Journal of Global Economic Analysis, Volume 10 (2025), No. 1, pp. 106-174.

Endogenous Technological Change Adapted to the CGE Framework

BY JAN WITAJEWSKI-BALTVILKS[a b]

We introduce endogenous technological change in a multi-sector recursive dynamic Computable General Equilibrium (CGE) model. We consider the optimization problem faced by technology firms in choosing the optimal level of R&D intensity given market conditions. R&D intensity determines the number of innovations and the speed of technological change in each sector. In addition, firms can choose the direction of technological change; for instance, when designing, they may opt for less energy-intensive but more capital-intensive production when the price of energy increases. We also differentiate between local innovations whose outcomes are constrained by the global technology frontier and innovations that transcend this frontier. The incorporation of endogenous technological change has a significant impact on policy simulations. Once it is considered, the model predicts that climate policy (in the form of a carbon tax) has a significantly larger impact on emissions in the long-run than in the short run. Finally, we demonstrate how intertemporal knowledge spillovers lead to path dependency. The introduction of a carbon tax at an early stage induces early low-carbon R&D and early know-how accumulation, which in turn leads to higher productivity in low-carbon sectors and lower long-run costs of decarbonization when compared to the scenario of postponed carbon tax.

JEL codes: C68, O33, Q55

Keywords: Endogenous growth; Directed technological change; Inertia; Path dependency; Climate policy

1. Introduction

Recursive-dynamic Computable General Equilibrium (CGE) models are frequently employed to simulate the path of economic variables following policy interventions. The major strength of this type of model in such simulations is that the predictions are derived from changes in the optimal choices of economic actors;

[a] University of Warsaw, Faculty of Economic Sciences and The World Bank (e-mail: jwitajewski@wne.uw.edu.pl).

[b] I am very grateful to my colleagues from the macroeconomic modelling team at the World Bank and the participants of the 2023 GTAP Annual Conference, the editor and the anonymous reviewers for their valuable comments, suggestions and support

therefore, the predictions are always consistent with microeconomic theory. However, typically, these models' predictions are built on the questionable assumption that in their choices, actors must take technological variables as given; nor do they have an opportunity to influence the future characteristics of technologies. This paper is an attempt to remedy this weakness without compromising the fundamental strength of the CGE model: its microeconomic foundations.

We propose a framework that fits the typical structure of CGE models and that allows us to take into account that the speed of technological change and the factor intensity of technologies will be impacted by changing market conditions. At the same time, we preserve the microeconomics-based nature of CGE models by deriving all relations from the predictions of the optimization of profit-maximizing R&D firms. We also consider that the blueprints produced by those firms are a non-rival good. This is an important consideration, because it implies that the accumulation and trade of knowledge (contained in blueprints) will be distinct from the accumulation and trade of capital goods.

From a methodological perspective, our study contributes to the literature by integrating insights from endogenous growth theory, evolutionary economics, and theories on induced and directed technological change, and adapting them to the CGE framework. Following the endogenous growth literature (Romer (1990), Romer (1989), Grossman and Helpman (1991), Aghion and Howitt (1992)), we introduce monopolistic competition between owners of blueprints, quality ladder progress (every new blueprint allows for an improvement in the productivity of production), and a business-stealing effect (the owner of a blueprint receives profit only until another firm develops a superior blueprint). Based on the literature on evolutionary economics (Aghion (2002), Goldfarb (2011), Young (1993)), we allow for the possibility that in all sectors innovation and progress are bounded by the progress of global general-purpose technology (GPT). Following the literature on induced technological change (Kennedy (1964),Samuelson (1965), Drandakis and Phelps (1966)) and directed technological change (Acemoglu (1998), Acemoglu et al. (2015), Acemoglu et al. (2012), Aghion et al. (2016)), we consider that technology firms may choose not only the level of their R&D effort (which determines the number of innovations and the speed of technological change) but also the factor intensity of new technologies (the direction of technological change).

We also propose some simplifications and adjustments to adapt the aforementioned approaches to the modeling constraints of a standard country-level multi-sector recursive-dynamic CGE model. We assume that decisions regarding R&D activity and technological progress take place in each sector of the economy separately. We also assume the exogeneity of GPT progress which bounds progress in individual sectors. This assumption is based on the premise that the progress of the GPT is likely to be global and independent of country-level policy interventions. We also assume that firms are forward-looking in the sense that they anticipate future profits from innovations. However, we assume that they are naive in the sense that they believe that the critical parameters in their optimization problem – interest rate, growth of sectoral revenue and innovation rate – will remain constant in the future.

We assume that R&D effort requires the use of investment goods, and that the productivity of R&D activities decreases over time. This assumption is necessary to ensure the existence of a balanced growth path. We also propose and discuss approximations that significantly reduce the number of equations in the model, thereby improving the model's tractability and reducing the computational time required.

The key reason for incorporating of endogenous technological change in CGE analysis is to improve the dynamic predictions of these models. The immediate response of the economy to policy intervention will differ from its response in subsequent periods when actors can choose the characteristics of future technologies. For instance, after an increase in the energy price (induced, for example, by a carbon tax), firms will initially have limited opportunities to replace energy with other factors of production, for example, capital. However, as time passes, they will devise several ways to increase the energy efficiency of their production. As a result, the drop in energy demand will be greater in the long run than in the short run, and the long- and short-run price elasticity of demand will also differ. Another example is the impact of carbon tax on the deployment of renewable energy sources (RES). An initial increase in demand for RES will incentivize technology firms to invest more R&D effort in that sector, thereby generating more future innovations, a gradual reduction in the cost of manufacturing and RES prices, and thus, a further increase in demand. In this case, an initial increase in the deployment of renewables would be gradually amplified in subsequent periods.

In this paper, we demonstrate the use of CGE model with endogenous technology change to assess the macroeconomic impacts of climate policy (in the form of a hypothetical carbon tax) in India. The model has three features that, according to Grubb et al. (2021), are essential for the evaluation of the dynamic effects of a carbon tax: induced technological change, inertia and path dependency. The simulation demonstrates that a carbon tax induces greater R&D effort in low-carbon sectors, representing induced technological change. The deployment of RES will increase gradually, with a greater response of emissions in the long run than in the short run, representing inertia. Finally, the earlier introduction of a carbon tax reduces the costs of climate policy in the latter period, representing path dependency.

2. Endogenous technological change in climate policy analysis

The role of technological progress in the low-carbon transition has motivated the modeling community to endogenize technological change in the economic analysis of transition. In the large-scale numerical models, two approaches were applied. The first approach Stehfest et al. (2014), Kriegler et al. (2017), Bosetti et al. (2009)) involved the incorporation of learning curves, which are simple log-linear relations between the costs of a given technology and the cumulative installed capacity of that technology. The microfoundation of this relation (first proposed by Arrow (1962)) requires the assumption that all productivity improvement is a byproduct of production; there is no intentional R&D effort motivated by the expectation of profit from innovations. Some modelers (e.g. Klaassen et al. (2005), Wiebe and Lutz (2016), Kittner et al. (2017)) addressed this drawback by introducing a two-factor learning curve, where the cost of technology is a log-linear function of cumulative installed capacity and cumulative investment in R&D (or patents). However, to the best of our knowledge, this log-linear relationship has never been derived from a microfounded model with an explicit market for innovations.

The second approach, which is employed in CGE models and some hybrid models, entails the incorporation of knowledge stock that is built up through R&D investment (Popp (2004); Bosetti et al. (2009); Sue Wing (2001); Wang et al. (2009)). Knowledge in these models is treated in the same way as capital and is subject to diminishing returns to scale at the regional level.

These models assume that there is a representative plant that combines knowledge with rival inputs using CES production function. In addition, the models implicitly assume that the central planner chooses to use o n e representative plant. However, when knowledge is a non-rival good, this assumption cannot be justified: choosing to use only one plant would be suboptimal. If the central planner were to divide the available rival input and use it in many plants, all characterized by the same CES production function, the total output would be higher than that of one plant using all rival input. This is because the rival inputs in CES production are subject to diminishing returns and because knowledge, as a non-rival good, does not have to be split across plants. Romer (1990) presents a more general argument: he shows formally that if a non-rival input has productive value, then output cannot be a constant-returns-to-scale function of all its inputs taken together.

The literature on Directed Technological Change (henceforth DTC; Acemoglu et al. (2012), Aghion et al. (2016), Hemous (2016); van den Bijgaart (2017)) offers numerous small-scale models that endogenize technology change with a Romer-style setup, which we adopt also in our study. This approach is fully microfounded and it allows for both, intentional R&D investment and path-dependency. Our study departs from DTC literature in two ways. Firstly, while DTC models typically assume a stylized production function (usually with two factors of production: capital and labor and two intermediate inputs: dirty and clean), our framework is consistent with any nested CES production function of a large-scale multi-sector computable general equilibrium model. Secondly, our framework is adjusted to the context of country-level analysis. For instance, it allows for the existence of a global technological frontier that constrains productivity improvement at the country level.

3. The microfounded model

In this section we describe the microfoundations of the endogenous technological change model and derive equations that can be incorporated into a CGE model. The final list of implementable equations is provided in Appendix A2, while their detailed mathematical derivations are provided in Appendix A3.

3.1 Demand for varieties

We assume that output in sector a comprises a continuum of varieties that are imperfect substitutes:

$$Q_{a,t} = \left(\int_{\Omega_{a,t}} \left(\lambda_a^{top} \left(u \right) q_{a,t} \left(u \right) \right)^{\frac{\sigma_a^{var}}{\sigma_a^{var}-1}} du \right)^{\frac{\sigma_a^{var}-1}{\sigma_a^{var}}} \tag{1}$$

where $\Omega_{a,t}$ is a set of varieties available in sector a at time t, $q_{a,t}\left(u\right)$ is the quantity of variety u and $\lambda_a^{top}\left(u\right)$ is the productivity parameter for variety u.

Following Young (1998), we distinguish between vertical and horizontal technological progress. With vertical innovation, technology firms develop a new and superior blueprint for an existing variety; we refer to this using the label '*improvement*'. In horizontal innovation, technology firms develop a variety that did not previously exist; we label this '*invention*'.

Let $IMP_{a,t}$ denote the improvement rate, i.e. the proportion of varieties existing at time $t-1$ that were improved at time t, and $INV_{a,t}$ denote the invention rate, that is, the number of newly invented varieties relative to the total number of varieties produced at time $t-1$. When $\sigma_a^{var} > 1$, there are new varieties added through invention and the total number of varieties increases. When $\sigma_a^{var} < 1$, every invention replaces two varieties, and the total number of varieties decreases.[1]. In this section we will consider only the case of $\sigma_a^{var} > 1$ to simplify the exposition of the model.[2]

We will define the innovation rate as the total number of new blueprints relative to the total number of varieties produced at time $t-1$, $INN_{a,t} = IMP_{a,t} + INV_{a,t}$.

Varieties that are improved or invented at the same point in time are identical. Therefore, in a discrete-time setting, these varieties can be grouped into cohorts. Throughout this paper, we will use the index τ to denote the time at which a variety was invented or last improved, and we will use t for any other reference to time. For example, $p_{\tau,t}$ represents the price at time t of varieties that were invented or last improved at time τ.

[1] For example, when $\sigma_a^{var} = 0$ all varieties are essential inputs in the production of the sector's output. Inventions reduce the number of essential inputs and thereby reduce the cost of sectoral production.

[2] In Appendix A1 we present a more general version of the model that allows for the consideration of the case of $\sigma_a^{var} < 1$.

The production function, (1), then becomes:

$$Q_{a,t} = \left(\sum_{\tau=0}^{t} N_{a,\tau,t} \left(\lambda_{a,\tau}^{top} q_{a,\tau,t}^{top} \right)^{\frac{\sigma_a^{var}}{\sigma_a^{var}-1}} \right)^{\frac{\sigma_a^{var}-1}{\sigma_a^{var}}} \tag{2}$$

where $q_{a,t,\tau}^{top}$ denotes the output of a variety representative for cohort τ, $N_{a,\tau,t}$ with $\tau > 0$ being the number of varieties in cohort τ and $N_{a,0,t}$ the number of varieties that were invented or improved before the first year of the simulation. We use $M_{a,t}$ to denote the total number of varieties at time t, $M_{a,t} = \sum_{\tau=0}^{t} N_{a,\tau,t}$.

As is clear below, technology firms that choose the variety they want to improve are indifferent between varieties. Hence, the probability that a given variety is improved is the same in every cohort and equal to $IMP_{a,t}$. Thus, over time, $N_{a,\tau,t}$ with $\tau > 0$ evolves according to the equation:

$$N_{a,\tau,t} = \begin{cases} INN_{a,t} M_{a,t-1} & \text{for } \tau = t \\ \left(1 - IMP_{a,t}\right) N_{a,\tau,t-1} & \text{for } \tau < t \end{cases} \tag{3}$$

Similarly, $N_{a,0,t}$ evolves as follows: $N_{a,0,t} = \left(1 - IMP_{a,t}\right) N_{a,0,t-1}$ where $N_{a,0,0}$ is normalized to unity.

Given the dynamics described above, the total number of processes evolves according to

$$M_{a,t} = \left(1 + INV_{a,t}\right) M_{a,t-1} \tag{4}$$

The output, $Q_{a,t}$, is assembled by a competitive firm. Its cost minimization implies that the demand for the representative variety in cohort τ is given by

$$q_{a,\tau,t}^{top} = \left(\lambda_{a,\tau}^{top} \right)^{\sigma_a^{var}-1} \left(\frac{P_{a,t}}{\hat{p}_{a,\tau,t}} \right)^{\sigma_a^{var}} Q_{a,t} \tag{5}$$

where $P_{a,t}$ is the price of output in sector a and $\hat{p}_{a,\tau,t}$ is the price of the variety of cohort τ.

The unit cost and price of output $Q_{a,t}$ are given by

$$P_{a,t} = \left(\sum_{\tau=0}^{t} N_{a,\tau,t} \left(\frac{\hat{p}_{a,\tau,t}}{\lambda_{a,\tau}^{top}} \right)^{1-\sigma_a^{var}} \right)^{\frac{1}{1-\sigma_a^{var}}} \tag{6}$$

3.2 Price of varieties

Varieties are produced by firms functioning under conditions of monopolistic competition. Those firms purchase a blueprint that allows them to manufacture the variety. Blueprints are excludable; that is, a firm with a blueprint has some monopoly power. In addition, we assume that the blueprint can be imitated costlessly, albeit imperfectly. If the blueprint is imitated, an imitating firm can produce the particular variety at a productivity lower than that allowed by the original blueprint by a factor of $\frac{1}{1+\gamma}$. This implies that the costs of the imitator are $(1+\gamma)$ times the marginal costs of a monopolist who owns the blueprint, and the markup charged by the monopolist cannot be higher than γ. We assume that $\gamma < g$, where g is the improvement of productivity between the two periods[3] and that for all sectors a, the elasticity of substitution between varieties, σ_a^{var}, is sufficiently low to ensure that the monopolists have no incentive to set a markup lower than γ.[4] The two assumptions imply that the markup chosen by the firm is equal to γ.

$$\hat{p}_{a,\tau,t} = (1+\gamma)\, p_{a,\tau,t}^{top} \tag{7}$$

where $p_{a,\tau,t}^{top}$ is the cost of manufacturing the representative variety in the cohort τ.

3.3 Manufacturing of varieties

To simplify the exposition in this section, we consider a setup with single-nest production function. In Appendices 2 and 3, we extend this setup to a more general framework consistent with any nested CES production function.

We assume that the output of each variety, $q_{a,\tau,t}^{top}$, is produced by combining multiple inputs, as shown below:

$$q_{a,\tau,t}^{top} = \left(\sum_{ch} \left(\alpha_{a,\tau}^{ch}\right)^{\frac{1}{\sigma_a^{top}}} \left(\lambda_{a,\tau}^{ch} q_{a,\tau,t}^{ch}\right)^{\frac{\sigma_a^{top}-1}{\sigma_a^{top}}} \right)^{\frac{\sigma_a^{top}}{\sigma_a^{top}-1}}$$

where $q_{a,\tau,t}^{top}$ represents the use of input ch, and $\alpha_{a,\tau}^{ch}$ is the technological parameter determining the intensity of input ch for variety τ.

The demand for the child input, $q_{a,\tau,t}^{ch}$, as a function of the demand for the variety, $q_{a,\tau,t}^{top}$, can be derived from the cost-minimization problem of a monopolist manufacturing the variety, as follows:

$$q_{a,\tau,t}^{ch} = \alpha_{a,\tau}^{ch} \left(\frac{p_{a,\tau,t}^{ch}}{p_{a,\tau,t}^{top}} \right)^{-\sigma_a^{top}} \left(\lambda_{a,\tau}^{ch}\right)^{\sigma_a^{top}-1} q_{a,\tau,t}^{top} \tag{8}$$

[3] This assumption ensures that the owner of the previous version of the blueprint has no incentive to continue producing under the old blueprint - they would rather imitate.
[4] Otherwise, if σ_a^{var} is high, and if firms take others' prices as given (as in the Bertrand price competition described in Markusen (2023)) the markup is determined by the conditions for the optimal choice of price in the profit maximization problem and, given the demand function (5) and given that the number of varieties is infinite, it is equal to $\frac{\sigma_a^{var}}{\sigma_a^{var}-1}$ (Dixit and Stiglitz (1977)).

where p^{ch} denotes the price of input, ch, and p^{top} denotes the unit cost of manufacturing the output of variety. The latter can be derived from the cost-minimization problem as follows:

$$p^{top}_{a,\tau,t} = \left(\sum_{ch \in nest(top)} \alpha^{ch}_{a,\tau} \left(\frac{p^{ch}_{a,\tau,t}}{\lambda^{ch}_{a,\tau}} \right)^{1-\sigma^{top}_a} \right)^{\frac{1}{1-\sigma^{top}_a}} \tag{9}$$

3.4 The Value of a Blueprint

We assume free entry into the market for varieties, meaning that any firm can commission a new blueprint. Consequently, the price of a blueprint is equal to its value, which is equal to the discounted stream of future cash flows. The cash flow in each period is the operating surplus, i.e., the difference between the firm's revenue and the manufacturing costs: $\left(\hat{p}_{a,\tau,t} - p^{top}_{a,\tau,t}\right) q^{top}_{a,\tau,t} = \gamma p^{top}_{a,\tau,t} q^{top}_{a,\tau,t}$, where γ is the markup rate (as defined in equation (7)). The discount factor is determined by the future interest rate, i^{RD}_t, and the future improvement rate, $IMP_{a,t}$. The improvement rate is incorporated into the discount factor to account for the business-stealing effect: a successful innovator will receive cash flow from the blueprint only until another successful innovator produces a superior blueprint that enhances the technology of the variety. At time t, this occurs with probability $IMP_{a,t}$. Thus, the value at time t of a blueprint invented at time τ is:

$$\sum_{T=t}^{\infty} \left(\prod_{s=t+1}^{T} \frac{1 - IMP_{a,s}}{1 + i^{RD}_s} \right) \gamma p^{top}_{a,t,T} q^{top}_{a,t,T}$$

We assume firms take into account what a blueprint will allow in terms of future production but that they are also myopic, in the sense that they assume that (i) Revenue will grow at an exogenous rate, denoted as $g^{exp}_{a,t}$ (the 'expected growth rate'), (ii) the probability that their blueprint will be replaced in the future is constant and given by the expected improvement rate for the next period, $IMP^{exp}_{a,t} \equiv IMP_{a,t+1}$, (iii) the interest rate is exogenous and constant, denoted by $i^{RD}_t = i^{RD}_0$. Additionally, we assume that blueprints become available for use one period after they are commissioned by monopolists. This implies that there is no cash flow generated at the time of commissioning. Consequently, the value of a blueprint commissioned at time t is given by:

$$\frac{\gamma p^{top}_{a,t+1,t+1} q^{top}_{a,t+1,t+1}}{1 + i^{RD}_0} + \frac{\left(1 - IMP^{exp}_{a,t}\right)}{\left(1 + i^{RD}_0\right)^2} \gamma p^{top}_{a,t+1,t+2} q^{top}_{a,t+1,t+2} + \ldots =$$

$$= \frac{\gamma p^{top}_{a,t,t} q^{top}_{a,t,t}}{\left(1 + i^{RD}\right) / \left(1 + g^{exp}_{a,t}\right) - \left(1 - IMP^{exp}_{a,t}\right)} \tag{10}$$

Due of the myopic nature of firms, there may be a discrepancy between expected profits and actual profits. Hence, we introduce an 'insurance' actor, who bears the difference between the two. In the MANAGE model, which is used for the numerical illustration in Section 5, this difference is assumed to be covered by a sector of 'enterprises'. This sector acts as an intermediary, collecting capital rents from firms in all sectors and transferring them to households. The difference between actual and expected R&D profit - whether positive or negative - is then added to these capital transfers.

