APPENDICES

A Solution of the model

The Lagrangian function for the planning problem is:

$$\mathcal{L} = \sum_{t=0}^{\infty} \beta^t W_t,$$

where

$$\begin{aligned} W_t = {} & U(Y_t - \psi \sum_{j=a,c,d} x_{jt} \, , \widetilde{F}(S_t)) \\ & + \pi_t \left[\left(Y_{ct}^{\frac{\varepsilon-1}{\varepsilon}} + Y_{dt}^{\frac{\varepsilon-1}{\varepsilon}} \right)^{\frac{\varepsilon}{\varepsilon-1}} - Y_t \right] + \sum_{j=a,c,d} \pi_{jt} \left[L_{jt}^{1-\alpha} A_{jt}^{1-\alpha} x_{jt}^{\alpha} - Y_{jt} \right] \\ & + w_t \left[1 - \sum_{j=a,c,d} L_{jt} \right] + v_t \left[1 - \sum_{j=a,c,d} s_{jt} \right] + \sum_{j=a,c,d} \mu_{jt} \left[\left(1 + \gamma \eta_j s_{jt} \right) A_{jt-1} - A_{jt} \right] \\ & + \omega_t \left[-\xi \left(Y_{dt-1} - Y_{at-1} \right) + (1+\delta) S_{t-1} - S_t \right] + \phi_t \left[Y_{dt} - Y_{at} \right]. \end{aligned}$$

Thus W_t is the undiscounted period t welfare, π_t is the social value of final production in period t, $w_t(v_t)$ is the shadow value of labor (research) in period t, μ_{jt} is the social value of productivity in sector j in period t, and ω_t is the social value of the environment in period t. When writing the quality index of the environment as a function of the stock of CO_2, we have subsumed the relationship through the increase in temperature, Δt.

The first-order condition (FOC) w.r.t. Y_t shows that $\pi_t = U_{ct}$. The FOC w.r.t. S_t shows that $U_{F_t} \widetilde{F}_t' = \omega_t - (1+\delta)\beta\omega_{t+1}$. This is a forward-looking equation that can be solved for ω_t, the social value of a one unit improvement of the environment in t, as:

$$\omega_t = \sum_{k=0}^{\infty} \beta^k (1+\delta)^k U_{F_k} \widetilde{F}_k'.$$

Improving the environment today thus generates a stream of future benefits.

We now solve for the remaining decision variables. First, note that because of the CES specification, both the clean and dirty inputs will be used in strictly positive quantities. For the clean input, we obtain the optimality condition:

$$MP_{ct} = \widehat{p}_{ct} \stackrel{\text{def}}{=} \frac{\pi_{ct}}{\pi_t},$$

i.e., the equality of its marginal product, MP_{ct}, with its social price. A similar condition holds for the dirty input, corrected for the environmental externality:

$$MP_{dt} = \widehat{p}_{dt} + \xi\beta\frac{\omega_{t+1}}{\pi_t} - \frac{\phi_t}{\pi_t}, \tag{8}$$

where $\widehat{p}_{dt} \stackrel{\text{def}}{=} \frac{\pi_{dt}}{\pi_t}$, the social price of the dirty input. The term $\xi\beta\frac{\omega_{t+1}}{\pi_t}$ is equivalent to a tax on the use of the dirty input in a decentralized solution. It ensures a more moderate use of the dirty input than the equality of MP_{dt} with $\widehat{p}_{dt}$ would call for. The extra term $\frac{\phi_t}{\pi_t}$ is due to abatement. Before interpreting it, we give the FOC w.r.t. Y_{at}:

$$-\pi_{at} + \xi\beta\omega_{t+1} - \phi_t \leq 0,$$

with equality when $Y_{at} > 0$. Dividing through by π_t and defining $\widehat{p}_{at} \stackrel{\text{def}}{=} \frac{\pi_{at}}{\pi_t}$, we can write this as:

$$\widehat{p}_{at} \geq \xi\beta\frac{\omega_{t+1}}{\pi_t} - \frac{\phi_t}{\pi_t}.$$

If any abatement is suboptimal, $Y_{at} = 0 < Y_{dt}$, and $\widehat{p}_{at} \geq \xi\beta\frac{\omega_{t+1}}{\pi_t}$; the social marginal cost of abatement is too high compared with its social marginal benefit. However, suppose that abatement is optimal, then either there is partial abatement, $0 < Y_{at} \leq Y_{dt}$, in which case $\widehat{p}_{at} = \xi\beta\frac{\omega_{t+1}}{\pi_t}$, or there is full abatement, $Y_{at} = Y_{dt}$, in which case $\widehat{p}_{at} \leq \xi\beta\frac{\omega_{t+1}}{\pi_t}$. In this last case, the social marginal benefit is larger than the social marginal cost, but the welfare programme is constrained by the fact that abatement can only apply to contemporaneous emissions, not to CO_2 emitted in previous periods (i.e., it is not possible to remove previously emitted CO_2 from the atmosphere). If this is the case, then social welfare may be increased by expanding dirty input production beyond the level where $MP_{dt} = \widehat{p}_{dt} + \xi\beta\frac{\omega_{t+1}}{\pi_t}$. Indeed, then:

$$MP_{dt} = \widehat{p}_{dt} + \widehat{p}_{at}. \tag{9}$$

CO_2 abatement is merely an additional social cost. Thus, we can conclude that:

$$MP_{dt} = \widehat{p}_{dt} + \min\{\widehat{p}_{at}, \xi\beta\frac{\omega_{t+1}}{\pi_t}\}, \tag{10}$$

$$= \widehat{p}_{dt} + \min\{\widehat{p}_{at}, \tau_t\widehat{p}_{dt}\}. \tag{11}$$

Having determined the optimality conditions for Y_{jt}, we now consider the use of labor and physical capital. Both inputs are required in positive amounts. For labor, the value of the marginal product of labor in the production of sector j must equal the social wage rate $\widehat{w}_t \stackrel{\text{def}}{=} \frac{w_t}{\pi_t}$:

$$\widehat{p}_{jt}MP_{Ljt} = \widehat{w}_t, \text{ or} \tag{12}$$

$$(1-\alpha)\widehat{p}_{jt}A_{jt}^{1-\alpha}\left(\frac{x_{jt}}{L_{jt}}\right)^{\alpha} = \widehat{w}_t. \tag{13}$$

Likewise, for machines:

$$\widehat{p}_{jt}MP_{xjt} = \psi, \text{ or} \tag{14}$$

$$\alpha\widehat{p}_{jt}A_{jt}^{1-\alpha}\left(\frac{L_{jt}}{x_{jt}}\right)^{1-\alpha} = \psi, \tag{15}$$

where ψ is the (exogenously given) amount of final goods necessary to build one machine.

Finally, we determine the allocation of scientists, and the production of knowledge. The FOC w.r.t. s_{jt} is:

$$\frac{\mu_{jt}}{\pi_t}\gamma\eta_j A_{jt-1} \leq \widehat{v}_t \stackrel{\text{def}}{=} \frac{v_t}{\pi_t},$$

with equality whenever $s_{jt} > 0$. The left-hand side (LHS) is the social price of sector j knowledge, $\frac{\mu_{jt}}{\pi_t}$, times the marginal knowledge production of an additional researcher. The RHS is the social wage rate of a researcher.

The final set of FOCs characterizes the allocation of productivity improvements in the different sectors across time. The FOC w.r.t. A_{jt} reads:

$$\widehat{p}_{jt}(1-\alpha)L_{jt}^{1-\alpha}\left(\frac{x_{jt}}{A_{jt}}\right)^{\alpha} = \frac{\mu_{jt}}{\pi_t} - \beta\frac{\mu_{jt+1}}{\pi_{t+1}}\frac{\pi_{t+1}}{\pi_t}(1+\gamma\eta_j s_{jt+1}).$$

The LHS is the value of the marginal product of newly acquired knowledge on the use of machines. Using (15), an optimal allocation of knowledge implies that the social price of sector j knowledge, $\frac{\mu_{jt}}{\pi_t}$, must evolve according to the rule:

$$(1-\alpha)\left(\frac{\alpha}{\psi}\right)^{\frac{\alpha}{1-\alpha}} L_{jt}\widehat{p}_{jt}^{\frac{1}{1-\alpha}} = \frac{\mu_{jt}}{\pi_t} - \beta\frac{\mu_{jt+1}}{\pi_{t+1}}\frac{\pi_{t+1}}{\pi_t}(1+\gamma\eta_j s_{jt+1}).$$

Multiplying through by $\pi_t A_{jt}$ and making use of $A_{jt+1} = (1+\gamma\eta_j s_{jt+1})A_{jt}$ give:

$$\mu_{jt}A_{jt} = (1-\alpha)\left(\frac{\alpha}{\psi}\right)^{\frac{\alpha}{1-\alpha}} L_{jt}\pi_t\widehat{p}_{jt}^{\frac{1}{1-\alpha}}A_{jt} + \beta\mu_{jt+1}A_{jt+1}.$$

