitution.

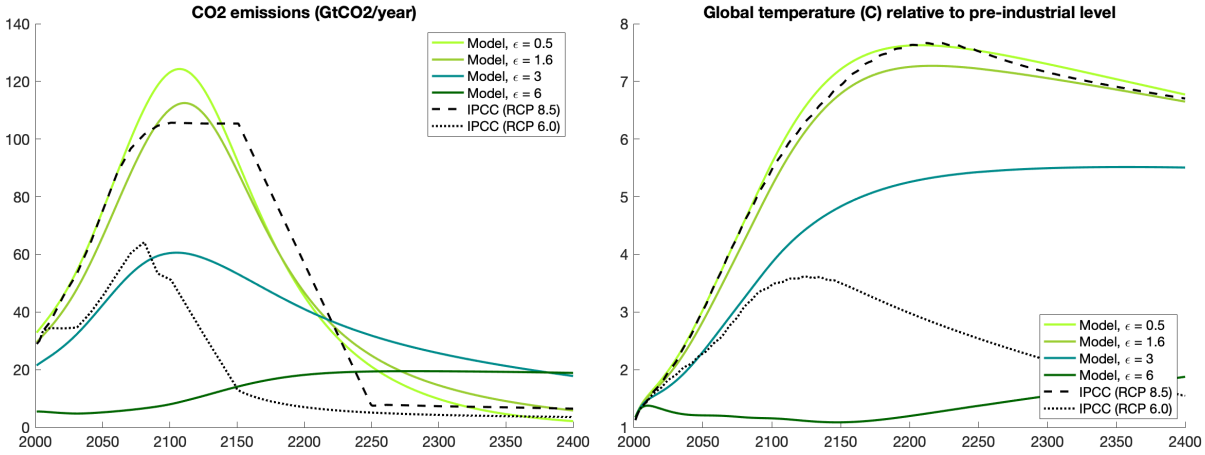

Figure 49: CO_2 emissions and global temperature with a clean energy subsidy of 75% across different elasticities of energy substitution.

the countries initially using more clean energy relatively to fossil fuels undergo greater declines in the composite price of energy. Those declines are larger, the higher the energy elasticity.

On average, developed countries tend to be more intensive in clean energy, attracting more households to those places. So, the relocation of persons towards the most productive places is stronger when energy sources are more substitutable, augmenting global real GDP and welfare, as shown in Figure 50.

The boost in innovation, due to the population allotment towards productive places, and the more protracted path for global temperature with high ε, rise the economic benefits of the environmental policy for the subsequent decades until they reach a peak. Afterwards, global real GDP and welfare benefits start to diminish, as the income increase reduces natality rates and so does global population. Table 10 summarizes the global real GDP and welfare gains of a clean energy subsidy of 75% under different degrees

of energy substitution and discount factors.

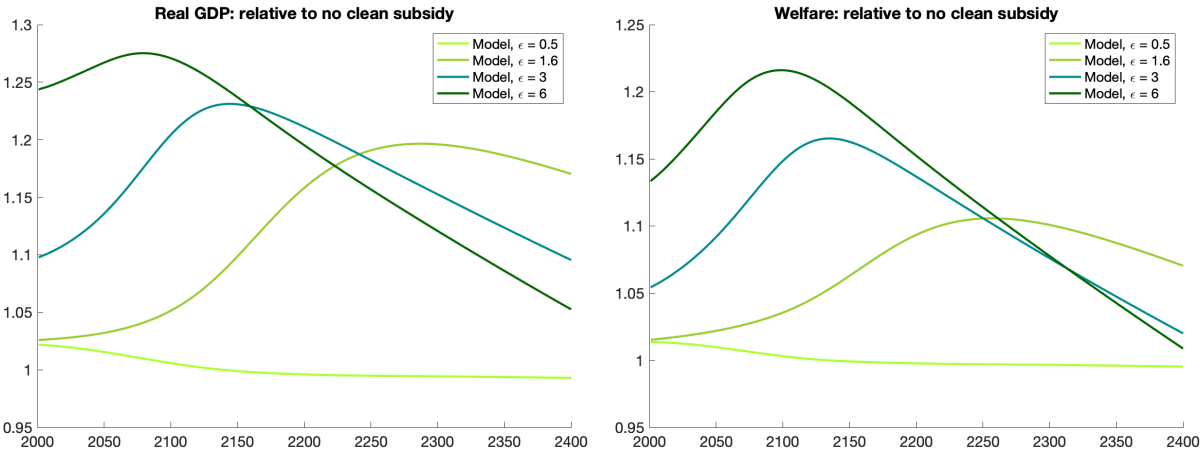

Figure 50: Real GDP and welfare with a clean energy subsidy of 75% across different elasticities of energy substitution.

	PDV of real GDP			Welfare		
	BGP gr	β=0.965	β=0.969	BGP gr	β=0.965	β=0.969
ε=0.5	2.871%	1.005	0.999	2.852%	1.003	0.998
ε=1.6	3.012%	1.094	1.044	2.989%	1.050	0.975
ε=3.0	2.992%	1.155	0.992	2.967%	1.089	0.934
ε=6.0	2.982%	1.186	0.960	2.954%	1.118	0.911

Table 10: PDV of real GDP and welfare gains with a clean energy subsidy of 75% under different elasticities of energy substitutions and discount factors.

F.5 Sensitivity to Carbon Deposit Size

In this subsection, we explore the robustness of the quantitative results when considering different sizes of carbon deposits. Schwerhoff and Stuermer (2020) argue that over the past three centuries, extraction of non-renewable resources has increased, but their real prices do not display an increasing trend. They predict that the interaction of innovation and geological properties allows for a constant real resource price on the Balanced Growth Path.[75]

To incorporate this conclusion in our analysis, we consider that the extraction cost function is characterized by the same parameters $(f_1, \cdots, f_5)$, but modify the carbon deposit size according to $maxCumCO2 \in$ {9,700 $GtCO_2$,19,500 $GtCO_2$,45,300 $GtCO_2$}, where 19,500 $GtCO_2$ denotes the benchmark calibration; 9,700

[75]Moreover, they also predict a constant growth rate of extraction, since innovation converts previously inaccessible lower grade deposits into economic recoverable reserves and since greater resource quantities are found in progressively lower grade deposits (Fundamental Law of Geochemistry).

$GtCO_2$ the CO_2 emissions for the next five centuries according to the RCP 6.0; and 45,300 $GtCO_2$ the stock of fossil fuels resources estimated in Table 4 of Gaedicke et al. (2020).

Figure 51 compares the extraction cost functions under different carbon deposit sizes, so that with larger carbon deposits we are able to reconcile the findings by Schwerhoff and Stuermer (2020) in terms of constant extraction costs for a sufficiently long period of time.

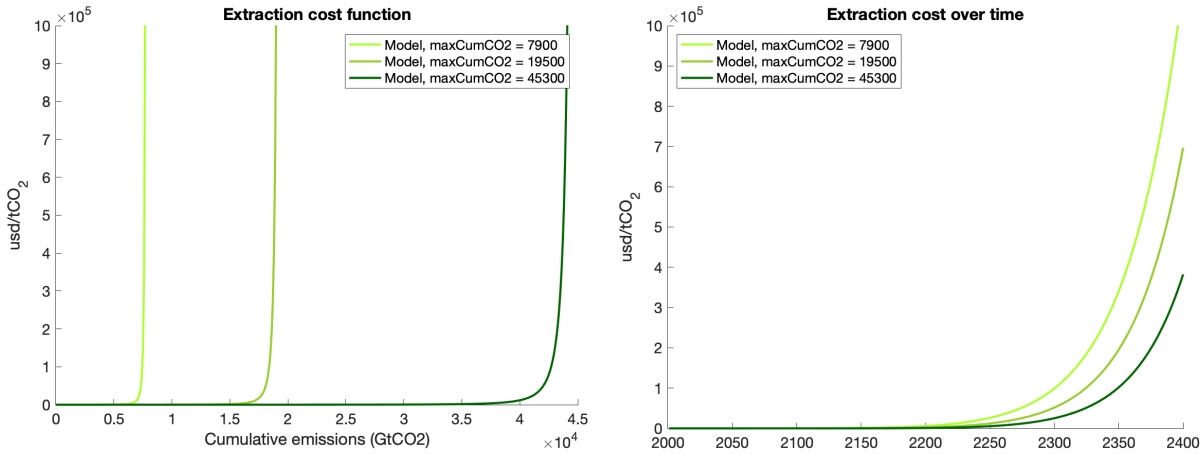

Figure 51: Extraction cost function across different sizes of carbon deposits.