3.5 Creation of blueprints

New blueprints are generated by technology firms that conduct R&D. The Poisson arrival rate of original blueprints, that is, the quantity of new blueprints in one period, is given by

$$R_{a,t}^{\phi^{RD}} \Gamma M_{a,t} \tag{11}$$

where $R_{a,t}$ is the total research intensity within activity a, and Γ represents the productivity of research, which we assume to be constant. The term $M_{a,t}$ captures the intertemporal spillover: the productivity of researchers increases with the number of past inventions. Parameter ϕ^{RD} reflects the size of the stepping-on-toes effect (Jones (1995), Jones and Williams (2000)): when $\phi^{RD} < 1$ greater research effort is associated with lower research productivity, because of the higher risk of researchers duplicating ideas. The value of parameter ϕ^{RD} can be, but does not have to be strictly smaller than one - it can also take the value of unity.

Both, the spillover effect (captured by the term $M_{a,t}$) and the stepping-on-toes effect (captured by the term $\frac{R_{a,t}^{\phi^{RD}}}{R_{a,t}}$) are the externalities: an individual firm that considers investing an additional unit of R&D intensity does not take into account that it will increase future R&D productivity (through an increase in $M_{a,t}$) or that it decreases the productivity of current researchers (through a decrease in $\frac{R_{a,t}^{\phi^{RD}}}{R_{a,t}}$). In other words, from the firm's perspective one additional unit of R&D activity increases the Poisson arrival rate by $R_{a,t}^{\phi^{RD}-1} \Gamma M_{a,t}$ where $R_{a,t}$ and $M_{a,t}$ are taken as given.

3.6 Zero profit condition and equilibrium innovation rate

The equilibrium level of R&D intensity in sector a is determined by the zero-profit condition for technology firms. Given the value of each successful blueprint and the Poisson arrival rate described above, the condition reads:

$$p_{RD,t}\Phi_{a,t} = R_{a,t}^{-\left(1-\phi^{RD}\right)}\Gamma\frac{\gamma p_{a,t,t}^{top}q_{a,t,t}^{top}}{\left(1+i^{RD}\right)\,/\,\left(1+g_{a,t}^{exp}\right)-\left(1-IMP_{a,t}^{exp}\right)} \tag{12}$$

where $p_{RD,t}\Phi_{a,t}$ is the unit cost of R&D effort.

Some blueprints will be dedicated to improvements and some to the inventions of new varieties. Firms are indifferent between improvements and inventions because all improved and new varieties are characterized by exactly the same productivity and, therefore, offer the same revenue and profit streams. For the same reason, technology firms will be indifferent between the varieties they can choose to improve. We assume that the share of blueprints dedicated to improvements and to inventions is constant:

$$\rho = \frac{INV_{a,t}}{INN_{a,t}} = \frac{INV_{a,t}}{INV_{a,t}+IMP_{a,t}} \tag{13}$$

Parameter ρ determines how much of technological progress can be attributed to vertical technological change (through improvements), which, in our setup, is bounded by the exogenous frontier described below (see section 3.7), and how much can be attributed to horizontal technological change (through inventions), which, in our setup, is unbounded and associated with path-dependency. We discuss the importance of this distinction in Section 4.

As noted above, new blueprints must be commissioned one period earlier. This means that their number at time t depends on the choice of R&D activity at time $t-1$. In combination with equation (3) and (11), this implies

$$M_{a,t-1}INN_{a,t} = R_{a,t-1}^{\phi^{RD}}M_{a,t-1}\Gamma \tag{14}$$

Given equation (13), we can then express the equilibrium innovation rate as:

$$INN_{a,t}^{exp} = R_{a,t}^{\phi^{RD}}\Gamma \tag{15}$$

where

$$INN_{a,t}^{exp} = INN_{a,t+1} \tag{16}$$

3.7 Productivity path

We assume that in each period, the productivity of each variety in each sector is bounded by the state of the global technological platform, which we label as the frontier. The technological platform could be interpreted as a set of GPTs, such as ICT and automatization. The frontier reflects the state of advancement of these technologies. In our model the position of the frontier at time t determines the highest potential productivity attainable for a variety at time t. The position and speed of advancement of the frontier can be sector-specific to reflect that the applicability of GPTs can be asymmetric across sectors.

We assume that the improvement of an existing variety allows firms to raise the variety's productivity to the level of the frontier and remain there until the next improvement. Similarly, when innovation results in the invention of a new variety, the productivity of that variety will be set at the level of the frontier at the time of the invention and remains at that level until it is improved by another firm. Hence, the productivity of all varieties that were improved or invented at time t (i.e. belonging to the cohort $\tau = t$) is:

$$\lambda_{a,t}^{top} = frontier_t \tag{17}$$

We assume that the frontier grows at the exogenous rate g:

$$frontier_t = (1+g)\, frontier_{t-1} \tag{18}$$

We set the value of $frontier_0$ at a level that ensures that, at $t = 1$, average productivity grows at the same rate as the frontier (see, Appendix A4).

The cost of production of sectoral output $P_{a,t}$ also depends on the number of varieties being produced (see equation (6)). Inventions always result in that cost dropping. When $\sigma_a^{var} \geq 1$, this is achieved via an increase in the total number of varieties; when $\sigma_a^{var} < 1$, it is achieved via a decrease in the number of varieties (every invention replaces two existing varieties). In Section 4.2 we compare and further discuss improvement- and invention-driven growth.

3.8 Directed search

For projects commissioned at time $t-1$ that generate blueprints at time t, technology firms can adjust the design to the prices observed at time t. Specifically, technology firms choose $\alpha_{a,\tau}^{ch}$ to minimize the expected costs of manufacturing varieties designed at time t:

$$\min_{\alpha_{a,t}^{ch}} \left(\sum_{ch \in nest(par)} \alpha_{a,t}^{ch} \left(\frac{p_{a,t,t}^{ch}}{\lambda_{a,t}^{ch}} \right)^{1-\sigma_a^{top}} \right)^{\frac{1}{1-\sigma_a^{top}}}$$

for every nest par, subject to the possibility constraint (see, Caselli and Coleman II (2006), Growiec (2008)):

$$\sum_{ch \in nest(par)} \left(\frac{\alpha_{a,t}^{ch}}{\beta_{a,ch}} \right)^{\omega_a^{top}} = 1 \tag{19}$$

where $\beta_{a,ch}$ and ω_a are parameters of the possibility constraint.[5]. To ensure that

[5] Throughout the paper we assume $\omega_a^{top} > \sigma_a^{top} - 1$ to avoid corner solution – see proposition 1 in Caselli and Coleman II (2006)

the technology firm has an incentive to minimize units of production, we assume that the imitator can also costlessly choose parameters $\alpha^{ch}_{a,t}$.

The first order conditions for this problem determine optimal choices of $\alpha^{ch}_{a,\tau}$ for cohort τ:

$$\alpha^{ch}_{a,\tau} = \left(\frac{p^{ch}_{a,\tau,\tau} / \lambda^{ch}_{a,\tau}}{p^{top}_{a,\tau,\tau}} \right)^{\sigma^{top}_a - \tilde{\sigma}^{top}_a} (\beta_{a,ch})^{\frac{1-\tilde{\sigma}^{top}_a}{1-\sigma^{top}_a}} \tag{20}$$

where $\tilde{\sigma}^{top}_a = \frac{\omega^{top}_a \sigma^{top}_a - 1}{\omega^{top} - 1}$. It can be shown that while σ^{top}_a governs the elasticity of demand for energy in the short-run, $\tilde{\sigma}^{top}_a$ governs the elasticity of demand in the long-run.

Again, in this formulation, we assume technology firms are myopic, believing that factor prices will remain constant.

3.9 Short and Long version of the model

The equations implemented in the CGE model are derived directly from the solution of the microfounded model presented in Sections 2.1-2.8, with some exceptions involving the use of approximations. The approximations are optional, hence there are two versions of the model: one in which we apply the approximations (the short version) and one in which we do not (the long version). The list of equations implemented in the short version is presented in appendix A2. That appendix also contains a discussion of the approximations. The list of equations implemented in the long version, together with detailed mathematical derivations is presented in Appendix A3. Appendix A6 presents a comparison of the carbon tax simulation results generated with the short version of the model (with approximations) and the long version of the model (without approximations).

4. The key features of the framework

4.1 Speed and direction of technological change

As outlined in the introduction, the primary purpose of our study is to improve the dynamic predictions of a CGE model by endogenizing technological parameters. We wish to describe an economic system where changes in market conditions that favor some technologies will result in their accelerated development and diffusion and the gradual restructuring of the economy. There are two ways in which the model we propose generates such dynamics: by allowing prices to influence (i) productivity growth and (ii) factor intensities.

The first of these is relatively straightforward. Any change in market conditions that increases the demand for the sector's output will result in an increase in its revenues. In our model, this will be associated with an expectation of higher profits and a higher value of potential innovations. This will, in turn, lead to a larger investment in R&D activity in this sector and a gradual increase in the portion of varieties in the vicinity of the frontier. This will lead to an acceleration of average productivity growth in the sector, a reduction in unit costs and a further increase in demand.

Perhaps the most obvious example of such a development is the impact of a carbon tax on the diffusion of RES: initially, the tax will lead to a single jump in the use of RES. In the long run, however, the initial increase in demand will induce efforts to innovate, perhaps in the form of efficient production processes or efficient value chains. As a consequence, the unit cost of RES will gradually fall, amplifying the initial impact of the tax.

However, this feature of the model can only be exploited when the sectors included are sufficiently disaggregated. The endogenous development of a single technology can only be captured if this technology is explicitly named in the model as one of the sectors of the economy. RES technology, which we use in the example above, will typically satisfy this condition; CGE models often include the key electricity generation technologies as separate sectors (Faehn et al. (2020), Peters (2016)). However, finding other examples is difficult. Most CGE models do not distinguish between alternative transport technologies or alternative production methods in heavy industry.

The second feature of our model, the adjustment of factor intensities, allows technological change to be endogenous even when alternative technologies are not explicitly named. After a change in prices of factors of production, firms will search for ways of producing their output using less of the factor that has become relatively more expensive.

An example that illustrates this would be an evolution of energy demand by the transport sector after an increase in the price of oil. We expect that, as time passes, more and more firms will consider how to deliver their services with less reliance on oil and more reliance on other inputs, such as electricity or capital.

Firms' choices are constrained by the possibilities of the global technological frontier (see equation (19)) and involve trade-offs: the choice of a lower intensity, $\alpha_{a,t}^{ch}$, for some inputs, requires an increase in intensity for other inputs. For instance, in the context of an increase in the price of energy, the trade-off is that, for new varieties, energy will be used more efficiently, and capital will be used less efficiently, compared to the situation absent the price shock. Importantly, this does not necessarily mean that, in new varieties, capital will be used less efficiently than in old varieties.

To understand this, note that the total efficiency in using capital (measured, for instance, by $\frac{\lambda_{a,\tau}^{top} q_{a,\tau.t}^{top}}{q_{a,\tau.t}^{capital}}$) is determined not only by $\alpha_{a,t}^{capital}$, which may decrease compared to $\alpha_{a,t-1}^{capital}$, but also by the productivity parameter, $\lambda_{a,t}^{top}$, which increases over time. As a result, unless the price change is very large, the total efficiency of use, will be higher in new varieties than in old varieties. In other words, the efficiency of all inputs will benefit from technological progress, although the rate of improvement will differ between inputs.

Note also that at the sectoral level, the change in the demand for factors will be gradual; in our setup firms can adjust the intensities for new varieties only. If the price increase for input *ch* takes place at time T, only varieties invented or improved after that time will be characterized by the low use intensity in respect of that input; the pre-shock varieties will be characterized by high intensity of use. However, the share of pre-shock varieties will decrease over time as a result of continued improvements and inventions. Hence, the intensity of use of the input at the sectoral level will gradually decrease. The speed of the decay of pre-shock varieties depends on the arrival rate of the new blueprints.

Finally, note that the arrival rate is also endogenous and will respond to policies. This can be illustrated by two examples. First, consider a shock that increases the demand for rail transport. The sector's revenue will increase, which will push up the innovation rate. A higher innovation rate will then increase the share of varieties with adjusted intensities. In this case, at the sectoral level, adjustment of input demand will be faster than if the innovation rate is constant.

The second example is the impact of a carbon tax on the cement sector. On the one hand, the tax will incentivize a more energy-saving choice of technology for new blueprints. On the other hand, the tax will increase the total cost of production in the sector, which will be reflected in a higher output price. If demand for the sector is inelastic, which is likely to be the case, the sector's revenue will increase. As in the previous example, this will increase the value of the blueprints, speed up innovation and speed up replacement of pre-shock (energy-intensive) varieties with post-shock (energy-saving) varieties.

4.2 Vertical (bounded) and horizontal (unbounded) technological change

Another feature of the model is the existence of a frontier that bounds the productivity of each variety at every given moment. The frontier is exogenous in the model. It can be interpreted as the maximum productivity of varieties set by the state of GPT (such as ICT or automatization), which is independent of market conditions or country-level policies. In our setup, every successful attempt to improve a variety brings its productivity $\lambda_{a,t}^{top}$ up to the level of the frontier. As a result, all improved varieties in a given sector exhibit the same productivity immediately after the improvement. In this respect, the model departs from the standard quality-ladder setup, where an innovation increases productivity by a fixed factor relative to the productivity of the variety preceding the innovation.

The introduction of a frontier is motivated by the local geographical scope of most CGE analyses. Typical CGE simulations involve policies or shocks that impact only a single country or a limited group of countries. The situations in which such policies could lead to, or prevent, major technological breakthroughs and have a significant impact on global GPT progress, are rare.

Policies or a change in local market conditions will influence a specific type of innovation, which we label "improvements" and which are sometimes referred to in the literature as secondary innovations (Aghion (2002)). Secondary innovations allow local firms to catch up with the global technological frontier. In the real world these could take the form of adopting new ideas invented by the leading technology countries, learning and exploitation of new possibilities offered by GPT progress, or adaptation of new solutions to local conditions, for example, local consumer preferences or geographical constraints.

To understand the role of improvements in our model and the implications of the presence of the frontier, consider the evolution of average productivity across varieties: $\overline{\lambda}_{a,t}^{top} = \sum_{\tau=0}^{t} \frac{N_{a,\tau,t}}{M_{a,t}} \lambda_{a,\tau}^{top}$. To simplify our argument, suppose that the number of varieties is constant (i.e., there are no inventions, $\rho = 0$) and normalized to unity. In this case, given equation (3), average productivity would evolve as follows:

$$\overline{\lambda}_{a,t}^{top} = IMP_{a,t} frontier_{a,t} + (1 - IMP_{a,t})\, \overline{\lambda}_{a,t-1}^{top} \tag{21}$$

The average productivity across varieties at every point in time will be a weighted sum of (i) the term that reflects the current state of the frontier and (ii) the average productivity from the previous period. The weights are determined by the sectoral improvement rate: a higher improvement rate results in more weight on the frontier.

The resulting dynamics are illustrated in the left panel of Figure 1, which shows the path of average productivity after a shock to the improvement rate. A shock that permanently increases the improvement rate will result in productivity growth that is faster than the growth of the frontier in the short-run (see the top left panel of Figure 1).

In the long-run, a permanent shock will result in a permanently lower distance to the frontier. However, the growth rate will return to the growth rate of the frontier. In fact, if improvement remains constant after the shock at the level IMP_a, the distance between the average productivity of the sector and the state of the frontier, $\frac{\overline{\lambda}_{a,t}^{top}}{frontier_{a,t}}$, will converge to a steady state given by $IMP_a \left(\frac{1+g_a}{g_a + IMP_a} \right)$, where g_a is the growth of the frontier. The setup prevents a situation in which a policy leads to a permanent increase in the growth of a sector, causing explosive behavior.

In the case of a shock that temporarily increases the improvement rate, average productivity growth will accelerate in the short run. After the shock is phase-out, it will decelerate and, in the long-run the distance between productivity and the frontier will be the same as before the shock (see the bottom left panel of Figure 1). In other words, temporary shocks do not have permanent consequences; there is no path-dependency.

Because in some applications the framework with bounded technological change may be inappropriate and path-dependency may be a desirable feature, we allow for another type of technological change that is unbounded. We do this by introducing the possibility that innovations result in 'inventions', that is, result in completely new varieties. In our model, invention is an outcome of R&D effort that is distinct from but perfectly correlated with improvement. The relation between the number of improvements and the number of inventions is governed by the parameter ρ. One possible interpretation of our setup is that when technology firms engage in R&D activity, there is some probability (given by $(1-\rho)\,R_{a,t}^{\phi^{RD}-1}\Gamma M_{a,t}$) that their unit of effort will lead to an improvement of existing varieties and some probability (given by $\rho R_{a,t}^{\phi^{RD}-1}\Gamma M_{a,t}$) that it will result in an invention of a completely new variety.

Growth driven by inventions has very different properties than growth driven by improvements. Our approach to modeling invention is very close to the one in Romer (1990). We assume that there is no bound on the total number of inventions. Moreover, we assume intertemporal spillover effects in equation (14), whereby a higher number of inventions in the past results in a higher productivity of researchers. In contrast to the case of improvements discussed above, the development of inventions will result in path dependency: a single temporary shock to the invention rate will result in permanent changes in average productivity.

This is illustrated in the right panels of Figure 1, where we consider the full model with productivity growth driven by both inventions and improvements ($1 > \rho > 0$). In this case, the permanent shock to the innovation rate increases the long-run growth of average productivity. The temporary shock increases the growth only in the short run; in the long run, the growth returns to the initial level. However, the distance to the frontier also remains smaller in the long-run, reflecting path-dependency.

5. Numerical illustration

In order to illustrate how endogenous technological change can influence CGE projections, we incorporate the equations described in Appendix A2 in the MANAGE model and run a series of simulations for climate-policy scenarios in India. We implement the short and long version of the model (the latter using the equations in Appendix A3).