The social value of acquired knowledge in sector j at time t is the value of A_{jt} priced at its marginal product plus the "standing on the shoulder of giants" effect (future knowledge builds on today's knowledge). Using the forward operator F, multiplying through by $\gamma\eta_j \frac{A_{jt-1}}{A_{jt}}$ and making use of $A_{jt} = (1+\gamma\eta_j s_{jt})A_{jt-1}$ result in:

$$
\begin{aligned}
\frac{\mu_{jt}}{\pi_t}\gamma\eta_j A_{jt-1} &= \frac{1}{\pi_t}\frac{\gamma\eta_j}{1+\gamma\eta_j s_{jt}}\frac{1}{1-\beta F}(1-\alpha)\left(\frac{\alpha}{\psi}\right)^{\frac{\alpha}{1-\alpha}} L_{jt}\pi_t \widehat{p}_{jt}^{\frac{1}{1-\alpha}} A_{jt}, \\
&= \frac{1}{\pi_t}\frac{\gamma\eta_j}{1+\gamma\eta_j s_{jt}}(1-\alpha)\left(\frac{\alpha}{\psi}\right)^{\frac{\alpha}{1-\alpha}} \sum_{k=0}^{\infty}\beta^k L_{jt+k}\pi_{t+k}\widehat{p}_{jt+k}^{\frac{1}{1-\alpha}} A_{jt+k},
\end{aligned} \tag{16}
$$

so that the social value of allocating an extra researcher to sector j is given by the discounted sum of future knowledge levels, appropriately valued and weighted.

Solving (15) for x_{jt} gives:

$$
x_{jt} = \left(\alpha\frac{\widehat{p}_{jt}}{\psi}\right)^{\frac{1}{1-\alpha}} A_{jt}L_{jt}, \tag{17}
$$

which can be plugged into (13) to yield the social price of sector j output, as a weighted average of the exogenous machine price, ψ, and the shadow price of labor, $\widehat{w}_{jt}$:

$$
\widehat{p}_{jt} = \frac{1}{\mathcal{A}}\frac{1}{A_{jt}^{1-\alpha}}\widehat{w}_t^{1-\alpha}\psi^{\alpha}, \tag{18}
$$

where $\mathcal{A} \stackrel{\text{def}}{=} \alpha^{\alpha}(1-\alpha)^{1-\alpha}$. Hence, at an optimum, $\widehat{p}_{jt}$ will equal the social marginal cost of sector j output.

Next, the FOCs for Y_{ct} and Y_{dt} can be used to relate these input levels to aggregate output, Y_t and the shadow prices of the inputs:

$$
Y_{ct} = Y_t \widehat{p}_{ct}^{-\varepsilon} \text{ and} \tag{19}
$$

$$
\begin{aligned}
Y_{dt} &= Y_t\left[\widehat{p}_{dt} + \min\{\widehat{p}_{at}, \tau_t\widehat{p}_{dt}\}\right]^{-\varepsilon} \\
&= Y_t\widehat{p}_{dt}^{-\varepsilon}\left[1+\min\{\frac{\widehat{p}_{at}}{\widehat{p}_{dt}}, \tau_t\}\right]^{-\varepsilon} \\
&= Y_t\widehat{p}_{dt}^{-\varepsilon}\left[1+\min\{\left(\frac{A_{dt}}{A_{at}}\right)^{1-\alpha}, \tau_t\}\right]^{-\varepsilon},
\end{aligned} \tag{20}
$$

where the last equality follows from (18). Making use of the final good production function, we obtain:

$$
1 = \widehat{p}_{ct}^{1-\varepsilon} + \widehat{p}_{dt}^{1-\varepsilon}\left[1+\min\{\left(\frac{A_{dt}}{A_{at}}\right)^{1-\alpha}, \tau_t\}\right]^{1-\varepsilon}.
$$

As $\widehat{p}_{jt}$ $(j = c, d, a)$ are given by (18):

$$
1 = \frac{1}{\mathcal{A}^{1-\varepsilon}}\widehat{w}_t^{\varphi}\psi^{\alpha(1-\varepsilon)}\left(A_{ct}^{-\varphi} + A_{dt}^{-\varphi}\left[1+\min\{\left(\frac{A_{dt}}{A_{at}}\right)^{1-\alpha}, \tau_t\}\right]^{1-\varepsilon}\right),
$$

where $\varphi \stackrel{\text{def}}{=} (1-\alpha)(1-\varepsilon)$.

Hence, we can solve for the social value of the wage rate:

$$\widehat{w}_t = \mathcal{A}^{\frac{1}{1-\alpha}} \psi^{-\frac{\alpha}{1-\alpha}} \left(A_{ct}^{-\varphi} + A_{dt}^{-\varphi} \left[1 + \min\{ \left(\frac{A_{dt}}{A_{at}} \right)^{1-\alpha}, \tau_t \} \right]^{1-\varepsilon} \right)^{-\frac{1}{\varphi}}$$

$$= \mathcal{A}^{\frac{1}{1-\alpha}} \psi^{-\frac{\alpha}{1-\alpha}} B_t, \tag{21}$$

thereby implicitly defining the "sector average" productivity parameter B_t as:

$$B_t \overset{\text{def}}{=} \left(A_{ct}^{-\varphi} + A_{dt}^{-\varphi} \left[1 + \min\{ \left(\frac{A_{dt}}{A_{at}} \right)^{1-\alpha}, \tau_t \} \right]^{1-\varepsilon} \right)^{-\frac{1}{\varphi}}. \tag{22}$$