Figure 52 illustrates that larger carbon deposits generate greater peaks in CO_2 emissions occurring progressively later in time, as the extraction cost function displays a larger flat part. Since the extraction cost functions rise sharply when the resource gets exhausted, carbon dioxide emissions eventually decline towards zero. Accordingly, the larger carbon deposits not only generates higher long-run temperatures, but also faster temperature increases in the short-run.

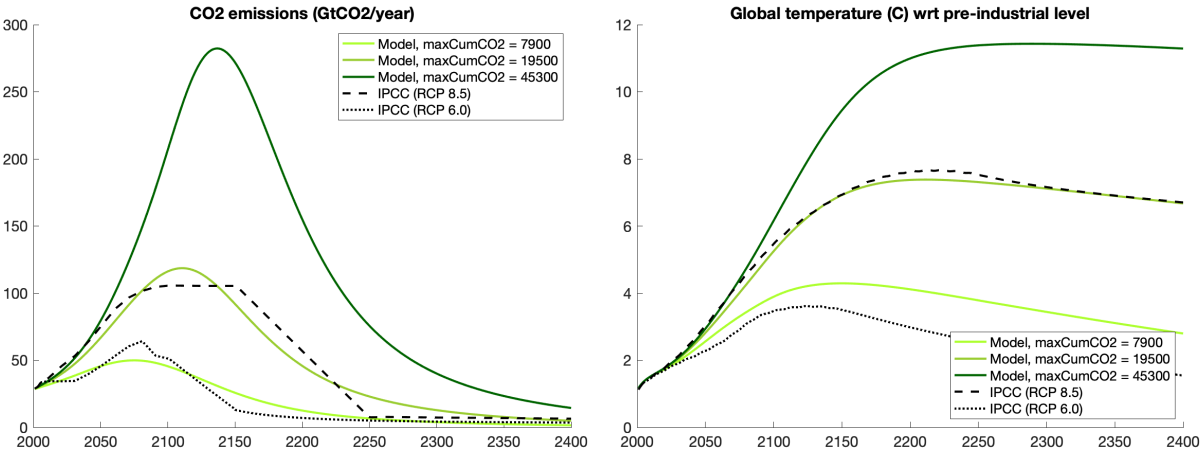

Figure 52: CO_2 emissions and global temperature across different sizes of carbon deposits.

The monotonically increase in the temperature evolution with greater carbon stocks in the ground amplifies economic losses over time, as depicted in Figure 53. In this sense, greater deposits of fossil fuels

widen the dispersion of welfare losses across regions, because warm (cold) regions face amenity and productivity deteriorations (improvements) with higher temperature levels, as shown in Figure 54.

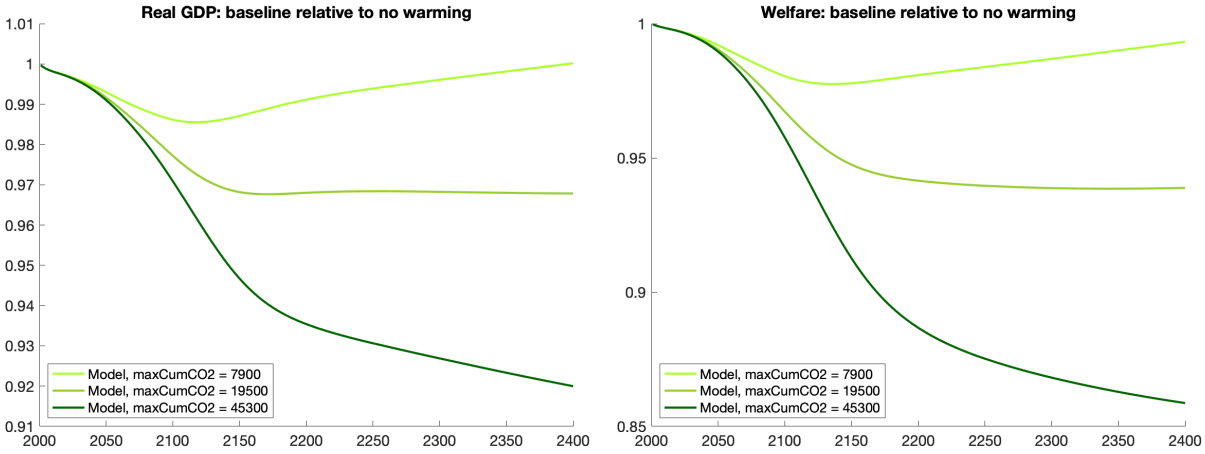

Figure 53: Real GDP and welfare losses over time across different sizes of carbon deposits.

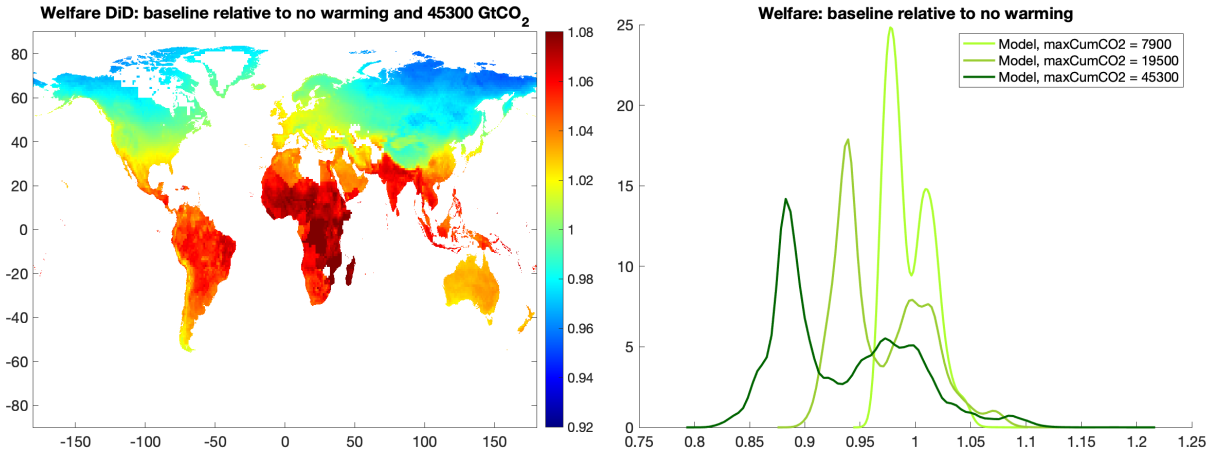

Figure 54: Welfare losses over space across different sizes of carbon deposits.

Now, we evaluate the sensitivity of environmental policies to different sizes of carbon deposits. First, we assess the effect of an spatial- and time-invariant carbon tax of 200%, keeping clean energy subsidies at 0%. Figure 55 displays in solid curves the evolution of global carbon dioxide emissions and temperature without abatement and in dashed curves those with the introduction of a free geoengineering technology in 2100. A carbon tax has the same effect on CO_2 emissions at impact, regardless of the deposit size, since the extraction cost curves have basically the same derivative when evaluated at zero.

