How Personal Experiences and Social Concerns Shape Green Innovation and Climate Policy

I n a u g u r a l d i s s e r t a t i o n

zur

Erlangung des PhD

der

Wirtschafts- und Sozialwissenschaftlichen Fakultät

der

Universität zu Köln

2026

Vorgelegt von

Marten Christian Ritterrath, Master of Science

Referent/in:	Prof. Anna Bindler PhD
Korreferent/in:	Prof. Dr. Felix Bierbrauer

Tag der Disputation: 20.07.2026

Contents

Introduction **10**

1 Inventors' Personal Experience of Natural Disasters and Green Innovation **16**

1.1 Introduction . . . 18
1.2 Data . . . 23
1.2.1 Patents . . . 23
1.2.2 Natural Disasters . . . 25
1.2.3 Analysis Sample - Patenting . . . 26
1.3 Empirical Strategy . . . 28
1.4 Effects on Innovation . . . 31
1.4.1 Mitigation vs. Adaptation . . . 32
1.4.2 Spillovers . . . 34
1.4.3 Patent Value . . . 35
1.4.4 Types of Inventors . . . 38
1.4.5 Inventor Movement . . . 39
1.4.6 Disaster Severity . . . 39
1.4.7 Survey Measure of Green Innovation . . . 41
1.5 Reasons to Innovate . . . 42
1.6 A Model of Disaster-Induced Green Innovation . . . 45
1.6.1 Model . . . 46
1.6.2 Empirical Results on Model Hypothesis . . . 54
1.7 Conclusion . . . 58

2 Welfare Weights and the Social Cost of Carbon **59**

2.1 Introduction . . . 60
2.2 Model . . . 63
2.2.1 The optimal carbon tax . . . 65
2.3 Main Result . . . 66
2.3.1 The Invariance Benchmark . . . 67
2.3.2 Equilibrium Brown Consumption . . . 69

2.3.3 Heterogeneous Preferences . . . 72
2.4 Dynamic Extension: When Does the Economy Become Green? . . . 74
2.4.1 Setup . . . 74
2.4.2 Incentive Compatibility . . . 76
2.4.3 The Planner's Problem . . . 76
2.4.4 Optimal Switching Time . . . 76
2.4.5 The switching condition in λ-Form . . . 77
2.4.6 Structural identity and within-phase channel . . . 78
2.4.7 Homogeneous Preferences and the Invariance Benchmark . . . 81
2.4.8 Heterogeneous Preferences . . . 83
2.5 Conclusion . . . 86

3 Morals and the Political Economy of Corrective Taxes 88
3.1 Introduction . . . 89
3.2 The Model . . . 95
3.2.1 Neoclassical consumers . . . 96
3.2.2 Socially responsible consumers . . . 97
3.3 Preferences over taxes . . . 98
3.3.1 Neoclassical consumers . . . 98
3.3.2 Socially responsible consumers . . . 100
3.4 Political equilibrium . . . 101
3.4.1 Country A: Neoclassical consumers only . . . 101
3.4.2 Country B: Socially responsible consumers only . . . 102
3.4.3 Country C: A mix of N- and S-agents . . . 102
3.5 Moral pressure . . . 106
3.6 Moral rebellion . . . 108
3.7 Survey . . . 117
3.7.1 Primitives . . . 118
3.7.2 Tax Preferences . . . 119
3.7.3 Moral Pressure . . . 120
3.8 Concluding Remarks . . . 123

4 Appendix for Chapter 1 135
4.1 Appendix - Theory . . . 135
4.1.1 Demand & Market-Clearing . . . 135
4.1.2 Proof of Proposition 1 . . . 137
4.1.3 Welfare Derivation and Proof of Proposition 2 . . . 138
4.1.4 Parameter Calibration - Comparative Statics . . . 139
4.2 Appendix - Empirical Analysis . . . 140

4.2.1 Alternative Estimator . . . 140
4.2.2 Patenting Results with Additional Controls . . . 141
4.2.3 Green innovation by technology class . . . 143
4.2.4 Subclasses of Y02 . . . 144
4.2.5 Most Severe-Disaster Severity . . . 146
4.2.6 Disaster Type . . . 147
4.2.7 Spillovers - Alternative Modeling . . . 148
4.2.8 Reasons to Innovate Robustness . . . 150
4.2.9 Alternative Measures of Green Innovation Robustness . . . 153
4.2.10 Alternative Competition Windows . . . 154
4.2.11 Alternative Green Good Window . . . 156
4.2.12 First Time & Repeat Inventors . . . 157
4.2.13 Past Patenting . . . 157
4.3 Appendix - Data . . . 158
4.3.1 Countries in the CompNet Dataset . . . 158
4.3.2 Community Innovation Survey . . . 160
4.3.3 Table of Green Goods for PRODCOM . . . 160

5 Appendix for Chapter 2 **172**
5.1 Basic Setup . . . 172
5.1.1 Lagrangian and First-Order Conditions . . . 172
5.1.2 Proof of Lemma 1 . . . 173
5.1.3 Proof of Proposition 1 . . . 174
5.1.4 Proof of Lemma 2.2 . . . 175
5.1.5 Proof of Proposition 2.4 . . . 177
5.1.6 Proof of Proposition 2.1 . . . 183
5.2 Appendix - Dynamic Extension . . . 184
5.2.1 Lagrangian and First-Order Conditions . . . 184
5.2.2 Derivation of the Switching FOC . . . 186
5.2.3 Appendix: Proof of Proposition 2.6 . . . 188
5.2.4 Numerical Calibration . . . 191

6 Appendix for Chapter 3 **193**
6.1 Appendix . . . 193
6.1.1 Proof of Proposition 3.1 . . . 193
6.1.2 Proof of Proposition 3.2 . . . 196
6.1.3 Proof of Proposition 3.3 . . . 197
6.1.4 Proof of Proposition 3.4 . . . 197
6.1.5 Proof of Proposition 3.5 . . . 198

6.1.6 Proof of Lemma 3.3 199
6.1.7 Proof of Proposition 3.6 199
6.1.8 Universalists 199
6.1.9 An alternative interpretation of the setup: Heterogeneity in incomes as opposed to hetereogeneity in the taste for dirty goods. . 201
6.1.10 Additional Numerical Illustrations 202
6.1.11 Empirical Appendix 203
6.1.12 Survey Instructions 206

7 Appendix for Software **236**

Software **236**

Curriculum Vitae **239**

Eidesstattliche Erklärung **242**

List of Figures

1.1 Patenting Activity in France and Germany over Time 25
1.2 Bivariate Map of Green Patent and Disaster Counts 27
1.3 Patenting following the Exposure to a Natural Disaster 32
1.4 Illustration of Distances to Region 35
1.5 Green innovation response to natural disaster exposure by patent characteristics 37
1.6 Innovation by Inventor Type and Mobility 38
1.7 Innovation Responses dependent on Competition Δ_j 51

2.1 **Homogeneous preferences: MRS by** V_h 69
2.2 **Heterogeneous preferences: MRS by** V_h 74
2.3 **Optimal switching time** T^* 87

3.1 Strong Waterbed Effect 104
3.2 Weak Waterbed Effect 105
3.3 Rebellion in Response to Moral Pressure 110
3.4 How Taxes Affect Moral Pressure and Moral Rebellion in Equilibrium . . 112
3.5 Preferred Tax Rates by Agent Type across Regimes 114
3.6 Political-Equilibrium Payoffs 115
3.7 Two-Dimensional Heterogeneity: Preferred Taxes and Equilibrium Payoffs 117

4.1 Innovation Responses dependent on market size η_j 140
4.2 Patenting following Exposure to a Natural Disaster — Estimator of Chaisemartin and D'Haultfœuille 2023; Chaisemartin and D'Haultfœuille 2024a 141
4.3 Patenting following exposure with technology trend controls 142
4.4 Patenting following exposure with GDP per capita controls 143
4.5 Green innovation response to natural disaster by technology class 144
4.6 Green innovation patterns in Germany and France by technology focus . 145
4.7 Patenting following the Exposure to the Most Severe Natural Disaster . . 147
4.8 Green Patenting Following Different Disaster Types 148
4.9 Green innovation response to natural disaster by past patenting 158

6.1 Weak Waterbed Effect under $\bar{\theta}_N > \bar{\theta}_S$ . . . 203
6.2 Introduction screen (1/2). . . . 207
6.3 Introduction screen (2/2). . . . 208
6.4 CAPTCHA screen. . . . 209
6.5 Prolific ID entry. . . . 209
6.6 Gender question. . . . 209
6.7 Education question (UK version). . . . 210
6.8 Education question (U.S. version). The education question was adapted to reflect the U.S. educational system. . . . 210
6.9 Income question (UK version). . . . 211
6.10 Income question (U.S. version). The income question was adapted to reflect U.S. income brackets (in U.S. dollars). . . . 212
6.11 Car type question. . . . 212
6.12 Car usage question (UK version). . . . 213
6.13 Car usage question (U.S. version). The car usage question was adapted to reflect U.S. driving patterns. . . . 213
6.14 Elicitation of θ. . . . 214
6.15 Elicitation of δ. . . . 214
6.16 Carbon tax information screen. . . . 215
6.17 Carbon tax comprehension check. . . . 216
6.18 Price adjustment under carbon tax. . . . 217
6.19 Price under carbon tax, hold-out version. . . . 218
6.20 Social cost of carbon information screen. . . . 219
6.21 Social cost of carbon comprehension check. . . . 220
6.22 Social cost of carbon adjustment screen. . . . 220
6.23 Social cost of carbon adjustment, hold-out version. . . . 221
6.24 Preferred tax level (UK baseline version). . . . 222
6.25 Preferred tax level, after the slider was moved. . . . 223
6.26 Preferred tax level (U.S. version). The preferred tax question was adapted to U.S. context (currency and reference points). . . . 224
6.27 Local tax based on SCC (UK version). . . . 225
6.28 Local tax based on SCC, after the slider was moved. . . . 225
6.29 Local tax based on SCC (U.S. version). The local SCC-based tax question was adapted to U.S. context. . . . 226
6.30 Local tax based on SCC, hold-out version. . . . 226
6.31 Global tax based on SCC (UK baseline version). . . . 227
6.32 Global tax based on SCC, after the slider was moved. . . . 228
6.33 Global tax based on SCC (U.S. version). The global SCC-based tax question was adapted to U.S. context. . . . 229

6.34 Global tax based on SCC, hold-out version. 229
6.35 Moral pressure, positive 230
6.36 Moral pressure, negative 230
6.37 Response to Moral Pressure 231
6.38 Price versus Harm 232
6.39 Reciprocity measure. 233
6.40 Clarity check. 234
6.41 Self-assessment screen. 235

List of Tables

1.1 Data Sources, Coverage, Periods, and Usage 23
1.2 Patenting Responses split by Adaptation vs. Mitigation 33
1.3 Spillovers of Neighboring Disasters 36
1.4 Patenting Effects by Disaster Severity 40
1.5 Effect of Natural Disasters on Green Innovation and Climate Affectedness 42
1.6 Effect of Cumulative Disaster Count on Reasons to Innovate 44
1.7 Competition and Green Product Split 56

3.1 Foundations — Correlations with Socially Responsible Behavior 119
3.2 Tax Preferences — Level 120
3.3 Support and Opposition to Moral Pressure — UK 121
3.4 Support and Opposition to Moral Pressure — US 122
3.5 Moral Pressure Attitudes and Tax Preferences 123

4.1 Y02 subclasses of climate change mitigation technologies 145
4.2 Regression by Y02 subclass 146
4.3 Spillovers of Neighboring Disasters - Alternative Windows 149
4.4 Spillovers of Neighboring Disasters - Unrestricted Sample 150
4.5 Effect of Cumulative Disaster Count on Reasons to Innovate (Robustness Check NUTS-2) 151
4.6 Effect of Cumulative Disaster Count on Reasons to Innovate (Robustness Check NUTS-3) 151
4.7 French Research Funding 152
4.8 Effect of Cumulative Disaster Count on Green Innovation and Climate Affectedness (Robustness Regional Fixed Effects) 154
4.9 Competition Split Above/Below Median 155
4.10 Green Product Split by Median 156
4.11 Response by First Time & Repeat Inventors 157
4.12 Comp Net TimpeSpans 159
4.13 Variable Definition - Factors driving green innovation 160
4.14 Green Good Table 161

6.1 Tax Preferences — Responsiveness to Local Damage Information 204
6.2 Tax Preferences — Responsiveness to Global Damage Information 204
6.3 Support and Opposition to Moral Pressure — UK (Ordered Probit) . . . 205
6.4 Support and Opposition to Moral Pressure — US (Ordered Probit) . . . 206

Introduction

In July 2021, the Ahr valley in western Germany flooded. More than 130 people died. The water tore through villages that had stood for centuries, reaching communities not typically associated with climate vulnerability. The federal election that followed put climate policy near the centre of the debate, and the new government legislated an ambitious climate package. Within four years, parts of it were being rolled back. The forces that shape these decisions remain poorly understood.

This dissertation studies two such forces. The first is personal experience, the influence of direct exposure to climate events on economic choices. The second is social concern, used here in two distinct senses that the chapters take up in turn. The positive sense concerns whether people accept a moral obligation to reduce their own carbon footprint beyond what material self-interest dictates, and how that obligation, where felt, shapes behavior in markets and at the ballot box. The normative sense concerns society's distributive judgments about whose welfare counts, and how those judgments shape the policies that economic theory recommends. Each chapter takes one of these forces and works through its consequences for a climate-relevant choice: the direction of inventive activity, the social cost of carbon, and the political equilibrium of corrective taxation. Each of these choices runs into a persistent difficulty. Societies underinvest in green R&D despite the high social returns to innovation, resist corrective pricing of carbon politically, and differ widely in their acceptance of climate-change risks.

Paper 1: Inventors' Personal Experience of Natural Disasters and Green Innovation. The first chapter, joint with Lisa Keding, asks whether inventors who personally experience natural disasters subsequently shift their research toward green technologies. We know surprisingly little about what drives inventors to pursue new ideas. A rich literature links innovation to market conditions such as competition (Aghion et al. 2005) and market size (Acemoglu 2002), but much less is known about inventor-level determinants. Addressing this gap is particularly relevant given persistent underinvestment in R&D despite its large social returns (Bloom et al. 2013; Azoulay et al. 2019; Jones and Summers 2020; Jones 2021), and the role of innovation in combating climate change (Acemoglu et al. 2012; Acemoglu et al. 2016).

The experience effects literature has shown that personal exposure to large shocks

reshapes household beliefs, preferences, and expectations (Malmendier and Nagel 2011; Malmendier and Nagel 2016; Giuliano and Spilimbergo 2025), yet has not asked whether such effects extend to the production side and shape the high-stakes decisions of inventors. We show that personal experience is one such determinant. Natural disasters locally increase the salience of climate change, shift expectations about the returns to green technologies, and redirect inventive effort accordingly. The innovations that result have global reach, being both highly cited and frequently patented in major markets around the world. The chapter thus documents a channel through which personal experience of natural disasters creates an endogenous response to climate change.

We exploit natural disasters as exogenous shocks to inventors' personal experience and link French and German inventors' home addresses to disaster records. Using event-study designs that exploit random variation in the timing of exposure, we find that one additional natural disaster increases green patenting by 8.1% relative to the sample mean. Five years after a disaster, the effect reaches 24%. Non-green innovation remains unaffected, and we find no evidence of sizable spatial spillovers, suggesting that personal experience rather than media coverage or broad policy changes drives the result. The green innovation response is concentrated in emission-reducing mitigation technologies, consistent with inventors responding to the broader issue of climate change rather than to the specific disaster they experience.

To explore the underlying mechanism, we link natural disasters to the German part of the Community Innovation Survey. Disaster exposure raises the likelihood that firms cite expected increases in green demand and anticipated environmental regulation as motives for green innovation. Exposed respondents do not report reputational concerns or increased government funding as drivers. These findings suggest that personal experience affects innovation choices by altering higher-order beliefs about consumer demand and voters' preferred regulation, and thus expected profitability.

We formalize this channel in a quality-ladder model building on Aghion et al. (2023) that disentangles profitability expectations from intrinsic motivation. The model predicts that only profit-motivated inventors respond more strongly in competitive markets, where price pressure makes green differentiation more valuable. We test this comparative static empirically and find that the innovation response to disaster exposure is significantly larger in competitive markets, consistent with profitability expectations as the central channel.

This chapter shows that personal experience of climate events partially addresses the underinvestment in R&D by shifting inventors' expectations about the returns to green technologies. The response is mediated by market structure and operates through a profit motive rather than through intrinsic motivation alone, which implies that competitive markets and belief in anthropogenic climate change are jointly necessary for personal experience to translate into green innovation.

