BERLIN SCHOOL OF ECONOMICS
DISCUSSION PAPERS

Discussion Paper #47

August 2024

The Political Economics of Green Transitions: Optimal Intertemporal Policy Response

Lorenz Dögnitz

Théo Konc

Linus Mattauch

The Political Economics of Green Transitions: Optimal Intertemporal Policy Response*

Lorenz Dögnitz[†], Théo Konc[‡], Linus Mattauch[§]

August 29, 2024

Abstract

Besley and Persson (2023) pioneer a political economy model of a green transition with changing preferences. Here we solve for the optimal policy intervention and find that the optimal tax on the polluting good starts high and is subsequently declining, to support the transition in preferences. We quantify the welfare loss of ignoring preference changes.

JEL classification: D62, H23, Q54

Keywords: Endogenous preferences, green transition, carbon tax, political economics, intertemporal optimisation

*We thank Tim Besley for support and Till Armbruster, Inge van den Bijgaart, Samuel Bowles, David Klenert, Katrin Schmelz, Sugandha Srivastav and Linchuan Xu for helpful comments. Konc and Mattauch thank the Robert Bosch foundation for financial support. The authors declare that they have no conflict of interest and have acknowledge all relevant funding.

[†]Technical University of Berlin

[‡]Wageningen University, Technical University of Berlin, Potsdam Institute for Climate Impact Research – member of Leibniz association

[§]Technical University of Berlin, Potsdam Institute for Climate Impact Research – member of Leibniz association

1 Introduction

For the analysis of green transitions – reducing the use of goods with negative externality over time – preference change can play an important role, as is documented by recent theoretical and empirical work in environmental economics and policy (Weinberger and Goetzke, 2010; van den Bijgaart, 2018; Bezin, 2019; Konc et al., 2021; Severen and Van Benthem, 2022; Kreps, 2023). An optimal green transition depends on the effects of the policy on preferences (Mattauch et al., 2022).

Besley and Persson (2023) pioneer a political economy model for green transitions with changing preferences. There are two types of preferences and two types of products, "green" and "brown", that is low-carbon and high-carbon. A share μ of citizens holds green preferences which increase the utility derived from green products and decrease the utility derived from brown products. Brown products have an environmental externality which lowers the utility of all citizens.

Policy-makers can implement a tax t on green goods and a tax T on brown goods. When the relative price of the brown good increases, consumers react by reducing their preferences for this good to maximize their utility. Besley and Persson (2023) model, as a base case, probabilistic electoral competition between parties, and show that voters will elect a party which optimizes present (but not intertemporal) welfare because parties cannot bind future legislation to higher tax rates. They assert that higher tax rates taking into account the preference transition would be welfare-superior over multiple periods, without deriving the optimal policy trajectory. Focusing on political economy, Besley and Persson (2023) do not consider intertemporal welfare effects of acknowledging that changes in preferences matter for a green transition.

In this paper, we provide a dynamic solution to the model of Besley and Persson (2023) and derive optimal policy. We compare the politically feasible tax rates, which are constant across time ("static"), with a socially optimal tax policy path, elucidating how far the political equilibrium is from the social optimum. We show the optimal tax starts high and then declines. The welfare loss of ignoring changes in preferences is 4.6% in the normalized and discounted welfare sum over 50 years.

2 Defining intertemporally optimal policies with changing preferences

We summarise the key elements from Besley and Persson (2023) needed for the subsequent analysis. Then, we analytically characterise the intertemporal social optimum before explaining the numerical approach to intertemporal optimization and the parameter choices.

2.1 Model and dynamic solution

Model of Besley and Persson (2023) Citizens with green or brown preferences have the following utility functions:

$$U_g = \frac{1}{1-\sigma}\left[\int_0^{\gamma}(1+g)^{\sigma}y_g(i)^{1-\sigma}di + \int_{\gamma}^{1}(1-g)^{\sigma}Y_g(i)^{1-\sigma}di\right] + x_g - \lambda\bar{Y} \tag{1}$$

$$U_b = \frac{1}{1-\sigma}\left[\int_0^{\gamma}y_b(i)^{1-\sigma}di + \int_{\gamma}^{1}Y_b(i)^{1-\sigma}di\right] + x_b - \lambda\bar{Y} \tag{2}$$

with substitution elasticity σ, preference shift g, demand for green product $i \in [0,\gamma]$ by the citizen with green preferences $y_g(i)$, demand for brown product $i \in [\gamma, 1]$ by the green citizen $Y_g(i)$, demands by citizens with brown preferences $y_b(i)$ and $Y_b(i)$, demand for numeraire good x_g, x_b, and externality $\lambda\bar{Y}$ with the total brown production $\bar{Y}$.

The budget constraint of a citizen with green preferences (similar with brown) is given by

$$R \geq x_g + \int_0^{\gamma}p(i)y_g(i)di + \int_{\gamma}^{1}P(i)Y_g(i)di \tag{3}$$

with income R, price for green goods $p(i)$, and price for brown goods $P(i)$.

Firms produce either brown or green goods, and are monopolists for their variety i. They maximize profits given the marginal costs χ, the additional marginal costs for green production ζ, and the taxes on green and brown products t and T. Green technology also comes with fixed adoption costs $m \cdot i$ that depend on the variety i produced by the firm.

In this paper, we focus on the welfare effects of different tax policy paths. Welfare Ω is given by the following expression:[1]

$$\Omega(\mu,\gamma,T) = \gamma(1+\mu g)w(t) + (1-\gamma)(1-\mu g)W(T) + I - \frac{\gamma^2 m}{2} \tag{4}$$

with

$$w(t) = \frac{1}{1-\sigma}\kappa(\zeta+t) - (\chi+\zeta)\kappa(\zeta+t)^{\frac{1}{1-\sigma}}$$

and

$$W(T) = \frac{1}{1-\sigma}\kappa(T) - (\chi+\lambda)\kappa(T)^{\frac{1}{1-\sigma}}$$

and

$$\kappa(z) = \left(\frac{\chi+z}{1-\sigma}\right)^{1-\frac{1}{\sigma}},$$

see also Besley and Persson (2023). Here, μ denotes the share of consumers holding green preferences, and γ the share of green varieties. $w(t)$ and $W(T)$ represent the components of

[1] We show the derivation in Appendix A.1.

welfare related to green and brown products, respectively. Ω, μ, γ, T are used here as functions in continuous time, and in discrete time in the next subsection.

To speed up the green transition, the tax on brown goods T can be increased or the tax on green goods t decreased. We optimize over T and hold t constant. Therefore, t is omitted as argument in the optimization.

Besley and Persson (2023) show that welfare-optimal tax rates *for the current period* are independent of μ and γ, but do not characterize the full social optimum. They find that those tax rates will be implemented by political parties seeking to win the election. Both taxes come with a negative part that compensates for the monopolistic market structure. The brown tax has a positive part correcting for the externality.

$$T = (1-\sigma)\lambda - \sigma\chi \qquad t = -\sigma(\chi + \zeta) \tag{5}$$

Intertemporal social optimum: analytical solution To solve the model dynamically and find the welfare-maximizing policy, we need to define the transition functions for the two state variables μ and γ. We assume that citizens and firms choose whether to go green or brown by looking at the current values of μ, γ, and T. In other words, households adapt their preferences if they observe that such a change is already beneficial in the current period. Firms follow a similar dynamics, and adjust their future production based on the current period's profits. This is a deviation from Besley and Persson (2023) who assume agents are forward-looking for one period. Our assumption simplifies the numerical intertemporal optimization, while being realistic: with sufficiently many time steps, changes in market conditions between periods will be small.

The change in μ depends linearly on the utility advantage of holding green preferences $\Delta = U_g - U_b$ with preference imitation speed parameter d. For $\mu \to 0$ or $\mu \to 1$, the change in preferences goes to zero. This value transition function is a simplification of the value transition function in Besley and Persson (2023)[2].

$$\dot{\mu} \equiv h(\mu, \gamma, T) = \mu(1-\mu) \cdot d \cdot \Delta(\gamma, T) \tag{6}$$

$$\Delta(\gamma, T) = \frac{\sigma g}{1-\sigma}\left[\gamma\kappa(\zeta + t) - (1-\gamma)\kappa(T)\right]. \tag{7}$$

Note that $\frac{\partial \Delta(\gamma,T)}{\partial T} > 0$, which means that an increased pollution tax improves the fitness of green preferences.

The technology transition in Besley and Persson (2023) happens instantaneously. The share of green firms in the next period are the firms that will make a higher profit with green production (this is due to linearly increasing fixed green costs $m \cdot i$ for firm $i \in [0, \gamma]$). In contrast, we assume that the technology transition does not happen instantaneously. This is consistent with a large

[2]See their appendix, Equation 35.

body of intertemporal economics models on the energy transition that captures inertia in capital stocks, innovation and production (Acemoglu et al., 2012; Bauer et al., 2013; Mattauch et al., 2015; Keppo et al., 2021). In other words, the transition depends on the current share of green firms γ. The change in γ is determined by the difference between the share of green firms with instant transition from Besley and Persson (2023)[3] and the current share γ. The transition speed is regulated by e.

$$\dot{\gamma}(\mu, \gamma, T) \equiv f(\mu, \gamma, T) = e \cdot \left[\frac{\sigma}{m}\left[(1+\mu g)\kappa(\zeta + t) - (1-\mu g)\kappa(T)\right] - \gamma\right] \tag{8}$$

Similarly, a higher tax T increases the fitness of green production.

We define the Hamiltonian to find the optimal policy path given the discount factor β:

$$H = \Omega(\mu(s), \gamma(s), T(s)) + \psi(s)h(\mu(s), \gamma(s), T(s)) + \alpha(s)f(\mu(s), \gamma(s), T(s)) \tag{9}$$

with $\dot{\mu} = h(\mu(s), \gamma(s), T(s))$, and $\dot{\gamma} = f(\mu(s), \gamma(s), T(s))$,

with s being time. For simplicity, we only consider the tax on pollution. The necessary conditions for an optimum are:

$$\frac{\partial \Omega}{\partial T} + \psi\frac{\partial h}{\partial T} + \alpha\frac{\partial f}{\partial T} = 0 \tag{10}$$

$$\dot{\psi} = -\left(\frac{\partial \Omega}{\partial \mu} + \psi(\frac{\partial h}{\partial \mu} - \beta) + \alpha\frac{\partial f}{\partial \mu}\right) \tag{11}$$

$$\dot{\alpha} = -\left(\frac{\partial \Omega}{\partial \gamma} + \psi\frac{\partial h}{\partial \gamma} + \alpha(\frac{\partial f}{\partial \gamma} - \beta)\right) \tag{12}$$

Equation (10) provides an insight about the role of changing preferences in this economy. Holding technological development constant, it implies that the marginal effect of the tax on instant welfare ($\frac{\partial \Omega}{\partial T}$) should be equal to the negative product of the shadow value of holding green preferences (ψ) and the increased speed at which preferences change with the tax ($\frac{\partial h}{\partial T}$). When the value of holding green preferences is positive ($\psi > 0$), the tax should be set such that $\frac{\partial \Omega}{\partial T} < 0$. In other words, the optimal trajectory reflects that higher tax and lower instantaneous welfare are justified by the effect of the tax on future green preferences, leading to higher future welfare.

Proposition 1. *The optimal tax trajectory is defined by:*

$$\Leftrightarrow \dot{T} = -\frac{h\frac{\partial \eta}{\partial T} + f\frac{\partial \phi}{\partial T} - \eta\frac{\partial h}{\partial T} - \phi\frac{\partial f}{\partial T}}{\frac{\partial^2 \Omega}{\partial T^2} + \psi\frac{\partial^2 h}{\partial T^2} + \alpha\frac{\partial^2 f}{\partial T^2}}. \tag{13}$$

with $\eta = \left(\frac{\partial \Omega}{\partial \mu} + \psi(\frac{\partial h}{\partial \mu} - \beta) + \alpha\frac{\partial f}{\partial \mu}\right)$ *and* $\phi = \left(\frac{\partial \Omega}{\partial \gamma} + \psi\frac{\partial h}{\partial \gamma} + \alpha(\frac{\partial f}{\partial \gamma} - \beta)\right)$. η *represents the overall marginal welfare effect of moving to green preferences and* ϕ *the overall marginal welfare effect of moving to green production.*

[3] See their Equation 21

Proof. See Appendix A.2.

Corollary 1. *The optimal tax is decreasing if:*

$$h\frac{\partial \eta}{\partial T} + f\frac{\partial \phi}{\partial T} \leq \eta\frac{\partial h}{\partial T} + \phi\frac{\partial f}{\partial T} \tag{14}$$

With $h > 0$, $\eta > 0$, $f > 0$, and $\phi > 0$, a sufficient condition for Eq. 14 is:

$$\begin{aligned} &h\frac{\partial \eta}{\partial T} \leq \eta\frac{\partial h}{\partial T} \textit{ and } f\frac{\partial \phi}{\partial T} \leq \phi\frac{\partial f}{\partial T} \\ \Leftrightarrow &\frac{\partial log(\eta)}{\partial T} \leq \frac{\partial log(h)}{\partial T} \textit{ and } \frac{\partial log(\phi)}{\partial T} \leq \frac{\partial log(f)}{\partial T} \end{aligned} \tag{15}$$

Proof. See Appendix A.3.

