MAKING CARBON TAXATION A GENERATIONAL WIN WIN

Laurence J. Kotlikoff
Felix Kubler
Andrey Polbin
Jeffrey D. Sachs
Simon Scheidegger

Making Carbon Taxation a Generational Win Win
Laurence J. Kotlikoff, Felix Kubler, Andrey Polbin, Jeffrey D. Sachs, and Simon Scheidegger
NBER Working Paper No. 25760
April 2019
JEL No. F0,F20,H0,H2,H3,J20

ABSTRACT

Carbon taxation has been studied primarily in social planner or infinitely lived agent models, which trade off the welfare of future and current generations. Such frameworks obscure the potential for carbon taxation to produce a generational win-win. This paper develops a large-scale, dynamic 55-period, OLG model to calculate the carbon tax policy delivering the highest uniform welfare gain to all generations. The OLG framework, with its selfish generations, seems far more natural for studying climate damage. Our model features coal, oil, and gas, each extracted subject to increasing costs, a clean energy sector, technical and demographic change, and Nordhaus (2017)'s temperature/damage functions. Our model's optimal uniform welfare increasing (UWI) carbon tax starts at $30 tax, rises annually at 1.5 percent and raises the welfare of all current and future generations by 0.73 percent on a consumption-equivalent basis. Sharing efficiency gains evenly requires, however, taxing future generations by as much as 8.1 percent and subsidizing early generations by as much as 1.2 percent of lifetime consumption. Without such redistribution (the Nordhaus "optimum"), the carbon tax constitutes a win-lose policy with current generations experiencing an up to 0.84 percent welfare loss and future generations experiencing an up to 7.54 percent welfare gain. With a six-times larger damage function, the optimal UWI initial carbon tax is $70, again rising annually at 1.5 percent. This policy raises all generations' welfare by almost 5 percent. However, doing so requires levying taxes on and giving transfers to future and current generations ranging up to 50.1 percent and 10.3 percent of their lifetime consumption. Delaying carbon policy, for 20 years, reduces efficiency gains roughly in half.

Laurence J. Kotlikoff
Department of Economics
Boston University
270 Bay State Road
Boston, MA 02215
and NBER
kotlikoff@gmail.com

Felix Kubler
University of Zurich
Plattenstrasse 32
CH-8032 Zurich
Switzerland
and Swiss Financial Institute
fkubler@gmail.com

Andrey Polbin
The Russian Presidential Academy
of National Economy
and Public Administration
82 Vernadskogo prosp 117517
Moscow Russian Federation
and The Gaidar Institute for Economic Policy
apolbin@gmail.com

Jeffrey D. Sachs
Center for Sustainable Development,
The Earth Institute
Columbia University
535 West 116th Street, MC 4327
New York, NY 10027
and NBER
sachs@columbia.edu

Simon Scheidegger
University of Lausanne Department
of Finance
Extranef 234
CH-1015 Lausanne
Switzerland
simon.scheidegger@gmail.com

1 Introduction

Climate change presents grave risks to current and future generations. The perils include drought, extreme storms, floods, a rise in sea level, intense heat, wildfires, pollution, desertification, the spread of disease, earthquakes, tsunamis, a rise in ocean acidity, the proliferation of insects, and mass extinctions. Anthropogenic global warming, associated with the human release of carbon into the atmosphere, is widely viewed as the primary cause of climate change. According to NASA[1], carbon dioxide in the atmosphere has increased by one third since 1950 and is now at its highest level in 650,000 years. Since 1880, the planet's average temperature has risen by 1.8 degrees, minimum levels of Arctic ice are declining by 12.8 percent per decade, 413 gigatonnes of ice sheets are melting annually, and the sea level may rise 8 feet by 2100. The 2018 U.S. Government's National Climate Assessment projected potentially massive costs to the economy, the eco-system, health, infrastructure, and the environment.[2] Recent estimates put 2100 climate damages at one quarter of global GDP.[3]

Inspired by Sachs (2014), this paper develops a large-scale OLG climate-change model to realistically assess the potential generational win-wins available from carbon taxation. The OLG model appears better suited for studying carbon taxation than the standard frameworks – the social planner and infinitely-lived agent models. In the OLG framework, generations are selfish. Hence, their imposition of negative externalities on future generations comes naturally. The OLG framework also highlights a key point that has been obscured in all but a few analyses of carbon taxation. *Carbon taxation, coupled with appropriate intergenerational redistribution, can make all current and future generations better off. Indeed, it can make them uniformly better off.*

Our OLG model has 55 overlapping generations. A single consumption good (corn) is produced with capital (unconsumed corn), labor, and energy. Energy is clean or dirty. Clean energy is produced using capital, labor, and fixed natural resources (e.g., windy areas), which is proxied by land. Labor and land are in fixed supply. Corn and clean energy experience technical change (TFP growth), which can proceed at permanently different rates. Apart from our energy supply side and OLG preferences, our model matches Nordhaus'(2017). In particular, it adopts Nordhaus' modeling of carbon emissions, temperature change, and temperature-induced economic damage. It also incorporates Nordhaus' projections of global population with appropriate

---

[1]https://climate.nasa.gov/

[2]https://nca2018.globalchange.gov/

[3]http://nymag.com/intelligencer/2017/07/climate-change-earth-too-hot-for-humans.html?gtm=topgtm=top. In one of our models below, which features an extremely large, but, unfortunately, highly plausible damage function, we produce climate damages of this order of magnitude.

assumptions about its distribution by birth cohort.

As in Golosov et al. (2014), we explicitly model dirty energies, in our case coal, oil, and natural gas.[4] Each has finite reserves and each is subject to increasing extraction cost. We use Auerbach and Kotlikoff's (1987) lump-sum redistribution authority (LSRA) to derive the largest uniform (across all current and future generations) welfare increasing (UWI) carbon tax, where welfare changes are measured as compensating consumption differentials. We also present results for two alternative means of distributing efficiency gains from controlling CO_2 emissions. The first allocates all efficiency gains uniformly to current generations (the born). The second allocates efficiency gains uniformly to future generations (the unborn). Depending on the size of damages, optimal carbon policy can depend on how efficiency gains are shared.

Calculating Pareto improvements is standard procedure for determining optimal policy responses to negative externalities. Anthropomorphic climate change is arguably the planet's worst negative externality. Yet ours appears to be the first large-scale study of Pareto improving carbon taxation. Standard integrated assessment models, e.g., Nordhaus (2017), maximize the welfare of a social planner (SP). Other models, e.g., Golosov et al. (2014), examine optimal taxation in infinitely-lived agent (ILA) models. There is no guarantee the SP's "optimal" carbon tax policy will achieve a generational win-win since it pays no heed to the initial intergenerational distribution of welfare. Indeed, the SP would, even in the absence of carbon damage, use carbon policy to redistribute intergenerationally to achieve what, to the SP, is a preferable distribution of welfare across generations. For its part, the infinitely-lived agent model relies, implicitly, on intergenerational altruism (see Barro (1974)). However, such altruism begs the question of why appropriate climate policy is not already in place.[5] The choice of OLG versus ILA frameworks would be of little importance for carbon policy if both frameworks produced identical or very similar policy prescriptions. This is not the case. Indeed, depending on the rate of time preference and the magnitude of climate damages, the two frameworks can produce dramatically different optimal carbon policies with the ILA policy potentially harming some generations to help others.

---

[4]Golosov et al. (2014) combine oil and natural gas.

[5]Yes, infinitely lived dynasties would try to free ride on other dynasties both within and across regions. But, as shown by Kotlikoff (1983) and Bernheim and Bagwell (1988), intermarriage between altruistic dynasties produces altruistic linkages across dynasties, which eliminate the free rider problem. One can also question the ILA framework on empirical grounds. See, in this regard, Altonji et al. (1997), Altonji et al. (1992), Hayashi et al. (1996), Abel and Kotlikoff (1994), and Gokhale et al. (1996).

1.1   Overview of Findings

Under our baseline, business-as-usual (BAU) calibration, climate damages are initially 0.2 percent of output, peaking in year 130 at 7.7 percent of output. Dirty energy represents 97 percent of total energy in year 0, 86 percent in year 50, 46 percent in year 100, and zero percent after year 130, when the value of additional dirty energy extraction exceeds its cost.[6] Relative to their initial stocks, BAU extraction reduces, by year 130, coal, oil, and gas reserves by 70, 75, and 85 percent, respectively. Hence, BAU entails burning most of the planet's fossil fuels. Our BAU simulation also predicts that gas production will rise over the next 20 years and then steadily decline, coal production will rise over the next 45 years and then sharply decline, and oil production will fall for 20 years, rise for the following 40 years, and then gradually fall through year 130. These surprising dynamics reflect the different extraction cost functions for gas, coal, and oil as well as the precise dynamics of the price of energy.

Given the significant uncertainty (see Lontzek et al. (2015)) surrounding climate damage, we also consider larger damage functions where the quadratic coefficient in Nordhaus (2017)'s damage function is multiplied by either 3 or 6. Below, we refer, admittedly loosely, to the three damages functions as 1x, 3x, and 6x damages.[7] As Lontzek et al. (2015) indicates, sensitivity analysis is hardly a perfect substitute for formally modeling climate-damage uncertainty in an OLG setting, which is our major near-term research goal.[8] In our 6x BAU simulation, damages are initially 1 percent of output, peaking at 18.8 percent in year 220. The BAU course of dirty energy production and the extent of reserve exhaustion are quite similar for all three (1x, 3x, and 6x) damage functions.

With our baseline, Nordhaus (2017) damage function, the optimal, uniform welfare increasing (UWI) carbon tax starts at $30, rises at 1.5 percent per year and raises the welfare of all current and future generations by 0.73 percent on a consumption-equivalent basis.[9] Sharing the efficiency gains from our two-part carbon policy (an initial tax plus its annual growth rate) evenly requires taxing future generations by as much as 8.1 percent and subsidizing early generations by as much as 1.2 percent of the present value of their remaining or total lifetime future

---

[6]Our model predicts what it simulates – a simultaneous end to the extraction of coal, oil, and gas.

[7]Nordaus(2017)' specification for damages as share of GDP is $D_t = 1 - \frac{1}{1+\pi_1 T_t^A + \pi_2 (T_t^A)^2}$ with parameters values $\pi_1 = 0$, $\pi_2 = 0.00236$, where the term T_t^A references the Celsius change since 1900 in global mean surface temperature. The reference, below, to an "mx Damage Function" means we set π_2 to 0.00236 times m.

[8]With uncertaintly, the optimal UWI carbon tax will be higher to insure future generations against extreme carbon damage. Indeed, a positive carbon tax would likely be warranted even were carbon damage, on average, negative.

[9]This is relative to consumption under BAU. I.e., each generation's utility gain is equivalent to raising their BAU consumption in each year of their remaining (in the case of current generations) or full lives (in the case of future generations) by 0.73 percent.

consumption. Without such redistribution, the carbon tax constitutes a win-lose policy with those now alive experiencing up to a 0.84 percent welfare loss and those born in the medium run (in this case, year 235) experiencing up to a 7.54 percent welfare gain. Those born in the long run benefit by 3.45 percent.

With 6x damages, the optimal UWI initial carbon tax is $70, again rising at 1.5 percent per year. This policy raises all generations' welfare by almost 5 percent. However, doing so entails lump-sum taxes as high as 50.1 percent and lump sum subsidies as high as 10.3 percent of lifetime consumption. Of note, the carbon tax sans redistribution achieves, in this case, close to a win-win with a) minor welfare losses for current and early generations, but b) gains reaching 45 percent for future generations. The UWI solution is, of course, just one of an infinite number of Pareto paths. For 1x damages, the optimal path of carbon taxation is robust to how efficiency gains are shared. For larger damage functions, the sharing method matters to optimal policy.

The optimal carbon tax materially alters the course of dirty energy production. For example, with 1x damages, coal production is zero for 25 years, positive for 40 years, and zero thereafter. This means the planet needs to cool down before it is efficient to reuse coal. However, the reuse of coal, when it occurs, is small. Indeed, our UWI, 1x damages solution entails burning, over time, only 10 percent of initial coal reserves. As indicated, under business as usual (BAU), i.e., no-policy, 70 percent of initial coal reserves are burnt. With 6x damages, UWI policy entails an immediate end to coal production. Pareto efficient carbon taxation also shortens the period during which dirty energy is produced from 130 years to 100 years in the 1x case and 80 years in the 6x case.

Limiting our carbon policy instruments to two (an initial carbon tax and its annual growth rate), which we do for computational convenience, may understate potential efficiency gains. But, as discussed, we explored whether having separate short- and long-term carbon-tax growth rates permit much higher values of the UWI efficiency gain. The answer is no. The optimal short- and long-term growth rates are identical.

We also calculate optimal two-part carbon policy for our model's sister ILA model – an ILA model that differs from our OLG model due solely to the assumption of intergenerational altruism. The optimal two-part UWI carbon policy in the OLG and ILA models are identical for the *Nordhaus* (2017) baseline 1.5 percent time preference rate and damage function (our 1x case). However, the two sets of policies differ dramatically if the time preference rate differs from 1.5 percent or the damage function is large. Indeed, we provide examples in which optimal carbon taxes a century from now differ by a factor of four between the two frameworks. Moreover, even, in our baseline calibration, where the time path of the optimal carbon tax is the same in both frameworks, the OLG UWI solution calls for intergenerational redistribution,

whereas the ILA model does not.[10]

We also consider whether delaying the implementation of optimal UWI carbon policy for 20 years materially reduces the UWI gain. The answer is yes – by 44.6 percent with baseline damages and by 55.4 percent when our damage function is six times larger.

Finally, we measure the uniform efficiency loss from imposing optimal carbon policy were damages, in fact, zero, which climate skeptics claim. The answer is the potential efficiency losses from taking precaution are smaller than the potential efficiency gains and far smaller in the 3x and 6x damage cases.

1.2 Organization of Paper

The climate literature has grown exponentially since Nordhaus (1979) seminal work. Section 2 reviews a very small portion of this literature with apologies to papers we under-emphasize or overlook. Section 3 presents our model, section 4 describes its calibration, section 5 explains our solution methods, section 6 presents our results, including their sensitivity to parameter values and the sharing of efficiency gains, section 7 compares optimal taxation in our model with that in an infinitely-lived agent (ILA) model where the models differ only along one dimension – intergenerational altruism, section 8 estimates the cost of imposing carbon policy if damages are, in fact, zero, and section 9 summarizes and concludes.

2 Literature Review

There is a vast and growing literature on the economics of climate change, much of it emanating from seminal contributions by Hotelling (1931), Solow (1974a,b) and Nordhaus (1979). The literature includes theoretical models, optimal tax models, and simulation models, including Nordhaus (1994, 2008, 2010), Nordhaus and Boyer (2000), Nordhaus (2017), Stern (2007), Metcalf (2010, 2014), Gurgel et al. (2011), Rausch et al. (2011), Manne et al. (1995), Plambeck et al. (1997), Tol (1997, 2002), Tol et al. (2003), and Ortiz et al. (2011). The literature also incorporates problems of coalition formation, e.g., Bréchet et al. (2011), Nordhaus (2015), Yang (2008), endogenous economic growth, e.g., van der Zwaan et al. (2002), Popp (2004) and Acemoglu et al. (2012), and stochastic damages, e.g., Lemoine and Traeger (2014), Lontzek et al. (2015), Cai et al. (2013), or Brock and Hansen (2017).

The Golosov et al. (2014) paper shows how to cast the problem into a standard calibrated

---

[10]This is not a problem for the ILA model since intergenerational redistribution is of no consequence in that framework. But if the underlying reality is that of selfish generations, the failure to compensate losing generations means the ILA prescription – Just tax carbon. – will produce a generational win-lose outcome.

setting and how to decentralize the planner's solution with carbon taxes. It also introduces more than one form of dirty energy and considers clean energy's production as limited due to a fixed input – two features we adopt in our analysis. Cai et al. (2013), Cai and Lontzek (2018), Lemoine and Traeger (2014) are major additions to the literature showing that optimal carbon tax rates can be considerably higher if the extent of future carbon damage is uncertain. The downside tail of the damage distribution can be particularly important given climate tipping points. These include losing much of the Amazon rain forest, faster onset of El Niño, the reversal of the Gulf Stream and other ocean circulatory systems, the melting of Greenland's ice sheet, the melting of Siberia's permafrost, and the collapse of the West Antarctic ice shelf.

There is also a growing literature considering regional effects of climate change and carbon taxation (see e.g. Nordhaus and Yang (1996), Cai et al. (2018), or Krusell and Smith Jr (2015). [11]

Early OlG models that consider resource-extraction and the environment include Howarth and Norgaard (1990), Howarth and Norgaard (1992), Burton (1993), Pecchenino and John (1994), John et al. (1995) and Marini and Scaramozzino (1995). Howarth and Norgaard (1990), Howarth (1991a,b), and Burton (1993).

Howarth and Norgaard (1990), posting a pure exchange model, and Howarth (1991b), using a two-period OLG model with capital, point out that policymakers can choose among an infinite number of Pareto efficient paths in the process of correcting negative environmental externalities. Gerlagh and Keyzer (2001), Gerlagh and van der Zwaan (2001) consider the choice among such Pareto paths and the potential use of trust-fund policies that provide future generations a share of the income derived from the exploitation of the natural resource. Gerlagh and van der Zwaan (2001) also point out that demographics can impact the set of efficient policy paths through their impact on the economy's general equilibrium.

Howarth (1991a) extended his important prior work to consider, in general terms, how to analyze economic efficiency in OLG models in the context of technological shocks. Howarth and Norgaard (1992) introduced damages to the production function from environmental degradation and studied the problem of sustainable development.[12] Rasmussen (2003) and Wendner (2001) examine the impact of the Kyoto Protocol on the future course of the energy sector. Wendner (2001) also considers the extent to which carbon taxes can be used to shore up Aus-

---

[11] Adding regions, with their own CO_2 emissions and damage functions, is another priority for our future research.

[12] An alternative approach to incorporating a negative environmental externality is including environmental quality directly in the utility function. Pecchenino and John (1994) and John et al. (1995) make this assumption in a discreet-time OlG model. Marini and Scaramozzino (1995) does the same in a continuous-time OLG framework. The problem of generational equity and sustainable development is also discussed by Batina and Krautkraemer (1999), Mourmouras (1991, 1993) in a model where energy is renewable.

tria's state pension system. Their papers feature large scale, perfect-foresight, single-country models. However, they omit climate damage.

The fact that OLG models do not admit unique solutions when it comes to allocating efficiency gains across agents, including agents born at different dates, has led some economists to introduce social welfare weights. Papers in this genre include Burton (1993), Calvo and Obstfeld (1988), Endress et al. (2014), Ansuategi and Escapa (2002), Howarth (1998), Marini and Scaramozzino (1995), Schneider et al. (2012), Lugovoy and Polbin (2016). These papers appear to emulate the SP solution.

Our paper is closely related to Bovenberg and Heijdra (1998, 2002), Heijdra et al. (2006). Their studies consider a continuous time Yaari-Blanchard model and they examine the use of debt policy to achieve Pareto improvements in the context of adverse climate change.[13] However, their models differ from ours in three important ways. First, they confine environmental damage to the utility function. Second, they do not model clean as well as dirty energy, with dirty energy exhausting in the future based on the speed of technological change in the clean energy sector as well as carbon policy. And third, their study doesn't provide a quantitative analysis of optimal policy. Rather it derives conditions on fundamentals, which ensure the existence of Pareto-improving tax-bond policies.