When analyzing the evolution of economic benefits of the carbon tax, the initial decline in production is identical across deposit sizes, leading to the same real GDP and welfare global losses at impact. However, over the next decades, welfare is higher the lower the carbon deposit, since the extraction cost curve displays a higher slope, amplifying the rise in fossil fuel price and thus the temperature decline. This situ-

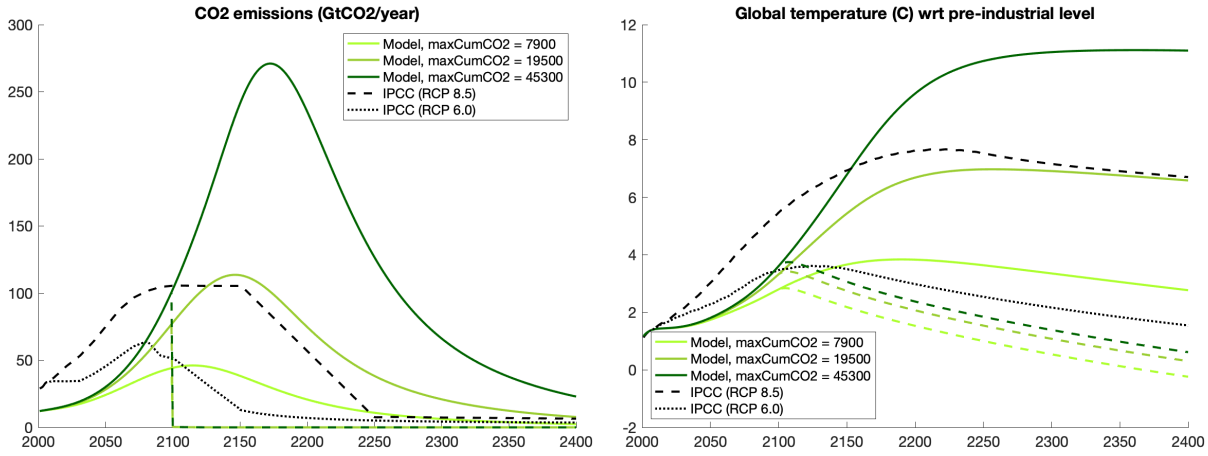

Figure 55: CO_2 emissions and global temperature with a carbon tax of 200% across different sizes of carbon deposits.

ation reverts after one century, because for the scenario with large carbon deposits, the level of cumulative extraction induces a sharp increase in the price of fossil fuels. In the long-run, the welfare benefits of the carbon tax across carbon deposits converge to the same trend.

When considering the introduction of a costless geoengineering technology in 2100, the welfare improvements due to the environmental policy in the long-run are greater with larger carbon deposits, because the difference in long-run temperature with and without the abatement is higher.[76] Tables 11 and 12 summarize the global real GDP and welfare gains for a carbon tax of 200%, and a carbon tax of 200% with an abatement technology under different discount factors and sizes of carbon deposits.

	PDV of real GDP			Welfare		
	BGP gr	β=0.965	β=0.969	BGP gr	β=0.965	β=0.969
$maxCumCO2$=9,700 $GtCO_2$	3.062%	0.984	1.044	3.044%	0.993	1.036
$maxCumCO2$=19,500 $GtCO_2$	3.053%	0.981	1.042	3.032%	0.993	1.033
$maxCumCO2$=45,300 $GtCO_2$	3.041%	0.980	1.037	3.016%	0.998	1.021

Table 11: Real GDP and welfare gains with a carbon tax of 200%, with no abatement, under different sizes of carbon deposits and discount factors.

Now, we examine the effect of a spatial- and time-invariant clean energy subsidy of 75%, keeping carbon taxes at 0%. Figure 57 shows that the clean subsidy yields minuscule reductions in the flow of CO_2 emissions and the path for global temperature, regardless of the size of fossil fuel deposits.

[76] Note that relative real GDP and welfare can be slightly lower in the abatement case for a couple of decades after the invention of the abatement technology. The reason is that the difference in temperatures between the benchmark scenario with and without abatement can be larger than the difference in temperature with and without abatement in the scenario with a carbon tax, depending on the second derivative of the temperature function at the time the abatement technology arrives. After a few decades, this effect is always dominated by the faster increases in temperature in the case without abatement.

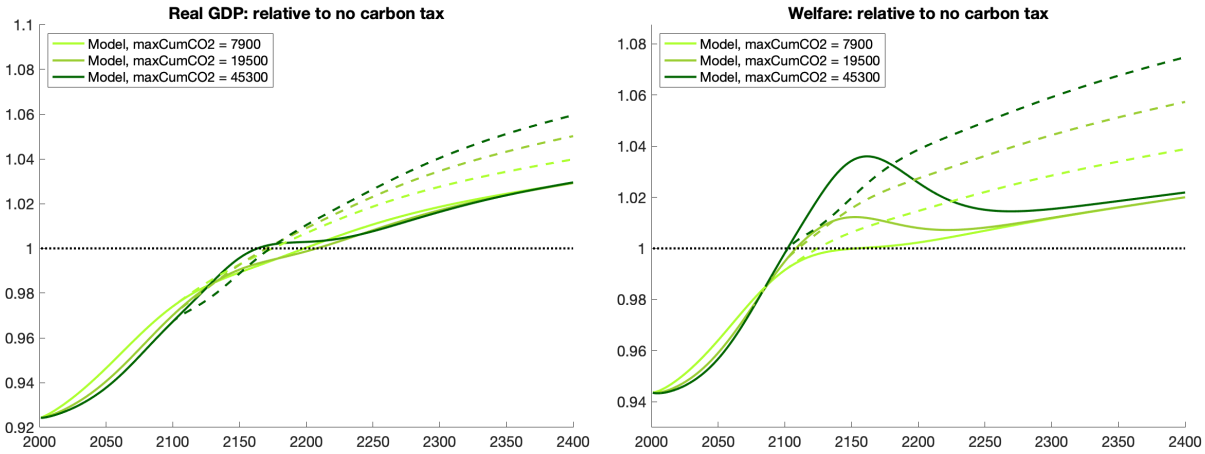

Figure 56: Real GDP and welfare with a carbon tax of 200% across different sizes of carbon deposits.

	PDV of real GDP			Welfare		
	BGP gr	β=0.965	β=0.969	BGP gr	β=0.965	β=0.969
$maxCumCO2$=9,700 $GtCO_2$	3.066%	0.988	1.060	3.052%	1.000	1.061
$maxCumCO2$=19,500 $GtCO_2$	3.065%	0.989	1.074	3.051%	1.006	1.082
$maxCumCO2$=45,300 $GtCO_2$	3.064%	0.989	1.087	3.050%	1.012	1.101

Table 12: PDV of real GDP and welfare gains with a carbon tax of 200%, with abatement, under different sizes of carbon deposits and discount factors.

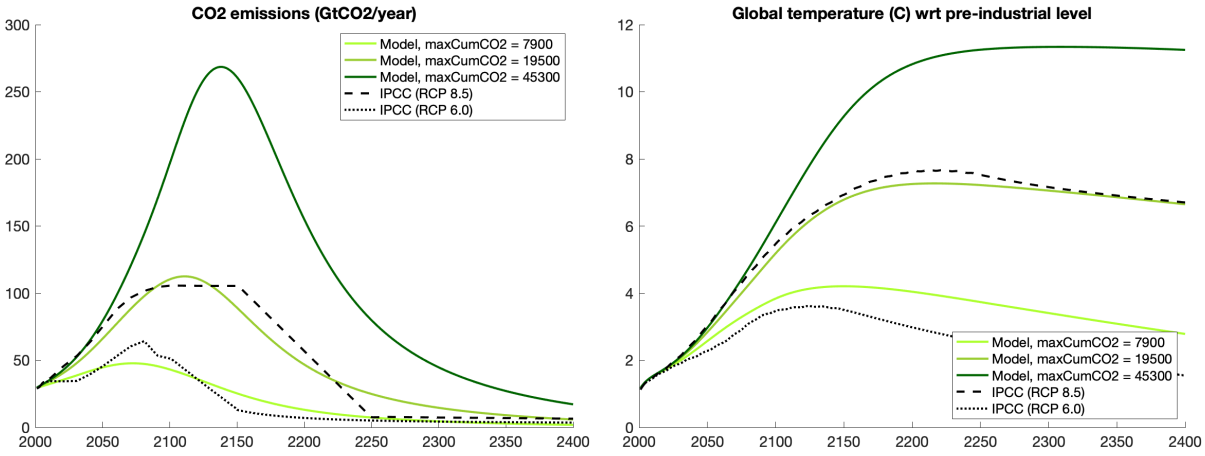

Figure 57: CO_2 emissions and global temperature with a clean energy subsidy of 75% across different sizes of carbon deposits.

The implementation of a green subsidy generates larger short-run benefits in terms of global real GDP and welfare the smaller the carbon deposit, as illustrated in Figure 58. Since extraction costs are higher the smaller the carbon deposit, a clean subsidy has a larger effect on the relative price of fossil fuels to clean energy. Consequently, more households migrate towards the clean-intensive places, which tend to be

high-productive places, rising global production and welfare.