Corollary 1 shows that the optimal tax is decreasing if the log-derivative of the preference transition is larger than the log-derivative of the welfare effects of the preference transition. The same must hold for the transition in technology. In other words, the optimal tax decreases when its overall welfare effects are smaller than its direct effect in accelerating the green transition. As we document in Section 3 initially in the transition this is not the case: there are early benefits to raising the tax slightly to catalyse the transition. Once a transition is under way, we confirm numerically that Corollary 1 holds.

2.2 Intertemporal optimization algorithm

For the numerical analysis, we optimize welfare in discrete time s over the time horizon S, given the discount factor β and the two transition functions $\gamma_{s+1}(\cdot)$, $\mu_{s+1}(\cdot)$ to derive an optimal taxation path.

$$\max_{\{T_s\}_{s=0}^{S}} \sum_{s=0}^{S} \beta^s \Omega(\mu_s, \gamma_s, T_s), \text{ s.t.} \tag{16}$$

$$\mu_{s+1}(\mu_s, \gamma_s, T_s) = \mu_s + \mu_s(1-\mu_s) \cdot d \cdot \Delta(\gamma_s, T_s), \tag{17}$$

$$\gamma_{s+1}(\mu_s, \gamma_s, T_s) = \begin{cases} 1, & if > 1 \\ \gamma_s + e \cdot \left[\frac{\sigma}{m}\left[(1+\mu_s g)\kappa(\zeta + t) - (1-\mu_s g)\kappa(T_s)\right] - \gamma_s\right], & if \in [0,1] \\ 0, & if < 0 \end{cases} \tag{18}$$

We use a backward induction algorithm to carry out the computation as the model is fully deterministic[4]. We start in the last period $s = S$ by maximizing the welfare given (μ_S, γ_S). We define $T_S^*(\mu_S, \gamma_S)$ as the optimal tax at (μ_S, γ_S). Similarly, $R_S^*(\mu_S, \gamma_S)$ is the maximum welfare at (μ_S, γ_S). The optimization is carried out for a grid of combinations (μ_S, γ_S) and the functions

[4]See Appendix B for the Python implementation.

R^*_S and T^*_S are interpolations between the grid points. Formally, we compute:

$$T^*_S(\mu_S, \gamma_S) = \text{argmax}\ \Omega(\mu_S, \gamma_S, T_S) \tag{19}$$

$$R^*_S(\mu_S, \gamma_S) = \Omega(\mu_S, \gamma_S, T^*_S(\mu_S, \gamma_S)). \tag{20}$$

For the preceding period s, function V, which is the current welfare and the discounted value of R^*_{s+1}, is maximised. The arguments $(\mu_{s+1}, \gamma_{s+1})$ of R^*_{s+1} are defined by the current (μ_s, γ_s) using the transition functions. The value of V given the optimal tax is stored in R^*_s, so R^* contains the discounted welfare sum of the current and all subsequent periods. This optimization is computed for all periods until $s = 0$.

$$\begin{aligned} T^*_s(\mu_s, \gamma_s) &= \text{argmax}\ V(\mu_s, \gamma_s, T_s) \\ R^*_s(\mu_s, \gamma_s) &= V(\mu_s, \gamma_s, T^*_s(\mu_s, \gamma_s)), \\ \text{with } V(\mu_s, \gamma_s, T_s) &= \Omega(\mu_s, \gamma_s, T_s) + \beta \cdot R^*_{s+1}(\mu_{s+1}(\mu_s, \gamma_s, T_s), \gamma_{s+1}(\mu_s, \gamma_s, T_s))\ \forall s \in [0, S[. \end{aligned}$$

After maximizing backwards, the optimal tax policy path given an initial (μ_0, γ_0) can be tracked forwards. It starts with $T^*_0(\mu_0, \gamma_0)$. Then, μ_{s+1} and γ_{s+1} are calculated using the transition functions and T^*_0. This is repeated for every period until $s = S$.

2.3 Simulation parameters

The choice of parameters for the simulation and the sensitivity analysis is given in Table 1. As Besley and Persson (2023) do not provide a numerical implementation of their model, we select values for the parameters, justified as follows.

Utility function The substitution elasticity is $\sigma \in [0, 1[$. For $\sigma \to 0$, every variety goes into the utility function linearly and the preference shift has no effect. For $\sigma \to 1$, the substitutability decreases and the preference shift gets stronger.

The preference shift of green consumers is denoted by g. With the chosen σ and g, green consumers get +22% utility from green products and -29% utility from brown products. The parameter d regulates the preference imitation speed, that is how fast people imitate rewarding preferences. We choose its value so that a full transition in preferences can happen during the simulation time of 60 years.

Production function The marginal cost of production is χ. It is chosen so that the market for green and brown products makes up around 5% of the economy given an endowment of $I = 1$. The additional marginal cost of green production ζ is chosen in relation to χ so that there is a green premium of $+33\%$. e is the technology imitation speed, which regulates the pace at which firms can adopt green technologies.

Symbol	Value (sensitivity analysis)	Name
σ	0.5 (0.49, 0.51)	substitution elasticity
g	0.5 (0.45, 0.55)	preference shift of green consumers
χ	3.0 (2.5, 3.5)	marginal cost of production
ζ	1.0 (0.75, 1.25)	additional marginal cost of green production
m	0.3 (0.25, 0.35)	fixed cost of green production for the most expensive firm
λ	8.0 (6.0, 10.0)	marginal damage of brown production (externality)
β	0.99 (0.97, 1)	discount factor
d	1.5 (0.0, 0.5, 1.0, 2.0)	preference imitation speed
e	0.25 (0.1, 0.4)	technology imitation speed
μ_0	0.25 (0.15, 0.35)	initial share of green citizens
γ_0	0.25 (0.15, 0.35)	initial share of green firms
I	1.0	endowment

Table 1: Choice of parameters

The fixed cost of green production for the most expensive firm is m. This is chosen so that the share of green firms γ stays in $]0, 1[$ for $\mu \in [0, 1]$, that is, both types of products remain on the market.

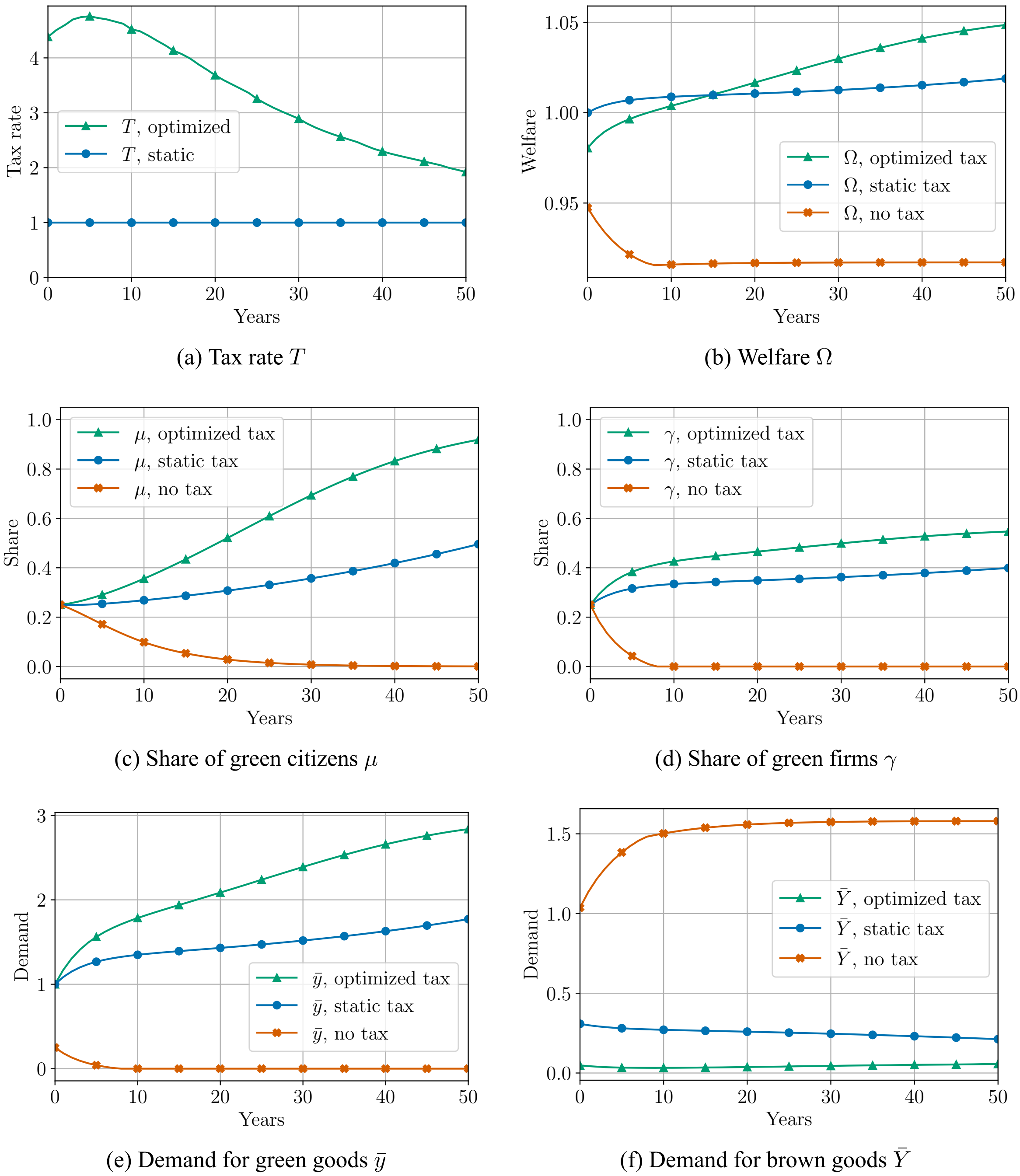
Finally, λ denotes the marginal damage of brown production. We choose $\lambda = 8$ so that the static brown tax (5) is positive, i.e. the distortion created by the environmental externality is larger than the one from the imperfect competition. We set μ_0 and γ_0 such that initially 25% of consumers hold green preferences and 25% of firms provide green goods.

3 Results

The main conceptual finding of our analysis is that, when preferences are endogenous, the optimal trajectory of a tax is non-monotonic. It starts high, somewhat increases and then declines. The intuition for this result is that a high tax level early accelerates the transition towards green preferences, and reduces the future welfare costs of taxing polluting goods. Failing to account for endogenous preferences yields a 4.6% decrease in the discounted sum of welfare.

Figure 1 (a), (b) shows the optimal trajectory of the pollution tax T and the social welfare Ω. Without a tax on pollution, the economy converges to the brown steady state quickly (Figure 1 (c), (d)): there are no green preferences nor green products ($\mu = 0,\ \gamma = 0$). Taxing the polluting good increases the share of green citizens and provides incentives for green firms to operate.

Initially, the optimal tax is about four times larger than the constant tax rate which is statically optimal ("politically feasible"). The optimal tax rate increases for a few periods, and later decreases. This trajectory can be explained by the effect of taxation on preferences. A higher initial tax is initially costly in terms of welfare , but also increases the share of citizens with green preferences in the next periods (Figure 1 (c)). This shift in preferences induced by the tax contributes to the reduction in pollution, and lowers the welfare costs of high pollution taxes in the future. As a result, the benefits of taxation are the highest during the first periods. Subse-



(a) Tax rate T

(b) Welfare Ω

(c) Share of green citizens μ

(d) Share of green firms γ

(e) Demand for green goods $\bar{y}$

(f) Demand for brown goods $\bar{Y}$

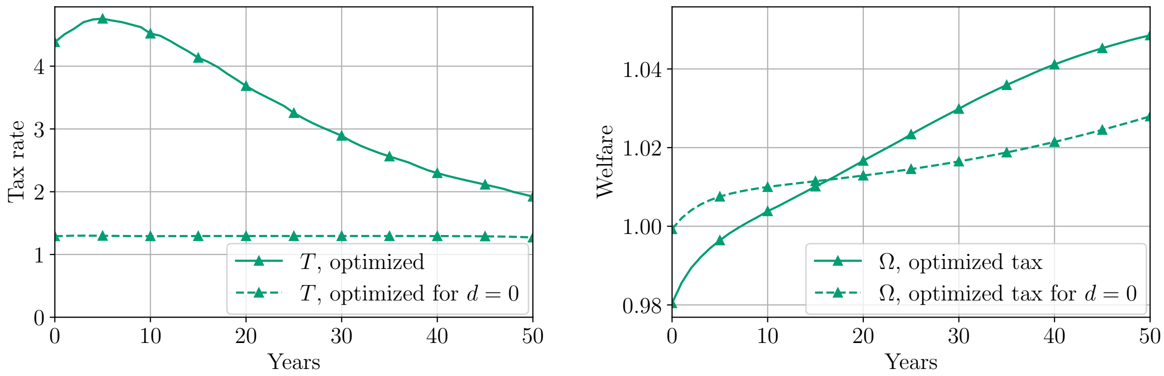
Figure 1: Simulation results

Note: Tax rate, welfare, and demand are normalized by their initial value with a static tax.

quently the tax decreases as the share of green citizens is high enough, so that demand for green goods increases even with a lower tax (Figure 1 (e)).