Paper 2: Welfare Weights and the Social Cost of Carbon. The second chapter, joint with Felix Bierbrauer and Niklas Taft, asks whether the social cost of carbon depends on society's distributive preferences between rich and poor. Cost-benefit analysis of climate policy relies on the social cost of carbon, the present value of future income losses from climate change. Prior work examines how this measure depends on the discount rate and on how we account for uncertainty about future damages (Stern 2007; Tol 2009; Nordhaus 2017; Känzig 2023). We show that it also depends on a distributive judgment.

We use a simple Mirrleesian setup. Two types of individuals differ in productive ability and consume a green good and a brown good. Aggregate consumption of the brown good creates an atmospheric externality. Our results show that optimal climate policy is equal across agents. In the optimum and everywhere on the Pareto frontier, all individuals have the same marginal rate of substitution between green and brown goods. We characterize the Pareto frontier using a mechanism design approach. Under the separability conditions of Atkinson and Stiglitz (1976), the standard prescription is that the Pigouvian correction equals the social cost of carbon (Cremer et al. 1998; Kaplow 2006) and implicitly is invariant to welfare weights. We show that this invariance rests on a specific combination of assumptions. Under homogeneous environmental preferences, linear damages, and a CES consumption aggregator, the optimal Pigouvian tax is constant along the entire Pareto frontier. As soon as one of these assumptions is relaxed, the location on the Pareto frontier matters for the level of the corrective tax.

Two channels generate the dependence. The first operates through the equilibrium level of brown consumption. Even when both types value clean air identically, the marginal damage from brown consumption depends on aggregate brown consumption, which itself depends on the planner's distributive choice. Under convex damages, a Rawlsian planner who compresses consumption to redistribute faces a lower marginal damage and implements a lower carbon tax. The second channel operates through heterogeneous environmental preferences. When the two types differ in how strongly they value clean air, the Pigouvian correction embeds a welfare-weighted average of the two damage valuations. A planner who places more weight on the type with weaker environmental concern implements a lower corrective tax. The two channels are logically distinct but reinforce one another when damages are convex and environmental preferences are heterogeneous.

We extend the analysis to a dynamic setting where the economy transitions from a brown to a green production technology at an endogenous switching time. The optimal timing of the green transition varies along the Pareto frontier in the same way as the static carbon tax. A planner who favors the environmentally less-concerned type delays the green transition and accepts a higher long-run pollution stock. A planner who favors the more-concerned type transitions earlier and stabilizes at a lower stock. This result contributes to the literature on integrated assessment and optimal dynamic carbon taxation (Nordhaus 1992; Golosov et al. 2014; Barrage 2020), which derives optimal

transition paths but takes the social welfare function as given. The question of when to abandon fossil fuels therefore embeds a distributional judgment that existing estimates of the optimal transition date leave implicit.

This chapter connects to the dissertation's broader theme by examining the normative sense of social concern. It shows that even within the standard separable framework of optimal taxation, where corrective and redistributive motives are cleanly separated, the corrective tax is not uniquely pinned down. The social cost of carbon varies with welfare judgment, and existing estimates are implicitly conditioned on a distributional choice that is never made explicit.

Paper 3: Morals and the Political Economy of Corrective Taxes. The third chapter, joint with Felix Bierbrauer, Mattias Polborn, and Georg Weizsäcker, studies the political economy of corrective taxes when consumers differ not only in their carbon footprints and damage estimates but also in whether they accept a moral obligation to reduce emissions beyond material self-interest. We model an economy with two groups. Neoclassical consumers take prices and total emissions as given and maximize individual consumer surplus. Socially responsible consumers internalize the social cost of carbon even without corrective taxes, modeled as group-rule-utilitarian agents who optimize collectively while taking the behavior of neoclassical consumers as given.

We derive preferences over corrective taxes for both groups. For neoclassical consumers, we adapt the approach from the political economy of redistributive taxation (Roberts 1977; Meltzer and Richard 1981). Each neoclassical consumer balances the mitigative effect of a Pigouvian tax and the distributive effect of rebated revenue against the loss of consumer surplus. For socially responsible consumers, we infer tax preferences from the value function of their group-rule-utilitarian optimization problem. The Pigouvian motive in their preferred tax applies only to the behavioral response of neoclassical consumers, since the socially responsible group's own behavior is already optimal. Whether the distributive component calls for taxes above or below the pure Pigouvian benchmark depends on whether socially responsible consumers have smaller carbon footprints than neoclassical ones on average. We show that an increase in the population share of socially responsible consumers can lead to lower equilibrium taxes and higher emissions, because the politically decisive neoclassical consumer adjusts her preferred tax downward for distributive reasons.

We then enrich the framework by allowing socially responsible consumers to exert moral pressure on neoclassical consumers to reduce their carbon footprints. The latter face a choice of whether to comply or to reject this pressure, a case we call moral rebellion. Moral pressure is a substitute for taxes. The lower the tax rate, the more intensely socially responsible consumers pressure neoclassical ones. When neoclassical consumers simply accept this pressure, equilibrium taxes rise and emissions fall relative

to the baseline without moral activism. However, when we allow moral rebellion with strategic complementarities, an equilibrium with widespread rejection of moral pressure can coexist with the compliant equilibrium. In the rebellious equilibrium, equilibrium taxes are lower than in the baseline case. We show that the rebellious equilibrium can be Pareto-inferior to the outcome that results when the options for moral pressure and moral rebellion are removed entirely.

We complement the theory with an online survey of representative US and UK samples. The survey documents three findings consistent with the model. Respondents who adjust their stated car usage in response to higher social costs of carbon, our proxy for socially responsible behavior, are more concerned about climate change, place less weight on carbon-intensive consumption, and prefer higher carbon taxes. This pattern holds conditional on the self-reported importance of polluting consumption and climate concern. Moreover, more socially responsible respondents approve of moral pressure while respondents for whom carbon-intensive consumption is important are more likely to view moral pressure as intrusive.

This chapter connects to the dissertation's theme of social concern in its positive sense by modeling heterogeneity in moral motivation and showing how it shapes both market behavior and political equilibrium outcomes. The result that moral activism can be counterproductive, generating moral rebellion that leaves all agents worse off, identifies a mechanism through which social concern interacts with the political process.

Synthesis. Taken together, the three chapters show that climate-relevant economic decisions depend on forces that are often neglected. Personal experience shifts the direction of inventive activity by changing what inventors expect about the returns to green technologies. Distributive judgments between rich and poor alter the social cost of carbon even within frameworks designed to separate corrective from redistributive policy. Moral heterogeneity shapes the political equilibrium of corrective taxation and can make well-intentioned social pressure counterproductive. The common finding is that the responses of individuals to climate change, whether through innovation, through market behavior, or through voting, are governed by beliefs, preferences, and moral commitments that vary across the population and change over time. Climate policy that abstracts from this heterogeneity misses both the endogenous forces that work in its favor and the political constraints that work against it.

Overview

Inventors' Personal Experience of Natural Disasters and Green Innovation

This chapter is co-authored with Lisa Keding and contains the current version of the working paper of the same title (Keding and Ritterrath 2025). The initial idea was jointly developed by both authors. I conceptualized and developed the theoretical model. The identification strategy exploiting natural disasters as exogenous variation was jointly conceived. I took the lead on the detailed econometric implementation, including the difference-in-differences design, robustness checks, and extended analyses. Lisa Keding was responsible for the survey-based analysis. Data cleaning was conducted jointly. I wrote the initial draft and shaped the overall conceptual narrative of the paper.

Welfare Weights and the Social Cost of Carbon

This chapter is co-authored with Felix Bierbrauer and Niklas Taft. The conception and operationalization of the research question as well as the introduction were jointly developed by all three co-authors. Niklas Taft and I jointly characterized the solutions to the underlying optimization problems, calibrated the model, and conducted a complete characterization of Pareto-efficient tax systems using numerical methods. I designed, and implemented the dynamic extension of the model, including the theoretical formulation of the dynamic framework, the derivation of the corresponding optimality conditions, and the numerical implementation and analysis.

Morals and the Political Economy of Corrective Taxes

This chapter is co-authored with Felix Bierbrauer, Mattias Polborn, and Georg Weizsäcker. The chapter combines a theoretical analysis with a survey study. The theory characterizes political equilibria under several scenarios. I calibrated the model and conducted the numerical analysis of these scenarios. This analysis was a prerequisite for completing the theoretical results. The survey was conceptualized jointly by all authors. I took the lead in its implementation and analysis.

Chapter 1

Inventors' Personal Experience of Natural Disasters and Green Innovation

Joint with Lisa Keding

Abstract

We show that personal experiences affect high-stakes economic decisions among inventors. Using matched patent and survey data from French and German inventors linked to natural disaster records, we exploit exogenous variation in disaster exposure, encompassing both direct and indirect exposure. Inventors who experience natural disasters subsequently produce 8.1% more green patents, primarily driven by emission-reducing mitigation technologies, while non-green innovation remains unaffected. The absence of sizable spatial spillovers highlights the importance of personal experience. The personal experience of natural disasters shapes innovation choices by altering profitability expectations through shifting higher-order beliefs about consumer demand and anticipated regulation. Embedding this channel in a formal model, we disentangle the role of expectations and intrinsic motivation. The model predicts, and the data confirm, that effects are strongest in competitive markets, where profit incentives matter most.

Acknowledgements. We thank Philippe Aghion, Antonin Bergeaud, Felix Bierbrauer, Anna Bindler, Pierre Boyer, Oana Branzei, Fabian Gaessler, Yves Le Yaouanq, Ulrike Malmendier, Lewin Nolden, Chris Roth, Sebastian Siegloch, Nico Voigtländer, Romain Wacziarg, Johannes Wohlfart, and Yanos Zylberberg for helpful comments. We also thank participants of the 2025 EEA Congress, Summer Conference of the Association of Environmental and Resource Economists, the ARCS Annual Conference, the 2024

RISE Workshop, the ARCS/Ivey PhD Sustainability Academy, ZEW/MaCCI Conference on the Economics of Innovation and Patenting, the Annual Meeting of the Academy of Management, the Annual Conference of the European Association of Environmental and Resource Economists, EAERE Summer School, European Meeting of the Urban Economics Association, Warwick Economics PhD Conference, Royal Economic Society Annual Conference, and RGS PhD Conference for their feedback. Further, we thank various seminar participants at the University of Cologne, ENSAE/CREST, UCLA Anderson, UC Berkeley, DIW, the University of Maastricht, and the Collège de France, PSE, INSEAD Innovation Seminar. We also thank Carl Benedikt Hüls for his excellent research assistance.

Lisa Keding: RWTH University — Supported by the Joachim Herz Foundation.
Marten C. Ritterrath: University of Cologne — Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2116/1-390838866; supported by the Joachim Herz Foundation and C-SEB.

1.1 Introduction

What drives inventors to pursue new ideas? Popular accounts often point to moments of personal frustration that make unmet needs visible: Netflix is linked to an anecdote about a $40 late DVD return fee, while BioLite emerged when its founders sought a cleaner, lighter wood-burning camp stove instead of carrying heavy fossil-fuel canisters.[1] Such stories suggest that invention may often arise from personal experience. A rich literature links innovation to market conditions such as competition (Aghion and Howitt 1992; Aghion et al. 2005) and market size (Acemoglu 2002; Acemoglu and Linn 2004), but much less is known about inventor-level determinants. Social background and childhood socioeconomic status influence who becomes an inventor (Einiö et al. 2025; Bell et al. 2019), but we know little about what ultimately drives their choices about what to innovate. Addressing this gap is particularly relevant given persistent underinvestment in R&D despite its large social returns (Bloom et al. 2013; Azoulay et al. 2019; Jones and Summers 2020; Jones 2021), and the role of innovation in combating climate change (Acemoglu et al. 2012; Acemoglu et al. 2016).

In this paper, we bring insights from the experience effects literature to the production side and explore how personal experiences shape inventors' choices. The experience effects literature has shown that households' personal experiences shape beliefs, preferences, and expectations (e.g., Malmendier and Nagel 2011; Malmendier and Nagel 2016; Giuliano and Spilimbergo 2025). Expectations, in particular about commercial success and profitability, matter for inventors, as pursuing innovation is inherently uncertain and risky (Rosenberg 1998; Bloom 2007; Manso 2011). Personal experiences plausibly play a role in forming these expectations.

We exploit natural disasters as exogenous shocks to inventors' personal experience. In this context, we refer to personal experience as the full range of possible exposure to a regional natural disaster. This includes being physically affected, witnessing destruction, interacting with affected individuals, or navigating the economic and social disruptions that disasters impose on local communities. This quasi-experimental variation allows us to identify causal effects on subsequent green innovation. Green technologies either mitigate or adapt to climate change. To explore the underlying mechanisms, the second part of the paper uses survey data to examine how disaster exposure influences inventors' expectations about consumer demand for green goods and environmental regulation. In the third part of the paper, we propose a formal model that incorporates our behavioral mechanism into a quality-ladder model, and disentangles the roles of intrinsic motivation and profitability expectations.

We begin our analysis by linking French and German inventors' home addresses to

[1] Based on interviews with Netflix's Reed Hastings (CBS News 2006) and BioLite's Alec Drummond and Jonathan Cedar (Bastone 2018).

natural disasters and use event-study designs to estimate the effects of natural disaster exposure on subsequent innovation. Our design exploits random variation in the timing of inventors' exposure to natural disasters. We use patent data from 1994 to 2014 with geocoded addresses from De Rassenfosse et al. (2019). Disasters are matched spatially using the Emergency Events Database (Guha-Sapir et al. 2022), supplemented with geolocations from the Geocoded Disasters dataset (Rosvold and Buhaug 2021). We identify green patents using an established classification that tags technologies aimed at climate change mitigation or adaptation. Most inventors in our sample are affiliated with firms (72%), while a substantial share (16.8%) operates independently.

Our results show that natural disaster exposure leads to a significant increase in green innovation. On average, one additional natural disaster increases green patenting by 8.1% relative to the sample average. This impact evolves dynamically over time. Notably, five years after a disaster, the number of green patents rises by 24%, with the effect gradually tapering off thereafter. This 24% increase corresponds to 0.61 additional green patents in the affected region. In terms of magnitude, our estimate is comparable to the effect of the European Emissions Trading System (EU-ETS), which increased green patenting among covered firms by about 10% (see Calel and Dechezleprêtre 2016).[2]

We find that disaster exposure only affects climate-related innovation, consistent with a salient link between personal experiences of natural disasters and climate change. Disaster exposure does not lead to significant changes in non-green innovation, suggesting limited crowding out of other inventive activity. Green innovation increases significantly more following more deadly disasters, consistent with the idea that the salience of climate-related risks increases with disaster fatalities (Demski et al. 2017; Kalatzi Pantera et al. 2023). The green innovation response appears tied to the broader issue of climate change rather than to the specific disasters inventors experience. Disaster exposure increases patenting in mitigation technologies by 8.5 %. These technologies aim to reduce greenhouse gas emissions and address the root causes of climate change, and are not tied to the type of event. In contrast, directly related adaptation technologies that address local climate impacts respond less strongly (4.4 % increase).

Inventors respond only when personally affected; we find no evidence of large spillovers from natural disasters to nearby regions. The absence of sizable spillovers suggests that the effect operates through local exposure rather than through channels that extend beyond affected areas, such as media coverage or broad policy changes. The innovations that result from these local personal experiences have global reach and value, being both highly cited and frequently triadic. Triadic patents are those filed in the United States, Europe, and Japan, and are often used as a proxy for the most valuable technologies (Dernis and Khan 2004; Rassenfosse and Pottelsberghe de la Potterie 2009).

[2]The two treatments differ in nature. The EU-ETS is a continent-wide, permanent policy, whereas disaster exposure is localized and transitory.

To explore the mechanisms behind this increase in green innovation, the second part of the paper links natural disasters to the German part of the Community Innovation Survey (CIS), which asks research personnel in German firms about recent innovations and their motivations. While the survey data cannot isolate the individual-level experience of each inventor, they allow us to observe how firms in disaster-affected areas report the drivers of their green innovations.

Natural disaster exposure alters higher-order beliefs (beliefs about others' beliefs) regarding consumers' climate change attitudes, their consumption preferences and voting behavior. Concretely, experiencing one additional disaster raises the likelihood of citing expected increases in green demand as a motive for green innovation by 0.87 percentage points. It also increases the likelihood of citing anticipated environmental regulation by 0.93 percentage points. Exposed respondents do not report reputational concerns or increased government funding and subsidies as drivers of their green innovation. As an additional check, we match French administrative data on government research funding to our disaster exposure measure and find no systematic increases in funding for affected regions. We find that economically more damaging disasters produce weaker innovation responses, which also speaks against physical destruction and reconstruction as the primary channel. Our results are consistent with a mechanism that primarily goes through expectations about the future profitability of green innovation. Natural disasters locally increase the salience of climate change, shifting consumption toward environmentally friendly products, and thus raise expectations of green demand and regulation.[3] As a result, inventors expect higher returns to green innovation and increase their R&D efforts accordingly.