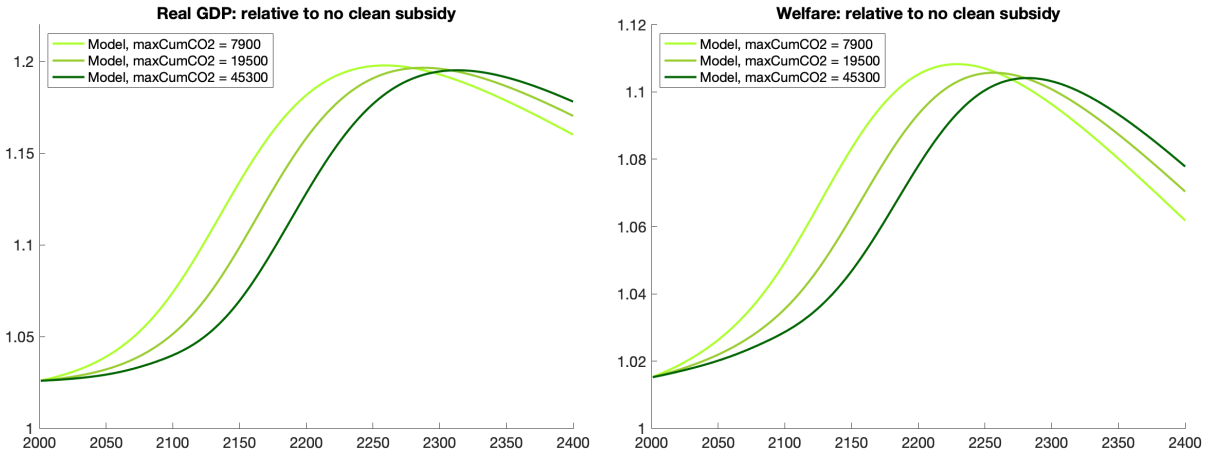

Figure 58: Real GDP and welfare with a clean energy subsidy of 75% across different sizes of carbon deposits.

Table 13 summarizes the global real GDP and welfare gains of a clean energy subsidy of 75% under different discount factors and sizes of carbon deposits. For low (high) discount factors, real GDP and welfare rise with smaller (greater) carbon deposits.

The Balanced Growth Path growth rate of real GDP and welfare, for both carbon taxes and clean subsidies, rise with lower carbon deposits, as the short-run high price of fossil fuels reduces income, increasing natality rates and consequently global population.

	PDV of real GDP			Welfare		
	BGP gr	β=0.965	β=0.969	BGP gr	β=0.965	β=0.969
$maxCumCO2$=9,700 $GtCO_2$	3.018%	1.106	1.029	2.999%	1.056	0.959
$maxCumCO2$=19,500 $GtCO_2$	3.012%	1.094	1.044	2.989%	1.050	0.975
$maxCumCO2$=45,300 $GtCO_2$	3.003%	1.084	1.059	2.975%	1.046	0.991

Table 13: PDV of real GDP and welfare gains with a clean energy subsidy of 75% under different sizes of carbon deposits and discount factors.

G Global Warming in the Worst-Scenario

This section shows the effects of global warming when considering that the damage functions $\Lambda^a(\cdot), \Lambda^b(\cdot)$ are described by the lower 95% confidence interval. Additionally, we quantify the role of environmental policies in improving real GDP and welfare in the most pessimistic scenario.

G.1 The Welfare Cost of Global Warming in the Worst-Scenario

Figure 59 displays the spatial configuration of fundamental amenities and productivities in the year 2200 in the worst-scenario relative to the counterfactual scenario with no warming damages. The most affected places face amenity and productivity losses of 17% and 70%, respectively. As a comparison, in the benchmark scenario, the most affected places experienced losses of 16% and 60%, respectively.[77]

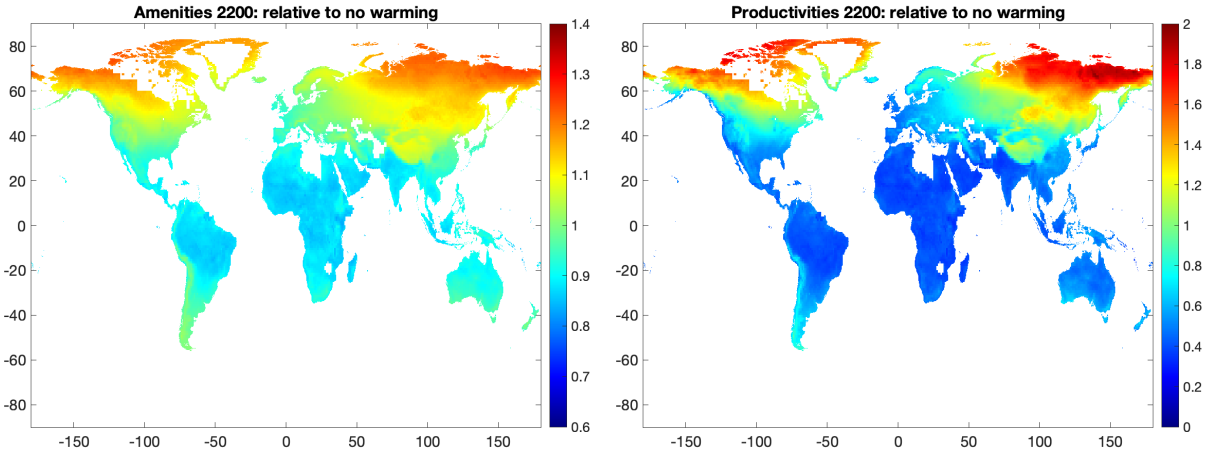

Figure 59: Losses in amenities and productivities from global warming in the year 2200 in the worst-scenario.

On average, the economy experiences welfare losses of 13% and real GDP losses of 7%, which are larger than those projected in the benchmark case: 6% and 3%, respectively. The most damaged zones in the planet face welfare losses of almost 20%, whereas a minuscule 0.02% fraction of population, located in arctic zones, experience welfare gains. Figures 60 and 61 present the spatial distribution of welfare and real GDP losses.

G.2 Environmental Policies in the Worst-Scenario

As shown in Figure 62, the implementation of a carbon tax has the same real GDP and welfare changes at impact, as in the benchmark scenario. However, as time evolves the slopes of both variables are higher than those of the baseline case, as CO_2 levies have a higher potential to improve real GDP and welfare. As a comparison, in the worst-scenario, for a tax of 200%, welfare losses become zero by the year 2076; whereas in the baseline case that event occurs 61 years later.

After those points in time, real GDP and welfare keep rising above the levels of the baseline case. When considering a tax of 200% and a discount factor of β =0.969, welfare rises by 3.5% in the worst-scenario and by 3% in the benchmark scenario.

Under the introduction of an abatement technology, the benefits for the economy are much higher, as

[77] Lower amenities and productivities in the worst-scenario translate into lower income, with respect to the benchmark case. Hence, natality rates augment, yielding higher levels of global population in the most pessimistic scenario.

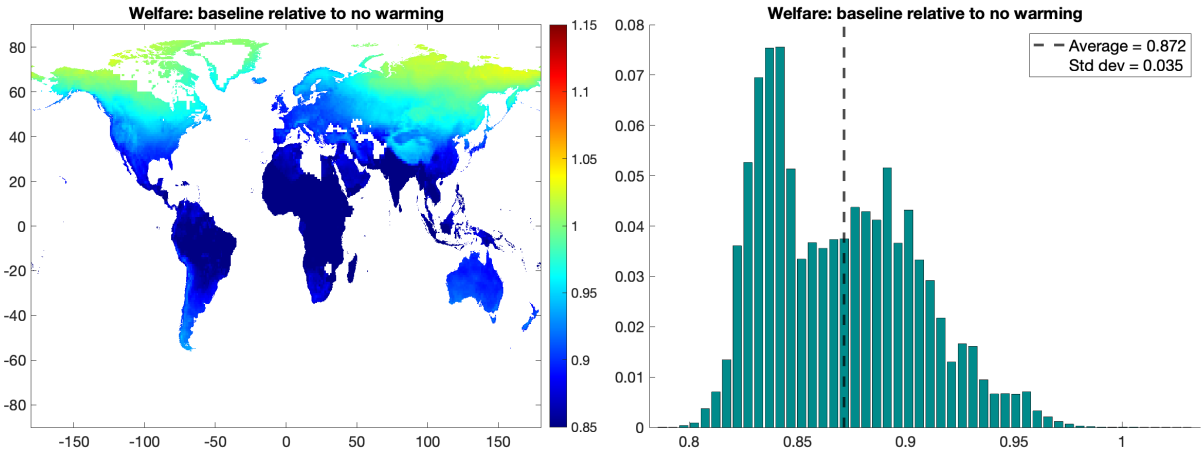

Figure 60: Welfare losses due to global warming in the worst-scenario.