Finally, the transition on production is slow because of relatively high additional costs for green technology $\zeta = 1$, $m = 0.3$. Optimal taxation leads to a higher share of green firms throughout the simulation (Figure 1 (d)).

We have shown that the interplay between pollution taxes and preferences leads to a non-monotonic optimal tax trajectory. To evaluate the policy-relevance of this result, we now com-



(a) Tax rate T (b) Welfare Ω

Figure 2: Comparison to a tax optimized without considering preference change
Note: Tax rate and welfare are normalized by their initial value with a static tax.

pare the social welfare from this optimal policy with the social welfare of a policy that would ignore the change in preferences. Concretely, we compute the optimal tax setting $d = 0$ (implying fixed preferences), and calculate the welfare resulting from applying this tax schedule when preferences are in fact endogenous ($d = 1.5$). In other words, we calculate the welfare costs of ignoring that preferences are endogenous.

We find that implementing a tax schedule that does not account for changing preferences leads to a decrease of 4.6% in the discounted sum of welfare over 50 years. We compute this value by subtracting welfare obtained by an optimal tax erroneously ignoring evolving preferences from the true intertemporal optimum. Figure 2 depicts the trajectory of a tax mistakenly optimized under the false assumption of fixed preferences compared to the fully optimal tax. The welfare losses are high in the long-term, as the transition to green preferences is slowed down. To put the numerical result in context, the Figure shows that in the long-term the welfare gain is approximately two percentage points per year even if the economy is calibrated so that only around 5% of GDP are related to green or brown production.

We perform a comprehensive sensitive analysis, see Appendix C. The non-monotonicity of the tax trajectory is robust to a large variety of parametrizations. As expected, assuming an easier shift to green preferences leads to higher taxes initially, as they imply larger welfare gains later on. This is especially the case when the elasticity σ, the preference shift g, or the preference imitation d are higher.

4 Conclusion

Our results show that when a green transition is understood to be about changing preferences in addition to changing production, an optimal tax on the product with the externality starts high and peaks early before it decreases over time. The reason is that the tax incentivizes the preference change. Adapting the model of Besley and Persson (2023) to study intertemporal optimality, we show the optimal tax is non-monotonic and much higher than the politically feasible tax. Implementing a policy that does not take into account changes in preferences in a world where preferences are endogenous leads to sizeable welfare losses.

An important limitation of the model of Besley and Persson (2023) we extend here is the assumption about separating “green” and “brown” products with explicit preferences: Electricity is a homogeneous good with most consumers indifferent whether it is produced by low-carbon sources, rather key application include choosing between combustion-engine cars and low-carbon transport modes, or between high- and low-carbon food choices. Furthermore, while in practical fiscal policy, other factors such as short-term growth objectives, social justice and lobbyism would dominate setting carbon prices, the model pioneered by Besley and Persson (2023) is natural to exhibit the policy implications of endogenous preferences in a green transition clearly.

Two broader implications flow from our result. First, if the reason a society cannot implement an intertemporal optimum due to self-interest of political competitors, delegation to an independent authority with a technical mandate such as a “climate central bank” could help (Grosjean et al., 2016; Mattauch and Srivastav, 2023) – the UK Committee on Climate Change is a real-world example. Second, it is apparently easy for politicians to forget that as a result of political reform, the preferences of their electorate will look different. If changing values necessitate high and declining tax rates to steer the transition in preferences, as we formally show, this implies the new argument that politicians enacting environmental reform and will be punished less for it at the ballot box than they might think when assuming fixed preferences.

References

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

Bauer, N., Mouratiadou, I., Luderer, G., Baumstark, L., Brecha, R., Edenhofer, O., and Kriegler, E. (2013). Global fossil energy markets and climate change mitigation – an analysis with remind. *Climatic Change*, 136.

Besley, T. and Persson, T. (2023). The political economics of green transitions. *The Quarterly Journal of Economics*, 138:1863–1906.

Bezin, E. (2019). The economics of green consumption, cultural transmission and sustainable technological change. *Journal of economic theory*, 181:497–546.

Grosjean, G., Acworth, W., Flachsland, C., and Marschinski, R. (2016). After monetary policy, climate policy: Is delegation the key to EU ETS reform? *Climate Policy*, 16(1):1–25.

Keppo, I., Butnar, I., Bauer, N., Caspani, M., Edelenbosch, O., Emmerling, J., Fragkos, P., Guivarch, C., Harmsen, M., Lefèvre, J., Gallic, T. L., Leimbach, M., McDowall, W., Mercure, J.-F., Schaeffer, R., Trutnevyte, E., and Wagner, F. (2021). Exploring the possibility space: taking stock of the diverse capabilities and gaps in integrated assessment models. *Environmental Research Letters*, 16(5):053006.

Konc, T., Savin, I., and van den Bergh, J. C. (2021). The social multiplier of environmental policy: Application to carbon taxation. *Journal of Environmental Economics and Management*, 105:102396.

Kreps, D. (2023). *Arguing About Tastes: Modeling How Context and Experience Change Economic Preferences*. Columbia University Press.

Mattauch, L., Creutzig, F., and Edenhofer, O. (2015). Avoiding carbon lock-in: Policy options for advancing structural change. *Economic Modelling*, 50:49–63.

Mattauch, L., Hepburn, C., Spuler, F., and Stern, N. (2022). The economics of climate change with endogenous preferences. *Resource and Energy Economics*, 69:101312.

Mattauch, L. and Srivastav, S. (2023). Principles of decarbonization politics. *Nature Climate Change*, pages 503–504.

Severen, C. and Van Benthem, A. A. (2022). Formative experiences and the price of gasoline. *American Economic Journal: Applied Economics*, 14(2):256–284.

van den Bijgaart, I. (2018). Too slow a change? deep habits, consumption shifts and transitory tax policy.

Weinberger, R. and Goetzke, F. (2010). Unpacking preference: How previous experience affects auto ownership in the united states. *Urban studies*, 47(10):2111–2128.

Appendix for Online Publication only

A Derivations

A.1 Derivation of indirect welfare

A.1.1 Utility and demand

The utility of green citizens is

$$U_g = \frac{1}{1-\sigma}\left[\int_0^\gamma (1+g)^\sigma y_g(i)^{1-\sigma} di + \int_\gamma^1 (1-g)^\sigma Y_g(i)^{1-\sigma} di\right] + x_g - \lambda\bar{Y}. \tag{21}$$

With symmetric firms and varieties this becomes

$$U_g = \frac{1}{1-\sigma}\left[\gamma(1+g)^\sigma y_g(i)^{1-\sigma} + (1-\gamma)(1-g)^\sigma Y_g(i)^{1-\sigma}\right] + x_b - \lambda\bar{Y}. \tag{22}$$

The Lagrangian for utility maximization with budget restriction is

$$\begin{aligned} L_g =& \frac{1}{1-\sigma}\left[\gamma(1+g)^\sigma y_g(i)^{1-\sigma} + (1-\gamma)(1-g)^\sigma Y_g(i)^{1-\sigma}\right] + x_g - \lambda\bar{Y} \\ &+ \alpha\left[R - x - \gamma p(i) y_g(i) - (1-\gamma)P(i)Y_g(i)\right]. \end{aligned} \tag{23}$$

Deriving the Lagrangian yields the demand of a green citizen for a green product $y_g(i)$ and for a brown products $Y_g(i)$.

$$\begin{aligned} \frac{\partial L_g}{\partial x_g} &= 1 - \alpha \overset{!}{=} 0 \\ \Rightarrow \alpha &= 1 \end{aligned}$$

$$\begin{aligned} \frac{\partial L_g}{\partial y_g(i)} &= \gamma(1+g)^\sigma y_g(i)^{-\sigma} - \alpha\gamma p(i) \overset{!}{=} 0 \\ \Rightarrow y_g(i)^{-\sigma} &= (1+g)^{-\sigma} p(i) \\ \Rightarrow y_g(i) &= (1+g)p(i)^{-\frac{1}{\sigma}} \end{aligned} \tag{24}$$

$$\begin{aligned} \frac{\partial L_g}{\partial Y_g(i)} &= (1-\gamma)(1-g)^\sigma Y_g(i)^{-\sigma} - \alpha(1-\gamma)P(i) \overset{!}{=} 0 \\ \Rightarrow Y_g(i)^{-\sigma} &= (1-g)^{-\sigma} P(i) \\ \Rightarrow Y_g(i) &= (1-g)P(i)^{-\frac{1}{\sigma}} \end{aligned} \tag{25}$$

Utility of brown citizens is obtained from (22) with $g = 0$.

$$U_b = \frac{1}{1-\sigma}\left[\gamma y_b(i)^{1-\sigma} + (1-\gamma)Y_b(i)^{1-\sigma}\right] + x_b - \lambda\bar{Y} \tag{26}$$

Inserting $g = 0$ into (24) and (25) yields the brown citizens' demands for green $y_b(i)$ and brown $Y_b(i)$ varieties.

$$y_b(i) = p(i)^{-\frac{1}{\sigma}} \tag{27}$$
$$Y_b(i) = P(i)^{-\frac{1}{\sigma}} \tag{28}$$

Summing up the demand from both types of citizens gives the aggregated demands for a green and a brown good.

$$\begin{aligned} y(i) &= \mu y_g(i) + (1-\mu)y_b(i) \quad | \text{ with (24), (27)} \\ &= \mu(1+g)p(i)^{-\frac{1}{\sigma}} + (1-\mu)p(i)^{-\frac{1}{\sigma}} \\ &= (1+\mu g)p(i)^{-\frac{1}{\sigma}} \end{aligned} \tag{29}$$

$$\begin{aligned} Y(i) &= \mu Y_g(i) + (1-\mu)Y_b(i) \quad | \text{ with (25), (28)} \\ &= \mu(1-g)P(i)^{-\frac{1}{\sigma}} + (1-\mu)P(i)^{-\frac{1}{\sigma}} \\ &= (1-\mu g)P(i)^{-\frac{1}{\sigma}} \end{aligned} \tag{30}$$

A.1.2 Profits and prices

The profit of green firm i is given by

$$\begin{aligned} \pi(i) &= (p(i) - (\chi+\zeta+t))y(i) - mi \quad | \text{ with (29)} \\ &= (p(i) - (\chi+\zeta+t))(1+\mu g)p(i)^{-\frac{1}{\sigma}} - mi \\ &= (1+\mu g)\left[p(i)^{1-\frac{1}{\sigma}} - (\chi+\zeta+t)p(i)^{-\frac{1}{\sigma}}\right] - mi. \end{aligned} \tag{31}$$

The firms set the price as monopolists for their variety i.

$$\begin{aligned} \frac{\partial\pi(i)}{\partial p(i)} = (1+\mu g)\left[\left(1-\frac{1}{\sigma}\right)p(i)^{-\frac{1}{\sigma}} + \frac{1}{\sigma}(\chi+\zeta+t)p(i)^{-\frac{1}{\sigma}-1}\right] &\overset{!}{=} 0 \quad | \cdot \sigma p(i)^{1+\frac{1}{\sigma}} \\ \Rightarrow (\sigma-1)p(i) &= -(\chi+\zeta+t) \\ \Rightarrow p(i) &= \frac{\chi+\zeta+t}{1-\sigma} \end{aligned} \tag{32}$$

Inserting this into (31) yields the profit of a green firm.

$$\begin{aligned}
\pi(i) &= (1+\mu g)\left[p(i)^{1-\frac{1}{\sigma}} - (\chi+\zeta+t)p(i)^{-\frac{1}{\sigma}}\right] - mi \\
&= (1+\mu g)\left[\left(\frac{\chi+\zeta+t}{1-\sigma}\right)^{1-\frac{1}{\sigma}} - (\chi+\zeta+t)\left(\frac{\chi+\zeta+t}{1-\sigma}\right)^{-\frac{1}{\sigma}}\right] - mi \\
&= (1+\mu g)\left[\left(\frac{\chi+\zeta+t}{1-\sigma}\right)^{1-\frac{1}{\sigma}} - (1-\sigma)\left(\frac{\chi+\zeta+t}{1-\sigma}\right)^{1-\frac{1}{\sigma}}\right] - mi \\
&= \sigma(1+\mu g)\left(\frac{\chi+\zeta+t}{1-\sigma}\right)^{1-\frac{1}{\sigma}} - mi \\
&= \sigma(1+\mu g)\kappa(\zeta+t) - mi
\end{aligned} \tag{33}$$

with

$$\kappa(z) = \left(\frac{\chi+z}{1-\sigma}\right)^{1-\frac{1}{\sigma}}. \tag{34}$$