5.1 MANAGE model

The MANAGE model is a CGE model developed by the World Bank to analyze country-level macroeconomic policies, such as carbon tax or fiscal reforms. The version built for India disaggregates the economy into 43 sectors calibrated using GTAP input-output data. Version 10 of the GTAP Data Base is described by Aguiar et al. (2019) and further details on construction of the database are provided by Aguiar et al. (2016). In addition the model uses GTAP-Power Data Base by Peters (2016), which extends the 'standard' GTAP Data Base (Aguiar et al. (2016) and

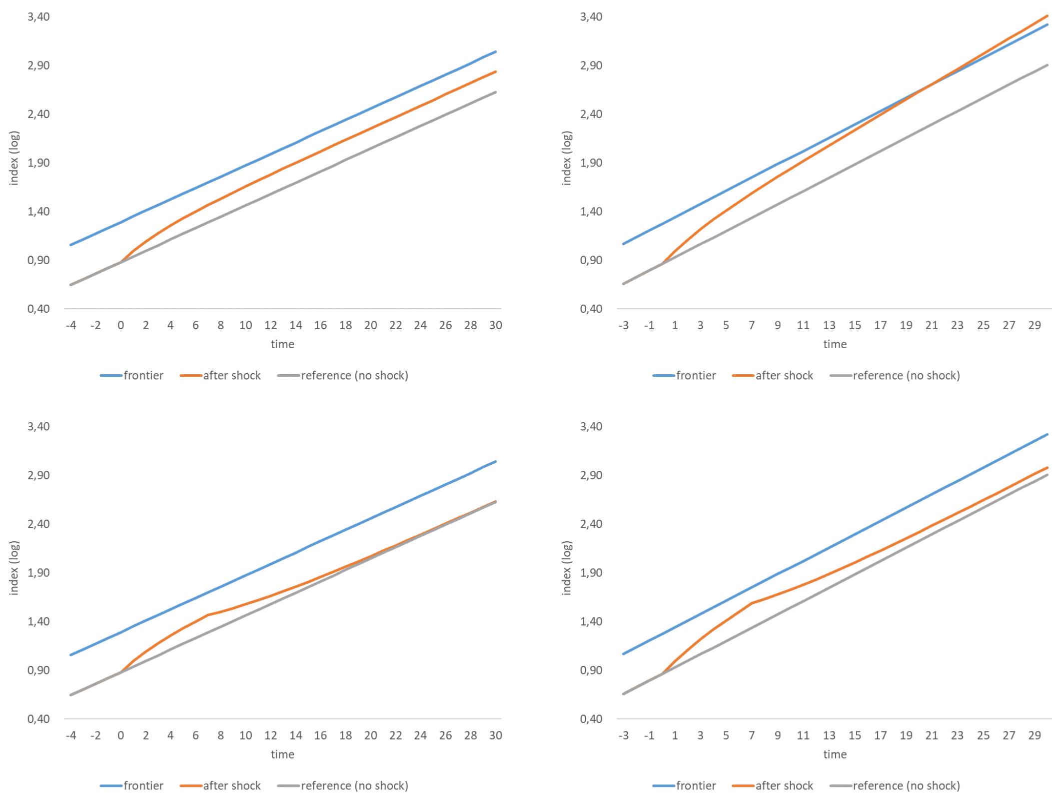

Figure 1. The impact of a change in the rate of arrival of blueprints on the path of efficiency of a sector.

Notes: The top panels show the impact of a permanent increase in the arrival rate, and the bottom panels show a temporary increase (of 7 periods). The left panels show the impact when inventions are switched off ($\rho = 0$), and only improvements (bounded technological change) are activated. The right panels show the impact when both inventions and improvements are activated. The efficiency of a sector is measured as the inverse of the cost of manufacturing the sectoral output ($\frac{1}{P_{a,t}}$). In each graph, the orange line shows the path of efficiency affected by the shock. The blue line is the hypothetical path if all varieties are at the frontier and the number of varieties grow at the initial rate. The gray line is the efficiency path if there are no shocks to the arrival of blueprints.

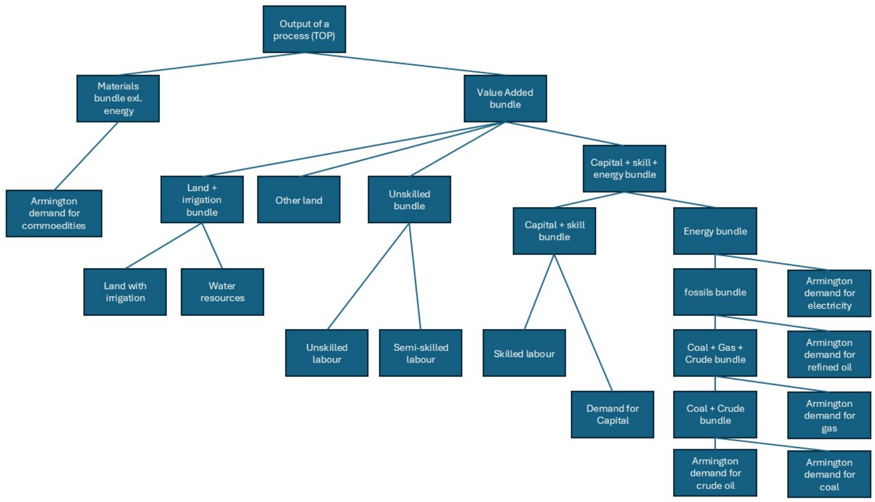

Figure 2. Production nests in the MANAGE model

Aguiar et al. (2019)) by disaggregating the electricity sector into 11 electricity generation technologies and transmission and distribution.

The model features a nested production function (see Figure 2). The most important nest, in the context of climate policy, is the one where energy is combined with a capital-skill composite. The latter is produced by combining skilled labor with the use of capital. In turn, energy is produced by combining electricity with a composite of fossil fuels. On the other side of the production function, capital-skill-energy is combined with unskilled labor and land to produce 'value-added' input. This is then combined with non-energy materials to produce the final output. The detailed model structure is described in World Bank (2024).

In the simulations, for all sectors, we endogenize the choice of share parameters for the CES nests that combine energy and the capital-skill composite. Essentially, this implies that firms are allowed to direct research efforts towards energy- or capital-saving change in production. In addition, we endogenize the speed of technological change (through endogenizing the innovation rate and R&D investment) in sectors that are carbon intensive (iron and steel and road transport) or related to electricity production (wind power generation, solar, hydro, coal, gas and nuclear).

5.2 Calibration

We start with the calibration of the directed search described in Section 3.8. In our numerical exercise, we introduce the distinction between the short- and long-run values of the elasticity of substitution for the nest that combines energy with the capital-skill composite.

For the short-run elasticity (parameter σ), Koetse et al. (2008) find the Morishima elasticity of substitution between energy and capital at the level of 0.4 for North America and 0.2 for Europe. We assume an elasticity of substitution between energy and the capital-skill composite of 0.3. For the long-run elasticity parameter

(parameter $\tilde{\sigma}$), Koetse et al. (2008) find the Morishima elasticity of substitution at 1.1 for North America and 0.8 for Europe. This suggests the elasticity should be close to unity. We set the value to 0.99. These values are applied in all sectors.

The baseline innovation rate (parameter $INN_{a,0}$) can be determined using historical records of the speed of input adjustment in response to price change. Given the climate-related context of our scenarios, we focus on the speed of adjustment in the use of energy. Maddala et al. (1997), who examine the dynamic response of energy demand to changes in energy price, suggest the coefficient on the lagged energy demand is in the range of 0.5 to 0.9, suggesting that $INN_{a,0}$ takes a value between 0.1 and 0.5. We pick the value of 0.3 for all sectors except those producing electricity (for these sectors, we choose $INN_{a,0} = 0.1$).[6]

Next, we calibrate the parameters that determine the behavior of the R&D sector (Sections 3.5-3.7). The markup value should match R&D expenditure relative to the value of production. In India, R&D expenditure constitutes approximately 0.7% of GDP, so we set the markup to 0.007. [7]

The responsiveness of the innovation rate to R&D expenditure, ϕ^{RD} can be derived from the elasticity of energy-related patents to changes in energy expenditure (assuming that patents are proportional to the total number of innovations and that R&D expenditure is proportional to sector revenue). Witajewski-Baltvilks et al. (2017) find that this elasticity takes a value between 0.5 and 1.1. We assume the elasticity takes the value of 0.7 for all sectors with an endogenous innovation rate.

In the model, the endogenous innovation rate can be activated for selection of sectors. This is achieved by setting $\phi^{RD} > 0$ for the sectors where innovation is to be activated, and $\phi^{RD} = 0$ for all other sectors. While the model theoretically allows the activation of the endogenous innovation rate for all sectors, doing so may increase the risk of numerical problems. Specifically, activating it in sectors that produce commodities with a high elasticity of demand can cause explosive behavior. In such cases, an increase in demand triggers an increase in innovation, which in turn causes prices to drop. This drop in prices further increases demand which again drives innovation.

We also assume a low elasticity of substitution between varieties, σ_a^{var} (set at the level of 0.3 for all sectors).

We assume that the frontier grows at a rate equal to the rate of long-run growth of the US economy, that is, 2%. The exception is the growth of the frontier in solar PV and in the wind sector, which we assume to be 4% per year.[8] Parameter ρ can be

[6] These sectors are relatively small. Higher values of $INN_{a,0}$ for these sectors cause numerical instability.

[7] data for the period 2012-2020 according to UNESCO (2023).

[8] NREL (2023) predicts a decline in the costs of these technologies of approx. 2% per year. If wages grow at the rate of 2%, the productivity of the sector must grow at approx. 4% per year

estimated using India's contribution to global technological progress. The global inflow of patents is approximately 3.5m per year;[9] India contributes to this with 77k patents per year. The global effort translates into 4% growth of the frontier for wind and solar, suggesting India's contribution translates into 4p.p.*77k/3.5m, which is equal to a 0.088 percentage-point contribution to the global annual growth of the global frontier. The productivity growth driven by domestic inventions is given by $\frac{\sigma_a^{var}-1}{\sigma_a^{var}} INV_{a,0} = \frac{\sigma_a^{var}-1}{\sigma_a^{var}} \rho * INN_{a,0}$ (see Equations (2) and (4)). This suggests that $\rho = 0.0038$. However, this simple calculation rests on the assumption that India's patents are as relevant to its economy as any foreign patent, which is unlikely. Instead, we assume that India's patents are six times more relevant for India than foreign patents and so we set $\rho = 0.022$. However, we also run sensitivity analysis with $\rho = 0.0038$ (see the results in Appendix A7).

5.3 Induced technological change and inertia

First, we use the model to illustrate how climate policy can induce carbon-saving R&D activity and to demonstrate the inertia in the economy's response. We examine the dynamics following the introduction of a permanent carbon tax set at 6 USD per tCO2. In all scenarios, the carbon tax revenue is allocated to increase government savings, leading to a rise in investment. We focus on reporting results for two sectors photovoltaic electricity generation (to showcase the impact on productivity and costs) and iron and steel (to demonstrate the shift toward energy-saving technologies).

For the simulation, we employ the long version of the model (with no approximations). The results for the short version (reported in Appendix A6) are qualitatively similar.

In the case of the photovoltaic sector, the carbon tax shock increases demand for output, resulting in higher revenue (top-left corner of Figure 3) and an increase in the value of new blueprints. This incentives an increase in R&D investment (of 30% in 2030; not reported in figures) and innovation rates (of 22%; top right corner of Figure 3). A higher innovation rate translates into more rapid productivity growth, reflected in reduced manufacturing costs (middle-left corner of Figure 3). Lower costs lead to a further increase in demand (middle-right corner of Figure 3), amplifying the initial impact and generating additional growth in R&D investment and innovation.

By 2040, the innovation rate would increase by about 30% compared to the BAU scenario. To put these figures into perspective, India's R&D sector generated an average of 60 patents per year between 2012-2015 (IRENA (2023)). Assuming this number remains constant in the BAU scenario, and that patents are proportional to the rate of innovation, the simulation suggests that the carbon tax would induce an additional 18 patents per year. While this increase could be significant from India's

[9] All data on patent counts refer to 2022 figures. The data were retrieved from WIPO (2023)

perspective, it would contribute only marginally to global trends. For comparison, the total number of PV patents worldwide increased from 25,000 in 2010 to 45,000 in 2022 (IRENA (2023)).

Regarding the iron and steel sector, the carbon tax motivates innovators to search for blueprints that reduce energy consumption, given an increase in its cost. This results in a lower energy-capital ratio for new varieties. However, at the aggregate level, as shown in the bottom-left corner of Figure 3, the change is gradual; in each period, new energy-efficient (post-shock) varieties replace only a fraction of the energy-intensive (pre-shock) varieties. Only in the long run does the sector converge toward its new energy-saving production.

The speed of convergence depends on what share of the pre-shock varieties is replaced by new varieties, which depends, in turn, on the innovation rate; in our simulation, this is also endogenous. An increase in energy prices raises the production costs of iron and steel, resulting in increased revenue and R&D investment (although the effect is small; R&D investment increases by 0.2%). This increase translates into a higher innovation rate, accelerating the convergence toward energy-efficient production. However, this outcome is highly sensitive to other parameters of the economy. For instance, in our simulation, we assume a low Armington elasticity (equal to 0.1), implying that domestic iron and steel are not easily substituted with foreign alternatives. With high elasticity, the high sector costs lead to a large drop in demand, a decrease in revenue, reduced R&D investment in iron and steel, and a slower rate of convergence.

Overall, the predictions of the model with endogenous technological change suggest that a carbon tax results in deeper emissions reductions than if we assume exogenous technological change (bottom-right corner of Figure 3). Importantly, this difference increases over time: in the short run, endogenous technological change has a minimal impact on emissions; in the long run, the impact becomes more pronounced.

Comparable inertia effects are present in the models that distinguish between new and old vintages of capital (see, e.g., the ENVISAGE model by Roson and van der Mensbrugghe (2012)). However, in those models, speed is determined by the share of capital installed in a given period (new vintage) in the total use of capital. This share is, in turn, determined by the demand for capital and the depreciation rate. As noted above, in our model, the speed is determined by the innovation rate, which is, in turn, determined by expected profits determined by sector revenue.

5.4 Path dependency

To demonstrate path dependency, we compare the results from two scenarios: gradual and late decarbonization. For the gradual scenario, we assume that in 2025, the government will begin setting a limit for emissions, gradually tightening this to halve emissions in the BAU by 2040. This necessitates a yearly limit reduction

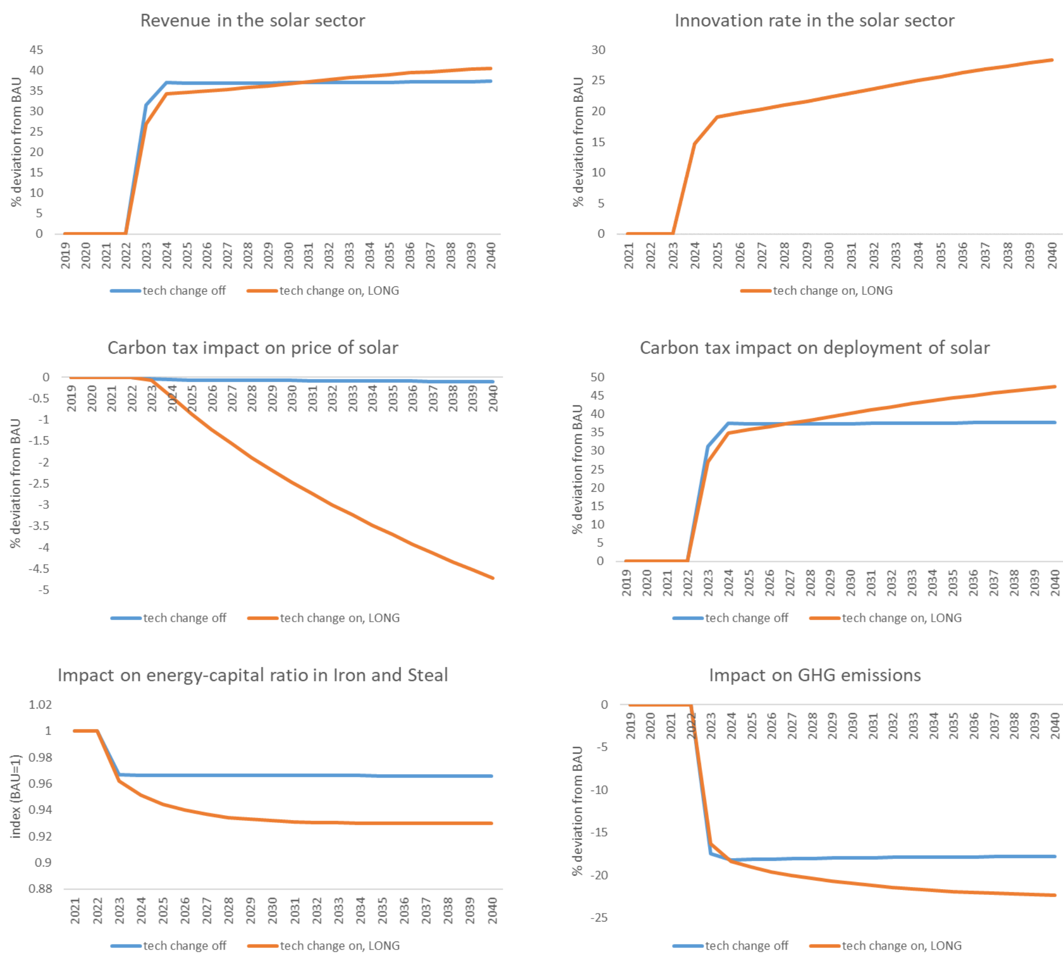

Figure 3. The impact of a carbon tax: Comparison of the results for the model with (tech change on) and without (tech change off) endogenous technological change.

Notes: The simulations with endogenous technological change are prepared using the long version of the model

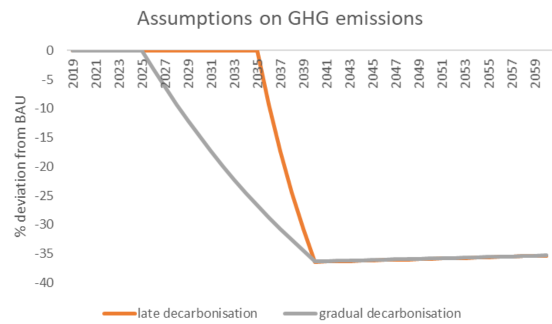

Figure 4. The assumptions on the path of the emissions limit set by the government in the gradual- and late-decarbonization scenarios

of 4.5% (compared to the BAU) from 2025 to 2040. In the late decarbonization scenario, the government delays its climate policy until 2035 and then follows a rapid-reduction path (13% reduction per year) to achieve a 50% level of emissions (compared to the BAU) by 2040. The emissions target assumptions are displayed in Figure 4. In both scenarios, the targets are achieved through a carbon tax covering all GHG emissions in the economy (the model computes the carbon tax necessary to achieve a given emission limit for each simulation year).

The simulation horizon extends to 2060. In both scenarios, the emissions limit remains at 50% of the BAU level after 2040. The results for 2040-2060 highlight the role of path dependency, illustrating how the short- and medium-term decarbonization path influences the cost of decarbonization in the long run. Due to the extended time horizon, we utilize the short version of the model for the simulations in this section.

The difference between the two scenarios is most evident in the cost path of photovoltaic energy. In both scenarios, we observe a reduction in costs driven by induced technological change (the left side of Figure 5). However, in the late scenario, this reduction is delayed; it accelerates in 2035. The acceleration occurs because the rapid emission reduction in that scenario requires a relatively high carbon tax, stimulating higher innovation rate. The manifestation of path dependency emerges after 2040: in both scenarios, costs continue to decrease due to the accumulation of know-how, which, through intertemporal knowledge spillover, facilitates increased innovation rate. Additionally, although the cost difference between the scenarios begins to decrease in 2040, it never disappears. In other words, the costs of PV in the gradual scenario also remain below those in the late scenario in the long run. The reason is that in the gradual scenario, the accelerated accumulation of know-how commences earlier, and the spillover effects are at work for a longer period. Consequently, by 2040, innovators in the gradual scenario have an advantage over those in the late scenario. The persistent difference between the two scenarios can be observed also in the path of GDP (the right side of Figure 5).

