Pareto-improving climate policy with heterogeneous abatement costs in the building sector

Matthias Kalkuhl[1,2], Maximilian Kellner[1], Noah Kögel[1,2], and Lennart Stern[1]

[1]Potsdam-Institute for Climate Impact Research, Potsdam DE
[2]University of Potsdam, Faculty of Economics and Social Sciences, Potsdam DE

March 19, 2025

Abstract

We build a dynamic model in which home owners decide when to switch to carbon-neutral heating. Agents differ with regard to carbon intensity and abatement costs, the latter being non-observable, private information. The heating-related investment model is nested in an overlapping generations Mirrlesian optimal taxation model with heterogeneous home ownership and labor productivity. We develop a compensation mechanism which guarantees a weak Pareto-improvement and averts potentially severe welfare losses due to horizontal heterogeneity when climate policy passes the cost-benefit test. The mechanism includes carbon pricing with category-based transfers, uniform ad-valorem subsidies on investments financed by public debt and an income tax adjustment proportional to mitigation benefits to service debt. We show that exact compensation of homeowners' dynamic abatement costs requires only minimal information: the interest rate and the fossil fuel price path. By means of exact compensation, our model utilizes the income-tax system to redistribute heterogeneous transformation costs between households according to any number of normative considerations without efficiency losses. We numerically illustrate welfare costs of uncompensated horizontal heterogeneity, optimal subsidy rates and income tax adjustments for Germany.

JEL codes: D61; D82; H21; H23; H24; Q54

Acknowledgements

We thank Rüdiger Bachmann, Christian Bayer, Justis Caron, Ottmar Edenhofer, Karsten Neuhoff, Matthias Paustian, Karen Pittel, Jan Steckel, Alex Stomper for their feedback and comments as well as Karolina Rütten and Felix Knopp for helpful data research and feedback. Max Kellner acknowledges financial support from the German Federal Ministry of Education and Research (BMBF) via "Kopernikus Projekt Ariadne" (FKZ 03SFK5J0), Noah Kögel acknowledges financial support from the German Federal Ministry of Education and Research (BMBF) via "KlimFi Projekt FEIRE". All opinions and errors are those of the authors and do not reflect the position of the BMBF or the project Kopernikus. Lennart Stern acknowledge financial support by the Werner Siemens Foundation within the research project CERES.

1 Introduction

In order to achieve nationally and internationally agreed upon climate targets, carbon emissions have to decrease considerably over the next decade (IPCC 2023). In the existing building stock, this primarily relates to decarbonizing cooling and heating of residential buildings and office spaces. Unlike the decarbonization of other sectors, such as energy provision, the transformation of the building sector requires direct action and participation by a large number of small, individual actors, mostly private households. Carbon pricing should, in principle, be an efficient instrument to achieve emission reductions: investments in mitigation take place where the marginal value of abating is the highest, while the regulator requires virtually no information about households and their property structure. A key property of Pigouvian pricing via a carbon tax is that it fulfills the Kaldor-Hicks criterion: After introducing a Pigouvian price, aggregate benefits outweigh the cost. Uncompensated carbon pricing can, however, impose substantial costs on individuals (Többen et al. 2023; Edenhofer et al. 2021; Berry 2019) with severe political-economic as well as normative implications: Citizens often oppose high carbon prices, partly due to the high salience of the costs imposed on them but also due to misperceptions about the true cost of carbon pricing (Douenne and Fabre 2022; Carattini et al. 2017). Apart from the direct burden on individuals themselves, distributional effects and concerns about equity also increase opposition against carbon pricing (Bergquist et al. 2022; Sommer et al. 2022).

The main focus of economic analyses addressing this problem has been on rebating revenues from carbon pricing, e.g., as uniform per-capita transfers, to cushion distributional effects or even create a progressive distributional outcome (Klenert et al. 2018; Klenert and Mattauch 2016). Embedding climate policy in an optimal taxation framework allows to derive optimal, welfare-maximizing adjustments of the income tax schedule in a more general way (van der Ploeg et al. 2024; Jacobs and van der Ploeg 2019). This strand of the literature has also been augmented to account for so-called horizontal distributional effects, i.e., unequal costs of carbon pricing within the same income group. Fischer and Pizer (2019) argue that horizontal equity effects reduce social welfare due to loss aversion as the losses for some households are not outweighted by (equal) gains for other households in the same income group. Hänsel et al. (2022) show that, even disregarding loss aversion, the concavity of the utility function creates a normative reason to address horizontal inequality in the standard optimal taxation model. Furthermore, they show in a calibrated optimal taxation model that non-linear energy taxes or subsidies for low-carbon technologies are second-best when (horizontal) heterogeneity in abatement costs is not observable by the regulator.

While the optimal taxation framework provides a consistent answer to how income taxes or transfers should be adjusted after introducing carbon pricing, it cannot guarantee a Pareto-improvement. Hence, some (or many) individuals may be worse-off despite aggregate welfare gains. In theory, to ensure that no agent is worse off due to carbon pricing, lump-sum transfers could be implemented to exactly offset the pricing burden. Yet, in practice, differentiated transfers on the scale of individual households appear informationally infeasible. For the transport sector, Sallee (2019) shows that creating a transfer scheme based on time-invariant observable characteristics cannot achieve a Pareto improvement due to large unobserved heterogeneity. Pigouvian taxation, in fact, creates losers. The key underlying reason for this finding is (unobserved) horizontal heterogeneity: if individuals differ only with regard to labor productivity, costs and benefits of policies would be income-dependent. In that case, an income-neutral tax reform, as suggested by Kaplow (2012), can ensure that no income group loses from introducing a Pigouvian tax.

These insights translate also to the building sector where the cost of climate policy depends on a complex set of individual and building-related characteristics on both the intensive (i.e., use patterns) and extensive (i.e., technology choice) margin. Differences in energy efficiency or building characteristics result in unequal burdens not only on the vertical (between income groups) but also on the horizontal dimension (within income groups).

In this paper, we develop a policy package that ensures a Pareto-improvement under minimal informational requirements as long as aggregate benefits of the policy exceed aggregate cost, i.e., as long as the Kaldor-Hicks criterion is met. The focus on Pareto-improvements is particularly attractive from a political-economy perspective. It should be much easier to agree on a policy if it creates no losers but only winners. In contrast, if a policy does not achieve a Pareto-improvement, the losers can be expected to prefer the status-quo and oppose it. The more heterogeneous the abatement costs, the larger the share of losers in the population becomes, making it also difficult for policy makers facing electoral competition to implement the respective policy. This is particularly problematic when future generations, who are not represented in present-day elections, are the main beneficiaries of a policy. Finally, when some individuals benefit from a policy while others lose, there can be polarized public debates about the fairness of cost distribution. Restricting attention to Pareto-improving policies preempts these problems by providing a universally beneficial way out of an inefficient allocation caused by market failures. Thus, our approach provides a climate policy design with minimal distributional conflicts.[1]

Besides the political-economy argument for Pareto-improvements, our model has another important property as it is capable of transferring the large cost heterogeneity within and between income groups from private households to the public budget without efficiency losses. This has three important advantages: First, it enables redistributing abatement costs in a transparent and flexible way through adjustments in the income-tax systems. Hence, accounting for various normative or political preferences, abatement costs can now be redistributed along the income dimension in a very flexible way according to what societies perceive as fair (Pareto-improvements are then only one special form of redistribution). Second, our approach eliminates uncertainty about individual abatement costs which might be a decisive factor for overestimating the cost of climate policy (Douenne and Fabre 2022). Third, uncertainty by itself constitutes a welfare loss due to risk aversion. Using German data, we compute that, if individuals are loss averse, perfectly compensating the vertical heterogeneity along income, yet, neglecting the horizontal dimension, still incurs an average welfare loss of 22% of pre-tax heating expenditures or an annual total of EUR 6.9 bln for a carbon price equal to the social cost of carbon in 2024 (EUR 182 per ton). Therefore, full compensation of unequal horizontal costs could also be understood as an insurance mechanism that is even welfare enhancing despite potentially high administrative complexity.

We employ a tractable dynamic model of decarbonization in the building sector where individuals differ with respect to the carbon intensity and abatement cost of their building stock. We develop a policy package consisting of five components: First, carbon pricing as the primary incentive for private households to switch to carbon-neutral heating technologies. Second, a category-based transfer is used to exactly offset the carbon price burden. The category-based transfer sorts households into groups which are homogeneous not only regarding their current carbon

[1] However, normative and political questions regarding the fairness of the status-quo as well as the distribution of the surplus generated by the policy remain unaddressed by our approach.

consumption but also their individually rational decarbonization times. Unlike a uniform lump-sum rebate or grandfathering based on historical emissions, this approach guarantees that compensation equals the direct (financial) burden associated with carbon pricing in a dynamic setting with endogenous investments. Importantly, the category-based compensation does not require the regulator to be informed about individual abatement *costs* ex-ante. Rather, it is sufficient that the *determinants* of cost heterogeneity are known resulting in a substantial informational advantage. In our model, abatement costs constitute the cost of preponing the investment to carbon-neutral heating. These costs cannot be compensated for by rebating the carbon price revenue. Thus, as the third component, a uniform ad-valorem subsidy on investments in and operating costs of carbon-neutral heating technologies is needed to offset abatement cost heterogeneity. Calculating the subsidy which exactly offsets abatement costs for (almost) arbitrary cost heterogeneity requires only knowledge about the discount rate and the fossil fuel price path. Information on individual building characteristics is not required. We show that this package ensures exact compensation whilst generating a budget deficit equal to the aggregate induced abatement cost. Fourth, this deficit is fully funded via additional public debt. Fifth, the intertemporal income tax schedule is adjusted in accordance with the income elasticity of climate damages to service public debt. By embedding our housing-sector model in a generalized Mirrlesian overlapping generations (OLG) model, we show that this five-component package ensures a Pareto-improvement as long as the discounted aggregate benefits of climate policy exceed its discounted aggregate cost. The regulator issues public debt to shift the cost of climate policy over time to where the benefits occur. This ensures that a Pareto-improvement can be achieved across multiple generations even if households are not dynastic.[2]

We numerically compute subsidy rates and income tax adjustments for German energy price data and the German income tax schedule. Empirical estimates suggest that the price of natural gas increases at about half of the real interest rate. In this case, an initial subsidy rate of 14% ensures exact compensation. Providing an incentive to decarbonize via the subsidy, allows for a carbon price that is 26% lower than the social cost of carbon to achieve the underlying climate target. In contrast, if the market price of natural gas is stagnant, subsidy rates should start at almost 40% and continue to increase over time. Income tax adjustments depend strongly on the income-elasticity of climate damages and the aggregate size of avoided climate damages.

Our mechanism can also address some public behavioral economy problems that are believed to reduce the acceptability of carbon pricing: The carbon price with the category-based transfer could be labeled as a *bonus-malus* (or feebate) system for buildings. Although economically equivalent, such a framing can have important advantages: (i) It avoids or reduces some of the negative associations citizens have regarding carbon pricing or carbon taxes, (ii) it avoids the problematic cost salience of carbon pricing, as the bonus-malus system would primarily display net costs to homeowners (which are close to zero) although marginal price signals continue to provide full incentives, (iii) by integrating pricing and compensation within one single instrument, it is less prone to citizens' believes or distrust in the use of carbon pricing revenues by governments.

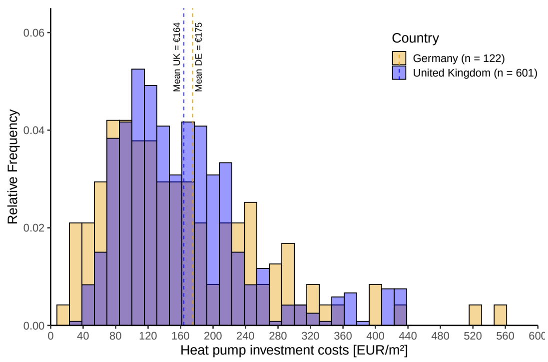
Our paper is organized as follows. In Section 2, we provide stylized evidence from the German building sector to illustrate the heterogeneity regarding both carbon price burden and retrofitting costs. In Section 3, we present the formal optimization problem of the private household. Section 4

[2]This decisive role of public debt is independent of the classical dynamic (in-)efficiency property of public debt in the standard OLG model (e.g., Blanchard 1985; Diamond 1965). Hence, our model is agnostic as to whether public debt increases dynamic efficiency or crowds out private capital.

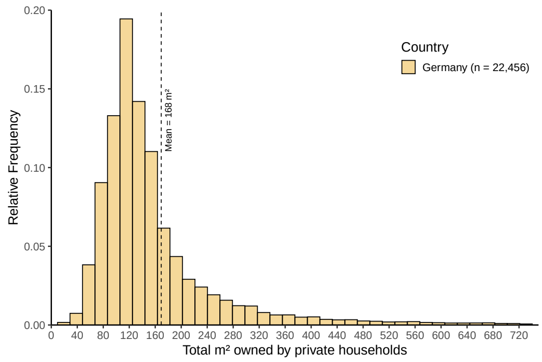
discusses the requirements for Pareto-efficiency and derives the components of our policy package. Extensions and generalizations of the main mechanism are examined in Section 5, followed by a calibrated numerical analysis in Section 6. The final section concludes the paper.

2 Cost heterogeneity in the building sector

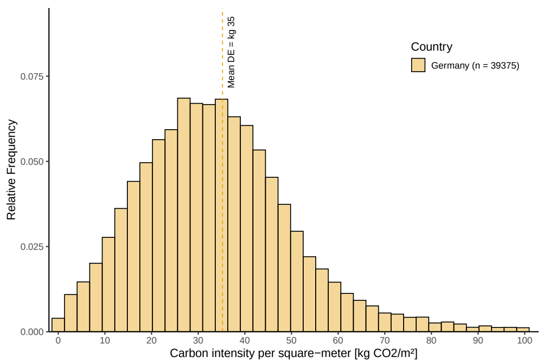
In this section, we identify several sources of heterogeneity related to the cost of decarbonizing the building sector, stemming from carbon pricing, retrofitting as well as installing and operating carbon-neutral heating technologies (heat pumps in particular). We employ stylized data from the representative, large-scale German sample survey on income and consumption (EVS) reporting 2018 household data, the representative German Heating and Housing Panel on retrofitting and energy consumption between 2000 and 2021 (RWI, see Frondel et al. 2023) as well as a field study on the feasibility and economic costs of heat pump installation in the United Kingdom in the period 2020-2022 (UK, see Energy Systems Catapult 2023).



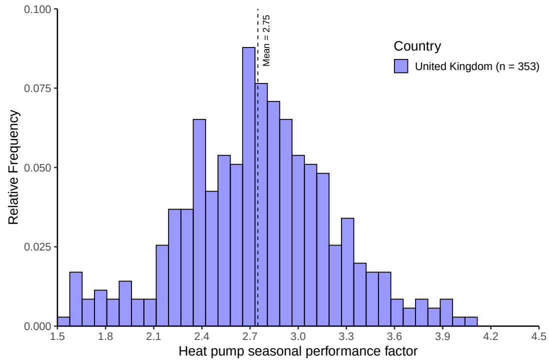
(a) Heat pump investment cost (Source: RWI, UK)



(b) Private residential ownership (Source: EVS)



(c) Carbon intensity (Source: EVS)



(d) Heat pump efficacy (Source: UK)

Figure 1: Distribution of drivers of cost heterogeneity in the building sector (Source: own depiction based on data from German Sample Survey on income and consumption (EVS), German Heating and Housing Panel (RWI), Energy Systems Catapult (UK))

Figure 1 illustrates the considerable variation in the determinants of these heterogeneities. According to Figure 1a, the investment cost of switching to a carbon-neutral heat pump may vary between less than EUR 20 to more than EUR 500 per square meter of floor area. Note that these values are sizable given an average annual basic rent of EUR 90 per square meter in Germany (Destatis 2023a). While larger buildings typically benefit from economies of scale, the cost also depends on which type of heat pump is installed (e.g., air-air v. air-ground), how easily it can be retrofitted into the building and regional variations in labor costs. Another factor of heterogeneity of particular relevance in Germany, where 54% of residential units are inhabited by tenants, 58% of which are owned by private landlords (Statistische Ämter des Bundes und der Länder 2019), is the distribution of ownership depicted in Figure 1b. In an efficient housing market, the cost and benefits of switching to carbon-neutral heating will be capitalized in rent prices (Myers 2019). However, in the short run, sticky prices, frictions or credit restrictions may affect landlords' willingness and ability to adopt green technologies at a large scale.

Additional heterogeneity is embedded in the operational costs of both fossil and carbon-neutral technologies. How strongly households are affected by carbon pricing, i.e., the opportunity cost of operating vintage fossil systems, is directly contingent on the carbon intensity of their respective residences as shown in Figure 1c. The heterogeneous burden due to carbon pricing primarily stems from variations in energy efficiency, heating technologies and preferences for room temperature[3]. In contrast, the cost of operating a heat pump is correlated to its efficacy, measured by the seasonal performance factor in Figure 1d, which is not only driven by energy efficiency and insulation. For instance, in two otherwise identical buildings, heat pump efficacy also depends on the size of the heating surfaces. Typically, larger surface areas (such as in floor heating) enables operation at lower supply temperatures and a higher efficacy (Khorasanizadeh et al. 2014).

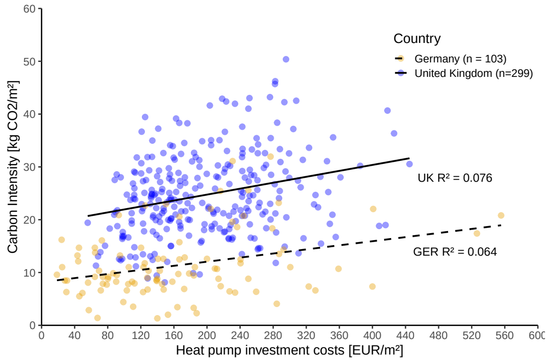
In principle, differentiated transfers based on a readily observable characteristic could still be employed for effective compensation if the heterogeneities presented in Figure 1 are (strongly) correlated. However, Figure 2 shows that none of the presented investment and operational costs appear to follow a shared distribution. For instance, buildings subject to high investment costs for heat pumps are neither particularly carbon intense prior to switching (Figure 2a) nor do we find a clear correlation with heat pump efficacy (Figure 2b). Furthermore, energy efficient buildings, i.e., those with low carbon intensities, do not display systematically higher heat pump efficacies (Figure 2c).[4] Finally, Figure 2d illustrates that some households with low to medium incomes also own substantial quantities of residential living space (either as owner-occupiers or landlords) which may give rise to issues of credit availability for financing switching costs.

These stylized empirical results suggest that Pareto-improving climate policy has to compensate for manifold heterogeneity. Relatively accessible observables such as income or energy efficiency alone are not sufficient to design transfers able to offset the costs induced by climate policy. However, this does not imply that informationally infeasible complete knowledge of all costs is required to achieve a Pareto improvement.