Additionally, the CIS data allow us to capture innovations beyond those recorded in the patent system, which we use to confirm that affected firms are more likely to introduce green products and engage in green process innovation. Exposure increases the likelihood of introducing green process innovation and green product innovation by roughly 4.3% and 4.7%, respectively. Moreover, self-reported climate change affectedness is robustly correlated with our natural disaster measure.

In the third part of the paper, we embed our behavioral mechanism in a formal theoretical model to examine how it interacts with market forces, allowing us to disentangle profitability expectations and intrinsic motivation. Building on Aghion et al. (2023), consumers value both the consumption utility and the carbon footprint of goods, for example, transportation and its associated emissions. Consumer preferences depend on beliefs about climate change. We extend this framework by modeling how inventors respond to the heightened salience of climate change. Inventors form expectations about the future profitability of green technologies and derive intrinsic utility from engaging in

[3]Disaster exposure increases green good demand & environmental policy preferences (Djourelova et al. 2024, Chae et al. 2025).

green research. Both expectations about profitability and intrinsic motives are shaped by local personal experiences with natural disasters. Inventors operate in markets with varying degrees of competition, which shapes how strongly they respond to changes in expected demand for green technologies. In monopolistic markets, inventors have little incentive to develop green alternatives, as incumbents already earn high profits. In contrast, in competitive markets, inventors face price pressure and can use green innovation to differentiate their products and escape competition.

Our model predicts that only inventors motivated by profitability expectations respond more strongly to natural disaster exposure when they operate in competitive markets. This comparative static provides a test of our proposed mechanism. Since market structure affects only the profit motive and not inventors' intrinsic motivation, a stronger innovation response in competitive markets would indicate that profit incentives play a central role. In contrast, if intrinsic motivation were the sole driver, inventors would respond equally regardless of the level of competition.

We test this comparative static empirically by matching our patent data with information on industry competition, and find evidence consistent with profitability expectations playing a central role. Following Aghion et al. (2023), we use inverse profit margins as a proxy for industry-country-year-level competition.[4] We find statistically significantly larger effects in markets with high levels of competition, in line with our model predictions for profit-oriented inventors. This highlights the importance of inventors' personal experiences in shaping profit expectations and ultimately affecting their innovation choices.

Our model predicts stronger responses in relatively larger green good markets, in line with the "building on the shoulders of giants" insight from the literature on market size and the direction of innovation (Acemoglu et al. 2012; Acemoglu et al. 2016), as well as research on innovation responses to market size more generally (Aghion et al. 2024). In larger markets, innovation becomes relatively cheaper because there is a richer base of existing knowledge on which to build.[5] We also take this to the data and find stronger effects of natural disaster exposure in product markets with larger green good shares.[6]

Related Literature & Contribution: Our central contribution is to causally identify personal experience as a driver of innovation. We highlight a novel channel: personal experiences influence inventors' profitability expectations. Disaster exposure shifts expectations about future green good demand and environmental regulation. It alters higher-order beliefs about consumers.

Our findings add to a small but growing literature that studies inventor-level drivers

[4]Our competition data come from CompNet (CompNet 2022).

[5]For example, inventing an induction stove is more feasible when the principles of magnetic induction are already well understood, reducing the cost and risk associated with developing new green technologies.

[6]We proxy for green good demand by the share of green goods in a product market, using data from PRODCOM and a list of green products from Bontadini and Vona (2023).

of innovation. Recent work has investigated how socioeconomic background affects who becomes an inventor, and documents unequal access to innovation careers (Aghion et al. 2017; Akcigit et al. 2017; Bell et al. 2019). There is also work that documents social and intrinsic motives to innovate (Stern 2004; Einiö et al. 2025). Our work discusses inventors' personal experiences as central to their choice about what to innovate. These experiences affect inventors' choices through changes in their beliefs about the returns to innovation.

Our work builds on the behavioral experience effects literature, which documents correlations between personal experiences and household expectations about inflation, recessions, house prices, and stock returns (Malmendier and Nagel 2011; Malmendier and Nagel 2016; Kuchler and Zafar 2019; Laudenbach et al. 2024). Giuliano and Spilimbergo (2025) summarize this literature. We contribute to the literature on experience effects in three ways. First, we move beyond correlations and show that experience effects causally influence high-stakes decisions. Second, we extend the literature to the production side of the economy by showing that personal experiences affect innovation choices. Third, we disentangle the channels through which personal experiences operate, distinguishing intrinsic motivation from profit expectations.

Our findings relate to the literature on higher-order belief formation and salience. Evidence on higher-order belief formation has mostly been limited to experimental studies and information treatments in surveys (Coibion et al. 2021).[7] Our findings provide novel evidence from an observational setting on how personal experiences of large shocks shape higher-order beliefs. These findings also connect to the literature discussing the role of salience in decision making (Bordalo et al. 2012; Bordalo et al. 2022). Prior work related to natural disasters shows that they heighten local climate change salience and change voting and consumption patterns (Gallagher 2014; Herrnstadt and Muehlegger 2014; Djourelova et al. 2024; Chae et al. 2025). We show that inventors respond to these local changes, and that the salience of climate change affects high-stakes innovation decisions.

Lastly, we contribute to the literature on the determinants of the direction of innovation. Green R&D responds to market structure and policy incentives such as competition, energy-price shocks, and carbon regulation (Aghion et al. 2005; Acemoglu et al. 2012; Acemoglu et al. 2016; Hassler et al. 2012; Calel and Dechezleprêtre 2016). There is a small literature that links the invention of adaptation technologies such as drought-resistant crops and air-conditioning to natural disasters (Miao and Popp 2014; Barreca et al. 2016; Moscona and Sastry 2023). However, how innovation responds in the crucial domain of mitigation, technologies that directly combat climate change by reducing emissions, is poorly understood. We address this gap and show that natural disaster exposure

[7]Higher-order beliefs shape decision-making and coordination in markets. This layering of expectations helps explain diverse phenomena, from fluctuations in economic activity and asset price movements to behavioral distortions like myopia and anchoring (Lorenzoni 2009; Banerjee et al. 2009; Angeletos and Huo 2021; Huo and Takayama 2025).

leads to inventive activity that goes beyond damage control and targets forward-looking abatement. Our results show that there is an endogenous response to increasingly severe climate change, which integrated assessment models neglect (see for example Ackerman et al. 2009; Cai 2020). Crucially, this response depends on consumers' green preferences and the market environment.

The rest of this paper is organized as follows. Section 1.2 describes our data. Section 1.3 describes our empirical approach. Section 1.4 presents our results on patenting. Section 1.5 discusses our proposed mechanism and provides survey evidence on inventors updating their higher-order beliefs. Section 1.6 starts with our theoretical model and provides empirical evidence for our proposed comparative statics. Section 1.7 concludes.

1.2 Data

To conduct our analysis we draw on a variety of different data sources. Table 1.1 gives an overview of the data, the geographic coverage, the time period, and the key variables we use.

Table 1.1: Data Sources, Coverage, Periods, and Usage

Data & Source	Countries	Time Period	Usage
Patents (PATSTAT)	France, Germany (also data for EPO, USPTO, JPO)	1994–2014	Patent indicators; technology classification; inventor geolocations
Natural Disasters (EM-DAT & GDIS)	France, Germany	1980–2018	Location; severity; type
Firm-level Survey (Mannheim Innovation Panel)	Germany	2009, 2015, 2021	Firm-level green innovation measures; stated reasons for green innovation
Competition (CompNet)	France, Germany	From 2000	Industry-level competition measure (available years vary by country – see Appendix 4.3.1)
Green Goods (Eurostat PRODCOM)	France, Germany	1995–2014	Industry-level production share of green products, green goods from Bontadini and Vona (2023)
Research Funding (ScanR)	France	1999–2023	Public research and innovation funding

1.2.1 Patents

To measure innovation, we use data on patent applications filed by inventors living in France and Germany from the European Patent Office's (EPO) PATSTAT database, covering the period from 1994 to 2014. Roughly 50% of patents filed at the EPO came

from EPO member states. Of those, 50% came from France and Germany. They are the two most active countries among EPO member states in terms of patenting. We have additional information on research funding for France, and a firm-level survey on innovation for Germany. Hence, we focus our analysis on France and Germany.

Given our focus on the personal experiences of inventors rather than on patents themselves, it is crucial to have comprehensive data on inventors' patenting behavior regardless of where they choose to file these patents. Because we study inventors residing in France and Germany rather than patents filed solely within these countries, we require access to patent records extending beyond the European Patent Office (EPO). For example, a French inventor personally affected by a natural disaster may choose to patent an innovation in the United States, anticipating greater commercial potential there. PATSTAT meets this requirement by providing extensive data that includes filings not only at the EPO and its member states' national offices but also at global patent offices such as the Japanese Patent Office (JPO) and the United States Patent and Trademark Office (USPTO). Patents filed in these three jurisdictions are often referred to as triadic patents, an established indicator of high-value inventions (see, e.g., Dernis and Khan 2004; Rassenfosse and Pottelsberghe de la Potterie 2009; Dechezleprêtre et al. 2017).

We are interested in green technologies, which are innovations addressing climate change through mitigation or adaptation. To identify green patents, we rely on PATSTAT's Cooperative Patent Classification (CPC) data, an extended version of the International Patent Classification (IPC). CPC features the Y02 classification, explicitly denoting technologies designed for mitigating climate change effects or adapting to its impacts. Utilizing this classification, we distinguish between mitigation patents, which are aimed at reducing the environmental impact of human activities, and adaptation patents, which are designed to help societies better cope with climate change. The detailed CPC classes available in PATSTAT also allow us to categorize patents broadly into technological groups such as agriculture, concrete and cement making, or combustion engines. This detailed classification enables us to construct precise indicators representing a region's technological specialization. We use this for a robustness check. Figure 1.1 illustrates patenting activity from 1994 to 2014 in France and Germany, distinguishing green from non-green patents. Green patents comprise approximately 7.7% of all patents filed during this period, and this share gradually increased over time.

We supplement PATSTAT with detailed information on the location of inventors and applicants. Specifically, we use data from De Rassenfosse et al. (2019), which provides precise coordinates for each inventor's and applicant's primary place of residence at the time of patent filing. This data roughly corresponds to city-level assignments, enabling us to link all patents in our sample to the location of their inventors. The data are available up to 2014, limiting our analysis to years up to and including 2014.

From the available patent documents, we restrict our sample to first filings, i.e., the

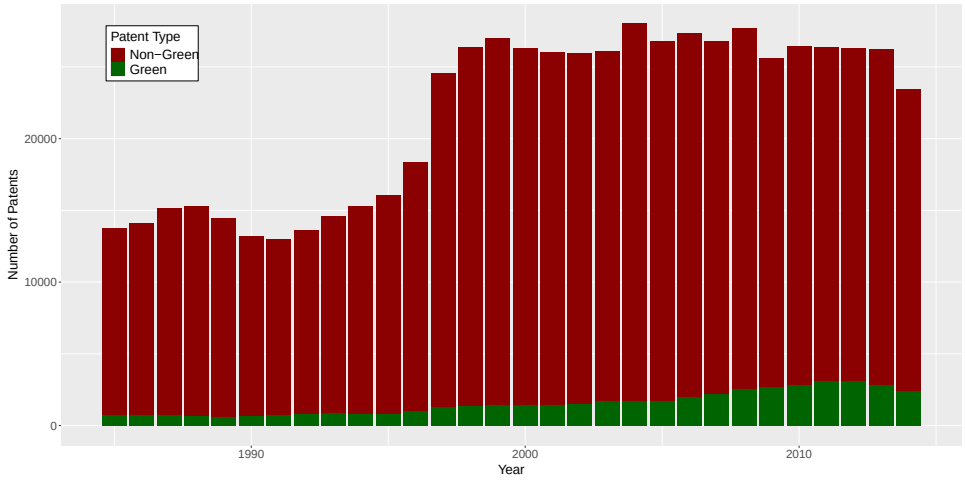


Figure 1.1: Patenting Activity in France and Germany over Time

Note: This figure depicts how patenting in France and Germany evolves over time until 2014. We split patents by green and non-green patents.

first time any application was made for a distinct invention within an EPO worldwide bibliographic data (DOCDB) simple patent family. All patents in a simple patent family are considered to cover the same technical content and share the same priority. In other words, they describe the same invention and represent the same technological advancement. A priority patent is the first patent filed for that specific invention. We use the priority date, the date of the first filing, as the year in which an invention was filed. Later claims or modifications to patent claims are excluded from our dataset, as we focus solely on original inventions. Thus, we do not count instances where an existing patent is subsequently filed in another jurisdiction. Similarly, we include only granted patents to ensure that what we measure is a true new invention.

1.2.2 Natural Disasters

We obtain information on natural disasters from the emergency events database (EM-DAT) published by the Centre for Research on the Epidemiology of Disasters (CRED) (see Guha-Sapir et al. 2022). In our analysis, we are interested in the emergence of innovations to combat and mitigate the consequences of climate change. We therefore only consider natural disasters that occur more frequently in France and Germany due to climate change. For France and Germany, these are floods, storms, extreme temperature events, and droughts (see Intergovernmental Panel on Climate Change (IPCC) 2023). The CRED includes a disaster in the database if it meets at least one of the following conditions: (a) a death toll of ten or more people, (b) there are at least 100 people affected by the disaster, (c) the disaster causes the declaration of a state of emergency, or (d) the affected country calls for international assistance.[8]

We complement the EM-DAT data with geolocations from the Geocoded Disaster

[8]See `https://doc.emdat.be/docs/protocols/entry-criteria/` for the precise inclusion criteria.

(GDIS) dataset (Rosvold and Buhaug 2021), which provides detailed spatial information on disasters from 1960 to 2018. GDIS assigns events to the most precise available administrative unit, ranging from national to subnational levels. In France and Germany, this includes up to three tiers of administrative divisions: régions and Länder (large federal or territorial states), départements and Regierungsbezirke (mid-level units akin to provinces), and arrondissements and Kreise (smallest administrative units, comparable to US counties). The arrondissement and Kreis levels correspond to the third Nomenclature of Territorial Units for Statistics (NUTS) level in France and the second NUTS level in Germany. There are 403 Kreise and 350 arrondissements in our data.

Until 1984, only a small number of disasters are geolocated below the first administrative level. We therefore only consider events from 1984 onward. For our event study, we use 10 years of lags and 4 years of leads, and use disaster data from the years 1984 to 2018. For our full sample, we end up with 150 distinct natural disasters, some of which affect multiple regions at once.

1.2.3 Analysis Sample - Patenting

We merge our patent and disaster data at the most granular administrative level in our data, arrondissements in France and Kreise in Germany. Throughout the analysis, we refer to this as the "regional" level or simply "region," which should not be confused with the French "région."

Our dataset includes approximately 520,000 patents, of which around 40,300 are classified as green. These patents were filed by approximately 1,385,000 and 110,000 inventors, respectively. We aggregate all patents by the region of their inventors. Since some patents have multiple inventors with addresses in different administrative areas, we assign each region a proportionate share of the patent. For instance, consider a patent i with three inventors, where two live in Region A and one in Region B. Patent P would then be attributed with a share of 2/3 to Region A and 1/3 to Region B. More generally, to calculate the count of all green patents in region l in year t, we sum over all patents i, weighting by the share of i's inventors residing in region l:

$$C(\mathbf{Y02}_{lt}) = \sum_{i}^{N} \left(\frac{[\mathbf{Y02}_{ilt} = 1]}{\sum_{l}^{L} [\mathbf{Y02}_{ilt} = 1]} \right) \tag{1.1}$$

where

$$[\mathbf{Y02}_{ilt} = 1] = \begin{cases} 1, & \text{if patent } i \text{ in year } t \text{ and region } l \text{ is green (Y02)} \\ 0, & \text{otherwise} \end{cases} \tag{1.2}$$

This yields a continuous (in fraction of counts) variable for the annual number of green

and non-green patents in each region. On average, 33.04 patents are granted annually in each region, 2.54 of which are green patents. To ensure comparability between green and non-green patents, we normalize the count of each type of patent in each region by its respective mean across all years t and all regions l:

$$P(\mathbf{Y02}_{lt}) = \frac{C(\mathbf{Y02}_{lt})}{\frac{1}{L}\sum_{l}^{L}\frac{1}{T}\sum_{t}^{T} C(\mathbf{Y02}_{lt})} \tag{1.3}$$

where L is the number of regions and T is the number of years, and equivalently for non-green patents. We adopt the same normalization procedure when aggregating across subclasses or when splitting the sample by e.g. competition.