To achieve our narrow objective, evaluating the potential win-win from carbon policy in a model that is as close as possible to the literature's gold standard – Nordhaus (2017), we ignore fiscal policies, such as income taxes, which distort labor supply and saving decisions. Including such polices (another future research task) would permit larger win-wins from carbon taxation were the carbon tax revenue used to reduce such distortionary taxes. The potential to achieve a "double carbon dividend" was first explored by Goulder (1995),

Rausch (2013) and Rausch and Yonezawa (2018) examine the distributional effects of carbon taxation in a large scale OLG model. They take as given that a reduction of CO2 emissions is required and do not model the welfare effects of climate change. Instead, they consider various forms of revenue recycling and various forms of carbon taxes and compare their impacts on current and future generations' welfare.

Finally, we should point out that our model builds on a long line of dynamic CGE OLG models, which, however, don't include climate change. Early work on such models include Summers (1981)'s, Auerbach and Kotlikoff (1983), Auerbach and Kotlikoff (1987), and Altig et al. (2001). These models have been extended over time to include multiple regions, multiple goods, and demographic change (see, e.g., Börsch-Supan et al. (2006), Kotlikoff et al. (2007),

---

[13]Karp and Rezai (2014) also considers a life-cycle model, but explores the degree to which policy-induced general equilibrium changes in factor and asset prices could affect a Pareto improvement with no direct redistribution across generations. In his model agents live for two periods.

Fehr et al. (2003), and Benzell et al. (2017)).

3 Our Model

This section presents our model. Section 3.1 discusses its firms and section 3.2, its households. Section 3.3 demonstrates how the climate is coupled to the economy and section 3.4 discusses long-run growth.

3.1 Firms

Output is produced via

$$Y_t = A_t K_{y,t}^{\alpha} L_{y,t}^{\beta} E_t^{1-\alpha-\beta}, \tag{1}$$

where Y_t is final output whose price is normalized to 1, and A_t, $K_{y,t}$, $L_{y,t}$, E_t reference total factor productivity and the three inputs used to produce this output – capital, labor, and energy. Profit maximization requires

$$\alpha A_t K_{y,t}^{\alpha-1} L_{y,t}^{\beta} E_t^{1-\alpha-\beta} = r_t + \delta, \tag{2}$$

$$\beta A_t K_{y,t}^{\alpha} L_{y,t}^{\beta-1} E_t^{1-\alpha-\beta} = w_t, \tag{3}$$

and

$$(1 - \alpha - \beta) A_t K_{y,t}^{\alpha} L_{y,t}^{\beta} E_t^{-\alpha-\beta} = p_t, \tag{4}$$

where r_t, δ, w_t and p_t reference the real interest rate, the capital depreciation rate, the real wage rate, and the price of energy, respectively.[14]

We posit four perfectly substitutable sources of energy: clean energy, S_t, oil, O_t, gas, G_t, and coal, $\mathcal{C}_t$. Hence, we can write total energy supply, E_t, as

$$E_t = S_t + \kappa_O O_t + \kappa_G G_t + \kappa_{\mathcal{C}} \mathcal{C}_t. \tag{5}$$

Each type of dirty energy is measured in units of CO_2. The parameters κ_O, κ_G, and $\kappa_{\mathcal{C}}$ are energy efficiency coefficients.

Production of clean energy obeys

$$S_t = B_t K_{s,t}^{\theta} L_{s,t}^{\varphi} H_t^{1-\theta-\varphi}, \tag{6}$$

where B_t, $K_{s,t}$, $L_{s,t}$, and H_t reference, respectively, the clean energy sector's productivity level

---

[14]As is standard, α and β each lie between zero and 1.

and its demands for capital, labor, and land. Land is fixed in supply. Profit maximization in the clean-energy sector requires

$$p_t \theta B_t K_{s,t}^{\theta-1} L_{s,t}^{\varphi} H_t^{1-\theta-\varphi} = r_t + \delta, \tag{7}$$

$$p_t \varphi B_t K_{s,t}^{\theta} L_{s,t}^{\varphi-1} H_t^{1-\theta-\varphi} = w_t, \tag{8}$$

and

$$p_t (1 - \theta - \varphi) B_t K_{s,t}^{\theta} L_{s,t}^{\varphi} H_t^{-\theta-\varphi} = n_t, \tag{9}$$

where n_t is the rental price of land.

Dirty energy producers, indexed by $M \in \{O, G, C\}$, have a finite amount of energy reserves, R_t^M. The costs of extracting these reserves are increasing in the cumulative amount extracted. We posit the following functional form for the extraction cost of dirt energy of type M per unit of dirty energy extracted:

$$c_t^M (R_t^M) = \left(\xi_1^M + \xi_2^M \left(R_0^M - R_t^M \right) + \xi_3^M \left(R_0^M - R_t^M \right)^2 + \xi_4^M \left(R_0^M - R_t^M \right)^3 + \left(\frac{1}{R_t^M} \right) \right) \tag{10}$$

The last term ensures that extraction costs approach infinity when reserves approach zero.

Dirty energy producing firms maximize market value, V_t^m, given by

$$V_t^M = \sum_{j=0}^{\infty} \left[\left(p_{t+j}^M - c_{t+j}^M (R_{t+j}^M) - \tau_{t+j} \right) M_{t+j} + \mathcal{T}_t^M \right] \left(\prod_{i=0}^{j} \frac{1}{1 + r_{t+i}} \right), \tag{11}$$

subject to

$$R_t^M = R_{t-1}^M - M_t, \tag{12}$$

$$-R_t^M \leq 0, \tag{13}$$

$$-M_t \leq 0, \tag{14}$$

where p_t^M is the price of a unit of dirty energy M at time t, τ_t is the absolute tax per unit of carbon levied at time t, and $\mathcal{T}_t^M$ is the lump-sum rebate of time-t carbon taxes to type M dirty energy producers.

The Kuhn–Tucker optimality conditions for this problem are

$$p_t^M - c_t^M (R_t^M) - \tau_t - \ell_t^M + \mu_t^M = 0 \tag{15}$$

and

$$\frac{\partial c_t^M(R_t^M)}{\partial R_t^M}M_t + \ell_t^M - \frac{\ell_{t+1}^M}{1+r_{t+1}} - \psi_t^M = 0, \tag{16}$$

where ℓ_t, ψ_t and μ_t are non-negative Lagrange multipliers for the restrictions in equations 12, 13 and 14, respectively.

The complementary slackness conditions are

$$M_t \mu_t^M = 0, \tag{17}$$

and

$$R_t^M \psi_t^M = 0. \tag{18}$$

The value of land, Q_t, equals the present value of future land rents.

$$Q_t = \sum_{j=0}^{\infty} n_{t+j} H \left(\prod_{i=0}^{j} \frac{1}{1+r_{t+i}} \right). \tag{19}$$

Equation 5 implies

$$p_t^M = \kappa_M p_t. \tag{20}$$

3.2 Households

3.2.1 Overlapping Generation Households

Each household lives for 55 periods. Households born at time t maximize utility defined by

$$U_t = \sum_{j=1}^{55} \frac{1}{(1+\rho)^j} \frac{C_{t+j-1,j}^{1-\sigma} - 1}{1-\sigma} \tag{21}$$

subject to

$$a_{t+1,j+1} = (1+r_t)a_{t,j} + w_t l_j + \mathbb{T}_{t,j} - C_{t,j}, \tag{22}$$

where $C_{t,j}$, $\mathbb{T}_{t,j}$, $a_{t,j}$, l_j correspond to consumption, transfers from the government, assets, and labor supply of generation j at time t, ρ is the time preference rate, and σ is the coefficient of relative risk aversion. Total household assets comprise physical capital, the value of dirty energy firms, the value of land, and government debt.

$$\sum_{j=1}^{55} P_{t,j} a_{t,j} = K_t + V_t^O + V_t^G + V_t^C + Q_t + \mathcal{D}_t, \tag{23}$$

where $P_{t,j}$ is the population of generation j at time t and $\mathcal{D}_t$ is government debt.

Government debt evolves according to the following equation:

$$\mathcal{D}_{t+1} = (1 + r_t)\mathcal{D}_t + \sum_{j=1}^{55} P_{t,j}\mathbb{T}_{t,j}. \tag{24}$$

In implementing the LSRA policy, we set the initial debt level $\mathcal{D}_0$ and $\mathbb{T}_{t,j}$ for $t > 0, j > 1$ to zero. Thus the government pays transfers (possibly negative) to unborn generations in the year they are born whereas current generations receive transfers in at $t = 0$.

Total supplies of capital and labor equal the sum of their sectoral demands.

$$K_t = K_{y,t} + K_{s,t}. \tag{25}$$

$$L_t \equiv \sum_{j=1}^{55} P_{t,j}l_j = L_{y,t} + L_{s,t}. \tag{26}$$

3.2.2   The Infinitely Lived Agent

The ILA maximizes

$$U_t = \sum_{t=1}^{\infty} P_t \frac{1}{(1+\rho)^t} \frac{C_t^{1-\sigma} - 1}{1 - \sigma} \tag{27}$$

subject to the flow budget constraint

$$K_{t+1} + V_{t+1}^O + V_{t+1}^G + V_{t+1}^{\mathcal{C}} + Q_{t+1} = (1 + r_t)(K_t + V_t^O + V_t^G + V_t^{\mathcal{C}} + Q_t) + w_t L_t - P_t C_t, \tag{28}$$

where $P_t = \sum_{j=1}^{55} P_{t,j}$ is total population at time t and C_t is per capita consumption.

3.3   Modeling Climate Change's Negative Externality

Following Nordhaus (1994, 2008, 2010) and Nordhaus and Yang (1996), we assume the following time-t TFP-damage function, D_t.

$$D_t = 1 - \frac{1}{1 + \pi_1 T_t^A + \pi_2 \left(T_t^A\right)^2}. \tag{29}$$

The term T_t^A references the change, since 1900, in global mean surface temperature measured

11

in Celsius. The damage function alters TFP according to

$$A_t = (1 - D_t)Z_t, \tag{30}$$

where Z_t is an exogenous path of TFP absent climate change.

We adopt Nordhaus' three-reservoir (the atmosphere, the upper ocean, and the lower ocean) temperature model. CO_2 concentration obeys

$$\begin{pmatrix} J_t^A \\ J_t^U \\ J_t^L \end{pmatrix} = \Phi^J \begin{pmatrix} J_{t-1}^A \\ J_{t-1}^U \\ J_{t-1}^L \end{pmatrix} + \begin{pmatrix} O_t + G_t + \mathcal{C}_t \\ 0 \\ 0 \end{pmatrix}, \tag{31}$$

where J_t^A, J_t^U, J_t^L are concentrations of CO_2 in atmosphere, upper oceans and lower oceans, and Φ^J is a matrix of parameters.

CO_2 in the atmosphere impacts radiative forcing, F_t, according to

$$F_t = \eta_1 \log \frac{J_t^A}{J^0}. \tag{32}$$

And radiative forcing influences temperature according to

$$\begin{pmatrix} T_t^A \\ T_t^L \end{pmatrix} = \Phi^T \begin{pmatrix} T_{t-1}^A \\ T_{t-1}^L \end{pmatrix} + \begin{pmatrix} \eta_2 F_t \\ 0 \end{pmatrix}, \tag{33}$$

where T_t^L is the Celsius change, since 1900, in the temperature of the the deep oceans.

Per Nordhaus' formulation, a share of atmospheric carbon is absorbed by the oceans with the remaining share remaining in the atmosphere forever. Thus, damages from emissions are hump-shaped, peaking and declining, but not falling to zero.

3.4 Long-run Growth

We assume technology improves according to

$$Z_t = Z_0 exp(g_Z t) \tag{34}$$

and

$$B_t = B_0 exp(g_B t). \tag{35}$$

In the long run, after all dirty energy reserves have been extracted, and climate change

damage has stabilized, output and clean energy grow at rates g_Y and g_S, determined by

$$g_Y = \frac{g_Z + (1 - \alpha - \beta)g_B}{1 - \alpha - \theta(1 - \alpha - \beta)} \tag{36}$$

and

$$g_S = g_B + \theta\frac{g_Z + (1 - \alpha - \beta)g_B}{1 - \alpha - \theta(1 - \alpha - \beta)}. \tag{37}$$

In addition, the prices of energy and land grow at rates g_P and g_N determined by

$$g_P = \frac{g_Z(1 - \theta) - g_B\beta}{1 - \alpha - \theta(1 - \alpha - \beta)} \tag{38}$$

and

$$g_N = \frac{g_Z + (1 - \alpha - \beta)g_B}{1 - \alpha - \theta(1 - \alpha - \beta)} \equiv g_Y. \tag{39}$$

It is easy to show that, along with the economy's balanced growth path, the wage rate grows at g_Y and the return to capital is constant. As the growth-rate equations make clear, g_Z can differ from g_B without preventing long-run balanced growth. Indeed, our model admits many different long-run balanced growth paths. These include steady states in which output grows faster or slower than energy supply. If energy supply grows at a slower rate than output, its price must fall through time.[15]

4 Calibration

Our calibration adheres, where possible, to Nordhaus (2017)[16]. We also calibrate some parameters based on Golosov et al. (2014). Deviations from Nordhaus (2017) reflect structural differences between the two models. To begin, our model features autonomous overlapping generations, whereas the DICE model posits an infinitely-lived SP. Our model has three dirty energies supplied at an increasing cost. It also has an explicit clean energy sector whose production, at a point in time, is limited by complementary natural resources. The DICE model posits a fixed supply of a single dirty energy, which can be extracted at zero cost. There is no explicit clean energy sector. The DICE model's time periods reference five years rather than our one year.[17] Emissions abatement arises, in our model, via the replacement, in the market place, of clean for dirty energy. In DICE-2016R (henceforth, DICE), the SP determines how much of the fixed stock of dirty energy reserves to burn and when to burn them before year

---

[15]Note, a long-run declining energy price is required to ensure dirty energy extraction ends at a finite date.

[16]Version DICE-2016R-091916ap. available at: https://sites.google.com/site/williamdnordhaus/dice-rice

[17]As Cai et al. (2012) make clear, the length, in years, of a period matters to optimal carbon policy.

500, at which point the DICE model ends. Instead of an end-time condition, our model features a well defined long-run balanced growth path.[18]

4.1 Supply-Side Parameters

The capital, labor, and energy shares in the final goods production function, equation(1), are set to .3, .66, and .04, respectively, as in Golosov et al. (2014). The capital, labor, and land shares in the production function for clean energy (equation (6) are set at .125, .275, and .6, respectively.[19] The depreciation rate, δ, is set at 10 percent. The technology growth rates, g_B and g_Z, for clean energy and output, are set to .0064 for g_Z and .0137 for g_B. These values produce Nordhaus (2017)'s assumed long-run annual output growth of 2 percent and long-run -0.5 percent annual decline in the price of energy (actually clean energy, since it is the only long-run energy source).

Initial world capital equals \$223 trillion as in the DICE model. As for the global workforce, the DICE model assumes that the world's population rises from 7.403 to 11.5 billion. We disaggregate DICE's population dynamics to accommodate our 55-generation population age structure. Figure 1 summarizes this disaggregation. We calibrate initial levels of TFP in each sector to accord with DICE's initial world GDP of \$105.5 trillion 2010 USD and the 2010 \$550 USD backstop cost per ton of CO_2. Our damage function parameters are those in DICE, namely $\pi_1 = 0$ and $\pi_2 = 0.00236$.

4.2 Preferences

We set σ, the risk aversion coefficient, to 1.45 as in DICE. We assume that households work full time for their first 40 years and then retire. Moreover, we incorporate the following age-efficiency labor-earnings profile.

$$l_j = l_0 e^{(4.47+0.033j-0.00067j^2)}, j \leq 40.^{20} \tag{40}$$

We also follow the DICE model in assuming, for our base case, a time preference rate, ρ, of .015.

---

[18]Although our model's growth path accommodates stable population growth, we follow UN projections, which assume the planet's population stabilizes in the medium run.

[19]I.e., we assume that 60 percent of clean energy output is paid to the fixed input, land, and the rest of clean energy output is distributed between labor and capital in the same proportion as in the final goods sector.

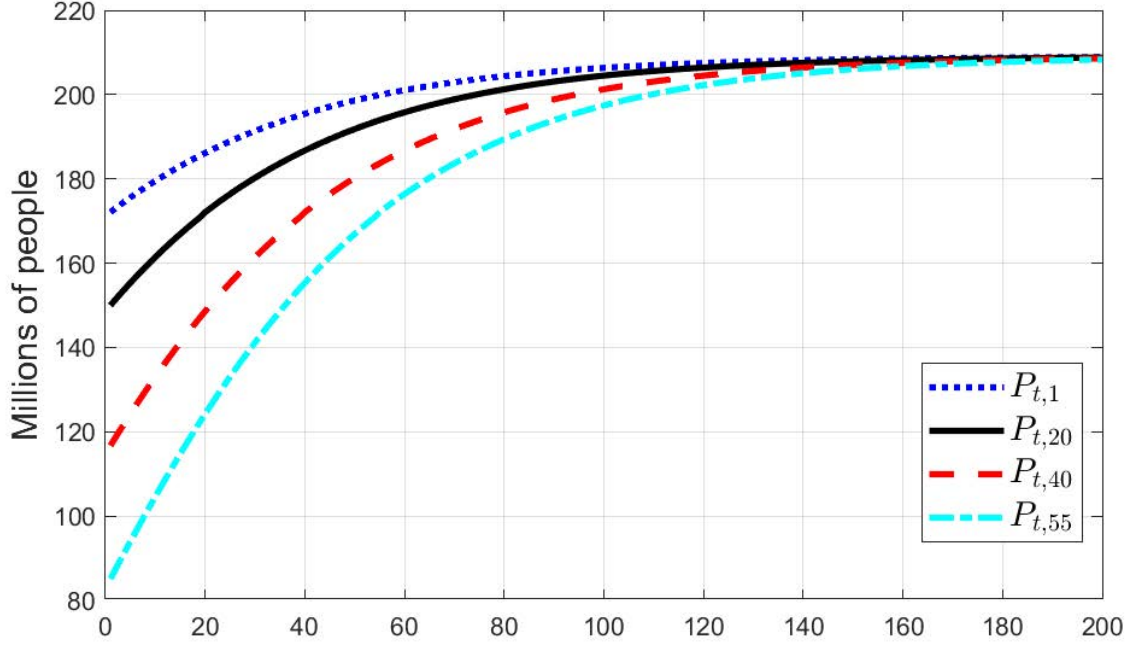[20]This is taken from Benzell et al. (2015).

Figure 1: Population dynamics

4.3 Dirty Energy Production-Sector Parameters

Based on McGlade and Ekins (2015), we calibrate global available oil reserves at 600 GtC, global available gas reserves at 400 GtC, and global available coal reserves at 2700 GtC. To calibrate the extraction cost functions for each of the fossil fuels, we fit third order polynomials to extraction cost data reported in McGlade and Ekins (2015).[21]

Finally, we calibrate our energy-efficiency parameters such that, initially, dirty energy constitutes 96 percent of global energy production. We normalize our oil efficiency coefficient, κ_O, to 1, set κ_G equal to 1.1, and κ_C to 0.35. Based on these values, the time-0 dirty-energy composition is 35 percent oil, 20 percent gas, and 45 percent coal, which accords with Boden et al. (2017).