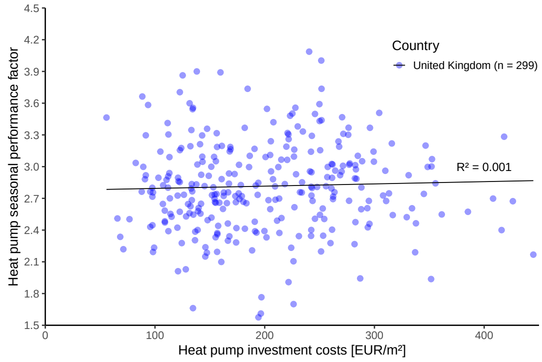
Moreover, addressing the horizontal dimension of heterogeneous transformation costs is not only relevant from a politico-economic perspective. Figure 3 illustrates the welfare loss induced by carbon pricing if revenues are fully rebated as income-dependent transfers. Even with exact compensation along the vertical income dimension for each income percentile, the remaining horizontal variation in the carbon intensity constitutes an uninsured risk: carbon pricing with full revenue recycling within each income percentile is a negative lottery for homeowners. With risk averse households, an initial carbon price of EUR 182 (the social cost of carbon in 2024) causes

[3]The energy savings potential of (deep) retrofits is highly case-specific and depends on the building material, circumstances, untertaken measures and owners' behavioral biases. Accordingly, the energy saving potential may vary between 20-80% as empirical evidence from Switzerland and Germany shows (Hondeborg et al. 2023; Galvin 2024). Residential energy demand decreases by about 6.9-8.2% if set-point temperature was reduced by 1°C in German residential buildings (Becker and Knoll 2011). Computer simulations for US family homes find elasticities of -5.4% and -2.7% (Moon and Han 2011).

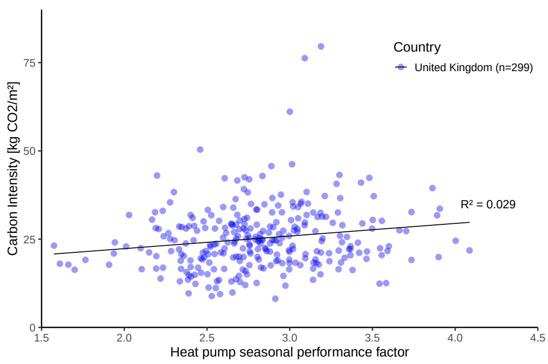
[4]Figure 2c is based on imputed carbon intensities, thus, underestimating the respective heterogeneity.



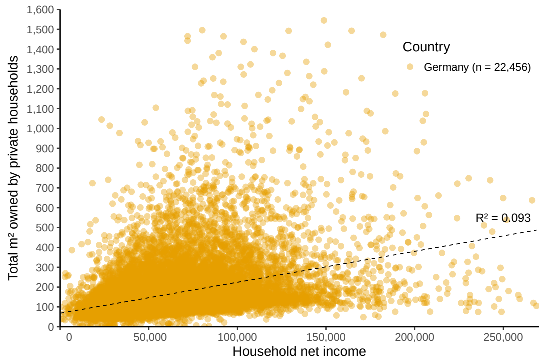
(a) Source: RWI & UK



(b) Source: UK



(c) Source: UK

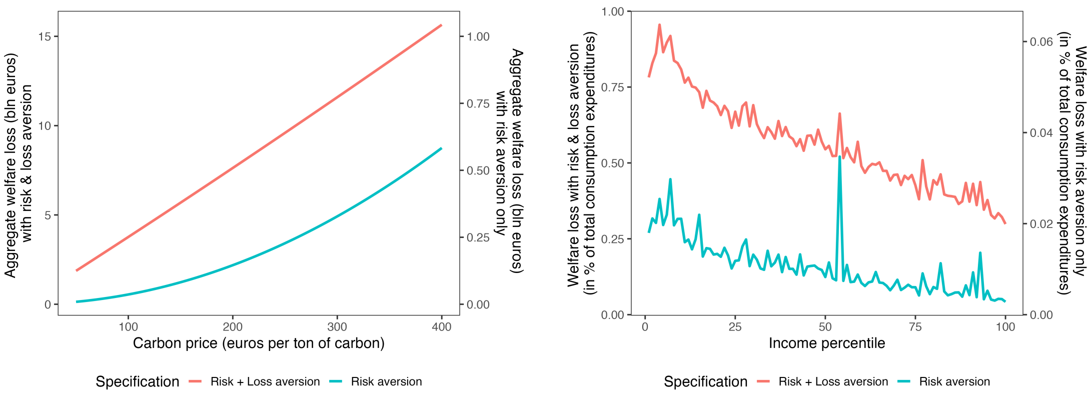


(d) Source: EVS

Figure 2: Correlation between drivers of cost heterogeneity in the building sector (Source: own depiction based on data from German Sample Survey on income and consumption (EVS), German Heating and Housing Panel (RWI), Energy Systems Catapult (UK))

a welfare loss of EUR 121 mln in the first year, multiplying to EUR 6.9 bln if we additionally consider loss aversion (see Figure 3a).[5] Hence, any transfer scheme eliminating the remaining horizontal variation, such as the category-based transfer developed below, generates a welfare gain as long as its annual excess administrative cost compared to the income-based transfer remains below this number. The welfare cost of heterogeneous carbon price burdens diminishes over time as households react to the carbon price and adopt carbon-neutral heating technologies. At the same time, switching technologies also induces horizontally heterogeneous costs. Due to a lack of readily available data, we cannot reliably estimate the associated welfare loss. Yet, based on the distributions depicted in Figure 1, the welfare cost may well be in the same range as the cost of horizontally heterogeneous carbon burden.

To see the implications for income inequality, we decompose the total welfare loss at a carbon price of EUR 182 per ton into relative consumption losses experienced by each percentile of the income distribution in Figure 3b. Despite the vertical heterogeneity associated with carbon pricing already being eliminated by the income-based rebate, a clearly regressive trend persists in the welfare losses. While the lowest percentiles face losses equivalent to up to 1% of their annual consumption if loss aversion is taken into account, the upper end of the income distribution only faces losses of about 0.3% of their consumption expenditures. Hence, from a distributional perspective, it appear also relevant to address the horizontal heterogeneity of costs related to climate policy.



(a) Aggregate welfare cost

(b) Welfare loss at EUR 182 carbon price

Figure 3: Welfare effects of imperfect carbon price compensation in Germany (Source: own calculations based on EVS 2018)

We now proceed to outline an integrated policy package with minimal informational requirements, still capable of exact compensation while maintaining efficient switching incentives.

3 The basic model

Assume that there are N households. For ease of exposition, suppose that each household i lives in an owner-occupied building. We disregard heterogeneities in real estate ownership and labor productivity (income) for the moment but introduce them in Section 4.3. Household i's heating

[5] We employ a loss aversion factor of $\lambda = 2$, implying that losses count twice as high as gains. Computations are detailed in Supplementary Appendix B.1.

related intertemporal utility is given by

$$u_i(\tau_t,\beta) = -\int_{t=0}^{s_i} \left\{ e^{-rt}(p_t^F + \tau_t)\varepsilon_i \right\} - e^{-rs_i}(1-\beta)C_{i,s_i} + T_i, \tag{1}$$

where the integral on the RHS of (1) denotes the total discounted cost of using a fossil-fired heating system until the switching period s_i. Employing the fossil technology induces operational costs equal to the market price of fossil fuels, p_t^F, and the carbon tax, τ_t, per consumed unit of fuel. How much fuel a household requires is contingent on its energy demand ε_i. We abstract from endogenous adjustments of the room temperature assuming that energy demand is solely defined by building characteristics, e.g., insulation, exterior surface and location. This approach is motivated by the fact that the price elasticity of heating demand is low in the short run, when only behavioral adjustments are possible. In the long run, when households can react to price changes by retrofitting and investing, substantially higher elasticities are observed.[6] The last term in (1), T_i, denotes total discounted public transfers which the household considers as given. Hence, household i's optimization problem boils down to choosing s_i such that the total cost of heating, i.e., the sum of the first two terms in (1), is minimized.

From s_i onward, the household switches to a carbon-neutral heating technology (e.g., a heat pump) which amounts to total switching costs C_{i,s_i}. The switching cost may be subsidized by the government at rate β and consists of four components

$$C_{i,s_i} = \underbrace{c_i^I}_{Insulation} + \underbrace{c_i^R}_{Radiators} + \underbrace{\frac{c_i^H}{1-e^{-rT^H}}}_{\text{Heat pump replaced every } T^H \text{ years}} + \underbrace{\int_{t=s_i}^{\infty} \left\{ e^{-r(t-s_i)} p_t^E \frac{\varepsilon_i(c_i^I)}{\delta_i(c_i^I, c_i^R, c_i^H)} \right\}}_{\text{Operational cost for running the heat pump}}. \tag{2}$$

The first two terms denote investment costs in insulation and radiators, where the former improves the energy efficiency of the building and the latter captures that household i may have to invest in larger radiators or floor heating.[7] Heat pumps are assumed to have a technical lifetime equal to T^H and installation cost, c_i^H, such that the third term captures the geometrical sum of all future replacement costs.[8] The last term on the RHS of (2) represents the operational cost of the heat pump contingent on the electricity price, p_t^E, energy demand, $\varepsilon(c_i^I)$, which is decreasing in insulation spending, and the efficacy of the heat pump, δ_i, i.e., its seasonal performance factor.

In general, C_{i,s_i} may be endogenous to the switching time, s_i. Consider, for instance, that the electricity price increases over time. Then, *ceteris paribus*, decreasing the carbon price will postpone s_i which also implies higher operational costs due to an increasing electricity price, p_t^E. In this case, it may also be optimal to re-adjust investments c_i^I and c_i^R to account for a higher electricity price. For the sake of tractability, we assume a constant price of electricity, p^E, implying that $C_{i,s_i} \equiv C_i$ as p^E may be factorized in (2) such that s_i no longer affects the switching cost.

[6] While Ruhnau et al. (2023) and Auffhammer and Rubin (2018) estimate demand short-run elasticities of about 20%, Buchsbaum (2023) find sixteen times larger elasticities in the long than in the short run.

[7] By subsuming these components in C_{i,s_i}, we implicitly assume that households concurrently invest in energy efficiency measures and carbon-neutral heating. In reality, it may be efficient to separate these investments which will be discussed in more detail in Section 5.4.

[8] Note that we consider vintages for the heat pump (infinitely lived technologies, $T^H \to \infty$, could be assumed without loss of generality), yet not for the fossil system. For the category-based transfer to function, we have to avoid bunching of switching times and, thus, implicitly assume that all fossil equipment is replaced before the end of its technical lifetime.

4 Pareto-improving climate policy

To identify Pareto-improving policies, we proceed as follows: First, we design a policy package that completely offsets all costs of climate policy for each homeowner. Second, we derive an income-tax adjustment scheme to finance compensation payments, accounting for the benefits of reduced climate damages. The cost of climate policy is contingent on a counterfactual reference point which we denote as the business-as-usual (BAU) scenario. In the BAU, the government adheres to the pre-existing policy package $(\tau^{BAU}, \beta^{BAU})$. Thus, Pareto-improvement requires

$$u_i(\tau_t, \beta) \geq u_i(\tau_t^{BAU}, \beta^{BAU}) \quad \forall i. \tag{3}$$

In addition, to a constant electricity price, $p^E = const.$, we assume that fossil energy prices grow with the market interest rate, r, as suggested for depletable resources by the Hotelling rule, i.e., $p_t^F = p_0^F e^{rt}$, for the case where extraction costs are negligible (this assumption will be relaxed in Section 5.1, where we show that our results also hold under fairly general assumptions about the fossil price). Together, these assumptions about the energy prices imply that (carbon-neutral) electricity becomes relatively cheaper in comparison to fossil fuels as time progresses. This is supported by evidence from the literature (Creutzig et al. 2023). Therefore, even in the BAU without climate policy intervention, all households will eventually transition from a fossil heating system to a carbon-neutral technology. In the BAU, total cumulative emissions over time amount to

$$E^{BAU} = \sum_i s_i^{BAU} \varepsilon_i. \tag{4}$$

Climate policy has the objective of reducing cumulative emissions at least (aggregate) cost to $E^* < E^{BAU}$ by providing an incentive to shift the households' individually rational switching time forwards such that $s_i^* < s_i^{BAU}$. For ease of exposition, suppose that fossil fuel prices and emission prices grow with the interest rate and that both emission taxes and subsidies are equal to zero in the BAU. Then, by substituting (1) for s_i^* and s_i^{BAU}, respectively, in (3) and imposing equality, we obtain

$$u_i^* - u_i^{BAU} = \underbrace{-s_i^* \tau_0^* \varepsilon_i}_{A} \underbrace{-C_i \left[(1-\beta^*)e^{-rs_i^*} - e^{-rs_i^{BAU}}\right]}_{B} + \underbrace{\left(s_i^{BAU} - s_i^*\right) p_0^F \varepsilon_i}_{C} + T_i^* \stackrel{!}{=} 0. \tag{5}$$

Shifting the individually rational switching time from s_i^{BAU} to s_i^* by means of climate policy has three effects on household utility. First, households are subject to an annual *carbon price burden*, $\tau_t \varepsilon_i$, as long as it is still optimal for them to use fossil fuels for heating. The net-present value of this burden amounts to term A in (5). It can also be understood as a 'stranded asset' cost as the carbon price applies to a past investment decision that cannot and should not be changed until the optimal switching time has come. Second, shifting the individual switching time forwards causes an additional cost of *foregone interest earnings* captured by term B. Each period that the technology transition occurs earlier, the household loses interest payments on the aggregate cost of switching, C_i.[9] Third, earlier switching also comes with a monetary benefit as the household avoids expenditures for fossil fuels for the duration $s_i^{BAU} - s_i^*$ as denoted by term C. These savings cannot

[9] Alternatively, this may be interpreted as the additional cost of taking out a loan to fund the investment earlier.

outweigh the foregone interest earnings in term B. Otherwise, it would already be individually rational to switch earlier under the BAU regime. Both components, B and C, are household-specific, where we denote the total, $B+C$, as household-specific *induced abatement costs*.

To reach exact compensation, the regulator has to choose τ_t, β and T_i such that all three effects in (5) aggregate to zero for each i. Under perfect information, the government could offset both cost components via an individual lump-sum transfer, T_i^{fb}. In this case, no subsidy is required and the climate policy package only consists of first-best carbon pricing and lump-sum transfers. As apparent from (5), determining T_i^{fb} requires *ex-ante* knowledge about the set of N individually rational switching times both in the BAU, s_i^{BAU}, and in the presence of climate policy, s_i^*, in addition to C_i and ε_i. Due to the heterogeneity illustrated in Section 2, the informational requirements for a first-best lump-sum compensation appear prohibitively high. We propose an alternative mechanism consisting of carbon prices, τ_t^*, a uniform subsidy rate, β^*, and, as the novelty of our design, category-based lump-sum transfers, T_i^θ. Our mechanism achieves exact compensation at far less stringent informational requirements than the first-best transfer, i.e., it is sufficient for the regulator to know the *determinants* of each household's cost structure *ex-ante* instead of explicit costs and switching times.

Note that exact compensation is not possible with a balanced budget even if first-best transfers are feasible. While the regulator earns revenues from carbon pricing amounting to the households' total carbon price burden, additional public funds are required to offset induced mitigation costs. Building on the income-neutral tax reform approach of Kaplow (2012), the deficit caused by climate policy can be funded via income or capital tax adjustments in a strictly Pareto-improving way under fairly general assumptions about the incidence of climate damages. Uninhibited climate change will cause substantial losses to productivity and welfare (Burke et al. 2015; Kalkuhl and Wenz 2020). Recent estimates put the Social Cost of Carbon in the range of USD 140-420 in 2030 depending on the discount rate (EPA 2023). The literature also shows that the benefits of the Paris agreement may outweigh the cost of climate policy (Glanemann et al. 2020; Hänsel et al. 2020). This implies that climate policy – when meeting the Kaldor-Hicks-criterion – constitutes a *potential* Pareto-improvement. In combination with an incentive-neutral redistributive policy, an *actual* Pareto-improvement can be achieved. We will discuss this in more detail in Section 4.3.

4.1 Category-based transfer

To ensure full compensation of the carbon price burden, the regulator categorizes all households into Θ bins depending on building characteristics. The characteristics determining category θ have to be chosen such that all buildings of type θ face the same energy demand, ε_i, and transformation cost, C_i, per square-meter of living space. This ensures two central properties of this mechanism. First, all households in one category are described by an identical carbon price burden per square-meter, $\tau_t \varepsilon_i / sqm_i$, in period t. Thus, refunding the revenues generated *within* category θ, i.e., $\sum_{i \in \theta} \tau_t \varepsilon_i$, on an equal per square-meter basis to all households $i \in \theta$ perfectly compensates for their carbon price burden. Since

$$\frac{\varepsilon_i}{sqm_i} = \frac{\varepsilon_j}{sqm_j} \quad \forall i \neq j \in \theta, \tag{6}$$

the rebate to household i may be designed in a way that it is only conditional on the emissions of all other households $j \neq i$ of type θ in order to prevent households from strategically adjusting their emissions. Hence, the perfectly offsetting rebate to household i in period t is defined as

$$T_{i,t}^{\theta} := T_{i\in\theta,t} = \psi_{t,\theta} \frac{sqm_i}{\sum_{j\in\theta} sqm_j} \sum_{j\in\theta:j\neq i} \tau_t \varepsilon_j \quad \forall t < s_i, \tag{7}$$

where $\psi_{t,\theta}$ is a parameter ensuring that carbon pricing revenues are fully disbursed in each period t within each category θ. Second, if all households in one bin face identical transformation costs relative to the energy demand, $\frac{C_i}{\varepsilon_i} = \frac{C_j}{\varepsilon_j}$ for all $i, j \in \theta$, our categorization ensures that all households switch to the renewable heating technology at the same point in time, s_θ. This implies that form $t \geq s_\theta$ onwards, both revenues from carbon pricing and lump-sum transfers in category θ drop to zero, i.e., no household is overcompensated after switching.

To illustrate the advantages of a category-based mechanism in comparison to a uniform lump-sum transfer or grandfathering of emission allowances, consider a scenario where equal shares of the population inhabit one of two different building types, both of which use fossil fuels for heating in $t = 0$. The first building type has already been retrofitted to a high energy efficiency standard with low energy demand. The other building type is poorly insulated and, accordingly, has higher energy demand. Due to $\varepsilon_1 < \varepsilon_2$, the first building type is subject to a lower carbon price burden than the second. Therefore, a uniform lump-sum transfer, $T_t^U = \tau_t(\varepsilon_1 + \varepsilon_2)/2$, overcompensates type 1, whereas the rebate to type 2 falls short of its burden. Grandfathering emission allowances, e.g., based on historical fuel consumption, can avoid this problem in the short run. However, this does not hold in a dynamic setting. Assuming that owners of type 1 buildings face lower investment costs relative to energy demand, $\frac{C_1}{\varepsilon_1}$, it will be individually rational for them to switch earlier than type 2 households. If the regulator wants to cut transfers to type 1 after switching, the reference point for grandfathering has to be reset in the future. This entails detrimental behavioral incentives as discussed in Section 5.4. Neither problem occurs in the category-based transfer as buildings in each bin are, by definition, identical and revenues from carbon pricing are rebated within each group so that cross-funding is ruled out by design.