From (21) and (18), the social prices of the two inputs as well as the price of abatement are then:

$$\widehat{p}_{jt} = \left(\frac{B_t}{A_{jt}} \right)^{1-\alpha} \quad (j = c, d, a). \tag{23}$$

Machine use in sector j can be obtained from (17) and (23):

$$x_{jt} = \widehat{p}_{jt}^{\frac{1}{1-\alpha}} \left(\frac{\alpha}{\psi} \right)^{\frac{1}{1-\alpha}} A_{jt} L_{jt} = \left(\frac{\alpha}{\psi} \right)^{\frac{1}{1-\alpha}} B_t L_{jt}, \tag{24}$$

and therefore the aggregate machine cost (the share of final good production used for capital) is:

$$AMC_t \overset{\text{def}}{=} \sum_{j=c,d,a} \psi x_{jt} = \psi \left(\frac{\alpha}{\psi} \right)^{\frac{1}{1-\alpha}} B_t \sum_{j=c,d,a} L_{jt} = \psi^{-\frac{\alpha}{1-\alpha}} \alpha^{\frac{1}{1-\alpha}} B_t, \tag{25}$$

where the last equality follows from the normalization of the labor supply to one.

To find the levels of production in the three sectors, we plug the solution for x_{jt} (24) into the production function, yielding:

$$Y_{jt} = A_{jt} L_{jt} \left(\alpha \frac{\widehat{p}_{jt}}{\psi} \right)^{\frac{\alpha}{1-\alpha}} = \left(\frac{\alpha}{\psi} \right)^{\frac{\alpha}{1-\alpha}} L_{jt} A_{jt}^{1-\alpha} B_t^{\alpha} \quad (j = c, d, a).$$

Therefore $L_{jt} = \left(\frac{\alpha}{\psi} \right)^{-\frac{\alpha}{1-\alpha}} A_{jt}^{\alpha-1} B_t^{-\alpha} Y_{jt}$, which allows us to write (16) as:

$$\frac{\mu_{jt}}{\pi_t} \gamma \eta_j A_{jt-1} = \frac{1}{\pi_t} \frac{\gamma \eta_j}{1 + \gamma \eta_j s_{jt}} (1-\alpha) \sum_{k=0}^{\infty} \beta^k A_{jt+k}^{\alpha} B_{t+k}^{-\alpha} Y_{jt+k} \pi_{t+k} \widehat{p}_{jt+k}^{\frac{1}{1-\alpha}}$$

$$= \frac{1}{\pi_t} \frac{\gamma \eta_j}{1 + \gamma \eta_j s_{jt}} (1-\alpha) \sum_{k=0}^{\infty} \beta^k \pi_{t+k} \widehat{p}_{jt+k} Y_{jt+k},$$

where the second equality follows from (23). This is expression (5) in the text.

On the other hand, (19) and (20) together with (23) give:

$$Y_{ct} = Y_t \left(\frac{B_t}{A_{ct}} \right)^{-\varepsilon(1-\alpha)}, \text{ and} \tag{26}$$

$$Y_{dt} = Y_t \left(\frac{B_t}{A_{dt}} \right)^{-\varepsilon(1-\alpha)} \left[1 + \min\{ \left(\frac{A_{dt}}{A_{at}} \right)^{1-\alpha}, \tau_t \} \right]^{-\varepsilon}. \tag{27}$$

The last three expressions now allow us to write the labor balance equation as:

$$A_{ct}^{-\varphi}B_t^{-(1-\varphi)}Y_t + A_{dt}^{-\varphi}B_t^{-(1-\varphi)}\left[1+\min\left\{\left(\frac{A_{dt}}{A_{at}}\right)^{1-\alpha},\tau_t\right\}\right]^{-\varepsilon}Y_t + Y_{at}A_{at}^{-(1-\alpha)}B_t^{-\alpha} = \left(\frac{\alpha}{\psi}\right)^{\frac{\alpha}{1-\alpha}}. \tag{28}$$

We now look at the three possibilities. The first is where there is full abatement, $Y_{at} = Y_{dt}$, such that $\min\{\left(\frac{A_{dt}}{A_{at}}\right)^{1-\alpha},\tau_t\} = \left(\frac{A_{dt}}{A_{at}}\right)^{1-\alpha}$. In that case:

$$Y_{at} = Y_{dt} = Y_t\left(\frac{B_t}{A_{dt}}\right)^{-\varepsilon(1-\alpha)}\left[1+\left(\frac{A_{dt}}{A_{at}}\right)^{1-\alpha}\right]^{-\varepsilon}.$$