---

[21]Let c^M stand for the cost of extracting dirty energy, M, then our regression is $c^M = \left(\xi_1^M + \xi_2^M \left(R_0^M - R^M \right) + \xi_3^M \left(R_0^M - R^M \right)^2 + \xi_4^M \left(R_0^M - R^M \right)^3 + \left(\frac{1}{R^M} \right) \right) + \epsilon$, where $\xi_1^O = 0.807$, $\xi_2^O = 0.355$, $\xi_3^O = -3.076e^{-04}$, $\xi_4^O = 1.260e^{-07}$, $\xi_1^G = 45.405$, $\xi_2^G = 0.142$, $\xi_3^G = -1.998e^{-04}$, $\xi_4^G = 2.304e^{-07}$, $\xi_1^C = 14.839$, $\xi_2^C = 0.013$, $\xi_3^C = -1.481e^{-07}$, $\xi_4^C = 5.106e^{-11}$.

15

4.4 Modeling Climate Change

Per Nordhaus (2017), we set initial atmospheric carbon concentration, J_t^A, at 851 GtC, the initial upper ocean carbon concentration, J_t^U, at 460 GtC, and the initial lower ocean carbon concentration, J_t^L, at 1740 GtC . We also set the initial lower ocean temperature change (relative to its 1900 value), T_t^L, at .0068 °C and the initial atmospheric temperature change (again, relative to its 1900 value), T_t^A, at 0.85 °C.

Cai et al. (2012) suggest how to calibrate our 1-year per period climate model to accord with DICE's 5-year per period climate model. Specifically, we simulate the DICE model with no limitation on emissions and record the resulting temperature path. Next, we interpolate DICE's 5-year per period temperature path over single years. Finally, we estimate parameters for our 1-year climate model that minimize the mean squared percentage error between our interpolated paths of climate variables over 5-year intervals in our 1-year system and DICE's original 5-year system. This produces the following values for the parameter matrices Φ^J and Φ^T as well as the constant η_2. Over 5-year periods our annual system produces climate outcomes that are almost identical to those in the DICE model.

$$
\Phi^J = \begin{pmatrix} 0.9763 & 0.0387 & 0 \\ 0.0237 & 0.9599 & 0.0003 \\ 0 & 0.0014 & 0.9997 \end{pmatrix}, \Phi^T = \begin{pmatrix} 0.9715 & 0.0017 \\ 0.0049 & 0.9951 \end{pmatrix}, \eta_2 = 0.0225. \tag{41}
$$

The parameters $\eta_1 = 3.6813$ and $\overline{J} = 588$ GtC in the radiative forcing equation (32) are those in the DICE model.

5 Computation Methods

This section presents our methods for solving our OLG model as well as its ILA counterpart. As in Auerbach and Kotlikoff (1987), our OLG model's overall solution combines solving, in an inner loop, microeconomic equations either analytically or numerically, depending on the time interval, while iterating, in an outer-loop, over macroeconomic aggregates. Our ILA model's solution combines outer-loop macro iteration with an inner-loop multiple shooting algorithm, where the inner loop is solved, analytically or numerically, depending on the time interval.

5.1 Our OLG Solution Method

Our method starts by guessing the time paths of the aggregate capital stock, K_t, and damages, D_t. When we apply the LSRA mechanism, we also guess the uniform welfare gain, λ. Next,

we jointly calculate, for each year, factor prices, the sectoral distributions of capital and labor, output of the final good, CO_2 emissions, the production of each type of energy, the prices of clean energy, oil, gas, and coal, the price of land, the prices of coal, oil, and gas reserves, and, where relevant, the path of generation-specific LSRA net taxes, with the LSRA net tax imposed on the oldest generation chosen to achieve intertemporal LSRA budget balance. From the paths of a factor and non-capital asset prices, we determine the time path of households' aggregate supply of capital. The new paths of capital and damages are weighted with the prior paths to obtain new guesses for these paths in the outer loop of our iteration. If the percentage change in welfare for the initial oldest generation differs from λ (or zero, in the case the initial old don't share in the efficiency gains), the guessed value of λ is adjusted to produce that outcome.[22] We continue the outer to inner to outer loop iteration until convergence.

Following Nordhaus (2017), we chose parameters for the right-hand-side of equation (38) that produce, in the long run, a decreasing price of energy. This guarantees that our model's economy will never extract all dirty energy reserves, i.e., extraction will endogenously stop at time T. Since dirty energy reserves will always be positive, ψ_t^M for all M will always equal zero. I.e., we can drop the complementary slackness conditions $R_t^M \psi_t^M = 0$ and the variables ψ_t^M. After T, we can analytically solve for the sectoral distribution of aggregate capital and labor using

$$K_{y,t} = \frac{\alpha}{\theta\,(1-\beta) + \alpha(1-\theta)} K_t, \tag{42}$$

$$K_{s,t} = \frac{(1-\alpha-\beta)\theta}{\theta\,(1-\beta) + \alpha(1-\theta)} K_t, \tag{43}$$

$$L_{y,t} = \frac{\beta}{\varphi\,(1-\alpha) + \beta(1-\varphi)} L_t, \tag{44}$$

and

$$L_{s,t} = \frac{\varphi(1-\alpha-\beta)}{\varphi\,(1-\alpha) + \beta(1-\varphi)} L_t. \tag{45}$$

In the solution algorithm, we set two dates, T_2 and T_1. T_2 is the year by which we assume the model has reached, to a very high degree of precision, its long-run balanced growth path. $T_1 > T$ is a year sufficiently high to assure that dirty energy extraction has stopped. $T_2 - T_1$ must be quite large because atmospheric carbon depreciates very slowly in our model. Our algorithm sets T_2 at 2999 (3000 years from t=0). We typically set T_1 to 149 in running the model.

For $t \leq T_1$, when each type of dirty energy is potentially being extracted, we need to

---

[22]See Auerbach and Kotlikoff (1987) for a description of this adjustment process.

solve the following system of equations, which include Kuhn-Tucker conditions permitting the extractions of some or all of the fossil fuels to equal zero between T_1 and T.

$$K_t = K_{y,t} + K_{s,t}, \tag{46}$$

$$L_t = L_{y,t} + L_{s,t}, \tag{47}$$

$$\alpha A_t K_{y,t}^{\alpha-1} L_{y,t}^{\beta} E_t^{1-\alpha-\beta} = r_t + \delta \tag{48}$$

$$\beta A_t K_{y,t}^{\alpha} L_{y,t}^{\beta-1} E_t^{1-\alpha-\beta} = w_t \tag{49}$$

$$(1-\alpha-\beta)A_t K_{y,t}^{\alpha} L_{y,t}^{\beta} E_t^{-\alpha-\beta} = p_t, \tag{50}$$

$$E_t = S_t + \kappa_O O_t + \kappa_G G_t + \kappa_C C_t. \tag{51}$$

$$S_t = B_t K_{s,t}^{\theta} L_{s,t}^{\varphi} H^{1-\theta-\varphi}, \tag{52}$$

$$p_t \theta B_t K_{s,t}^{\theta-1} L_{s,t}^{\varphi} H^{1-\theta-\varphi} = r_t + \delta, \tag{53}$$

$$p_t \varphi B_t K_{s,t}^{\theta} L_{s,t}^{\varphi-1} H^{1-\theta-\varphi} = w_t, \tag{54}$$

$$\kappa_O p_t - c_t^O(R_t^O) - \tau_t - \ell_t^O + \mu_t^O = 0, \tag{55}$$

$$\kappa_G p_t - c_t^G(R_t^G) - \tau_t - \ell_t^G + \mu_t^G = 0, \tag{56}$$

$$\kappa_C p_t - c_t^C(R_t^C) - \tau_t - \ell_t^C + \mu_t^C = 0, \tag{57}$$

$$\frac{\partial c_t^O(R_t^O)}{\partial R_t^O} O_t + \ell_t^O - \frac{\ell_{t+1}^O}{1+r_{t+1}} = 0, \tag{58}$$

$$\frac{\partial c_t^G(R_t^G)}{\partial R_t^G} G_t + \ell_t^G - \frac{\ell_{t+1}^G}{1+r_{t+1}} = 0, \tag{59}$$

$$\frac{\partial c_t^C(R_t^C)}{\partial R_t^C} C_t + \ell_t^C - \frac{\ell_{t+1}^C}{1+r_{t+1}} = 0, \tag{60}$$

$$R_t^O = R_{t-1}^O - O_t, \tag{61}$$

$$R_t^G = R_{t-1}^G - G_t, \tag{62}$$

$$R_t^C = R_{t-1}^C - C_t, \tag{63}$$

$$O_t \mu_t^O = 0, \tag{64}$$

$$G_t \mu_t^G = 0, \tag{65}$$

$$C_t \mu_t^C = 0. \tag{66}$$

In solving this system of equations, we provide initial conditions for oil, gas and coal reserves as well as the zero-value, for for $t \geq T_1$, terminal conditions for the shadow values of dirty energy reserves, $\ell_{T_1+1}^O$, $\ell_{T_1+1}^G$ and $\ell_{T_1+1}^C$.

To facilitate our solution, we used the square norm of the Fischer-Burmeister function and replaced equations 64 - 66 by

$$\left(O_t + \mu_t^O - \sqrt{(O_t)^2 + (\mu_t^O)^2} \right)^2 = 0, \tag{67}$$

$$\left(G_t + \mu_t^G - \sqrt{(G_t)^2 + (\mu_t^G)^2} \right)^2 = 0, \tag{68}$$

and

$$\left(C_t + \mu_t^C - \sqrt{(C_t)^2 + (\mu_t^C)^2} \right)^2 = 0. \tag{69}$$

For example, equation 67 equals zero iff $O_t \geq 0$, $\mu_t^O \geq 0$ and $O_t \mu_t^O = 0$. Thus we can numerically solve the system of equations (46) - (63) and (67) - (69) with interior-point methods. Our solver is the standard Matlab fmincon function with the 'interior-point' option and no bounds on variables. We compute the Hessian and gradients and provide them to fmincon as sparse matrices, which dramatically speeds up convergence. Our algorithm can be formalized in the steps stated in Alg. 1.

Algorithm 1 Algorithm for solving the OLG model

Require: $T_1, T_2, B_t, A_t^{guess}, K_t^{guess}, tol, tol_K, tol_A$

1: $K_t \leftarrow K_t^{guess}, A_t \leftarrow A_t^{guess}$
2: **repeat**
3:   Solve system of Eq. 46 - 63, 67 - 69 for $0 \leq t \leq T_1$ with 'fmincon' with tolerance tol assuming that after T_1 extraction of dirty energy is zero
4:   Solve system of Eq. 46 - 63, 67 - 69 for $T_1 < t \leq T_2$ analitically assuming that after T_1 extraction of dirty energy is zero
5:   $K_t \leftarrow K_t^{guess}, A_t \leftarrow A_t^{guess}$ construct K_t^{guess} using household optimal behavior with given prices, construct A_t^{guess} using dirty energy emissions and climate block equations
6: **until** $\|K_t - K_t^{guess}\| < tol_K, \|A_t - A_t^{guess}\| < tol_A$

5.2   Our ILA Solution Method

The ILA model consists of equations (46) - (63), (67) - (69) plus

$$\frac{C_{t+1}}{C_t} = \left(\frac{1 + r_{t+1}}{1 + \rho} \right)^{1/\sigma} \tag{70}$$

and

$$Y_t = K_{t+1} - (1 - \delta)K_t + P_t C_t + c_t^{\mathcal{C}}(R_t^{\mathcal{C}})\mathcal{C}_t + c_t^O(R_t^O)O_t + c_t^G(R_t^G)G_t. \tag{71}$$

Equation (70) is the standard ILA consumption-growth formula. Equation (71) is the ILA's flow budget. To solve the model, we guess the paths of damages, solve equations (46) - (63) and (67) - (71) using multiple shooting (see (Lipton et al., 1982)), and then update our guessed path of damages based on the derived path of output. Our shooting algorithm divides the transition path into m 50 year-long sub periods, $t_0 < t_1 < .. < t_{m-1} < T_2$ and guesses the capital stock at dates $t_1, .., t_{m-1}$, denoted $k_{t_1}, .., k_{t_{m-1}}$.

Assuming that the model reaches its steady state at T_2, we shoot $C_{t_{m-1}}$ with initial condition $k_{t_{m-1}}$, targeting steady state consumption at date T_2. Then we shoot $C_{t_{m-2}}$ with initial condition $k_{t_{m-2}}$, targeting consumption $C_{t_{m-1}}$ and so on. After each consumption shooting round, we update $k_{t_1}, .., k_{t_{m-1}}$ and repeat, using simple bisection, until convergence. As with the OLG solution, we set t_1 high enough, typically at 150 years, to ensure dirty energy extraction ends by that date. Our shooting prior to t_1 uses fmincon to solve the relevant system of equations, namely (46) - (63) and (67) - (71). For intervals between t_1 and T_2, the equations are solved analytically for given paths of consumption and capital.

5.3 Accuracy of the Solutions

All results reported below are based on highly accurate numerical solutions. We solve the OLG model setting the error tolerance – step 6 in Alg. 1 – at 10^{-6}. The residuals in our ILA solutions are of order $\sim 10^{-9}$.

6 Results

We first examine the magnitude of climate damage absent carbon policy and show the remarkable ability of such policy to mitigate the problem. Second, we present optimal UWI carbon policy for our base case and consider its sensitivity to the damage function, the time preference rate, and the long-run growth rate of the price of energy. Third, we examine the size of the intergenerational redistribution required to achieve the UWI win-wins. Fourth, we consider how optimal carbon taxation, including requisite intergenerational redistribution, differs depending on how efficiency gains are distributed across generations. Fifth, we examine the economic precision of our optimal tax solution. Next, we consider the efficiency costs of implementing the optimal carbon tax but with a delay of 20 years. These efficiency costs include those arising from Sinn (2008)'s Green Paradox. Sixth, we compare optimal UWI carbon policy in our OLG

model with optimal carbon policy in the analogous ILA model. We point out in this subsection that our two-part carbon policy – an initial tax and an annual growth rate – appears to achieve the UWI and ILA optimums were we to consider more fiscal instruments. Seventh, we consider the efficiency cost of implementing optimal carbon taxation if the carbon skeptics are correct and there are no damages from climate change. The Appendix presents figures showing the evolution of all key variables for different damage functions under Business as Usual, implementation of the carbon tax policies that maximize UWI gains, but leaving out the policies' associated LSRA transfers and taxes, and implementation of the carbon policies, inclusive of the LSRA transfers and taxes, which maximize UWI gains.

6.1 Damages Under Business as Usual

Under the BAU-scenario, CO_2 concentration in the atmosphere increases over the next 100 years to about 900 Gt (see figures 4-6 in the Appendix). It then stabilizes and decreases slightly over the subsequent century. The associated increase in surface temperature over the proximate century is 4.7 degrees Celsius and almost 6 degrees at its peak. This is consistent with the more optimistic scenarios used in climate science, in particular with so-called representative carbon pathways (RCPs) designed RCP4.5 and RCP6, but not with the pessimistic RPC8.5[23] However, as emphasized by Weitzman (2007), Pindyck (2013), Lontzek et al. (2015), and others, the extent and timing of temperature rise under BAU is highly uncertain. Again, modeling this uncertainty is at the top of our future research agenda. Here, however, we simply seek to understand the Pareto efficiency gains from carbon taxation in a framework that closely mimics the standard DICE model.

Table 1 shows BAU damages as a percent of GDP for selected years assuming 1x, 3x, and 6x damage functions. It also shows the size of damages but under optimal UWI policy. With BAU, damages are initially 0.17, 0.51, and 1.01 percent of GDP for the 1x, 3x, and 6x damage-function cases, respectively. In year 200, damages in the three cases are 7.70, 18.77, and 29.55 percent of output. Paradoxically, with higher damages, the temperature increase is actually smaller. In the 6x case, for example, the peak temperature increase is 4.5 degrees in year 99. There are two reasons. First, damages are so large in the 6x case that they limit future output, which limits future emissions. Second, the clean energy sector is assumed to operate with no damage, giving it a greater comparative advantage the larger the extent of the damages. The table also shows that the optimal UWI carbon policy is remarkably effective in reducing peak damages. Peak damages are 2.92 percent, not 7.72 percent of GDP in the 1x case. They are 5.52 percent, not 18.78 percent of GDP in the 3x case. Moreover, they are 8.62 percent, not 29.55

---

[23]See Rogelj et al. (2011)

percent of GDP in the 6x case. Thus, the UWI policy reduces peak damages by 62.2 percent, 70.6 percent, and 70.8 percent in the 1x, 3x, and 6x cases, respectively.

Assuming 1x damages, atmospheric CO_2 rises under the optimal UWI policy slightly over the next 70 years and then decreases. The reason? CO_2 concentration stays around 5000 Gt under the policy rather than rising to 9000 Gt. Moreover, the rise in temperature is considerably slower than under BAU. Despite the drastic reduction in CO_2 emissions, the average surface temperature still rises by about 4 degrees over the next 125 years under the optimal UWI carbon policy. Nevertheless, given our damage function, the resulting damages are much smaller. As opposed to an almost 8 percent long-run permanent decline in TFP, the long-run TFP reduction is limited to roughly 3 percent. This long-run difference of 5 percent of GDP provides the scope for our calculated UWI gain.[24] Assuming 6x damages under optimal taxation, total CO_2 in the atmosphere stays below 4000 Gt. The associated increase in temperature is only about 2.5 degrees – close to the goal of the Paris accord (see figures 10-12 in the Appendix).

Table 1: Damages as Percent of GDP under BAU and Optimal UWI, by Year

	t=0	t=10	t=50	t=100	t=150	t=200	t=250	t=300
1x, BAU	0.17	0.29	1.77	5.08	7.28	7.70	7.72	7.65
1x, UWI	0.17	0.27	1.03	2.45	2.87	2.92	2.89	2.83
3x, BAU	0.51	0.87	5.09	13.49	18.01	18.77	18.79	18.62
3x, UWI	0.51	0.79	2.56	4.93	5.48	5.52	5.44	5.34
6x, BAU	1.01	1.72	9.58	23.05	28.69	29.55	29.53	29.27
6x, UWI	1.01	1.55	4.66	7.99	8.61	8.62	8.49	8.31

BAU - Business as Usual, UWI - Uniform Welfare Increase, 1x, 3x, and 6x references models with damage functions equal to the Nordhaus (2017) function, 3 times that function, and 6 times that function.

6.2   Optimal Uniform Welfare Improving Carbon Policy

Table 2 presents our optimal UWI carbon policies under alternative assumptions. The second row is our base case. It assumes Nordhaus (2017)'s 1.5 percent time preference rate and his long-run negative .5 percent energy-price growth rate. The optimal initial energy tax, τ_0, is $30 per cubic ton of CO_2, rising at a rate, g_τ, of 1.5 percent per year. With 3x and 6x damage functions, the initial optimal carbon tax is $50 and $70, respectively. However, the tax's optimal growth rate remains 1.5 percent.