Lemma 1 states formally sufficient conditions for the category-based transfers to achieve exact compensation without distorting optimal switching times. Essentially, this results in a 'perfectly differentiated command-and-control' scheme where the regulator induces each homeowner to switch at the individually optimal time consistent with target E.

Lemma 1. *Suppose that buildings are sorted in a way such that all buildings in a category face the same relative transformation cost, C_i/ε_i, and relative emission intensity, ε_i/sqm_i. Then, the individually optimal time to switch to carbon-neutral heating is identical for all buildings within a category. Moreover, in equilibrium, the category-based transfer will be identical to the carbon price burden for each building at each point in time.*

The ability of a category-based transfer to completely offset the carbon price burden critically depends on how accurately buildings are sorted into each bin. If, in the optimum, all buildings in a category are perfectly identical regarding the conditions in Lemma 1, our mechanism is able to exactly offset the carbon price burden. The more heterogeneous a category becomes, e.g., due to faulty sorting, the less precise the compensation. To ensure effective binning, the category-based transfer requires arguably more information than either a uniform rebate or grandfathering. Nonetheless, there is no need for *ex-ante* knowledge about the carbon price burden or investment costs. It is sufficient to identify characteristics which are informative about cost structures such as building type, age, insulation or heating technology. Finally, defining bins based on past emissions has the same informational requirements as grandfathering, but still reduces heterogeneity compared to a simple grandfathering approach as switching-time heterogeneity is accounted for. With a category-based transfer in place, it is also possible to combine carbon price payments and the rebate into a single payment leading to the following lemma:

Lemma 2. *The carbon price with category based transfers is equivalent to a bonus-malus system where in each category homeowner i pays* $BM_i^\theta = \tau_i \varepsilon_i - T_i^\theta$. *The bonus-malus results in zero payments if buildings are perfectly sorted into categories and homeowners act in line cost-minimization principles.*

The bonus-malus system has three key political economy advantages. First, the system only relies on an internal shadow price for emissions to compute the net tax or transfer, respectively, after deducting the category-based transfer. Thus, explicit mention of the carbon price is not necessary, avoiding negative associations and limited support associated with the mention of the term 'carbon pricing' (Douenne and Fabre 2022; Sommer et al. 2022). Second, the system circumvents the high salience of carbon prices and public attention attached to price increases, as only the net payments, BM_i^θ, become salient which either are very small or equal to zero if Lemma 2 holds. Third, directly combining carbon pricing and rebates into a single payment foregoes the risk that citizens understand a separate rebate as windfall rather than as a compensation for the carbon price (Souleles 1999), while also preempting mistrust in public institutions to actually rebate the revenues.

However, an integrated bonus-malus system may suffer efficiency losses due to behavioral biases. In particular, Ito (2014) shows that a considerable share of consumers are prone to employ average instead of marginal costs in their decision making which would lead to late or insufficient retrofitting and decarbonization investments in the case at hand. This effect, in turn, is attenuated by positive signaling externalities associated with the category-based transfer. While households following an average-cost optimization or acting under incomplete information may delay their switching time for too long, other fully-informed household within the same category will decarbonize at the individually rational time. As soon as these households switch, they will incur a bonus at the expense of not-yet decarbonized households who then run a malus, i.e., exact compensation no longer occurs if some households miss the optimal switching time. This creates an information spillover from fully-informed households to those with behavioral biases or incomplete information, signaling via the malus payment that it is time to invest in a carbon-neutral technology. The signal becomes especially salient when carbon price and rebate are combined into a single net payment.

4.2 Carbon price paths and subsidies for efficiency and exact compensation

Next, we define a policy package (τ_t, β) which ensures that both the cumulative emission target, E, is achieved at minimal cost and the Pareto-criterion in (3) is met.

Since the carbon budget may be treated equivalently to a depletable resource, Hotelling's rule tells us that the target, E, is achieved at minimal cost by a tax rate that grows with the interest rate, $\tilde{\tau}_t = \tau_0(\tilde{\beta}, E)e^{rt}$. Here, $\tau_0(\beta, E)$ denotes the initial carbon price such that τ_t implements E when combined with the uniform ad-valorem subsidy rate, $\beta \geq 0$. Taxation according to Hotelling's rule does not require additional subsidies in order to reach E, i.e., $\tilde{\beta} = 0$. The category-based transfer in (7) fully offsets the carbon price burden. However, households are still subject to induced mitigation costs as given by the term $(B-C)$ in (5). Without additional transfers or subsidies, the pure Hotelling price path violates the Pareto-criterion by exactly the foregone interest earnings. In Appendix A.1, we derive the following result for the efficient carbon price path:

Lemma 3. *A policy package consisting of a positive ad-valorem subsidy rate, β, and the carbon price path*

$$\tau_t = \left[(1-\beta)\tau_0(0,E) - \beta p_0^F\right]e^{rt}, \tag{8}$$

achieves the same efficient allocation, $s_i^* = s_i^{fb}$ for all i, as a first-best policy consisting of Hotelling carbon prices and (informationally infeasible) differentiated lump-sum transfers.

Together with household i's first-order condition with regard to s_i, (8) can be used to solve for the individually optimal switching time

$$s_i = \frac{1}{r}\log\left[\frac{rC_i}{\varepsilon_i\left(\tau_0(0,E)+p_0^F\right)}\right], \tag{9}$$

as long as household i is not at the corner solution of switching immediately in $t=0$. The identity regarding the allocation between the Hotelling price path, $\tilde{\tau}_t$, and the policy package with $\tau_t, \beta > 0$ conveniently allows us to express (9) as a function of only the initial carbon price, $\tau_0(0,E)$ absent subsidies. Without loss of generality, we assume that the BAU carbon price is equal to zero, i.e., $\tau_t^{BAU} = 0$ for all t, while the pre-existing subsidy rate may be positive, i.e., $\beta^{BAU} > 0$.[10] In Appendix A.2, we show:

Lemma 4. *Suppose that the revenues from carbon pricing are fully rebated via a category-based transfer as defined by Lemma 3. Then, for a constant electricity price, p^E, and a fossil fuel price growing with the interest rate, $p_t^F = p_0^F e^{rt}$, the policy package consisting of the tax rate, τ_t^*, as defined by Lemma 1 and the constant subsidy rate*

$$\beta^* = \frac{\frac{\tau_0(0,E)}{p_0^F} - \log\left[\left(1-\beta^{BAU}\right)\left(1+\frac{\tau_0(0,E)}{p_0^F}\right)\right]}{\left(1+\frac{\tau_0(0,E)}{p_0^F}\right)} \tag{10}$$

achieves the emission target, E, at minimal cost whilst simultaneously ensuring that all homeowners i attain utility exactly equal to their BAU levels.

Lemma 4 states that for each emission target, E, there exists one unique ad-valorem subsidy rate, β^*, ensuring that each individual household i attains exactly the same utility in the presence of climate policy as in the BAU scenario, i.e., $u_i^* = u_i^{BAU}$. This also implies that the associated tax path, τ_t^*, is uniquely identified by (8). By applying a uniform subsidy rate to all types of costs associated with switching from fossil fuel to a carbon-neutral heating technology, we provide the same incentive to all households to move their switching time forwards regardless of their individual cost structure in (2). This becomes apparent by computing the time by which the switching

[10] We choose this approach for ease of notation. Our subsequent results will also hold for $\tau_t^{BAU} > 0$, as long as the revenues from BAU carbon pricing are already accounted for, i.e., assuming that they are not available to offset the cost induced by pursuing a more ambitious climate policy target, $E < E^{BAU}$.

time is brought forward by the policy which equals

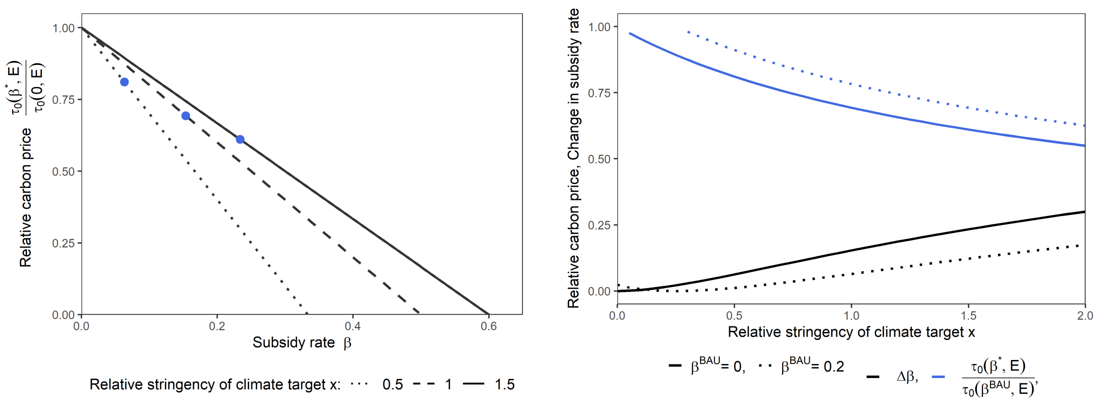
$$\Delta s = \Delta s_i \equiv s_i^{BAU} - s_i^* = \frac{1}{r} \log \left[\frac{\tau_0(0,E) + p_0^F}{\tau_0(0,E^{BAU}) + p_0^F} \right], \tag{11}$$

and is identical for all households i. The uniform subsidy substantially reduces the informational requirements of the mechanism as no information on individuals' costs are needed. The heterogeneity in the building stock may imply that one building has high investment cost for the heat pump itself (e.g., because piping has to be retrofitted while energy efficiency is already high) when, for another building, earlier switching is primarily inhibited by high investments in insulation or new radiators. It is also optimal to subsidize the operational cost of a heat pump, i.e., electricity used for heating, even though individual rationality ensures that no household switches their heating type before electricity becomes cheaper than fossil fuels. However, consider a building which is already well insulated and can easily be retrofitted with a heat pump. This building will have an early switching time in the BAU and gains few incentives to adopt earlier when only investment costs are subsidized. In order to also reduce this household's switching time by Δs_i and maintain Pareto-efficiency, it is necessary to further reduce the relative price of electricity.

As indicated by (8), the ad-valorem subsidy can substitute for carbon pricing in order to achieve a given emission target. Subsidies reduce the cost of switching to a carbon-neutral technology, leading to earlier individually optimal switching times, s_i, thus, lower cumulative emissions. Formally, the relation between the initial carbon price in the presence of positive subsidies, $\tau_0(\beta, E)$, and the respective price without subsidies, $\tau_0(0,E)$, can be expressed as

$$\frac{\tau_0(\beta,E)}{\tau_0(0,E)} = 1 - \beta \left(1 + \frac{1}{x}\right), \tag{12}$$

where $x = \tau_0(0,E)/p_0^F$ denotes the relative markup on the initial fossil fuel price required to achieve E when $\beta = 0$. Thus, for any given fossil price, p_0^F, a higher value of x corresponds to a tighter emission budget, E. Figure 4a illustrates the inverse relation between carbon price and



(a) Effect of subsidy rate on carbon price required to achieve climate target x

(b) Optimal combination of carbon price and subsidy rate to achieve climate target x

Figure 4: Interaction between carbon price and ad-valorem subsidy (Source: own depiction)

ad-valorem subsidy rate for different ambition levels, x. For instance, if the gross price of fossil fuels has to double to achieve the emission target without subsidies ($x = 1$), the same target could also be reached with a 50% lower carbon price, i.e., $\tau_0(\beta,E)/\tau_0(0,E) = 0.5$, if accompanied by an ad-valorem subsidy equal to $\beta = 0.25$ or even no price increase at all if the subsidy rate is set to $\beta = 0.5$. However, climate policy should not just implement any arbitrary point on this line. From Lemma 4, we know that, for each ambition level, E (or x, respectively, for given p_0^F), a unique combination of carbon price and ad-valorem subsidy ensures exact compensation at least cost (indicated by the blue points in Figure 4a). Substituting (10) into (12), the optimal initial carbon price, $\tau_0(\beta^*, E)$, relative to $\tau_0(0, E)$ can be expressed as

$$\frac{\tau_0(\beta^*, E)}{\tau_0(\beta^{BAU}, E)} = \frac{\log\left\{(1-\beta^{BAU})(1+x)\right\}}{x - \beta^{BAU}(1+x)}, \tag{13}$$

for $\beta^{BAU} \geq 0$. Figure 4b illustrates optimal combinations of $\tau_0(\beta^*, E)/\tau_0(\beta^{BAU}, E)$ and $\Delta\beta = \beta^* - \beta^{BAU}$ for $\beta^{BAU} \in \{0, 0.2\}$. For low ambition levels (i.e., low values of x), the emission target is achieved primarily via carbon pricing. With higher values of x and more ambitious targets, climate policy requires progressively earlier switching, increasing the induced mitigation cost. To offset the induced mitigation cost and enable exact compensation, the ad-valorem subsidy has to rise, subsequently leading to lower carbon prices in the optimal policy package. The substitution effect effect becomes less pronounced if there already are pre-existing subsidies in the BAU.

4.3 Public debt and income tax adjustments

The policy package described in Proposition 4 includes a positive ad-valorem subsidy rate which is constant over time, i.e., the regulator has to subsidize investments in energy efficiency and carbon-neutral heating technologies as well as the associated operational cost ad infinitum. The positive subsidy rate ensures that homeowners are indifferent between business-as-usual and the proposed policy package. However, the policy package is not revenue neutral. In Appendix A.3, we show:

Lemma 5. *The ad-valorem subsidy rate corresponding to exact compensation, β^*, as identified by Lemma 4, is associated with a discounted intertemporal budget deficit equal to*

$$D = \left\{ \frac{\tau_0(0,E)}{p_0^F} - \log\left[\left(1-\beta^{BAU}\right) \left(1 + \frac{\tau_0(0,E)}{p_0^F}\right) \right] - \frac{\beta^{BAU}}{1-\beta^{BAU}} \right\} \frac{p_0^F}{r} \mathbb{E}(\varepsilon_i), \tag{14}$$

where the expectations operator $\mathbb{E}(\varepsilon_i)$ denotes aggregate emissions if the mass of homes is normalized to one. This deficit exactly corresponds to the aggregate induced mitigation cost of achieving the (cumulative) emission target, E, at least cost.

Since the revenues from carbon pricing are already earmarked for the category-based transfer, the subsidy causes a total discounted fiscal deficit equal to D. Recall that we assume a carbon price equal to zero in the BAU scenario. The deficit, D, denotes the *additional* fiscal burden caused by increasing the subsidy rate by $\Delta\beta = \beta - \beta^{BAU}$. Whenever the status-quo subsidy rate is positive, i.e., $\beta^{BAU} > 0$, the regulator already incurs a deficit in the BAU.

Following Kaplow (2012), we show that the regulator can implement a combination of public debt and income tax adjustments to fund D and redistribute the remaining surplus to achieve a Pareto-improvement. Hence, climate policy can only be Pareto-improving if it passes the cost-benefit test, i.e., benefits in terms of avoided climate damages have to exceed the induced mitigation cost. Some assessments indicate that this might even the case for the ambitious (global) climate targets of the Paris Agreement (Glanemann et al. 2020; Hänsel et al. 2020). As the bulk of D will be composed of short-run investment costs, whereas benefits accrue long into the future, subsidies have to be funded by issuing additional public debt which will be repaid later when the benefits can be skimmed via income taxation.

To formally derive this result, we embed our optimal switching model in a Mirrleesian income tax model, by assuming that individual i's overall intertemporal utility equals

$$V_i\left[U_i(c_i, y_i), b_i, U_{-i}(c_{-i}, y_{-i})\right]. \tag{15}$$

The individual's total utility is determined by "self-regarding utility", U_i, bequests, b_i, and all other individuals' utility, U_{-i}.[11] Self-regarding utility, U_i, is determined by the consumption path, c_i, the individual production path, y_i, (more precisely, gross production before accounting for climate damages) and bequests, b_i. In the classical Mirrlees model, y_i equals the product of labor effort and productivity, i.e., $\omega_i l_i$.[12] In our application, it is not necessary to make any specific functional assumptions regarding the individual's production function. It is sufficient to note that U_i amounts to the maximal intertemporal utility individual i can achieve for given paths of consumption and production, i.e., U_i implicitly captures the disutility of labor.

For ease of exposition, we assume deterministic lifespans with death occurring at time T_i. Thus, the individual's bequest, b_i, is equal to whatever wealth has accumulated by T_i, i.e.,

$$b_i = e^{rT_i}\left\{w_i(b_{-i}) + u_i(\tau,\beta) + \int_{t=0}^{T_i} e^{-rt}\left[y_{i,t} - \mu_{t,e}(y_{i,t}) - c_{i,t} - \Gamma_t\left(y_{i,t} - \mu_{t,e}(y_{i,t})\right)\right]dt\right\}, \tag{16}$$

where w_i denotes the net present value of all wealth inherited from other individuals. Bequests can be split between all other surviving agents in any way. We assume that wealth yields the certain real interest rate, r. All forms of wealth are subsumed into w_i, including housing wealth, except for heating related investment and expenditures which are already accounted for in u_i. Thus, the actual wealth (in terms of NPV at at the time i is born) amounts to $w_i(b_{-i} + u_i(\tau,\beta))$.

In practice, the regulator will not be able to observe the counterfactual income absent climate damages, y_t. Hence, the income tax function, Γ_t, is levied on the individual's income net climate damages, i.e., $y_t - \mu_{t,e}(y_t)$. In line with the existing literature, we model climate damages, $\mu_{t,e}(y_t)$, as a function of private income, y_i, and the path of emissions, $e(t)$. This approach is sufficiently flexible to account for climate impacts that reduce overall total factor productivity of an economy (e.g., as in Nordhaus 1993) as well as productivity in various sectors which, in turn, allows to account for income-sensitive damages. Assume that climate damages have an income elasticity, ϕ, such that

$$\mu_{t,e}(y_t) = y_t^{\phi}\tilde{\mu}_t(e) \quad \forall \quad y_t, e, \tag{17}$$

where $\tilde{\mu}_t(e)$ denotes any function with e as its only argument. Additionally, we define the average marginal income tax rate on the interval, $[y_t - \mu_{t,e^{BAU}}(y_t), y_t - \mu_{t,e}(y_t)]$, as $m(z_t)$. As seen from (16), individual i's total utility depends on other agents' past actions (e.g., ancestors deciding about their bequests to i) and other agents' future actions (e.g., i deciding about how much to bequest to their

[11] As a special case of (15), individual i may attribute (more) weight to their direct descendants' self-regarding utility than for other individuals. For our application, it is sufficient to leave the model in its most general form.

[12] Both ω_i and l_i are unobserved by the government, only y_i is observed. This motivates why we use an income tax rather than individualized taxes. Analogously, and more generally, the use of the income tax can be motivated by the fact that the government does not know the individual utility function, $U_i(c_i, y_i, b_i)$, which indirectly captures how costly in terms of utility it is for individual i to achieve production y (e.g., in terms of the effort and hours required).

offspring). To solve the optimization problem, we invoke the concept of Subgame Perfect Nash Equilibria deriving the following result in Appendix A.4:

Lemma 6. *Consider a small open economy which can issue external public debt at the interest rate r. Suppose that the regulator implements the unique optimal policy package (τ_t^*, β^*) to achieve the climate target, E, with exact compensation at least cost associated with deficit D. Then, the regulator can adjust the income tax schedule in comparison to the BAU by*

$$\Delta\Gamma\left[y_t - y_t^{\phi}\tilde{\mu}_t(e)\right] = [1-m(z_t)]y_t^{\phi}\left[\tilde{\mu}_t\left(e^{BAU}\right) - \tilde{\mu}_t(e)\right], \tag{18}$$

in order to skim all benefits of avoided climate damages to fund D and generate a surplus if

$$\sum_i \int_{t=0}^{\infty} e^{-rt}\left[\mu_{t,e^{BAU}}(y_i^{BAU}) - \mu_{t,e}(y_i^{BAU})\right]dt > D, \tag{19}$$

i.e., if the net present value of the monetary benefits of climate policy exceeds the cost in terms of aggregate induced mitigation cost.