The profits of brown firms are given by

$$\begin{aligned}
\Pi(i) =& (P(i)-(\chi+T))Y(i) \quad | \text{ with (30)} \\
=& (P(i)-(\chi+T))(1-\mu g)P(i)^{-\frac{1}{\sigma}} \\
=& (1-\mu g)\left[P(i)^{1-\frac{1}{\sigma}} - (\chi+T)P(i)^{-\frac{1}{\sigma}}\right].
\end{aligned} \tag{35}$$

Profit maximization yields the price for brown varieties.

$$\begin{aligned}
\frac{\partial \Pi(i)}{\partial P(i)} = (1-\mu g)\left[\left(1-\frac{1}{\sigma}\right)P(i)^{-\frac{1}{\sigma}} + \frac{1}{\sigma}(\chi+T)P(i)^{-\frac{1}{\sigma}-1}\right] \stackrel{!}{=} 0 \quad | \cdot \sigma P(i)^{1+\frac{1}{\sigma}} \\
\Rightarrow (\sigma-1)P(i) = -(\chi+T) \\
\Rightarrow P(i) = \frac{\chi+T}{1-\sigma}
\end{aligned} \tag{36}$$

Inserting this into (35) yields the profit of a brown firm.

$$\begin{aligned}
\Pi(i) &= (1-\mu g)\left[P(i)^{1-\frac{1}{\sigma}} - (\chi+T)P(i)^{-\frac{1}{\sigma}}\right] \\
&= (1-\mu g)\left[\left(\frac{\chi+T}{1-\sigma}\right)^{1-\frac{1}{\sigma}} - (\chi+T)\left(\frac{\chi+T}{1-\sigma}\right)^{-\frac{1}{\sigma}}\right] \\
&= (1-\mu g)\left[\left(\frac{\chi+T}{1-\sigma}\right)^{1-\frac{1}{\sigma}} - (1-\sigma)\left(\frac{\chi+T}{1-\sigma}\right)^{1-\frac{1}{\sigma}}\right] \\
&= \sigma(1-\mu g)\left(\frac{\chi+T}{1-\sigma}\right)^{1-\frac{1}{\sigma}} \\
&= \sigma(1-\mu g)\kappa(T)
\end{aligned} \tag{37}$$

A.1.3 Externality

The negative external effect $\lambda\bar{Y}$ which is part of the utility function depends on the total production of brown goods $\bar{Y}$.

$$\begin{aligned}
\lambda\bar{Y} &= \lambda \int_{\gamma}^{1} Y(i)\, di \quad \text{| with (30)} \\
&= \lambda(1-\gamma)(1-\mu g)P(i)^{-\frac{1}{\sigma}} \quad \text{| with (36)} \\
&= \lambda(1-\gamma)(1-\mu g)\left(\frac{\chi+T}{1-\sigma}\right)^{-\frac{1}{\sigma}} \\
&= \lambda(1-\gamma)(1-\mu g)\kappa(T)^{\frac{1}{1-\sigma}} \qquad (38)
\end{aligned}$$

A.1.4 Tax revenues

$$\begin{aligned}
G &= t\int_{0}^{\gamma} y(i)di + T\int_{\gamma}^{1} Y(i)di \\
&= t\gamma y(i) + T(1-\gamma)Y(i) \quad \text{| with (29), (30)} \\
&= t\gamma(1+\mu g)p(i)^{-\frac{1}{\sigma}} + T(1-\gamma)(1-\mu g)P(i)^{-\frac{1}{\sigma}} \quad \text{| with (32), (36)} \\
&= t\gamma(1+\mu g)\left(\frac{\chi+\zeta+t}{1-\sigma}\right)^{-\frac{1}{\sigma}} + T(1-\gamma)(1-\mu g)\left(\frac{\chi+T}{1-\sigma}\right)^{-\frac{1}{\sigma}} \\
&= t\gamma(1+\mu g)\kappa(\zeta+t)^{\frac{1}{1-\sigma}} + T(1-\gamma)(1-\mu g)\kappa(T)^{\frac{1}{1-\sigma}}. \qquad (39)
\end{aligned}$$

A.1.5 Welfare

Reordering the budget constraint of green citizens with income R yields their demand for the numeraire good x_g.

$$\begin{aligned}
x_g &= R - \int_{0}^{\gamma} p(i)y_g(i)di - \int_{\gamma}^{1} P(i)Y_g(i)di \\
&= R - \gamma p(i)y_g(i) - (1-\gamma)P(i)Y_g(i) \quad \text{| with (24), (25)} \\
&= R - \gamma(1+g)p(i)^{1-\frac{1}{\sigma}} - (1-\gamma)(1-g)P(i)^{1-\frac{1}{\sigma}} \quad \text{| with (32), (36), (34)} \\
&= R - \gamma(1+g)\kappa(\zeta+t) - (1-\gamma)(1-g)\kappa(T). \qquad (40)
\end{aligned}$$

Inserting the demands for green and brown goods (24), (25), and for the numeraire (40) into the utility function of green citizens (22) yields

$$\begin{aligned}
U_g &= \frac{1}{1-\sigma}\left[\gamma(1+g)^{\sigma}y_g(i)^{1-\sigma} + (1-\gamma)(1-g)^{\sigma}Y_g(i)^{1-\sigma}\right] + x_g - \lambda\bar{Y} \\
&= \frac{1}{1-\sigma}\left[\gamma(1+g)\kappa(\zeta+t) + (1-\gamma)(1-g)\kappa(T)\right] + x_g - \lambda\bar{Y} \\
&= \frac{\sigma}{1-\sigma}\left[\gamma(1+g)\kappa(\zeta+t) + (1-\gamma)(1-g)\kappa(T)\right] + R - \lambda\bar{Y}. \qquad (41)
\end{aligned}$$

Similarly, the utility of brown citizens given by

$$U_b = \frac{\sigma}{1-\sigma}\left[\gamma\kappa(\zeta+t) + (1-\gamma)\kappa(T)\right] + R - \lambda\bar{Y}. \tag{42}$$

The income R consists of endowment I, tax revenue (39), and profits (33), (37).

$$\begin{aligned}
R &= I + G + \int_0^\gamma \pi(i)di + \int_\gamma^1 \Pi(i)di \\
&= I + t\gamma(1+\mu g)\kappa(\zeta+t)^{\frac{1}{1-\sigma}} + T(1-\gamma)(1-\mu g)\kappa(T)^{\frac{1}{1-\sigma}} \\
&\quad + \gamma\sigma(1+\mu g)\kappa(\zeta+t) - \frac{\gamma^2 m}{2} + (1-\gamma)\sigma(1-\mu g)\kappa(T)
\end{aligned} \tag{43}$$

The welfare function consists of the utilities of all citizens, using (41), (42), (43), (38).

$$\begin{aligned}
\Omega &= \mu U_g + (1-\mu)U_b \\
&= \mu\frac{\sigma}{1-\sigma}\left[\gamma(1+g)\kappa(\zeta+t) + (1-\gamma)(1-g)\kappa(T)\right] \\
&\quad + (1-\mu)\frac{\sigma}{1-\sigma}\left[\gamma\kappa(\zeta+t) + (1-\gamma)\kappa(T)\right] + R - \lambda\bar{Y} \\
&= \frac{\sigma}{1-\sigma}\left[\gamma(1+\mu g)\kappa(\zeta+t) + (1-\gamma)(1-\mu g)\kappa(T)\right] + R - \lambda\bar{Y} \\
&= \frac{\sigma}{1-\sigma}\left[\gamma(1+\mu g)\kappa(\zeta+t) + (1-\gamma)(1-\mu g)\kappa(T)\right] \\
&\quad + I + t\gamma(1+\mu g)\kappa(\zeta+t)^{\frac{1}{1-\sigma}} + T(1-\gamma)(1-\mu g)\kappa(T)^{\frac{1}{1-\sigma}} \\
&\quad + \gamma\sigma(1+\mu g)\kappa(\zeta+t) - \frac{\gamma^2 m}{2} + (1-\gamma)\sigma(1-\mu g)\kappa(T) \\
&\quad - \lambda(1-\gamma)(1-\mu g)\kappa(T)^{\frac{1}{1-\sigma}} \\
&= \gamma(1+\mu g)\left[\left(\frac{\sigma}{1-\sigma}+\sigma\right)\kappa(\zeta+t) + t\kappa(\zeta+t)^{\frac{1}{1-\sigma}}\right] \\
&\quad + (1-\gamma)(1-\mu g)\left[\left(\frac{\sigma}{1-\sigma}+\sigma\right)\kappa(T) + (T-\lambda)\kappa(T)^{\frac{1}{1-\sigma}}\right] + I - \frac{\gamma^2 m}{2}
\end{aligned} \tag{44}$$

The first part in square brackets is

$$\begin{aligned}
w(\zeta+t) &= \left(\frac{\sigma}{1-\sigma}+\sigma\right)\kappa(\zeta+t) + t\kappa(\zeta+t)^{\frac{1}{1-\sigma}} \\
&= \frac{(2-\sigma)\sigma}{1-\sigma}\kappa(\zeta+t) + \frac{1-\sigma}{1-\sigma}(\chi+\zeta+t)\kappa(\zeta+t)^{\frac{1}{1-\sigma}} - (\chi+\zeta)\kappa(\zeta+t)^{\frac{1}{1-\sigma}} \\
&= \frac{(2-\sigma)\sigma}{1-\sigma}\kappa(\zeta+t) + (1-\sigma)\kappa(\zeta+t) - (\chi+\zeta)\kappa(\zeta+t)^{\frac{1}{1-\sigma}} \\
&= \frac{1}{1-\sigma}\kappa(\zeta+t) - (\chi+\zeta)\kappa(\zeta+t)^{\frac{1}{1-\sigma}}.
\end{aligned} \tag{45}$$

The second part in square brackets is

$$\begin{aligned}
W(T) &= \left(\frac{\sigma}{1-\sigma}+\sigma\right)\kappa(T)+(T-\lambda)\kappa(T)^{\frac{1}{1-\sigma}} \\
&= \frac{(2-\sigma)\sigma}{1-\sigma}\kappa(T)+\frac{1-\sigma}{1-\sigma}(\chi+T)\kappa(T)^{\frac{1}{1-\sigma}}-(\chi+\lambda)\kappa(T)^{\frac{1}{1-\sigma}} \\
&= \frac{(2-\sigma)\sigma}{1-\sigma}\kappa(T)+(1-\sigma)\kappa(T)-(\chi+\lambda)\kappa(T)^{\frac{1}{1-\sigma}} \\
&= \frac{1}{1-\sigma}\kappa(T)-(\chi+\lambda)\kappa(T)^{\frac{1}{1-\sigma}}. \qquad (46)
\end{aligned}$$

A.1.6 Green utility advantage

The difference in utility of holding green instead of brown preferences is given by the utility advantage.

$$\begin{aligned}
\Delta &= U_g - U_b \quad | \text{ with (41), (42)} \\
&= \frac{\sigma g}{1-\sigma}\left[\gamma\kappa(\zeta+t)-(1-\gamma)\kappa(T)\right] \qquad (47)
\end{aligned}$$

A.1.7 Total market size

The total production is given by

$$\mu x_g + (1-\mu)x_b + \gamma y(i) + (1-\gamma)Y(i) \qquad (48)$$

with demand for green and brown products (29), (30) and demand for the numeraire (40) (with R given in (43)).

A.2 Proof of Proposition 1

With continuous time and infinite horizon, we solve the following current value Hamiltonian:

$$H = \Omega(\mu(s),\gamma(s),T(s)) + \psi(s)h(\mu(s),\gamma(s),T(s)) + \alpha(s)f(\mu(s),\gamma(s),T(s)) \qquad (49)$$

with $\dot{\mu} = h(\mu(s),\gamma(s),T(s))$, and $\dot{\gamma} = f(\mu(s),\gamma(s),T(s))$

$\dot{\mu}$ and $\dot{\gamma}$ are time derivatives.