---

[24]The crucial reason for the relatively large increase in temperature permitted under the optimal policy compared, for example, with the Paris accord, which vowed to limit the rise in surface temperature to 1.5 to 2 degrees, lies in the Nordhaus (2017) damage function. With this function, a 4-degree increase in average temperature leads to major, but still moderate long-run damages.

The UWI efficiency gain, λ, is 0.73 percent. With 3x damages, it is 2.58 percent. Moreover, with 6x damages, it is 4.69 percent. Even a 0.73 percent gain is significant. It is equivalent to adopting no policy but raising, under BAU, each current and future generation's consumption in each future year by 0.73 percent. The close to 5 percent efficiency gains for all generations for all time is massive, reflecting the severity of damages in the 6x case. The first row of table 2 considers the lowest time preference rate, namely 0.8 percent, for which our solution method converges for all three damage functions. Our optimal two-part carbon policy does not change in the 1x case when we set ρ to 0.8 rather than 1.5. The initial tax, but not the tax's growth rate, does rise with higher damages – from 50 to 60 in the 3x damage case and from 70 to 80 in the 6x damage case.

There is considerable sensitivity in the base case to increases in ρ to 3 percent, with the base-case initial tax dropping from \$30 to \$10 and the tax's annual growth rate doubling to 3 percent. Higher damage functions entail lower initial taxes as well, but no higher growth rates.

The size of ρ makes a significant difference in the size of the UWI efficiency gains. In the 1x case, the gain is 1.07 percent, 0.73 percent, and 0.23 percent when ρ is 0.8 percent, 1.5 percent, and 3.0 percent, respectively. In the 6x case, the corresponding gains are 5.58 percent, 4.69 percent, and 4.64 percent. The last three rows of the table show that the smaller or larger percentage annual declines in energy's long-run price make no difference to the optimal policy and no material difference to the size of UWIs.

Table 2: Optimal Uniform Welfare Improving Carbon Policy

ρ	g_P	1x Damage Function			3x Damage Function			6x Damage Function		
		τ_0	g_τ	λ	τ_0	g_τ	λ	τ_0	g_τ	λ
0.8%	-0.5%	\$30	1.5%	1.07%	\$60	1.5%	3.25%	\$80	1.5%	5.58%
1.5%	-0.5%	\$30	1.5%	0.73%	\$50	1.5%	2.58%	\$70	1.5%	4.69%
3%	-0.5%	\$10	3%	0.23%	\$40	1.5%	1.26%	\$50	1.5%	4.64%
1.5%	-0.1%	\$30	1.5%	0.70%	\$50	1.5%	2.52%	\$70	1.5%	4.58%
1.5%	-1%	\$30	1.5%	0.72%	\$50	1.5%	2.60%	\$70	1.5%	4.74%
1.5%	-2%	\$30	1.5%	0.65%	\$50	1.5%	2.46%	\$70	1.5%	4.61%

ρ references the time preference rate, g_P the long-run annual percentage change in the price of energy, τ_0 the initial carbon tax, g_τ the growth rate of the carbon tax, and λ the welfare gain measured as the equivalent percentage increase in consumption under BAU needed to achieve the welfare gain under the optimal UWI policy.

6.3 Size of Intergenerational Redistribution Required to Achieve UWI

Table 3 considers, under the heading "Transfers", the size of the net lump-sum transfers required, in the 1x, 3x, and 6x cases to affect each case's uniform percentage rise in welfare. The net transfers are expressed as a percentage of remaining lifetime consumption in the case of current cohorts and as a percent of the present value of total lifetime consumption in the case of those born in the future. The columns with the headings λ show the welfare changes that would arise if we impose the optimal carbon policy but engage in no intergenerational redistribution.

Consider, first, the results in the third column. They show that if we just impose a 30 carbon tax and let it rise at 1.5 percent annual, but do not redistribute across generations to achieve the associated 0.73 percent UWI, we end up with losers as well as winners. In this case, first, older generations benefit due to the rise in the price of energy, which increases the value of their dirty energy reserves, leaving them better off despite paying more for energy in their remaining years. For example, those in their last year of life at the time of the policy experience a 4.52 percent welfare gain. For future generations, there are also far more substantial welfare gains than 0.73 percent. Indeed, those born in the year 200 are better off by 7.50 percent. However, the failure to redistribute hurts initial young generations and those born in the first three or so decades after the policy is imposed. Their maximum welfare loss amounts to roughly 1 percent.

Column two shows the net transfers needed to transform what would otherwise be a generational win-lose outcome to a uniform win-win in our baseline case. There are several interesting findings here. First, the transfers to all living generations are positive. They are highest for the oldest current generations who will not live long enough to experience much if any of the benefits from reduced climate damage.[25] Second, those born in the first few decades after the policy begins also receive compensation. Again, their loss in welfare from facing higher energy prices exceeds their benefits from lower climate damage. However, the compensation does not decline uniformly from oldest to youngest among the living and near-term newborn generations. For example, the net transfer is 0.55 percent for those 15 years old in a year 0, 1.17 percent for those born in a year 0, and 0.80 percent for those born in a year 15. For those born 35 and more years after the reform, the uniform win-win policy entails rising and then declining net taxes. The net tax peaks at roughly 8 percent for those born two to three centuries into the future. In the very long run, it declines to less than 4 percent. The requisite net tax is smaller in the long run due to the slow, but steady depreciation of atmospheric CO_2. Hence, generations born far into the future have much less to gain from reducing carbon emissions since even under BAU, CO_2 emissions end by roughly year 125. Since such long-run generations gain less than

---

[25]Asset prices are adversely affected by the rise in interest rates associate with LSRA debt crowding out caital.

generations born, say, in year 200, they cannot be taxed as highly in achieving the uniform welfare improvement.[26]

As the 3x and 6x results show, future generations must, but are also able to compensate current generations and those about to be born to a much higher extent when climate change is an even bigger problem. The higher damage functions mean far higher carbon taxes and, consequently, far higher short-term energy prices (see figures 7-9 and 10-12 in the Appendix). With a 6x damage function, generations born in year 200 face a 50.13 percent tax, again measured as a share of their BAU consumption. However, it would, relative to BAU, leave them and all other generations, on balance, 4.69 percent better off. Of course, future generations would greatly prefer letting earlier generations impose the UWI optimal carbon policy, but without any compensation. For newborns in year 200, this would produce a 45.04 percent utility increase – almost ten times the utility gain when efficiency gains are evenly distributed to each and all. The magnitude of the transfers and taxes needed to achieve the UWI with 6x damages speaks to the political challenges in its implementation. Today's youngsters would, for example, need to receive a roughly 10 percent of BAU consumption transfer each year through the rest of their lives, whereas those born some four decades later would be required to pay taxes of the same magnitude.

Our LSRA mechanism can be viewed as a non-distortionary deficit policy in which the government cuts taxes in a lump sum manner for early current and newborn generations based on the figures in the table and then raises the net tax, again, as indicated, to service the associated debt. The size of the debt to GDP ratio required to effect the win-win is significant even for the 1x damage case. The LSRA's debt to GDP ratio is 0.52 in year 50, 0.78 in year 100, 0.82 in year 200, 0.70 in the year 1000, stabilizing at 0.48 in the long run in 1x damage function case. The LSRA's debt to GDP ratio is 1.03 in year 50, 1.22 in year 100, 1.19 in year 200, 1.11 in the year 1000, stabilizing at 0.89 in the long run in 3x damage function case. Moreover, the LSRA's debt to GDP ratio is 1.31 in year 50, 1.34 in year 100, 1.24 in year 200, 1.22 in the year 1000, stabilizing at 1.12 in the long run in 6x damage function case.

6.3.1 Sensitivity of Optimal Carbon Policy to Intergenerational Sharing of Efficiency Gains

Table 4 considers the sensitivity of optimal carbon policy to alternative ways to share efficiency gains. In addition to first row's UWI , we consider allocating all efficiency gains uniformly to either the born (those alive in year zero) or the unborn (those born after year zero). In the 1x

---

[26]This statement is relative to the conventional formulation of climate damage. It would not hold were we to incorporate, as we intend in future work, climate tipping points that impose permanent damage on the economy.

case, the optimal carbon policy is invariant to the manner of efficiency-gain sharing. In the 3x and 6x cases, the initial tax, but not the carbon tax growth rate, depends on the sharing rule. The largest discrepancy arises in the 6x case. When the gains are allocated to the born, the optimal initial tax is \$60. It is one third, i.e., \$20 larger when the gains are shared just among the unborn.[27] One would expect welfare gains to be more significant when spread over fewer generations. This is what table 2 shows for the 1x and 3x cases. In the 1x case, the welfare gains are 0.73 percent with uniform sharing, 1.79 percent with sharing among the born, and 1.20 percent with sharing among the unborn case. In the 6x case, the corresponding percentages are 4.69, 8.90, and 9.22. In this case, the larger gain to the unborn may reflect the larger size of the climate damage in the BAU. Table 5 compares LSRA transfers and taxes under the three methods of sharing the efficiency gains. Both the transfers and taxes are largest when the efficiency gains are distributed uniformly to the born, next largest when they are distributed uniformly to both the born and unborn, and smallest when they are distributed to the unborn.[28]

---

[27]This appears to reflect the lower path of interest rates when the efficiency gains are distributed just to the unborn. The switch from distributing the gains from the born to the unborn leads the LSRA to accumulate a smaller debt, which means more capital and a lower interest rate at each future date. This makes the present value cost of compensating initial generations (the born) lower, permitting a higher initial carbon tax.

[28]There are many factors at play in determining these figures, including the extent of damages, the degree of capital's crowding out from the LSRA debt policy, and the size of each cohort's consumption, since the taxes and transfers gains are measured as consumption equivalents.

Table 3: LSRA Net Transfer as Share of the Cohort's Present Value of Remaining or Total Lifetime Consumption and Policy-Induced Percentage Welfare Change that Would Occur Absent Net Transfers

	1x Damage Function		3x Damage Function		6x Damage Function	
Birth year	Transfers	No Tranfers λ	Transfers	No Tranfers λ	Transfers	No Transfers λ
-54	1.17	4.52	1.96	5.36	2.51	6.13
-45	1.14	4.50	2.03	5.29	2.72	6.05
-35	1.09	3.17	2.52	3.67	3.58	4.31
-25	0.97	0.97	3.75	1.04	5.63	1.41
-15	0.55	-0.28	6.01	-0.40	9.38	-0.07
-5	0.60	-0.83	6.46	-0.87	10.26	-0.40
0	1.17	-0.89	1.96	-0.82	2.51	-0.22
5	1.09	-0.94	1.60	-0.69	1.78	0.14
15	0.80	-0.82	0.48	0.08	-0.37	1.75
25	0.37	-0.51	-0.99	1.27	-3.24	4.10
35	-0.15	-0.08	-2.75	2.73	-6.72	7.00
45	-0.73	0.40	-4.77	4.45	-10.78	10.45
55	-1.36	0.92	-7.07	6.43	-15.41	14.40
100	-5.33	4.56	-19.37	17.37	-38.53	34.40
200	-8.05	7.50	-26.39	24.23	-50.13	45.04
300	-7.98	7.48	-26.06	24.05	-49.40	44.60
400	-7.74	7.29	-25.19	23.38	-47.67	43.32
500	-7.45	7.06	-24.16	22.59	-45.67	41.80
1000	-5.96	5.83	-18.98	18.49	-35.61	34.07
1500	-4.81	4.88	-15.06	15.33	-28.06	28.13
2000	-4.02	4.21	-12.39	13.15	-22.94	24.04
2500	-3.49	3.75	-10.60	11.67	-19.55	21.29

The UWIs (the λ values with transfers) are 0.73, 2.58, and 4.69 percent with 1x, 3x, and 6x damage functions, respectively. Transfers reference the LSRA net payment expressed as a present value of the cohort's remaining or total present value of lifetime consumption under the optimal policy, which are needed, under the optimal policy, to deliver the UWI gain relative to the cohort's remaining or lifetime utility under BAU. No-Transfers λ references the welfare gain measured as the equivalent percentage increase in consumption under BAU needed to achieve the welfare gain under the optimal UWI policy were the policy to be imposed without its LSRA net payments.

Table 4: Optimal Carbon Tax Policy Under Alternative Ways to Share Efficiency Gains*

	1x Damage Function			3x Damage Function			6x Damage Function		
	τ_0	g_τ	λ	τ_0	g_τ	λ	τ_0	g_τ	λ
Uniform	30$	1.5%	0.73%	50$	1.5%	2.59%	70$	1.5%	4.69%
Born	30$	1.5%	1.79%	40$	1.5%	5.44%	60$	1.5%	8.90$
Unborn	30$	1.5%	1.20%	50$	1.5%	4.70%	80$	1.5%	9.22%

*$\rho = 1.5\%$, $g_P = -0.5\%$

Table 5: LSRA Net Transfer at Time 0 or at Birth as Share of the Present Value of Consumption under Alternative Redistribution Policies (in percents)

	1x Damage Function			3x Damage Function			6x Damage Function		
Birth year	UWI	UWB	UWU	UWI	UWB	UWU	UWI	UWB	UWU
-54	1.17	1.93	0.62	1.96	3.74	0.02	2.51	5.07	-0.86
-45	1.14	1.79	0.68	2.03	3.54	0.39	2.72	4.91	-0.12
-35	1.09	1.81	0.58	2.52	4.13	0.88	3.58	5.77	0.79
-25	0.97	2.16	0.13	3.75	6.33	1.36	5.63	8.75	1.64
-15	0.55	2.92	-1.16	6.01	11.17	1.46	9.38	15.22	1.83
-5	0.60	3.24	-1.31	6.46	12.27	1.25	10.26	16.97	1.54
0	1.17	1.95	0.61	1.96	3.79	-0.02	2.51	5.12	-0.93
5	1.09	0.15	1.68	1.60	-1.72	4.03	1.78	-3.98	6.75
15	0.80	-0.08	1.36	0.48	-2.70	2.83	-0.37	-6.07	4.49
25	0.37	-0.46	0.90	-0.99	-4.06	1.30	-3.24	-8.89	1.56
35	-0.15	-0.96	0.36	-2.75	-5.73	-0.47	-6.72	-12.33	-1.92
45	-0.73	-1.54	-0.22	-4.77	-7.72	-2.43	-10.78	-16.42	-5.91
55	-1.36	-2.19	-0.83	-7.07	-10.05	-4.62	-15.41	-21.22	-10.36
100	-5.33	-6.26	-4.75	-19.37	-22.37	-16.43	-38.53	-44.98	-32.79
200	-8.05	-9.02	-7.43	-26.39	-29.52	-23.18	-50.13	-57.07	-43.91
300	-7.98	-8.96	-7.36	-26.06	-29.22	-22.86	-49.40	-56.37	-43.16
400	-7.74	-8.71	-7.12	-25.19	-28.36	-22.01	-47.67	-54.64	-41.44
500	-7.45	-8.42	-6.83	-24.16	-27.34	-21.01	-45.67	-52.61	-39.45
1000	-5.96	-6.90	-5.36	-18.98	-22.19	-15.99	-35.61	-42.41	-29.56
1500	-4.81	-5.74	-4.23	-15.06	-18.27	-12.19	-28.06	-34.69	-22.18
2000	-4.02	-4.93	-3.44	-12.39	-15.58	-9.61	-22.94	-29.43	-17.21
2500	-3.49	-4.39	-2.92	-10.60	-13.78	-7.89	-19.55	-25.93	-13.93

UWI - uniform welfare improving, UWB - uniform welfare improving of born generations, UWU - uniform welfare improving of unborn generations

6.3.2 Economic Precision

Tables 6, 7, and 8 tell us about the cost of setting carbon policies that differ from the absolute optimum.[29] For all three damage functions, the message is clear. The growth rate of the tax is more important than its initial value. For example, an initial $10 tax will deliver close to the same UWI gain or at least a relatively high UWI gain if its annual growth rate is 4.5 percent. Moreover, starting the initial tax $10 below the optimum will not materially impact the size of the uniform efficiency gain. With 6x damages, the initial tax can start $30 from its optimum and still deliver close to 98 percent of the UWI gain. What about setting the optimal initial tax, but keeping it fixed? In the 1x case, this lowers the UWI gain by 27.4 percent. In the 6x case, it lowers it by only 4.1 percent.

The three tables also point to a surprisingly effective alternative carbon policy, namely starting with a high carbon tax and having it decline through time to zero. Take the 6x case. An initial $90 tax declining at 1.5 percent per year delivers a UWI gain of 3.809 percent. This is within 20 percent of the gain from the optimal policy.

Why would a high initial carbon tax that falls through time make sense? The answer is that there are two ways to limit the burning of fossil fuels. The first is to make the net-of-tax value of burning coal, oil, and gas sufficiently low relative to their extraction costs that reducing production makes sense. This is the case of a low, but rising carbon tax. It does not immediately shut off CO_2 emissions. It slows them down and shuts them off sooner than would otherwise be the case. This first mechanism involves limiting the burn over time to an increasing degree. The second mechanism is to give dirty energy producers an incentive to slow down their burning of fossil fuels not by making the return to production lower, but by raising the gain from deferring production. An initially quite high, but a declining carbon tax, provides the incentive to produce less in the present and more in the future. Implementing a high initial, but declining tax may be politically more palatable than starting with a lower tax that rises. However, the transfers to initial and early generations and the taxes on future generations needed to achieve the UWI optimum is very similar to what arises when the tax starts low, but rises.

6.3.3 The Efficiency Cost of Delay

As stressed by Sinn's Green Paradox Sinn (2012) and the large literature, e.g, Van der Ploeg and Withagen (2015), it spawned, a delay in implementing carbon policy can lead dirty energy producers to "use it or lose it," thus accelerating the extraction of fossil fuels.[30] Unlike a model

---

[29]In searching over optimal policy, we used a grid with a step of $ 10 per ton of CO_2 for the initial tax and a step of 1.5% for it's growth rate.

[30]This point was stressed, recently, in Kotlikoff et al. (2016) in conjunction with the Paris Accord.