To gain an intuition about the income tax adjustment in (18), consider an emissions path that is only marginally below the BAU path. Then, mitigated damages equal $d\mu_t = \tilde{\mu}_t\left(e^{BAU}\right) - \tilde{\mu}_t(e)$. Defining $z(y_t) = y_t - y_t^{\phi}\tilde{\mu}_t(e)$, we can rewrite (18) as

$$\Delta\Gamma(z_t) = [1-m(z_t)]y(z_t)^{\phi}d\mu_t. \tag{20}$$

For marginal changes in climate damages, $d\mu_t$, the marginal income tax rate at the pre-tax income, z_t, remains the same as in the BAU. This does not hinge on assumptions about the functional form. In general, the proportion $m(z_t)$ of benefits due to avoided marginal damages is automatically skimmed via the BAU income tax schedule as pre-tax incomes increase with mitigation. The objective of an income tax adjustment is to capture the remaining $1-m(z_t)$ share of benefits. Differentiating (20) yields

$$\Delta\Gamma'(z_t) = \left[-m'(z_t)y(z_t)^{\phi} + [1-m(z_t)]\phi y(z_t)^{\phi-1}y'(z_t)\right]d\mu_t. \tag{21}$$

If climate damages are not correlated to income, i.e., $\phi = 0$, the income tax adjustment $\Delta\Gamma(z_t) = [1-m(z_t)]d\mu_t$ is equivalent to a uniform lump-sum tax of $[1-m(0)]d\mu_t < 0$ and a concurrent adjustment of the marginal tax rate of $\Delta\Gamma'(z_t) = -m'(z_t)d\mu_t$. In contrast, $\phi > 0$ requires $\Delta\Gamma'(0) = 0$ such that lump-sum taxation is no longer optimal. Instead the marginal income tax adjustment in (21) can be approximated as

$$\Delta\Gamma'(z_t) \approx \left[-m'(z_t)y(z_t)^{\phi} + [1-m(z_t)]\phi y(z_t)^{\phi-1}\right]d\mu_t, \tag{22}$$

assuming that climate damages account for a relatively small fraction of income and $y'(z_t) \to 1$.

Lemma 3 and 4 ensure that the policy package (τ_t^*, β^*) leaves every household exactly as well-off as in the BAU regarding heating-related costs, i.e., $u_i(\tau^*, \beta^*) = u_i(0, \beta^{BAU})$ for all i. Similarly, the income tax adjustment in Lemma 6 is designed such that the individually optimal path, (c_i, y_i), also remains unchanged in comparison to the BAU, i.e., overall $U_i^* = U_i^{BAU}$ for all i. Hence, the integrated reform consisting of the climate policy package and the income tax adjustment is distribution neutral such that every individual i is indifferent between the reform and the BAU.

However, if climate policy passes the cost-benefit test, i.e., if (19) in Lemma 6 is fulfilled, the income tax adjustment not only funds the deficit associated with exact compensation but also incurs the budget surplus

$$S=\sum_i \int_{t=0}^{\infty} e^{-rt}\left[\mu_{t,e^{BAU}}(y_i^{BAU})-\mu_{t,e}(y_i^{BAU})\right]dt-D>0, \tag{23}$$

leading to the following result:

Corollary 1. An integrated reform consisting of category-based transfers, $T_{i,t}^{\theta}$, rebating the revenues from the carbon price, τ_t^*, the ad-valorem subsidy, β^*, associated with a debt-funded public deficit, D, and the income tax adjustment, $\Delta\Gamma_t$, which services public debt, results in a strict Pareto improvement if the surplus, S, stemming from the income tax reform is rebated as a uniform lump-sum transfer.

Corollary 1 summarizes how our proposed mechanism achieves Pareto-improving climate policy despite many-dimensional heterogeneity in the building sector. The uniform lump-sum transfer of the surplus is employed here as a straightforwards means to show that the surplus can be re-distributed in order to achieve a strict Pareto improvement. The same result can be achieved by other rebate schemes, e.g., regressive income-based transfers. Alternatively, S may also be used to reduce pre-existing external debt.

5 Discussion and Extensions

In this section we discuss a number of important extensions, generalizations and modifications of the mechanism. We first relax the assumptions on Hotelling prices for fossil energy and carbon, showing that perfect compensation is possible for a broad set of general fossil energy and carbon price paths (including the case where the carbon price equals the social cost of carbon). Again we find that in these more general cases perfect compensation is possible without the need to know individuals' costs or even the cost distribution. Calculating the optimal subsidy and tax paths only requires knowledge of the discount rate, r, the growth rate of the fossil energy price in the BAU, g, the envisaged pre-reform carbon price path (i.e., the SCC) and the pre-reform subsidy rate. We further discuss additional aspects like endogeneous thermostat setting, accounting for other domains of climate damages and adverse incentives of grandfathering schemes, among others.

5.1 Non-Hotelling fossil price paths

In the main analysis, we assume that the price of fossil fuels grows with the interest rate, i.e., it moves along the Hotelling price path. For expositional purposes, we relax this assumption in two steps. First, if the fossil price grows at a constant positive rate, $g>0$, with $g\neq r$, it can be shown that the cost-effective policy package achieving exact compensation for all households consists of the ad-valorem subsidy path

$$\beta_t^*=\frac{g\left(e^{t(g-r)}+x\right)\left[\left(\frac{1}{xe^{t(r-g)}+1}\right)^{r/g}-1\right]+rx}{(r-g)\left(e^{(g-r)t}+x\right)}, \tag{24}$$

and the complementary tax rate

$$\tau_t^* = \frac{p_0\left(ge^{gt}+rxe^{rt}\right)\left[\left(e^{gt}+xe^{rt}\right)\left(\frac{1}{xe^{t(r-g)}+1}\right)^{r/g}-e^{gt}\right]}{(g-r)\left(e^{gt}+xe^{rt}\right)}. \tag{25}$$

When the fossil price does not grow at the interest rate, the optimal subsidy rate is no longer constant over time as it compensates for the difference between the growth rate of the fossil fuel price, g, and the interest rate, r.[13] For $g \to r$, we can apply L'Hôpital's rule to (24) to see that, in the limit, β_t approaches the constant rate β derived in the main analysis. Another special case of particular interest is a stationary fossil fuel price, i.e., $g \to 0$. By calculating the limits of (24) and (25), we find that the optimal tax and subsidy paths approach

$$\lim_{g\to 0}\beta_t = \frac{1}{1+\frac{e^{-rt}}{x}} \quad \text{and} \quad \lim_{g\to 0}\tau_t = \frac{p_0^F}{1+\frac{e^{-rt}}{x}}. \tag{26}$$

When $p_t^F = const.$, any household that does not switch immediately in $t=0$ will never adopt the carbon-neutral technology in the BAU, especially if $p_0^F < p_0^E$ and $\beta^{BAU}=0$. In this case, climate policy is the only means to reduce emissions over time.

5.2 Social cost of carbon, general fossil and carbon price paths

Our results are applicable to an even less restrictive set of both fossil and carbon price paths if we no longer confine ourselves to deriving closed-form expressions for β_t^* and τ_t^*. Most notably, this enables us to drop the assumption that carbon prices follow a Hotelling price path. While an efficient means to govern an exogenously specified carbon budget, carbon prices may also be determined by cost-benefit evaluation, e.g., the social cost of carbon as employed by the EPA (2023).

Denote by τ_t^{SCC} the carbon price as given by the social cost of carbon in t. It is sufficient that the carbon and fossil price paths fulfill three conditions for

$$\frac{du_i}{ds_i} = -\left(p_{s_i}^F + \tau_{s_i}^{SCC}\right)\varepsilon_i + rC_i = 0, \tag{27}$$

to characterize the unique efficient switching time of household i. First, gross heating expenditures in the initial period must not exceed the investment cost, i.e., $(p_0^F+\tau_0^{SCC})\varepsilon_i \le rC_i$, such that $du_i/ds_i|_{s_i=0} \ge 0$. Otherwise, households are in a corner solution switching immediately in $s_i=0$. Second, $\lim_{s_i\to\infty}(p_{s_i}^F+\tau_{s_i}^{SCC})\varepsilon_i > \sup rC_i$ ensures that du_i/s_i becomes negative for sufficiently large (i.e., late) switching times, s_i. Hence, if households delay switching for too long, operating the fossil heating system for one more period exceeds the interest cost of investing. Third, if the gross fossil price, $p_t^F+\tau_t^{SCC}$, is increasing over time, du_i/ds_i is monotonically decreasing. Together, these conditions ensure that it is individually rational for each household i to switch to the carbon-neutral technology at a finite point in time, i.e., $0 \le s_i^* < \infty$ for all i. Additionally, if the set of efficient switching times is bounded above, i.e., if the term

$$\frac{\frac{d}{ds_i}(p_{s_i}^F e^{-rs_i})}{(p_{s_i}^F+\tau_{s_i}^{SCC})e^{-rs_i}}, \tag{28}$$

[13]These results are formally derived in Supplementary Appendix B.2.

is bounded above, a unique combination of β_t and τ_t^{SCC} exists achieving exact compensation at least costs. In supplementary appendix B.3, we provide a detailed derivation of (28) for the an analogous case with Hotelling carbon and Non-Hotelling fossil price paths. We also display the differential equation that describes the the perfectly compensating subsidy rate and see that it only requires aggregate price information and no information on individuals' costs.

5.3 Endogenous thermostat setting and rebound effect

The main analysis assumes that households only react to climate policy via adjusting their optimal switching time. Especially in the short run before investing in a carbon-neutral technology, households may also react to carbon pricing by reducing the thermostat setting, i.e., by lowering the room temperature. However, as a positive ad-valorem subsidy requires a lower carbon price to achieve a given climate target, this incentive becomes diluted. After switching, the incentive to conserve energy remains weakened by the ad-valorem subsidy on the operational cost of heat pumps, i.e., electricity price. The aggregate efficiency loss equals the weighted sum of efficiency losses before switching (scales with carbon reduction squared) and after switching (scales with subsidy squared).

To obtain a numeric estimate of the efficiency losses, suppose that the households' short-run reaction is governed by the price elasticity of heating demand equal to $e_D < 0$. Then, the combination of carbon price and subsidy, $(\tau_t(\beta^*, E), \beta^*)$, derived above under the implicit assumption that $e^D = 0$ creates a deadweight loss compared to the pure carbon price absent subsidies, $\tau_t(0, E)$, as households increase their energy consumption.

This welfare loss is given by the Harberger triangle formula:

$$\frac{e_D}{2} (\beta^*)^2 p_t^F \varepsilon_i, \tag{29}$$

where we substitute $\beta^* = [x - \log(1+x)]/(1+x)$. For a price elasticity of $e_D = -0.2$ (Auffhammer and Rubin 2018), at a relative stringency of climate policy equal to $x = 0.59$ (see Section 6) and an interest rate of 2%, we get that the welfare costs due to insufficient incentives to reduce thermostat settings is initially approximately 0.063 percent of initial heating expenditure.

Assuming the cost of heat pump electricity is lower than the fossil price in $t = 0$, we can derive an upper bound for the discounted total of the deadweight loss due to inefficiently high heating consumption as a share of the discounted aggregate cost of decarbonizing the building sector as given by

$$\frac{1}{2} \frac{[x - \log(x)]}{1+x} \left(1 + r \frac{E}{\mathbb{E}(\varepsilon_i)}\right) e_D, \tag{30}$$

where the expectations operator, $\mathbb{E}(\varepsilon_i)$, denotes the average heating demand of buildings while still operating fossil equipment which is equivalent to aggregate emissions if we normalize the mass of buildings to one.[14] For a price elasticity of $e_D = -0.2$ (ibid.), at a relative stringency of climate policy equal to $x = 0.59$ (see Section 6), an interest rate of 2% and assuming that annual emissions in $t = 0$ amount to 9.5% of the remaining carbon budget, i.e. $\frac{E}{\mathbb{E}(\varepsilon_i)} = 0.095$ (based on emissions targets for the building sector in the German Climate change Act), the deadweight loss due the inefficiently low incentives to save energy is roughly bounded at less than 1% of the total

[14] For a detailed derivation see Supplementary Appendix B.4.

cost of the transformation.

5.4 Further extensions

Damage domains. Climate damages may not accrue as a uniform shock on total factor productivity. Climate damages can also affect individual utility via alternative channels, e.g., altruistic preferences for future generations, impacts on the quality of amenities or varied effects on prices contingent on how sensitive production of a specific product is to climate change. In Supplementary Appendix B.5, we consider a model where, instead of uniform damages to output, $\mu(y)$, we allow for heterogeneous climate-related effects on prices for differentiated products and consider expenditures on an emission offsetting technology to account for the willingness to pay for amenities and altruistic preferences. We show that the results in Lemma 6 continue to hold if the income tax adjustment, $\Delta\Gamma$, is adapted to skim all benefits of avoided climate damages in this scenario. Climate damages can also accrue on capital and wealth by depreciating the property value of real estate depending on its susceptibility to warming-induced extreme weather (e.g., flooding and storms). Assuming mandatory insurance against extreme weather, mitigation benefits would be monetized in the form of reduced insurance premiums. Consequently, a tax on insurance premiums may be suitable to skim the benefits of climate policy.

Grandfathering. In Section 4.1, we argue that our proposal of a category-based transfer is superior to uniform transfers and grandfathering of emission allowances as it is able to perfectly mimic first-best compensation. While it is straightforward to see that a uniform transfer cannot account for the heterogeneity in the building sector, the same observation is less obvious in the case of grandfathering. From a static perspective, grandfathering based on historic emissions is also capable of perfectly compensating the carbon price for each household as we abstract from the intensive use margin (i.e., heating demand is predetermined and fixed by ε_i). However, grandfathering becomes gradually less accurate over time if we consider that households dynamically switch to the carbon-neutral technology. To avoid overcompensating households after switching at the expense of households still paying the carbon price, the reference period used for calculating the grandfathered transfer has to be reset regularly after the mechanism has been introduced. If households anticipate this, they have a strategic incentive to delay switching to also benefit from the transfer after the next reset. In Supplementary Appendix B.6, we show that grandfathering cannot achieve exact compensation unless the switching cost, C_i, is perfectly correlated with energy intensity, ε_i, while also requiring a higher carbon price to account for strategic inefficiencies.

Separable retrofitting investments. As a simplifying assumption in (2), retrofitting investments always take place at the time when the household switches its heating technology. The results for our main specification with $g = r$ can be shown to extend to the case where investments in insulation are separable from switching heating systems and endogenously timed. In this case, the same ad-valorem subsidy, β^*, should apply to any energy efficiency-enhancing investment regardless of whether it coincides with s_i or occurs independently.

Credit cost. The spectrum of subsidizable expenditures should be expanded if cost heterogeneity is also driven by credit ratings. If a household has to take out a loan to finance retrofitting at the individually rational switching time, the interest rate may be subject to socio-economic or building-related characteristics. In order to offset this heterogeneity, interest costs have to be treated as an additional cost component in (2) and subsidized at the rate, β^*. In practice, this might require an opt-in subsidy scheme to account for the household-specific risk premium. Alternatively, the regulator could offer retrofitting loans at the fixed (risk-free) market rate to all households, thereby, indirectly subsidizing households which would have to pay a higher interest rate to private lenders.

Inconvenience cost. Homeowners and tenants experience inconvenience costs related to retrofitting (e.g., opportunity and information costs, noise and dirt). If these costs are proportional to the monetary cost of retrofitting,[15] our results continue to hold under a slightly different interpretation. The ad-valorem subsidy derived above now amounts to the *effective* subsidy rate which has to increase by the proportion of inconvenience costs to arrive at the same allocation. E.g., if the optimal subsidy rate in (10) equals 10% and inconvenience costs amount to 50% of the monetary cost, then a *nominal* subsidy rate of 15% is required.

Income-tax adjustments without Pareto improvements. Throughout the paper we derive a policy package ensuring (weak) Pareto-improvements. On the one hand, from a political economy perspective, Pareto-improving policies should not face opposition from any political camp, thus, suffer much smaller acceptance penalties than current climate policy approaches. Additionally, even if a compensation scheme perfectly offsets the vertical inequality, horizontal heterogeneity in transformation costs remains, acting as a random shock within income groups. This shock is arbitrarily distributed, depending on each household's existing building stock, such that welfare is reduced by the size of the risk premium households are willing to pay in order to eliminate this arbitrary shock and restore horizontal equity. The weakly pareto-improving combination of carbon pricing and ad-valorem subsidies derived above increases welfare by the cumulative amount of these risk premiums. On the other hand, the Pareto perspective is also expedient from a methodological approach, enabling us to derive incentive-compatible and distribution neutral policies without having to make any implicit or explicit normative assumptions about welfare and distributional preferences. However, by perfectly compensating households in the first step, our approach is flexible and readily adaptable to account for other normative concepts. As the category-based transfer and ad-valorem subsidies transfer the aggregate cost of decarbonizing the building stock to the government budget, all distributional effects are then relegated to how the deficit, D, is funded.

6 Calibrated model

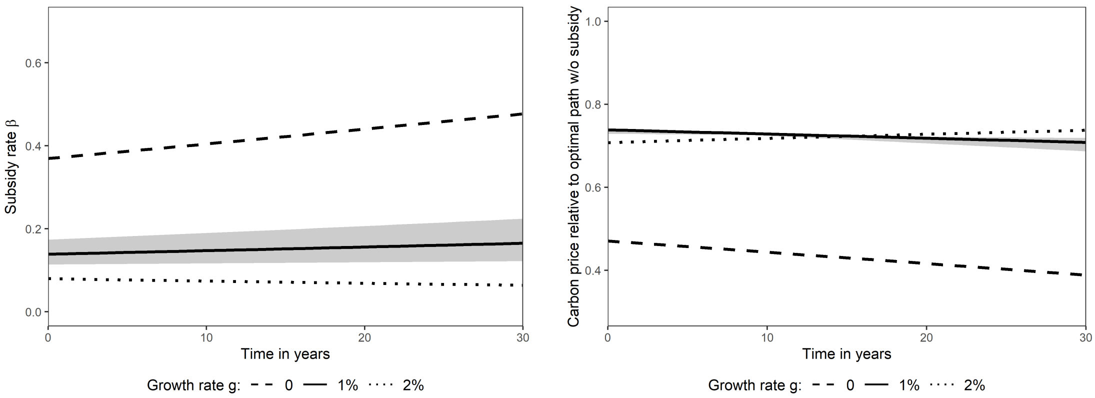
6.1 Optimal ad-valorem subsidy and carbon price paths

To simulate the optimal paths of the ad-valorem subsidy and the carbon price, respectively, we have to pin down the initial fossil fuel price, its growth rate, the interest rate and the carbon price required in the absence of additional subsidies.