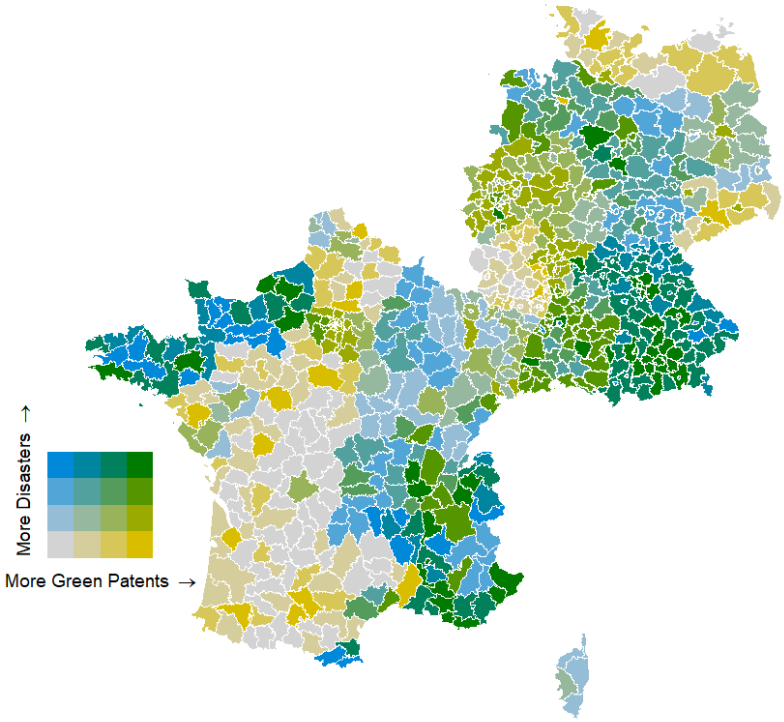


Figure 1.2: Bivariate Map of Green Patent and Disaster Counts

Note: This map depicts the bivariate variation in green patenting (shades of yellow) and natural disasters (shades of blue). Darker tones signify more patents (disasters) respectively. Shades of green indicate regions with high levels of both green patents and natural disasters. Grey regions have a very low number of patents and disasters. Only green patents are displayed and we pool all years.

Natural disasters are reported at either the first-, second-, or third-order administrative level. To ensure consistent spatial coverage, we assign each disaster reported at the first- or second-order level to all corresponding third-order areas within the respective administrative boundary. For instance, if an extreme temperature event is reported in the German state Hessen, all 26 Kreise within Hessen are coded as being exposed during this period. Our sample includes 150 natural disasters in total. Broken down by type, there are 64 floods, 63 storms, 20 extreme temperature events, and 3 droughts.

Figure 1.2 visualizes the geographic variation in both disaster exposure and green patenting activity across the regions in our sample. Regions shaded in yellow are char-

acterized by elevated levels of green patenting but relatively few natural disasters, while those shaded in blue have experienced many disasters but exhibit limited green innovation. Green-shaded regions display both high disaster exposure and high green patenting, indicating potential alignment between environmental shocks and green innovation. In contrast, gray areas denote regions with neither significant disaster exposure nor notable green patenting activity.

1.3 Empirical Strategy

This section outlines our empirical strategy for estimating the effect of natural disasters on innovation. Innovation is a gradual process, as the path from idea to patentable prototype often spans years. To capture these dynamics, we adopt an event-study design, which allows us to observe the dynamics of patenting following a natural disaster. Our baseline event-study specification, applied to the data, is presented in the equation:

$$P(\mathbf{Y02}_{lt}) = \sum_{s=-5, s\neq -1}^{11} \beta_s D_{l,t}^{s} + \lambda_{c(l),t} + \lambda_l + \epsilon_{lt} \tag{1.4}$$

with

$$D_{l,t}^{s} = \begin{cases} \sum_{s=-\infty}^{-5} d_{l,t-s} & \text{if } s = -5 \\ d_{l,t-s} & \text{if } -5 < s < 11 \\ \sum_{s=11}^{\infty} d_{l,t-s} & \text{if } s = 11 \end{cases} \tag{1.5}$$

where $P(\mathbf{Y02}_{lt})$ is the normalized count of patents (as specified in equation (1.3)), and $d_{l,t-s}$ is the count of natural disasters experienced by region l in year $t-s$. The reference period is the year prior to disaster exposure. Following McCrary (2007) and the formal definition of Schmidheiny and Siegloch (2023), we bin all periods that are more than 10 years in the past or more than 4 years in the future. We include region fixed effects λ_l and country-by-year fixed effects $\lambda_{c(l),t}$, where $c(l)$ denotes the country that region l belongs to.

Region level fixed effects (λ_l) control for region-specific natural disaster risk characteristics and account for differences, such as one region being more accustomed to floods than another. Although the exact timing of natural disasters is random, some regions experience such events more frequently, as shown in Figure 1.2: coastal and Alpine areas face higher disaster risks than interior regions. Inventors might choose locations based on these regional risks, potentially causing selection bias. Region fixed effects mitigate this concern by capturing region-specific disaster exposure and institutional characteristics that are constant over time. Additionally, country-by-year fixed effects ($\lambda_{c(l),t}$) account for trends in overall disaster risk and differing innovation patterns between France and

Germany.

Our coefficients of interest, β_s, estimate the average change in green innovation in a region at event time s, relative to the year before the disaster ($s = -1$), controlling for time-invariant regional characteristics and differential trends across countries. Identification comes from cross-regional variation in disaster timing. Because our natural disaster data are only available at the regional level, our estimates resemble an intent-to-treat effect. If a flood impacts only part of a region (such as a valley), we still treat all inventors in that region as exposed. To the extent that some of these inventors were not personally affected, this attenuates the estimated effect and yields a conservative lower bound. However, personal experience of a disaster extends beyond direct physical exposure to include witnessing destruction, interacting with affected individuals, and navigating the economic and social disruptions that disasters impose on local communities. Regional-level treatment assignment may therefore capture much of the relevant exposure, particularly in the relatively small administrative units in our sample. We cluster standard errors at the regional level, which is the level of treatment variation.

One remaining identification concern is that climate change may induce differential trends in underlying disaster risk across regions. If these trends are observable to inventors, they may select into regions accordingly, generating potential selection bias based on heterogeneous regional risk trajectories. We argue that the absence of such selection is plausible in our setting, as the regional impact of climate change is inherently difficult to predict, even for climate scientists (Hulme et al. 1999). Moreover, the literature on migration patterns following natural disasters mostly documents out-migration of skilled individuals (Boustan et al. 2020), and inventors are highly skilled (Bell et al. 2019). If out-migration of inventors occurred in our setting, our estimates would be biased downward and suffer from attenuation. If the opposite was true, and inventors moved to affected regions, we would overestimate the effect. The underlying assumption of our work is that inventors do not select into regions based on regional differences in the trend of natural disasters.

We observe the universe of patent applications, which allows us to identify a subset of inventors who previously filed in the same region. In Section 1.4.5, we show that our results remain robust when restricting to this subsample (see Figure 1.6b), which alleviates concerns about inventor selection. A remaining caveat is that we observe inventors only at the time of filing, so unobserved moves between filings cannot be entirely ruled out. Additionally, in Section 1.4.7, we show that self-reported climate change affectedness is robustly correlated with our natural disaster measure.

In our context, the stable unit treatment value assumption (SUTVA) implies that there are no unmodeled spillovers between regions. We will later explicitly model spillovers to neighboring regions and show that natural disaster exposure only marginally affects

directly adjacent regions.[9]

A concern in our event study framework is the potential presence of heterogeneous treatment effects. Standard two-way fixed effects (TWFE) estimators, as used in our baseline specification, implicitly average treatment effects across groups that may differ in the timing and magnitude of treatment, potentially leading to biased or misleading estimates when treatment effects are heterogeneous. In our setting, regions are exposed to different numbers of disasters at different points in time, and the effect of an additional disaster may vary depending on prior exposure. This dynamic poses a challenge. Regions used as controls at a given point may themselves become treated in subsequent periods, and their treatment effects may not be comparable to the newly treated units.

To address this concern, in Section 4.2.1 of the Appendix we adopt the estimator proposed by Chaisemartin and D'Haultfœuille (2023) and Chaisemartin and D'Haultfœuille (2024a). This estimator constructs control the group by conditioning on treatment history. Specifically, regions that have experienced the same cumulative number of disasters up to $t-1$ form the control group, and those newly treated at t are compared against those that remain untreated at that time. Over time, as more units receive treatment, the control group shrinks. Our baseline model on the other hand, maintains a fixed control group. While the alternative estimator yields somewhat larger and more persistent effects (Figure 4.2 in Section 4.2.1 in the Appendix), the results are qualitatively consistent with those from our preferred TWFE specification (Figure 1.3 in Section 1.4). This robustness across estimation strategies provides reassurance that our main findings are not driven by bias due to heterogeneous treatment effects, even if each approach carries its own limitations. We mainly report the TWFE results because they provide more conservative estimates, and also potentially offer greater external validity due to a more realistic and stable control group.

We are also interested in the long-run average effect that one additional disaster has on green innovation in a region l. To estimate this effect, we use the following collapsed difference-in-differences equation:

$$P(\mathbf{Y02}_{lt}) = \beta \left(\sum_{s=0}^{\infty} d_{l,t-s} \right) + \lambda_{c(l),t} + \lambda_l + \epsilon_{lt}, \tag{1.6}$$

where $\sum_{s=0}^{\infty} d_{l,t-s}$ represents the cumulative number of past natural disasters. The parameter of interest, β, estimates the average effect that one additional disaster has on the number of green patents in a region.

To summarize, our identifying variation comes from the random timing of severe natural disasters across regions, which, conditional on regional fixed effects, is plausibly

[9]The effect is 1/8 of our baseline estimate for the effect on the directly exposed region. See section 1.4.2 for our results on spillovers. In general, positive spillovers would mean attenuation of our baseline estimates.

exogenous and allows us to compare changes in green innovation in affected regions before and after disasters relative to unaffected regions. Our data contain only severe natural disasters. Therefore, we caution that not all exposure to the forces of nature induces changes in inventor behavior.

1.4 Effects on Innovation

Figure 1.3 presents our results when estimating our event-study specification (1.4) for green and non-green innovation. Year 0 represents the partially treated year.[10] While the initial effect is small, we observe a large and significant impact two years after the natural disaster, with the effect peaking five years after the event. Five years after natural disaster exposure, green patenting is 24% higher than in unaffected regions. Subsequently, the effect diminishes over time, becoming insignificant ten years after the natural disaster. We interpret the inverted U-shape of the innovation response as stemming from the fact that innovation takes time. Natural disasters trigger an impulse towards inventive activity, with the resulting innovations materializing in the subsequent years. This pattern aligns with earlier literature, which suggests that the salience and behavioral response to natural disaster exposure tend to fade over time (see Gallagher 2014). The initial lag is consistent with innovation taking time. We find no significant change in non-green innovation, consistent with limited crowding out of other inventive activity.

The magnitude of our effects is substantial. In comparison, Calel and Dechezleprêtre (2016) find a 10% increase in green patents among firms covered by the EU Emissions Trading System (EU-ETS) in its early years, using the same measure of green innovation we employ. Relative to their estimates, we observe a larger short-run spike in patenting five years after a disaster, while our long-run effect, an 8.1% increase relative to the sample average (Table 1.2), is of similar magnitude but slightly smaller. The oil crisis led to a 3% rise in green patenting (Popp 2002; Hassler et al. 2012), and a 10% increase in fuel prices generated a 10% increase in green innovation (Aghion et al. 2016). By this benchmark, natural disasters elicit a large and statistically significant response from local inventors. That said, these comparisons should be interpreted cautiously. The EU-ETS and oil price shocks are systemic, partly predictable, and in the case of the EU-ETS permanent, whereas disaster exposure is localized, transitory, and inherently unpredictable.

The event study plots indicate flat pre-trends, with pre-treatment coefficients closely centered around zero, supporting the validity of the parallel trends assumption. To address concerns about heterogeneous treatment effects, in Section 4.2.1 we re-estimate our main green innovation results using the estimator proposed by Chaisemartin and D'Haultfœuille (2023) and Chaisemartin and D'Haultfœuille (2024a). Section 4.2.2 in

[10] If region l experienced a natural disaster in June, only patents filed in the months after could potentially be influenced by the natural disaster.

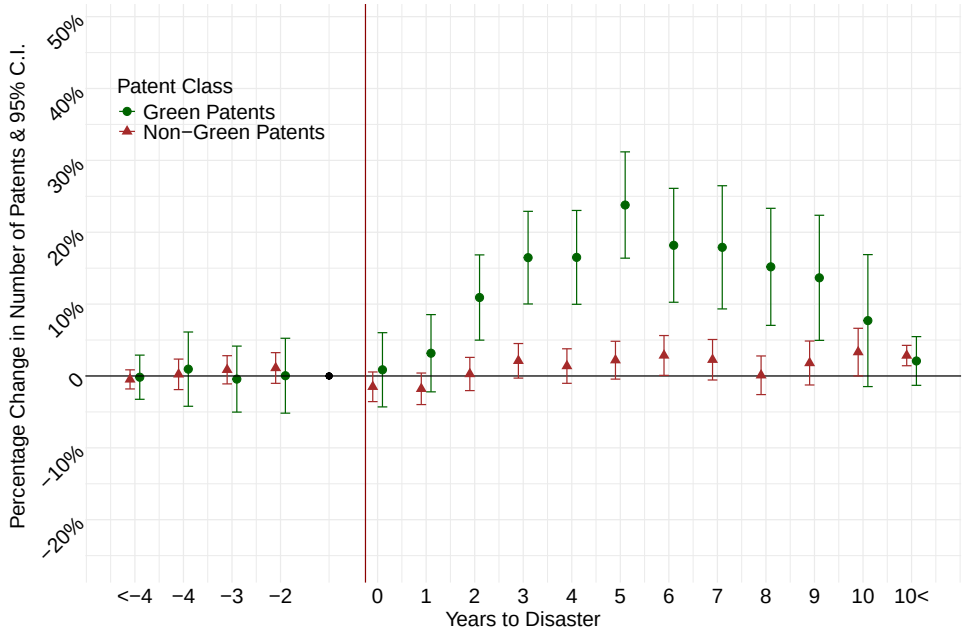


Figure 1.3: Patenting following the Exposure to a Natural Disaster

Note: This figure depicts the results for our baseline event-study specification, where we compare patenting in regions exposed to natural disasters to unaffected regions. We plot two separate regressions for green and non-green patents. Percentages are relative to the overall sample averages of green/non-green patents per year per region: (2.54) and (30.5) respectively. An increase in green patenting of 24% corresponds to 0.61 additional green patents. Standard errors are clustered on the region level, and confidence intervals are drawn for the 95% interval.

the Appendix additionally reports results including controls for regional technology class composition and regional GDP. We omit these from our baseline specification since both variables may themselves be affected by disaster exposure or outcomes of the additional green patents, rendering them potentially endogenous controls. Results look qualitatively the same.

In Section 4.2.3 of the Appendix, we show that our results are consistent across a broad range of CPC classes and are not driven by any single technological domain. To assess this, we replicate our baseline specification separately for green patents grouped by the broad CPC class they belong to, based on the primary technology classification of each patent. The post-disaster response is remarkably similar across classes B (Performing Operations; Transporting), C (Chemistry; Metallurgy), F (Mechanical Engineering; Lighting; Heating), G (Physics), and H (Electricity). The effects are more volatile in D (Textiles; Paper) and E (Fixed Constructions), while in A (Human Necessities), which includes agriculture, the response is comparatively weak or absent.

1.4.1 Mitigation vs. Adaptation

Do inventors primarily adapt to a changing environment (see, for instance, Miao and Popp 2014 and Moscona and Sastry 2023), or do their inventions combat the causes of climate change? We investigate this by exploring the subcategories of green patents.

We split the sample of green patents based on their purpose, differentiating between adaption to climate change or mitigation of climate change. Specifically, we use the Y02A

class, "technologies for adaptation to climate change," and all the other Y02 subclasses which relate to mitigation. Mitigation technologies for climate change are inventions that reduce greenhouse gas emissions or enhance carbon removal from the atmosphere. Examples include renewable energy (like solar and wind), electric vehicles, carbon capture and storage, and energy-efficient buildings.[11] Table 1.2 presents the estimates for our difference-in-differences specification (1.6).

In the long run, one additional natural disaster increases patenting in mitigation technologies by 8.5% compared to the sample average, and patenting in adaptation technologies by 4.4%. Mitigation technologies are crucial in order to reduce long-run emissions. These coefficients are also statistically significantly different.[12] There are roughly 10 times more mitigation patents than adaptation patents in our sample. Section 4.2.4 of the Appendix presents our results when estimating our event-study specification. We find that the pattern for mitigation patents largely mirrors our main results, while the effect on adaptation technologies is comparatively muted.