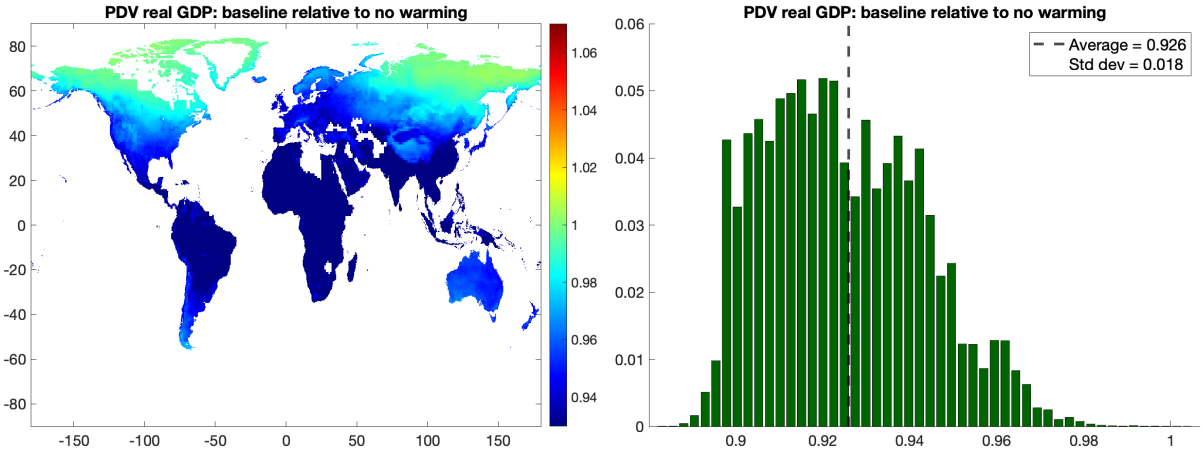

Figure 61: Real GDP losses due to global warming in the worst-scenario.

illustrated by the dashed curves in Figure 62. By the year 2400, global average welfare (real GDP) gains with the geoengineering technology in the most pessimistic scenario reach a value of 8.8% (7.1%); whereas in the benchmark scenario, they are 4.5% (4.8%).

Table 14 quantifies the gains in global average real GDP and welfare from the implementation of taxes of 50%, 100% and 200% on the use of fossil fuels with respect to a scenario with no CO_2 levies, when no geoengineering technology is introduced. Table 15 performs a similar comparison, but taking into account the presence of a free abatement technology in 2100.

As clean energy subsidies barely affect the temperature path, their effects on the economy are very similar with respect to those of the benchmark case. Table 16 shows the values of global average real GDP and welfare gains when enforcing subsidies on the use of clean energy of 25%, 50% and 75%.

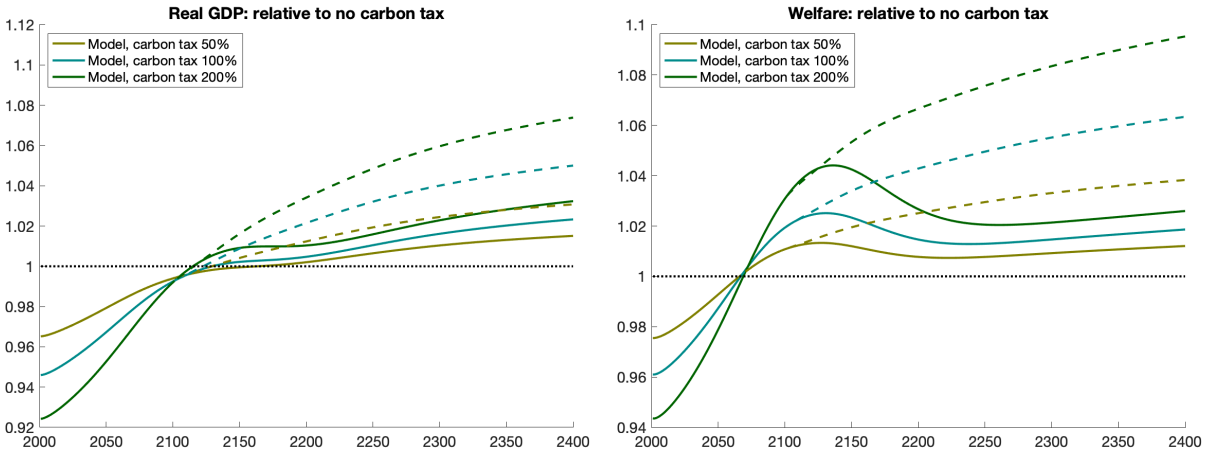

Figure 62: Real GDP and welfare gains under different carbon taxes in the worst-scenario, when considering the introduction of an abatement technology in 2100.

	PDV of real GDP			Welfare		
	BGP gr	β=0.965	β=0.969	BGP gr	β=0.965	β=0.969
τ=0%	3.053%	1	1	3.030%	1	1
τ=50%	3.057%	0.995	1.020	3.033%	1.004	1.017
τ=100%	3.059%	0.993	1.031	3.034%	1.006	1.026
τ=200%	3.061%	0.992	1.043	3.035%	1.011	1.036

Table 14: PDV of real GDP and welfare gains under different carbon taxes and discount factors in the worst-scenario.

	PDV of real GDP			Welfare		
	BGP gr	β=0.965	β=0.969	BGP gr	β=0.965	β=0.969
τ=0%	3.062%	1	1	3.046%	1	1
τ=50%	3.068%	1.002	1.039	3.051%	1.015	1.047
τ=100%	3.071%	1.004	1.065	3.054%	1.025	1.079
τ=200%	3.074%	1.008	1.097	3.057%	1.038	1.119

Table 15: PDV of real GDP and welfare gains under different carbon taxes and discount factors in the worst-scenario, when considering the introduction of an abatement technology in 2100.

H Adaptation

This section deepens on the relevance of adaptation channels in shaping the economic consequences of global warming. More specifically, we compare the temporal and spatial dimension of warming losses in real GDP when considering economies with higher migration, commercial and innovation frictions. Those patterns display large similarities with those of welfare discussed in Section 5.

	PDV of real GDP			Welfare		
	BGP gr	β=0.965	β=0.969	BGP gr	β=0.965	β=0.969
s=0%	3.053%	1	1	3.030%	1	1
s=25%	3.050%	1.011	1.008	3.027%	1.007	1.000
s=50%	3.043%	1.032	1.018	3.019%	1.020	0.995
s=75%	3.022%	1.093	1.036	2.995%	1.051	0.971

Table 16: PDV of real GDP and welfare gains under different clean energy subsidies and discount factors in the worst-scenario.

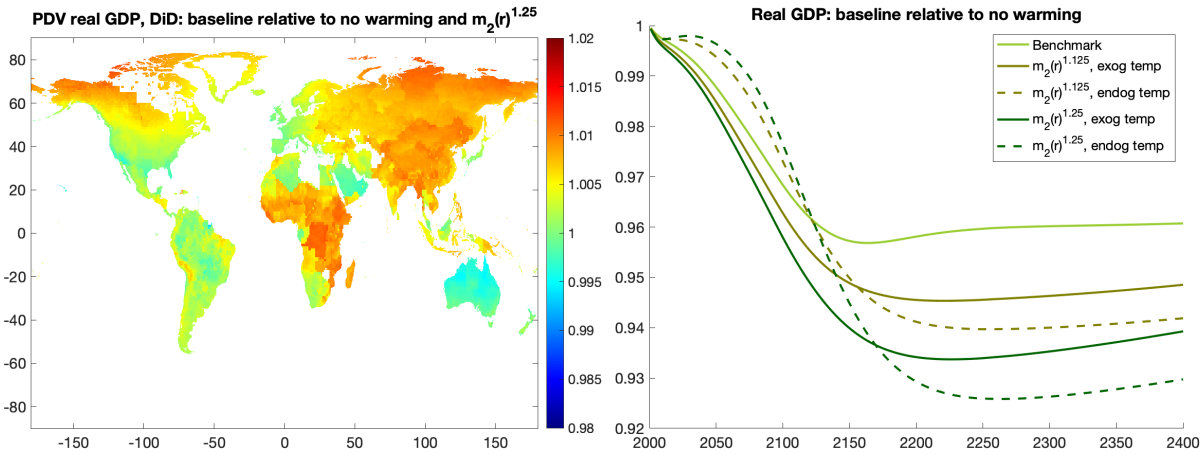

Figure 63: Real GDP across different migration costs.