The initial carbon price, $\tau_0(0,E)$, is calibrated to EUR 182 per ton of CO_2e in 2024 and increasing to EUR 283 per ton in 2054 (using 2020 real prices and exchange rates) to match the

[15] This assumption appears plausible as more invasive investments, e.g., refitting piping or floor heating, tend to be substantially more expensive, time consuming and impairing than merely placing a heat pump in the front yard.

Social Cost of Carbon (SCC) reported in EPA (2023) for a 2% discount rate. We set the interest rate equal to the discount rate, i.e., $r = 0.02$. As the price hikes during the energy price crisis appear to be temporary shocks with energy prices already returning to their pre-crisis levels, we employ the average consumer price for natural gas in Germany between 2008 and 2022 as the initial fossil price, i.e., $p_0^F = 6.2$ EUR cts per kWh (Destatis 2023b). Given the specific carbon intensity of natural gas, $\tau_0(0,E) = 182$ EUR leads to a relative markup on the market price of $x = 0.59$. To determine the growth rate of the fossil price, we first annualize the US commodity price increases in Jacks (2019) such that the price for natural gas has grown until 2015 at an average rate of 2.4% since 1950 and 1.5% since 1975.[16] Analogously, Jordà et al. (2019) report average rates of returns on German public bonds until 2015 of 3.7% since 1950 and 4.2% since 1980. Thus, the historical growth rate of the natural gas price amounts to 65% and 36%, respectively, of the real interest rate on government bonds, with 50.5% the average between these two numbers. To examine a set of analytically instructive and empirically relevant growth rates, we choose $g = \alpha r$, where $\alpha \in \{0, 0.5, 1\}$.



(a) Ad-varlorem subsidy (b) Carbon price

Figure 5: Optimal path of the ad-valorem subsidy rate and the corresponding carbon price (relative to optimal carbon price absent subsidies) for different realizations of fossil price growth rate for $r = 2\%$. Line types reflect the set of values for α whereas the shaded area depicts the range of g between historic plausible values of 36% to 65% of r (Source: own depiction).

Figure 5a shows that both the initial level of the perfectly offsetting subsidy rate and its time path crucially depend on the growth rate of the fossil fuel price.[17] If the fossil price is stagnant ($g = 0$), the BAU incentive to decarbonize is weak or non-existent such that climate policy has to achieve substantially earlier switching times. To maintain exact compensation, an initial subsidy rate of almost 40% is required. With higher fossil price growth rates, it becomes individually rational to switch earlier already in the BAU such that less stringent climate policy intervention is necessary to reach a given emission target. As a result, the initial subsidy rate also decreases as less compensation is required to ensure a (weak) Pareto improvement. For g=0.01, we find an initial optimal subsidy rate of approximately 14%.

Additionally, we find that whether the optimal subsidy rate increases over time depends on g relative to the average growth rate of the SCC, $h = 0.015$. If the fossil price growth rate falls

[16] We employ US commodity prices, assuming that, after decoupling from Russian supply in the mid-run, international LNG and US price developments will be indicative for price growth rates in Germany.

[17] In Supplementary Appendix B.7, we provide functional forms as well as numerical results for alternative specifications of the optimal carbon price path, including cost-effectiveness estimates. The results in this section are also reproducible using the code provided above.

short of h, the BAU incentive to decarbonize, p_t^F, increases more slowly than the policy-induced incentive, $\tau_t(0,E)$. Thus, the optimal subsidy rate has to increase over time to compensate for the increasing gap between market price and the social cost of carbon in order to maintain exact compensation. In contrast, if the market price of fossil fuels grows faster than the social cost of carbon, the BAU incentive becomes relatively more relevant over time and the ad-valorem subsidy has to compensate for a decreasing gap between utility in the BAU and under climate policy. Hence, we observe the strongest increase of more than 10pp over 30 periods for a constant fossil price. For $h > g = 0.01$, the effect on the optimal ad valorem subsidy is more subdued, yet, still increasing by 2.6pp. If fossil prices increase faster than the SCC, as in the case of $g = 0.02$, the subsidy rate decreases by 1.6pp over 30 periods.

Figure 5b depicts how the optimal carbon price is adjusted relative to the SCC (which coincides with the optimal carbon price absent subsidies) such that the combined policy still attains the same climate target, E. For a stagnant fossil price ($g = 0$), the regulator has to employ a high subsidy rate to ensure exact compensation, allowing for a substantially lower carbon price, decreasing from 47% of the SCC in the initial period to 39% over the course of 30 years. For positive fossil price growth rates ($g = 0.01$ and $g = 0.02$) closer to the growth rate of the SCC, lower ad-valorem subsidies are required for exact compensation, implying that the carbon price can only be decreased to around 70% of the SCC fluctuating by 3pp between the first and the final period. In the case of $g = 0.01$, which mandates increasing subsidy rates, the complementary carbon price decreases over time. The opposite effect occurs for $g = 0.02$.

How the optimal carbon price reacts to the ad-valorem subsidy is governed by both a *level* and a *slope* effect. The former effect implies that, at a given time t, a positive subsidy rate provides an incentive to switch, allowing for a lower carbon price, i.e., $\partial \tau_t(\beta_t, E)/\partial \beta_t < 0$. The latter effect captures that an increasing subsidy rate creates an opposing incentive to postpone switching to benefit from a higher subsidy later on (analogously, a decreasing subsidy path induces preponing) which has to be compensated for by increasing the carbon price, i.e., $\partial \tau_t(\beta_t, E)/\partial \dot{\beta}_t > 0$. If the subsidy path is steep enough, the slope effect can outweigh the level effect and even call for a carbon price exceeding the SCC to prevent inefficient postponement.

6.2 Income tax adjustments

Next, we compute by how much the income tax can be adjusted in order to skim all benefits from climate policy. As stated by Lemma 6, this requires knowledge of how damages are distributed between income groups as well as the damage function, μ, both under the BAU emissions path, e^{BAU}, and the emissions path, e, consistent with the cumulative emission target. For illustrative purposes, we assume that the climate target is set such that aggregate avoided climate damages are equal to 5% of the total income tax base in Germany. The respective benefits are assigned to income levels based on the income elasticity of damages, ϕ. We approximate the pre-tax income distribution by a log-normal distribution that matches the pre-tax Gini coefficient for market incomes without transfers of 0.50 (OECD 2024). The status-quo income tax is calibrated to match the German tax schedule with a tax-free allowance of EUR 11,000 and a tax rate progressively increasing from 14 to 45%.

Regarding the choice of the income elasticity of climate damages, ϕ, there is little empirical evidence so far. A common choice in integrated assessment models is to assume climate damages to be proportional to output (Nordhaus 1993; Golosov et al. 2014; Dietz and Venmans 2019), implying an income elasticity of one. Heterogeneity with respect to climate damages can result from different baseline exposures or different vulnerabilities (Hsiang et al. 2019). Heterogeneity in baseline exposure stems from the fact that individuals living in already adverse climate conditions might suffer overproportionally from further climatic changes due to the convexity of the damage function. This type of heterogeneity explains why many poor countries are expected to suffer strongly from climate damages as they already live in hot regions (Kalkuhl and Wenz 2020; Burke et al. 2015). We abstract from this type of heterogeneity in our model but could incorporate it by allowing for place-specific adjustments in the income tax schedule (see also the analogue discussion of damage domains in Section 5.4). Alternatively, heterogeneity in vulnerability may be a result of income-dependent preferences for climate-sensitive goods, services and amenities (e.g., food, health) or liquidity constraints (e.g., low-income households do not get a loan to retrofit their homes). Gilli et al. (2024) estimate within-country income elasticities of climate damages of 0.64 by evaluating income-climate response functions. However, this estimate could also capture within-country heterogeneity in exposure to some extent. Additionally, this estimate disregards adaptation costs, wealth impacts (e.g., on land values) as well as non-market impacts, particularly on amenities, mortality or ecosystem services – which is valued higher as income increases. For example, using a spatial equilibrium model with endogenous relocation choices, Wrenn (2024) finds that high-skilled individuals have a substantially higher willingness-to-pay to reduce disaster and extreme weather risk than low-skilled individuals. His figures on the willingness-to-pay versus the average income of the respective skill-level suggest an income elasticity of 2.9. Based on these considerations, we compute the income tax adjustment for income elasticities equal to $\phi \in \{0.1, 0.5, 1, 2\}$ with $\phi = 1$ being our central estimate and $\phi = 2$ a reasonable high estimate.

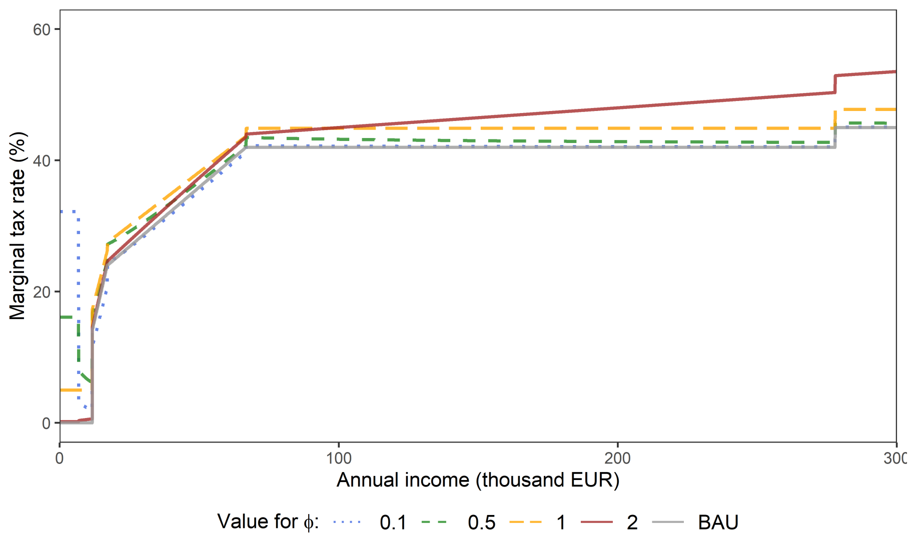


Figure 6: Income tax adjustment recapturing the benefits of avoiding climate damages in the amount of 5% of pre-tax incomes (Source: own depiction)

The results are reported in Figure 6. We find that the tax adjustment skimming all benefits of avoided climate damages is largely governed by the income elasticity of damages. As already noted in the discussion of Lemma 6, the benefit is skimmed via a uniform lump-sum tax if damages are very inelastic and largely uncorrelated with the income level. In the numerical simulation ($\phi = 0.1$), this is represented by a tax rate of more than 30% on initial incomes while leaving the existing tax schedule virtually untouched.[18] In contrast, if damages are proportional to income ($\phi = 1$), it is optimal to increase the marginal tax rate by 3pp on average at all income levels.[19] If climate damages increase with income ($\phi = 2$), it is optimal to implement a progressive income tax increase. This becomes particularly apparent for pre-tax incomes between EUR 67,000 and EUR 278,000 where the current German system amounts to a flat tax.[20]

7 Conclusion

In this paper, we develop a comprehensive policy package that achieves efficient mitigation in the building sector while ensuring a Pareto-improvement at minimal informational requirements. The regulator does neither need to know individual abatement costs nor individual switching times nor the respective distributions. Full compensation of carbon pricing is only contingent on knowledge about historical carbon intensities and the *determinants* of cost heterogeneity (or switching time heterogeneity). This information is sufficient for category-based transfers that exactly offset the direct costs of carbon pricing. Equivalently, carbon pricing and category-based transfers can be implemented as a single instrument via a bonus-malus system. The latter has a strong political appeal due to lower cost salience. Compensating heterogeneous dynamic abatement costs – i.e., the investment and operating costs of switching to a carbon-neutral heating system – through ad-valorem subsidies is informationally even less demanding: optimal subsidies can be calculated using discount rates and fossil fuel price paths.

Since the revenues from carbon pricing are depleted by the category-based transfer, the uniform *ad-valorem* subsidy generates a deficit which has to be tax-funded, e.g., via an income tax increase. When climate policy passes the cost-benefit test, the income tax schedule can be adjusted in a distribution-neutral way to skim the benefits of avoided climate damages. As the fiscal cost of subsidies and the benefit of avoided climate damages do not occur simultaneous, temporary debt funding ensures that no generation is worse-off. Thus, revenues from the income tax adjustment are first used to service public debt. The remaining surplus can then be redistributed to attain a strict Pareto-improvement compared to the reference scenario without climate policy.

Our analytical framework has some noteworthy limitations. First, we disregard the intensive use margin of adjusting the thermostat and ventilation behaviour. Price elasticities of heating demand suggest that this effect is of minor importance in the long run. Based on empirical evidence on demand adjustments, dead-weight losses are likely less then 1% of the overall costs of climate policy. Second, our model implicitly assumes that retrofitting investments affect utility exclusively via their impact on heating costs. While this indirect effect may be justified in case of injecting insulation into walls or replacing heat piping, other (visible) retrofitting activities such as replacing windows or facades can also have a direct effect on utility, e.g., by increasing the visual appeal of the building. The deadweight loss associated with over-subsidizing visual improvements is quadratic in the subsidy rate and could be quantified empirically based on the respective price elasticities. Hence, ad-valorem subsidies should be targeted carefully and exclude items

[18] To depict a lump-sum tax in the marginal income tax schedule, we assume that all households earn an income equal to at least the social security minimum of EUR 6,500 such that tax adjustments below this level do not affect labor supply. The lump-sum is collected via a uniform tax rate on this income range.

[19] The German income tax schedule consists of ‘progression tiers’ leading to discontinuities in the marginal tax rate, e.g., at EUR 17,000 and EUR 67,000 pre-tax income. This also explains the discontinuities in Figure 6.

[20] As climate damages, μ, themselves are a function of income, y, (22) indicates that it is, in principle, possible to obtain a regressive optimal income tax adjustment for very large values of ϕ.

that provide non-energy related co-benfits to homeowners. Third, the advantage of our mechanism over alternative refund options critically depends on the regulator's ability to sort households into appropriate categories. The larger the heterogeneity within one category, the less precise the compensation for the carbon price burden. This does not only relate to *status-quo* fossil fuel consumption, but also the individually rational time of switching. If households within one bin do not all switch at approximately the same time, our transfer mechanism will also dynamically incur over- and under-compensation. Imprecise categories may also increase the political pressure to reset the mechanism which then gives rise to inefficient strategic incentives. In practice, this limitation could prove the most challenging to overcome.

Besides these limitations, our mechanism provides a way forward to overcome one of the key political obstacles in implementing carbon pricing. While we have focused here on Pareto-improving reforms, our mechanism is capable of redistributing the costs of climate policy through the income tax system in a very flexible way. By eliminating horizontal costs completely and shifting abatement costs to the government budget at no efficiency loss, it provides the foundations for a transparent public debate on how to distribute the costs of climate policy. This allows for compromise between different political camps aiming at distributing costs in a progressive way or aiming at preserving the status-quo (weak Pareto improvement). Without elimination of horizontal cost heterogeneity, climate policy will always create losers that perceive some form of arbitrariness from bearing these costs.

Future research can build on and enrich our formal findings by simulating household-specific optimal switching times in a bottom-up approach. This would generate two main benefits. First, a simulation based on representative household data enables us to identify household and building-related characteristics decisive for optimal switching, information required for sorting buildings into suitable bins for the category-based rebate of carbon pricing revenues. Second, the simulation can be adjusted for second-best policy paths and market inefficiencies observed in reality to determine the impacts on renewable technology adoption and compliance with climate targets. Most importantly, a simulation model should incorporate credit and income constraints which prevent low-income households from optimal switching. We provide evidence for the substantial horizontal heterogeneity in the transformation cost. This horizontal effect can also dynamically exacerbate vertical income inequality. While richer households adopt technologies at the individually rational time, low-income households may have to hold off on investments and, thus, experience increasing under-compensation of the carbon price burden over time, effectively implying an income transfer from poor households, subject to a fossil lock-in, to high incomes.

Replication. Code and data for replication are available at https://doi.org/10.5281/zenodo.14933610.

References

Auffhammer, Maximilian and Edward Rubin (2018). *Natural Gas Price Elasticities and Optimal Cost Recovery Under Consumer Heterogeneity: Evidence from 300 million natural gas bills*. Working Paper 24295. National Bureau of Economic Research.

Becker, Martin and Peter Knoll (2011). *Energieeffizienz durch Gebäudeautomation mit Bezug zur DIN V 18599 und DIN EN 15232*. Kurzzusammenfassung der Studie. Hochschule Biberach.

Bergquist, Magnus, Andreas Nilsson, Niklas Harring, and Sverker C. Jagers (2022). "Meta-analyses of fifteen determinants of public opinion about climate change taxes and laws". In: *Nature Climate Change* 12.3, pp. 235–240.

Berry, Audrey (2019). "The distributional effects of a carbon tax and its impact on fuel poverty: A microsimulation study in the French context". In: *Energy Policy* 124, pp. 81–94.

Blanchard, Olivier J. (1985). "Debt, deficits, and finite horizons". In: *Journal of Political Economy* 93.2, pp. 223–247.

Buchsbaum, Jesse (2023). *Are consumers more responsive to prices in the long run? Evidence from electricity markets*. Working Paper.

Burke, Marshall, Solomon M. Hsiang, and Edward Miguel (2015). "Global non-linear effect of temperature on economic production". In: *Nature* 527 (7577), pp. 235–239.

Carattini, Stefano, Andrea Baranzini, Philippe Thalmann, Frédéric Varone, and Frank Vöhringer (2017). "Green Taxes in a Post-Paris World: Are Millions of Nays Inevitable?" In: *Environmental and Resource Economics* 68 (1), pp. 97–128.

Creutzig, Felix, Jérôme Hilaire, Gregory Nemet, Finn Müller-Hansen, and Jan C. Minx (2023). "Technological innovation enables low cost climate change mitigation". In: *Energy Research & Social Science* 105, p. 103276.

Destatis (2023a). *Mikrozensus 2022*. https://www.destatis.de/DE/Methoden/Qualitaet/Qualitaetsberichte/Bevoelkerung/mikrozensus-2022.pdf?__blob=publicationFile. accessed: 16th July 2024.

— (2023b). *Preise - Daten zur Energiepreisentwicklung - Lange Reihen von Januar 2005 bis Januar 2023*. https://www.destatis.de/DE/Themen/Wirtschaft/Preise/Publikationen/Energiepreise/energiepreisentwicklung-pdf-5619001.pdf?__blob=publicationFile. accessed: 16th July 2024.

Diamond, Peter A. (1965). "National debt in a neoclassical growth model". In: *American Economic Review* 55.5, pp. 1126–1150.

Dietz, Simon and Frank Venmans (2019). "Cumulative carbon emissions and economic policy: in search of general principles". In: *Journal of Environmental Economics and Management* 96, pp. 108–129.

Douenne, Thomas and Adrien Fabre (2022). "Yellow Vests, Pessimistic Beliefs, and Carbon Tax Aversion". In: *American Economic Journal: Economic Policy* 14.1, pp. 81–110.