Table 6: Welfare Gains with 1x Damage Function.

g_τ / τ_0	-1.5%	0%	1.5%	3.0%	4.5%
10	0.165	0.254	0.451	0.623	0.612
20	0.268	0.422	0.682	0.695	0.524
30	0.336	0.527	**0.726**	0.602	0.362
40	0.381	0.586	0.661	0.509	0.2
50	0.409	0.612	0.629	0.397	0.032
60	0.425	0.614	0.579	0.272	-0.142
70	0.431	0.599	0.513	0.138	-0.317
80	0.429	0.572	0.433	-0.004	-0.493
90	0.42	0.536	0.343	-0.151	-0.667

Table 7: Welfare Gains with 3x Damage Function

g_τ / τ_0	-1.5%	0%	1.5%	3.0%	4.5%
10	0.5	0.741	1.207	1.69	1.939
20	0.87	1.314	1.996	2.38	2.418
30	1.151	1.733	2.433	2.548	2.399
40	1.372	2.032	2.578	2.518	2.295
50	1.548	2.236	**2.585**	2.458	2.171
60	1.689	2.369	2.573	2.376	2.032
70	1.802	2.449	2.542	2.276	1.881
80	1.891	2.49	2.494	2.162	1.725
90	1.961	2.505	2.432	2.036	1.566

with zero extraction costs, our increasing extraction costs places a natural brake on the Green Paradox problem, but it doesn't eliminate it. To assess the magnitude of the problem in our model, we calculated the optimal two-part UWI carbon policy, but starting in year 20 rather than in year 0. Even though the policy start date is later, the UWI gains are provided, as before, to all generations alive or still to be born as of time 0. In the case of 1x damages, the year-20 20 UWI optimal tax starts at $50 ($20 higher than the starting tax were the optimal policy to start at time 0) and grows thereafter at 1.5 percent (the same rate as starting at time 0). This is higher by $10 than the optimal tax would have been in year 20 were it enacted starting in year 0. For the 6x damage case, the optimal tax starting in year 20 is $140 ($70 larger than the optimal year-0 $70 value), again growing at the same 1.5 percent. This is $46 higher than its year-20 value under the optimal UWI policy starting in year 0. The message, then, is that when carbon taxation is delayed, the tax needs to start at a higher value than would otherwise

Table 8: Welfare Gains with 6x Damage Function

τ_0 \ g_τ	-1.5%	0%	1.5%	3.0%	4.5%
10	0.922	1.336	2.038	2.757	3.237
20	1.625	2.389	3.433	4.146	4.387
30	2.168	3.161	4.281	4.568	4.505
40	2.602	3.711	4.585	4.609	4.475
50	2.953	4.089	4.641	4.61	4.409
60	3.239	4.338	4.674	4.58	4.318
70	3.47	4.494	**4.685**	4.524	4.207
80	3.658	4.586	4.675	4.448	4.084
90	3.809	4.635	4.648	4.355	3.952

be the case.

Thanks to waiting too long to limit emissions and to the exacerbating influence of the Green Paradox, UWI gains are reduced considerably by delay. In the 1x case, the UWI gain is 0.402 percent versus 0.726 percent. In the 6x case, the UWI gain is 2.09 percent versus 4.69 percent. In percentage terms, a 20-year delay reduces the efficiency gains by 44.6 percent in the 1x case and 55.4 percent in the 6x case. The message here is that waiting too long to address climate change is not only inefficient. It also reaches an economic tipping point where the UWI gains are potentially too small to achieve consensus on taking action.

7 Comparison of Optimal Carbon Policy – OLG Versus ILA

Standard integrated assessment models, e.g. Nordhaus (1994, 2008, 2010) and Golosov et al. (2014), determine the optimal time path of CO_2 emissions that maximize the welfare of a social planner (SP). The SP in these studies may be proxying for a model of infinitely lived agents (ILAs) whose government has a sufficient fiscal instrument to achieve the SP solution. An example, in this regard, is Golosov et al. (2014) who derives (decentralize) the optimal ILA carbon-tax policy based on the solution to the corresponding SP's problem. The ILA framework implicitly assumes that all currently living agents are homogenous and, as mentioned, altruistic in the sense that consumption of their children enters their utility function.[31]

One can also obviously consider a setting in which generations are selfish, but are controlled

---

[31]Indeed, the standard ILA formulation assumes that the utility of children enters the parent's utility function in the same way as their consumption.

by an SP. The SP would effect an efficient solution, but not necessarily a Pareto improvement relative to the no-SP equilibrium. Alternatively, one can contemplate a Pareto-constrained social planner who achieves alternative Pareto-improving, efficient equilibriums, such as our maximum UWI solution. In contrast to our UWI solution, which is limited due to the consideration of just two policy instruments, an SP would, in effect, have an unlimited number of instruments at her disposal. As in Golosov et al. (2014), one could take the SP's UWI solution as a benchmark and examine whether it can be decentralized with time-varying carbon and generation-specific lump-sum taxes. We reserve this exercise for future research.

This said we have explored, for our base case, using three rather than two fiscal instruments to improve on our UWI efficiency gain. The three instruments are an initial carbon tax, a short-run (50-year) growth rate, and a long-run (post year 50) growth rate. We find no change in the optimal UWI policy, i.e., the optimal short-term growth rate is the same as the long-term growth rate. The ILA optimal carbon policy is also unchanged when we add a short-term growth rate.[32]

Given the frequent reliance on the ILA for analyzing optimal carbon taxation, a natural question is whether the optimal two-part carbon policy differs in our OLG and ILA models, where the ILA model differs from the OLG model only through its assumption of intergenerational altruism. Table 9 presents our comparison. Consider first rows 3 and 4. They show optimal UWI OLG and ILA carbon policy assuming the 1.5 percent time preference rate and damage function assume by Nordhaus (2017). The optimal two-part policies are identical, namely a $30 initial tax growing annually at 1.5 percent.[33] However, since the ILA solution does not include the LSRA, it doesn't ensure a Pareto improvement. This, indeed, is what we see in the third column of table 3 – benefits to older and future generations at a cost to young and early generations. Moreover, the UWI OLG and ILA policies differ significantly when time preference rates are larger or smaller or when the damage function is more significant. Consider, for example, the 1x case with a time preference rate of 0.8 percent. The ILA optimal tax, in this case, is perpetually one third larger than the OLG maximizing UWI carbon tax. For a higher, 3 percent time preference rate, the OLG and ILA initial optimal taxes are both $10, but the tax rises at 3 percent annually in the OLG case and at 1.5 percent annually in the ILA case. After 50 years, the OLG carbon tax is twice the ILA carbon tax. After a century, it is 4.4 times larger.

---

[32]In the 1x damage case, the initial tax remains at $30 and the optimum short-term growth rate as well as long-term growth rates are both 1.5 percent, i.e., the extra policy instrument makes no difference. With 6x damages, the two-part policy entailed a $90 initial tax growing at 3 percent per year. With three instruments, the initial tax is again $90 with a short-term 3.0 percent growth rate followed by a 1.5 percent long-term growth rate. But the welfare gain with the third instrument is 7.2256 percent compared with 7.2255 percent without it. This trivial difference is within the range of computation precision.

[33]This surely reflects, in part, the fact that our BAU delivers an interest rate path that's comparable to that in the ILA equilibrium.

As another example of the significant differences in carbon policy prescriptions emanating from the two models, consider the 6x case with a time preference rate of 0.8 percent. The optimal ILA carbon tax starts at $140 and rises at 3 percent annually. In the OLG model, the carbon tax starts at $80 and rises at 1.5 percent annually. Once again, the difference in the two optimal carbon taxes rises dramatically through time.

One reason the OLG and ILA optimal policies may differ so much is their different general equilibrium paths. In the ILA, the economy's long-run interest rate is pegged, in significant part, by the rate of time preference. In the OLG model, the time preference rate matters, but so does generational policy, including the LSRA's debt accumulation policy.

As Table 10, the LSRA's debt policy raises the long-term interest rate. Take our base case. The interest rate starts at 4.3 percent and under BAU ends up at 4.0 percent. However, when we maximize the UWI gain, the long run interest rate is 5.4 percent. In the corresponding ILA simulations, the long-run interest rates are 4.5 percent. For the case of very large damages, intergenerational transfers and, thus, LSRA debt, must be very large to compensate current generations for the much higher requisite carbon taxation. With ρ at 1.5 percent and a 6x damage function, the OLG long-run interest rate is 8.5 percent. It is 4.1 percent under BAU and 4.5 percent under ILA BAU and ILA optimal policy. Clearly, the rise in the interest rate, which must be paid on the LSRA debt, naturally limits how large the carbon taxes can be today.

So far we have shown that the choice of model, OLG versus ILA, will, in general, make a significant difference in determining optimal carbon policy. However, will it matter to the intergenerational distribution of welfare? The answer is yes, as shown in figures 2 and 3. These figures show the impact on the time path of current and future generations' utilities of imposing ILA "optimums" on the OLG economy with no redistribution across generations to offset utility losses. If damages are 1x, the optimal ILA solution, when the ILA agents are presumed to have very low time preference rates, e.g., 0.1 percent, can materially hurt young and early generations. With 6x damages, the potential harm to such generations is smaller, but non trivial.

Table 9: Optimal Carbon Policy – OLG Versus ILA

Model	ρ	1x Damage Function			3x Damage Function			6x Damage Function		
		τ_0	g_τ	λ	τ_0	g_τ	λ	τ_0	g_τ	λ
OLG	0.8%	$30	1.5%	1.07%	$60	1.5%	3.25%	$80	1.5%	5.58%
ILA	0.8%	$40	1.5%	1.30%	$70	3%	5.72%	$140	3%	12.38%
OLG	1.5%	$30	1.5%	0.73%	$50	1.5%	2.58%	$70	1.5%	4.69%
ILA	1.5%	$30	1.5%	0.49%	$70	1.5%	3.15%	$90	3%	7.23%
OLG	3%	$10	3%	0.23%	$40	1.5%	1.26%	$50	1.5%	4.64%
ILA	3%	$10	1.5%	0.07%	$40	1.5%	0.99%	$40	3%	2.76%

ρ references the time preference rate, g_P the long-run annual percentage change in the price of energy, τ_0 the initial carbon tax, g_τ the growth rate of the carbon tax, and λ the welfare gain measured as the equivalent percentage increase in consumption under BAU needed to achieve the welfare gain under the optimal UWI policy in the OLG case or the ILA optimal in the ILA case.

Table 10: Interest Rates (percent per year)

	t=0	t=10	t=50	t=100	t=150	t=200	t=250
OLG, BAU, 1x DF, $\rho = 0.8\%$	4.3	4.0	3.3	3.3	3.5	3.6	3.6
OLG, OP, 1x DF, $\rho = 0.8\%$	4.2	4.1	3.8	4.4	4.9	5.0	5.0
ILA, BAU, 1x DF, $\rho = 0.8\%$	4.3	3.6	3.4	3.5	3.7	3.7	3.7
ILA, OP, 1x DF, $\rho = 0.8\%$	4.1	3.5	3.5	3.6	3.7	3.7	3.7
OLG, BAU, 1x DF, $\rho = 1.5\%$	4.3	4.3	3.7	3.8	4.0	4.0	4.0
OLG, OP, 1x DF, $\rho = 1.5\%$	4.2	4.4	4.2	4.8	5.3	5.4	5.4
ILA, BAU, 1x DF, $\rho = 1.5\%$	4.3	4.1	4.1	4.2	4.4	4.4	4.5
ILA, OP, 1x DF, $\rho = 1.5\%$	4.2	4.1	4.2	4.3	4.4	4.5	4.5
OLG, BAU, 1x DF, $\rho = 3\%$	4.4	5.3	5.0	5.2	5.4	5.4	5.4
OLG, OP, 1x DF, $\rho = 3\%$	4.3	5.4	5.2	5.8	6.2	6.3	6.3
ILA, BAU, 1x DF, $\rho = 3\%$	4.4	5.3	5.6	5.7	5.9	6.0	6.0
ILA, OP, 1x DF, $\rho = 3\%$	4.3	5.3	5.7	5.8	6.0	6.0	6.0
OLG, BAU, 6x DF, $\rho = 0.8\%$	4.2	3.9	2.9	2.9	3.4	3.5	3.6
OLG, OP, 6x DF, $\rho = 0.8\%$	3.9	4.4	5.0	7.1	8.2	8.4	8.4
ILA, BAU, 6x DF, $\rho = 0.8\%$	4.2	3.4	2.9	3.0	3.6	3.7	3.7
ILA, OP, 6x DF, $\rho = 0.8\%$	3.8	3.4	3.5	3.7	3.7	3.7	3.7
OLG, BAU, 6x DF, $\rho = 1.5\%$	4.2	4.2	3.3	3.4	3.9	4.0	4.1
OLG, OP, 6x DF, $\rho = 1.5\%$	3.9	4.6	5.3	7.3	8.3	8.5	8.5
ILA, BAU, 6x DF, $\rho = 1.5\%$	4.2	4.0	3.6	3.7	4.3	4.4	4.5
ILA, OP, 6x DF, $\rho = 1.5\%$	3.9	4.0	4.0	4.4	4.4	4.5	4.5
OLG, BAU, 6x DF, $\rho = 3\%$	4.3	5.2	4.5	4.8	5.3	5.4	5.4
OLG, OP, 6x DF, $\rho = 3\%$	4.0	5.5	6.2	8.1	9.0	9.2	9.2
ILA, BAU, 6x DF, $\rho = 3\%$	4.3	5.1	5.1	5.3	5.9	6.0	6.0
ILA, OP, 6x DF, $\rho = 3\%$	4.1	5.1	5.5	5.9	6.0	6.0	6.0

*NP - No Policy, OP - Optimal Policy, DF - Damage Function

Optimal tax in OLG without the LSRA being operational
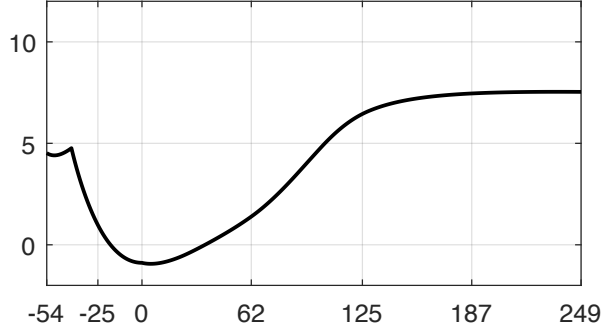
Optimal tax in OLG with the LSRA being operational
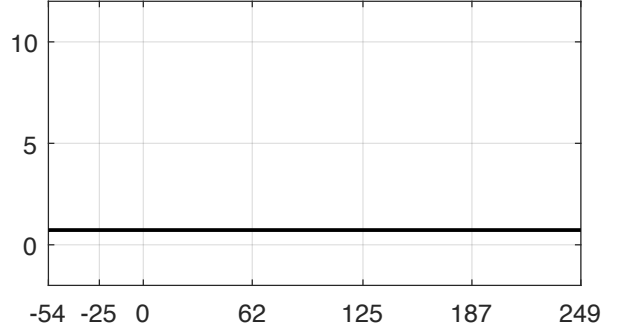
Optimal tax in ILA with $\rho = 0.1\%$
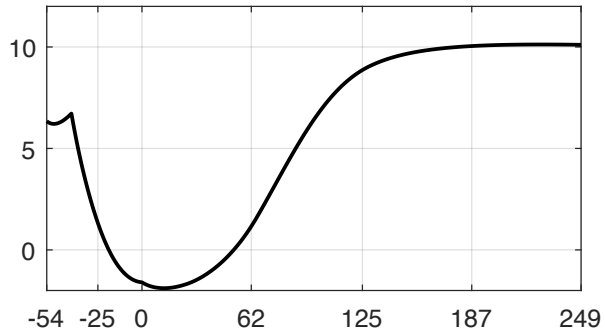
Optimal tax in ILA with $\rho = 1.5\%$
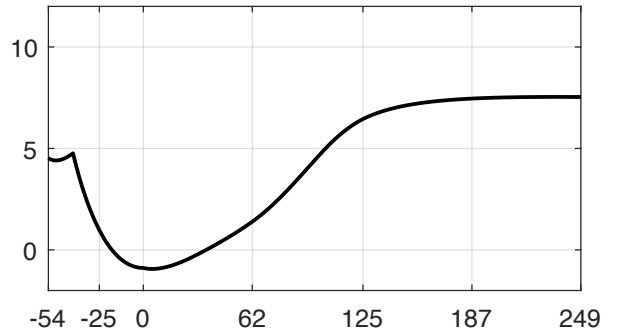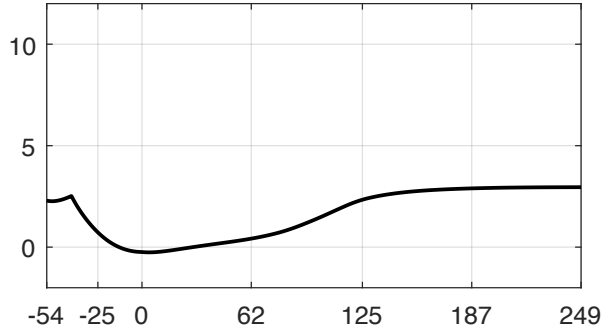

Optimal tax in ILA with $\rho = 3\%$

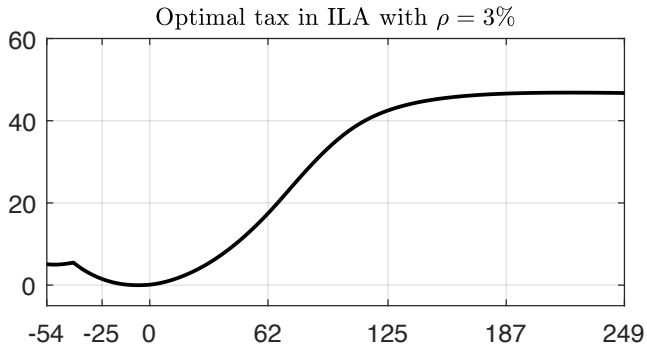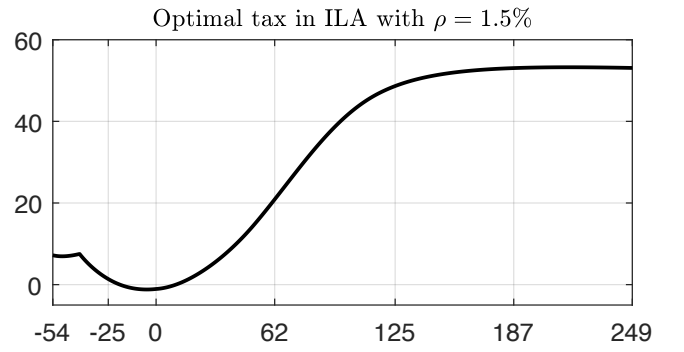Figure 2: Welfare Gains, OLG with Optimal Taxation vs ILA Solution with 1x Damage Function
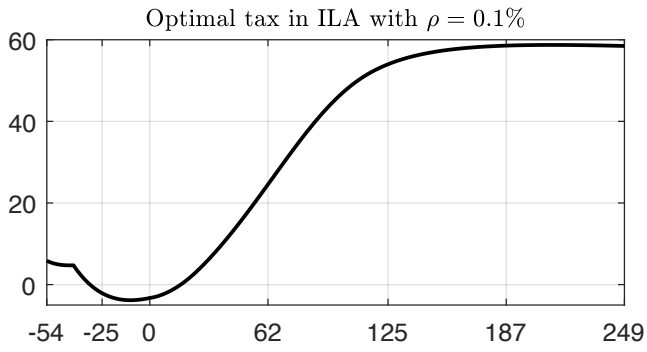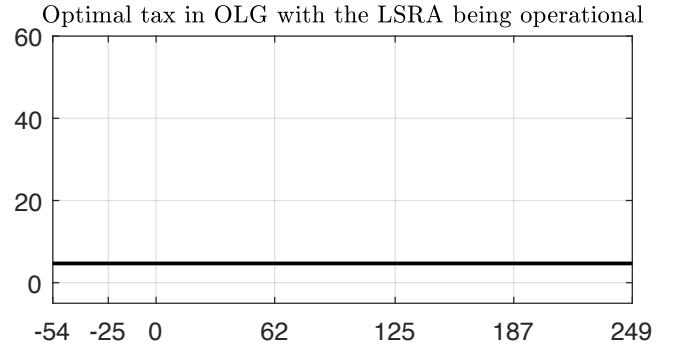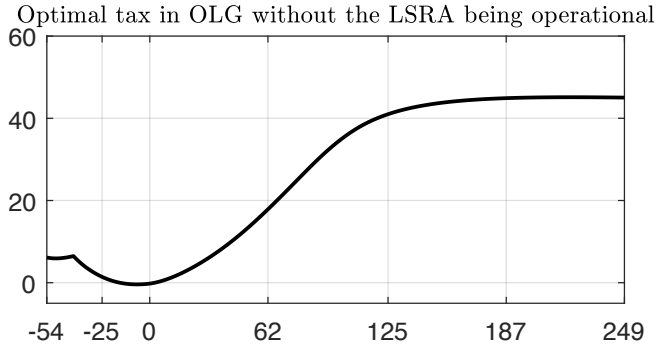
Figure 3: Welfare Gains, OLG with Optimal Taxation vs ILA Solution with 6x Damage Function

8 The Cost of Imposing Carbon Policy If Damages Are, in Fact, Zero

Despite the overwhelming scientific evidence of climate change, climate skeptics abound. Hence, it seems worthwhile to address their view that imposing carbon taxation would be costly with no benefit. If damagers were, indeed, zero and the government imposed the 1x UWI policy, the uniform, across all generations, efficiency loss would equal 0.54 percent. With 3x policy, it would equal 0.87 percent. And with 6x policy, it would equal 1.04 percent. These values are all smaller or far smaller than the far more likely respective gains of 0.73 percent, 2.58 percent, and 4.69 percent.