Making use of these values for Y_{at} and Y_{dt} in the labor balance equation (28) reduces the latter to:

$$\left(\frac{\alpha}{\psi}\right)^{-\frac{\alpha}{1-\alpha}}B_t^{-(1-\varphi)}B_t^{-\varphi}Y_t = 1,$$

so that

$$\begin{aligned}
Y_t^{FA} &= \left(\frac{\alpha}{\psi}\right)^{\frac{\alpha}{1-\alpha}}B_t,\\
Y_{ct}^{FA} &= \left(\frac{\alpha}{\psi}\right)^{\frac{\alpha}{1-\alpha}}B_t^{\varphi+\alpha}A_{ct}^{1-(\varphi+\alpha)},\\
Y_{dt}^{FA} &= \left(\frac{\alpha}{\psi}\right)^{\frac{\alpha}{1-\alpha}}B_t^{\varphi+\alpha}A_{dt}^{1-(\varphi+\alpha)}\left[1+\left(\frac{A_{dt}}{A_{at}}\right)^{1-\alpha}\right]^{-\varepsilon}, \text{ and}\\
Y_{at}^{FA} &= Y_{dt}^{FA}.
\end{aligned} \tag{29}$$

In the second case, there is partial abatement such that $0 < Y_{at} < Y_{dt}$ and $\min\{\left(\frac{A_{dt}}{A_{at}}\right)^{1-\alpha},\tau_t\} = \left(\frac{A_{dt}}{A_{at}}\right)^{1-\alpha} = \tau_t$. Now (28) becomes:

$$\left\{A_{ct}^{-\varphi}B_t^{-(1-\varphi)}Y_t + A_{dt}^{-\varphi}B_t^{-(1-\varphi)}\left[1+\tau_t\right]^{-\varepsilon}Y_t + Y_{at}A_{at}^{-(1-\alpha)}B_t^{-\alpha}\right\} = \left(\frac{\alpha}{\psi}\right)^{\frac{\alpha}{1-\alpha}},$$

yielding:

$$Y_t^{PA} = \left\{\left(\frac{\alpha}{\psi}\right)^{\frac{\alpha}{1-\alpha}} - Y_{at}A_{at}^{-(1-\alpha)}B_t^{-\alpha}\right\}\frac{B_t^{1-\varphi}}{\left[A_{ct}^{-\varphi}+A_{dt}^{-\varphi}\left[1+\tau_t\right]^{-\varepsilon}\right]}, \tag{30}$$

$$Y_{ct}^{PA} = Y_t^{PA}\left(\frac{B_t}{A_{ct}}\right)^{-\varepsilon(1-\alpha)}, \tag{31}$$

$$Y_{dt}^{PA} = Y_t^{PA}\left(\frac{B_t}{A_{dt}}\right)^{-\varepsilon(1-\alpha)}\left[1+\tau_t\right]^{-\varepsilon}. \tag{32}$$

For this to be compatible with partial abatement, we need $Y_{at} \le Y_{dt}^{PA}$, which can be shown to be equivalent with:

$$Y_{at} \le \left(\frac{\alpha}{\psi}\right)^{\frac{\alpha}{1-\alpha}}\frac{A_{dt}^{\varepsilon(1-\alpha)}B_t^{\alpha}\left[1+\tau_t\right]^{-\varepsilon}}{\left\{A_{ct}^{-\varphi}+A_{dt}^{-\varphi}\left[1+\tau_t\right]^{1-\varepsilon}\right\}}. \tag{33}$$

Given partial abatement is optimal, we have $\left(\frac{A_{dt}}{A_{at}}\right)^{1-\alpha} = \tau_t$, and this condition reduces to:

(34)
$$Y_{at} \leq \left(\frac{\alpha}{\psi}\right)^{\frac{\alpha}{1-\alpha}} \frac{A_{dt}^{\varepsilon(1-\alpha)} B_t^{1-\varepsilon(1-\alpha)}}{[1+\tau_t]^{\varepsilon}}.$$

In the third case, there is no abatement: $Y_{at} = 0$ and $\min\{\left(\frac{A_{dt}}{A_{at}}\right)^{1-\alpha}, \tau_t\} = \tau_t$. The equilibrium value for Y_t is then found by setting $Y_{at} = 0$ in (30):