Edenhofer, Ottmar, Matthias Kalkuhl, and Christina Roolfs (2021). "Carbon Pricing and Revenue Recycling: An Overview of Vertical and Horizontal Equity Effects for Germany". In: *CESifo Forum* 22 (5), pp. 10–14.

Energy Systems Catapult (2023). *Electrification of Heat Demonstration Project: Heat Pump Performance Cleansed Data, 2020-2022*. [data collection]. UK Data Service. SN: 9050, DOI:

http://doi.org/10.5255/UKDA-SN-9050-1. Contains public sector information licensed under the Open Government Licence v2.0.

Environmental Protection Agency (2023). *Report on the social cost of Greenhouse gases: Estimates incorportating recent scientific advances*. https://www.epa.gov/environmental-economics/scghg [Last accessed 29 September 2024].

Fischer, Carolyn and William A. Pizer (2019). "Horizontal Equity Effects in Energy Regulation". In: *Journal of the Association of Environmental and Resource Economists* 6.S1, S209–S237.

Frondel, Manuel, Andreas Gerster, Kathrin Kaestner, Michael Pahle, Antonia Schwarz, Puja Singhal, and Stephan Sommer (2023). *Wärme- und Wohnen-Panel - Welle 1. German Heating and Housing Panel*. Version 1. RWI – Leibniz Institute for Economic Research. Dataset. https://doi.org/10.7807/ghhp:building:v1.

Galvin, Ray (2024). "Deep energy efficiency renovation of Germany's residential buildings: is this as economically viable as Germany's policymakers and popular promoters often claim?" In: *Energy Efficiency* 17.5, p. 47.

Gilli, Martino, Matteo Calcaterra, Johannes Emmerling, and Francesco Granella (2024). "Climate change impacts on the within-country income distributions". In: *Journal of Environmental Economics and Management* 127, p. 103012.

Glanemann, Nicole, Sven N. Willner, and Anders Levermann (2020). "Paris Climate Agreement passes the cost-benefit test". In: *Nature Communications* 11.110.

Golosov, Mikhail, John Hassler, Per Krusell, and Aleh Tsyvinski (2014). "Optimal taxes on fossil fuel in general equilibrium". In: *Econometrica* 82.1, pp. 41–88.

Hänsel, Martin C., Moritz A. Drupp, Daniel J. A. Johansson, Frikk Nesje, Christian Azar, Mark C. Freeman, Ben Groom, and Thomas Sterner (2020). "Climate economics support for the UN climate targets". In: *Nature Climate Change* 10 (8), pp. 781–789.

Hänsel, Martin C., Max Franks, Matthias Kalkuhl, and Ottmar Edenhofer (2022). "Optimal carbon taxation and horizontal equity: A welfare-theoretic approach with application to German household data". In: *Journal of Environmental Economics and Management* 116, p. 102730.

Hondeborg, Dianne, Benedict Probst, Ivalin Petkov, and Christof Knoeri (2023). "The effectiveness of building retrofits under a subsidy scheme: Empirical evidence from Switzerland". In: *Energy Policy* 180, p. 113680.

Hsiang, Solomon, Paulina Oliva, and Reed Walker (2019). "The distribution of environmental damages". In: *Review of Environmental Economics and Policy* 13.1.

IPCC, The Intergovernmental Panel on Climate Change (2023). *Sixth Assessment Report*. https://www.ipcc.ch/assessment-report/ar6/ [Last accessed 28 December 2023].

Ito, Koichiro (2014). "Do consumers respond to marginal or average price? Evidence from nonlinear electricity pricing". In: *American Economic Review* 104.2, pp. 537–563.

Jacks, David S (2019). "From boom to bust: A typology of real commodity prices in the long run". In: *Cliometrica* 13.2, pp. 201–220.

Jacobs, Bas and Frederick van der Ploeg (2019). "Redistribution and pollution taxes with nonlinear Engel curves". In: *Journal of Environmental Economics and Management* 95, pp. 198–226.

Jordà, Òscar, Katharina Knoll, Dmitry Kuvshinov, Moritz Schularick, and Alan M Taylor (2019). "The rate of return on everything, 1870–2015". In: *The Quarterly Journal of Economics* 134.3, pp. 1225–1298.

Kalkuhl, Matthias and Leonie Wenz (2020). "The impact of climate conditions on economic production: Evidence from a global panel of regions". In: *Journal of Environmental Economics and Management* 103, p. 102360.

Kaplow, Louis (2012). "Optimal Control of Externalities in the Presence of Income Taxation". In: *International Economic Review* 53 (2), pp. 487–509.

Khorasanizadeh, H., G.A. Sheikhzadeh, A.A. Azemati, and B. Shirkavand Hadavand (2014). "Numerical study of air flow and heat transfer in a two-dimensional enclosure with floor heating". In: *Energy and Buildings* 78, pp. 98–104.

Klenert, David and Linus Mattauch (2016). "How to make a carbon tax reform progressive: The role of subsistence consumption". In: *Economics letters* 138, pp. 100–103.

Klenert, David, Linus Mattauch, Emmanuel Combet, Ottmar Edenhofer, Cameron Hepburn, Ryan Rafaty, and Nicholas Stern (2018). "Making carbon pricing work for citizens". In: *Nature Climate Change* 8.8, pp. 669–677.

Moon, Jin Woo and Seung-Hoon Han (2011). "Thermostat strategies impact on energy consumption in residential buildings". In: *Energy and Buildings* 43.2-3, pp. 338–346.

Myers, Erica (2019). "Are home buyers inattentive? Evidence from capitalization of energy costs". In: *American Economic Journal: Economic Policy* 11.2, pp. 165–188.

Nordhaus, William D. (1993). "Optimal Greenhouse-Gas Reductions and Tax Policy in the "DICE" Model". In: *The American Economic Review* 83 (2), pp. 313–317.

OECD (2024). *Income distribution database. OECD Data Explorer*. https://data-explorer.oecd.org/ [Last accessed 17 July 2024].

Pietzcker, Robert et al. (Nov. 2021). *Notwendige CO2-Preise zum Erreichen des europäischen Klimaziels 2030*. Ariadne Hintergrund. Kopernikus-Projekt Ariadne, Potsdam-Institut für Klimafolgenforschung (PIK).

Ruhnau, Oliver, Clemens Stiewe, Jarusch Muessel, and Lion Hirth (2023). "Natural gas savings in Germany during the 2022 energy crisis". In: *Nature Energy* 8 (6), pp. 621–628.

Sallee, James M (2019). *Pigou Creates Losers: On the Implausibility of Achieving Pareto Improvements from Efficiency-Enhancing Policies*. Working Paper 25831. National Bureau of Economic Research.

Simmons, George F (2016). *Differential equations with applications and historical notes*. CRC Press.

Sommer, Stephan, Linus Mattauch, and Michael Pahle (2022). "Supporting carbon taxes: The role of fairness". In: *Ecological Economics* 195, p. 107359.

Souleles, Nicholas S (1999). "The response of household consumption to income tax refunds". In: *American Economic Review* 89.4, pp. 947–958.

Statistische Ämter des Bundes und der Länder (2019). *Wohnen in Deutschland – Zusatzprogramm des Mikrozensus 2018, Tabelle 8*. https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fwww.destatis.de%2FDE%2FThemen%2FGesellschaft-Umwelt%2FWohnen%2FPublikationen%2FDownloads-Wohnen%2Fwohnen-in-deutschland-5122125189005.xlsx%3F__blob%3DpublicationFile&wdOrigin=BROWSELINK [Last accessed 02 January 2024].

Többen, Johannes, Peter-Paul Pichler, Ingram S Jaccard, Kurt Kratena, Daniel Moran, Heran Zheng, and Helga Weisz (2023). "Unequal carbon tax impacts on 38 million German households: assessing spatial and socio-economic hotspots”. In: *Environmental Research: Climate* 2.4, p. 045006.

van der Ploeg, Frederick, Armon Rezai, and Miguel Tovar Reanos (2024). “Public Finance Options for Recycling Carbon Tax Revenue in Germany”. In: *FinanzArchiv: Public Finance Analysis* 80.1, pp. 39–69.

Wrenn, Douglas H (2024). “The Effect of Natural Disasters and Extreme Weather on Household Location Choice and Economic Welfare”. In: *Journal of the Association of Environmental and Resource Economists* 11.5, pp. 1101–1134.

A Appendix: Proofs

A.1 Proof of Lemma 3

Household i's utility in (1) may be rewritten such that the optimization problem reads

$$\max_{s_i} u_i = -(1-\beta)\left[\int_{t=0}^{s_i}\left\{e^{-rt}\frac{p_t^F+\tau_t}{1-\beta}\varepsilon_i\right\}+e^{-rs_i}C_i\right]+T_i^\theta. \tag{31}$$

Hence, the first-order condition with regard to s_i

$$\frac{\partial u_i}{\partial s_i}=\frac{p_t^F+\tau_t}{1-\beta}\varepsilon_i-rC_i=0, \tag{32}$$

is affected by climate policy via

$$\frac{p_t^F+\tau_t}{1-\beta}. \tag{33}$$

The policy path in (33) leads always to the same allocation as the Hotelling price path, $\tilde{\tau}_t = \tau_0(0,E)e^{rt}$, if and only if:

$$\begin{aligned}\frac{p_t^F+\tau_t}{1-\beta} &= p_t^F+\tau_0(0,E)e^{rt},\\ \Rightarrow \tau_t &= (1-\beta)\tau_0(0,E)e^{rt}-\beta p_t^F,\end{aligned} \tag{34}$$

where we substitute $p_0^F e^{rt}$ for p_t^F to obtain (8). In turn, by substituting (8) into (32), the first-order condition may be solved for the individually optimal switching time

$$\begin{aligned}(1-\beta)e^{rs_i}(\tau_0(0,E)+p_0^F)\varepsilon_i &= (1-\beta)rC_i\\ \Rightarrow s_i &= \frac{1}{r}\log\left[\frac{rC_i}{(\tau_0(0,E)+p_0^F)\varepsilon_i}\right]\end{aligned}$$

A.2 Proof of Lemma 4

We want to choose the constant ad-valorem subsidy rate such that each household i obtains exactly the same utility, u_i^*, when optimizing under the climate policy package (τ_t,β) as in the BAU scenario, i.e., $u_i^* = u_i^{BAU}$ for all i. Consider that the category-based transfer in (7) exactly compensates for the carbon price in each period t such that

$$T_i^\theta = \int_{t=0}^{s_i} e^{-rt}\tau_t\varepsilon_i, \tag{35}$$

where T_i^θ denotes the discounted intergral of $T_{i,t}^\theta$ until s_i. Hence, (1) may be written as

$$u_i = -p_0^F s_i\varepsilon_i - e^{-rs_i}(1-\beta)C_i. \tag{36}$$

Substituting (9) for s_i in (36) yields

$$u_i^* = -\frac{p_0^F}{r}\log\left(\frac{rC_i}{\varepsilon_i\left[\tau_0(0,E)+p_0^F\right]}\right)\varepsilon_i-\frac{\varepsilon_i\left[\tau_0(0,E)+p_0^F\right]}{r}(1-\beta). \tag{37}$$

Since we assume that carbon prices are equal to zero in the BAU scenario, both the associated burden and the category-based transfer amount to zero in this case. Analogously, substituting (9) for s_i^{BAU} in (36) results in

$$u_i^{BAU} = -\frac{p_0^F}{r}\log\left(\frac{rC_i}{\varepsilon_i\left[\tau_0(0,E^{BAU})+p^F\right]}\right)\varepsilon_i - \frac{\varepsilon_i\left[\tau_0(0,E^{BAU})+p_0^F\right]}{r}(1-\beta^{BAU}). \tag{38}$$

Using $\tau_t^{BAU} = 0$ in (8), we obtain

$$\tau_0(0,E^{BAU}) = \frac{\beta^{BAU}}{1-\beta^{BAU}}p_0^F, \tag{39}$$

such that $\tau_t(0,E^{BAU})e^{rt}$ defines the counterfactual carbon price path which would achieve the same cumulative emission level, E^{BAU}, without additional subsidies as β^{BAU} does without additional carbon taxes. Equating (37) and (38) while substituting (39) yields

$$\log\left[(1-\beta^{BAU})\left(1+\frac{\tau_0(0,E)}{p_0^F}\right)\right] = (1-\beta)\left(1+\frac{\tau_0(0,E)}{p_0^F}\right)-1 \tag{40}$$

Which can be solved for β to receive (10). Now, by substituting β^* into (8), we obtain the complementary tax path, τ_t^*, which completes the proof of Lemma 4.

A.3 Proof of Lemma 5

The category-based transfer, T_i^θ, is designed such that it completely refunds the revenues from carbon pricing, i.e., these two components of the mechanism are revenue-neutral following from

$$\int_{i=1}^{N} T_i^\theta = \int_{i=1}^{N}\int_{t=0}^{s_i} e^{-rt}\tau_t\varepsilon_i. \tag{41}$$

Thus, the fiscal deficit, D, induced by the policy path (τ_t,β) is contingent on the additional cost of increasing the subsidy rate from β^{BAU} to β. Normalizing the mass of households, N, to 1, we can use the expectations operator, $\mathbb{E}$, to express the deficit in comparison to the BAU as

$$D = \beta\mathbb{E}\left\{e^{-rs_i}C_i)\right\} - \beta^{BAU}\mathbb{E}\left\{e^{-rs_i^{BAU}}C_i\right\}. \tag{42}$$

By substituting (9) for s_i and s_i^{BAU} as well as (10) and (39) into (42), we obtain (14). To see that D coincides with the aggregate induced mitigation cost, consider that D equals the intertemporal sum of subsidy payments to all households i. According to Lemma 1, the category-based transfer completely offsets the carbon price burden identified by term A in (5). Lemma 4 defines the subsidy such that it ensures exact compensation for each individual households implying that subsidy payments have to coincide with $(B+C)$ in (5), i.e., the induced mitigation cost. This completes the proof of Lemma 5.

A.4 Proof of Lemma 6

Consider individual i's bequest in the BAU

$$b_i^{BAU} = e^{rT_i}\Bigg\{ w_i(b_{-i}) + u_i(0,\beta^{BAU}) + \int_{t=0}^{T_i} e^{-rt}\Big(y_t - \mu_{t,e^{BAU}}(y_t) - c_t - \Gamma_t^{BAU}\left[y_t - \mu_{t,e^{BAU}}(y_t)\right]\Big)dt\Bigg\}. \tag{43}$$

Contrast this with the individual's bequest function under the climate policy package, (τ_t^*,β^*), and adjusted income tax schedule, Γ_t^*

$$b_i^* = e^{rT_i}\left\{ w_i(b_{-i}) + u_i(\tau_t^*,\beta^*) + \int_{t=0}^{T_i} e^{-rt}\Big(y_t - \mu_{t,e}(y_t) - c_t - \Gamma_t^*\left[y_t - \mu_{t,e}(y_t)\right]\Big)dt\right\}, \tag{44}$$

where the income tax schedule which completely captures the income effect of deviating from the BAU emissions path, e^{BAU}, to e is given by

$$\Gamma_t^*[y_t - \mu_{t,e}(y_t)] = \Gamma_t^{BAU}[y_t - \mu_{t,e^{BAU}}(y_t)] + \mu_{t,e^{BAU}}(y_t) - \mu_{t,e}(y_t). \tag{45}$$

The tax schedule in (45) depends on the unobserved counterfactual income before climate damages, y_t. We use this formulation of the income tax schedule as a shorthand in the proof which is identical to the observable tax path

$$\begin{aligned}\Gamma_t^*[y_t - \mu_{t,e}(y_t)] = \Gamma_t^{BAU}\left[Y_t\big(y_t - \mu_{t,e}(y_t),e\big)\right] - \mu_t\left[Y_t\big(y_t - \mu_{t,e}(y_t),e\big),e^{BAU}\right] \\ +\mu_t\left[Y_t\big(y_t - \mu_{t,e}(y_t),e\big),e^{BAU}\right] - \mu_t\left[Y_t\big(y_t - \mu_{t,e}(y_t),e\big),e\right],\end{aligned} \tag{46}$$

where Y_t denotes the unique income before climate damages fulfilling $Y_t(z,e) - \mu_{t,e}(Y_t) = z$. Substituting (17) into (45), defining m as the average marginal income tax rate on the interval $[y_t - \mu_{t,e^{BAU}}(y_t), y_t - \mu_{t,e}(y_t)]$ and taking the difference between Γ_t^* and Γ_t^{BAU} yields the income tax adjustment in (18).

Substituting (45) into (44), we obtain $b_i^* = b_i^{BAU}$. In particular, this implies that each other individual $j \neq i$ makes the same choices as in the BAU such that individual i's set of feasible choices (c_i, y_i, b_i) remains unaffected by the income tax adjustments. Thus, if a Subgame Perfect Nash Equilibrium exists in the BAU, the optimal decisions of each i remain unchanged and the solution in the case with an income tax adjustment remains a Subgame Perfect Nash Equilibrium.

As a result, the change in income tax revenues can be evaluated at the BAU income path, y_i^{BAU}. Since the adjusted income tax in (45) is designed to skim all monetary benefits in terms of averted climate damages stemming from reducing e^{BAU} to e, and Lemma 5 states that the deficit, D, amounts to the induced mitigation cost, the income tax reform achieves a budget surplus if it passes the cost-benefit test in (19). This completes the proof of Lemma 6.

B Supplementary Appendix

B.1 Welfare cost of imperfect carbon price compensation

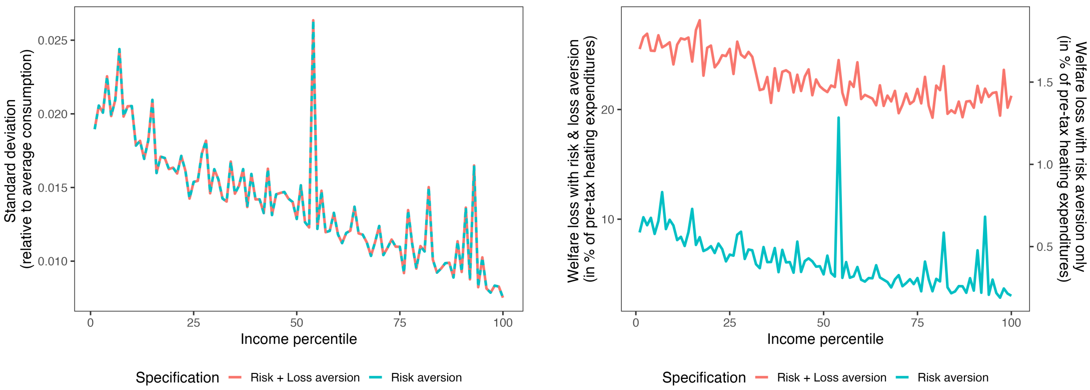
To approximate a transfer which perfectly compensates the vertical heterogeneity in the carbon price burden, we assume that a lump-sum transfer equal to the average burden per percentile of the income distribution is available. Hence, the net burden of carbon pricing on consumption, dc_i, amounts to

$$dc_i = -\tau \left[\varepsilon_i - \frac{1}{N_p} \sum_{i \in N_{p(j)}}^{N_{p(j)}} \varepsilon_i \right], \tag{47}$$

where N_p denotes the number of households per percentile j and ε_i gives the household-specific, heating-related carbon consumption. The net burden remains subject to horizontally heterogeneous carbon price costs within the income percentile. Normalizing status-quo consumption, c_i^0, and the respective marginal utility of consumption, $u'(c_i^0)$, to one each, we can use a second-order Taylor expansion to express the relative intra-percentile welfare loss as

$$RWL_j(\tau) = \frac{\eta}{2}\sigma_j^2 + (\lambda - 1)\Pr(d\tilde{c}_{i,j} < 0) \left[E(d\tilde{c}_{i,j}) + \frac{\eta}{2}\tilde{\sigma}_j^2 \right], \tag{48}$$

where σ_j is the standard deviation of dc within percentile j relative to average consumption in the percentile, $\bar{X}_j$. The parameters η and $\lambda \geq 1$ denote the degrees risk and loss aversion, respectively. If $\lambda = 1$, the second term in (48) cancels out and we only consider the effect of risk aversion on the welfare loss caused by imperfect compensation along the horizontal dimension. Alternatively, for $\lambda > 1$, the effects on the subset of households experiencing a loss in consumption (i.e., $dc_i < 0$, with the respective variables denoted by a tilde) have a more pronounced effect on the welfare evaluation. We use representative German household data (sample survey on income and consumption, EVS 2018) to estimate the consumption effect of a carbon price of $\tau = 182$ (the SCC for 2024 in 2020 prices) and calibrate the model parameters to $\eta = 1$ and $\lambda = 2$.