Table 1.2: Patenting Responses split by Adaptation vs. Mitigation

	Dependent variable:		
	All Green	Mitigation	Adaptation
	(1)	(2)	(3)
Cumulative Count	0.081***	0.085***	0.044***
	(0.009)	(0.010)	(0.013)
Country-Year F.E.	Yes	Yes	Yes
Region F.E.	Yes	Yes	Yes
Wald-test p-value		0.0117	
Sample Mean	2.54	2.32	0.22
Observations	15,813	15,813	15,813
R^2	0.737	0.721	0.513
Adj. R^2	0.723	0.706	0.487

Note: This table gives the results for our baseline regression for all green patents (1), and split by mitigation and adaptation in columns (2) and (3). Cumulative Count is the count of past natural disasters. We construct a Wald-test of the form $W = \frac{(\hat{\beta}_{\text{eq1}} - \hat{\beta}_{\text{eq2}})^2}{\text{Var}(\hat{\beta}_{\text{eq1}} - \hat{\beta}_{\text{eq2}})}$, where: $\text{Var}(\hat{\beta}_{\text{eq1}} - \hat{\beta}_{\text{eq2}}) = \text{Var}(\hat{\beta}_{\text{eq1}}) + \text{Var}(\hat{\beta}_{\text{eq2}}) - 2 \cdot \text{Cov}(\hat{\beta}_{\text{eq1}}, \hat{\beta}_{\text{eq2}})$. We can reject the Null hypothesis $H_0 : \beta_{\text{eq1}} = \beta_{\text{eq2}}$ against the alternative ($H_1 : \beta_{\text{eq1}} \neq \beta_{\text{eq2}}$) with the reported p value. Standard errors are clustered on the region level and are reported in parentheses. P-values are as follows: *p<0.1; **p<0.05; ***p<0.01

Inventors thus not only invent technologies protecting against the adverse effects of climate change, but they patent ideas that help combat climate change itself. Mitigation technologies are not directly tied to natural disasters. The strong effects we observe

[11] See Appendix Section 4.2.4 for an overview of all the Y02 classes used in this analysis.

[12] Using a Wald-test, we test whether these coefficients are statistically significantly different and can reject the null hypothesis of equality with a p-value of 0.0117.

for these technologies suggest that inventors respond not only to the immediate threat of disasters, but also by developing innovations with broader applications in everyday products. This implies that they not only recognize rising risks, such as increased flood frequency, but also perceive greater value in technologies that reduce GHG emissions. In Section 1.5, we show that these expectations about greater value, at least in part, stem from inventors' higher-order beliefs about consumer preferences about green consumption and environmental policy.

In Appendix Section 4.2.4, we additionally present results for all the subclasses in isolation. Estimating our baseline difference-in-differences specification (1.6) separately for each of these subclasses, we find that the coefficient on the cumulative count of past disasters is consistently positive and statistically significant across all subclasses. Of particular interest for mitigation are the Y02E class, "reduction of greenhouse gas (GHG) emissions related to energy generation, transmission, or distribution," and the Y02T class, "climate change mitigation technologies related to transportation," as they cover the most polluting activities. Additionally, we present event study estimates for these two subclasses separately. The patterns mirror our main results, emphasizing that inventors react across different industries, even in the most polluting ones.

1.4.2 Spillovers

If personal experience is the primary channel through which natural disasters affect innovation, then we should observe little to no impact in neighboring regions that were not directly exposed. To test this, we examine potential spillovers by estimating the effect of disaster exposure in adjacent regions on local patenting activity. For each region, we calculate the number of natural disasters in neighboring regions. Figure 1.4 depicts our different distance bands within which we consider a region to be a neighbor. Put differently, we estimate the effect of natural disasters in the black-shaded region on patenting in neighboring regions. Red-shaded areas are regions whose borders are closer than 50km, while orange-shaded areas are closer than 100km away, and yellow shaded areas are closer than 150km away. To avoid including directly exposed regions in the control group, we remove all regions that experienced natural disasters in the past five years.

We estimate our difference-in-differences specification (1.6) and depict the results in Table 1.3. Natural disasters only marginally affect patenting in directly adjacent regions, leading to a long-run increase in patenting of about 1% in adjacent regions. Relative to the results presented in Table 1.2, the magnitude of the increase is approximately eight times smaller. When we move to regions that are 100km or 150km away, we find precisely estimated point estimates close to zero. There is no sizeable effect of natural disasters on regions that are more than 50km away. The lack of large spillovers is striking and underlines that the direct personal experience of the inventor is the driving force behind

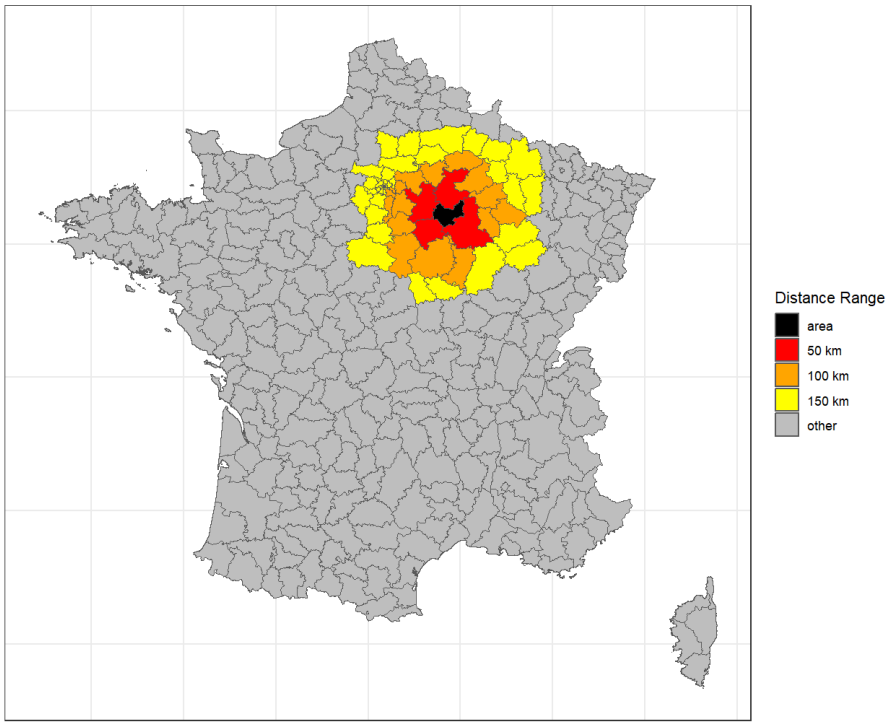


Figure 1.4: Illustration of Distances to Region

Note: This figure depicts the different distances at which we consider an area to be a neighbor of the black-shaded region in the center. Red-shaded areas are regions whose borders are closer than 50km to the area of interest, while orange-shaded areas are closer than 100km away, and yellow-shaded areas are closer than 150km away. We show only France for simplicity.

our results.

While our approach of excluding affected regions from the control group ensures clean identification, it also systematically excludes regions that experience disasters more frequently. This exclusion may introduce some bias, as effects are primarily identified from regions with lower disaster exposure. In section 4.2.7 in the Appendix, we also show our results when we exclude regions that have been affected in the past three, four, six, seven years respectively. We also have one specification where we include all affected regions in the control group, thus not restricting our sample. Results are qualitatively the same for both types of robustness checks.[13]

1.4.3 Patent Value

To assess whether natural disaster exposure spurs economically valuable innovation, we examine the value of resulting patents. A key concern is that observed increases in patenting may reflect low-quality or hastily conceived inventions. To address this, we use

[13]When we exclude all regions affected in the past six, and seven years, we find slightly larger and more imprecise estimates for the 50km range. When we do not restrict the control group, the results remain very similar for the 50km range. We find very small but significant negative effects of neighboring natural disasters that are more than 50km away. This is due to the control group now containing regions that are affected by natural disasters. We essentially invert our main regression. In this specification, we compare an area unaffected by a natural disaster with areas that are affected.

Table 1.3: Spillovers of Neighboring Disasters

	Dependent variable:		
	$P(Y02_{lt})$		
	(1)	(2)	(3)
Cumulative Count (50km)	0.0101***		
Neighboring Disasters	(0.0038)		
Cumulative Count (100km)		0.0001	
Neighboring Disasters		(0.0005)	
Cumulative Count (150km)			−0.0005
Neighboring Disasters			(0.0003)
Country-Year F.E.	Yes	Yes	Yes
Region F.E.	Yes	Yes	Yes
Observations	4,125	4,125	4,125
R^2	0.8495	0.8489	0.8489
Adjusted R^2	0.8200	0.8192	0.8193

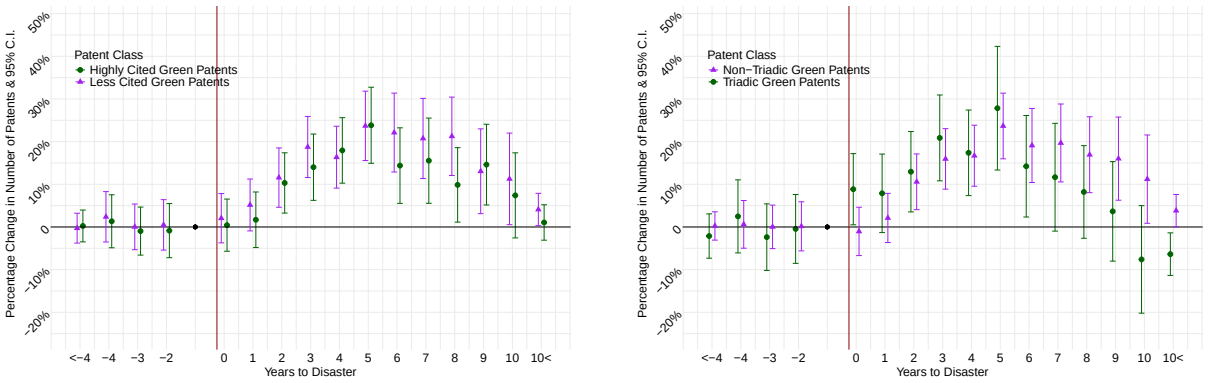
Note: This table shows our regressions results for spillovers from adjacent regions. "Cumulative Count Neighboring Disasters" is the count of past natural disasters in neighboring regions. Which regions are considered as "neighbors" depends on the distance threshold as shown in Figure 1.4. From these regressions, we exclude all regions that themselves experienced a disaster in the past 5 years. Standard errors are clustered on the region level and are reported in parentheses. P-values are as follows: *p<0.1; **p<0.05; ***p<0.01

citation counts, a proxy for patent value, to gauge the technological and commercial relevance of disaster-induced innovations. Prior work shows that more highly cited patents are more socially valuable (Trajtenberg 1990), command higher market valuations (Hall et al. 2005), and are more likely to be sold (Harhoff et al. 1999). More recent studies further affirm their reliability as indicators of patent quality (Jaffe and Rassenfosse 2017).

We investigate whether our findings result from the invention of high- and/or low-value patents by examining the effects for patents with high and low citation counts separately. We split the sample based on patents that received citations above or below the median within their respective groups. Given that a patent published in 1995 is likely to have more citations than one published in 2005, and that a patent for a toothbrush may attract a different number of citations compared to one on quantum computing, we compare patents within the same CPC class j (e.g., CPC class C for Chemistry) and published in the same year t. Let the group of patents belonging to CPC class j published in year t be denoted by G_{jt}. For all such groups, we then compute the median number of citations, denoted by $\tilde{G}_{jt}$. Since a patent might belong to multiple CPC classes (for instance, j and k), we define it as having above-median citations if:

$$\text{Citations}_{it} > \frac{\tilde{G}_{jt} + \tilde{G}_{kt}}{2}. \tag{1.7}$$

Figure 1.5a plots the results when estimating our event-study in equation (1.4) for both samples.



(a) Patenting Activity by Citation

(b) Triadic vs Non-Triadic Patenting

Figure 1.5: Green innovation response to natural disaster exposure by patent characteristics

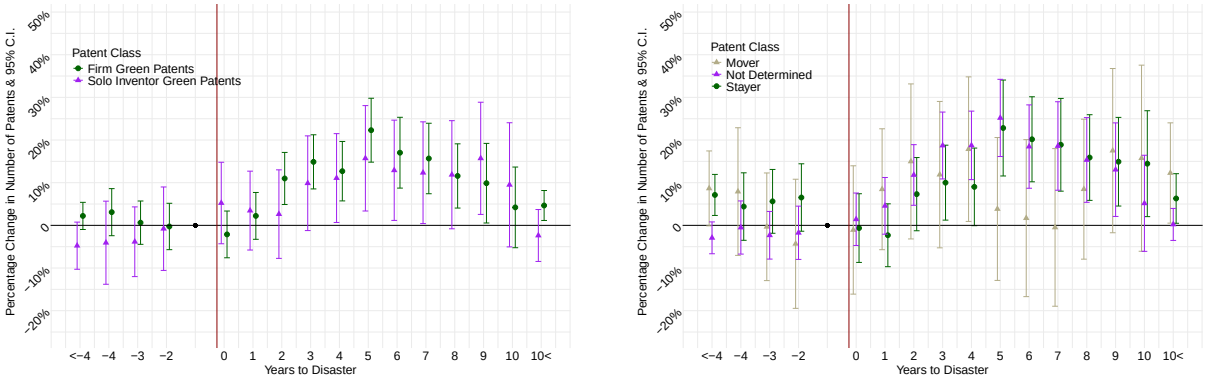
Note: This figure presents results from our event-study analysis of green patenting following natural disasters, split by different patent characteristics. Panel (a) differentiates patents by citations (above vs below median citations), with a sample mean of 1.271 (highly cited) and 1.253 (less cited) green patents per region-year. Panel (b) differentiates between Triadic and Non-Triadic green patents, with sample means of 0.503 and 2.05, respectively. Standard errors are clustered at the region level, and confidence intervals represent the 95% confidence level.

Both regressions show a positive, significant effect on subsequent green patents, as observed in our baseline results in Figure 1.3. Moreover, there does not seem to be a significant difference between patents with different citation counts. Our results suggest that natural disasters stimulate patenting activity regardless of whether patents are highly cited or not, indicating broad-based innovation rather than targeted low-value activity.

Another commonly used way to measure patent value is whether patents belong to a triadic family. A triadic patent is a patent filed at the European Patent Office (EPO), the Japanese Patent Office (JPO), and the United States Patent and Trademark Office (USPTO). Patents of such nature are usually quite valuable, as filing multiple patents in vastly different jurisdictions is, first of all, expensive, and secondly, implies that their technical content is economically valuable in some of the biggest markets on earth. We then use this indicator to estimate our event study for these triadic patents. Figure 1.5b plots our results. We find a similar pattern as in our baseline regression. Significant effects on triadic patents highlight that local experiences generate innovations with global market relevance, showing that inventors' responses are economically substantive and internationally applicable. As triadic patents are filed all over the world, these findings also alleviate the concern that our results are driven by any effect the natural disaster might have on the patent examiner. It is unlikely that a disaster in the south of France will influence the examiner at the USPTO.

1.4.4 Types of Inventors

To better understand who drives the observed green innovation response to natural disasters, we disaggregate our analysis by inventor type, using detailed information on applicants and inventors recorded in PATSTAT. This is possible because PATSTAT includes standardized names for both inventors and applicants, along with information on applicant type, such as whether the applicant is a firm or an individual.[14] In our sample, 72% of patents are filed by firms and 16.8% by individual inventors. For the remaining 11.2%, the applicant type cannot be determined.[15]



(a) Firms vs Individual Inventors

(b) Stayers vs Movers vs Not Categorized

Figure 1.6: Innovation by Inventor Type and Mobility

Note: This figure presents results from our event-study analysis of green patenting following natural disasters, split by inventor characteristics. Panel (a) differentiates patents filed by firms (sample average: 1.82 patents per region-year) and solo inventors (sample average: 0.43 patents per region-year). Panel (b) differentiates inventors by their mobility status: Stayers (sample average: 0.58), Movers (sample average: 0.1), and Not Determined (sample average: 1.86). Standard errors are clustered at the region level, and confidence intervals represent the 95% confidence level.

We estimate equation (1.4) separately by inventor type, using the count of all patents filed by either firms or individuals as the outcome. As patents can be filed by an individual and a firm jointly, when focusing on individual patent holders, we only keep patents exclusively filed by individuals. We compare the patenting behavior of exposed and unexposed inventors, separately for those affiliated with firms and those working independently. Figure 1.6a shows the event-study results. Given the larger number of patents filed by firms, estimates for firms are more precise. Natural disasters affect both inventors embedded in firms and those working independently or in small teams.

In Section 4.2.12 of the Appendix, we show that personal experience of natural disasters increases green patenting along both the intensive and extensive margins, affecting both first-time and repeat inventors. We classify an inventor as a repeat inventor when they have previously shown up as an inventor in the patent system. Any inventor that files additional patents only afterward or only once is classified as a first-time inventor. We classify a patent as coming from a repeat inventor if at least one of its inventors has previously patented. Notice that the majority (roughly 83%) of patents are filed by at

[14] This classification in PATSTAT is based on data from ECOOM (K.U. Leuven).