9 Summary and Conclusion

Large-scale modeling of optimal carbon policy has been framed, almost exclusively, in two ways. The first invokes an omnipotent social planner whose preferences about the treatment of different generations need bear no connection to actual preferences of either those alive or those yet to be born. The second assumes that households are infinitely lived, ostensibly due to altruistic concerns for their descendants. Each approach is problematic. Neither "optimums" need produce a Pareto improvement if cohorts are selfish. Hence, relying on either approach can transform the debate about efficient carbon policy from one of a generational win-win to one of a generational win-lose.

The appropriate framework for studying the carbon externality is surely that of selfish cohorts who quite naturally ignore the costs their CO_2 emissions impose on today's and tomorrow's children. In focusing on generational win-wins, principally the carbon policy producing the highest uniform welfare gain for all generations, our paper confines optimal carbon taxation to the province of economics. I.e., we leave aside carbon policies that are "just" based on the preferences of an assumed social planner or those of agents, all of whom supposedly care about future generations, but somehow collectively fail to assist them.

When we calibrate the emissions/temperature/damages relationships per Nordhaus (2017), the optimal carbon tax is 30 per cubic ton of CO_2 rising at 1.5 percent per year. With a six-times worse damage function, the optimal initial carbon tax is 70, also growing at 1.5 percent per year. In the analogous ILA model, the optimal 1x damage function tax is also 30 rising at 1.5 percent. With a 6x damage function, it is 90, growing at 3 percent annually. Hence, optimal OLG UWI and ILA carbon policies can be very different. Moreover, even in the 1x

case, where the carbon tax path is the same, implementing the ILA optimum, which entails no intergenerational redistribution, will not be Pareto improving if cohorts are selfish.

In our OLG optimal carbon policy calculations, which include lump-sum redistribution, the uniform welfare gains with 1x and 6x damage functions are 0.73 percent and 4.69 percent, respectively, measured as consumption compensating variations in the business-as-usual equilibrium. Achieving these win-win outcomes does, however, require significant redistribution from future winners and transfers to early losers. Indeed, with 6x damages, the net taxes on future generations can reach up to half of their annual consumption. This is testimony to the magnitude of the negative carbon externality.

We have, as mentioned, many important first-order issues to explore in future work – issues that have already been examined in SP and ILA frameworks and even some steady-state OLG models. We have excluded these factors here to keep our model simple and to focus on our key objective – calculating optimal, Pareto-improving carbon taxation and its associated intergenerational redistribution in a large-scale OLG model that incorporates standard (Nordhaus (2017)) assumptions about the relationship between carbon emissions, temperature increase and economic damages.

Our goals for future research are, first and foremost, incorporating uncertainty and deriving state–specific carbon tax and state-specific intergenerational redistribution that uniformly raise all generations' expected utilities. Second is adding multiple regions together with region-specific damages, some of which may be negative, as well as regional free riding. Third is accounting for carbon taxation's regressivity. Doing so requires expanding the model to include multiple goods, some of which are energy intensive as well as differentially consumed by the poor. It also requires within-cohort labor earnings heterogeneity. Fourth is incorporating pre-existing fiscal distortions whose reduction, via the carbon policy, may permit even larger Pareto improvements – the so-called carbon double dividend.

With inter-generational, intra-generational, inter-regional heterogeneity, pre-existing distortions, and limitations on lump-sum (equivalently, inframarginal) taxes and transfers, achieving realistic win-win solutions requires careful design of tax-transfer schemes where such schemes rely either exclusively or primarily on extant fiscal instruments. But calculating such taxes and transfers within large-scale OLG models seems critical for identifying carbon policy changes that are of universal benefit and, therefore, far more likely to be adopted on a planet-wide basis.

References

Abel, A. and Kotlikoff, L. J. (1994). Intergenerational altruism and the effectiveness of fiscal policy: New tests based on cohort data. *Savings and bequests*, pages 31–42.

Acemoglu, D., Aghion, P., Bursztyn, L., and Hemous, D. (2012). The environment and directed technical change. *The American Economic Review*, 102(1):131–166.

Altig, D., Auerbach, A. J., Kotlikoff, L. J., Smetters, K. A., and Walliser, J. (2001). Simulating Fundamental Tax Reform in the United States. *American Economic Review*, pages 574–595.

Altonji, J. G., Hayashi, F., and Kotlikoff, L. J. (1992). Is the extended family altruistically linked? direct tests using micro data. *The American Economic Review*, pages 1177–1198.

Altonji, J. G., Hayashi, F., and Kotlikoff, L. J. (1997). Parental altruism and inter vivos transfers: Theory and evidence. *Journal of political economy*, 105(6):1121–1166.

Ansuategi, A. and Escapa, M. (2002). Economic growth and greenhouse gas emissions. *Ecological Economics*, 40(1):23–37.

Auerbach, A. J. and Kotlikoff, L. J. (1983). National Savings, Economic Welfare, and the Structure of Taxation. In *Behavioral Simulation Methods in Tax Policy Analysis*, pages 459–498. University of Chicago Press.

Auerbach, A. J. and Kotlikoff, L. J. (1987). *Dynamic fiscal policy*, volume 11. Cambridge University Press Cambridge.

Barro, R. J. (1974). Are government bonds net wealth? *Journal of political economy*, 82(6):1095–1117.

Batina, R. G. and Krautkraemer, J. a. (1999). On sustainability and intergenerational transfers with a renewable resource. *Land Economics*, 75(2):167–184.

Benzell, S. G., Goryunov, E., Kazakova, M., Kotlikoff, L. J., LaGarda, G., Nesterova, K., and Zubarev, A. (2015). Simulating russia's and other large economies' challenging and interconnected transitions. Technical report, National Bureau of Economic Research.

Benzell, S. G., Kotlikoff, L. J., and LaGarda, G. (2017). Simulating business cash flow taxation: An illustration based on the "better way" corporate tax reform. Technical report, National Bureau of Economic Research.

Bernheim, B. D. and Bagwell, K. (1988). Is everything neutral? *The Journal of Political Economy*, 96(2):308–338.

Boden, T. A., Andres, R. J., and Marland, G. (2017). Global, regional, and national fossil-fuel co2 emissions (1751 - 2014) (v. 2017).

Börsch-Supan, A., Ludwig, A., and Winter, J. (2006). Ageing, pension reform and capital flows: A multi-country simulation model. *Economica*, 73(292):625–658.

Bovenberg, A. and Heijdra, B. J. (1998). Environmental tax policy and intergenerational distribution. *Journal of Public Economics*, 67(1):1–24.

Bovenberg, A. L. and Heijdra, B. J. (2002). Environmental abatement and intergenerational distribution. *Environmental and Resource Economics*, 23(1):45–84.

Bréchet, T., Gerard, F., Tulkens, H., et al. (2011). Efficiency vs. stability in climate coalitions: a conceptual and computational appraisal. *Energy Journal*, 32(1):49–75.

Brock, W. A. and Hansen, L. P. (2017). Wrestling with uncertainty in climate economic models.

Burton, P. S. (1993). Intertemporal preferences and intergenerational equity considerations in optimal resource harvesting. *Journal of Environmental Economics and Management*, 24(2):119–132.

Cai, Y., Brock, W., Xepapadeas, A., and Judd, K. (2018). Climate policy under cooperation and competition between regions with spatial heat transport. Technical report, National Bureau of Economic Research.

Cai, Y., Judd, K. L., and Lontzek, T. S. (2012). Continuous-time methods for integrated assessment models. Technical report, National Bureau of Economic Research.

Cai, Y., Judd, K. L., and Lontzek, T. S. (2013). The social cost of stochastic and irreversible climate change. Technical report, National Bureau of Economic Research.

Cai, Y. and Lontzek, T. (2018). The social cost of carbon with economic and climate risks. *Journal of Political Economy*.

Calvo, G. A. and Obstfeld, M. (1988). Optimal time-consistent fiscal policy with finite lifetimes. *Econometrica*, 56(2):411–432.

Endress, L. H., Pongkijvorasin, S., Roumasset, J., and Wada, C. A. (2014). Intergenerational equity with individual impatience in a model of optimal and sustainable growth. *Resource and Energy Economics*, 36(2):620–635.

Fehr, H., Jokisch, S., and Kotlikoff, L. (2003). The Developed World's Demographic Transition - The Roles of Capital Flows, Immigration, and Policy. (10096).

Gerlagh, R. and Keyzer, M. A. (2001). Sustainability and the intergenerational distribution of natural resource entitlements. *Journal of Public Economics*, 79(2):315–341.

Gerlagh, R. and van der Zwaan, B. (2001). The effects of ageing and an environmental trust fund in an overlapping generations model on carbon emission reductions. *Ecological Economics*, 36(2):311–326.

Gokhale, J., Kotlikoff, L. J., and Sabelhaus, J. (1996). Understanding the postwar decline in us saving: A cohort analysis. Technical report, National Bureau of Economic Research.

Golosov, M., Hassler, J., Krusell, P., and Tsyvinski, A. (2014). Optimal taxes on fossil fuel in general equilibrium. *Econometrica*, 82(1):41–88.

Goulder, L. H. (1995). Environmental taxation and the double dividend: a reader's guide. *International tax and public finance*, 2(2):157–183.

Gurgel, A. C., Paltsev, S., Reilly, J., Metcalf, G., et al. (2011). An analysis of us greenhouse gas cap-and-trade proposals using a forward-looking economic model. *Environment and Development Economics*, 16(2):155–176.

Hayashi, F., Altonji, J., and Kotlikoff, L. (1996). Risk-sharing between and within families. *Econometrica*, 64(2):261–294.

Heijdra, B. J., Kooiman, J. P., and Ligthart, J. E. (2006). Environmental quality, the macroeconomy, and intergenerational distribution. *Resource and Energy Economics*, 28(1):74–104.

Hotelling, H. (1931). The economics of exhaustible resources. *The journal of political economy*, pages 137–175.

Howarth, R. B. (1991a). Intergenerational competitive equilibria under technological uncertainty and an exhaustible resource constraint. *Journal of Environmental Economics and Management*, 21(3):225–243.

Howarth, R. B. (1991b). Intertemporal equilibria and exhaustible resources: an overlapping generations approach. *Ecological Economics*, 4(3):237–252.

Howarth, R. B. (1998). An overlapping generations model of climate-economy interactions. *The Scandinavian Journal of Economics*, 100(3):575–591.

Howarth, R. B. and Norgaard, R. B. (1990). Intergenerational resource rights, efficiency, and social optimality. *Land Economics*, 66(1):1–11.

Howarth, R. B. and Norgaard, R. B. (1992). Environmental valuation under sustainable development. *American Economic Review*, 82(2):473–477.

John, A., Pecchenino, R., Schimmelpfennig, D., and Schreft, S. (1995). Short-lived agents and the long-lived environment. *Journal of Public Economics*, 58(1):127–141.

Karp, L. and Rezai, A. (2014). The political economy of environmental policy with overlapping generations. *International Economy Review*, 55(3):711–733.

Kotlikoff, L. J. (1983). Altruistic linkages within the extended family, a note. *Sloan Foundation Proposal*.

Kotlikoff, L. J., Polbin, A., and Zubarev, A. (2016). Will the paris accord accelerate climate change? Technical report, National Bureau of Economic Research.

Kotlikoff, L. J., Smetters, K., and Walliser, J. (2007). Mitigating america's demographic dilemma by pre-funding social security. *Journal of monetary Economics*, 54(2):247–266.

Krusell, P. and Smith Jr, A. A. (2015). Climate change around the world. In *Work in progress, Walras-Bowley Lecture. Econometric Society 2015 World Congress, Montréal*.

Lemoine, D. and Traeger, C. (2014). Watch your step: optimal policy in a tipping climate. *American Economic Journal: Economic Policy*, 6(1):137–66.

Lipton, D., Poterba, J., Sachs, J., and Summers, L. (1982). Multiple shooting in rational expectations models. *Econometrica*, 50(5):1329–1333.

Lontzek, T. S., Cai, Y., Judd, K. L., and Lenton, T. M. (2015). Stochastic integrated assessment of climate tipping points indicates the need for strict climate policy. *Nature Climate Change*, 5(5):441–444.

Lugovoy, O. and Polbin, A. (2016). On intergenerational equity of greenhouse gases emissions abatement. *The Journal of the New Economic Association*, 31(3).

Manne, A., Mendelsohn, R., and Richels, R. (1995). Merge: A model for evaluating regional and global effects of ghg reduction policies. *Energy policy*, 23(1):17–34.

Marini, G. and Scaramozzino, P. (1995). Overlapping generations and environmental control. *Journal of Environmental Economics and Management*, 29(1):64–77.

McGlade, C. and Ekins, P. (2015). The geographical distribution of fossil fuels unused when limiting global warming to 2 c. *Nature*, 517(7533):187.

Metcalf, G. E. (2010). Investment in energy infrastructure and the tax code. *Tax Policy and the Economy*, 24:1–33.

Metcalf, G. E. (2014). Using the tax system to address competition issues with a carbon tax. *National Tax Journal*, 67(4):779–806.

Mourmouras, A. (1991). Competitive equilibria and sustainable growth in a life-cycle model with natural resources. *The Scandinavian Journal of Economics*, 93(4):585–591.

Mourmouras, A. (1993). Conservationist government policies and intergenerational equity in an overlapping generations model with renewable resources. *Journal of Public Economics*, 51(2):249–268.

Nordhaus, W. D. (1979). *Efficient use of energy resources*. Yale University Press,New Haven, CT.

Nordhaus, W. D. (1994). *Managing the global commons: the economics of climate change*. MIT press Cambridge, MA.

Nordhaus, W. D. (2008). *A Question of Balance: Weighing the Options on Global Warming Policies*. Yale University Press, New Haven, CT.

Nordhaus, W. D. (2010). Economic aspects of global warming in a post-copenhagen environment. *Proceedings of the National Academy of Sciences*, 107(26):11721–11726.

Nordhaus, W. D. (2015). Climate clubs: overcoming free-riding in international climate policy. *The American Economic Review*, 105(4):1339–1370.

Nordhaus, W. D. (2017). Revisiting the social cost of carbon. *Proceedings of the National Academy of Sciences*, page 201609244.

Nordhaus, W. D. and Boyer, J. (2000). *Warming the world: economic modeling of global warming*. MIT press Cambridge, MA.

Nordhaus, W. D. and Yang, Z. (1996). A regional dynamic general-equilibrium model of alternative climate-change strategies. *The American Economic Review*, pages 741–765.

Ortiz, R. A., Golub, A., Lugovoy, O., Markandya, A., and Wang, J. (2011). Dicer: A tool for analyzing climate policies. *Energy Economics*, 33:S41–S49.

Pecchenino, R. and John, A. (1994). An overlapping generations model of growth and the environment. *The Economic Journal*, 104(427):1393–1410.

Pindyck, R. S. (2013). Climate change policy: what do the models tell us? *Journal of Economic Literature*, 51(3):860–72.

Plambeck, E. L., Hope, C., and Anderson, J. (1997). The page95 model: Integrating the science and economics of global warming. *Energy Economics*, 19(1):77–101.

Popp, D. (2004). Entice: endogenous technological change in the dice model of global warming. *Journal of Environmental Economics and management*, 48(1):742–768.

Rasmussen, T. N. (2003). Modeling the economics of greenhouse gas abatement: an overlapping generations perspective. *Review of Economic Dynamics*, 6(1):99–119.

Rausch, S. (2013). Fiscal consolidation and climate policy: An overlapping generations perspective. *Energy Economics*, 40:S134 – S148. Supplement Issue: Fifth Atlantic Workshop in Energy and Environmental Economics.

Rausch, S., Metcalf, G. E., and Reilly, J. M. (2011). Distributional impacts of carbon pricing: A general equilibrium approach with micro-data for households. *Energy Economics*, 33:S20–S33.

Rausch, S. and Yonezawa, H. (2018). The intergenerational incidence of green tax reform. *Climate Change Economics*, 9(01):1840007.

Rogelj, J., Hare, W., Lowe, J., Van Vuuren, D. P., Riahi, K., Matthews, B., Hanaoka, T., Jiang, K., and Meinshausen, M. (2011). Emission pathways consistent with a 2 c global temperature limit. *Nature Climate Change*, 1(8):413.

Sachs, J. D. (2014). Climate change and intergenerational well-being. *The Oxford handbook of the macroeconomics of global warming*, pages 248–259.

Schneider, M. T., Traeger, C. P., and Winkler, R. (2012). Trading off generations: equity, discounting, and climate change. *European Economic Review*, 56(8):1621–1644.

Sinn, H.-W. (2008). Public policies against global warming: a supply side approach. *International Tax and Public Finance*, 15(4):360–394.

Sinn, H.-W. (2012). *The green paradox: a supply-side approach to global warming.* MIT press.

Solow, R. M. (1974a). The economics of resources or the resources of economics. *The American Economic Review*, 64(2):1–14.

Solow, R. M. (1974b). Intergenerational equity and exhaustible resources. *The Review of Economic Studies*, 41:29–45.

Stern, N. H. (2007). *The economics of climate change: the Stern review.* Cambridge University Press.

Summers, L. H. (1981). Capital Taxation and Accumulation in a Life Cycle Growth Model. *The American Economic Review*, 71(4):533–544.

Tol, R. S. (1997). On the optimal control of carbon dioxide emissions: an application of fund. *Environmental Modeling & Assessment*, 2(3):151–163.

Tol, R. S. (2002). Welfare specifications and optimal control of climate change: an application of fund. *Energy Economics*, 24(4):367–376.

Tol, R. S., Heintz, R. J., and Lammers, P. E. (2003). Methane emission reduction: An application of fund. *Climatic Change*, 57(1-2):71–98.

Van der Ploeg, F. and Withagen, C. (2015). Global warming and the green paradox: A review of adverse effects of climate policies. *Review of Environmental Economics and Policy*, 9(2):285–303.

van der Zwaan, B. C., Gerlagh, R., Schrattenholzer, L., et al. (2002). Endogenous technological change in climate change modelling. *Energy economics*, 24(1):1–19.

Weitzman, M. (2007). Role of uncertainty in the economics of catastrophic climate change.