Deriving $\dot{T}$ Necessary conditions:

$$\frac{\partial \Omega}{\partial T} + \psi(s)\frac{\partial h}{\partial T} + \alpha(s)\frac{\partial f}{\partial T} = 0 \tag{50}$$

$$\dot{\psi} = -\left(\frac{\partial \Omega}{\partial \mu} + \psi(s)(\frac{\partial h}{\partial \mu} - \beta) + \alpha(s)\frac{\partial f}{\partial \mu}\right) \tag{51}$$

$$\dot{\alpha} = -\left(\frac{\partial \Omega}{\partial \gamma} + \psi(s)\frac{\partial h}{\partial \gamma} + \alpha(s)(\frac{\partial f}{\partial \gamma} - \beta)\right) \tag{52}$$

The boundary conditions are

$$\lim_{S\to\infty} \psi(S) = 0 \tag{53}$$

$$\lim_{S\to\infty} \alpha(S) = 0 \tag{54}$$

Time-differentiating Eq. 50 yields:

$$\frac{d}{ds}\left(\frac{\partial \Omega}{\partial T}\right) + \frac{d}{ds}\left(\psi\frac{\partial h}{\partial T}\right) + \frac{d}{ds}\left(\alpha\frac{\partial f}{\partial T}\right) = 0 \tag{55}$$

$$\begin{aligned}
&\Leftrightarrow \frac{\partial^2 \Omega}{\partial \mu \partial T}h + \frac{\partial^2 \Omega}{\partial \gamma \partial T}f + \frac{\partial^2 \Omega}{\partial T^2}\dot{T} + \dot{\psi}\frac{\partial h}{\partial T} + \psi\left(\frac{\partial^2 h}{\partial \mu \partial T}h + \frac{\partial^2 h}{\partial \gamma \partial T}f + \frac{\partial^2 h}{\partial T^2}\dot{T}\right)\\
&+ \dot{\alpha}\frac{\partial f}{\partial T} + \alpha\left(\frac{\partial^2 f}{\partial \mu \partial T}h + \frac{\partial^2 f}{\partial \gamma \partial T}f + \frac{\partial^2 f}{\partial T^2}\dot{T}\right) = 0
\end{aligned} \tag{56}$$

Substituting $\dot{\alpha}$ and $\dot{\psi}$ by Eqs.52 and 51:

$$\begin{aligned}
\Leftrightarrow &\frac{\partial^2 \Omega}{\partial \mu \partial T}h + \frac{\partial^2 \Omega}{\partial \gamma \partial T}f + \frac{\partial^2 \Omega}{\partial T^2}\dot{T} - \left(\frac{\partial \Omega}{\partial \mu} + \psi(\frac{\partial h}{\partial \mu} - \beta) + \alpha\frac{\partial f}{\partial \mu}\right)\frac{\partial h}{\partial T}\\
&- \left(\frac{\partial \Omega}{\partial \gamma} + \psi\frac{\partial h}{\partial \gamma} + \alpha(\frac{\partial f}{\partial \gamma} - \beta)\right)\frac{\partial f}{\partial T} + \psi\left(\frac{\partial^2 h}{\partial \mu \partial T}h + \frac{\partial^2 h}{\partial \gamma \partial T}f + \frac{\partial^2 h}{\partial T^2}\dot{T}\right)\\
&+ \alpha\left(\frac{\partial^2 f}{\partial \mu \partial T}h + \frac{\partial^2 f}{\partial \gamma \partial T}f + \frac{\partial^2 f}{\partial T^2}\dot{T}\right) = 0
\end{aligned} \tag{57}$$

Finally, we solve for $\dot{T}$:

$$\begin{aligned}
&\left(\frac{\partial^2 \Omega}{\partial T^2} + \psi\frac{\partial^2 h}{\partial T^2} + \alpha\frac{\partial^2 f}{\partial T^2}\right)\dot{T} + \frac{\partial^2 \Omega}{\partial \mu \partial T}h + \frac{\partial^2 \Omega}{\partial \gamma \partial T}f\\
&- \left(\frac{\partial \Omega}{\partial \mu} + \psi(\frac{\partial h}{\partial \mu} - \beta) + \alpha\frac{\partial f}{\partial \mu}\right)\frac{\partial h}{\partial T}\\
&- \left(\frac{\partial \Omega}{\partial \gamma} + \psi\frac{\partial h}{\partial \gamma} + \alpha(\frac{\partial f}{\partial \gamma} - \beta)\right)\frac{\partial f}{\partial T}\\
&+ \psi\left(\frac{\partial^2 h}{\partial \mu \partial T}h + \frac{\partial^2 h}{\partial \gamma \partial T}f\right) + \alpha\left(\frac{\partial^2 f}{\partial \mu \partial T}h + \frac{\partial^2 f}{\partial \gamma \partial T}f\right) = 0
\end{aligned} \tag{58}$$

With $\eta = \left(\frac{\partial \Omega}{\partial \mu} + \psi(\frac{\partial h}{\partial \mu} - \beta) + \alpha\frac{\partial f}{\partial \mu}\right)$ and $\phi = \left(\frac{\partial \Omega}{\partial \gamma} + \psi\frac{\partial h}{\partial \gamma} + \alpha(\frac{\partial f}{\partial \gamma} - \beta)\right)$, we have:

$$\Leftrightarrow \dot{T} = -\frac{h\frac{\partial \eta}{\partial T} + f\frac{\partial \phi}{\partial T} - \eta\frac{\partial h}{\partial T} - \phi\frac{\partial f}{\partial T}}{\frac{\partial^2 \Omega}{\partial T^2} + \psi\frac{\partial^2 h}{\partial T^2} + \alpha\frac{\partial^2 f}{\partial T^2}}. \tag{59}$$

A.3 Proof of Corollary 1

To prove that $\dot{T} \leq 0 \Leftrightarrow h\frac{\partial \eta}{\partial T} + f\frac{\partial \phi}{\partial T} \leq \eta\frac{\partial h}{\partial T} + \phi\frac{\partial f}{\partial T}$, it is sufficient to show that that the denominator of Eq.59 is negative, i.e. $\frac{\partial^2 \Omega}{\partial T^2} + \psi\frac{\partial^2 h}{\partial T^2} + \alpha\frac{\partial^2 f}{\partial T^2} < 0$.

This is true by assumption as functions Ω, h, and f are concave in T.

B Optimization Code

```
import numpy as np
from scipy.optimize import minimize
from scipy.interpolate import interpn
directory_data = r'C:\Users\Lorenz\data'
```

Listing 1: File setup

```
σ = 0.5              # substitution elasticitiy
g = 0.5              # preference shift
χ = 3               # marginal brown cost
ζ = 1               # additional marginal green cost
m = 0.3              # fixed cost (m*i)
λ = 8               # marginal damage of the externality
β = 0.99             # discount factor
d = 1.5              # value transition speed
e = 0.25             # technolgy transition speed
μ0 = 0.25            # initial share of green citizens
γ0 = 0.25            # initial share of green firms
I = 1               # endowment
gridsize = 50        # grid size for the discretization of μ and γ
periods = 60         # optimization is performed for 60 periods, but 50 are shown
# Deviations for sensitivity analysis
σ_up = 0.51
σ_down = 0.49
g_up = 0.55
g_down = 0.45
χ_up = 3.5
χ_down = 2.5
ζ_up = 1.25
ζ_down = 0.75
m_up = 0.35
m_down = 0.25
λ_up = 10
λ_down = 6
β_up = 1
β_down = 0.97
d_up = 2
d_down = 1
e_up = 0.4
e_down = 0.2
μ0_up = 0.35
μ0_down = 0.15
γ0_up = 0.35
γ0_down = 0.15
```

Listing 2: Setting of parameters

```python
parameters_orig = np.array([σ,g,χ,ζ,m,λ,β,d,e,μ0,γ0])      # base case parameters not to be changed
parameters_unicode = ["σ","g","χ","ζ","m","λ","β","d","e","μ0","γ0"]
parameters_tex = ["\sigma","g","\chi","\zeta","m","\lambda",r"\beta","d","e","\mu_0","\gamma_0"]

periods_array = np.arange(periods)      # from 0 to periods-1
μ_array = np.linspace(0,1,gridsize)     # all μ values
γ_array = np.linspace(0,1,gridsize)     # all γ values


def initialization():                   # sets the parameter variables according to the array "parameters"

    global σ,g,χ,ζ,m,λ,β,d,e,μ0,γ0,filename,t,T_static

    σ = parameters[0]
    g = parameters[1]
    χ = parameters[2]
    ζ = parameters[3]
    m = parameters[4]
    λ = parameters[5]
    β = parameters[6]
    d = parameters[7]
    e = parameters[8]
    μ0 = parameters[9]
    γ0 = parameters[10]

    filename = f'σ{σ} g{g} χ{χ} ζ{ζ} m{m} λ{λ} β{β} d{d} e{e} μ{μ0} γ{γ0} I{I} gridsize{gridsize} periods{periods}'

    t = -σ*(χ+ζ)                # green tax is fixed to the static green tax
    T_static = (1-σ)*λ-σ*χ      # static brown tax
```

Listing 3: Initialization of initial parameters and function initialization() for changed parameter sets

```python
def Ω(T,μ,γ):                   # negative welfare function (to use minimization instead of maximization)
    return -(γ*(1+μ*g)*w() + (1-γ)*(1-μ*g)*W(T) + I - γ**2 * m/2)

def w():                        # part of the welfare function
    return κ(ζ+t)/(1-σ) - (χ+ζ) * κ(ζ+t)**(1/(1-σ))

def W(T):                       # part of the welfare function
    return κ(T)/(1-σ) -  (χ+λ) * κ(T)**(1/(1-σ))

def V(T,μ,γ,period,R):          # negative welfare of the current and all future periods
    return Ω(T,μ,γ) - β * interpn((μ_array,γ_array),R[period+1,:,:],(μ_(T,μ,γ),γ_(T,μ,γ)),method='linear')

def μ_(T,μ,γ):                  # transition function for share of green citizens
    return min(max(μ + μ*(1-μ)*d*Δ(T,γ), 0), 1)
```

```
def Δ(T,γ):                    # green utility advantage
    return σ*g/(1-σ) * (γ * κ(ζ+t) - (1-γ) * κ(T))

def γ_(T,μ,γ):                 # transition function for share of green firms
    return min(max(γ + e * ((σ/m)*((1+μ*g)*κ(ζ+t) - (1-μ*g)*κ(T)) - γ), 0), 1)

def κ(x):
    return ((χ+x)/(1-σ))**(1-(1/σ))

def y(μ,γ):                    # aggregated demand for green goods \bar{y}
    return γ * (1+μ*g) * ((χ+ζ+t)/(1-σ))**(-1/σ)

def Y(T,μ,γ):                  # aggregated demand for brown goods \bar{Y}
    return (1-γ) * (1-μ*g) * ((χ+T)/(1-σ))**(-1/σ)
```

Listing 4: Model functions

```
def optimization():

    R = np.zeros((periods,gridsize,gridsize))     # welfare of current and all future periods with optimal brown
    ↪ tax, given period, μ, and γ
    T = np.zeros((periods,gridsize,gridsize))     # optimal brown tax T, given period, μ, and γ

    period = periods-1                  # last period: optimize Ω instead of V
    print("Calculating period",period)
    for i in range(gridsize):           # every μ
        for j in range(gridsize):       # every γ
            res = minimize(Ω, x0=(T_static), args=(μ_array[i],γ_array[j]), bounds=[(0,20)], tol=1e-12) #tol=1e-10
            R[period,i,j] = -res.fun    # optimal welfare given μ, γ
            T[period,i,j] = res.x       # optimal T given μ, γ
    period -= 1

    while period >= 0:
        print("Calculating period",period)
        for i in range(gridsize):       # every μ
            for j in range(gridsize):   # every γ
                res = minimize(V, x0=(T[period+1,i,j]), args=(μ_array[i],γ_array[j],period,R), bounds=[(0,20)],
                ↪ tol=1e-12)
                R[period,i,j] = -res.fun    # optimal welfare of current and future periods given μ, γ
                T[period,i,j] = res.x       # optimal T given μ, γ
        period -= 1

    return T
```

Listing 5: Welfare optimization using backward induction

```
def tracking():

```

```
global t

""" Optimized tax """

track = np.zeros((periods,6))       # axis 1: [0]: μ, [1]: γ, [2]: T, [3]: Ω, [4]: y, [5]: Y
period = 0
μ = μ0
γ = γ0

while period < periods:
    track[period,0] = μ
    track[period,1] = γ
    track[period,2] = interpn((μ_array,γ_array),T[period,:,:],(μ,γ),method='linear')
    track[period,3] = -Ω(track[period,2],μ,γ)
    track[period,4] = y(μ,γ)
    track[period,5] = Y(track[period,2],μ,γ)
    μ = μ_(track[period,2],μ,γ)
    γ = γ_(track[period,2],μ,γ)
    period += 1

np.save(f'{directory_data}\{filename} Track',track)


""" Static tax """

track = np.zeros((periods,6))
period = 0
μ = μ0
γ = γ0

while period < periods:
    track[period,0] = μ
    track[period,1] = γ
    track[period,2] = T_static
    track[period,3] = -Ω(track[period,2],μ,γ)
    track[period,4] = y(μ,γ)
    track[period,5] = Y(track[period,2],μ,γ)
    μ = μ_(track[period,2],μ,γ)
    γ = γ_(track[period,2],μ,γ)
    period += 1

np.save(f'{directory_data}\{filename} Track_static',track)


""" No tax """

t_temp = t
t = 0
```

```
    track = np.zeros((periods,6))
    period = 0
    μ = μ0
    γ = γ0

    while period < periods:
        track[period,0] = μ
        track[period,1] = γ
        track[period,2] = 0
        track[period,3] = -Ω(0,μ,γ)
        track[period,4] = y(μ,γ)
        track[period,5] = Y(track[period,2],μ,γ)
        μ = μ_(0,μ,γ)
        γ = γ_(0,μ,γ)
        period += 1

    np.save(f'{directory_data}\{filename} Track_zero',track)

    t = t_temp
```