[15] We omit results for this unclassified group.

least one repeat inventors. Both first-time and repeat inventors react to natural disaster exposure, with the response being significantly larger among repeat inventors.

In Appendix Section 4.2.13, we further explore how inventors' past patenting activities influence their innovation response to natural disasters. Specifically, we differentiate between inventors with extensive prior patent filings ("large filers") and those with fewer previous patents ("small filers"), additionally distinguishing between inventors with substantial past engagement in environmentally harmful ("brown") technologies versus those with less or no prior involvement. Overall, the observed innovation responses are relatively similar across these classifications; however, large patent filers consistently exhibit stronger responses than small patent filers.

1.4.5 Inventor Movement

To address potential concerns about endogenous inventor selection, we examine whether green innovation responses differ between inventors who remain in their region and those who move. For a limited sub-sample of inventors, we know if they were previously recorded as living in the same region, or if they have moved. We can thus compare the evolution of patents for inventors who either stayed, moved, or for whom we do not have such information. We know that roughly 23% of inventors did not move, 4% moved, and we lack information to pinpoint the moving status for the rest (73%). Figure 1.6b plots our results. For inventors who did not move, results remain similar to our baseline findings. Results are noisy for movers, likely because we have limited statistical power due to the low number of movers in our sample. For those inventors for which we cannot determine their status, results are very similar to our baseline findings. Our findings alleviate some of the concerns about inventor selection into affected regions, as inventors who did not move equally respond to natural disaster exposure.

1.4.6 Disaster Severity

To explore whether the magnitude of natural disasters shapes the innovation response, we examine how green patenting varies with disaster severity, distinguishing between more and less deadly events as well as more and less economically damaging ones. Table 1.4 presents estimates from our baseline difference-in-differences specification (1.6), comparing the effects of natural disasters on green innovation separately by disaster severity. In columns (1) and (2), we contrast the effects of disasters with death tolls above the median column (1) and economic damages above the median column (2) to those with lower severity. Previous work has shown that disasters involving significant loss of life create greater salience of climate-related risks among the public (Eisensee and Strömberg 2007; Kalatzi Pantera et al. 2023; Demski et al. 2017).

Table 1.4: Patenting Effects by Disaster Severity

	$P(Y02_{lt})$	
	Split by Median	
	Deaths	Damages
	(1)	(2)
Cumulative Count: Above Median	0.1082***	0.0592**
Most Severe Disasters	(0.0179)	(0.0299)
Cumulative Count: Below Median	0.0532***	0.0852***
Less Severe Disasters	(0.0226)	(0.0115)
Country-Year F.E.	Yes	Yes
Region F.E.	Yes	Yes
P-Value: Coef. Difference	0.1337	0.4733
Observations	15,813	15,813
R^2	0.7374	0.7373
Adjusted R^2	0.7235	0.7234

Note: This table compares disasters with above and below median level disaster severity as measured by either deaths or damages. In Column (1) disasters are split along the median based on their number of deaths. In column (2) disasters are split along the median on the monetary value of damages. We test for difference in coefficients between severe and less severe disasters using a Wald test of the form $W = \frac{(\hat{\beta}_{\text{top}} - \hat{\beta}_{\text{bot}})^2}{\text{Var}(\hat{\beta}_{\text{top}} - \hat{\beta}_{\text{bot}})}$, where $\text{Var}(\hat{\beta}_{\text{top}} - \hat{\beta}_{\text{bot}}) = \text{Var}(\hat{\beta}_{\text{top}}) + \text{Var}(\hat{\beta}_{\text{bot}}) - 2 \cdot \text{Cov}(\hat{\beta}_{\text{top}}, \hat{\beta}_{\text{bot}})$. We test the null hypothesis $H_0 : \beta_{\text{top}} = \beta_{\text{bot}}$ against the two-sided alternative $H_1 : \beta_{\text{top}} \neq \beta_{\text{bot}}$, and report the corresponding p-value. Standard errors are clustered on the region level and are reported in parentheses. P-values are as follows: *p<0.1; **p<0.05; ***p<0.01

We find that highly deadly disasters have stronger effects on green innovation compared to less deadly disasters, although these differences are not statistically significant. This is in line with our proposed mechanism, as larger increases in public salience of climate change should translate into inventors forming higher monetary expectations, thereby prompting a stronger innovation response.[16]

In contrast, our analysis of economically damaging disasters shows weaker innovation responses for the most severe disasters, although these differences are not statistically significant. In column (2), the effect for the most damaging disasters is smaller than that for the less destructive disasters. This result is in line with literature documenting that large-scale economic disruptions from disasters destroy productive capital and critical infrastructure, which may in turn impede innovation by limiting firms' capacity to invest and adapt (see Peters et al. 2024; Le et al. 2024).

[16]We also examine the effects of the single most severe disaster experienced by a region, which yields larger but statistically noisier coefficients. These results are reported in Section 4.2.5 of the Appendix. In addition, we explore heterogeneity by disaster type. Due to the small number of drought events in our data (only three cases), we exclude them from this analysis. For extreme temperature events, floods, and storms, we find positive effects on green innovation that broadly mirror the patterns shown in Figure 1.3, albeit with greater noise. Results are reported in Section 4.2.6 of the Appendix.

1.4.7 Survey Measure of Green Innovation

Patent data are frequently used to measure innovation; however, there are some limitations e.g. not all innovations are patentable, and not all inventors opt to patent their innovations. To further underline our findings, we repeat our analysis using an alternative indicator for green innovation based on survey data from the German part of the Community Innovation Survey (CIS), a biennial firm-level survey covering innovation activity. While exclusively focusing on firms is a drawback, the majority of green innovation occurs within a firm structure, with approximately 72% of patents in our data being filed by inventors embedded within firms, compared to only 16.8% filed by individuals.[17] The core questionnaire of the CIS captures general innovation behavior, selected waves also include a dedicated module on environmental innovation. These green innovation questions were introduced in 2009 and repeated in 2015 and 2021. Survey responses are provided by individuals within firms who are familiar with the innovation process or directly involved in R&D. However, because we did not conduct the survey ourselves and the CIS does not record respondent identifiers or roles, we cannot observe who exactly within the research department completed the questionnaire. Our analysis sample comprises 18,425 firm-year observations from these three survey waves. The CIS is structured as a repeated cross-section, and firms cannot be tracked across waves due to the absence of consistent firm identifiers. The CIS contains region-level information of the firms' location, which we use to link firms to our natural disaster data.

We construct three alternative indicators of green innovation. The first captures the implementation of internal environmental innovations within firms, such as technologies that reduce energy, material, or water use, lower emissions or pollution, substitute fossil fuels with renewables, or introduce safer materials. On average, 48.3% of firms report the adoption of such internal green processes. The second indicator focuses on the introduction of new or significantly improved products or services offering environmental benefits, such as facilitating recycling, reducing pollution, or extending product life. On average, 34.8% of firms report such product- or service-based green innovations. Both of these are dummies that are equal to one if a firm indicated that they introduced one of these innovations. The third measure combines both indicators.[18]

In Table 1.5 (columns 1-3), we regress all three indicators of green innovation on the count of past natural disasters. One additional past natural disaster increases the likelihood of introducing process innovation and green products by roughly 4.3% and 4.7% relative to the sample average, respectively.[19] The analysis supports our prior

[17] For the remaining 11.2%, we are unable to determine the type of inventor. See Section 1.4.4 for details on how inventor types are classified.

[18] Section 4.2.9 in the Appendix describes the exact procedure how we construct our different measures for green innovation.

[19] This corresponds to the 2.08% and 1.63% percentage points increase in Table 1.5.

findings and highlights the positive and significant effect of natural disaster exposure on green innovation. We do robustness checks in Section 4.2.9 in the Appendix, where we include NUTS-2 and region-level (NUTS-3) fixed effects. Results remain robust with NUTS-2 fixed effects. With region-level fixed effects, point estimates are similar but lose statistical significance due to limited within-region variation of our treatment across only three survey waves.

Table 1.5: Effect of Natural Disasters on Green Innovation and Climate Affectedness

	Green Innovation Combined	Within-firm Process Innovation	Green Products	Climate Affectedness
	(1)	(2)	(3)	(4)
Cumulative Count	2.16***	2.08***	1.63***	5.58***
	(0.283)	(0.287)	(0.237)	(0.691)
Firm Size F.E. (employment)	Yes	Yes	Yes	Yes
Revenue	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	No (Single Wave)
Industry F.E. (2-digit NACE)	Yes	Yes	Yes	Yes
Observations	15,395	15,426	15,226	4,873
R^2	0.629	0.591	0.451	0.582
Adj. R^2	0.627	0.589	0.448	0.576

Note: This table reports the effect of the cumulative number of past natural disasters on several survey outcomes. Column (2) reports results for within-firm process innovation, columns (3) for a firm introducing new green products, and column (1) for both of these combined. Column (4) reports the effect on firms' self-reported climate affectedness. All models include firm size (based on employment dummies), revenue controls, year fixed effects, and 2-digit NACE industry fixed effects. Standard errors are clustered at the regional (Kreis) level. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Additionally, we examine whether firms exposed to natural disasters report experiencing their effects. In the last wave of the survey, firms were asked how important various climate-related impacts were between 2018 and 2020. We create a dummy variable equal to one if the item "Impact of extreme weather conditions" (e.g., transport disruptions, storm damage, flooding, drought) was rated as high, medium, or low importance. Firms in disaster-affected areas are significantly more likely to report experiencing climate-related impacts, providing a strong first stage for our analysis (see column (4) in Table 1.5).

1.5 Reasons to Innovate

In this section, we investigate how inventors' personal experiences of natural disasters affect their innovation choices. We draw on additional survey items in the CIS environmental innovation module. Conditional on reporting the introduction of an innovation with environmental benefits, firms were asked to assess the importance of various potential drivers. These include existing environmental regulations, anticipated future regulations or taxes, voluntary standards or best practices within the industry, current or expected market demand, government funding or subsidies, and reputational concerns. The precise question text reads as:

> "During [the past two years], how important were the following factors in driving your enterprise's decision to introduce innovations with environmental benefits?"

Each factor generates a dummy equal to one if it was rated "low," "medium," or "high" importance, and zero if deemed "not relevant" (see Appendix 4.3.2 for details).

The primary drivers of green innovations among the surveyed firms are existing environmental regulations, with 63.3% of firms identifying this factor as significant. Voluntary actions or standards for environmental best practices within their sector were noted as important by 57.9% of firms. Anticipated future regulations or taxes motivated 55.6% of the firms, while current or expected market demand for environmental innovations influenced 49.1%. Lastly, 40.7% of firms cited government grants and subsidies as a key motivating factor.

We then estimate the effect of natural disaster exposure on each of these drivers using

$$Y_{ilkt} = \beta \left(\sum_{s=0}^{\infty} d_{l,t-s} \right) + \gamma_1 \, \mathbf{S}_{it} + \gamma_2 \, R_{it} + \gamma_3 \, \lambda_t + \gamma_4 \, \lambda_k + \epsilon_{iltk}, \tag{1.8}$$

where Y_{ilkt} indicates whether firm i in region l in industry k at time t rated the factor as relevant; $\sum_{s=0}^{\infty} d_{l,t-s}$ is the count of natural disasters prior to questioning; $\mathbf{S}_{it}$ is a vector of firm-size dummies for medium and large firms, with small firms being the reference group; R_{it} is firm i's revenue in year t; λ_t and λ_k are year and two-digit NACE industry fixed effects; and ϵ_{iltk} is an idiosyncratic error term clustered at the NUTS-3 (Kreis) level. The firm-size dummies $\mathbf{S}_{it}$ are based on the number of employees. Specifically, we differentiate small firms with less than 50 employees, medium firms with 50-249 employees, and large firms employing more than 249 individuals. Our coefficient of interest β estimates the effect one additional past natural disaster has on a firm's probability of stating a factor as relevant for their green innovation decision. The comparison is between firms that have recently introduced a green innovation and have been affected by a natural disaster, and those firms that have introduced a green innovation but have not been affected by a natural disaster. Table 1.6 gives our results.

In column (1) of Table 1.6, a one-unit increase in the count of past natural disasters raises the probability of mentioning expected future regulation by 0.932 percentage points on a baseline mean of 63.33 percent, a highly significant effect at the 1 percent level. Column (2) shows a similarly strong response for expected demand. Each additional disaster increases the likelihood of reporting anticipated market demand as a motivator by 0.865 percentage points, relative to a 55.64 percent mean. Firms also become 0.748 percentage points more likely to point to existing environmental regulations (mean 40.66 percent). By contrast, the effects on public funding and reputation concerns are small and statistically indistinguishable from zero. Voluntary standards show a sizable 1.140

Table 1.6: Effect of Cumulative Disaster Count on Reasons to Innovate

	Dependent variable:					
	Expected Regulation (1)	Expected Demand (2)	Existing Regulation (3)	Public Funding (4)	Voluntary Standard (5)	Reputation (6)
Cumulative Count	0.932***	0.865***	0.748***	0.169	1.140***	0.391
	(0.266)	(0.263)	(0.263)	(0.269)	(0.218)	(0.313)
Firm Size F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Revenue	Yes	Yes	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Industry F.E. (2-digit)	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Y	63.33	55.64	40.66	49.07	57.95	61.47
Observations	8,787	8,754	8,862	8,776	8,782	5,423
R^2	0.548	0.489	0.607	0.418	0.543	0.667
Adj. R^2	0.544	0.484	0.603	0.413	0.539	0.662

Note: This table reports the effect of disaster exposure on the drivers of innovation as measured in the survey. One additional past disaster increases the probability of a firm mentioning these factors by β percentage points (e.g. for column 1 by β=0.932 percentage points). Outcomes are: Expected Regulation = expected future environmental rules; Expected Demand = expected future market demand for green innovation; Existing Regulation = current environmental rules/charges; Public Funding = government grants/subsidies; Voluntary Standard = voluntarily joining a standard for green practices (e.g. environmental/organic label); Reputation = firm's reputation concerns. Standard errors clustered at the region (NUTS-3) level and are reported in parentheses. Significance levels: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

percentage-point increase (57.95 percent mean). Taken together, these results suggest that natural disasters most strongly amplify firms' expectations about future regulation and market demand, with more modest or no effects on public funding and reputational motives.

We interpret the effects on expected regulation and expected demand as firms expecting consumers to increasingly value green alternatives in the future. Pro-environmental preferences can be expressed not only through consumers purchasing decisions but also through political support for stricter environmental policies. In essence, experiencing a natural disaster changes higher-order beliefs (beliefs about the beliefs of others) about consumers' climate change beliefs and their valuation of green consumption and green policy. In turn, this expectation of stronger consumer environmental consciousness leads firms to foresee greater pecuniary returns to investing in green innovation. Prior research documents that natural disaster exposure shifts beliefs toward greater support for environmental policy. Dechezleprêtre et al. (2025a) and Djourelova et al. (2024) show that disasters increase the salience of environmental issues. Owen et al. (2012) and Osberghaus and Fugger (2022) find that personal disaster experience heightens perceived climate risks and support for environmental regulation.

Our findings also point to firms increasing their awareness of existing environmental regulation. A potential reason is the increased salience of such policies. An increase in local climate change salience leads firms to pay closer attention to related topics, such as environmental regulation.

The positive effect on firms citing voluntary standards as primary reasons for innovation can plausibly be attributed to two channels in particular. First, firms might join voluntary standards due to the signaling value these standards have for consumers. Be-

ing able to label your products as e.g. "micro-plastic-free" can attract consumers (Agatz et al. 2021, Duckworth et al. 2022). Second, firms may engage in green innovation due to intrinsic motivation. Intrinsic motivation refers to the idea that there is no reward for an activity other than the activity itself. This includes acting based on ethical convictions or long-term sustainability goals, even in the absence of financial rewards. In our setting, exposure to natural disasters may shift firms' internal priorities in this direction. We view reputational concerns as being potentially aligned with such non-monetary motivations.

In Section 4.2.8 in the Appendix, we additionally include NUTS-2-level and region-level fixed effects.[20] Results remain robust when we include NUTS-2 level fixed effects. When including region-level fixed effects, our point estimates remain similar, but become statistically insignificant apart from for the effect on "Expected Regulation". When including region-level fixed effects we lose a lot of our underlying disaster variation, as we only have three survey waves.

Natural disaster exposure does not affect firms' self-reported take-up of public funding schemes. To further assess whether government research support could explain the observed increase in green innovation, we analyze administrative data on French public R&D funding (see Section 4.2.8 in the Appendix). We find no significant changes in the number of grants or funding levels in disaster-affected regions, further underlining that the innovation response is not driven by targeted public subsidies.