(a) Standard deviation

(b) Relative welfare loss

Figure 7: Remaining heterogeneity in the carbon price burden after percentile-based lump-sum rebate (Source: own calculations based on EVS 2018)

Figure 7a shows that the standard deviation of the consumption change, dc, relative to average consumption within the percentile is decreasing in income. This effect is predominantly driven by the fact that higher incomes spend a smaller share of their total consumption on energy and, by extension, the carbon price. Translating the standard deviation into welfare losses, Figure 7b reveals the sizable effect of loss aversion on the outcome. If only risk aversion is considered, the relative welfare loss ranges between 0.2 and 1.3% (averaging at 0.4%) of pre-tax heating expenditures in the status quo. With a loss aversion parameter of $\lambda = 2$, the welfare loss increases to approximately 25% of pre-tax heating expenditures in the lower income percentiles and remains above 19% in the highest percentile.

To compute the aggregate welfare loss across all percentiles for varying carbon price levels as depicted in Figure 3a, we weigh the relative welfare loss in (48) by the average consumption expenditures per percentile (in 2020 prices) and the number of households per percentile to obtain

$$WL = N_p \sum_{j}^{100} RWL_j(\tau)\bar{X}_j. \tag{49}$$

B.2 Optimal policy under diverging fossil price growth rate and interest rate

First, note that if the fossil price grows at the constant rate $g > 0$, it follows that $p_t^F = p_0^F e^{gt}$. For the Hotelling carbon price without subsidies, the first-order condition of (1) with regard to s_i can we written as

$$\frac{du_i}{ds_i} = e^{-rs_i}\left(p_0^F e^{gs_i} + \tau_0(0,E)e^{rs_i}\right)\varepsilon_i\left[\frac{r}{p_0^F e^{gs_i} + \tau_0(0,E)e^{rs_i}}\frac{C_i}{\varepsilon_i} - 1\right] = 0. \tag{50}$$

For a policy package consisting of the adjusted Hotelling carbon price, τ_t, as defined in (8) and a time-variant subsidy path, β_t, the first-order condition reads

$$\frac{du_i}{ds_i} = e^{-rs_i}(p_0^F e^{gs_i} + \tau_{s_i})\varepsilon_i\left[\frac{\dot{\beta}_{s_i} + (1-\beta_{s_i})r}{p_0^F e^{gs_i} + \tau_{s_i}}\frac{C_i}{\varepsilon_i} - 1\right], \tag{51}$$

where $\dot{\beta}_t$ denotes the time derivative of β_t. Thus, for any given C_i/ε_i, each household's individually efficient choice of s_i under the policy package, (τ_t, β_t), coincides with the optimal decision under the pure Hotelling carbon price, $\tau_0(0,E)e^{rt}$, if

$$\frac{\dot{\beta}_{s_i} + (1-\beta_{s_i})r}{p_{s_i}^F + \tau_{s_i}} = \frac{r}{p_{s_i}^F + \tau_0(0,E)e^{rs_i}}, \tag{52}$$

implying that the policy package also attains E at least costs as s_i is efficient for all i. Assuming $\beta^{BAU} = 0$ for the sake of notational brevity, (51) can be solved for

$$s_i^{BAU} = \frac{1}{g}\log\left(\frac{C_i r}{\varepsilon_i p_0^F}\right). \tag{53}$$

By use of the envelope theorem, the partial derivative of (1) with regard to C_i at the individually efficient s_i yields

$$\frac{d}{dC_i}(\sup_{s_i} u_i) = -(1-\beta_{s_i})e^{-rs_i} + \frac{dT_i^\theta}{dC_i},$$

where we also take the derivative of T_i^θ with regard to C_i as the category-based transfer depends on the C_i, since this co-determines the category of house i. Since the transfer perfectly offsets the carbon price burden, we substitute $T_i^\theta = e^{-rs_i}\tau_{s_i}\varepsilon_i$. For compensation to be exact, $\frac{d}{dC_i}(\sup_{s_i} u_i -$

$u_{i,BAU}) = 0$ has to be fulfilled for all C_i and ε_i which, by use of (53), equals

$$-(1-\beta_{s_i})e^{-rs_i} + \frac{ds_i}{dC_i}\tau_{s_i}\varepsilon_i e^{-rs_i} + \left(\frac{\varepsilon_i p_0^F}{C_i r}\right)^{\frac{r}{g}} = 0. \tag{54}$$

Differentiating (51) with regard to C_i yields

$$\frac{1}{\varepsilon_i} = \left(\frac{g}{r}p_0^F e^{gs_i} + \tau_0(0,E)e^{rs_i}\right)\frac{ds_i}{dC_i}, \tag{55}$$

which may be substituted into (54) to solve for

$$\tau_{s_i} = \left[(1-\beta_{s_i}) - \left(\frac{\varepsilon_i p_0^F}{C_i r}\right)^{\frac{r}{g}} e^{rs_i}\right]\left(\frac{g}{r}p_0^F e^{gs_i} + \tau_0(0,E)e^{rs_i}\right). \tag{56}$$

Now, we can eliminate the tax rate in efficiency condition (52) to get

$$\frac{\dot{\beta}_{s_i}}{r} + (1-\beta_{s_i}) = \frac{p_{s_i}^F + \left[(1-\beta s_i) - \left(\frac{\varepsilon_i p_0^F}{C_i r}\right)^{\frac{r}{g}} e^{rs_i}\right]\left(\frac{g}{r}p_0^F e^{gs_i} + \tau_0(0,E)e^{rs_i}\right)}{p_{s_i}^F + \tau_0(0,E)e^{rs_i}}. \tag{57}$$

Using (50), this expression is rearranged to obtain the differential equation

$$\beta_{s_i} - (1+xe^{(r-g)s_i})\frac{\dot{\beta}_{s_i}}{r-g} = \left(\frac{1}{1+xe^{(r-g)s_i}}\right)^{\frac{r}{g}}\left(\frac{g}{r-g} + \frac{r}{r-g}xe^{(r-g)s_i}\right) - \frac{g}{r-g}, \tag{58}$$

where $x = \tau_0(0,E)/p_0^F$. To identify an initial condition for β_0, consider a household which is induced by climate policy to switch in $t = 0$ without being in a corner solution. Utlity equals

$$u_i^* = -(1-\beta_0)C_i. \tag{59}$$

By employing (53) and $\beta^{BAU} = 0$, we rewrite utility in the BAU as

$$u_i^{BAU} = \frac{1}{r-g}\left[\left(\frac{\varepsilon_i p_0^F}{C_i r}\right)^{\frac{r-g}{g}} - 1\right]p_0^F\varepsilon_i - \left(\frac{\varepsilon_i p_0^F}{C_i r}\right)^{\frac{r}{g}} C_i, \tag{60}$$

where we use $\frac{\varepsilon_i p_0^F}{C_i r} = \frac{p_0^F}{p_0^F + \tau_0(0,E)} = \frac{1}{1+x}$ and $\varepsilon_i = \frac{rC_i}{p_0^F + \tau_0(0,E)}$ from (51) to further simplify to

$$u_i^{BAU} = -\left[\frac{r}{r-g}\frac{1}{1+x} - \frac{g}{r-g}\left(\frac{1}{1+x}\right)^{\frac{r}{g}}\right]C_i. \tag{61}$$

Exact compensation of household i choosing a (weakly) interior solution $s_i^* = 0$ is achieved when $u_i^* = u_i^{BAU}$ such that by equating (59) and (61), we can solve for β_0

$$\beta_0 = 1 - \frac{r}{r-g}\frac{1}{1+x} + \frac{g}{r-g}\left(\frac{1}{1+x}\right)^{\frac{r}{g}}. \tag{62}$$

Importantly, this condition neither involves C_i nor ε_i. Solving the differential equation in (58) for β_{s_i} with the initial condition (62) yields

$$\beta_{s_i} = \frac{g\left(e^{s_i(g-r)}+x\right)\left[\left(\frac{1}{xe^{s_i(r-g)}+1}\right)^{r/g}-1\right]+rx}{(r-g)\left(e^{(g-r)s_i}+x\right)}, \tag{63}$$

which equals (24) if we set $s_i = t$. To obtain the corresponding tax rate in (25), we substitute (24) into efficiency condition (52) and solve for τ_t.

B.3 Generalization with minimal assumptions about the fossil price path

If we lift the requirement of explicit analytical solutions for τ_t and β, the results in Section 5.1 may be generalized even further. In fact, it is sufficient that $rC_i > [p_0^F + \tau_0(0,E)]\varepsilon_i$ for all i and the term

$$\frac{\frac{d}{ds_i}\left(p_{s_i}^F e^{-rs_i}\right)}{p_{s_i}^F e^{-rs_i} + \tau_0(0,E)}, \tag{64}$$

is bounded above such that a unique policy package (τ_s, β_s) exists which achieves the emission target E at minimal cost and exact compensation. To show this, first solve the efficiency condition (52) for

$$p_{s_i}^F + \tau_{s_i} = \left[\frac{\dot{\beta}_{s_i}}{r} + (1-\beta_{s_i})\right](p_{s_i} + \tau_0(0,E)e^{rs_i}), \tag{65}$$

and differentiate (51) with regard to C_i with the more general fossil price path to obtain

$$\frac{1}{\varepsilon_i} = \left(\frac{\dot{p}_{s_i}^F}{r} + \tau_0(0,E)e^{rs_i}\right)\frac{ds_i}{dC_i}. \tag{66}$$

Now we can substitute (66) into (54) and solve for

$$\tau_{s_i} = \left(1-\beta_{s_i} - \left(\frac{\varepsilon_i p_0^F}{C_i r}\right)^{\frac{r}{g}} e^{rs_i}\right)\left(\frac{\dot{p}_{s_i}^F}{r} + \tau_0(0,E)e^{rs_i}\right), \tag{67}$$

which, in turn, is used to eliminate the tax rate from the efficiency condition (52) in order to express it as the differential equation

$$\frac{\dot{\beta}_{s_i}}{r} = \frac{p_{s_i}^F + (1-\beta_{s_i})\left(\frac{\dot{p}_{s_i}^F}{r} - p_{s_i}^F\right) - \left(\frac{p_0^F}{p_{s_i}^F + \tau_0(0,E)e^{rs_i}}\right)^{\frac{r}{g}} e^{rs_i}\left(\frac{\dot{p}_{s_i}^F}{r} + \tau_0(0,E)e^{rs_i}\right)}{p_{s_i}^F + \tau_0(0,E)e^{rs_i}}. \tag{68}$$

So far, we derived a differential equation for β_t and simplified it to (68). This equation is equivalent to

$$\frac{d}{dC_i}(\sup_{s_i} u_i - u_{i,BAU}) = 0 \quad \forall\, C_i, \varepsilon_i. \tag{69}$$

Next, we have to pin down an initial condition for β_0. Recall that the individually optimal switching time is always purely a function of ε_i/C_i. Thus we can write $s(\varepsilon_i/C_i)$ for the individually rational switching time with climate policy, (τ_t, β_t) and, respectively, $s^{BAU}(\varepsilon_i/C_i)$ in the BAU with $\tau^{BAU} = \beta^{BAU} = 0$. Consider a household which is induced by climate policy to switch at time $t = 0$, without being at a corner solution there. The corresponding first-order condition reads

$$s^{-1}(0) = \frac{r}{p_0 + \tau_0(0,E)}. \tag{70}$$

Thus, utility is equal to

$$u_i = -(1-\beta_0)C_i. \tag{71}$$

In the BAU, the same household attain a utility level of

$$u_i^{BAU} = -C_i \left(\int_{t=0}^{s_i^{BAU}(\varepsilon_i/C_i)} e^{-rt} p_t^F \frac{\varepsilon_i}{C_i} + e^{-rs_i^{BAU}(\varepsilon_i/C_i)} \right). \tag{72}$$

For exact compensation, (71) and (72) have to equated which may then be solved for

$$\beta_0 = 1 - \left(\int_{t=0}^{s_i^{BAU}(\varepsilon_i/C_i)} e^{-rt} p_t^F \frac{\varepsilon_i}{C_i} + e^{-rs_i^{BAU}(\varepsilon_i/C_i)} \right). \tag{73}$$

We argue that if this equation holds for some i with $s_i(\varepsilon_i/C_i) = 0$, then it will hold for all such i. To see why this is the case, recall that the individually rational switching time given the fixed policy path, $s_i(\varepsilon_i/C_i)$, is actually fully determined by ε_i/C_i. In particular, all i with $s_i(\varepsilon_i/C_i) = 0$ have the same $s_i^{BAU}(\varepsilon_i/C_i)$ and, thus, for all these i, the right hand side of (73) is identical. If (73) holds for one such i it will hold for all such i.

Additionally, if (68) holds, we can deduce that $\sup_{s_i} u_i(C_i, \varepsilon_i) = u_{i,BAU}(C_i, \varepsilon_i)$ for all C_i. Thus, if we have a subsidy schedule, β_t, such that both (68) and (73) hold, then we will be guaranteed that $\sup_{s_i} u_i(C_i, \varepsilon_i) = u_{i,BAU}(C_i, \varepsilon_i)$ for all C_i, ε_i. To complete the proof, we argue that the differential equation has a unique solution for a given initial condition for β_0. To see why this is the case, we rewrite (68) as

$$\frac{\dot{\beta}_{s_i}}{r} = \frac{\frac{d}{ds_i}\left(p_{s_i}^F e^{-rs_i}\right)}{p_{s_i}^F e^{-rs_i} + \tau_0(0,E)} \beta_{s_i} + \frac{p_{s_i}^F + \left(\frac{\dot{p}_{s_i}^F}{r} - p_{s_i}^F\right) - \left(\frac{\dot{p}_{s_i}^F}{r} + \tau_0(0,E)e^{rs_i}\right)\left(\frac{p_0}{p_{s_i} + \tau_0(0,E)e^{rs_i}}\right)^{\frac{r}{g}} e^{rs_i}}{p_{s_i}^F + \tau_0(0,E)e^{rs_i}}. \tag{74}$$

Given our assumption that $\frac{\frac{d}{ds_i}(p_{s_i}^F e^{-rs_i})}{p_{s_i}^F e^{-rs_i} + \tau_0(0,E)}$ is bounded above, the right hand side of (74) is Lipschitz continuous. The claim now follows from Simmons (2016), Theorem B, Section 70, Chapter 13.

B.4 Deadweight loss of subsidizing

Exact compensation via the ad-valorem subsidy reduces the carbon price consistent with climate target E to

$$\tau_t^* = \tau_t(0,E) - \beta^*(\tau_t(0,E) + p_t^F). \tag{75}$$

The subsidy-induced relative reduction in the tax mark-up on fossil price follows as

$$\frac{\tau_t(0,E)-\tau_t^*}{p_t^F+\tau_t(0,E)}=\beta^*, \tag{76}$$

effectively reducing the incentive to turn down the thermostat at time t. Assuming that $e_D\beta^*$ is sufficiently small, the loss of social surplus caused by higher fossil consumption is approximated by the Harberger triangle formula as

$$\frac{e_D}{2}\left(\beta^*\right)^2 p_t^F\varepsilon_i, \tag{77}$$

where we substitute $\beta^*=[x-\log(1+x)]/(1+x)$ and take the discounted value of the loss until the household's optimal switching time s^* to obtain the upward distortion in fossil fuel consumption

$$DWL^F=\int_{t=0}^{s^*}e^{-rt}\frac{e_D}{2}\left(\frac{x-\log(1+x)}{1+x}\right)^2 p_t^F\varepsilon_i dt=\frac{e_D}{2}\left(\frac{x-\log(1+x)}{1+x}\right)^2 p_0^F\varepsilon_i s_i^*. \tag{78}$$

After switching to a heat pump, the operational cost, i.e., price of heat pump electricity, is also distorted downwards by the proportion β^*. Thus, assuming that the cost of heat pump electricity is lower than the fossil price in $t=0$, we can proceed analogously to (78) to obtain the discounted value of the loss of social surplus due to excessively high electricity consumption after switching given by

$$DWL^E=\int_{t=s_i^*}^{\infty}e^{-rt}\frac{e_D}{2}\left(\frac{x-\log(1+x)}{1+x}\right)^2 p_t^F\varepsilon_i=\frac{1}{r}e^{-rs_i^*}\frac{e_D}{2}\left(\frac{x-\log(1+x)}{1+x}\right)^2 p_0^F\varepsilon_i. \tag{79}$$

Across all households, the cumulative losses DWL^F+DWL^E amount to

$$\frac{e_D}{2}\left(\frac{x-\log(1+x)}{1+x}\right)^2 p_0^F\mathbb{E}\left[\varepsilon_i(s_i^*+\frac{e^{-rs_i^*}}{r})\right]. \tag{80}$$

Dividing by discounted aggregate transformation costs in (14) yields

$$\frac{1}{2}\frac{x-\log(1+x)}{1+x}\frac{\mathbb{E}[\varepsilon_i(s_i^*+\frac{e^{-rs_i^*}}{r})]}{\mathbb{E}[\frac{\varepsilon_i}{r}]}e_D\leq\frac{1}{2}\frac{x-\log(1+x)}{1+x}\left(1+\frac{rE}{\mathbb{E}[\varepsilon_i]}\right)e_D, \tag{81}$$

resulting in an upper bound for the deadweight loss caused by the ad-valorem subsidy as a share of the total cost of decarbonizing the building sector.