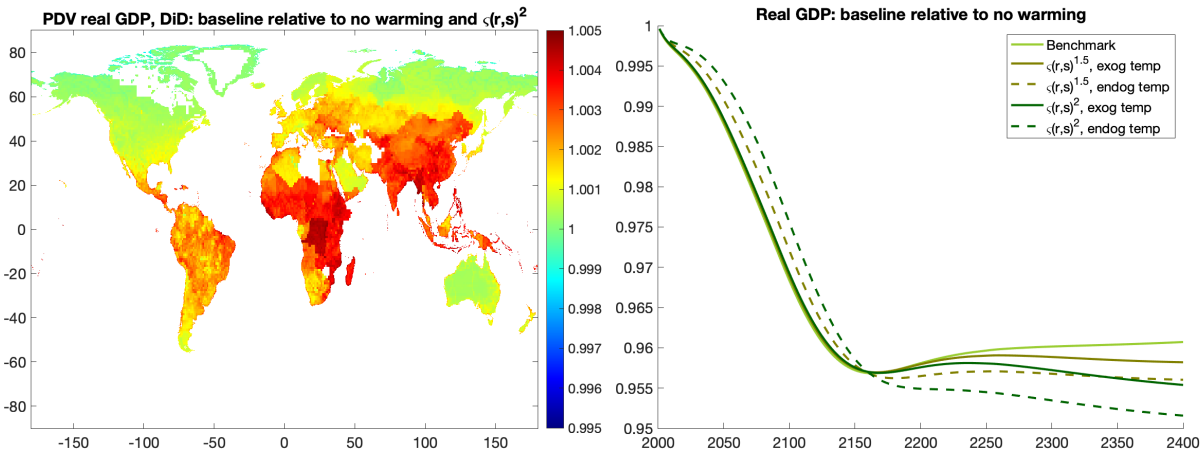

Figure 64: Real GDP across different iceberg trade costs.

I Additional Results regarding Environmental Policies

In this section we delve into the economic response over time and space of the implementation of carbon taxes, assess the welfare benefits of joint carbon taxes and clean energy subsidies, provide additional

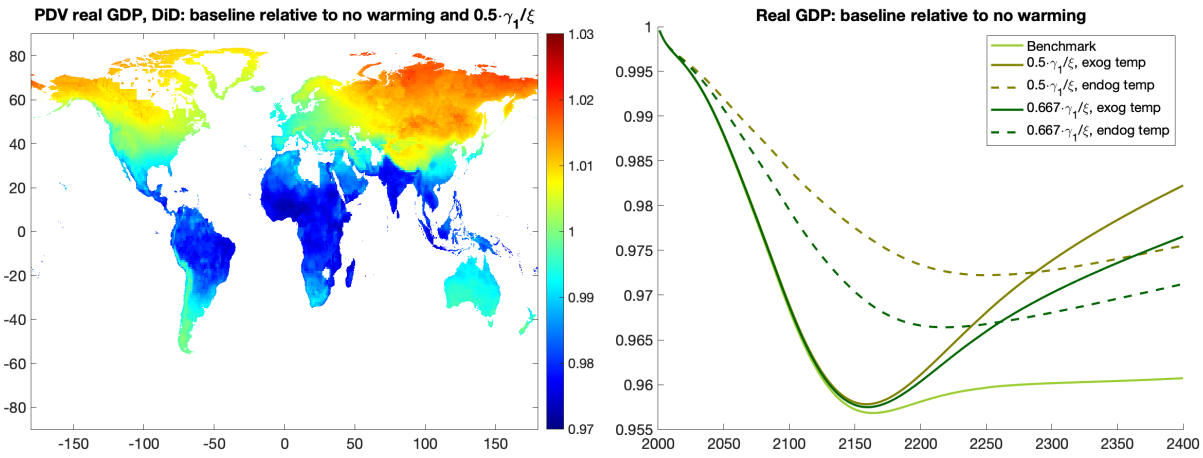

Figure 65: Real GDP across different innovation costs.

results regarding the introduction of an abatement technology in 2100 and contrast the results when the geoengineering innovation becomes available one century later.

I.1 Temporal and Spatial Evolution of Carbon Taxes

At impact, the implementation of a carbon tax of 200% rises the price of energy faced by firms in every cell of the world. However, these increases are heterogeneous: places in which fossil fuels are relatively expensive, like Canada and Europe, face smaller increases in the total energy price, as shown in Figure 66.

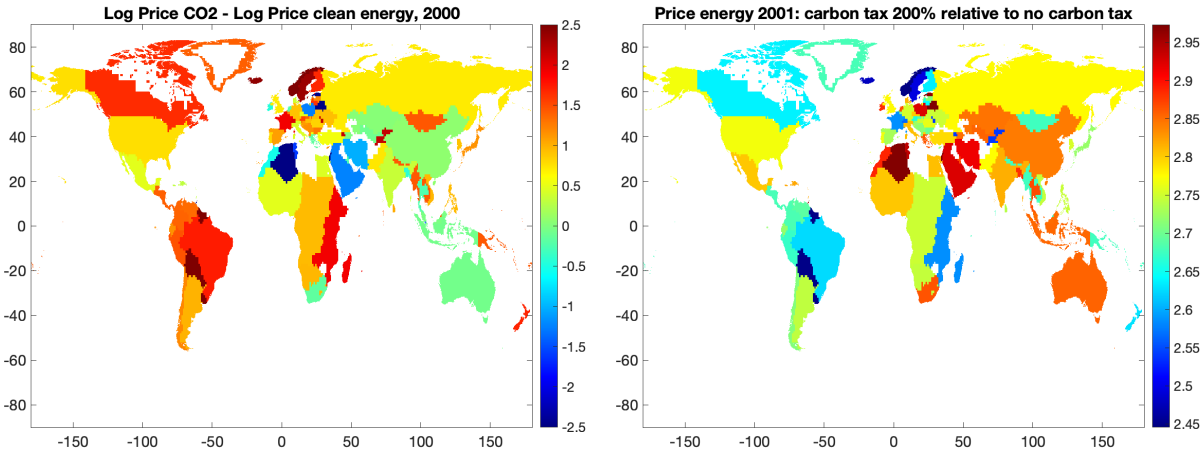

Figure 66: Relative price in fossil fuels and clean energy in the year 2000, and energy price increase with a carbon tax of 200% in the year 2001.

Even though at impact all regions are damaged from the implementation of a CO_2 levy, the places with lower increases in the total price of energy suffer less, as they attract more households, rising current productivity through agglomeration externalities. Figure 67 compares real GDP and welfare in the year

2001 under the implementation of a CO_2 tax of 200% with respect to a scenario with no environmental policy.

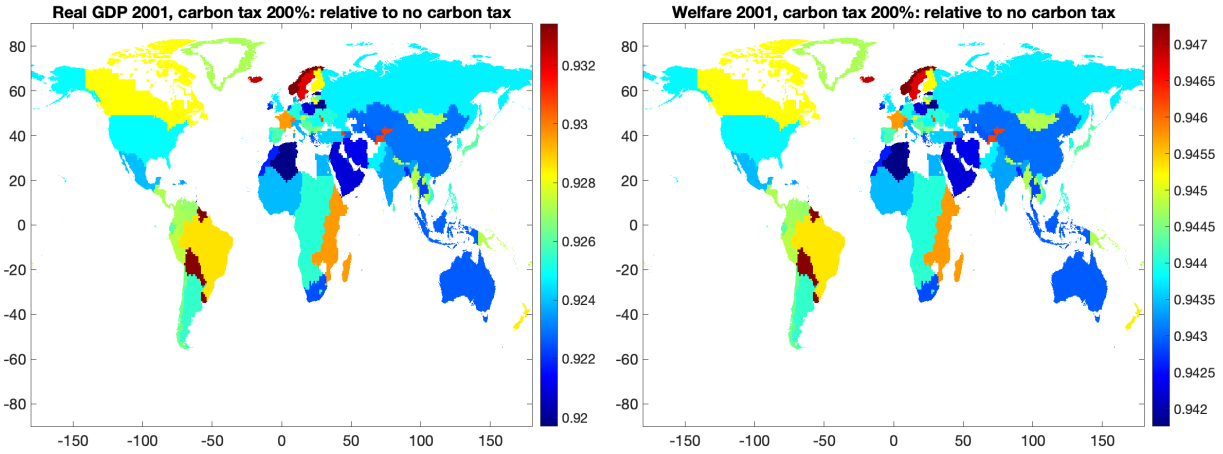

Figure 67: Local real GDP and welfare effects of a carbon tax of 200% in the year 2001.