Listing 6: Forward tracking of paths for optimized, static and no tax

```
# Optimization for the base case
print("Optimization for the base case.")
parameters = parameters_orig.copy()
initialization()
filename_orig = filename
T = optimization()
tracking()

# Optimization for sensitivity analysis
for i in range(11):
    for j in ["up","down"]:
        print(f"Optimization for {j}-deviation in {parameters_unicode[i]}.")
        parameters[i] = globals()[parameters_unicode[i]+"_"+j]          # before it is parameters = parameters_orig, change of one parameter, e.g. to σ_up
        initialization()                            # sets the global parameter variables according to the array parameters
        T = optimization()
        tracking()
        parameters = parameters_orig.copy()
```

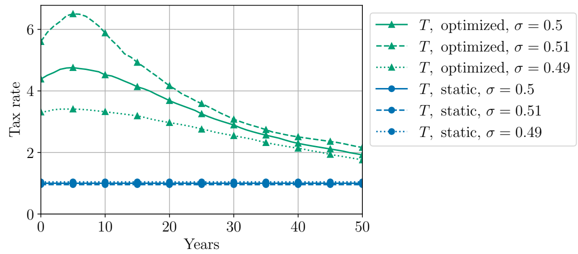
Listing 7: Call of initialization, optimization, and tracking functions for the base case and the deviations for sensitivity analysis

C Sensitivity analysis

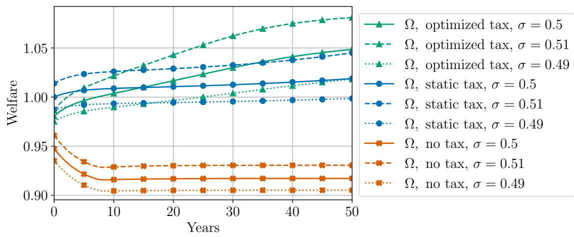
The following figures show deviations of the model parameters.

For some parameters, an *increase* leads to a higher peak in the optimal tax rate, a faster green transition, and higher levels of welfare. These are substitution elasticity σ, preference shift g, externality λ, discount factor β, value transition speed d, and technology transition speed e.

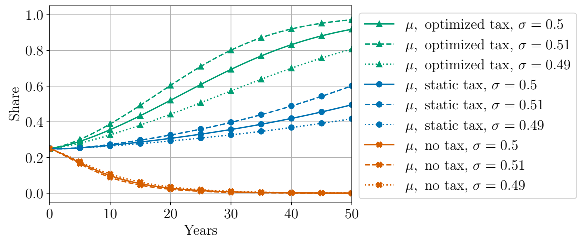
For the other parameters, the opposite is the case: a *decrease* leads to a higher peak in the optimal tax rate, a faster green transition, and higher levels of welfare. They are marginal costs χ, green marginal costs ζ, and green fixed costs m.



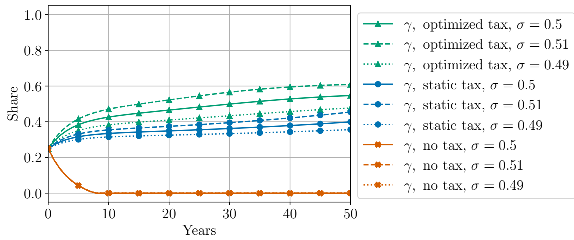
(a) Tax rate T



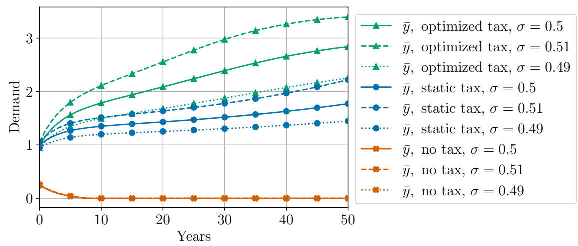
(b) Welfare Ω



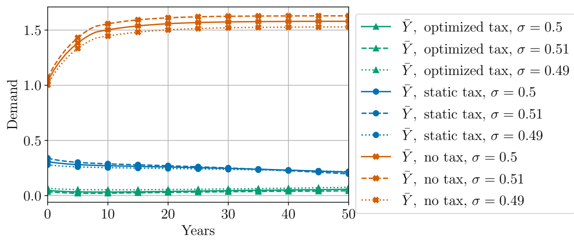
(c) Share of green citizens μ



(d) Share of green firms γ



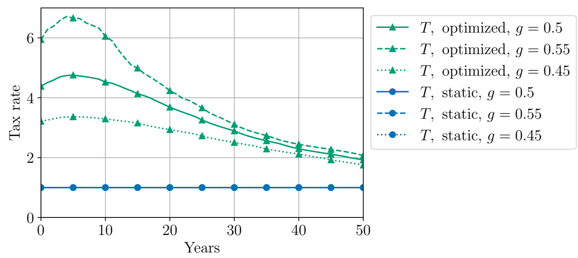
(e) Demand for green goods $\bar{y}$



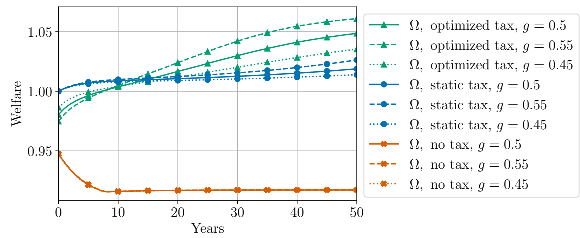
(f) Demand for green goods $\bar{y}$

Figure 3: Change in substitution elasticity σ

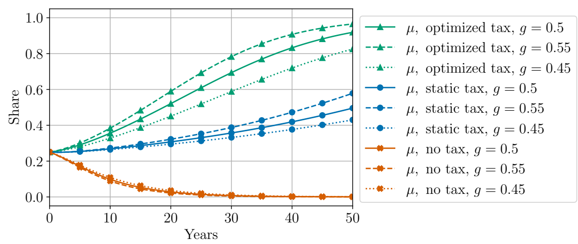
Note: Tax rate, welfare, and demand are normalized by their initial value with a static tax and the basecase parameters.



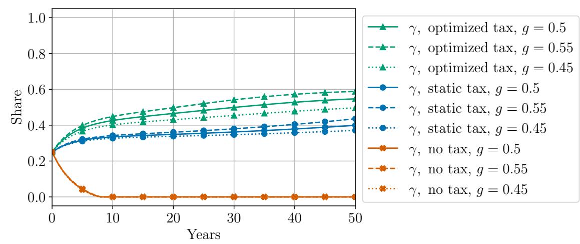
(a) Tax rate T



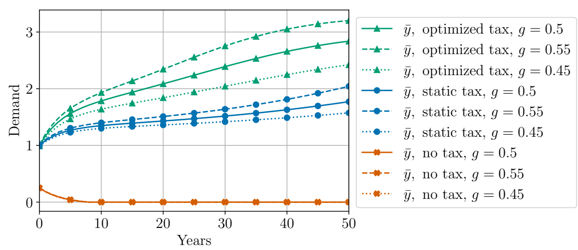
(b) Welfare Ω



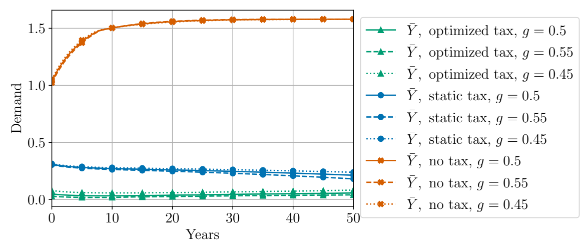
(c) Share of green citizens μ



(d) Share of green firms γ



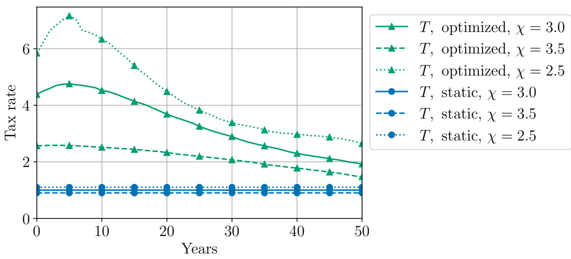
(e) Demand for green goods $\bar{y}$



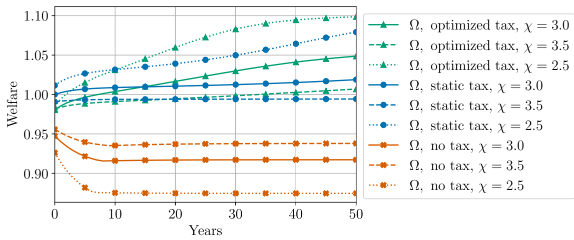
(f) Demand for green goods $\bar{y}$

Figure 4: Change in preference shift g

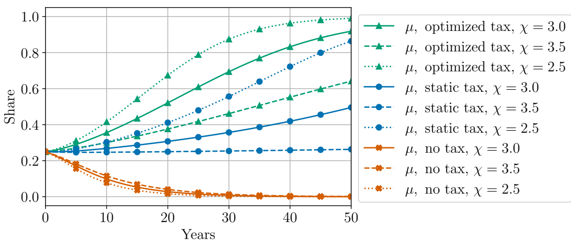
Note: Tax rate, welfare, and demand are normalized by their initial value with a static tax and the basecase parameters.



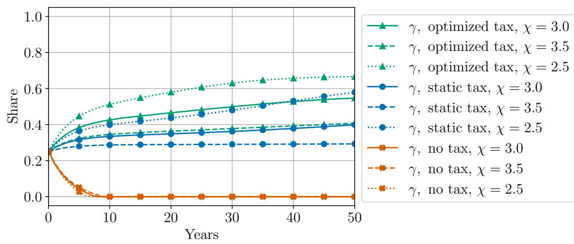
(a) Tax rate T



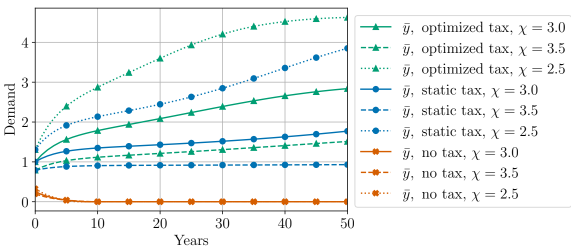
(b) Welfare Ω



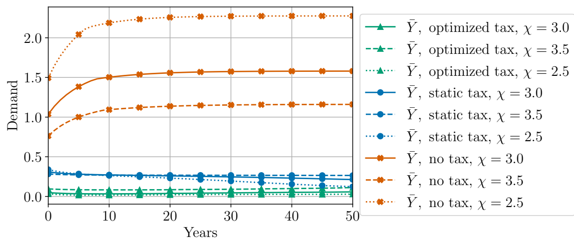
(c) Share of green citizens μ



(d) Share of green firms γ



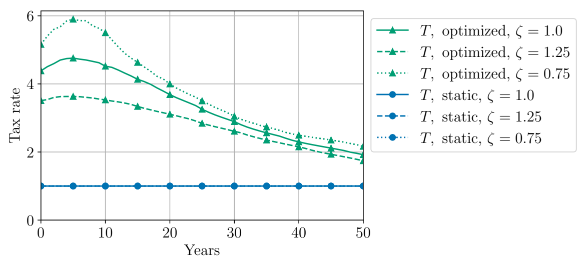
(e) Demand for green goods $\bar{y}$



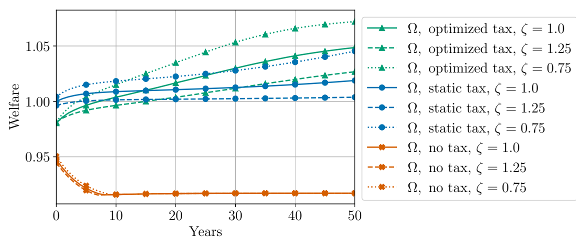
(f) Demand for green goods $\bar{y}$

Figure 5: Change in marginal costs χ

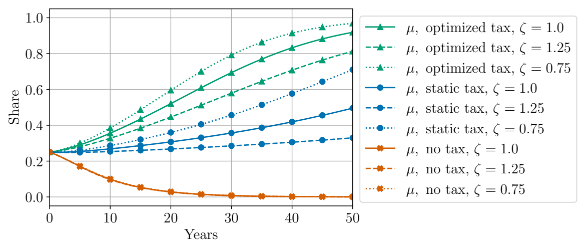
Note: Tax rate, welfare, and demand are normalized by their initial value with a static tax and the basecase parameters.