(35)
$$Y_t^{NA} = \left(\frac{\alpha}{\psi}\right)^{\frac{\alpha}{1-\alpha}} \frac{B_t^{1-\varphi}}{\left[A_{ct}^{-\varphi} + A_{dt}^{-\varphi}[1+\tau_t]^{-\varepsilon}\right]},$$

and

$$Y_{ct}^{NA} = \left(\frac{\alpha}{\psi}\right)^{\frac{\alpha}{1-\alpha}} \frac{B_t^{\alpha} A_{ct}^{\varepsilon(1-\alpha)}}{\left[A_{ct}^{-\varphi} + A_{dt}^{-\varphi}[1+\tau_t]^{-\varepsilon}\right]},$$
$$Y_{dt}^{NA} = \left(\frac{\alpha}{\psi}\right)^{\frac{\alpha}{1-\alpha}} \frac{B_t^{\alpha} A_{dt}^{\varepsilon(1-\alpha)}}{\left[A_{ct}^{-\varphi} + A_{dt}^{-\varphi}[1+\tau_t]^{-\varepsilon}\right]} [1+\tau_t\}]^{-\varepsilon}, \text{ and}$$
$$Y_{at}^{NA} = 0.$$

When solving the model, we search for a sequence $\{Y_{at}, \tau_t, s_{ct}, s_{dt}\}_{t=0}^{T}$ (where T is large) that maximizes:

$$\sum_{t=0}^{T} \beta^t U(Y_t^{PA}(Y_{at}) - AMC_t, \widetilde{F}((1+\delta)S_{t-1} - \xi(Y_{dt-1} - Y_{at-1}))),$$

subject to the equality constraints (25), (31), (32), $s_{at} = 1 - s_{ct} - s_{dt}$, $A_{jt} = (1 + \gamma\eta_j s_{jt}) A_{jt-1}$ (all t and j), the non-linear inequality constraint (33), the non-negativity constraint $Y_t^{PA}(Y_{at}) - AMC_t \geq 0$, and with the initial productivity levels A_{j0} given.

B Calibration of the model

Without any policy intervention in the base period, the laissez-faire levels for clean and dirty input production are:

$$Y_{c0} = \left(\frac{\alpha}{\psi}\right)^{\frac{\alpha}{1-\alpha}} \frac{B_0^{\alpha} A_{c0}^{\varepsilon(1-\alpha)}}{A_{c0}^{-\varphi} + A_{d0}^{-\varphi}}, \text{ and } Y_{dt} = \left(\frac{\alpha}{\psi}\right)^{\frac{\alpha}{1-\alpha}} \frac{B_0^{\alpha} A_{d0}^{\varepsilon(1-\alpha)}}{A_{c0}^{-\varphi} + A_{d0}^{-\varphi}},$$

where $B_0 \stackrel{\text{def}}{=} \left(A_{c0}^{-\varphi} + A_{d0}^{-\varphi}\right)^{-\frac{1}{\varphi}}$, and $\varphi \stackrel{\text{def}}{=} (1-\varepsilon)(1-\alpha)$. This system can be solved for A_{c0} and A_{d0}:

$$A_{d0} = \left(\frac{\alpha}{\psi}\right)^{-\frac{\alpha}{1-\alpha}} Y_{d0} \left[1 + \left(\frac{Y_{d0}}{Y_{c0}}\right)^{\frac{1-\varepsilon}{\varepsilon}}\right]^{\frac{\alpha+\varphi}{\varphi}},$$
$$A_{c0} = \left(\frac{\alpha}{\psi}\right)^{-\frac{\alpha}{1-\alpha}} Y_{c0} \left[1 + \left(\frac{Y_{c0}}{Y_{d0}}\right)^{\frac{1-\varepsilon}{\varepsilon}}\right]^{\frac{\alpha+\varphi}{\varphi}}.$$

As in AABH, we have used the values for world primary energy production by energy carrier during the period 2002–2006 (EIA, 2008, Table 11.1) and doubled them. Dirty carriers (coal, natural gas, crude oil, and natural gas plant liquids) yield 3786 QBTU, while clean carriers (solar and wind power, nuclear electric power, hydroelectric power, geothermal, electricity generation from wood and waste) provide 615 QBTU.[25] Under the assumptions that $\alpha = \frac{1}{3}, \varepsilon = 3$ (and therefore $\varphi = -\frac{4}{3}$), $\rho = 0.015$, and the normalization $\psi = \alpha^2$ (Acemoglu et al., 2012), we obtain the following estimates for A_{d0}, A_{c0}, and B_0: $A_{d0} = 2658$, $A_{c0} = 1072$, and $B_0 = 3232$.

C Long run growth scenarios

In this section we inquire about the optimal R&D policy in the steady state. We distinguish between a steady state where there is no abatement and one where there is full abatement. We omit the time index and denote the growth rate in variable y as $\dot{y}$, i.e., $\dot{y} = \frac{\mathrm{d}\log y}{\mathrm{d}t}$.