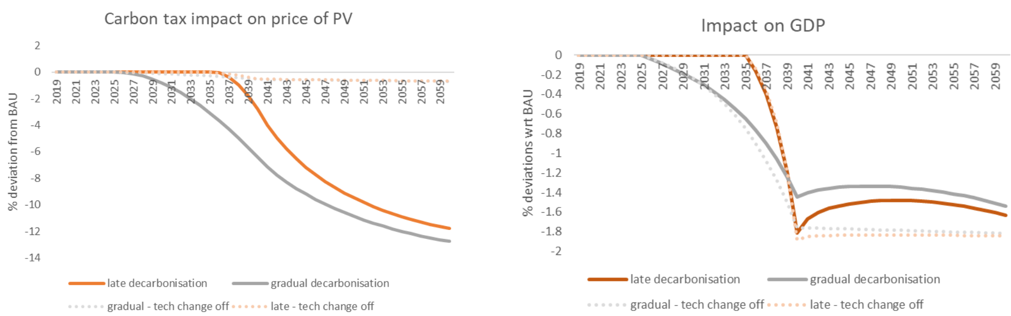

Figure 5. The impact of gradual and late decarbonization; simulations performed using the short version of the model.

As expected, these persistent effects almost vanish if we assume the Indian R&D sector makes a limited contribution to global frontier technological progress (i.e. low values of parameter ρ). The results for these simulations are provided in Appendix A7.

The impact of path dependency is also evident in macroeconomic outcomes. The right panel of Figure 5 depicts the GDP loss (with respect to the BAU) under the gradual and late decarbonization scenarios. The carbon tax exerts a distortionary effect in both cases, leading to GDP losses. The earlier initiation of the tax in the gradual scenario results in earlier GDP reductions. However, under the late scenario, in 2040, the GDP loss significantly exceeds that in the gradual scenario due to inertia; in the late scenario, very rapid decarbonization means firms lack sufficient time to develop new technologies, resulting in relatively high transition costs. In the years after 2040, firms continue to innovate and adjust technologies, thereby mitigating the decline in GDP. Nevertheless, GDP loss in the late scenario remains large relative to that in the gradual scenario, signifying path dependency, as explained above.

6. Discussion of assumptions and limitations

In this section, we outline and discuss the key assumptions underlying the most important predictions of our model, and compare them with alternative setups. We divided the assumptions into two categories: those related to market forces that induce R&D by and those concerning the dynamic effects of innovations.

6.1 Market induced R&D

The transmission of the changes in demand into changes in R&D in the proposed setup relies on three core assumptions: (i) innovations do not occur automatically or as a byproduct of production but require intentional R&D efforts., (ii) R&D investments are driven by firms' expectations of recovering R&D costs through future operational surpluses, (iii) the supply of R&D resources at the sectoral level is not inelastic.[10]

The first assumption may fail if innovation arises from learning by doing, as described by Arrow (1962) and Young (1993). In this case, productivity improvements may occur simply as a byproduct of machine usage. Under this hypothesis, the innovation rate is a function of capital use, which could lead to very different dynamics compared to our model. For instance, an increase in unit production costs in our setup would raise total costs and sectoral revenue (assuming demand for the sector's output is inelastic), increase dedicated R&D resources, and boost innovation. In contrast, in a 'learning by doing' setup, increased costs would reduce output and capital use, thus reducing innovation.

Another case where the first assumption may fail is when productivity improvements stem from the purchase of foreign technologies embodied in imported goods. For example, import of more fuel-efficient vehicles could reduce production cost across sectors that rely on those vehicles. If local vehicle producers are unable to adopt these more efficient technologies, the rate of productivity improvement would depend on trade policy. High tariffs, for instance, could impede productivity growth. Meanwhile, our model would incorrectly predict that productivity growth will continue (unless tariffs lead to a reduction in sectoral revenues).

The second assumption - that R&D investment is profit-driven - could fail when firms expect zero profits due to intense competition and low markups, even when demand is high. It could also fail in situations where there is insufficient competition: when there are strong institutional barriers to entry, firms do not invest in new technologies because they cannot capture the market even with superior technologies (e.g., the case of Japan as described by McKinsey & Company (2000)). In these cases, the link between market conditions and R&D effort would be broken, leading to predictions that differ from those outlined in our paper. Technological change in such cases might instead be driven by external factors, such as government programs.

Finally, a potential failure of the third assumption arises when firms wish to increase R&D expenditure in response to changes in demand but in equilibrium the R&D effort in a given sector is unaltered because the R&D resource is sector-specific and its supply is inelastic. For example, this might occur if the required input for local R&D activities is an expertise of workers with an extensive experience in a given sector, a sector-specific talent or entrepreneurship of local population (e.g.

[10] The first and second assumptions were introduced and discussed in the literature on endogenous growth, particularly in the works of Romer (1989),Romer (1990). The assumption regarding sectoral R&D supply and the shifting of R&D effort has been explored in the DTC literature (Acemoglu (1998), Acemoglu et al. (2015), Acemoglu et al. (2012), Aghion et al. (2016)). In the context of countries that are not on the frontier, Nelson and Phelps (1966) argued that adaptation and secondary innovations require resources such as human capital, an argument later incorporated in endogenous technological change models by Howitt and Mayer-Foulkes (2005) and Aghion and Howitt (2009).

the number of passionate gardeners).

6.2 Dynamics of technological change

The dynamic predictions of the model - such as a differences in the economy's response to shocks in the immediate, short, and long run - rely on assumptions regarding: (i) the presence of exogenous frontier and spillover effects, (ii) the inertia of the system and the factors that determine the speed of adjustment, and (iii) firms' expectations.

In our model, the growth of the technological frontier is fully exogenous and independent of the innovation rate. This assumption may break down in the presence of spillover effects within the sector. Consider, for example, a scenario where spillovers are the sole driver of progress at the frontier. Suppose that each local (region- and sector-specific) innovation contributes to local pool of knowledge:

$$KNW_{a,t} - KNW_{a,t-1} = \xi * KNW_{a,t-1} * INN_{a,t} \tag{22}$$

(where $\xi * KNW_{a,t-1}$ represents the contribution of each innovation) and the growth of knowledge drives the growth of the frontier, so that:

$$g_{a,t} = \frac{KNW_{a,t} - KNW_{a,t-1}}{KNW_{a,t-1}} = \xi * INN_{a,t}$$

If the contribution is treated as an externality, the optimization problem faced by research firms would remain unchanged, and the equilibrium innovation rate would still be determined by the demand for the sector's output. In this case, a shift in demand would lead to a corresponding shift in the innovation rate which would, in turn, accelerate the progress of the frontier. Permanent changes in demand would result in permanent changes to the growth rate of the frontier, in contrast to the predictions outlined in Section 4.2.

Second, consider the case where the progress of the frontier is driven not only by within-sector spillover effects but also by inter-sectoral and inter-regional spillovers. In this case, the pool of knowledge benefits not only from local innovations but also from innovations occurring in other sectors and other regions. In that case, equation (22) would need to be modified as follows:

$$KNW_{a,t} - KNW_{a,t-1} = (EXTa, t + \xi * INN_{a,t}) * KNW_{a,t-1}$$

where $EXT_{a,t}$ represents the external contribution to knowledge (from other sectors and regions), and the growth rate of the frontier would be given by $EXT_{a,t} + \xi * INN_{a,t}$. In this scenario, a permanent change in demand would still lead to a permanent change in the growth rate of productivity. However, the change in the growth rate would be less than proportional to the change in local innovation rates. The growth rate would be also determined by the external contribution to local knowledge (represented by $EXT_{a,t}$), which, for example, could depend on innovation rates in other sectors or countries in geographical, economic or cultural proximity to India.

Spillover effects are accounted for in some models that treat knowledge or firms' R&D efforts as inputs in a CES production function (e.g., the WITCH model by Bosetti et al. (2009), and various versions of GEM-E3 model by Karkatsoulis et al. (2014) and Fragkos et al. (2020), as discussed in section 2. Karkatsoulis et al. (2016) uses the GEM-E3 model to evaluate the cost of climate change policy and finds that the cost is significantly reduced when technological change is endogenized, which aligns with the predictions shown in Figure 5. However, the authors do not specify how much of this effect can be attributed to the presence of spillover effects. Importantly, spillover effects imply that low-carbon technological change induced in one region will lead to emissions reductions in other regions, i.e., a negative carbon leakage. This effect is captured, for instance, in the version of the GTAP-E model developed by Gerlagh and Kuik (2014). Verification of this hypothesis using the model presented in our paper would require its implementation in a multi-region model, along with careful parametrization and calibration of spillover effects. This lies beyond the scope of this paper but represents an interesting avenue for future research.

Another critical feature of the setup is its inertia: technological parameters governing the intensity of factor use can only be altered for those varieties that were invented or improved during a given period. Comparable inertia effects are present in models that distinguish between new and old vintages of capital (e.g., the ENVISAGE model by Roson and van der Mensbrugghe (2012); the ENV-Linkages model by Château et al. (2014); and the EPPA model by Chen et al. (2022)), although they reflect different narratives. In these models, the opportunity to adjust factor intensities arises from the installation of new machines, rather than from R&D activity. This difference translates also into a difference in predictions: in models with vintages, the speed at which factor intensities adjust at the sectoral level is determined by the share of capital installed during a given period (the 'new vintage') relative to the total use of capital. This share is, in turn, determined by the demand for capital and, consequently, by the demand for sectoral output. As noted earlier, in our model, the speed of adjustment is determined by the innovation rate, which itself is determined by expected profits and sectoral revenue.

Another recent refinement in CGE dynamics is the exploration of setups involving forward-looking agents. For example, Babiker et al. (2009) show that predicted consumption losses induced by GHG mitigation targets are smaller in the forward-looking version of the EPPA model than in the recursive-dynamic version. One possible explanation is that, in the latter model, consumers are able to adjust their consumption and investment profiles to minimize the loss.

As shown in Section 5, endogenizing technological change works in the same direction, i.e., reduces the negative impacts of climate mitigation policies. Therefore, one would expect that in a model combining the endogenous technology with forward-looking behavior of consumers, the two effects would add up, leading to significantly lower costs of mitigation policies compared to models that do not include these features. Perhaps more interestingly, the two features are likely to interact. For instance, in the 'late' scenario considered in Section 5.4, if firms have perfect foresight (and assuming the business-stealing effect is sufficiently low, so that firms expect long-term profits), they would anticipate a sharp increase in the revenue from low-carbon technologies in 2030s and begin investing in R&D much earlier.

7. Conclusions

We have developed a modeling framework that endogenizes technological change within a standard CGE model. This framework allows us to consider three critical mechanisms for evaluating the dynamic effects of climate policy (Grubb et al. (2021)): policy-induced innovation, inertia, and path dependency. The framework is fully micro-founded, with all equations relating investment in R&D to other economic variables derived from the optimality conditions of innovating firms. Explicitly modeling the market for innovations improves the transparency of the assumptions.

The framework is built on insights from theories of endogenous growth (Romer (1990); Grossman and Helpman (1991); Aghion and Howitt (1992)) and directed technological change (Acemoglu (1998); Acemoglu et al. (2015); Acemoglu et al. (2012); Aghion et al. (2016)), but it is tailored for country-level analysis. The framework distinguishes between two innovation types: improvements and inventions. An increase in the rate of improvement reduces the distance to the global technology frontier, bolstering short-run growth without affecting long-run growth. Conversely, an increase in the invention rate allows higher sectoral growth in the short and long term.

The model illustrates how carbon prices can induce technological change, ultimately reducing mitigation costs. The inertia of endogenous technological change, captured in our model, implies that the short-run impact of a carbon tax on emissions will be limited, manifesting fully only in the long run. Lastly, if the country's R&D sector is sufficiently large to generate inventions that significantly contribute to the progress of the global technological frontier, a change in policy timing can have a permanent impact on productivity, costs, sectoral prices and macroeconomic outcomes. For example, in scenarios where an early signal is absent, and the government opts for delayed and abrupt GHG reduction, the drop in GDP is more pronounced than in a case where the reduction in GHG emissions is gradual and commences early. The model suggests that faster action can prevent lock-in effects that impede transformation.

In the future, this framework could be extended to incorporate new features. It would be valuable, for example, to account for inter-sectoral spillover effects or expand the model to a global setting to include inter-regional spillover effects. An important limitation of the current framework is that firms' decisions regarding the intensity and direction of innovation are based on the economy's current state. An alternative setting could consider the impact of expectations influenced by policy signals. For instance, the announcement of a high carbon tax in a decade could increase the expected future value of the market for low-carbon energy sources, prompting innovation in the present. Additionally, future research could focus on developing econometric models to refine the calibration of critical parameters, such as the short- and long-run elasticities of substitution between capital and energy, baseline improvement and invention rates, and the elasticities of the innovation production function.

Acknowledgment

This study is funded by Climate Support Facility of the World Bank within the framework of the Project 'MTI Global - Macro Modeling of Climate Change - COVID-19 Green Recovery Analytical Tools & Methodologies'. The findings, interpretations, and conclusions expressed in this paper are that of authors and do not necessarily reflect the views of the World Bank, the Executive Directors of the World Bank or the governments they represent. The World Bank does not guarantee the accuracy of the data included in this work.

References

Acemoglu, D. 1998. "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality." *The Quarterly Journal of Economics*, 113(4): 1055–1089.

Acemoglu, D., P. Aghion, L. Bursztyn, and D. Hemous. 2012. "The Environment and Directed Technical Change." *American Economic Review*, 102(1): 131–166.

Acemoglu, D., et al. 2015. "Localised and Biased Technologies: Atkinson and Stiglitz's New View, Induced Innovations, and Directed Technological Change." *The Economic Journal*, 125(583): 443–463.

Aghion, P. 2002. "Schumpeterian Growth Theory and the Dynamics of Income Inequality." *Econometrica*, 70(3): 855–882.

Aghion, P., A. DechezleprÃªtre, D. Hemous, R. Martin, and J. Van Reenen. 2016. "Carbon Taxes, Path Dependency, and Directed Technical Change: Evidence from the Auto Industry." *Journal of Political Economy*, 124(1): 1–51.

Aghion, P., and P. Howitt. 1992. "A Model of Growth through Creative Destruction." *Econometrica*, 60(2): 323–351.

Aghion, P., and P. Howitt. 2009. *The Economics of Growth*, 1st ed. MIT Press.

Aguiar, A., M. Chepeliev, E. Corong, R. McDougall, and D. van der Mensbrugghe. 2019. "The GTAP Data Base: Version 10." *Journal of Global Economic Analysis*, 4(1): 1–27. doi:10.21642/JGEA.040101AF.

Aguiar, A., B. Narayanan, and R. McDougall. 2016. "An Overview of the GTAP 9 Data Base." *Journal of Global Economic Analysis*, 1(1): 181–208. doi:10.21642/JGEA.010103AF.

Arrow, K.J. 1962. "The Economic Implications of Learning by Doing." *Review of Economic Studies*, 29: 155–173.

Babiker, M., A. Gurgel, S. Paltsev, and J. Reilly. 2009. "Forward-looking versus recursive-dynamic modeling in climate policy analysis: A comparison." *Economic Modelling*, 26(6): 1341–1354. doi:10.1016/j.econmod.2009.06.009.

Bosetti, V., E. De Cian, A. Sgobbi, and M. Tavoni. 2009. "The 2008 WITCH Model: New Model Features and Baseline." Fondazione Eni Enrico Mattei (FEEM), Sustainable Development Papers No. 55284.

Caselli, F., and W.J. Coleman II. 2006. "The World Technology Frontier." *American Economic Review*, 96(3): 499–522. doi:10.1257/aer.96.3.499.

Château, J., R. Dellink, and E. Lanzi. 2014. "An Overview of the OECD ENV-Linkages Model: Version 3." OECD Publishing, OECD Environment Working Papers No. 65. doi:10.1787/5jz2qck2b2vd-en.

Chen, Y.H.H., S. Paltsev, A. Gurgel, J.M. Reilly, and J. Morris. 2022. "The MIT EPPA7: A Multisectoral Dynamic Model for Energy, Economic, and Climate Scenario Analysis." MIT Joint Program on the Science and Policy of Global Change, Joint Program Report Series No. Report 360. http://globalchange.mit.edu/publication/17777.

Dixit, A.K., and J.E. Stiglitz. 1977. "Monopolistic Competition and Optimum Product Diversity." *American Economic Review*, 67(3): 297–308.

Drandakis, E.M., and E.S. Phelps. 1966. "A Model of Induced Invention, Growth, and Distribution." *Economic Journal*, 76: 823–840.

Faehn, T., G. Bachner, R. Beach, J. Château, S. Fujimori, M. Ghosh, M. Hamdi-Cherif, E. Lanzi, S. Paltsev, T. Vandyck, B. Cunha, R. Garaffa, and K. Steininger. 2020. "Capturing Key Energy and Emission Trends in CGE models: Assessment of Status and Remaining Challenges." *Journal of Global Economic Analysis*, 5(1): 196–272.

Fragkos, P., K. Fragkiadakis, and L. Paroussos. 2020. "Low-Carbon R&D Can Boost EU Growth and Competitiveness." *Energies*, 13. doi:10.3390/en13195236.

Gerlagh, R., and O. Kuik. 2014. "Spill or leak? Carbon leakage with international technology spillovers: A CGE analysis." *Energy Economics*, 45: 381–388. doi:10.1016/j.eneco.2014.07.017.

Goldfarb, B. 2011. "Economic Transformations: General Purpose Technologies and Long-Term Economic Growth. By Richard G. Lipsey, Kenneth I. Carlaw, and Clifford T. Bekar. Oxford: Oxford University Press, 2005." *The Journal of Economic History*, 71(3): 820–823. doi:10.1017/S0022050711002099.

Grossman, G.M., and E. Helpman. 1991. "Quality Ladders in the Theory of Growth." *Review of Economic Studies*, 58(1): 43–61.

Growiec, J. 2008. "A new class of production functions and an argument against purely labor-augmenting technical change." *International Journal of Economic Theory*, 4(4): 483–502.

Grubb, M., C. Wieners, and P. Yang. 2021. "Modeling myths: On DICE and dynamic realism in integrated assessment models of climate change mitigation." *WIREs Climate Change*, 12: e698. doi:10.1002/wcc.698.

Hemous, D. 2016. "The dynamic impact of unilateral environmental policies." *Journal of International Economics*, 103: 80–95.

Howitt, P., and D. Mayer-Foulkes. 2005. "R&D, Implementation, and Stagnation: A Schumpeterian Theory of Convergence Clubs." *Journal of Money, Credit and Banking*, 37(1): 147–177.

IRENA. 2023. "International Standards and Patents in Renewable Energy - IRENA INSPIRE." https://inspire.irena.org/Pages/home.aspx, Accessed: 2024-11-13.

Jones, C.I. 1995. "R&D-Based Models of Economic Growth." *Journal of Political Economy*, 103(4): 759–784.

Jones, C.I., and J.C. Williams. 2000. "Too Much of a Good Thing? The Economics of Investment in R&D." *Journal of Economic Growth*, 5(1): 65–85.

Karkatsoulis, P., P. Capros, P. Fragkos, L. Paroussos, and S. Tsani. 2016. "First-mover advantages of the European Union's climate change mitigation strategy." *International Journal of Energy Research*, 40. doi:10.1002/er.3487.

Karkatsoulis, P., N. Kouvaritakis, L. Paroussos, P. Fragkos, and P. Capros. 2014. "Modification of GEM-E3 technological innovation module." SIMPATIC Working Paper, Report No. 18.

Kennedy, C.M. 1964. "Induced Bias in Innovation and the Theory of Distribution." *The Economic Journal*, 74: 541–547.

Kittner, N., F. Lill, and D. Kammen. 2017. "Energy storage deployment and innovation for the clean energy transition." *Nature Energy*, 2: 17125. doi:10.1038/nenergy.2017.125.