As time evolves, and as a consequence of the carbon tax, warm regions avoid higher temperatures, whereas cold places are not able to achieve more suitable temperatures for residing and producing. Figure 68 shows the evolution of real GDP and welfare after one century. These maps suggest that the effect of temperature outweighs that of the initial distribution of relative prices.

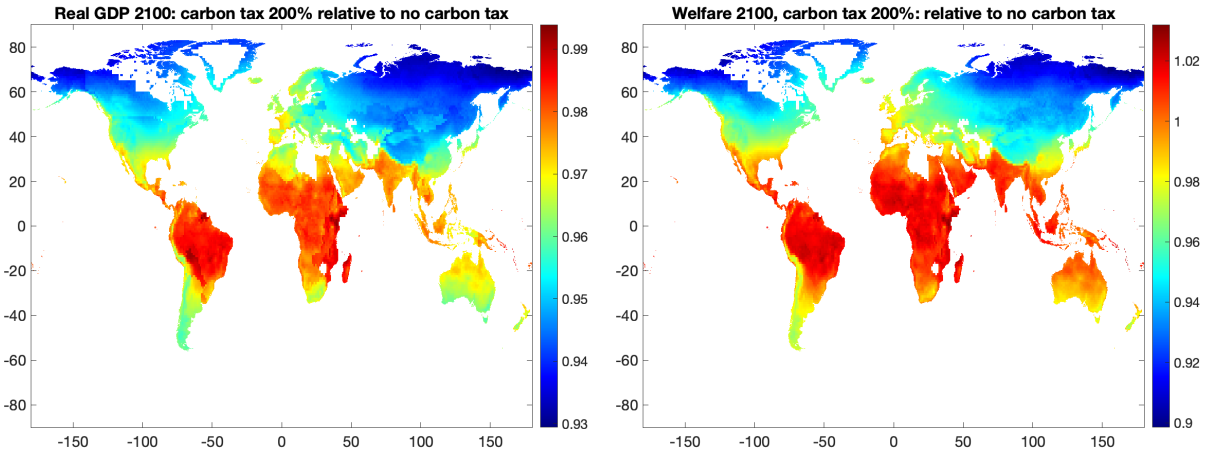

Figure 68: Local real GDP and welfare effects of a carbon tax of 200% in the year 2100.

After another century, the welfare effects preserve the same spatial pattern. However, real GDP has a different spatial configuration, as displayed in Figure 69. The places that have the highest use of fossil fuels with respect to clean energy are projected to undergo losses in real GDP by the year 2200. This is attributed to the large tax burden originated from the high use of fossil fuels.

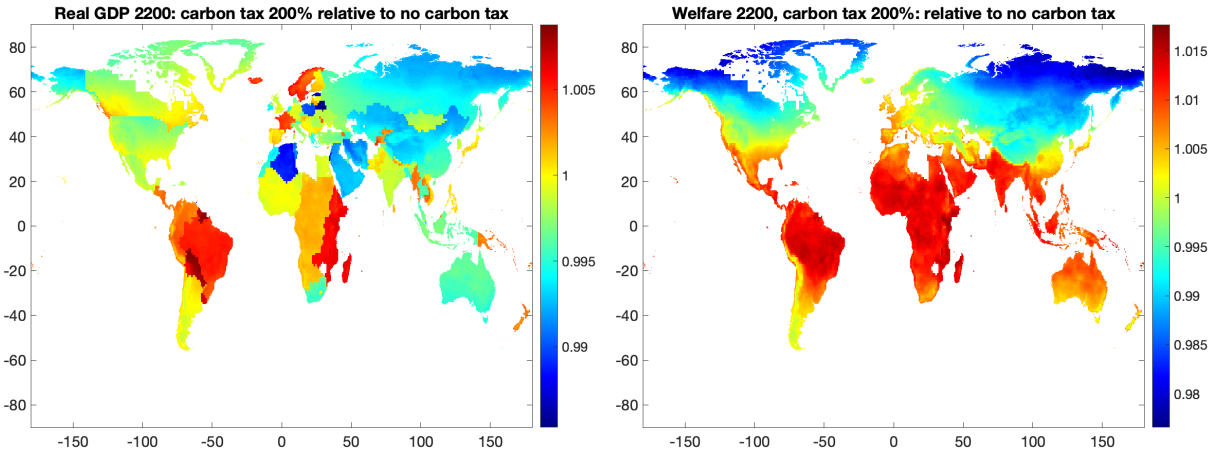

Figure 69: Local real GDP and welfare effects of a carbon tax of 200% in the year 2200.

I.2 Carbon Taxes and Clean Energy Subsidies

Tables 17 and 18 present the global average real GDP and welfare gains arising from the interactions between carbon taxes and clean energy subsidies, respectively, considering discount factors of $\beta = 0.965$ and $\beta = 0.969$. Additionally, Table 19 shows their growth rates in the Balanced Growth Path.

	PDV of real GDP, $\beta = 0.965$				PDV of real GDP, $\beta = 0.969$			
	s=0%	s=25%	s=50%	s=75%	s=0%	s=25%	s=50%	s=75%
τ=0%	1	1.011	1.032	1.094	1	1.009	1.021	1.044
τ=50%	0.991	1.003	1.024	1.087	1.019	1.027	1.037	1.055
τ=100%	0.987	0.980	1.020	1.083	1.030	1.037	1.046	1.060
τ=200%	0.981	0.993	1.015	1.079	1.042	1.048	1.055	1.064

Table 17: PDV of real GDP gains under different carbon taxes, clean energy subsidies and discount rates.

	Welfare, $\beta = 0.965$				Welfare, $\beta = 0.969$			
	s=0%	s=25%	s=50%	s=75%	s=0%	s=25%	s=50%	s=75%
τ=0%	1	1.007	1.020	1.050	1	1.000	0.996	0.975
τ=50%	0.997	1.004	1.017	1.048	1.016	1.015	1.009	0.983
τ=100%	0.995	1.003	1.015	1.047	1.024	1.023	1.016	0.987
τ=200%	0.993	1.001	1.014	1.046	1.033	1.031	1.022	0.989

Table 18: Welfare gains under different carbon taxes, clean energy subsidies and discount rates.

	Real GDP, BGP gr				Welfare, BGP gr			
	s=0%	s=25%	s=50%	s=75%	s=0%	s=25%	s=50%	s=75%
τ=0%	3.043%	3.040%	3.034%	3.012%	3.024%	3.020%	3.012%	2.989%
τ=50%	3.048%	3.045%	3.038%	3.016%	3.028%	3.024%	3.016%	2.992%
τ=100%	3.050%	3.047%	3.040%	3.018%	3.030%	3.026%	3.017%	2.992%
τ=200%	3.053%	3.050%	3.043%	3.019%	3.032%	3.028%	3.019%	2.993%

Table 19: Balanced-Growth-Path growth rate of real GDP and welfare under different carbon taxes and clean energy subsidies.

I.3 Carbon Taxes and Abatement in 2100

Figure 70 presents, in solid curves, the global average real GDP and welfare gains from the enforcement of different levels of carbon taxes, when no geoengineering technology arises, as in Figures 19 and 22. The dotted curves evaluate the benefits of carbon taxes under the introduction of a costless abatement technology in 2100 with respect to the benchmark scenario that considers no environmental policy and no abatement technology.[78] Table 20 summarizes the global real GDP and welfare gains of the implementation of different carbon taxes and the introduction of the abatement technology, with respect to the absence of both policies. The economic benefits overcome those of Table 3.