When there is no abatement, $\tau < \left(\frac{A_d}{A_a}\right)^{1-\alpha}$ and the productivity index (22) reduces to

$$B(\ell) \stackrel{\text{def}}{=} \left(A_c^{-\varphi} + A_d^{-\varphi}\left[1+\tau\right]^{\ell-\varepsilon}\right)^{-\frac{1}{\varphi}} \text{ with } \ell = 1.$$

Then from (35), it follows that

$$Y^{NA} = \left(\frac{\alpha}{\psi}\right)^{\frac{\alpha}{1-\alpha}} \frac{B(1)^{1-\varphi}}{B(0)^{1-\varphi}}.$$

Since $AMC = \left(\frac{\alpha}{\psi}\right)^{\frac{\alpha}{1-\alpha}} B(1)$ and since $C = Y - AMC$, the growth rate in consumption is given by

$$\begin{aligned}
\dot{C} &= \frac{Y}{C}\dot{Y} - \frac{AMC}{C}A\dot{M}C \\
&= \left[\kappa_c(1) - \frac{Y}{C}\varphi\left(\kappa_c(1) - \kappa_c(0)\right)\right]\dot{A}_c \\
&+ \left[\kappa_d(1) - \frac{Y}{C}\varphi\left(\kappa_d(1) - \kappa_d(0)\right)\right]\dot{A}_d \\
&- \left[\left(\frac{1}{1-\alpha} - \frac{Y}{C}\right)\kappa_d(1) + \frac{Y}{C}\varepsilon\left(\kappa_d(1) - \kappa_d(0)\right)\right]\frac{\tau}{1-\tau}\dot{\tau},
\end{aligned}$$

where $\kappa_c(\ell) \stackrel{\text{def}}{=} \frac{A_c^{-\varphi}}{A_c^{-\varphi} + A_d^{-\varphi}(1+\tau)^{\ell-\varepsilon}}$ and $\kappa_d(\ell) \stackrel{\text{def}}{=} 1 - \kappa_c(\ell)$.

Tedious manipulations then show that $\frac{C}{Y} - (1-\alpha) = \alpha\kappa_d(1)\frac{\tau}{1+\tau}$. And since $1 - \frac{\kappa_d(0)}{\kappa_d(1)} = \frac{\tau}{1+\tau}\kappa_c(0)$, the growth rate in C can be written as

$$\begin{aligned}
\dot{C} = &\left[\kappa_c(1) - \frac{Y}{C}\varphi\left(\kappa_c(1) - \kappa_c(0)\right)\right]\dot{A}_c + \left[\kappa_d(1) - \frac{Y}{C}\varphi\left(\kappa_d(1) - \kappa_d(0)\right)\right]\dot{A}_d \\
&- \frac{Y}{C}\kappa_d(1)\left[\frac{\alpha}{1-\alpha}\kappa_d(1) + \varepsilon\kappa_c(0)\right]\frac{\tau}{1+\tau}\dot{\tau}.
\end{aligned}$$

Logarithmic differentiation of $\kappa_c(\ell)$ yields

$$\dot{\kappa}_c(\ell) = (-\varphi)\,\kappa_d(\ell)\left[\dot{A}_c - \dot{A}_d - \frac{\tau}{1-\tau}\dot{\tau}\right].$$

Thus when all researchers are allocated to the clean energy sector ($s_c = 1, s_d = 0$), $\dot{A}_c = \gamma\eta$ and $\dot{A}_d = 0$. Then $\kappa_c(\ell) \to 1$ and $\kappa_d(\ell) \to 0$. Since $\kappa_d(1) \to 0$, the term with $\dot{\tau}$ in the expression for $\dot{C}$ will vanish. Hence $\dot{C} \to \dot{A}_c = \gamma\eta$. On the other

[25] The corresponding values for 2002–2006, adopted by AABH were 1893.25 and 307.77, respectively. We double the world primary energy production figures to obtain the same reference sectoral marginal costs of production as in AABH (cf. MC_{d0} and MC_{c0} on p. 7 in Section 4).

hand, if all researchers would be allocated to the dirty energy sector, $\dot{A}_d = \gamma\eta$ and $\dot{A}_c = 0$. Then $\kappa_d(\ell) \to 1$ while $\kappa_c(\ell) \to 0$, and the long term growth rate in consumption is given by

$$\dot{C} = \gamma\eta - \frac{Y}{C}\frac{\alpha}{1-\alpha}\frac{\tau}{1+\tau}\dot{\tau}.$$

Since the tax rate has to grow to contain the environmental consequences of dirty energy production, the growth rate is less than in the clean R&D scenario. Thus when abatement is too expensive, it is optimal to switch to a clean R&D policy, reconciling long term growth with environmental concerns.