Klaassen, G., A. Miketa, K. Larsen, and T. Sundqvist. 2005. "The impact of R&D on innovation for wind energy in Denmark, Germany and the United Kingdom." *Ecological Economics*, 54(2-3): 227–240.

Koetse, M.J., H.L.F. de Groot, and R.J.G.M. Florax. 2008. "Capital-energy substitution and shifts in factor demand: A meta-analysis." *Energy Economics*, 30(5): 2236–2251.

Kriegler, E., N. Bauer, A. Popp, F. Humpenöder, M. Leimbach, J. Strefler, L. Baumstark, B.L. Bodirsky, J. Hilaire, D. Klein, et al. 2017. "Fossil-fueled development (SSP5): An energy and resource intensive scenario for the 21st century." *Global Environmental Change*, 42: 297–315. doi:10.1016/j.gloenvcha.2016.05.015.

Maddala, G.S., R.P. Trost, H. Li, and F. Joutz. 1997. "Estimation of Short-Run and Long-Run Elasticities of Energy Demand from Panel Data Using Shrinkage Estimators." *Journal of Business & Economic Statistics*, 15(1): 90–100. doi:10.2307/1392078.

Markusen, J.R. 2023. "Incorporating Theory-Consistent Endogenous Markups into Applied General-Equilibrium Models." *Journal of Global Economic Analysis*, 8(2): 60–99.

McKinsey & Company. 2000. "Why the Japanese Economy is not Growing: Micro

Barriers to Productivity Growth." McKinsey Global Institute, Washington, D.C., Report.

Nelson, R.R., and E.S. Phelps. 1966. "Investment in Humans, Technological Diffusion, and Economic Growth." *The American Economic Review*, 56(1/2): 69–75.

NREL. 2023. "Annual Technology Baseline." https://atb.nrel.gov/, Accessed on 22 December 2023.

Peters, J.C. 2016. "The GTAP-Power Data Base: Disaggregating the Electricity Sector in the GTAP Data Base." *Journal of Global Economic Analysis*, 1(1): 209–250. doi:10.21642/JGEA.010104AF.

Popp, D. 2004. "ENTICE: Endogenous Technological Change in the DICE Model of Global Warming." *Journal of Environmental Economics and Management*, 48(1): 742–768.

Romer, P.M. 1989. "Increasing Returns and New Developments in the Theory of Growth." National Bureau of Economic Research, NBER Working Papers No. 3098.

Romer, P.M. 1990. "Endogenous Technological Change." *Journal of Political Economy*, 98(5, Part 2): S71–S102. doi:10.1086/261725.

Roson, R., and D. van der Mensbrugghe. 2012. "Climate Change and Economic Growth: Impacts and Interactions." *International Journal of Sustainable Economy*, 4(3): 270–285.

Samuelson, P.A. 1965. "A Theory of Induced Innovations along Kennedy-Weisacker Lines." *Review of Economics and Statistics*, 47: 343–356.

Stehfest, E., D. van Vuuren, T. Kram, and L. Bouwman. 2014. "Integrated Assessment of Global Environmental Change with IMAGE 3.0: Model Description and Policy Applications." Netherlands Environmental Assessment Agency (PBL), Report.

Sue Wing, I. 2001. "Induced Technical Change in Computable General Equilibrium Models for Climate-Change Policy Analysis." Ph.d. thesis, Massachusetts Institute of Technology.

UNESCO. 2023. "UIS.Stat Bulk Data Download Service." https://data.worldbank.org, Accessed on 21 December 2023.

van den Bijgaart, I. 2017. "The unilateral implementation of a sustainable growth path with directed technical change." *European Economic Review*, 91: 305–327.

Wang, K., C. Wang, and J. Chen. 2009. "Analysis of the Economic Impact of Different Chinese Climate Policy Options Based on a CGE Model Incorporating Endogenous Technological Change." *Energy Policy*, 37(8): 2930–2940. doi:10.1016/j.enpol.2009.03.023.

Wiebe, K.S., and C. Lutz. 2016. "Endogenous Technological Change and the Policy Mix in Renewable Power Generation." *Renewable and Sustainable Energy Reviews*, 60: 739–751.

WIPO. 2023. "Intellectual Property Facts and Figures." https://www.wipo.int/en/ipfactsandfigures/patents, Accessed on 20 December 2023.

Witajewski-Baltvilks, J., E. Verdolini, and M. Tavoni. 2017. "Induced Technological Change and Energy Efficiency Improvements." *Energy Economics*, 68(S1): 17–32.

World Bank. 2024. "MANAGE-WB - The Mitigation, Adaptation and New Technologies Applied General Equilibrium Model of the World Bank. Model Documentation and User Guide." The World Bank, Report, December. https://thedocs.worldbank.org/en/doc/77351105a334213c64122e44c2efe523-0500072021/related/MANAGE-WB-Documentation.pdf.

Young, A. 1993. "Invention and Bounded Learning by Doing." *Journal of Political Economy*, 101(3): 443–472. doi:10.1086/261882.

Young, A. 1998. "Growth without Scale Effects." *Journal of Political Economy*, 106(1): 41–63.

Appendix A. Equations in section 2 for the case of $\sigma_a^{var} < 1$

In the text we have considered the case when processes are gross substitutes, i.e. $\sigma_a^{var} > 1$. In this case every invention adds one variety and does not affect the remaining varieties. In this appendix we will describe a more general version of the model that allows σ_a^{var} to take negative values. When $\sigma_a^{var} < 1$ productivity at the sectoral level decreases with the number of varieties. Therefore for this case we assume that every invention adds one new variety and, at the same time makes two other varieties redundant, so that total number of varieties required for production decreases by one. We do not consider the case of $\sigma_a^{var} = 1$, because it significantly complicates the model.

Below we list all equations from the text that must be modified to account for the case of $\sigma_a^{var} < 1$. To facilitate implementation, we introduce a parameter κ that takes the value of $\kappa = 1$ if $\sigma_a^{var} < 1$ and $\kappa = 0$ if $\sigma_a^{var} > 1$.

First, equation (3) must be replaced with:

$$N_{a,\tau,t} = \begin{cases} (INV_{a,t} + IMP_{a,t})\, M_{a,t-1} & \text{for } \tau = t \\ (1 - IMP_{a,t} - 2\kappa INV_{a,t})\, N_{a,\tau,t-1} & \text{for } \tau < t \end{cases} \tag{A.1}$$

Similarly, $N_{a,0,t}$ evolves over time according to:

$$N_{a,0,t} = (1 - IMP_{a,t} - 2\kappa INV_{a,t})\, N_{a,0,t-1} \tag{A.2}$$

where $N_{a,0,0} = 1$

Second, we must amend the expression for the value of the blueprint. As described in section 2, a successful innovator will receive the cash flow from the blueprint only until another successful innovator produces a better blueprint that improves the technology of the variety; at time t, this happens with probability $IMP_{a,t}$. In addition, when $\sigma_a^{var} < 1$, a new invention can make the old variety redundant. Since we assumed invention replaces two old varieties with one, the probability that an old variety becomes obsolete at time t is $2IMP_{a,t}$.

We make the following assumption regarding firm's expectations: When computing the value of a blueprint, firms assume that (i) the probability that their blueprint is replaced in the future is constant and given by the expected improvement rate next period, $IMP_{a,t}^{exp} \equiv IMP_{a,t+1}$, and (ii) that the probability that their blueprint becomes redundant in the future is constant and given by $2\kappa INV_{a,t}^{exp} \equiv 2\kappa INV_{a,t+1}$. Therefore equation (10) must be replaced with:

$$\frac{\gamma p_{a,t+1,t+1}^{top} q_{a,t+1,t+1}^{top}}{1 + i_0^{RD}} + \frac{\left(1 - IMP_{a,t}^{exp} - 2\kappa INV_{a,t}^{exp}\right)}{\left(1 + i_0^{RD}\right)^2} \gamma p_{a,t+1,t+2}^{top} q_{a,t+1,t+2}^{top} + ... =$$

$$\left(\frac{1+g_{a,t}^{exp}}{1+i_0^{RD}}+\frac{\left(1-IMP_{a,t}^{exp}-2\kappa INV_{a,t}^{exp}\right)\left(1+g_{a,t}^{exp}\right)^2}{\left(1+i_0^{RD}\right)^2}+...\right)\gamma p_{a,t,t}^{top}q_{a,t,t}^{top}=$$

$$=\frac{\gamma p_{a,t,t}^{top}q_{a,t,t}^{top}}{\left(1+i^{RD}\right)/\left(1+g_{a,t}^{exp}\right)-\left(1-IMP_{a,t}^{exp}-2\kappa INV_{a,t}^{exp}\right)} \tag{A.3}$$

and so equation (12) becomes:

$$p_{RD,t}\Phi_{a,t}=R_{a,t}^{-\left(1-\phi^{RD}\right)}\Gamma\frac{\gamma p_{a,t,t}^{top}q_{a,t,t}^{top}}{\left(1+i^{RD}\right)/\left(1+g_{a,t}^{exp}\right)-\left(1-IMP_{a,t}^{exp}-2\kappa INV_{a,t}^{exp}\right)} \tag{A.4}$$

Appendix B. Implementation in the CGE framework

In this appendix, we describe the set of equations that constitute the 'short version' of the model. We begin by providing an overview of implementation strategy, including the key equations, parameters, and variables. In section A2.2, we replace the setup of a single-nest CES production function with a more general production function, in order to align our setup with that of large-scale CGE models. In section A2.3, we introduce additional definitions to facilitate implementation. Section A2.4 presents and discusses in detail the approximations applied in the short version. In Section A2.5, we describe the necessary changes to the calibration of the model. Finally, in Section A2.6, we provide a detailed description of the blocks of equations that constitute the short version of the model.

The equations are derived from the solution of the microfounded model presented in Section 2, but we refrain from presenting extensive derivations here. Instead, these derivations are described in Appendix A3, which also includes the list of equations implemented in the long version of the model.

A2.1 Overview

The model can be implemented within a standard recursive-dynamic CGE framework. It was originally designed for a single-country model, but the framework is also consistent with a multi-country model.[11] The model does not impose specific requirements regarding sectoral disaggregation. Endogenous technological change can be activated for a selection of sectors or for all sectors[12]. The directed technological change component of the model (described in section 3.8) was designed for a nested CES production function; there are no preconditions regarding the structure of the nesting.

The model can be implemented in five steps, which we outline below. We also list the variables, parameters, and equations required for implementation in Tables 1 - 3. A detailed description of the equations is provided in the section following this overview.

Step 1: The first step involves adding an additional index (e.g. denoted by v) to every variable related to production. In the short (default) version of the model, $v = (new, old)$, which implies that there will be two production functions for each sector: one for new and one for old varieties.[13] To allocate total output across new and old varieties in a sector, equations (B.19) and (B.20) can be used (see Table 2 or the detailed description of equations in Section A2.6).

Step 2: Monopolistic competition requires the introduction of a wedge between unit costs of production and output prices. This can be achieved using equation (B.22). The operating surplus is allocated to finance R&D investment; therefore, in the equation that balances the supply and demand for investment, we need to include R&D investment on the demand side (see equation (B.39) in Table 3 and in Section A2.6).

These changes also require adjustments to the calibration of share parameters in the model. Specifically: (i) for each sector, the values of compensation for physical capital in the calibration year can be computed by subtracting the value of operating surplus from the compensation for capital reported in national statistics, (ii) for

[11] Note, however, that incorporating international and trade effects would require extending the model to account for international spillovers and the diffusion of technological progress through trade - see the discussion in Section 6.1 (for trade effects) and 6.2 (for spillover effects).

[12] However, activating endogenous technological change for sectors producing commodities with very high elasticity of demand may cause explosive behavior: an increase in demand could cause greater innovation and cause prices to drop, which in turn may lead to a further sharp increase in demand.

[13] In some models (e.g., the original version of MANAGE, ENVISAGE, ENV-Linkages), a similar structure is used to differentiate between vintages of capital. The simplest implementation in these models would involve retaining the index v and reinterpreting its elements: instead of 'new' and 'old' vintages, it would now represent 'new' and 'old' varieties.

one of the actors receiving capital income, the value of that income in the calibration year can be calculated by deducting the operating surplus from the reported income in the national statistics, and (iii) the savings of that actor must be reduced by the value of the operating surplus. The details and logic of these changes are explained in the subsection A2.5 below.

Step 3: Due to myopia of actors, total operating surplus could be either higher or lower than R&D expenditure. The difference must be financed by one of the actors. The ideal candidate for this would be the actor receiving large share of capital income. In the model, we must add the operating surplus and subtract R&D expenditure from the income of that actor (see equation (B.38) in Table 2 or in Section A2.6).

Step 4: The next step is to link the R&D intensity variable to the sectoral revenue variable (see equation (B.37)), and then link the innovation rate to R&D intensity (see equations (B.33) - (B.36)).

Step 5: After solving the model for each year, the innovation rate is used to update the productivity parameter (see equations (B.23) - (B.25)). Additionally, predictions for prices and input demand are used to update the optimal choice of productivity for inputs (see equation (B.31)). This updated choice is then used to adjust the share parameters in the production function (see equation (B.32)).

PRE-EXISTING VARIABLES	
$P_{a,t}$,$Q_{a,t}$	Sectoral output and its price
NEW VARIABLES	
$Q_{a,t}^{top,new}$, $Q_{a,t}^{top,old}$	Output of new and old varieties
$P_{a,t}^{top,new}$, $P_{a,t}^{top,old}$	Price of new and old varieties
$\lambda_{a,t}^{top,new}$, $\lambda_{a,t}^{top,old}$	Productivity of new and old varieties
$INN_{a,t}$	Innovation rate
$INN_{a,t}^{exp}$	Expected innovation rate next period
$IMP_{a,t}$	Improvement rate
$INV_{a,t}$	Invention rate
$M_{a,t-1}$	Number of varieties
$R_{a,t}$	R&D effort
$\alpha_{a,ch,t}^{new}$, $\alpha_{a,ch,t}^{old}$	Reduced form factor intensity for new and old varieties
$\alpha_{a,t}^{ch}$	Firm's choice of factor intensities for new varieties
PRE-EXISTING PARAMETERS	
$b_{a,ch,t}$	[check in A2] Calibrated factor intensity for new varieties
NEW PARAMETERS	
σ_a^{var}	Elasticity of substitution across varieties
γ	Markup
$frontier_t$	Level of advancement of the frontier
g_t	Growth rate of the frontier
ρ	Ratio of invention rate to innovation rate.
Γ	[check]
i^{RD}	Interest rate applied to R&D projects
$g_{a,t}^{exp}$	Expected growth of sectoral revenue

Table B.1. Variables and parameters

Equation	Ref. to section A2.6
Distribution of production across new and old varieties	
$Q_{a,t}^{top,new} = (IMP_{a,t} + INV_{a,t}) \frac{M_{a,t-1}}{\lambda_{a,t}^{top,new}} \left(\frac{P_{a,t}\lambda_{a,t}^{top,new}}{(1+\gamma)P_{a,t}^{top,new}} \right)^{\sigma_a^{var}} Q_{a,t}$	(B.19)
$Q_{a,t}^{top,old} = (1 - IMP_{a,t}) \frac{M_{a,t-1}}{\lambda_{a,t}^{top,old}} \left(\frac{P_{a,t}\lambda_{a,t}^{top,old}}{(1+\gamma)P_{a,t}^{top,old}} \right)^{\sigma_a^{var}} Q_{a,t}$	(B.20)
$P_{a,t}Q_{a,t} = (1+\gamma)\left(Q_{a,t}^{top,old} P_{a,t}^{top,old} + Q_{a,t}^{top,new} P_{a,t}^{top,new} \right)$	(B.22)
Vertical technological progress	
$\lambda_{a,t}^{top,new} = frontier_t$	(B.23)
$frontier_t = (1+g)\, frontier_{t-1}$	(B.24)
$(1 + INV_{a,t-1}) \left(\lambda_{a,t}^{top,old} \right)^{\sigma_a^{par}-1} =$ $(IMP_{a,t-1} + INV_{a,t-1}) \left(\lambda_{a,t-1}^{top,new} \right)^{\sigma_a^{par}-1}$ $+ (1 - IMP_{a,t-1}) \left(\lambda_{a,t-1}^{top,old} \right)^{\sigma_a^{par}-1}$	(B.25)
Horizontal technological progress	
$M_{a,t} - M_{a,t-1} = INV_{a,t} M_{a,t-1}$	(B.26)
Direction of technological change	
$\alpha_{a,t}^{ch} = \left(\lambda_{a,t}^{ch}\right)^{1-\sigma_a^{par}} \left(\frac{P_{a,t}^{ch,new}}{P_{a,t}^{par,new}} \right)^{\sigma_a^{par}} \frac{Q_{a,t}^{ch,new}}{Q_{a,t}^{par,new}}$	(B.31)
$(1 + INV_{a,t-1})\, \alpha_{a,t}^{ch,old} =$ $INN_{a,t-1}\alpha_{a,t-1}^{ch} + (1 - IMP_{a,t-1})\, \alpha_{a,t-1}^{ch,old}$	(B.32)

Table B.2. Blocks of endogenous technological change equations - Part 1. Equations presented here are for the case of $\sigma^{var} > 1$. A detailed description of the equations is presented in Section A2.6 and the derivations are presented in Appendix A3

Equation	Ref. to section A2.6
Linking improvements and inventions to R&D effort	
$INN^{exp}_{a,t} = R^{\phi^{RD}}_{a,t} \left(\frac{INN_{a,0}}{R^{\phi^{RD}}_{a,0}} \right)$	(B.33)
$INN_{a,t} = INN^{exp}_{a,t-1}$	(B.34)
$INV_{a,t} = \rho INN_{a,t}$	(B.35)
$IMP_{a,t} = (1-\rho)\, INN^{exp}_{a,t}$	(B.36)
Linking R&D effort to profits and revenues	
$R^{1-\phi^{RD}}_{a,t} p_{RD,t} \Phi_{a,t} = \frac{\Gamma \gamma P^{top,new}_{a,t} Q^{top,new}_{a,t}}{(1+i^{RD})/\left(1+g^{exp}_{a,t}\right) - \left(1 - IMP^{exp}_{a,t}\right)}$	(B.37)
Adjustment of preexisting equations.	
Enterprise income equation	(B.38)
Investment-savings balance equation	(B.39)
Capital accumulation equation	(B.40)

Table B.3. Blocks of endogenous technological change equations - Part 2. A detailed description of the equations is presented in Section A2.6 and the derivations are presented in Appendix A3

A2.2 General Nested CES Production Function

In Section 3, we presented a simplified model that assumes a simple single-nest CES production function for varieties, in order to simplify the exposition. The full version of the model, as presented in Appendices A2 and A3, can be incorporated into models with any nested production function. In this subsection, we replace the simplified setup from Section 3.3 with a more general nested production function involving multiple nests. This setup was developed by the World Bank CGE team and is described in detail in the MANAGE model documentation (see the supplementary material to this article).

For each nest, we label the output of the nest as the 'parent' (*par*) and the input as the 'child' (*ch*). We assume that each nest has only one output, so each nest corresponds to one parent, and each parent is associated with a subset of children, $nest\,(par)$. The production function for the variety can then be written as a set of equations:

$$q^{par}_{a,\tau,t} = \left(\sum_{ch \in nest(par)} \left(\alpha^{ch}_{a,\tau}\right)^{\frac{1}{\sigma^{par}_a}} \left(\lambda^{ch}_{a,\tau} q^{ch}_{a,\tau,t}\right)^{\frac{\sigma^{par}_a - 1}{\sigma^{par}_a}} \right)^{\frac{\sigma^{par}_a}{\sigma^{par}_a - 1}}$$

,

for every *par* in the set of parents. The technological parameters $\alpha^{ch}_{a,\tau}$ and $\lambda^{ch}_{a,\tau}$ determine the input intensity of *ch* in the nest *par* for variety τ. In this paper, we use $\alpha^{ch}_{a,\tau}$ to control the intensity of input use (see the 'Direction of Technological Change' section below). $\lambda^{ch}_{a,\tau}$ can be set to unity, but in some applications of the CGE model, it can serve as a useful shifter of input productivity.