Wendner, R. (2001). An applied dynamic general equilibrium model of environmental tax reforms and pension policy. *Journal of Policy Modeling*, 23(1):25–50.

Yang, Z. (2008). *Strategic bargaining and cooperation in greenhouse gas mitigations: An integrated assessment modeling approach.* MIT Press.

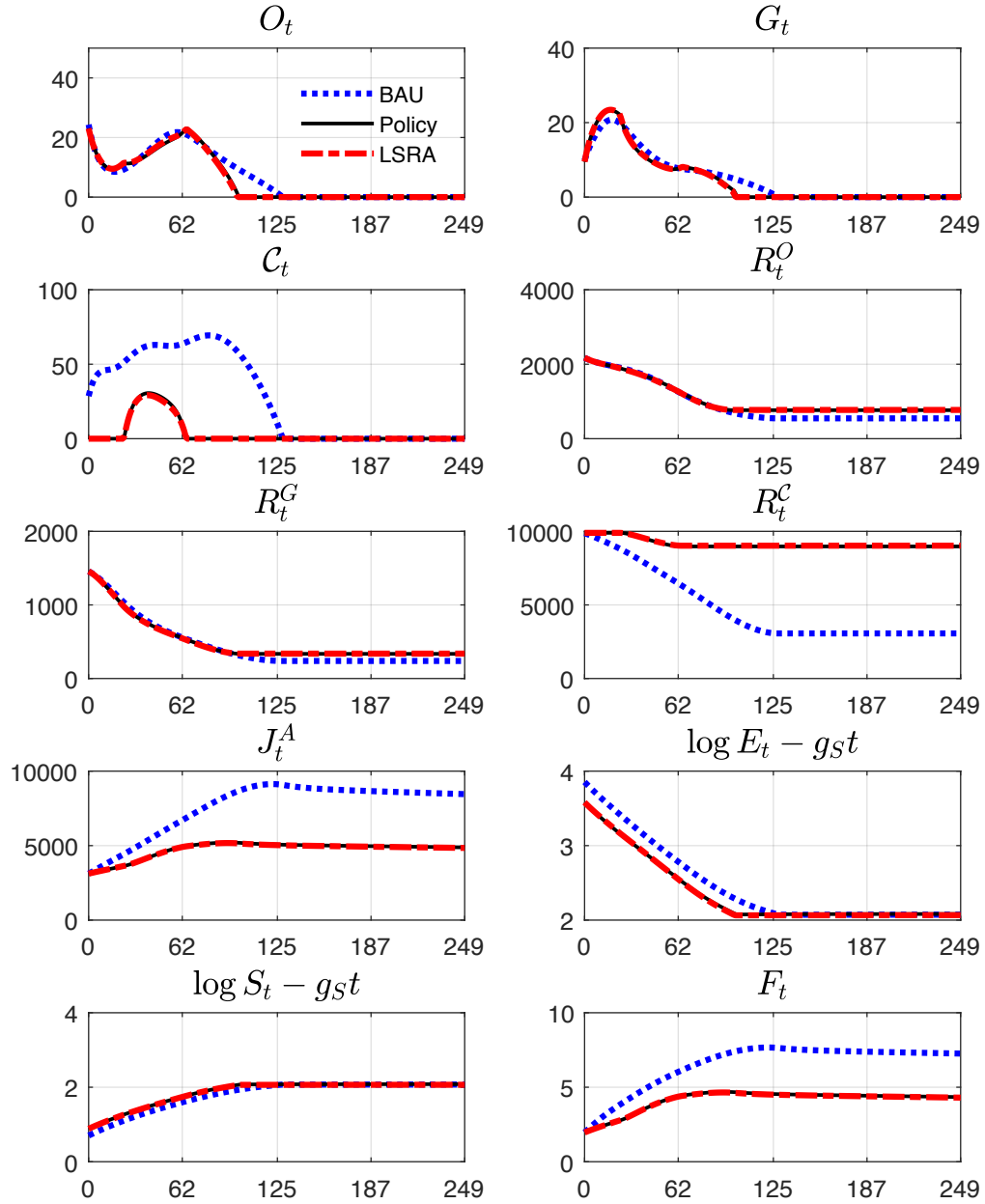# Appendix 1   Optimal Tax Policy with 1x Damage Function

Figure 4: Optimal Tax Policy with 1x Damage Function, Part 1*

*O_t - oil, G_t - gas, $\mathcal{C}_t$ -coal, R_t^O -oil reserves, R_t^G - gas reserves, $R_t^{\mathcal{C}}$ -coal reserves, J_t^A - CO_2 concentration in atmosphere (all 7 in $GtCO_2$), S_t - clean energy production,E_t - total energy consumption, F_t - radiative forcing, g_Y - steady state growth rate of output, g_S - steady state growth rate of clean energy.
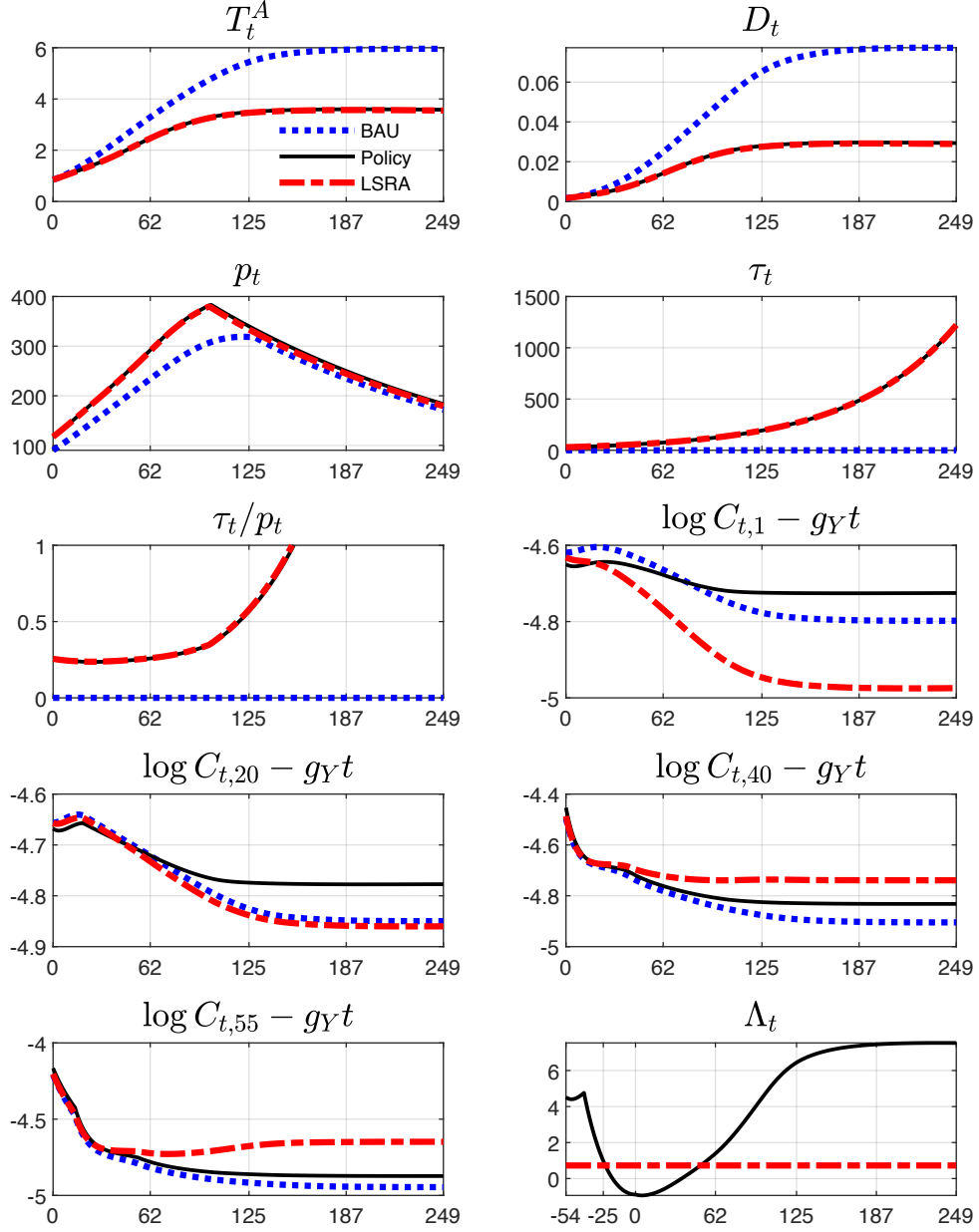
Figure 5: Optimal Tax Policy with 1x Damage Function, Part 2

*T_t^A - global mean surface temperature, D_t - damages, p_t - price of clean energy and oil per tCO_2, τ_t -absolute tax per tCO_2, $C_{t,j}$ -consumption of generation j at time t, g_Y - steady state growth rate of output, Λ_t - percentage change in consumption of households that produces the same change in the welfare of a generation born in period t compared to no policy scenario.
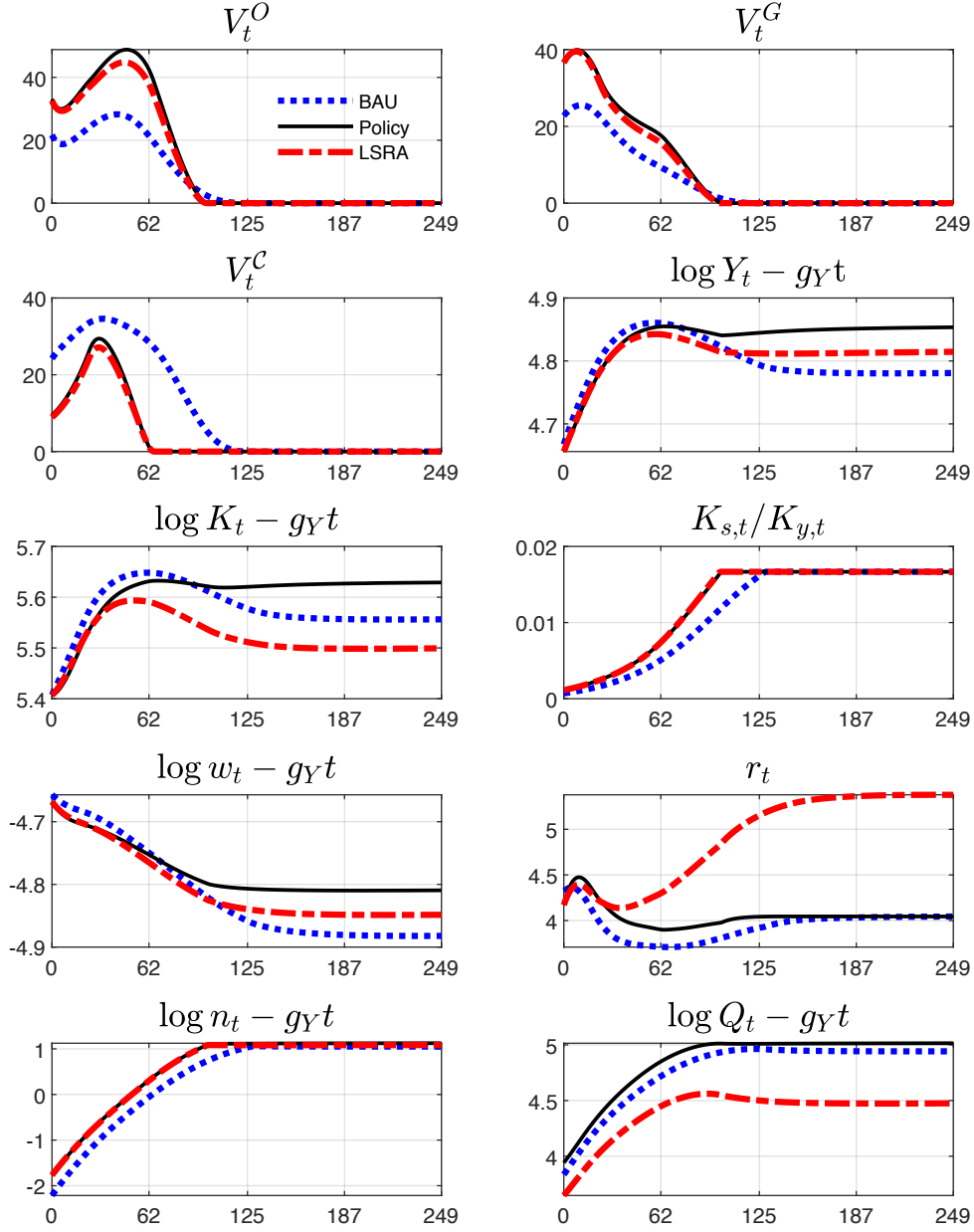
Figure 6: Optimal Tax Policy with 1x Damage Function, Part 3

* V_t^O - value of oil company, V_t^G - value of gas company, V_t^C - value of coal company, Y_t - final output, K_t - total capital, $K_{s,t}$ - clean energy sector's capital, $K_{y,t}$ - final goods production sector's capital, w_t - real wage rate, r_t - real interest rate, n_t - rental price of land, Q_t - value of land, g_Y - steady state growth rate of output.

Appendix 2   Optimal Tax Policy with 3x Damage Function
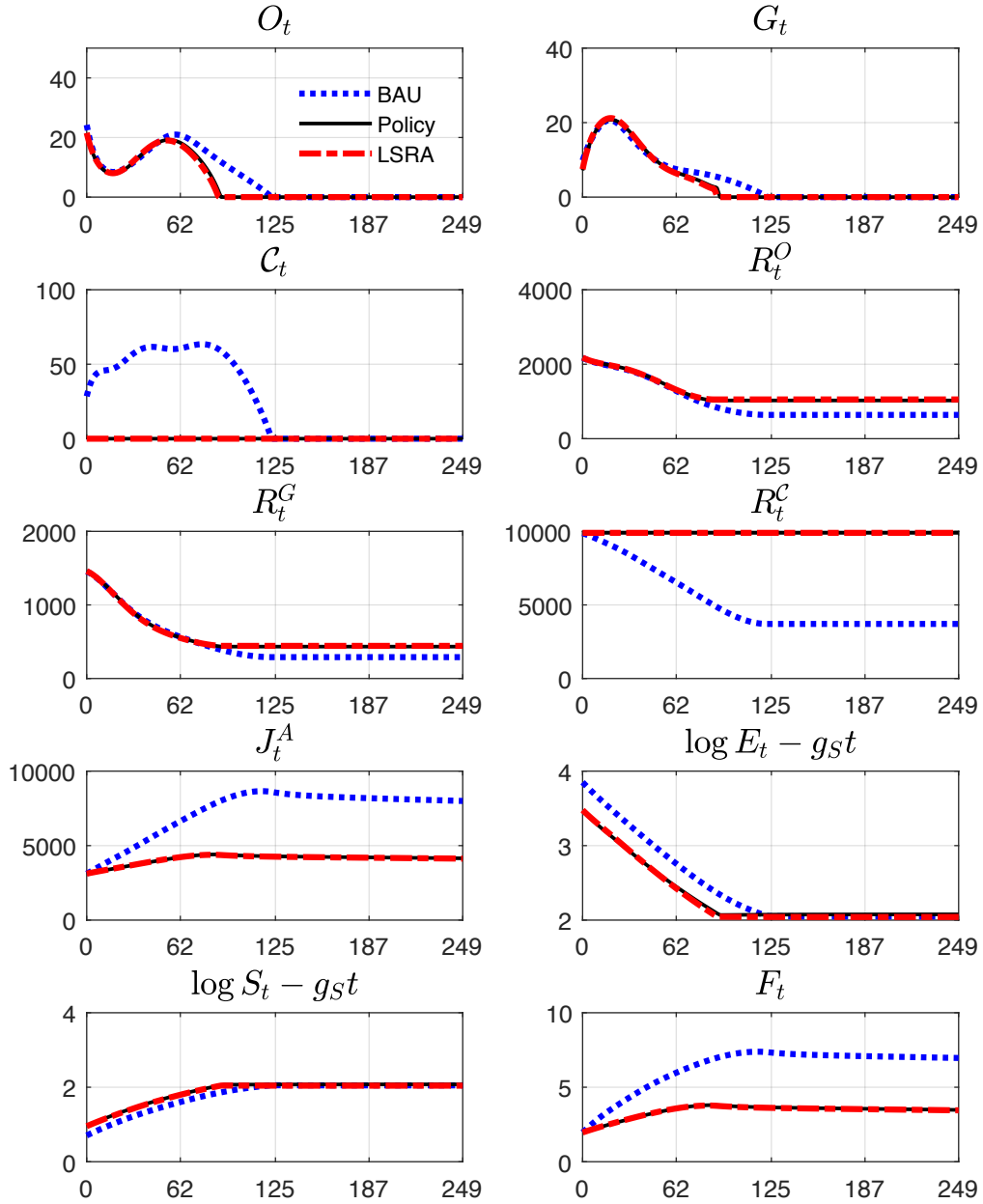
Figure 7: Optimal Tax Policy with 3x Damage Function, Part 1*

*O_t - oil, G_t - gas, $\mathcal{C}_t$ -coal, R_t^O -oil reserves, R_t^G - gas reserves, $R_t^{\mathcal{C}}$ -coal reserves, J_t^A - CO_2 concentration in atmosphere (all 7 in GtCO_2), S_t - clean energy production,E_t - total energy consumption, F_t - radiative forcing, g_Y - steady state growth rate of output, g_S - steady state growth rate of clean energy.
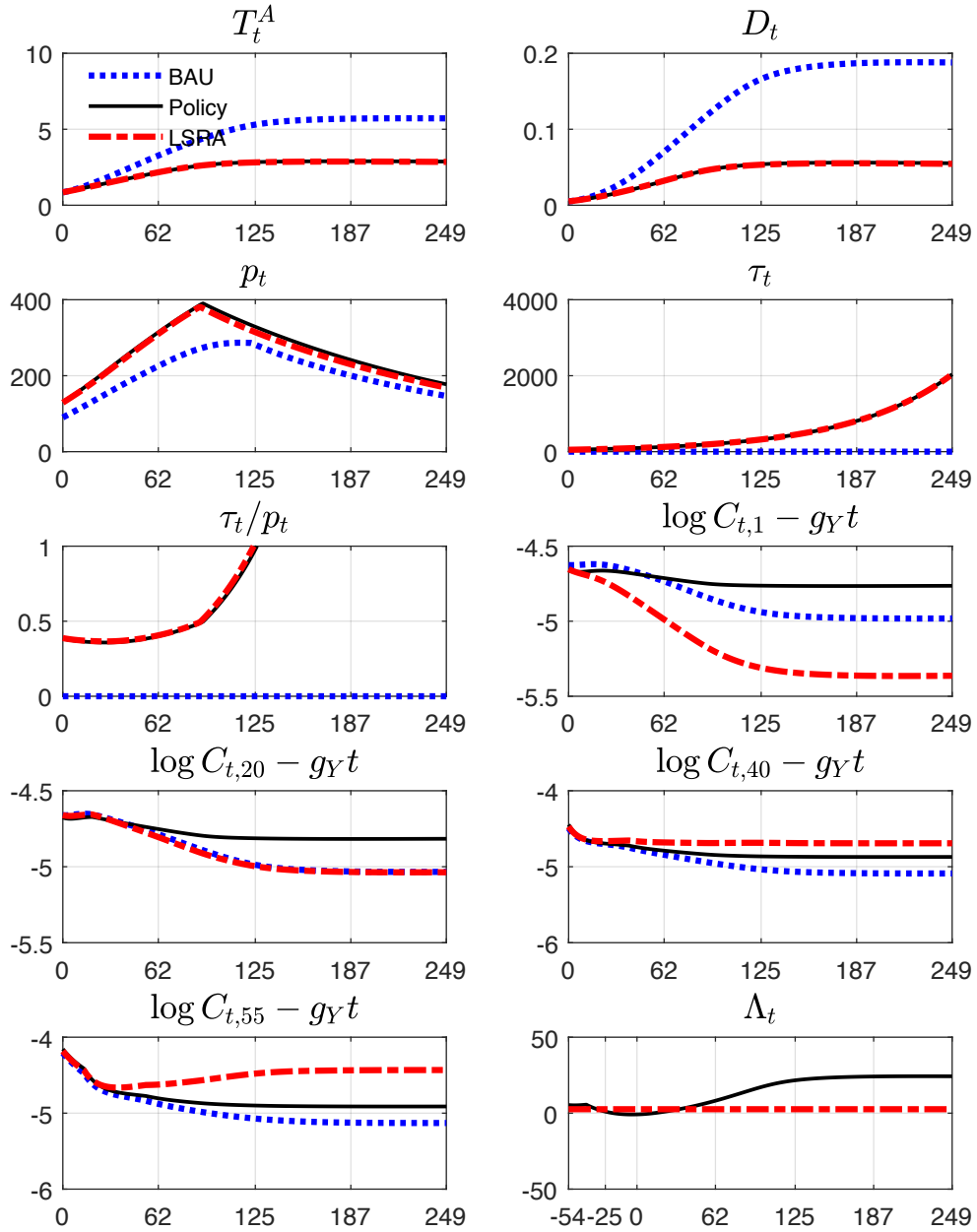
Figure 8: Optimal Tax Policy with 3x Damage Function, Part 2