When the abatement technology is advanced enough, abatement becomes sufficiently cheap for cleaned dirty energy to be a viable alternative to renewable energy. When abatement takes place, the productivity index (22) can be written as

$$B^A = \left(A_c^{-\varphi} + A_{da}^{-\varphi}\right)^{-\frac{1}{\varphi}},$$

where

$$A_{da} \stackrel{\text{def}}{=} \left[A_d^{-(1-a)} + A_a^{-(1-\alpha)}\right]^{-\frac{1}{1-\alpha}}.$$

With full abatement, final good production is given by the first line in (29). Since the aggregate machine cost is also proportional to B^A (cf (25)) consumption will grow at the same rate as B^A. Maximizing the growth rate of consumption then amounts to maximizing the growth rate of B^A, which is given by

$$\dot{B}^A = \gamma\eta\left\{\chi s_c + (1-\chi)[(1-\zeta)s_d + \zeta s_a]\right\},$$

where

$$\chi \stackrel{\text{def}}{=} \frac{A_c^{-\varphi}}{A_c^{-\varphi} + A_{da}^{-\varphi}}, \text{ and } \zeta \stackrel{\text{def}}{=} \frac{A_a^{-(1-\alpha)}}{A_d^{-(1-\alpha)} + A_a^{-(1-\alpha)}}.$$

Logarithmic differentiation of χ gives

$$\dot{\chi} = (1-\dot{\chi})(-\varphi)[\gamma\eta s_c - \dot{A}_{da}(s_{da})],$$

where $s_{da} = 1 - s_c$ and $\dot{A}_{da}(s_{da})$ is the growth rate in efficiency parameter A_{da} when a fraction s_{da} of researchers are in an optimal way allocated to dirty energy and abatement technology.

Logarithmic differentiation of ζ gives

$$\dot{\zeta} = (1-\zeta)(1-\alpha)\gamma\eta[s_d - s_a].$$

The problem of maximizing the long term growth in consumption is then

$$\max_{s_c, s_{da}} \gamma\eta\chi s_c + (1-\chi)\dot{A}_{da}(s_{da})$$
$$\text{s.t. } s_c + s_{da} = 1,$$

where

$$\dot{A}_{da}(s_{da}) = \max_{s_d, s_a}\left((1-\zeta)s_d + \zeta s_a\right)\gamma\eta$$
$$\text{s.t. } s_d + s_a = s_{da}.$$

We first solve the second problem. If $s_d < \frac{s_{da}}{2} < s_a$, then $\dot{\zeta} < 0$, so that ζ converges to zero and $\dot{A}_{da}$ converges to $s_d\gamma\eta (< \frac{s_{da}}{2}\gamma\eta)$. On the other hand, if $s_d > \frac{s_{da}}{2} > s_a$ then $\dot{\zeta} > 0$, so that ζ will approach 1 and $\dot{A}_{da}$ converges to $s_a\gamma\eta (< \frac{s_{da}}{2}\gamma\eta)$. Finally, if $s_d = s_a = \frac{s_{da}}{2}$, then ζ remains constant and $\dot{A}_{da} = \frac{s_{da}}{2}\gamma\eta$. Therefore, the optimal R&D strategy is to allocate the fraction s_{da} evenly across the dirty energy sector and the abatement sector: $\dot{A}_{da}(s_{da}) = \frac{s_{da}}{2}\gamma\eta$.

Now we can solve the former problem:

$$\max_{0<s_c<1} \dot{B}^{FA} = \gamma\eta\chi s_c + (1-\chi)\frac{1-s_c}{2}\gamma\eta = \gamma\eta\left(\frac{3\chi-1}{2}s_c + \frac{1-\chi}{2}\right)$$

The growth in χ can now be written as

$$\begin{aligned}\dot{\chi} &= (1-\chi)(-\varphi)[\gamma\eta s_c - \frac{1-s_c}{2}\gamma\eta] \\ &= -\varphi(1-\chi)\gamma\eta\left(\frac{3}{2}s_c - \frac{1}{2}\right).\end{aligned}$$

Since $\varepsilon > 1$, $\varphi < 0$. If $s_c > \frac{1}{3}$, then $\dot{\chi} > 0$ so that χ will converge to 1 and $\dot{B}^{FA}$ converges to $\gamma\eta s_c$. If $s_c < \frac{1}{3}$, then $\dot{\chi} < 0$ so that χ will converge to 0 and $\dot{B}^{FA}$ to $\frac{1-s_c}{2}\gamma\eta$. Finally, if $s_c = \frac{1}{3}$ then χ remains constant and $\dot{B}^{FA}$ converges to $\frac{1}{3}\gamma\eta$. Therefore, consumption growth is a convex function of s_c ($\vee$-shaped), with two local maxima: (i) $s_c = 0$ and $s_d = s_a = \frac{1}{2}$ giving a LT growth rate of $\frac{\gamma\eta}{2}$, and (ii) $s_c = 1$ and $s_d = s_a = 0$ giving a LT growth rate of $\gamma\eta$. The faster growth rate in the last case goes at the cost of a higher LT temperature increase. So there is a trade-off. For $\varepsilon = 3$ and $\varphi = -\frac{4}{3}$, both policies are about equally good when $MC_a = .30$, as in Figure 4.