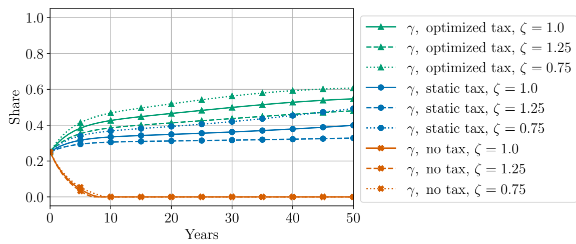
(a) Tax rate T



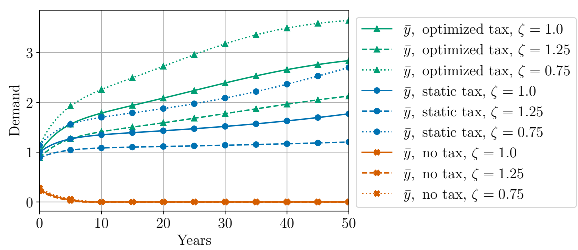
(b) Welfare Ω



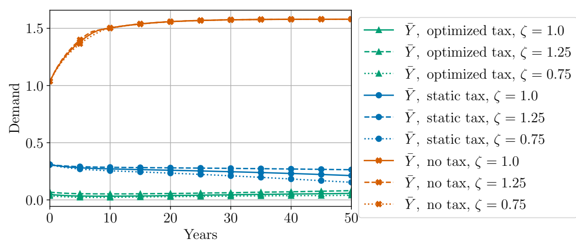
(c) Share of green citizens μ



(d) Share of green firms γ



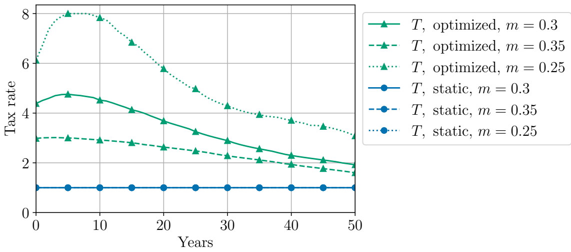
(e) Demand for green goods $\bar{y}$



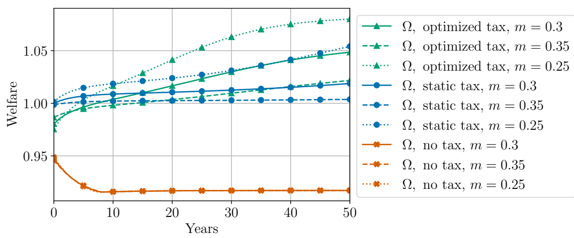
(f) Demand for green goods $\bar{y}$

Figure 6: Change in green marginal costs ζ

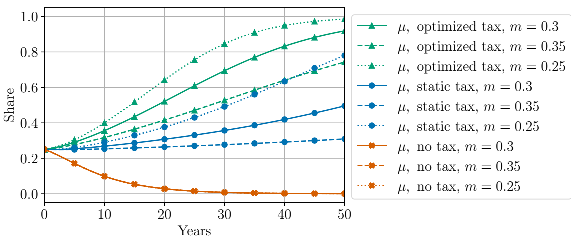
Note: Tax rate, welfare, and demand are normalized by their initial value with a static tax and the basecase parameters.



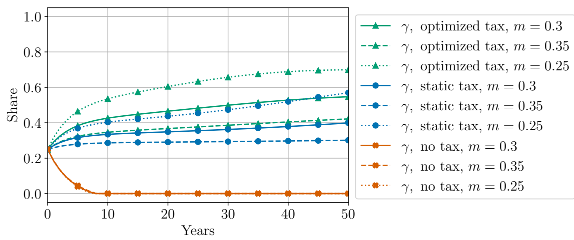
(a) Tax rate T



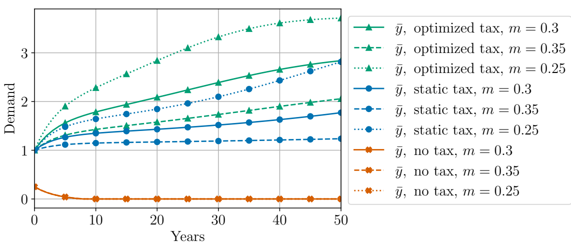
(b) Welfare Ω



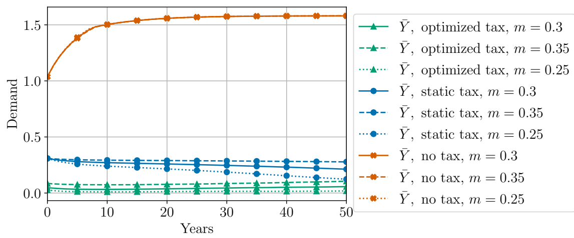
(c) Share of green citizens μ



(d) Share of green firms γ



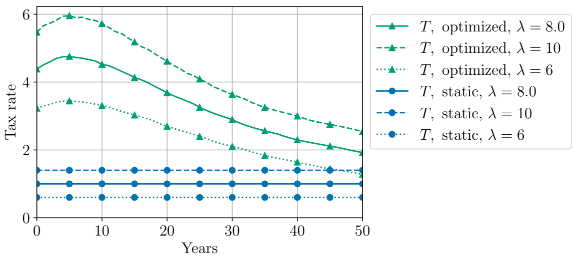
(e) Demand for green goods $\bar{y}$



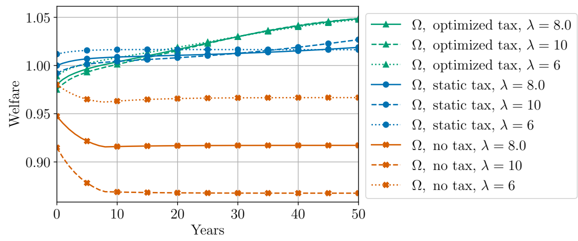
(f) Demand for green goods $\bar{y}$

Figure 7: Change in green fixed costs m

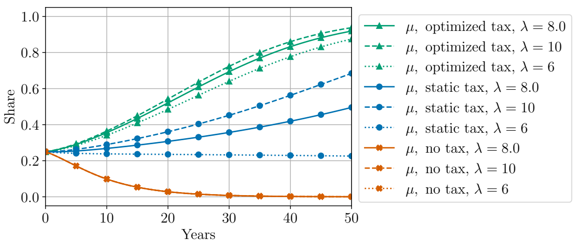
Note: Tax rate, welfare, and demand are normalized by their initial value with a static tax and the basecase parameters.



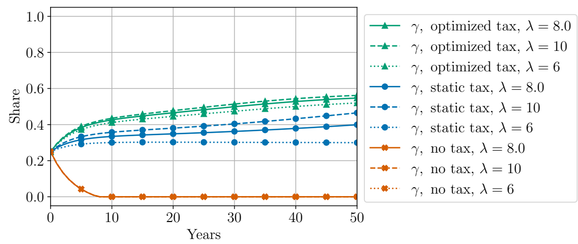
(a) Tax rate T



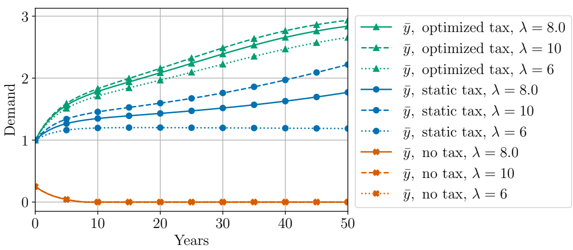
(b) Welfare Ω



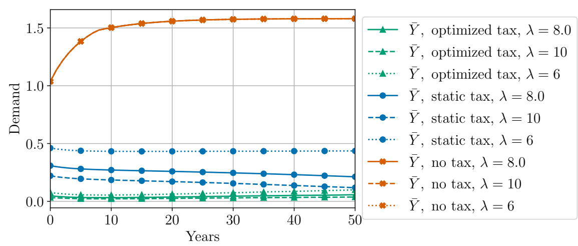
(c) Share of green citizens μ



(d) Share of green firms γ



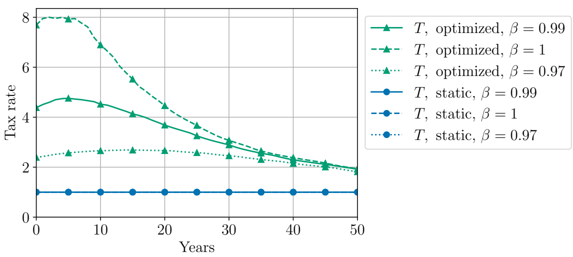
(e) Demand for green goods $\bar{y}$



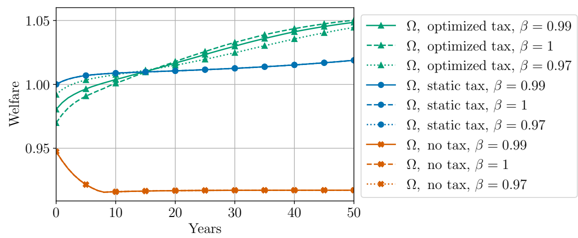
(f) Demand for green goods $\bar{y}$

Figure 8: Change in externality λ

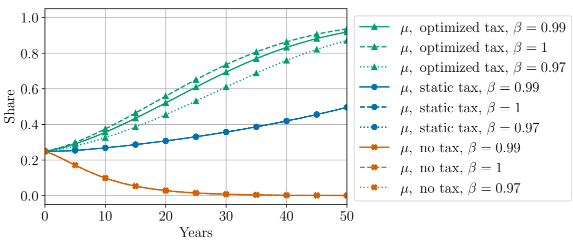
Note: Tax rate, welfare, and demand are normalized by their initial value with a static tax and the basecase parameters.