*T_t^A - global mean surface temperature, D_t - damages, p_t - price of clean energy and oil per tCO_2, τ_t -absolute tax per tCO_2, $C_{t,j}$ -consumption of generation j at time t. g_Y - steady state growth rate of output, Λ_t - percentage change in consumption of households that produces the same change in the welfare of a generation born in period t compared to no policy scenario.
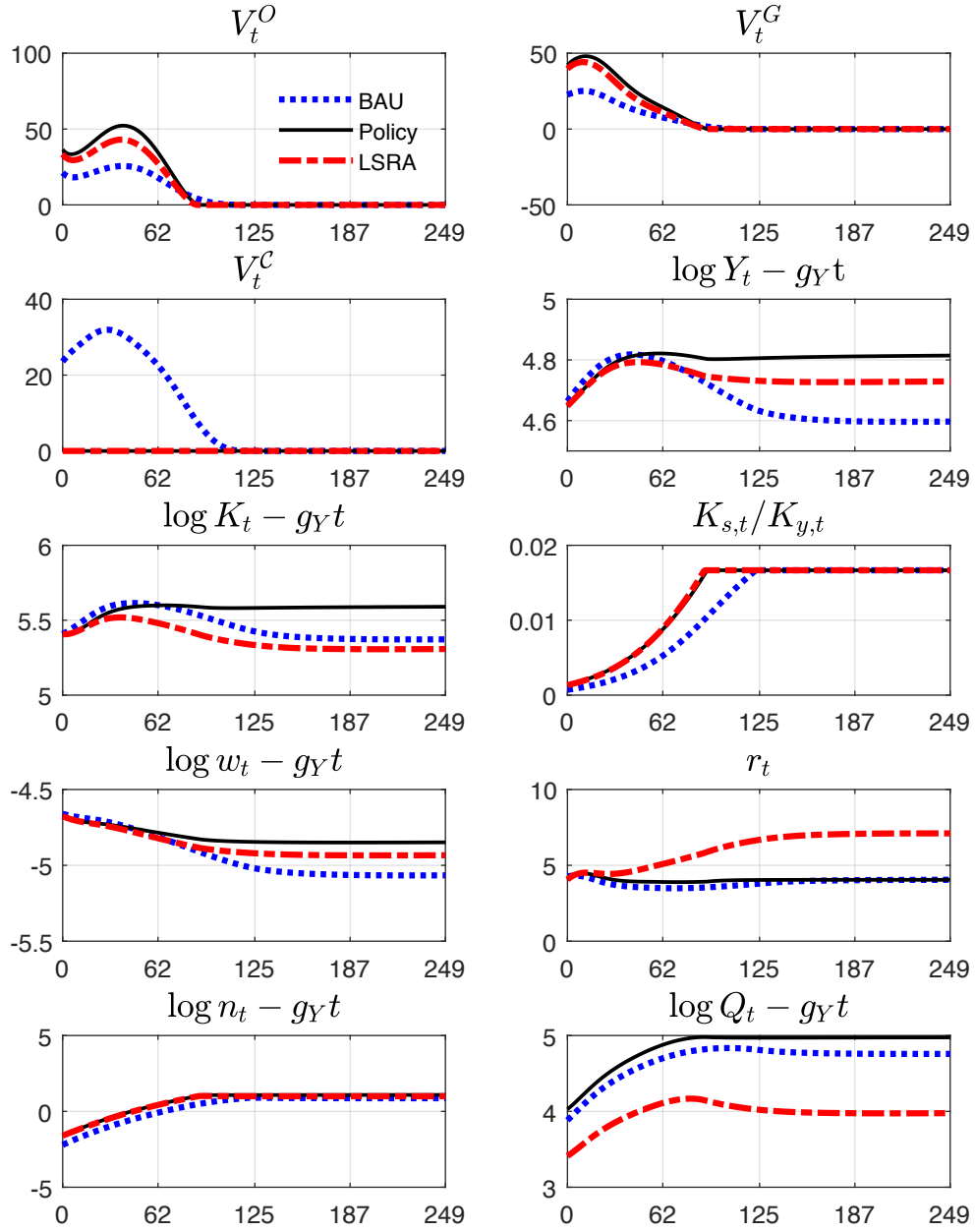
Figure 9: Optimal Tax Policy with 3x Damage Function, Part 3

* V_t^O - value of oil company, V_t^G - value of gas company, $V_t^{\mathcal{C}}$ - value of coal company, Y_t - final output, K_t - total capital, $K_{s,t}$ - clean energy sector's capital, $K_{y,t}$ - final goods production sector's capital, w_t - real wage rate, r_t - real interest rate, n_t - rental price of land, Q_t - value of land, g_Y - steady state growth rate of output.

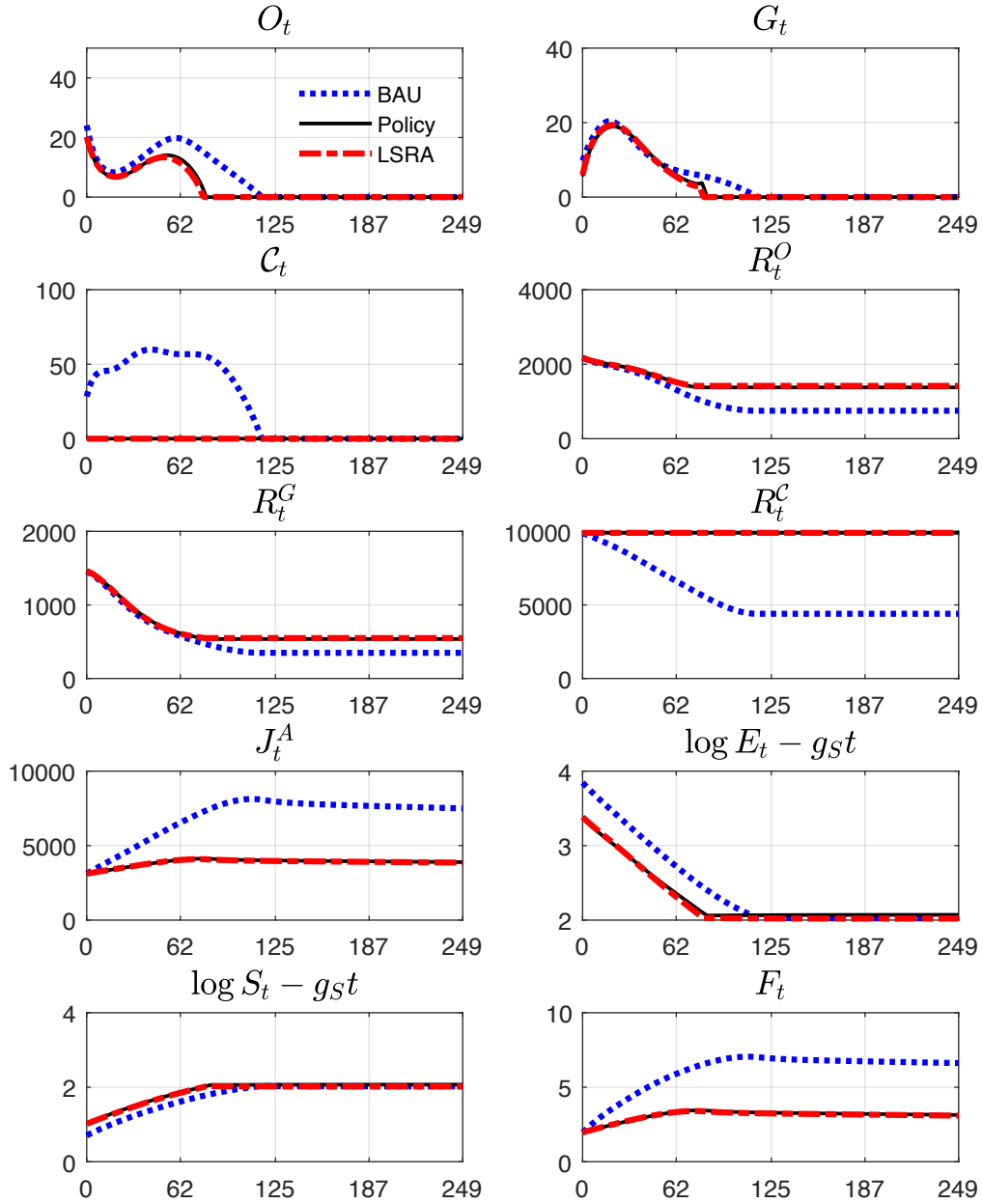# Appendix 3   Optimal Tax Policy with 6x Damage Function

Figure 10: Optimal Tax Policy with 6x Damage Function, Part 1*

*O_t - oil, G_t - gas, $\mathcal{C}_t$ -coal, R_t^O -oil reserves, R_t^G - gas reserves, $R_t^{\mathcal{C}}$ -coal reserves, J_t^A - CO_2 concentration in atmosphere (all 7 in GtCO_2), S_t - clean energy production,E_t - total energy consumption, F_t - radiative forcing, g_Y - steady state growth rate of output, g_S - steady state growth rate of clean energy.
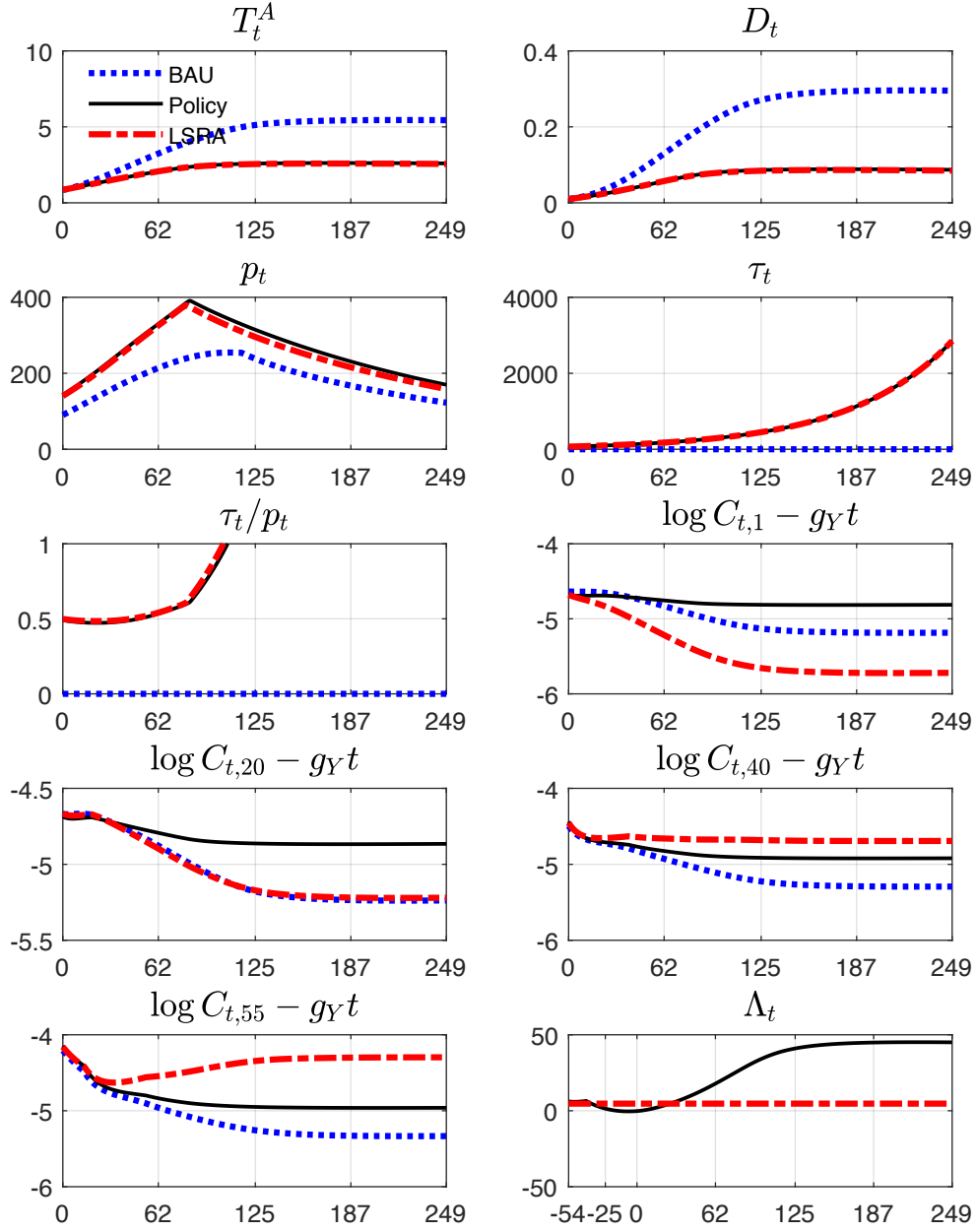
Figure 11: Optimal Tax Policy with 6x Damage Function, Part 2

*T_t^A - global mean surface temperature, D_t - damages, p_t - price of clean energy and oil per tCO_2, τ_t -absolute tax per tCO_2, $C_{t,j}$ -consumption of generation j at time t. g_Y - steady state growth rate of output, Λ_t - percentage change in consumption of households that produces the same change in the welfare of a generation born in period t compared to no policy scenario.
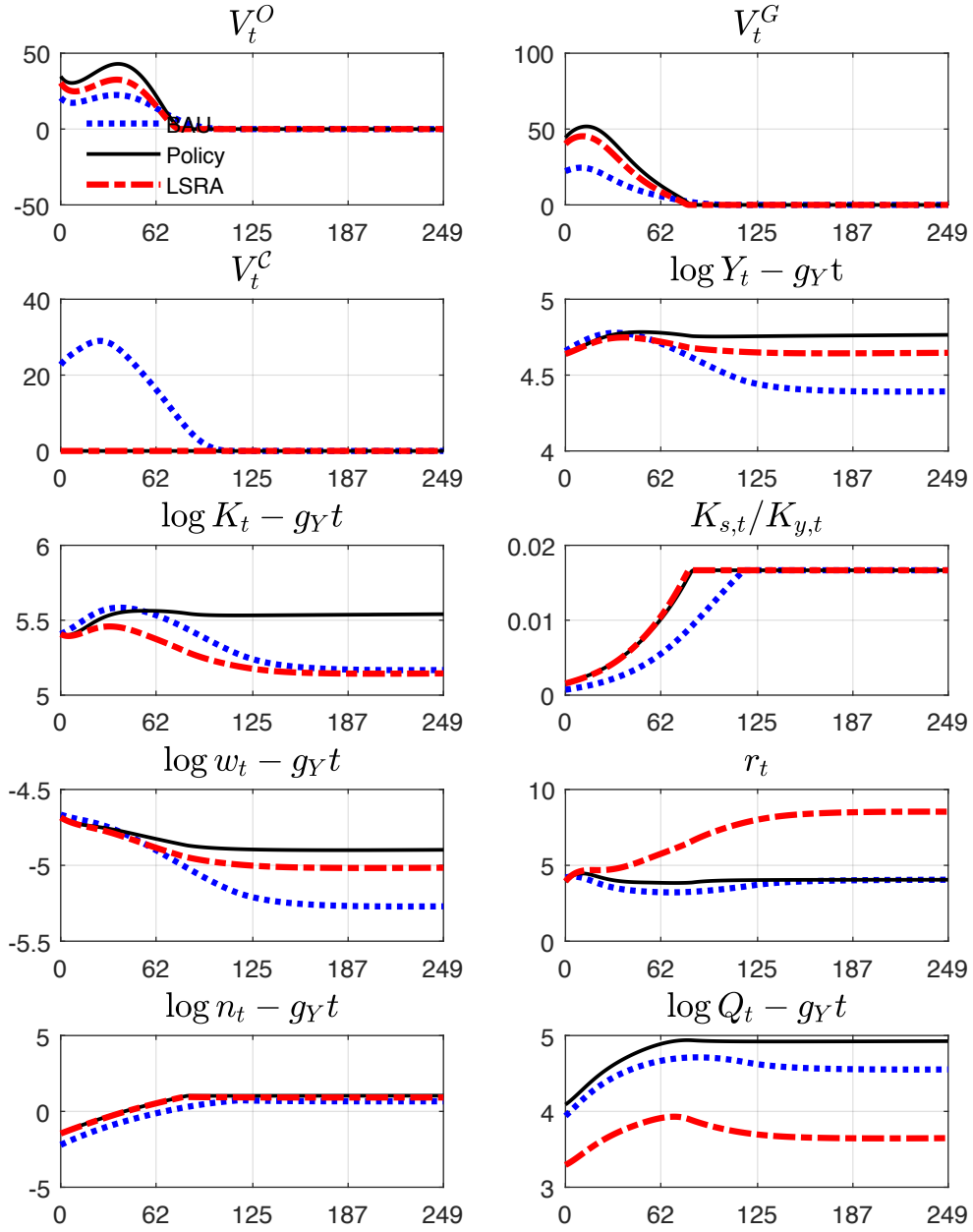
Figure 12: Optimal Tax Policy with 6x Damage Function, Part 3

* V_t^O - value of oil company, V_t^G - value of gas company, V_t^C - value of coal company, Y_t - final output, K_t - total capital, $K_{s,t}$ - clean energy sector's capital, $K_{y,t}$ - final goods production sector's capital, w_t - real wage rate, r_t - real interest rate, n_t - rental price of land, Q_t - value of land, g_Y - steady state growth rate of output.