We focus on one particular mechanism: exposure to natural disasters increases the local salience of climate change (other channels may operate alongside or in addition to this one). This increase in salience can have multiple effects on inventors and inventive firms. It shapes local inventors' expectations regarding environmental policy and the demand for green goods. These expectations are driven by inventors' higher-order beliefs about the climate change beliefs of consumers and voters. Increased salience further leads to an increase in attention to climate change and potentially affects inventors' and firms' intrinsic motivation.

1.6 A Model of Disaster-Induced Green Innovation

Our empirical results establish that the personal experience of a natural disaster increases green innovation and that this response operates through shifts in inventors' expectations about future green demand and regulation. However, the reduced-form estimates cannot separate two distinct channels through which disaster-induced salience may operate. Inventors may innovate because they expect green technologies to be more profitable, or

[20] In Germany, the NUTS-2 level corresponds to current and former "Bezirke", which are administrative regions that are below the "Länder". The city states of Berlin, Hamburg, and Bremen, as well as the federal states of Brandenburg, Mecklenburg-Western Pomerania, Schleswig-Holstein, Saxony-Anhalt, Saarland, and Thuringia, do not have any NUTS-2 level subdivision and are thus themselves NUTS-2 level areas.

because they derive intrinsic utility from pursuing environmentally beneficial research. In this section, we embed both channels in a formal theoretical framework that allows us to disentangle them by exploiting their differential interaction with market structure. We adapt and extend the framework proposed by Aghion et al. (2023), whose core insight is that green innovation responds to consumer demand for environmentally friendly products. A key feature of their model is that market structure matters. Firms facing intense product market competition benefit more from "escaping" competition by developing green products that differentiate them from incumbents.

We add a behavioral channel through which inventors form expectations about the future profitability of green technologies in response to natural disasters, encompassing both anticipated shifts in consumer demand and expected changes in environmental regulation. We introduce uncertainty about future profitability and allow inventors to derive intrinsic utility from pursuing environmentally beneficial innovation. Both profitability expectations and intrinsic motives vary across regions and are shaped by local natural disaster exposure. The model yields a testable comparative static: if profitability expectations drive the innovation response, then inventors in more competitive markets should respond more strongly to disaster exposure, since competition amplifies the returns to green differentiation. In contrast, if intrinsic motivation were the sole driver, the response would be independent of market structure. This extension is motivated by our empirical finding in Table 1.3, which shows limited evidence of geographic spillovers from neighboring disasters, suggesting that behavioral responses are highly localized.

1.6.1 Model

Inventors choose R&D investments aiming to maximize expected profits. Once innovations have realized, they produce with their respective technologies and compete for consumers. Revenues are distributed as wages to production and R&D workers, and net profits are redistributed to consumers, who also own firms as shareholders. There is a continuum of horizontally differentiated goods indexed by $j \in [0, 1]$. For each variety, two duopolists and a competitive fringe supply otherwise identical products that differ only in their emissions intensity embodied in production. Producing one unit of the good with environmental quality $q_{j,f}$ generates $x_{j,f} = 1/q_{j,f}$ units of carbon emissions. Labor is the sole input, supplied perfectly elastically at a wage normalized to one. The marginal labor requirement per unit of output equals a constant $c > 0$.

The representative consumer derives utility from variety consumption but also experiences private disutility from their carbon footprint. These preferences could arise, for instance, from social image concerns or a general sense of responsibility toward the environment. When purchasing $y_{j,f}$ units from each firm $f \in \mathcal{F}_j$ in sector j, the period-t

utility is given by:

$$U_t = \int_0^1 \ln \bar{y}_{j,t}\, dj, \quad \text{with} \quad \bar{y}_{j,t} = \int_{f \in \mathcal{F}_j} y_{j,f,t}\, q_{j,f,t}^{\delta_t}\, df. \tag{1.9}$$

where $\bar{y}_{j,t}$ is the quality-adjusted consumption of good j, purchasable from various firms $f \in \mathcal{F}_j$. The value of a green product depends positively on the stringency of environmental regulation, as stricter policies raise the relative costs of non-green alternatives. The parameter δ captures how individuals value their own private consumption and express their political preferences regarding environmental policy. Empirical evidence suggests exposure to natural disasters shifts local preferences towards stronger environmental policies.

Varieties j are imperfect substitutes. Within each variety, all demand will be allocated to the firm offering the highest quality-to-price ratio, q^{δ}/p. Logarithmic preferences imply that expenditure is uniform across all varieties. For a formal derivation of this result, refer to Section 4.1.1 in the Appendix. We assume that consumer demand is non-local. Once a product is patented, it is marketed globally.[21] This assumption enhances model tractability and is also grounded in the legal interpretation of patent rights. The Paris Convention for the Protection of Industrial Property (1883), which has been adopted almost universally, stipulates that an inventor who patents a product in one country has a 12-month window during which they can apply for protection in other contracting states. These subsequent applications are granted the same priority date as the original filing. This provision facilitates easier entry into international markets without risking loss of intellectual property rights to third parties. Even if an inventor decides not to patent their invention in some countries, the same invention cannot be patented there by others and is instead regarded as public knowledge. Note that the model separates the product market, which is global, from the process of belief formation, which is local. One might ask why local signals dominate when products are marketed globally. The key distinction is that inventors form expectations about future consumer valuations based on salient local experiences, even though the resulting innovations compete in global markets. This is consistent with our empirical finding of limited geographic spillovers (Table 1.3). Disaster exposure shifts local beliefs without propagating to neighboring regions through product market channels. Our results on triadic patents (Figure 1.5b) confirm that disaster-induced green innovations are indeed patented across major markets, indicating that locally formed beliefs translate into inventions that compete globally. We assume a market structure in which each sector features a duopoly, composed of two competing innovators, alongside a competitive fringe. The fringe consists of firms that

[21] Although this assumption is strong, we demonstrate that our main results remain valid when restricting the analysis to globally marketed patents. See Section 1.4.3 for results focusing exclusively on triadic patents, which are filed globally.

do not invest in innovation and continue producing the previous-generation good, which is one step behind the technological frontier. These goods are γ times more polluting than those of the duopolists and thus less attractive to environmentally conscious consumers. The presence of the fringe disciplines the market by limiting the pricing power of the duopolists. If the duopolists were to charge a price exceeding the marginal cost of the fringe good, adjusted for its lower environmental quality, consumers would switch to the cheaper, albeit dirtier, alternative. As a result, the duopolists, who have a quality advantage, cannot extract monopoly rents beyond what the quality differential justifies.

The quality of a green good y_j evolves according to $q_j = \gamma^{k_j}$, where $\gamma > 1$ denotes the step size of a green innovation, and k_j is the cumulative number of past innovations in variety j. Intuitively, each successful innovation improves the environmental performance of a good by a factor of γ. Innovation arises from directed R&D effort. Inventors can choose to exert research effort $z_j \in [0, 1]$, incurring convex costs of $\kappa z_j^2/2$ units of labor.[22] With probability z_j, the investment succeeds, improving the quality of the good by a factor of γ in the next period. With probability $1 - z_j$, the attempt fails, and no technological progress is achieved. Upon a successful innovation, the inventor receives a patent that grants it a temporary edge over its rival. To capture the entire market, a successful inventor engages in limit pricing, setting the price just low enough to undercut her competition, specifically at $p_M = \gamma^{\delta} c$. This allows the innovator to behave as a de facto monopolist for one period. After that, the patent expires, the quality gap closes, and market competition resumes.

An important aspect of an inventor's decision to innovate is how much she expects consumers to value green products. Consumer valuation of green goods δ evolves over time, and there exists local uncertainty regarding the future valuation level. The global level of δ could, for instance, depend on the degree of global exposure to climate change. Inventors are local and form expectations about consumers' valuation of green goods as a weighted combination of a global prior ρ (common across all locations) and local events D_l:

$$E_l[\delta] = \varphi\rho + (1 - \varphi)\phi D_l, \tag{1.10}$$

where ϕ denotes the size of the local shock. Inventors' expectations can fall whenever they are unaffected by natural disasters. We define the average expectation of consumer valuation for green goods across locations as:

$$\overline{E_l[\delta]} = \int_l E_l[\delta] f(l)\, dl = \varphi\rho + (1 - \varphi) \int_l \phi D_l f(l)\, dl = \varphi\rho + (1 - \varphi)\phi \overline{D_l}$$

where $f(l)$ denotes the probability density function over regions l, and $\overline{D_l}$ is the average

[22] Convex innovation costs are a plausible assumption, as reducing the environmental impact of goods becomes increasingly difficult. For example, designing a plane that consumes slightly less fuel is much easier than creating one that emits no CO_2 at all.

level of disaster exposure. We denote this benchmark as $\overline{E}[\cdot]$, which reflects the average belief held across all locations. Inventors' expectations may deviate from this average depending on their local exposure. In regions recently affected by natural disasters, inventors may hold higher expectations about the future valuation of green goods ($E_l[\delta] > \overline{E[\delta]}$), while inventors in unaffected areas may hold lower expectations ($E_l[\delta] < \overline{E[\delta]}$).

Inventors derive utility both from the profits of their innovation activity and from the intrinsic satisfaction of pursuing research. For tractability, we assume the factors are linearly separable in the utility function.

$$U_l^I(z_{lj}) = \alpha \Pi(z_{lj}) + \mu_l z_{lj}$$

where the parameter α captures the weight on monetary rewards, while μ_l reflects the inventor's intrinsic motivation, specifically the non-pecuniary utility derived from engaging in green innovation. Intrinsic motivation may stem from ethical concerns, a sense of moral responsibility, or personal interest in mitigating climate change. Importantly, μ_l may vary across regions and can be shaped by local experiences, such as exposure to natural disasters. Inventors with a high μ_l are thus more likely to engage in green innovation for its own sake, independent of financial incentives. At the time of investing in research, inventors form expectations over output and profitability, conditional on successful innovation:

$$E_l[y_{Mj}] = \frac{1}{E[p_{Mj}]} = \frac{1}{\gamma^{E_l[\delta]}c}, \quad \text{and} \quad E_l[\pi_{Mj}] = 1 - \frac{1}{\gamma^{E_l[\delta]}}, \tag{1.11}$$

Local expectations $\gamma^{E_l[\delta]}$ shape the profit component of utility and are critical for investment decisions. A formal derivation of expected demand and profits is provided in Section 4.1.1 of the Appendix. In markets where no innovation takes place, the duopolists engage in price competition. If they can collude perfectly, they charge the monopoly price and share profits equally. They are constrained by the competitive fringe. In contrast, under full competition, firms bid prices down to marginal cost. Following Aghion et al. (2005), we model the intensity of competition as $\Delta_j \in [1/2, 1]$, where $\Delta_j = 1$ corresponds to Bertrand competition and $\Delta_j = 1/2$ reflects full collusion. Duopoly profits thus depend on the degree of competition in sector j, and are a fraction of expected monopoly profits:

$$E_l\left[\pi_{Dj}(\Delta_j)\right] = (1 - \Delta_j) E_l\left[\pi_{Mj}\right]. \tag{1.12}$$

Given this structure, the locally expected price under imperfect competition is:

$$E_l\left[p_j(\Delta_j)\right] = \frac{c}{1 - 2(1 - \Delta_j)E_l[\pi_{Mj}]} = \frac{c}{1 - 2(1 - \Delta_j)\left(1 - \gamma^{-E_l[\delta]}\right)} \in \left[c, E_l[p_{Mj}]\right], \tag{1.13}$$

and the corresponding expected output is:

$$E_l\left[y_j(\Delta_j)\right] = \frac{1}{E_l[p_j(\Delta_j)]} = \frac{1}{c}\left[1 - 2(1-\Delta_j)\left(1-\gamma^{-E_l[\delta]}\right)\right] \in \left[E_l[y_{Mj}], \frac{1}{c}\right]. \tag{1.14}$$

A local inventor in region l and sector j with an R&D opportunity maximizes

$$\max_{z_{lj}\in[0,1]} U^I(z_{lj}, \Delta_j, E_l[\delta], \mu_l) = \alpha\left(z_{lj}E_l\left[\pi_{Mj}\right] + (1-z_{lj})E_l\left[\pi_{Dj}(\Delta_j)\right] - \frac{\kappa}{2}z_{lj}^2\right) + \mu_l\, z_{lj}.$$

where successful innovation yields monopoly profits, and failure yields the duopoly profits dependent on the level of competition. The stronger the competition is, the larger the benefit of escaping competition by innovating becomes. Competition acts as a wedge between the profits of a successful inventor and the profits that inventors can reap in the status quo. The first-order condition of a local inventor with respect to the research rate is:

$$z_{lj} = \min\left\{\frac{E_l\left[\pi_{Mj}\right] - E_l\left[\pi_{Dj}(\Delta_j)\right]}{\kappa} + \frac{\mu_l}{\alpha\kappa}, 1\right\}. \tag{1.15}$$

Using equations (1.11) and (1.12) in (1.15), we get:

$$z_{lj}(\Delta_j, E_l[\delta], \mu_l) = \underbrace{\frac{\Delta_j\, E_l\left[\pi_{Mj}\right]}{\kappa}}_{\text{Monetary Incentives}} + \underbrace{\frac{\mu_l}{\alpha\kappa}}_{\text{Intrinsic Motivation}} = \frac{\Delta_j}{\kappa}\left(1 - \frac{1}{\gamma^{E_l[\delta]}}\right) + \frac{\mu_l}{\alpha\kappa} \tag{1.16}$$

The optimal research rate consists of two additive components. The first part reflects the monetary incentives to innovate, and the second captures the inventor's intrinsic motivation. The monetary incentive term increases with the intensity of competition Δ_j and with local beliefs about the profitability of green goods, captured by $E_l[\delta]$. Intuitively, when competition is intense (i.e., Δ_j is high), duopoly profits are low, making the gains from obtaining monopoly status through successful innovation more attractive. Similarly, higher expected consumer valuation for green goods leads to higher expected monopoly profits, further incentivizing R&D. Formally, innovation effort is increasing in both arguments: $\frac{\partial z_{lj}}{\partial \Delta_j} > 0, \quad \frac{\partial z_{lj}}{\partial E_l[\delta]} > 0$. Moreover, these forces are complements, as shown by the positive cross-derivative: $\frac{\partial^2 z_{lj}}{\partial \Delta_j \partial E[\delta]} > 0$. Competition and demand expectations thus reinforce each other. In monopolistic markets, inventors already earn high rents from non-green products and respond little to shifts in consumer preferences. In contrast, under competitive pressure, inventors can escape price competition by innovating. When environmental quality matters to consumers, green innovation becomes a path to monopoly. Expected profits rise with both stronger demand expectations and higher competition.

The second term in Equation (1.16) reflects intrinsic motivation, scaled by the inventor-specific parameter μ_l. Importantly, this component is unaffected by market competition, as: $\frac{\partial^2 z_{lj}}{\partial \Delta_j\, \partial \mu_l} = 0$. This implies that inventors motivated purely by intrinsic factors (such as

environmental concern or personal satisfaction) invest in green innovation regardless of the competitive landscape. Competition only plays a role for those inventors who value monetary rewards. Figure 1.7 visualizes these comparative statics.[23]

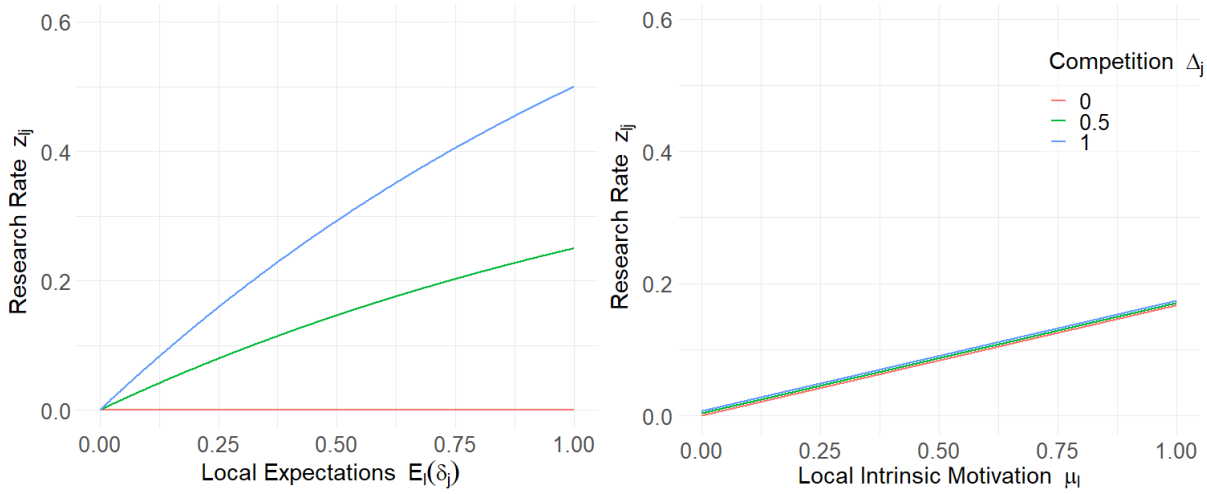


Figure 1.7: Innovation Responses dependent on Competition Δ_j

Note: The figure on the left plots inventors' innovation response dependent on their expectations $E_l[\delta]$ and the level of competition when intrinsic motivation plays no role ($\alpha \to \infty$, which drives the intrinsic motivation term $\mu_l/(\alpha\kappa) \to 0$, equivalent to setting $\mu_l = 0$). The figure on the right plots innovation responses only dependent on inventors' intrinsic motivation μ_l, i.e. for $\delta = 0$.