One can describe the entire nested structure by moving from the top to the bottom. There is a single top nest producing a parent called top. The children for that nest may include intermediate goods, materials (i.e., outputs of other sectors), and factors of production. When a child is an intermediate good, it must be a parent in one of the lower nests. Each of these lower nests has its own children, which may again include intermediate goods, materials, and factors of production.

In our model, each element in the set *ch* is assigned to exactly one element in the set *par*, meaning that a child cannot be used as an input in more than one nest. An example of the nesting structure we use to generate our illustrative simulations is given in Section 5.1.

The demand for the child, $q^{ch}_{a,\tau,t}$, as a function of the demand for the parent, $q^{par}_{a,\tau,t}$, can be derived from the cost-minimization problem of a monopolist who manufactures a variety, as follows:

$$q^{ch}_{a,\tau,t} = \alpha^{ch}_{a,\tau} \left(\frac{p^{ch}_{a,\tau,t}}{p^{par}_{a,\tau,t}} \right)^{-\sigma^{par}_a} \left(\lambda^{ch}_{a,\tau}\right)^{\sigma^{par}_a - 1} q^{par}_{a,\tau,t} \tag{B.1}$$

,

where p^{ch} denotes the cost of the child, *ch*, and p^{par} denotes the cost of manufacturing the output of the parent, *par*.

The unit cost of the output can also be derived from the cost-minimization problem:

$$p_{a,\tau,t}^{par} = \left(\sum_{ch \in nest(par)} \alpha_{a,\tau}^{ch} \left(\frac{p_{a,\tau,t}^{ch}}{\lambda_{a,\tau}^{ch}} \right)^{1-\sigma_a^{par}} \right)^{\frac{1}{1-\sigma_a^{par}}} \tag{B.2}$$

A2.3 Additional definitions

We introduce several additional variables to facilitate implementation. First, we introduce the variable $Q_{a,t}^{top,new}$, which represents the output of all varieties improved at time t:

$$Q_{a,t}^{top,new} = N_{a,t,t} q_{a,t,t}^{top} \tag{B.3}$$

.

The variable $\lambda_{a,t}^{top,new}$, represents the productivity of varieties invented or improved at time t:

$$\lambda_{a,t}^{top,new} = \lambda_{a,t}^{top}$$

$P_{a,t}^{top,new}$ represents the unit cost of production (excluding markups) of $Q_{a,t}^{top,new}$:

$$P_{a,t}^{top,new} = p_{a,t,t}^{top} \tag{B.4}$$

.

For the lower nests, let $Q_{a,t}^{ch,new}$ be the total input ch demanded by the new varieties:

$$Q_{a,t}^{ch,new} = N_{a,t,t} q_{a,t,t}^{ch} \tag{B.5}$$

.

and let $\lambda_{a,t}^{ch,new}$ be the productivity of input ch for the new varieties

$$\lambda_{a,t}^{ch,new} = \lambda_{a,t}^{ch}$$

.

For consistency of notation, we also let $\alpha_{a,t}^{ch,new}$ stand for the share parameter in the case of new varieties:

$$\alpha_{a,t}^{ch,new} = b_{a,ch,t} = \left(\beta_{a,ch}\right)^{\frac{1-\tilde{\sigma}_a^{par}}{1-\sigma_a^{par}}} \tag{B.6}$$

Similarly, let $P_{a,t}^{ch,new}$ be the unit cost of production of ch (when ch is an intermediate good) or a price of ch (when ch is a factor of production):

$$P_{a,t}^{ch,new} = p_{a,t,t}^{ch}$$

Next, we define a price index for old varieties, that is varieties that were invented or last improved at time $\tau \leq t-1$ as follows:

$$P_{a,t}^{top,old} = \left(\sum_{\tau}^{t-1} \frac{N_{a,\tau,t}}{\left(1 - IMP_{a,t} - 2\kappa INV_{a,t}\right) M_{a,t-1}} \left(p_{a,\tau,t}^{top}\right)^{1-\sigma_a^{top}} \right)^{\frac{1}{1-\sigma_a^{top}}} \tag{B.7}$$

We define the productivity index:

$$\lambda_{a,t}^{top,old} = \left(\sum_{\tau}^{t-1} \left[\left(\lambda_{a,\tau}^{top} \right)^{\sigma_a^{var}-1} \frac{N_{a,\tau,t}}{\left(1 - IMP_{a,t}^{exp} - 2\kappa INV_{a,t}^{exp}\right) M_{a,t-1}} \right] \right)^{\frac{1}{\sigma_a^{var}-1}} \tag{B.8}$$

and the demand for output of the old varieties as

$$Q_{a,\tau,t}^{top,old} = \sum_{\tau}^{t-1} N_{a,\tau,t} q_{a,\tau,t}^{top}$$

.

The demand for inputs, *ch*, generated by old varieties is defined as:

$$Q_{a,\tau,t}^{ch,old} = \sum_{\tau}^{t-1} N_{a,\tau,t} q_{a,\tau,t}^{ch} \tag{B.9}$$

We define their productivity index as:

$$\lambda_{a,t}^{ch,old} = \left(\sum_{\tau}^{t-1} \left[\left(\lambda_{a,\tau}^{ch} \right)^{\sigma_a^{par}-1} \frac{N_{a,\tau,t}}{\left(1 - IMP_{a,t} - 2\kappa INV_{a,t}\right) M_{a,t-1}} \right] \right)^{\frac{1}{\sigma_a^{par}-1}} \tag{B.10}$$

The associated price index is expressed as follows:

$$P_{a,t}^{ch,old} = \left(\sum_{\tau}^{t-1} \frac{N_{a,\tau,t}}{\left(1 - IMP_{a,t} - 2\kappa INV_{a,t}\right) M_{a,t-1}} \left(p_{a,\tau,t}^{ch} \right)^{1-\sigma_a^{ch}} \right)^{\frac{1}{1-\sigma_a^{ch}}} \tag{B.11}$$

.

Further, let $\alpha_{a,t}^{ch,old}$ stand for the weighted average of share parameters of old varieties.

$$\alpha_{a,t}^{ch,old} = \sum_{\tau}^{t-1} \frac{N_{a,\tau,t}}{\left(1 - IMP_{a,t} - 2\kappa INV_{a,t}\right) M_{a,t-1}} \alpha_{a,\tau}^{ch} \left(\frac{\lambda_{a,\tau}^{ch}}{\lambda_{a,t}^{ch,old}} \right)^{\sigma_a^{par}-1} \tag{B.12}$$

.

Finally, let define

$$\sigma_a^{par,old} = \sigma_a^{par}$$

and

$$\sigma_a^{par,new} = \tilde{\sigma}_a^{par} = \frac{\omega_a^{par} \sigma_a^{par} - 1}{\omega_a^{par} - 1}$$

.

A2.4 Approximations for the short version of the model

In this section, we discuss the possibility of introducing several approximations related to the demand and manufacturing costs of old varieties, that is, varieties that were not improved or invented in the current period. The introduction of approximations is optional: in our numerical model the user can choose whether to include them or not. We label the model including approximations the short version and the one without these as the long version. The introduction of approximations has important advantages and equally important disadvantages, which we discuss towards the end of this subsection.

In both the long and the short version the firms' choices regarding varieties just invented (i.e., varieties represented by $\tau = t$) are consistent with firms' cost minimization and can be described by equations (5), (7), (B.1) and (B.2). In the long version the same is true for all varieties. In the short version for all old varieties (i.e., those in cohorts $\tau \leq t-1$), we compute manufacturing costs by replacing the variety-specific cost of manufacturing in equation (B.2), $p^{ch}_{a,\tau,t}$, with the corresponding cost index, $P^{ch,old}_{a,t}$, which is constant across old varieties. In other words, we replace equation (B.2) with equation

$$p^{par}_{a,\tau,t} = \left(\sum_{ch \in nest(par)} \alpha^{ch}_{a,\tau} \left(\frac{P^{ch,old}_{a,t}}{\lambda^{ch,old}_{a,\tau}} \right)^{1-\sigma^{par}_a} \right)^{\frac{1}{1-\sigma^{par}_a}} \tag{B.13}$$

.

This, in combination with equation (B.11), implies that the price index of manufacturing nests for old varieties can be expressed as:

$$\left(P^{par,old}_{a,t}\right)^{1-\sigma^{par}_a} = \sum_{ch \in nest(par)} \left(P^{ch,old}_{a,t}\right)^{1-\sigma^{par}_a} \alpha^{ch,old}_{a,t} \left(\lambda^{ch,old}_{a,t}\right)^{\sigma^{par}_a - 1} \tag{B.14}$$

In addition, we compute the price of the sector output by replacing $p^{top}_{a,\tau,t}$ for $\tau \leq t-1$ in equation (6) with $P^{top,old}_{a,t}$. Then equation (6) becomes:

$$P_{a,t} = \frac{N_{a,t,t} p^{top}_{a,t,t} q^{top}_{a,t,t} + \sum_{\tau}^{t-1} N_{a,\tau,t} P^{top,old}_{a,t} q^{top}_{a,\tau,t}}{Q_{a,t}} \tag{B.15}$$

Similarly, we compute demand by replacing the variety-specific cost of manufacturing, $p^{ch}_{a,\tau,t}$ and $p^{par}_{a,\tau,t}$, and outputs, $q^{par}_{a,\tau,t}$, in equation (B.1) with their corresponding indices, $P^{ch,old}_{a,t}$, $P^{par,old}_{a,t}$ and $Q^{par,old}_{a,t}$:

$$q^{ch}_{a,\tau,t} = \tag{B.16}$$

$$\alpha^{ch}_{a,\tau}\left(\frac{P^{ch,old}_{a,t}}{P^{par,old}_{a,t}}\right)^{-\sigma^{par}_a}\left(\lambda^{ch}_{a,\tau}\right)^{\sigma^{par}_a-1}\frac{Q^{par,old}_{a,t}}{\left(1-IMP_{a,t}-2\kappa INV_{a,t}\right)M_{a,t-1}}$$

For the top nest, we compute the total demand for varieties $\tau \leq t-1$ (originally, equation (5)) using the following:

$$q^{top}_{a,\tau,t} = \left(\lambda^{top,old}_{a,t}\right)^{\sigma^{var}_a-1}\left(\frac{P_{a,t}}{(1+\gamma)\,P^{top,old}_{a,t}}\right)^{\sigma^{var}_a} Q_{a,t} \tag{B.17}$$

Discussion of approximations

The approximation is exact when all old varieties are identical. If we consider a scenario where, initially, all varieties are identical, but after an economic shock, they begin to vary, the predictions of the short version of the model will ignore several effects. The most important of these are: (i) changes in the shares of different old varieties after changes in factor prices and (ii) correlation between input price and its intensity of use across varieties. We elaborate on these two effects below and, in Appendix A5, set out further mathematical detail of all effects that are ignored by making the approximation. In addition, in Appendix A6 we present a comparison of the carbon tax simulation results generated with the short version of the model (with approximations) and the long version of the model (without approximations).

By introducing the approximations, we ignore that after a change in factor prices, the resulting change in manufacturing costs and demand will differ across old varieties. Consider, for example, a carbon tax introduced at time T_1. All pre-shock varieties, that is, all varieties improved or invented at time $\tau < T_1$, will be characterized by high energy intensity (high α^{energy}) and post-shock varieties (with $\tau \geq T_1$) will be characterized by lower energy intensity (lower α^{energy}).

Now suppose that at time $T_2 > T_1$, there is another increase in energy prices. in the long version of the model, varieties with $\tau < T_1$ will immediately become less competitive, and their production will decrease relative to the production of varieties with $\tau \geq T_1$. In the short version we ignore this adjustment. By doing so we overestimate the production by relatively expensive varieties and, as a result, overestimate the increase in price $P^{top,old}_{a,t}$ after the second shock.

Notice that this is only a problem for the medium distance between T_1 and T_2. For the short distance, at time T_2 almost all varieties are similar (the design of almost all varieties is not adjusted to the shock that happened at T_1), and their behavior is very close to the average. In the long run, all varieties are again similar (all designs are adjusted to the shock of T_1) and their behavior is close to average.

Notice also that approximation does not mean that, following a price shock, firms make no adjustments in demand for particular varieties. First, even when we make the approximation, we distinguish between new and old varieties. Because new varieties are, in general, more productive and their intensity parameters ($\alpha^{ch}_{a,t}$'s) are better adjusted to current market prices, they will have lower manufacturing costs and higher demand than old varieties (see equations (B.3) and (5)). Second, the demand for old varieties will also react to changes in factor prices; these will be transmitted through the change in the average price index $P^{top,old}_{a,t}$. Approximation does imply, however, that changes in all old varieties will be driven by the same price index $P^{top,old}_{a,t}$ (see equation (B.17)).

The second important effect that we ignore by approximating is related to the computation of the average price index for old varieties in the nested production function. In the long version, we compute this index by weighting price $p^{ch}_{a,\tau,t}$ with its intensity $\alpha^{ch}_{a,\tau}$ for each variety and then summing over all old varieties (see equations (B.2) and (B.11)). In general, in nested production function, the price of inputs can vary across varieties. Similarly, the intensity of use can also vary across varieties because of technological progress and because of different firms' choices of direction of technological change in different points in time. In the long version we consider this variation and we account for the potential correlation between input prices and input intensities across old varieties.

In the short version we compute the index by multiplying average intensity of use of input ch, $\alpha^{ch,old}_{a,t}\left(\lambda^{ch,old}_{a,t}\right)^{\sigma^{par}_{a}-1}$, by average price of input ch, $P^{ch,old}_{a,t}$ (see equation (B.14))[14]. This is correct only if the correlation between $\alpha^{ch}_{a,\tau}$ and $p^{ch}_{a,\tau,t}$ is zero. This assumption fails, for instance, if varieties that are characterised by higher cost of manufacturing of input ch, have lower intensity use of that input. In that case, to compute the true average price we would need to weight each higher price $p^{ch}_{a,\tau,t}$ with a lower intensity $\alpha^{ch}_{a,\tau}$ before summing over all old varieties. When we ignore this weighting, we are miscomputing the average price of old varieties (in the case of a negative correlation between $p^{ch}_{a,\tau,t}$ and $\alpha^{ch}_{a,\tau}$, we are over-estimating this price).

The advantage of the short version is a significantly improved tractability with respect to the long version. All key productivity and intensity parameters can be expressed simply as a linear dynamic equation that resembles the AR1 process (see for instance equation (B.25) below). The short version also reduces the number of effects that drive the results, which allows for a better exposition of effects driven purely by the choices of technology firms. In addition, the short version allows us to significantly reduce the number of equations and ease the computational burden of the model.

[14] Additionally, we ignore the role of the exponents $\left(1-\sigma^{par}_{a}\right)$. We discuss this in Appendix A5

In the next subsection, we describe the set of implementable equations for the short version of the model. The implementation of the long version is described in Appendix A3.

A2.5 Recalibration

Monopolistic competition requires the introduction of a wedge between unit costs of production and output prices, which necessitates changes in the calibration of share parameters in the model.

We assume that the compensation for capital by sector, as reported in national statistics, consists of two components: compensation for physical capital and the operating surplus that firms allocate to finance their R&D activities. To account for this, we can make a simple adjustment at the beginning of the calibration code. Specifically, the values for the compensation of physical capital in the calibration year should be calculated as the difference between the total compensation for capital in national statistics and the value of the operating surplus. The operating surplus is defined as the revenue of the sector times the markup rate.

Additionally, for one of the actors receiving capital income, we need to replace the value of that income with the difference between the income reported in national statistics and the operating surplus. Since this reduces the income of that actor, we must also reduce their savings by the value of the operating surplus. The total demand for investment goods does not change.

The intuition behind these changes is that each sector transfers a portion of its revenue (the operating surplus) to finance purchase of investment goods for R&D purposes, bypassing other actors in the economy. An example of the recalibration code is provided in the supplementary material.

A2.6 Implemented equations

We implement the module in seven blocks of equations, as described below. The detailed derivations are provided in Appendix A3. Here we limit our description to the outcome of these manipulations.

Demand and manufacturing costs of varieties

Equation (B.18) determines the manufacturing costs for each nest.

$$P_{a,t}^{par,v} = \left(\sum_{ch} \alpha_{a,t}^{ch,v} \left(P_{a,t}^{ch,v} / \lambda_{a,t}^{ch,v} \right)^{1-\sigma_a^{par,v}} \right)^{\frac{1}{1-\sigma_a^{par,v}}} \tag{B.18}$$

where $v \in \{new, old\}$

Equations (B.19) and (B.20) describe the demand for new and old varieties. The lower costs of production for new varieties will increase their contribution to the final output production of the sector. The demand also depends on the number of new and old varieties. An increase in innovation rate will result in an increase in the contribution of new varieties. Finally, the contribution of new varieties is higher when they are more productive, that is, when they have a larger $\left(\lambda_{a,t}^{top,new}\right)^{\sigma_a^{var}-1}$.

$$\begin{gathered} Q_{a,t}^{top,new} = \\ \left(IMP_{a,t} + INV_{a,t}\right) M_{a,t-1} \left(\lambda_{a,t}^{top,new}\right)^{\sigma_a^{var}-1} * \\ \left(\frac{P_{a,t}}{(1+\gamma) P_{a,t}^{top,new}} \right)^{\sigma_a^{var}} Q_{a,t} \end{gathered} \tag{B.19}$$

and

$$\begin{gathered} Q_{a,t}^{top,old} = \\ \left(1 - IMP_{a,t} - 2\kappa INV_{a,t}\right) M_{a,t-1} \left(\lambda_{a,t}^{top,old}\right)^{\sigma_a^{var}-1} * \\ \left(\frac{P_{a,t}}{(1+\gamma) P_{a,t}^{top,old}} \right)^{\sigma_a^{var}} Q_{a,t} \end{gathered} \tag{B.20}$$

where κ is a parameter that takes the value of $\kappa = 1$ if $\sigma_a^{var} < 1$ and $\kappa = 0$ if $\sigma_a^{var} > 1$.

The demand for inputs can be expressed using the standard rule for nested CES production functions:

$$Q_{a,t}^{ch,v} = \alpha_{a,t}^{ch,v} \left(\lambda_{a,t}^{ch,old}\right)^{\sigma_a^{par,v}-1} \left(\frac{P_{a,t}^{ch,v}}{P_{a,t}^{par,v}}\right)^{-\sigma_a^{par,v}} Q_{a,t}^{par,v} \tag{B.21}$$

.

Cost of sectoral output

The manufacturing costs of sector a can be calculated from the total cost of new and old varieties scaled by the markup:

$$P_{a,t}Q_{a,t} = (1+\gamma)\left(Q_{a,t}^{top,old}P_{a,t}^{top,old} + Q_{a,t}^{top,new}P_{a,t}^{top,new}\right) \tag{B.22}$$

We can then use equations (B.19), (B.20) and (B.22) to express the price as a function of $M_{a,t}$ and productivity parameters, $\lambda_{a,t}^{top,new}$ and $\lambda_{a,t}^{top,old}$:

$$P_{a,t} = (1+\gamma)\, M_{a,t-1}^{\frac{1}{1-\sigma_a^{var}}} \left((1-((1-\rho)+2\kappa\rho)\, INN_{a,t}) \left(\frac{P_{a,t}^{top,old}}{\lambda_{a,t}^{top,old}}\right)^{1-\sigma_a^{var}} + INN_{a,t}\left(\frac{P_{a,t}^{top,new}}{\lambda_{a,t}^{top,new}}\right)^{1-\sigma_a^{var}} \right)^{\frac{1}{1-\sigma_a^{var}}}$$

Hence, a decline in manufacturing cost at the sectoral level will be driven by (i) growth in productivity, $\lambda_{a,t}^{top,new}$ and $\lambda_{a,t}^{top,old}$ and (ii) growth in the total number of varieties, $M_{a,t-1}$, when $\sigma_a^{var} > 1$ or a decline in $M_{a,t-1}$, when $\sigma_a^{var} < 1$.