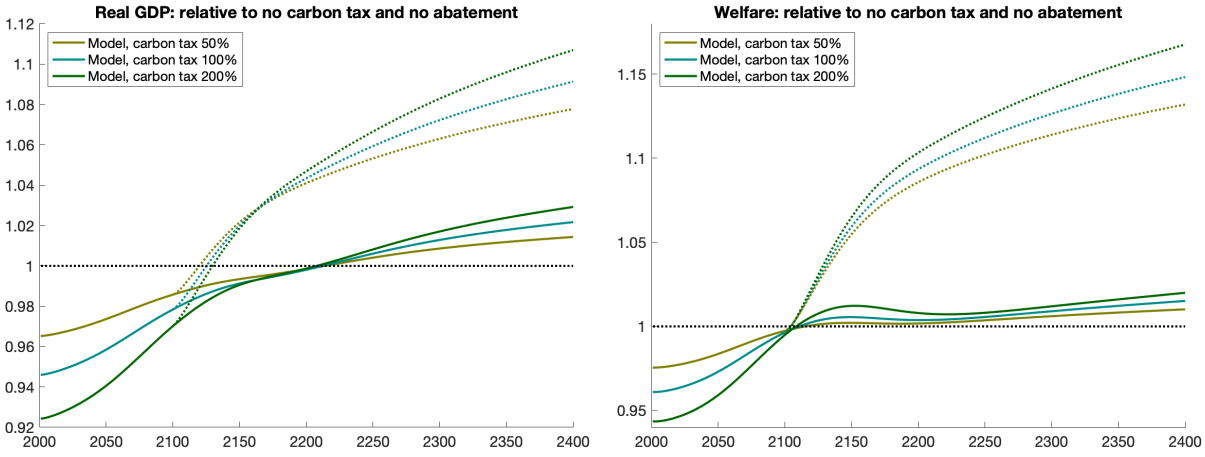

Figure 70: Real GDP and welfare under different carbon taxes, when considering the introduction of an abatement technology in 2100.

[78] The dashed curves of Figure 22 are calculated as the ratio of welfare (or real GDP) with carbon taxes and abatement with respect to zero carbon taxes and abatement. The dotted curves of Figure 70 are calculated as the ratio of welfare (or real GDP) with carbon taxes and abatement with respect to zero carbon taxes and no abatement.

	PDV of real GDP			Welfare		
	BGP gr	β=0.965	β=0.969	BGP gr	β=0.965	β=0.969
τ=0%	3.052%	1	1	3.037%	1	1
τ=50%	3.058%	1.016	1.098	3.043%	1.046	1.152
τ=100%	3.061%	1.014	1.119	3.046%	1.048	1.176
τ=200%	3.065%	1.011	1.145	3.051%	1.049	1.206

Table 20: PDV of real GDP and welfare under different carbon taxes and discount factors, when considering the introduction of an abatement technology in 2100.

I.4 Carbon Taxes and Abatement in 2200

We extend the analysis of the introduction of an abatement technology and assess its benefits in terms of real GDP and welfare when this innovation arises in the year 2200, rather than in the year 2100, as in Section 6.2. A century of delay in geoengineering advances provides more modest beneficial effects for the economy, as a higher share of the total stock of CO_2 has already been released to the atmosphere and, therefore, the reduction in long-run temperature is lower, as shown in Figure 71. Furthermore, the delay of this technology reduces the differences in the steady state temperature across carbon taxes, attenuating the benefits of stronger environmental policies.

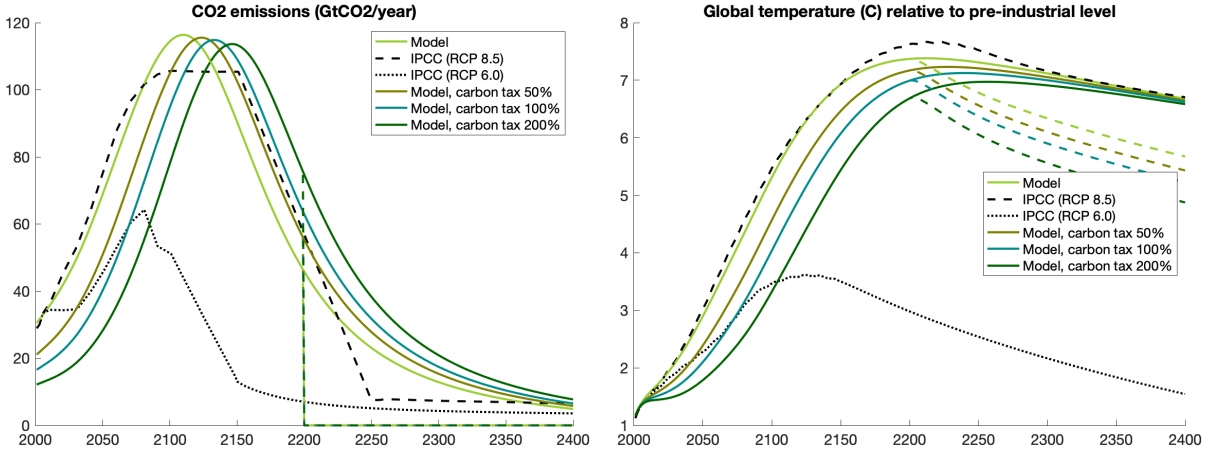

Figure 71: CO_2 emissions and global temperature under different carbon taxes, when considering the introduction of an abatement technology in 2200.

Figure 72 and Table 21 perform a similar analysis to Figure 22 and Table 3. A century of delay in geoengineering advances reduces welfare and real GDP gains in 2.3% and 1.5% for a tax of 200% and a discount factor of $\beta = 0.969$.

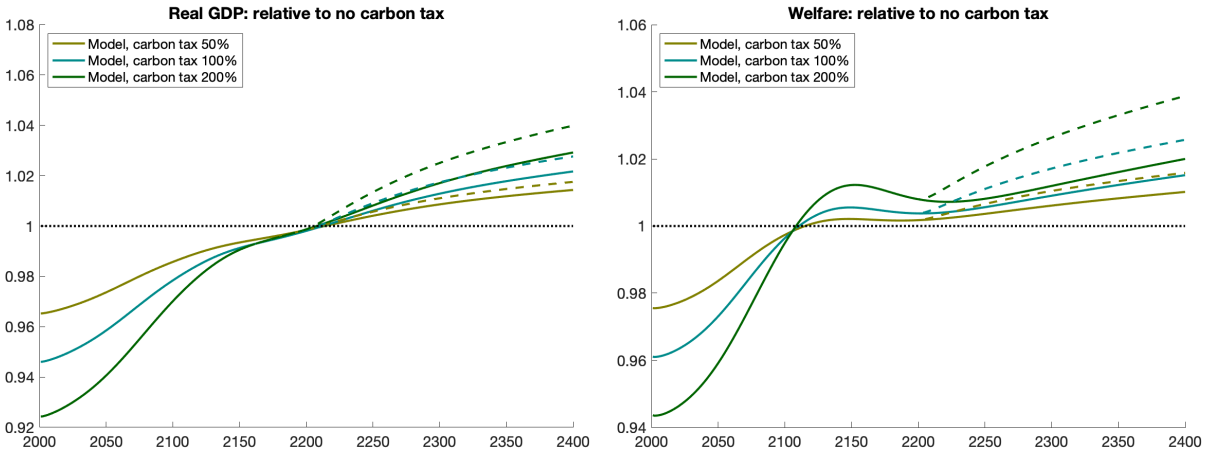

Figure 72: Real GDP and welfare under different carbon taxes, when considering the introduction of an abatement technology in 2200.

	PDV of real GDP			Welfare		
	BGP gr	β=0.965	β=0.969	BGP gr	β=0.965	β=0.969
τ=0%	3.047%	1	1	3.030%	1	1
τ=50%	3.052%	0.992	1.024	3.034%	0.999	1.024
τ=100%	3.055%	0.988	1.039	3.037%	0.998	1.039
τ=200%	3.059%	0.984	1.059	3.041%	0.998	1.059

Table 21: PDV of real GDP and welfare under different carbon taxes and discount factors, when considering the introduction of an abatement technology in 2200.