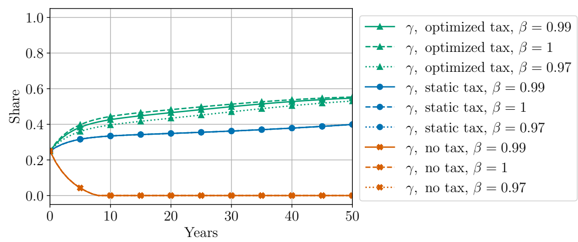
(a) Tax rate T



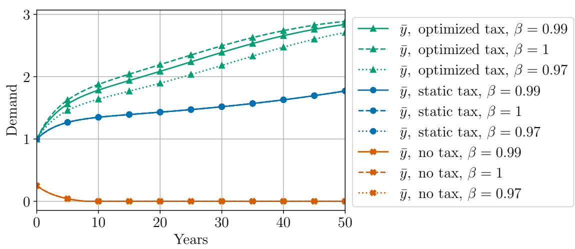
(b) Welfare Ω



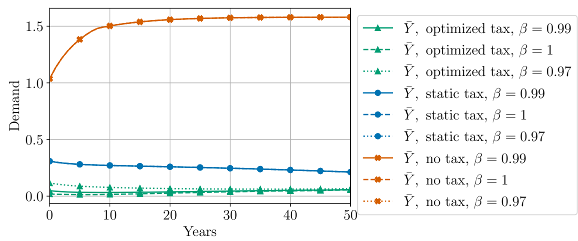
(c) Share of green citizens μ



(d) Share of green firms γ



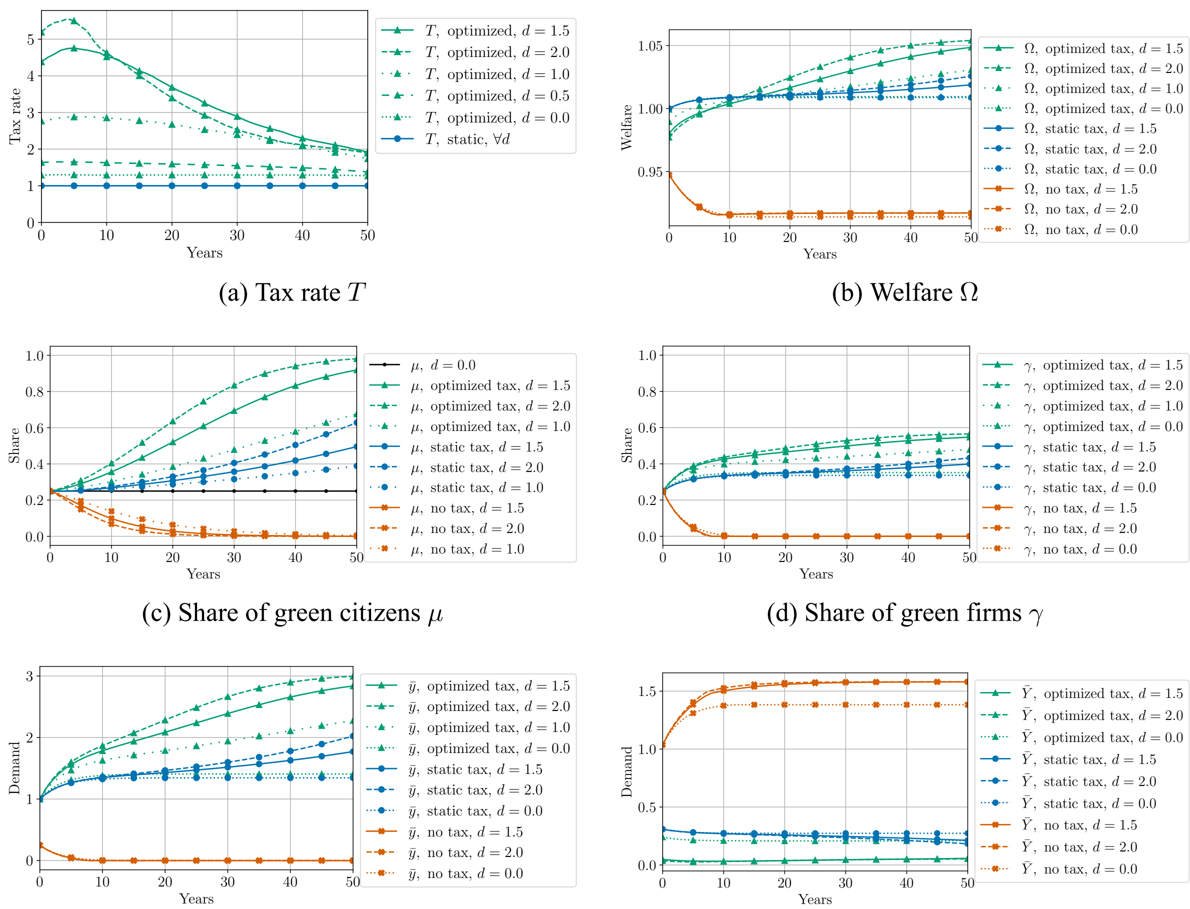
(e) Demand for green goods $\bar{y}$



(f) Demand for green goods $\bar{y}$

Figure 9: Change in discount rate β

Note: Tax rate, welfare, and demand are normalized by their initial value with a static tax and the basecase parameters.



(a) Tax rate T

(b) Welfare Ω

(c) Share of green citizens μ

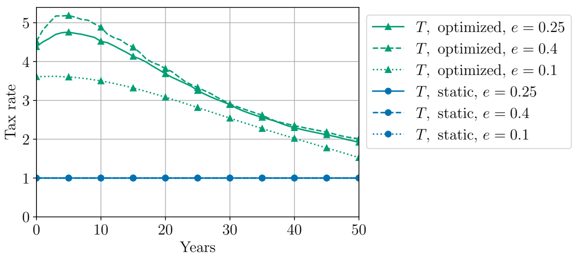
(d) Share of green firms γ

(e) Demand for green goods $\bar{y}$

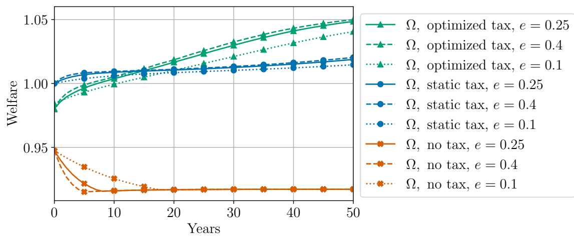
(f) Demand for green goods $\bar{y}$

Figure 10: Change in value transition speed d

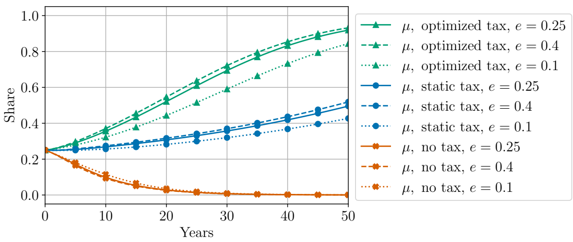
Note: Tax rate, welfare, and demand are normalized by their initial value with a static tax and the basecase parameters.



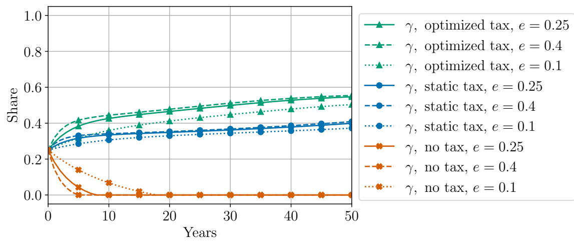
(a) Tax rate T



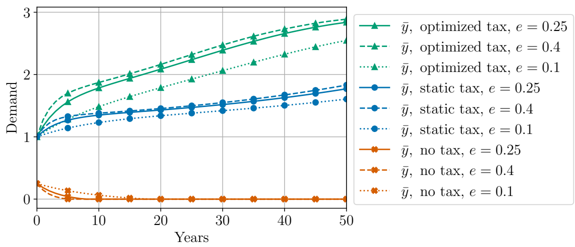
(b) Welfare Ω



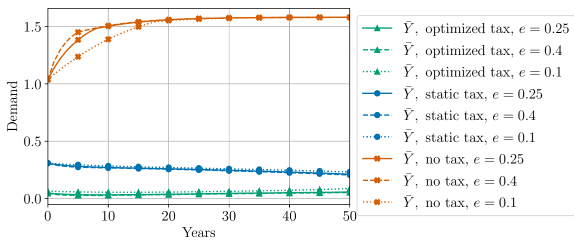
(c) Share of green citizens μ



(d) Share of green firms γ



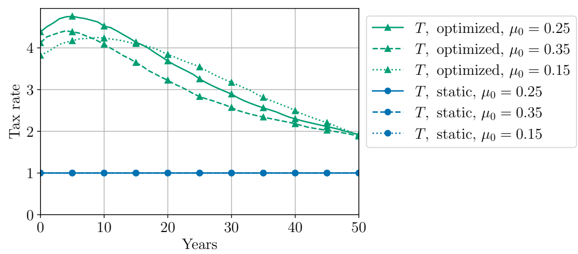
(e) Demand for green goods $\bar{y}$



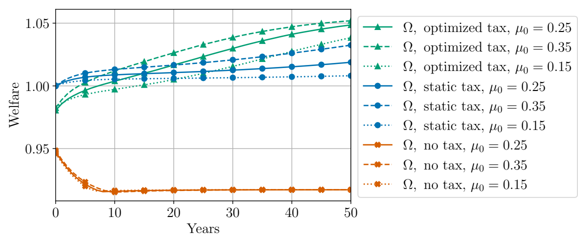
(f) Demand for green goods $\bar{y}$

Figure 11: Change in technology transition speed e

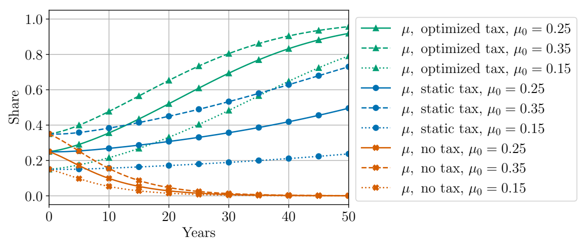
Note: Tax rate, welfare, and demand are normalized by their initial value with a static tax and the basecase parameters.



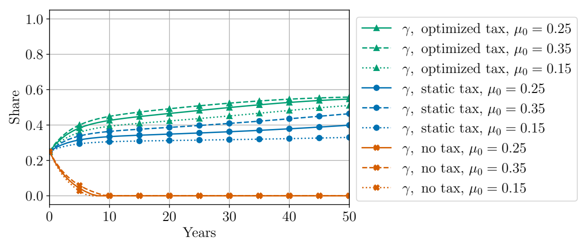
(a) Tax rate T



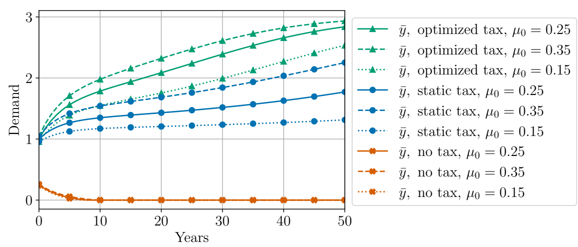
(b) Welfare Ω



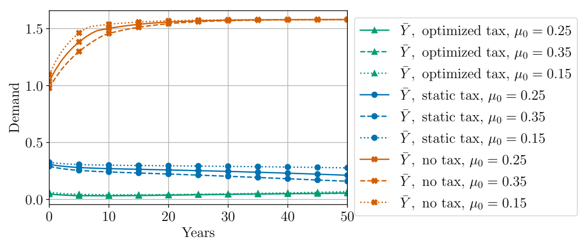
(c) Share of green citizens μ



(d) Share of green firms γ



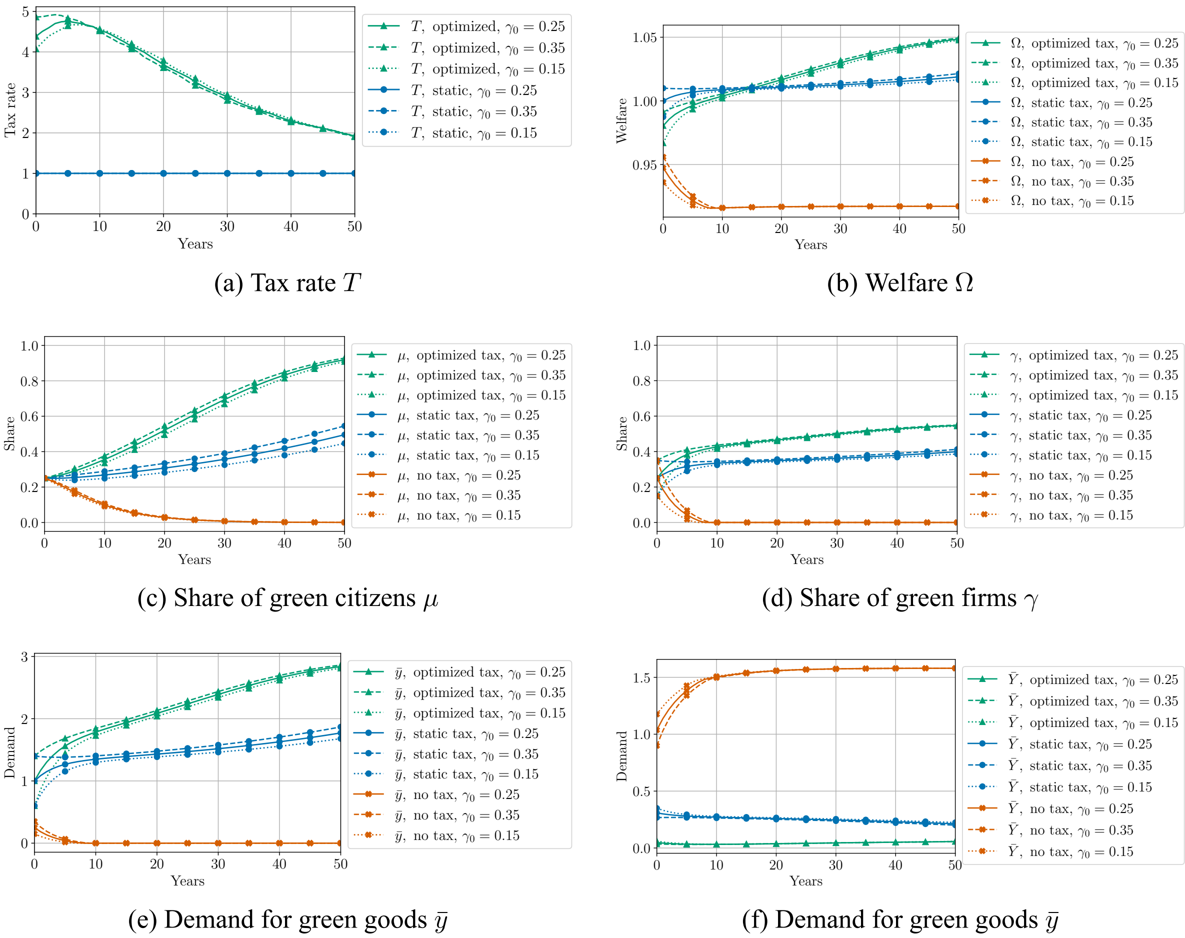
(e) Demand for green goods $\bar{y}$



(f) Demand for green goods $\bar{y}$

Figure 12: Change in initial share of green citizens μ_0

Note: Tax rate, welfare, and demand are normalized by their initial value with a static tax and the basecase parameters.



(a) Tax rate T

(b) Welfare Ω

(c) Share of green citizens μ

(d) Share of green firms γ

(e) Demand for green goods $\bar{y}$

(f) Demand for green goods $\bar{y}$

Figure 13: Change in initial share of green firms γ_0

Note: Tax rate, welfare, and demand are normalized by their initial value with a static tax and the basecase parameters.