Productivity path

Given the settings specified with equations (17) and (18), all new varieties are simply assigned the same productivity as the frontier, which grows at an exogenous rate of g. Thus, the path of $\lambda_{a,t}^{top,new}$ can be summarised through equations (B.23) and (B.24):

$$\lambda_{a,t}^{top,new} = frontier_t \tag{B.23}$$

$$frontier_t = (1+g)\, frontier_{t-1} \tag{B.24}$$

Meanwhile, under the assumptions described in Section 3.2, the equation describing the path of $\lambda_{a,t}^{top,old}$, can be reduced to a simple relation that resembles the AR1 process:

$$\left(1+(1-2\kappa)\,INV_{a,t-1}\right)\left(\lambda^{top,old}_{a,t}\right)^{\sigma^{par}_{a}-1} = \left(IMP_{a,t-1}+INV_{a,t-1}\right)\left(\lambda^{top,new}_{a,t-1}\right)^{\sigma^{par}_{a}-1} + \left(1-IMP_{a,t-1}-2\kappa INV_{a,t-1}\right)\left(\lambda^{top,old}_{a,t-1}\right)^{\sigma^{par}_{a}-1} \tag{B.25}$$

Evaluating this equation in reverse, the productivity measure, $\lambda^{top,old}_{a,t}$, can also be expressed as the weighted average of the productivity of individual variety cohorts weighted by cohort size.

Finally, using equations (3), we can express the growth (decline) of the total number of varieties, $M_{a,t}$, as a function of the invention rate:

$$\frac{M_{a,t}-M_{a,t-1}}{M_{a,t-1}} = \begin{cases} INV_{a,t} & \text{if } \sigma^{var}_{a} \geq 1 \\ -INV_{a,t} & \text{if } \sigma^{var}_{a} < 1 \end{cases} \tag{B.26}$$

.

Direction

As explained in Section 2.6, firms can choose the factor intensity of production for new varieties.

In the case of general nested CES production function, $\tilde{\sigma}^{top}_{a}$ and σ^{top}_{a} in (20) can be replaced with $\tilde{\sigma}^{par}_{a}$ and σ^{par}_{a}. Then the expression in (20) implies that in the year of invention (i.e., for $t=\tau$),

$$q^{ch}_{a,t,t} = b_{a,ch,t}\left(\frac{p^{ch}_{a,t,t}}{p^{par}_{a,t,t}}\right)^{-\tilde{\sigma}^{par}_{a}}\left(\lambda^{ch}_{a,t}\right)^{\tilde{\sigma}^{par}_{a}-1} q^{par}_{a,t,t} \tag{B.27}$$

and

$$p^{par}_{a,t,t} = \left(\sum_{ch} b_{a,ch,t}\left(\frac{P^{ch}_{a,t.t}}{\lambda^{ch}_{a,t}}\right)^{1-\tilde{\sigma}^{par}_{a}}\right)^{\frac{1}{1-\tilde{\sigma}^{par}_{a}}} \tag{B.28}$$

,

where $b_{a,ch,t} = (\beta_{a,ch,t})^{\frac{1-\tilde{\sigma}^{par}_{a}}{1-\sigma^{par}_{a}}}$.

In the model, we use equation (B.27) to set firms' demand for inputs for new varieties. We must thus set the factor intensity coefficient to $b_{a,ch,t}$:

$$\alpha^{new}_{a,ch,t} = b_{a,ch,t} \tag{B.29}$$

.

Next, notice that given equations (20) and (B.27), $\alpha_{a,t}^{ch}$ can be expressed as:

$$\alpha_{a,t}^{ch}\left(\lambda_{a,t}^{ch}\right)^{\sigma_a^{par}-1} = \left(\frac{p_{a,t,t}^{ch}}{p_{a,t,t}^{par}}\right)^{\sigma_a^{par}} \frac{q_{a,t,t}^{ch}}{q_{a,t,t}^{par}} \tag{B.30}$$

In the case when ch is an intermediate product, its price, $p_{a,t,t}^{ch}$, is determined by the optimal choice of $\alpha_{a,t}^{ch}$ in the lower nests (for example, if the intermediate good $ch1$ is produced with inputs $ch2 \in nest\,(ch1)$, then $p_{a,t,t}^{ch1}$ depends on the choice of $\alpha_{a,t}^{ch2}$).

Given the definitions we introduced in the subsection 'Additional definitions', we can express equation (B.30) as:

$$\alpha_{a,t}^{ch} = \left(\lambda_{a,t}^{ch}\right)^{1-\sigma_a^{par}} \left(\frac{P_{a,t}^{ch,new}}{P_{a,t}^{par,new}}\right)^{\sigma_a^{par}} \frac{Q_{a,t}^{ch,new}}{Q_{a,t}^{par,new}} \tag{B.31}$$

.

Finally, we can use this information to update the factor intensity of old varieties:

$$\begin{aligned}\left(1+\left(1-2\kappa\right)INV_{a,t-1}\right)\alpha_{a,t}^{ch,old} = \\ INN_{a,t-1}\alpha_{a,t-1}^{ch} \\ +\left(1-IMP_{a,t-1}-2\kappa INV_{a,t-1}\right)\alpha_{a,t-1}^{ch,old}\end{aligned} \tag{B.32}$$

Speed of innovation

The speed of innovation is determined endogenously. The innovation rate is determined by the research production function (equation (15)):

$$INN_{a,t}^{exp} = R_{a,t}^{\phi^{RD}}\left(\frac{INN_{a,0}}{R_{a,0}^{\phi^{RD}}}\right) \tag{B.33}$$

.

where $INN_{a,0}$ and $R_{a,0}^{\phi^{RD}}$ are the innovation rate and research intensity in the base year, which are exogenous.

The relation between $INN_{a,t}^{exp}$, $INN_{a,t}$, $IMP_{a,t}$ and $INV_{a,t}$ is specified by equations (13) and (16) and can be summarized as follows:

$$INN_{a,t} = INN_{a,t-1}^{exp} \tag{B.34}$$

$$INV_{a,t} = \rho INN_{a,t} \tag{B.35}$$

$$IMP_{a,t} = (1-\rho)\, INN^{exp}_{a,t} \tag{B.36}$$

.

Following the logic in Section 2.4, firms invest in costly R&D because they expect a stream of profit that translates into a positive value for each blueprint. The return to R&D decreases with an increase in research intensity for two reasons: reduction of the R&D success rate at the sectoral level (captured by the term $R^{1-\phi^{RD}}_{a,t}$ in equation (12)) and the business-stealing effect (captured by the term $\left(1 - IMP^{exp}_{a,t} - 2\kappa INV^{exp}_{a,t}\right)$). The equilibrium level of research intensity is determined by the zero-profit condition (derived from equation (12)) that equates return to R&D with its cost.

$$R^{1-\phi^{RD}}_{a,t} p_{RD,t}\Phi_{a,t} = \frac{\Gamma\gamma p^{top}_{a,t,t} q^{top}_{a,t,t}}{(1+i^{RD})\,/\,\left(1+g^{exp}_{a,t}\right) - \left(1 - IMP^{exp}_{a,t} - 2\kappa INV^{exp}_{a,t}\right)} \tag{B.37}$$

R&D finance equations

Equation (12) implies that the expected profit from all R&D projects across all sectors is given by:

$$\sum_a R^{\phi^{RD}}_{a,t}\Gamma \frac{\gamma p^{top}_{a,t,t} q^{top}_{a,t,t}}{(1+i^{RD})\,/\,\left(1+g^{exp}_{a,t}\right) - \left(1 - IMP^{exp}_{a,t} - 2\kappa INV^{exp}_{a,t}\right)}$$

As noted earlier, we assume that there is no perfect foresight. Firms base their R&D decisions on information about current profits in their sector, the expected growth rate (an exogenous parameter), and the obsolescence rate.

Actual profit from all past innovations is given by:

$$\sum_a \gamma \left(XPV^{old}_{a,t} PXV^{old}_{a,t} + XPV^{new}_{a,t} PXV^{new}_{a,t}\right)$$

.

As firms are myopic, their expected profits may differ from the actual profits. Any discrepancy is covered by the 'enterprises' sector, which acts as an insurance mechanism. We assume that no actor in the economy has perfect foresight, and that the value of the blueprints reflects the best available knowledge regarding expected profits. In the model, the enterprises sector collects all capital rents from sectoral firms and transfers them to households. The difference between actual and expected R&D profits (which could be either positive or negative) is added to these capital transfers. In the MANAGE model, this is implemented by adjusting the equation that determines the income of enterprises:

$$\text{Enterprises Income}_t = \text{Compensation on capital}$$

$$+\sum_a \gamma \left(XPV_{a,t}^{old} PXV_{a,t}^{old} + XPV_{a,t}^{new} PXV_{a,t}^{new} \right)$$

$$-\sum_a P_{RD,t} \Phi_{a,t} R_{a,t} \tag{B.38}$$

R&D Demand for Commodities

We assume that R&D activity requires the use of investment goods, which are purchased by technology firms. This assumption can be easily modified. Alternatively, we could assume that R&D requires the use of skilled labour or some R&D-specific factor of production ('entrepreneurship'), which could be supplied inelastically by households. In this case, R&D would not compete for resources with production firms. R&D intensity would remain endogenous at the sectoral level, but fixed at the aggregate level.

If R&D requires only the use of investment goods, the cost of increasing R&D intensity by one unit is: $P_{RD,t} = p_t^{investment}$, where $p_t^{investment}$ is the price of the investment good. The total value of investment allocated to R&D projects across all sectors is then:

$$INV_t^R \equiv \sum_a p_t^{investment} \Phi_{a,t} R_{a,t}$$

The price of investment good, $p_t^{investment}$, is determined by the cost of manufacturing faced by a producer of investment goods, as in the standard CGE framework.

The demand for investment dedicated to R&D projects is added to the demand for investment dedicated to accumulation of physical capital. The supply of investment is determined by the amount of available savings. In the case of MANAGE model, this is accounted for in the investment-saving balance equation:

$$\text{Total Investment}_t = \text{sum of domestic and foreign savings}$$

$$+ \text{sum of stock changes} + INV_t^R \tag{B.39}$$

Total investment variable is then used to determine total demand for investment goods. However, we must account for the fact that investment in R&D activity does not contribute to the accumulation of physical capital. Therefore, it must be deducted from the investment in the capital accumulation equation. In the MANAGE model, this correction is done as follows:

$$\text{Capital Stock}_t = (1 - \text{depreciation}) * \text{Capital Stock}_{t-1}$$

$$\frac{\text{Total\_Investment} - INV_t^R}{p_t^{investment}} \tag{B.40}$$

Appendix C. Derivation of implementable equations

A3.1 Additional definitions for the long version

In order to facilitate the exposition and implementation of the long version of the model, we introduce some additional definitions. First, we define variable $\hat{Q}^{old,top}_{a,t}$ as:

$$\hat{Q}^{old,top}_{a,t} = \frac{\sum_{\tau}^{t-1} p^{top}_{a,\tau,t} q^{top}_{a,\tau,t}}{P^{old,top}_{a,t}} \tag{C.1}$$

.

In addition, we define a variable, $\hat{q}^{top}_{a,\tau,t}$, which represents the output of varieties invented at time τ, relative to the value of $\hat{Q}^{old,top}_{a,t}$:

$$\hat{q}^{top}_{a,\tau,t} = N_{a,\tau,t} q^{top}_{a,\tau,t} / \hat{Q}^{old,top}_{a,t} \tag{C.2}$$

.

Further, $\hat{q}^{ch}_{a,\tau,t}$, represents the demand for the intermediate good or factor of production, ch, of varieties invented at time τ, relative to the demand generated by all varieties invented before t, $Q^{old,ch}_{a,t}$

$$\hat{q}^{ch}_{a,\tau,t} = N_{a,\tau,t} q^{ch}_{a,\tau,t} / Q^{old,ch}_{a,t} \tag{C.3}$$

.

A3.2 Demand and manufacturing costs of variety outputs and inputs

Manufacturing cost

*The long version

For the new varieties, the relation can be established using equation (B.28):

$$P_{a,t}^{par,new} = \left(\sum_{ch} b_{a,ch,t} \left(P_{a,t}^{ch,new} / \lambda_{a,t}^{ch,new} \right)^{1-\tilde{\sigma}_a^{par}} \right)^{\frac{1}{1-\tilde{\sigma}_a^{par}}} \tag{C.4}$$

.

We can then use equation (B.6) to obtain the same expression as in (B.18).

For the old varieties, $P_{a,t}^{top,old}$ is defined in equation (B.7):

$$\frac{P_{a,t}^{top,old}}{\lambda_{a,t}^{top,old}} = \left(\sum_{\tau}^{t-1} \frac{N_{a,\tau,t}}{(1-x_{a,t})\, M_{a,t-1}} \left(p_{a,\tau,t}^{top} / \lambda_{a,\tau,t}^{top} \right)^{1-\sigma_a^{var}} \right)^{\frac{1}{1-\sigma_a^{var}}} \tag{C.5}$$

.

In the long version, $p_{a,\tau,t}^{par}$ (including $p_{a,\tau,t}^{top}$), are determined using equation (B.2):

$$p_{a,\tau,t}^{par} = \left(\sum_{ch} \alpha_{a,ch,\tau} \left(\lambda_{a,\tau,t}^{ch} \right)^{\sigma_a^{par}-1} \left(p_{a,\tau,t}^{ch} \right)^{1-\sigma_a^{par}} \right)^{\frac{1}{1-\sigma_a^{par}}} \tag{C.6}$$

.

*The short version

In the short version, the derivation of $P_{a,t}^{par,new}$ follows exactly the same derivations as in the long version. $P_{a,t}^{par,old}$ (including $P_{a,t}^{top,old}$), can be derived using equations (B.7) and (B.13)

As a result, we obtain equation (B.18) (restated below for convenience):

$$P_{a,t}^{par,v} = \left(\sum_{ch} \alpha_{a,ch,t}^{v} \left(P_{a,t}^{ch,v} / \lambda_{a,t}^{ch,v} \right)^{1-\sigma_a^{par,v}} \right)^{\frac{1}{1-\sigma_a^{par,v}}}$$

.

Demand for outputs

*The long version

$Q_{a,t}^{top,new}$ can be expressed using equations (5), (B.4), (7), (B.3) and (3). We obtain equation (B.19) as follows:

$$Q_{a,t}^{top,new} = INN_{a,t} M_{a,t-1} \left(\lambda_{a,t}^{top,new} \right)^{\sigma_a^{var}-1} \left(\frac{P_{a,t}}{(1+\gamma)\, P_{a,t}^{top,new}} \right)^{\sigma_a^{var}} Q_{a,t} \tag{C.7}$$

.

In order to express $\hat{Q}^{top,old}_{a,t}$, we first use equations (C.1) and (5), (7) to obtain:

$$P^{top,old}_{a,t}\hat{Q}^{top,old}_{a,t} = \left(1 - IMP^{exp}_{a,t} - 2\kappa INV^{exp}_{a,t}\right) M_{a,t-1} \left(\frac{P_{a,t}}{1+\gamma}\right)^{\sigma^{var}_a} *$$

.

$$*Q_{a,t} \sum_{\tau}^{t-1} \frac{N_{a,\tau,t}}{\left(1 - IMP^{exp}_{a,t} - 2\kappa INV^{exp}_{a,t}\right) M_{a,t-1}} \left(p^{top}_{a,\tau,t}/\lambda^{top}_{a,\tau,t}\right)^{1-\sigma^{var}_a}$$

.

Then, we use equation (C.5) to obtain

$$\hat{Q}^{top,old}_{a,t} = \left(1 - IMP^{exp}_{a,t} - 2\kappa INV^{exp}_{a,t}\right) M_{a,t-1}*$$

$$* \left(\lambda^{top,old}_{a,t}\right)^{\sigma^{var}_a-1} \left(\frac{P_{a,t}}{(1+\gamma)\, P^{top,old}_{a,t}}\right)^{\sigma^{var}_a} Q_{a,t} \tag{C.8}$$

.

*The Short version

In the short version, the derivation of $Q^{top,new}_{a,t}$ follows exactly the same derivations as in the long version. Hence equation (C.7) is also implemented.

$Q^{top,old}_{a,t}$ can be derived using equations (B.9) and (B.17):

$$Q^{top,old}_{a,\tau,t} = \left(1 - IMP^{exp}_{a,t} - 2\kappa INV^{exp}_{a,t}\right) M_{a,t-1}$$

$$* \left(\frac{P_{a,t}}{(1+\gamma)\, P^{top,old}_{a,t}}\right)^{\sigma^{var}_a} \left(\lambda^{top,old}_{a,t}\right)^{\sigma^{var}_a-1} Q_{a,t}$$

..

Demand for intermediate goods

The long version

In the case of new varieties, the demand for intermediate goods for varieties is determined using equations (B.27), (B.5), (B.4) and (B.6)

$$Q^{ch,new}_{a,t} = \alpha^{ch,new}_{a,t} \left(\lambda^{ch,new}_{a,t}\right)^{\tilde{\sigma}^{par}_a-1} \left(\frac{P^{ch,new}_{a,t}}{P^{par,new}_{a,t}}\right)^{-\tilde{\sigma}^{par}_a} Q^{par,new}_{a,t} \tag{C.9}$$

.

In the case of old varieties, for all $par \neq top$, the demand is determined using equations (C.3) and (B.9):

$$Q_{a,t}^{ch,old} = \alpha_{a,ch,t}^{ch,old} \left(\lambda_{a,t}^{ch,old}\right)^{\sigma_a^{par}-1} \left(\frac{P_{a,t}^{ch,old}}{P_{a,t}^{par,old}}\right)^{-\sigma_a^{par}} Q_{a,t}^{par,old} \tag{C.10}$$

,

where

$$\hat{\alpha}_{a,ch,t}^{old} = \sum_{\tau}^{t-1} \left[\alpha_{a,\tau}^{ch} \left(\frac{\lambda_{a,\tau,t}^{ch}}{\lambda_{a,t}^{ch,old}}\right)^{\sigma_a^{par}-1} \hat{q}_{a,\tau,t}^{par} \left(\frac{p_{a,\tau,t}^{ch}}{p_{a,\tau,t}^{par}}\right)^{-\sigma_a^{par}} \left(\frac{P_{a,t}^{ch,old}}{P_{a,t}^{par,old}}\right)^{\sigma_a^{par}} \right] \tag{C.11}$$

.

For $par = top$, we follow exactly the same steps, except that we replace $Q_{a,t}^{par,old}$ with $\hat{Q}_{a,t}^{par,old}$:

$$Q_{a,t}^{ch,old} = \hat{\alpha}_{a,t}^{ch,old} \left(\lambda_{a,t}^{ch,old}\right)^{\sigma_a^{par}-1} \left(\frac{P_{a,t}^{ch,old}}{P_{a,t}^{par,old}}\right)^{-\sigma_a^{par}} \hat{Q}_{a,t}^{par,old} \tag{C.12}$$

Notice that we can express $\hat{q}_{a,\tau,t}^{par}$ using equations (C.3), (5) and (C.8):

$$\hat{q}_{a,\tau,t}^{top} = \frac{N_{a,\tau,t}}{(1-x_{a,t})\, M_{a,t-1}} \left(\frac{P_{a,t}^{top,old}}{p_{a,\tau,t}}\right)^{\sigma_a^{var}} \left(\frac{\lambda_{a,t}^{top,old}}{\lambda_{a,\tau,t}^{top}}\right)^{1-\sigma_a^{var}} \tag{C.13}$$

for the top nest and, using equation (C.3),

$$\hat{q}_{a,\tau,t}^{ch} = \frac{\alpha_{a,ch,\tau} \left(\lambda_{a,\tau,t}^{ch}\right)^{\sigma_a^{par}-1}}{\alpha_{a,ch,t}^{old}} \left(\frac{p_{a,\tau,t}^{ch}}{p_{a,\tau,t}^{par}}\right)^{-\sigma_a^{par}} \left(\frac{p_{a,t}^{ch,old}}{p_{a,t}^{par,old}}\right)^{\sigma_a^{par}} \hat{q}_{a,\tau,t}^{par} \tag{C.14}$$

for the nests with $par \neq top$.