B.5 Alternative specifications of climate damages

In the main analysis, we assume that climate damages reduce real output, i.e., negatively affect productivity. Three alternative channels of climate change-induced damages are considered in this extension. First, products may be subject to heterogeneous price changes as the vulnerability to climate change varies between sectors (e.g., agricultural productivity may increase reducing food prices while other products become more expensive). This is effect is represented in this extension by assuming that individual i's consumption, c_i, is a vector of k differentiated produces with respective prices, p_k. Second, climate change can have adverse impacts on amenities (e.g., lower quality of life due to excessive heat). Third, individuals' utility may also decrease as a result of (partial) altruism towards future generations which will experience the aforementioned effects (more severely). The last two impacts are accounted for by introducing spending on an offsetting technology, q, where the willingness to pay for emission offsetting coincides with the total marginal benefit of amenities and altruistic preferences. Consider individual utility of the form

$$u_i\left[v(c_i, e - q_i, q_i), y_i\right], \tag{82}$$

subject to the budget constraint

$$y_i - \Gamma(y_i) = c_i p + q_i p_q, \tag{83}$$

where p denotes the vector of prices, p_k for the vector of consumption, c_i, and p_q is the price vector for the path of emission offsets, q_i, respectively. By defining indirect utility

$$V(y, \Gamma, p, p_q, e) = \sup_{(c,q): y - \Gamma(y) = cp + qp_q} v(c, e - q, q), \tag{84}$$

we can identify the income tax schedule, $\Gamma_{(p,p_q,e)}$, which completely skims the benefits of avoided climate damages implicitly via the condition

$$V(y, \Gamma_{(p,p_q,e)}(y), p, p_q, e) = V(y, \Gamma^{BAU}, p^{BAU}, p_q^{BAU}, e^{BAU}). \tag{85}$$

By defining

$$U_i(y_i)_{(\Gamma, p, p_q, e)} := u_i\left[V(y_i, \Gamma, p, p_q, e), y_i\right], \tag{86}$$

condition (85) is equivalent to

$$U_i(y_i)_{(\Gamma_{(p,p_q,e)}, p, p_q, e)} = U_i(y)_{(\Gamma^{BAU}, p^{BAU}, p_q^{BAU}, e^{BAU})} \quad \forall (p, \rho, e), y_i, \tag{87}$$

implying that the optimal labor decision, y_i of each individual i remains undisitorted by the updated income tax schedule, $\Gamma_{(p,p_q,e)}$, in comparison to the BAU. By means of total differentiation, we derive how the income tax schedule has to change for a marginal change in the emission target. Consider that price paths p and p_q are functions of the emissions path e, such that

$$\frac{d}{de}\Gamma_{(p(e), p_q(e), e)}(y) = c(y,p)\frac{dp}{de} + \frac{v_2[c(y,p), e - q(y), q(y)]}{\frac{\partial}{\partial c_k} v[c(y,p), e - q(y), q(y)] p_k^{-1}} + q(y)\frac{dp_q}{de}. \tag{88}$$

Assuming that a change in the emissions path, e, does not affect the price for the offsetting technology, we can rewrite (88) as

$$\frac{d}{de}\Gamma_{(p(e), p_q, e)}(y) = c(y,p)\frac{dp}{de} + \frac{v_2[c(y,p), e - q(y), q(y)]}{\frac{\partial}{\partial c_k} v[c(y,p), e - q(y), q(y)] p_k^{-1}}, \tag{89}$$

where the first term on the RHS captures the change in expenditures assuming that a marginal change in p does not affect the composition of the consumption bundle, c. The second term equates to the willingness to pay for a marginal reduction in e for the sake of reducing emissions themselves apart from the fact that emission abatement also affects consumption prices, p_k. Hence, the second term relates to the marginal benefit of avoided damages with regards to amenities and altruism. The marginal income tax adjustment skimming all benefits of avoided climate damages follows as

$$\Delta\Gamma(y_i^{BAU}) = \Gamma_{(p(e),p_q,e)}(y_i^{BAU}) - \Gamma_{(p(e^{BAU}),p_q,e^{BAU})}(y_i^{BAU})$$
$$= \int_{\tilde{e}=e^{BAU}}^{e} c\left(y_i^{BAU}, p(\tilde{e})\right)\frac{dp}{d\tilde{e}} + \frac{v_2\left[c(y_i^{BAU},p), \tilde{e}-q(y_i^{BAU}), q(y_i^{BAU})\right]}{\frac{\partial}{\partial c_k}v\left[c(y_i^{BAU},p), \tilde{e}-q(y_i^{BAU}), q(y_i^{BAU})\right]p_k^{-1}}. \quad (90)$$

Similar to Lemma 6, the reform package consisting of a marginal emission reduction induced by τ_t^* and β_t^* and the corresponding income tax adjustment passes the cost-benefit test if $\int_{i\in I}\Delta\Gamma(y_i^{BAU})$, i.e., the monetized total net benefits, exceed the cost of mitigation.

Given the implicit assumptions about preferences in (82), these marginal effect of reducing emissions only depends on an individual's income which is why we can simply add the partial effects to the income tax schedule and exactly offset any effect on utilities. Then, labor supply incentives remain invariant, such that the labor supply decisions will in fact remain unchanged. For non-marginal reforms, we can carry out this construction in small steps. To compute this exactly, we will need to know how consumption choices adjust in response to price changes and how the marginal willingness to pay for emission reductions changes with emissions.

B.6 Alternative compensation via grandfathering based on historical emissions

Consider the case where there are only two types of household which differ both in ε_i and C_i, i.e., the carbon price burden varies between households as does the individually rational switching time. As long as both types still use fossil fuels, grandfathering based on ε_i ensures that each household is perfectly compensated for their carbon price burden. Yet, as soon as the first type switches to a heat pump, their carbon price burden drops to zero whereas the compensation is still tethered to historical emissions and, thus, positive. Consequently, the first type will be over-compensated at the expense of the second type which now finances the rebate to all types. In principle, this effect can be mitigated by resetting the reference point for grandfathering on a regular basis. However, if such resets are anticipated, households have a strategic incentive to inefficiently delay decarbonizing investments.

The net present value of transfers in the grandfathering scheme amounts to

$$T_i^g = \frac{\varepsilon_i}{E(\varepsilon_j)}\int_{t=0}^{\infty}\tau_t E(1_{s_j>t}\varepsilon_j), \quad (91)$$

where $1_{s_j>t}$ is a unit vector indicating whether household j is still using fossil fuels in period t. The expectations operator is taken over all households with their mass normalized to 1. In this application, ε_i coincides with household i's historical emissions. Unless the regulator resets the reference point, i will receive a transfer even after switching to a carbon-neutral technology. If the reference point is reset at any point in time, $0<\tilde{t}<\infty$, the fraction in (91) is updated to $1_{s_i>\tilde{t}}\varepsilon_i/E(1_{s_j>\tilde{t}}\varepsilon_j)$ such that only households which have not yet decarbonized by $\tilde{t}$ will be eligible to the transfer from then onward.

First, we show that exact compensation is generally not possible when transfers are based on a grandfathering rule. Assume for now that switching costs are a function of household i's switching time, s_i, as well as the rest of the population's switching times, s_{-i}, i.e., $C(s_i,s_{-i})$. This dependency may stem from time-variant prices for retrofitting, heat pumps and electricity adjusting to episodes of high or low aggregate demand for carbon-neutral technologies. Then, using (91), individual household utility can be rewritten as

$$\frac{u_i}{\varepsilon_i} = -\left(\int_{t=0}^{s_i} e^{-rt}(p_t^F + \tau_t) + (1-\beta_{s_i})e^{-rs_i}\frac{C_i(s_i, s_{-i})}{\varepsilon_i} \right) + \frac{1}{E(\varepsilon_j} \int_{t=0}^{\infty} \tau_t E(1_{s_j>t}\varepsilon_j). \tag{92}$$

Since each household is infinitesimally small relative to the set of all households , $E(1_{s_j>t}\varepsilon_j)$ is invariant to s_i. Similarly, a single household's switching decision has no effect on prices such that the reduced function, $s_i \mapsto C(s_i, s_{-i})/\varepsilon_i$, is identical for i for any given allocation, $s = (s_j)_{j\in I}$. Then, the RHS of (92) is identical for all i, such that $\sup_{s_i} u_i/\varepsilon_i$ and $\sup_{s_i} u_i^{BAU}/\varepsilon_i$ are also the same for all i. Consequently, a policy package consisting of carbon price, ad-valorem subsidy and grandfathered transfers can ensure $\sup_{s_i} u_i = \sup_{s_i} u_i^{BAU}$ for all households i, i.e., achieve exact compensation, if the transfer is sufficiently large. However, this requires that all heterogeneity in C_i is explained by the respective household's carbon intensity, ε_i. Otherwise, the RHS of (92) will not be the same for all i. As outlined in Section 2, switching costs in the building sector are subject to many-dimensional, weakly correlated heterogeneities turning the requirement $C_i \sim \varepsilon_i$ hard to justify. Among households with similar emission intensities, ε_i, those with high switching costs, C_i, will typically prefer to switch later. Grandfathering will dynamically under-compensate these households.

Second, we address the strategic incentives to delay switching under a grandfathering scheme. As more and more households switch to the carbon neutral technology and carbon pricing revenues are increasingly redistributed from households still using fossil fuels to decarbonized households, the regulator faces political pressure to reset the reference point as time progresses. Suppose that there is a fixed annual probability, λ, that the reference point is reset to the current-period emission intensities, $1_{s_i>t}\varepsilon_i$. This introduces incentives to postpone switching. Consider household i which could switch at time t or delay the investment by dt to $t+dt$. By delaying the investment, i incurs additional expenditures for fossil fuel equal to $(p_t^F + \tau_t)\varepsilon_i dt$. At the same time, i reaps additional interest earnings amounting to $rC_i dt$. These first two components drive the individually rational switching decision even under exactly compensating category-based transfers or absent any transfers. Strategic considerations stemming from grandfathering are introduced by the expected future transfer payments. If the reference point is reset in the interval between t and $t+dt$, postponing the investment entitles household i to future transfers. In expectation, future transfer payments amount to

$$\lambda dt \frac{\varepsilon_i}{E(1_{s_j>t}\varepsilon_j)} \int_{s=t}^{\infty} e^{-\lambda(s-t)} E(1_{s_j>s}\varepsilon_j)\tau_t, \tag{93}$$

where we already substitute the Hotelling carbon price path, $\tau_s = \tau_t e^{r(s-t)}$. The exponential under the integral, $e^{-\lambda(s-t)}$, denotes the probability that the reference point will not be reset again between $t+dt$ and s. Assuming that, each year, a fixed proportion of the households still using fossil fuels weighted by their respective emission intensities, ζ, switches to carbon-neutral heating, the emission-weighted share of households which has not yet switched by s is given by

$$F(s) = 1 - E(1_{s_j>s}\varepsilon_j) = 1 - e^{-\zeta s}, \tag{94}$$

which may be substituted into (93) to obtain

$$\frac{\lambda}{\lambda+\zeta}dt\varepsilon_i\tau_t. \tag{95}$$

Aggregating all additional cost and gains associated with postponing the investment by dt results in net expected gains of

$$-\left[p_t^F+\tau_t\left(1-\frac{\lambda}{\lambda+\zeta}\right)\right]\varepsilon_i dt + rC_i dt. \tag{96}$$

Strategic delays of switching erode $\lambda/(\lambda+\zeta)$ of the carbon price. Hence, the regulator has to increase the carbon price by a fraction of $(\lambda+\zeta)/\zeta$ in order to achieve the same switching incentives as in the absence of compensation via grandfathering. For instance, suppose that $\zeta=0.05$ and $\lambda=0.1$, i.e., 5% of all households still using fossil fuels switch annually while the reference point is reset once every 10 years on average. Then, the carbon price has to increase by three times to compensate for the strategic incentives introduced by grandfathering.

B.7 Calibrated model: alternative carbon price paths

In the main analysis, we provide numerical estimates of the optimal ad-valorem subsidy path and the corresponding tax rates for a climate target, E, consistent with the Social Cost of Carbon at a 2% discount rate. Employing the SCC has the advantage of knowing ex-ante that a policy package based on these carbon prices will pass the cost-benefit criterion. In the following, we provide additional numerical results for the SCC at higher and lower discount rates as also provided by EPA (2023). Alternatively, we can calibrate our model to match carbon prices consistent with the more ambitious EU climate targets employing optimal abatement cost estimates from the integrated assessment model REMIND-EU (Pietzcker et al. 2021). However, since these prices are derived via a cost-effectiveness approach, it is not immediately obvious that the policy package will pass the cost-benefit test, i.e., whether its benefits exceeding the aggregate cost of decarbonization. As the REMIND-EU model implements the EU climate target of net carbon neutrality by 2045 at least cost, it requires higher carbon prices than the SCC, resulting in a higher initial mark-up on the fossil price of $x=0.65$ (compared to $x=0.59$ for the SCC at $r=0.02$). This induces earlier switching across all households. To compensate for the increased abatement cost, a higher ad-valorem subsidy is required to achieve exact compensation.

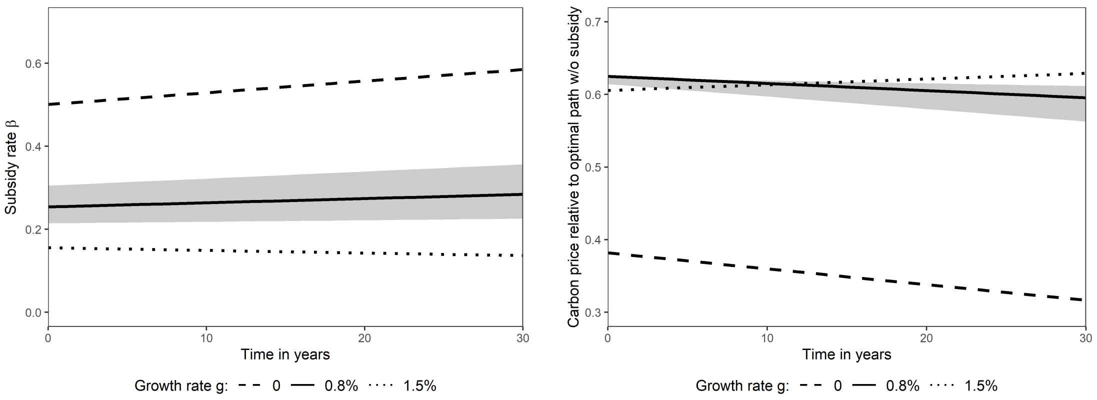
In either case, we can derive the optimal carbon price as

$$\tau_t(\beta_t,E)=-\frac{1}{r}\left(e^{gt}g+e^{ht}hx\right)\left[-1+e^{rt}\left(\frac{1}{e^{gt}+e^{ht}x}\right)^{r/g}\beta_t\right]. \tag{97}$$

While we can no longer derive a closed form solution for the optimal ad-valorem subsidy when employing a carbon price deviating from the Hotelling rule, the time derivative of β_t is given as

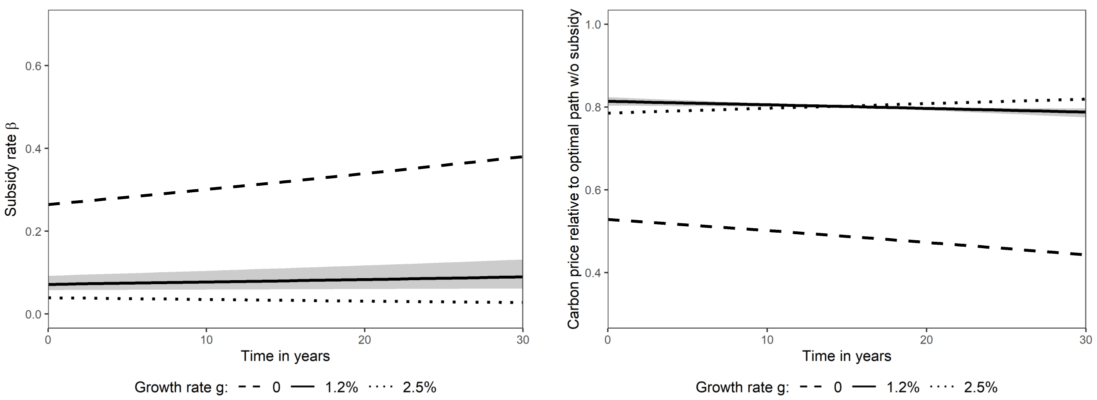
$$\dot{\beta}_t=\frac{1}{e^{gt}+e^{ht}x}\Bigg[e^{gt}g+e^{ht}hx-e^{ht}rx-\left(\frac{1}{e^{gt}+e^{ht}x}\right)^{r/g}\left(e^{(g+r)t}g+e^{(h+r)t}hx\right) \\ +\left(e^{gt}(r-g)+e^{ht}(r-h)x\right)\beta_t\Bigg], \tag{98}$$

illustrating that the slope of β_t depends on $g \gtrless h$, where h denotes the growth rate of the carbon price for $\beta = 0$, i.e., $\tau_t(0, E)$.



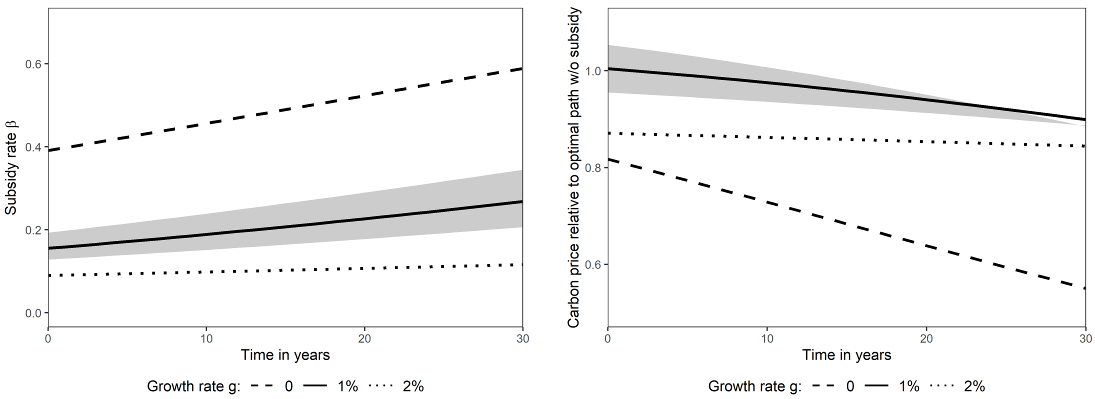
(a) Ad-varlorem subsidy (b) Carbon price

Figure 8: **SCC carbon price at 1.5% interest rate.** Optimal path of the ad-valorem subsidy rate and the corresponding carbon price (relative to optimal carbon price absent subsidies) for different realizations of fossil price growth rate. Line types reflect the set of values for α whereas the shaded area depicts the range of g between historic plausible values of 36% to 65% of g (Source: own depiction)



(a) Ad-varlorem subsidy (b) Carbon price

Figure 9: **SCC carbon price at 2.5% interest rate.** Optimal path of the ad-valorem subsidy rate and the corresponding carbon price (relative to optimal carbon price absent subsidies) for different realizations of fossil price growth rate. Line types reflect the set of values for α whereas the shaded area depicts the range of g between historic plausible values of 36% to 65% of g (Source: own depiction)



(a) Ad-varlorem subsidy

(b) Carbon price

Figure 10: **REMIND-EU carbon price at 2% interest rate.** Optimal path of the ad-valorem subsidy rate and the corresponding carbon price (relative to optimal carbon price absent subsidies) for different realizations of fossil price growth rate. Line types reflect the set of values for α whereas the shaded area depicts the range of g between historic plausible values of 36% to 65% of g (Source: own depiction)