The short version

In the short version, the intermediate demand for new varieties is derived in the same way as in the long version. In the case of old varieties, the demand is obtained by applying definitions (B.9) and (B.12) to equation (B.16):

$$Q_{a,t}^{ch,old} = \alpha_{a,t}^{ch,old} \left(\lambda_{a,t}^{ch,old}\right)^{\sigma_a^{par}-1} \left(\frac{P_{a,t}^{ch,old}}{P_{a,t}^{par,old}}\right)^{-\sigma_a^{par}} Q_{a,t}^{par,old}$$

In this equation, $\alpha_{a,t}^{ch,old}$ acts as an average intensity of use of factor ch across old varieties.

Price of sectoral output

The long version

Once the demand functions are established, we can determine the price of output, $P_{a,t}$, as follows:

$$P_{a,t}Q_{a,t} = (1+\gamma)\left(Q_{a,t}^{top,old}P_{a,t}^{top,old} + \hat{Q}_{a,t}^{top,new}P_{a,t}^{top,new}\right) \tag{C.15}$$

.

This will set exactly the same price as equation (6).

The Short version

In the short version, the output price is determined by equation (B.15). Combined with equations (B.3), (B.4) and (B.9), gives

$$P_{a,t} = \frac{P_{a,t}^{top,new}Q_{a,t}^{top,new} + \sum_{\tau}^{t-1} P_{a,t}^{top,old}Q_{a,t}^{top,old}}{Q_{a,t}}$$

.

Hence, the price can be again determined as follows:

$$P_{a,t}Q_{a,t} = (1+\gamma)\left(Q_{a,t}^{top,old}P_{a,t}^{top,old} + Q_{a,t}^{top,new}P_{a,t}^{top,new}\right).$$

A3.3 Productivity path

Equations describing the path of $\lambda_{a,t}^{top,new}$ follow directly from equations (B.23) and (B.24):

$$\lambda_{a,t}^{top,new} = frontier_t \tag{C.16}$$

$$frontier_t = (1+g)\, frontier_{t-1} \tag{C.17}$$

The updating rule for $\lambda_{a,t}^{top,old}$ can be derived from from (B.8) and (3) as follows:

$$\begin{aligned}(1+INV_{a,t-1})\left(\lambda_{a,t}^{top,old}\right)^{\sigma_a^{var}-1} &= (INV_{a,t-1}+IMP_{a,t-1})\left(\lambda_{a,t-1}^{top,new}\right)^{\sigma_a^{var}-1} \\ &\quad + \left(1-IMP_{a,t-1}^{exp}-2\kappa INV_{a,t-1}^{exp}\right)\left(\lambda_{a,t-1}^{top,old}\right)^{\sigma_a^{var}-1}\end{aligned}$$

A3.4 Direction

As explained in Section 2, firms can choose the factor intensity of production for new varieties. Given firm choices regarding inputs $\frac{Q^{ch,new}_{a,\tau}}{Q^{par,new}_{a,\tau}}$, we can recover firms choices of $\alpha^{ch}_{a,\tau}$ (see equation (B.30)); alphas chosen for a variety τ do not change over time. They are determined using equation (B.30):

$$\alpha^{ch}_{a,\tau}\left(\lambda^{ch}_{a,\tau}\right)^{\sigma^{par}_a-1}=\left(\frac{PXFAC^{ch,new}_{a,\tau}}{PXFAC^{par,new}_{a,\tau}}\right)^{\sigma^{par}_a}\frac{XFAC^{ch,new}_{a,\tau}}{XFAC^{par,new}_{a,\tau}} \tag{C.18}$$

.

$$\alpha^{ch}_{a,\tau}=\left(\lambda^{ch}_{a,\tau}\right)^{1-\sigma^{par}_a}\left(\frac{P^{ch,new}_{a,\tau}}{P^{par,new}_{a,\tau}}\right)^{\sigma^{par}_a}\frac{Q^{ch,new}_{a,\tau}}{Q^{par,new}_{a,\tau}} \tag{C.19}$$

Finally, we can use this information to update the factor intensity of the old varieties. The updating rule is different for the short and long versions. We describe them separately below:

The long version

The factor intensity parameter was derived above (see equation (C.11)) and is restated here for convenience:

$$\hat{\alpha}^{old}_{a,ch,t}=\sum_{\tau}^{t-1}\left[\alpha^{ch}_{a,\tau}\left(\frac{\lambda^{ch}_{a,\tau,t}}{\lambda^{ch,old}_{a,t}}\right)^{\sigma^{par}_a-1}\hat{q}^{par}_{a,\tau,t}\left(\frac{p^{ch}_{a,\tau,t}}{p^{par}_{a,\tau,t}}\right)^{-\sigma^{par}_a}\left(\frac{P^{ch,old}_{a,t}}{P^{par,old}_{a,t}}\right)^{\sigma^{par}_a}\right]$$

.

The short version

In the short version, the updating rule can be derived directly from (B.12) and (3) as follows:

$$\alpha^{ch,old}_{a,t}\left(\lambda^{ch,old}_{a,t}\right)^{\sigma^{par}_a-1}=\frac{INV_{a,t-1}+IMP_{a,t-1}}{1+INV_{a,t-1}}\alpha^{ch}_{a,t-1}\left(\lambda^{ch}_{a,t-1}\right)^{\sigma^{par}_a-1}$$

$$+\frac{1-IMP^{exp}_{a,t-1}-2\kappa INV^{exp}_{a,t-1}}{1+INV_{a,t-1}}\alpha^{ch,old}_{a,t-1}\left(\lambda^{ch,old}_{a,t-1}\right)^{\sigma^{par}_a-1}$$

Appendix D. Normalization of the initial state of the frontier

In this appendix, we describe the choice of the value of $frontier_0$ as follows. First, we define the average productivity in a sector as

$$\overline{\lambda}_{a,t}^{top} = \sum_{\tau=0}^{t} \frac{N_{a,\tau,t}}{M_{a,t}} \lambda_{a,\tau}^{top} = \frac{N_{a,t,t}}{M_{a,t}} \lambda_{a,t}^{top} + \frac{M_{a,t-1}}{M_{a,t}} \left(1 - IMP_{a,t}\right) \overline{\lambda}_{a,t-1}^{top}$$

.

Second, we set $\overline{\lambda}_{a,0} = 1$. We then set the value of $frontier_0$ at the level that ensures that, at $t = 1$, average productivity grows at the frontier growth rate:

$$\frac{\overline{\lambda}_{a,1}^{top} - \overline{\lambda}_{a,0}^{top}}{\overline{\lambda}_{a,0}^{top}} = \left(\frac{N_{a,t,t}}{M_{a,t}} \lambda_{a,t}^{top} + \frac{M_{a,t-1}}{M_{a,t}} \left(1 - x_{a,t}\right) \overline{\lambda}_{a,0}^{top} \right) - 1 = g$$

,

which requires

$$frontier_0 = \frac{(1+g)\left(1 + INV_{a,0}\right) - \left(1 - IMP_{a,0}\right)}{(1+g)\, INN_{a,0}}$$

Appendix E.

In this appendix we present in details what approximations are sufficient to obtain equations (B.13), (B.15), (B.16) and (B.17).

Suppose we consider a nested CES production function, where $ch0$ is an input for $ch1$, $ch1$ is an input for $ch2$, etc. Then, the demand for the bottom input, $ch0$ can be expressed as:

$$q_{a,\tau,t}^{ch0} = \alpha_{a,\tau}^{ch0} \left(p_{a,\tau,t}^{ch0}\right)^{-\sigma_a^{ch1}} \left(\lambda_{a,\tau,t}^{ch0}\right)^{\sigma_a^{ch1}-1} \left(p_{a,\tau,t}^{ch1}\right)^{\sigma_a^{ch1}-\sigma_a^{ch2}} \alpha_{a,\tau}^{ch1} \left(\lambda_{a,\tau,t}^{ch1}\right)^{\sigma_a^{ch2}-1} [...] *$$

$$* \left(p_{a,\tau,t}^{top}\right)^{\sigma_a^{top}-\sigma_a^{var}} \left(\lambda_{a,\tau}^{top}\right)^{\sigma_a-1} \left(P_{a,t}\right)^{\sigma_a^{var}} Q_{a,t}$$

Under the independence of $\alpha_{a,\tau}^{ch} \left(\lambda_{a,\tau,t}^{ch}\right)^{\sigma_a^{ch}-1}$ across nests:

$$E\left[q_{a,\tau,t}^{ch0}\right] = E\left[\alpha_{a,\tau}^{ch0} \left(p_{a,\tau,t}^{ch0}\right)^{-\sigma_a^{ch1}} \left(\lambda_{a,\tau,t}^{ch0}\right)^{\sigma_a^{ch1}-1}\right] *$$

$$E\left[\left(p_{a,\tau,t}^{ch1}\right)^{\sigma_a^{ch1}-\sigma_a^{ch2}} \alpha_{a,\tau}^{ch1} \left(\lambda_{a,\tau,t}^{ch1}\right)^{\sigma_a^{ch2}-1}\right] [...] *$$

$$*E\left[\left(p_{a,\tau,t}^{top}\right)^{\sigma_a^{top}-\sigma_a^{var}} \left(\lambda_{a,\tau}^{top}\right)^{\sigma_a^{var}-1}\right] \left(P_{a,t}\right)^{\sigma_a^{var}} Q_{a,t}$$

Ignoring the dependence between $\alpha_{a,ch,\tau} \left(\lambda_{a,\tau,t}^{ch}\right)^{\sigma_a^{par}-1}$ and $p_{a,\tau,t}^{par}$,

$$E\left[q_{a,\tau,t}^{ch0}\right] = E\left[\alpha_{a,\tau}^{ch0} \left(\lambda_{a,\tau,t}^{ch0}\right)^{\sigma_a^{ch1}-1}\right] E\left[\left(p_{a,\tau,t}^{ch0}\right)^{-\sigma_a^{ch1}}\right] *$$

$$*E\left[\alpha_{a,\tau}^{ch1} \left(\lambda_{a,\tau,t}^{ch1}\right)^{\sigma_a^{ch2}-1}\right] E\left[\left(p_{a,\tau,t}^{ch1}\right)^{\sigma_a^{ch1}-\sigma_a^{ch2}}\right] [...] *$$

$$*E\left[\left(p_{a,\tau,t}^{top}\right)^{\sigma_a^{top}-\sigma_a^{var}} \left(\lambda_{a,\tau}^{top}\right)^{\sigma_a^{var}-1}\right] \left(P_{a,t}\right)^{\sigma_a^{var}} Q_{a,t}$$

Ignoring that demand is a nonlinear function of prices: $E\left[\left(p_{a,\tau,t}^{ch1}\right)^{\sigma_a^{ch1}-\sigma_a^{ch2}}\right] \approx \left(E\left[p_{a,\tau,t}^{ch1}\right]\right)^{\sigma_a^{ch1}-\sigma_a^{ch2}}$ and $E\left[p_{a,\tau,t}^{ch2}\right] \approx \left(E\left[\left(p_{a,\tau,t}^{ch2}\right)^{1-\sigma_a^{ch2}}\right]\right)^{\frac{1}{1-\sigma_a^{ch2}}} = P_{a,t}^{old,ch2}$:

$$E\left[q^{ch0}_{a,\tau,t}\right] = E\left[\alpha_{a,ch0,\tau}\left(\lambda^{ch0}_{a,\tau,t}\right)^{\sigma^{ch1}_a - 1}\right]\left(\frac{P^{ch0,old}_{a,t}}{P^{ch1,old}_{a,t}}\right)^{-\sigma^{ch1}_a} *$$

$$E\left[\alpha_{a,ch1,\tau}\left(\lambda^{ch1}_{a,\tau,t}\right)^{\sigma^{ch2}_a - 1}\right]\left(\frac{P^{ch1,old}_{a,t}}{P^{ch2,old}_{a,t}}\right)^{-\sigma^{ch2}_a} [...] *$$

$$E\left[\left(\lambda^{top}_{a,\tau}\right)^{\sigma_a - 1}\right]\left(\frac{P^{top,old}_{a,t}}{P_{a,t}}\right)^{\sigma^{top}_a - \sigma^{var}_a} Q_{a,t}$$

Now, notice that

$$E\left[q^{top}_{a,\tau,t}\right] = E\left[\left(\lambda^{top}_{a,\tau}\right)^{\sigma^{var}_a - 1}\right]\left(\frac{P^{top,old}_{a,t}}{P_{a,t}}\right)^{\sigma^{top}_a - \sigma^{var}_a} Q_{a,t}$$

$$E\left[q^{ch2}_{a,\tau,t}\right] = E\left[\alpha_{a,ch2,\tau}\left(\lambda^{ch2}_{a,\tau,t}\right)^{\sigma^{ch3}_a - 1}\right]\left(\frac{P^{ch2,old}_{a,t}}{P^{ch3,old}_{a,t}}\right)^{-\sigma^{ch3}_a} E\left[q^{top}_{a,\tau,t}\right]$$

etc.
Hence, the expression above could be simplified to

$$E\left[q^{ch0}_{a,\tau,t}\right] = E\left[\alpha_{a,ch0,\tau}\left(\lambda^{ch0}_{a,\tau,t}\right)^{\sigma^{ch1}_a - 1}\right]\left(\frac{P^{ch0,old}_{a,t}}{P^{ch1,old}_{a,t}}\right)^{-\sigma^{ch1}_a} E\left[q^{ch1}_{a,\tau,t}\right]$$

which can be written as:

$$Q^{ch0}_{a,t} = E\left[\alpha_{a,ch0,\tau}\left(\lambda^{ch0}_{a,\tau,t}\right)^{\sigma^{ch1}_a - 1}\right]\left(\frac{P^{ch0,old}_{a,t}}{P^{ch1,old}_{a,t}}\right)^{-\sigma^{ch1}_a} Q^{ch1}_{a,t}$$

The same argument could be used for higher nests.
Regarding the prices: $p^{ch1}_{a,\tau,t}$ can be expressed using,

$$\left(p^{ch1}_{a,\tau,t}\right)^{1-\sigma^{ch1}_a} = \sum_{ch0}\left(p^{ch0}_{a,\tau,t}\right)^{1-\sigma^{ch1}_a} \alpha_{a,ch0,\tau}\left(\lambda^{ch0}_{a,\tau,t}\right)^{\sigma^{ch1}_a - 1}$$

$p^{ch2}_{a,\tau,t}$ can be expressed using,

$$\left(p^{ch2}_{a,\tau,t}\right)^{1-\sigma^{ch2}_a} = \sum_{ch1}\alpha_{a,ch1,\tau}\left(\lambda^{ch1}_{a,\tau,t}\right)^{\sigma^{ch2}_a - 1}\left(\sum_{ch0}\alpha_{a,ch0,\tau}\left(\frac{p^{ch0}_{a,\tau,t}}{\lambda^{ch0}_{a,\tau,t}}\right)^{1-\sigma^{ch1}_a}\right)^{\frac{1-\sigma^{ch2}_a}{1-\sigma^{ch1}_a}}$$

etc. Assuming independence of $\alpha_{a,\tau}^{ch}\left(\lambda_{a,\tau,t}^{ch}\right)^{\sigma_a^{ch}-1}$ across nests:

$$\left(P_{a,t}^{ch2,old}\right)^{1-\sigma_a^{ch2}} \equiv E\left[\left(p_{a,\tau,t}^{ch2}\right)^{1-\sigma_a^{ch2}}\right] = \sum_{ch1} E\left[\alpha_{a,ch1,\tau}\left(\lambda_{a,\tau,t}^{ch1}\right)^{\sigma_a^{ch2}-1}\right] E\left[\left(p_{a,\tau,t}^{ch1}\right)^{1-\sigma_a^{ch2}}\right]$$

Now, using again the approximation, $E\left[p_{a,\tau,t}^{ch1}\right] \approx \left(E\left[\left(p_{a,\tau,t}^{ch1}\right)^{1-\sigma_a^{ch1}}\right]\right)^{\frac{1}{1-\sigma_a^{ch1}}}$, we get:

$$\left(P_{a,t}^{ch2,old}\right)^{1-\sigma_a^{ch2}} = \sum_{ch1} E\left[\alpha_{a,ch1,\tau}\left(\lambda_{a,\tau,t}^{ch1}\right)^{\sigma_a^{ch2}-1}\right] P_{a,t}^{ch1}$$

Appendix F. Comparison of the carbon tax simulation results for the short and long version

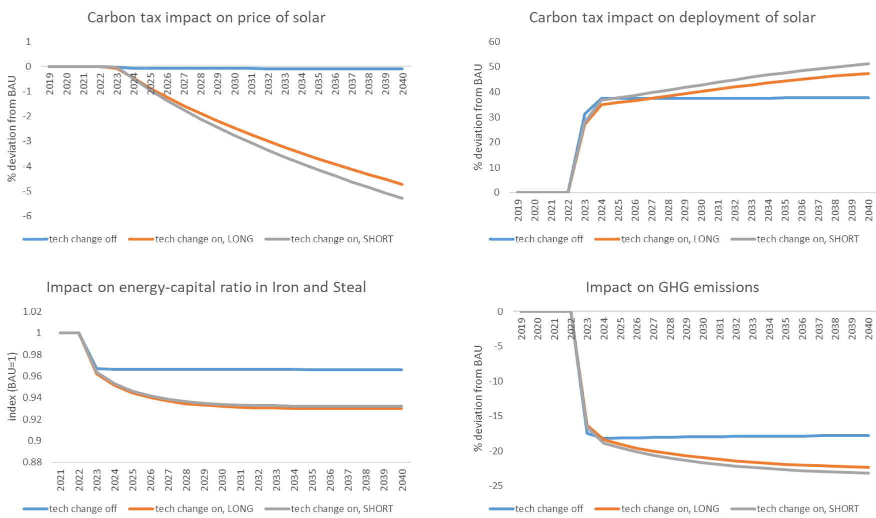

Figure F.1. The impact of a Carbon tax - comparison of the results for the model without (tech change off) endogenous technological change, the long version and the short version of the model with technological change.

Appendix G. Comparison of the carbon tax simulation for different invention-improvement rates

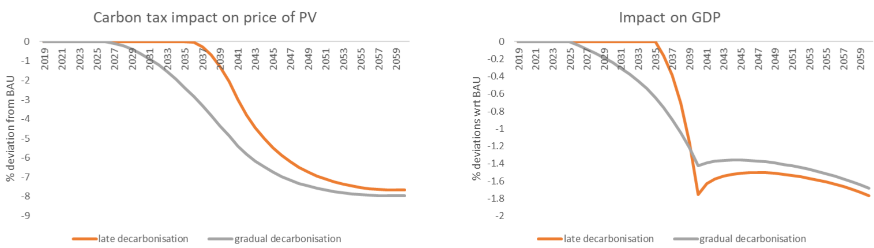

Figure G.1. The impact of gradual and late decarbonizaiton. The simulations were performed using the short version of the model with $\rho = 0.0038$