Hypothesis 1: Inventors facing fiercer competition (large Δ_j) increase their research output more whenever their profitability expectations increase due to exposure to a natural disaster.

In line with these comparative statics, we formulate the above empirically testable hypothesis. We can test Hypothesis 1 explicitly by comparing the response of inventors facing high levels of competition with those facing low levels of competition. We confirm this empirically in Table 1.7 in Section 1.6.2, which provides compelling evidence that shifts in higher-order beliefs significantly influence inventors' innovation decisions. Intuitively, because competition only affects pecuniary incentives, differences in competition levels should impact only those inventors driven by monetary rewards. Purely intrinsically motivated inventors would respond to exposure irrespective of the level of competition (see the right-hand panel of Figure 1.7).

Generally, the overall (private) innovation rate in the economy is the average across all sectors j and locations l, which is equivalent to the fraction of sectors where innovation is successful

$$\underline{z} \equiv \int_j \int_l z_{lj}(\Delta_j, E_l[\delta], \mu_l) f(l) dl f(j) dj.$$

Let

$$\overline{z} \equiv \int_j z_j(\Delta_j, \overline{E[\delta]}, \overline{\mu}) f(j) dj$$

[23]The figure is illustrative and highlights qualitative mechanisms. Quantitative implications depend on the specific calibration. For details, see Section 4.1.4 in the Appendix.

denote the research rate that would be achieved if all local inventors had the same (average across regions l) expectation on future environmental profitability $E_l[\delta] = \overline{E[\delta]}\ \forall l$ and intrinsic motivation $\mu_l = \overline{\mu}\ \forall l$. We can then compare $\overline{z}$ with the aggregate private research rate $\underline{z}$.

Assumption 1: Assume that the $E_l[\delta]$ are not degenerate, i.e.,

$$\exists\, l \neq l' \quad \text{with} \quad E_l[\delta] \neq E_{l'}[\delta].$$

Assumption 1 simply states that not all regions have identical expectations regarding future environmental demand. At least two regions must differ in expectations, reflecting realistic variation in the exposure to natural disasters.

Proposition 1: Under Assumption 1, we get that:

$$\begin{aligned}\overline{z} &= \int_j \left(\frac{\Delta_j}{\kappa} \left(1 - \frac{1}{\gamma^{\overline{E[\delta]}}} \right) + \frac{\overline{\mu}}{\alpha\kappa} \right) f(j)dj > \\ \underline{z} &= \int_j \int_l z_{lj}(\Delta_j, E_l[\delta], \mu_l) f(l)dl f(j)dj = \int_j \int_l \left(\frac{\Delta_j}{\kappa} \left(1 - \frac{1}{\gamma^{E_l[\delta]}} \right) + \frac{\mu_l}{\alpha\kappa} \right) f(l)dl f(j)dj \end{aligned} \tag{1.17}$$

The research rate achieved under average expectations is larger than the research rate achieved when expectations are heterogeneous across regions. See Section 4.1.2 in the Appendix for the proof. Intuitively, because the innovation production function is concave in expectations (via the function $1 - \gamma^{-E[\delta]}$), Jensen's inequality implies that dispersed beliefs reduce aggregate innovation relative to uniform beliefs at the same mean. Heterogeneous salience of climate change thus raises the aggregate cost of achieving a given level of green innovation.

Building on the Shoulders of Giants & Market Size

We further extend the model to incorporate market size effects, drawing on the literature on directed technical change, which emphasizes that the profitability of innovation increases with market size. This feature is often referred to as "building on the shoulders of giants" (Acemoglu 2002; Acemoglu 2007; see also Acemoglu et al. 2012 for an application to green technologies). To do so, we explicitly model the cost of research as depending on the size of the market for green goods $K(\eta_j)$ with $\frac{\partial K}{\partial \eta_j} < 0$. We then get that for any level of research effort $z_j \leq 1$, investing $K(\eta_j) z_j^2/2$ units of labor yields, with probability z_j, a green innovation. Similar to equation (1.16), we have that the optimal private research rate is chosen according to:

$$z_{lj}(\Delta_j, E_l[\delta], \mu_l, \eta_j) = \frac{\Delta_j E_l\left[\pi_{Mj}\right]}{K(\eta_j)} + \frac{\mu_l}{\alpha K(\eta_j)} = \frac{\Delta_j}{K(\eta_j)} \left(1 - \frac{1}{\gamma^{E_l[\delta]}} \right) + \frac{\mu_l}{\alpha K(\eta_j)}.$$

The optimal private research rate increases in the size of the market $\frac{\partial z_{lj}}{\partial \eta_j} > 0$. Additionally, inventor expectations about the profitability of a green good $E_l[\delta]$ and the market size of the green good η_j are complements: $\frac{\partial^2 z_{lj}}{\partial \eta_j \partial E_l[\delta]} > 0$. Intuitively, inventors in large markets face lower innovation costs and respond more readily to increased expectations about green profitability and increased intrinsic motivation. We plot this comparative static in Section 4.1.4 in the Appendix. This allows us to formulate an additional hypothesis:

Hypothesis 2: Innovation responses to natural disasters increase with the size of the green good market $\frac{\partial^2 z_{lj}}{\partial \eta_j \partial E_l[\delta]} > 0$ and $\frac{\partial^2 z_{lj}}{\partial \eta_j \partial \mu_l} > 0$. Put differently, the increase in innovation following exposure to natural disasters is larger in markets where the green technology market is bigger.

Welfare

We now state our welfare result. The social planner maximizes quality-adjusted consumption net of an emission externality weighted by $\psi > 0$, subject to a constraint that inventor utility remains non-negative, weighted by $\lambda > 0$. The planner can choose the research rate in every sector, and forms expectations using the population-average belief $\hat{E}[\delta] \equiv \overline{E_l[\delta]}$, so that expected monopoly output and profits are $\hat{E}[y_{M_j}] = 1/(\gamma^{\overline{E[\delta]}} c)$ and $\hat{E}[\pi_{M_j}] = 1 - 1/\gamma^{\overline{E[\delta]}}$, respectively. We state the result using the full model including the market size extension developed in Section 1.6.1. The full welfare derivation is in Section 4.1.3 of the Appendix.

Proposition 2: Under Assumption 1, the socially optimal research rate satisfies:

$$
\begin{aligned}
z_j^*(\Delta_j, \overline{E[\delta]}, \eta_j) = & \underbrace{\overline{z_j}(\Delta_j, \overline{E_l[\delta]}, \eta_j)}_{\text{average expectations}} \\
& + \frac{1}{\lambda K(\eta_j)} \left[\underbrace{\ln\left[\frac{\gamma^{\overline{E[\delta]}} \hat{E}[y_{M_j}]}{\hat{E}[y(\Delta_j)]}\right]}_{\text{collusion loss}} + \underbrace{\psi\left[\hat{E}[y(\Delta_j)] - \hat{E}[y_{M_j}]/\gamma\right]}_{\text{emission reduction}} \right] \\
& > \underbrace{\underline{z_j}(\Delta_j, E_l[\delta], \eta_j)}_{\text{local expectations}}
\end{aligned}
\tag{1.18}
$$

The socially optimal research rate exceeds the private rate for two reasons beyond the belief heterogeneity captured in Proposition 1. We can show that the collusion loss term is ≥ 0 and the emission reduction term is > 0. Imperfect competition (the collusion loss term) leads to private inventors underinvesting because duopoly rents reduce the incentive to escape competition through innovation. Additionally, private inventors do not internalize the emission reduction that green innovation delivers (the emission reduction term, scaled by the social cost of emissions ψ). Both wedges are amplified in larger

markets (lower $K(\eta_j)$). Together, these results provide a rationale for policy interventions that raise climate change salience in unaffected regions or subsidize clean innovation where private incentives fall short. Our welfare results also imply that more private green innovation is likely not socially "wasteful" even if local expectations about future green good demand are wrong. See Section 4.1.3 in the Appendix for the full welfare derivation and the proof of Proposition 2.

The main takeaways of our model are the following. First, inventors facing stronger competition increase their research effort more in response to higher profitability expectations after natural disaster exposure. Since competition only affects pecuniary incentives and not intrinsic motivation, evidence supporting Hypothesis 1 would show that shifts in higher-order beliefs about future profits play a key role in driving innovation decisions. Second, heterogenous salience of climate change leads to higher aggregate costs for the same research output than homogenous salience of climate change would. This is due to research costs being convex. Third, a larger green good market induces stronger responses to changes in inventors intrinsic utility and expectations about consumers' valuation of green goods. Fourth, the overall private research rate in the economy is lower than the socially optimal research rate.

1.6.2 Empirical Results on Model Hypothesis

In this section, we test how market conditions and changes in inventors' expectations interact. Our findings reveal that first, disaster-induced changes in inventors monetary expectations matter for their innovation. Second, a well-functioning market is essential to ensure that inventors respond to changes in the salience of climate change.

Competition

We first empirically test Hypothesis 1 of the model, that states if inventors are primarily profit motivated, they should exhibit stronger green innovation responses in more competitive industries. To briefly summarize the intuition behind this hypothesis: a monopolist does not have profit incentives to pursue green innovation, as green product differentiation does not increase her profits above the monopoly profits she already enjoys. An inventor in a competitive environment, on the other hand, stands to gain substantial monetary gains from differentiating their product away from the competition.

To empirically test this hypothesis, we measure competition using industry-level profit margins from CompNet (2022), following Aghion et al. (2023). Higher profit margins reflect lower competition. A key advantage of this measure is that it captures international competition, unlike the Herfindahl–Hirschman index, which is country-specific. The CompNet database consolidates administrative firm-level data across European countries and reports aggregated indicators at various levels of industry and geography. For

France and Germany, the data are available from 2003 and 2001, respectively. Both CompNet and PATSTAT use the European Classification of Economic Activities (NACE Rev. 2), allowing us to link patent filings to industry-level competition at the 2-digit level, the most granular available in CompNet. For patents associated with multiple industries, we use industry weights to compute a composite measure. For each patent, we calculate the associated profit margin $M(i)$ of patent i as:

$$M_{it} = \sum_c \omega_{ic} \sum_k w_{ik} \times \frac{(margin_{kc,t} + margin_{kc,t-1})}{2}$$

where ω_{ic} represents the share of patent i's inventors living in country c, and w_{ik} denotes the weight with which the patent belongs to a specific industry k. Lastly, $\frac{margin_{kc,t}+margin_{kc,t-1}}{2}$ is the average profit margin of industry k in country c during the year of filing and the prior year. The profit margin M_{it} of a patent i is thus the weighted average of the profit margins faced by its inventors at the time of invention and the year prior. For example, if a patent related to the automotive industry was filed in 2004 by one French and one German inventor, the associated profit margin would be the mean of the profit margins for both the German and French automotive industries in 2003 and 2004. In Appendix section 4.2.10, we show our results for a 1-year and 3-year window of the profit margin. Our results are robust to the use of these alternative windows.

Instead of splitting the sample based on the overall median level of competition across all patents, we conduct the split within industries. This avoids comparing structurally distinct sectors, such as the highly competitive LED industry and the less competitive airline industry, that may differ for reasons unrelated to competition intensity. Instead, we compare patents within a given industry during periods of relatively high and low competition. We define the "high-competition" group as patents with above-median competition within their industry, and the "low-competition" group as those below the within-industry median.

To implement this, we calculate a patent-specific benchmark competition level (BMC_i), based on the median competition level of each 2-digit NACE industry across all years. For each patent i, we compute:

$$BMC_i = \sum_c \omega_{ic} \sum_k w_{ik} \times \text{median(margin)}_{kc}$$

where $\text{median(margin)}_{kc}$ denotes the median competition level in industry k and country c, computed over the entire sample period. The weights ω_{ic} and w_{ik} reflect the country and industry affiliations of patent i, respectively. If a patent's observed competition level M_{it} exceeds its benchmark BMC_i, it is classified into the high-competition group; otherwise, it is assigned to the low-competition group. Since this procedure leads to different sample

averages across groups, we normalize the outcome variable by the sample mean within each group to ensure comparability of results.

After splitting, we aggregate patents at the regional level separately for the high- and low-competition samples, resulting in two distinct region-by-year panels. For each sample, we estimate a difference-in-differences regression that compares changes in patenting activity before and after natural disasters across affected and unaffected regions. This allows us to separately identify the effect of natural disaster exposure on patenting for inventors in highly competitive and less competitive markets. By comparing the estimated disaster effects across the two groups, we can assess whether stronger pecuniary incentives (i.e., greater competition) amplify inventors' responses to disaster exposure. Columns (1) and (2) in Table 1.7 show our results.

Table 1.7: Competition and Green Product Split

	Dependent variable: $P(Y02_{lt})$			
	Competition Cutoff		Greenness Cutoff	
	High-Competition	Low-Competition	Above Median	Below Median
	(1)	(2)	(3)	(4)
Cumulative Count	0.101***	0.004	0.085***	0.061***
	(0.022)	(0.033)	(0.011)	(0.009)
Country-Year F.E.	Yes	Yes	Yes	Yes
Region F.E.	Yes	Yes	Yes	Yes
Wald-test p-value	0.0376		0.039**	
Sample Mean	1.9854	1.284	1.335	1.3248
Observations	8,283	8,283	14,307	14,307
R^2	0.652	0.534	0.623	0.787
Adj. R^2	0.617	0.486	0.601	0.775

Note: This table reports the results for our tests of the model's comparative statics. Columns (1)-(2) split the sample based on competition, while columns (3)-(4) split based on the greenness of the industry's products. Cumulative count is the count of past natural disasters. The Wald-tests examine if the coefficient for cumulative disaster count significantly differs between splits. We construct a Wald-test of the form $W = \frac{(\hat{\beta}_{\text{eq1}} - \hat{\beta}_{\text{eq2}})^2}{\text{Var}(\hat{\beta}_{\text{eq1}} - \hat{\beta}_{\text{eq2}})}$, where: $\text{Var}(\hat{\beta}_{\text{eq1}} - \hat{\beta}_{\text{eq2}}) = \text{Var}(\hat{\beta}_{\text{eq1}}) + \text{Var}(\hat{\beta}_{\text{eq2}}) - 2 \cdot \text{Cov}(\hat{\beta}_{\text{eq1}}, \hat{\beta}_{\text{eq2}})$. We can reject the Null hypothesis $H_0 : \beta_h = \beta_l$ against the alternative ($H_1 : \beta_h > \beta_l$) with the reported p values. Competition is measured as the average across filing year and the previous year. Greenness is calculated as average green product share across filing year and the previous year. Standard errors are clustered at the region level and reported in parentheses. Significance levels: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Inventors respond significantly more strongly to natural disaster exposure when operating in highly competitive environments. This finding provides strong support for our hypothesis and is consistent with the predictions of our theoretical model. The difference in response between high- and low-competition settings is both economically and statistically significant; we reject the null of coefficient equality with a p-value of 0.0376** (see Table 1.7 and the accompanying Wald test).

The lack of a response in low-competition environments suggests that pecuniary incentives play a central role in driving the observed innovation effects. If responses were primarily driven by intrinsic motivation, we would expect to see effects even in less competitive settings. Monetary incentives matter for how innovation responds to climate change. Our results highlight the importance of functioning competitive markets in en-

abling innovation to respond to climate-related shocks.

Green Good Demand

We next test Hypothesis 2, which predicts that a larger green goods market amplifies inventors' response to natural disaster exposure, as market size reduces the cost of innovation.

To empirically capture green market size, we use data from PRODCOM, a Eurostat database that reports annual production values for over 4,000 manufactured goods in Europe. Each product is classified using an 8-digit PRODCOM code, with the first four digits aligned to NACE industry codes. We identify green goods using the taxonomy developed by Bontadini and Vona (2023), which refines earlier lists compiled by the WTO and OECD. We update their list to reflect changes in PRODCOM codes over time (see Appendix 4.3.3) and restrict to data from 1995 to 2014 for France and Germany. For each industry-year cell, we compute the share of green goods by production value:

$$\text{Green Share}_{jt} = \frac{\sum_g y_{jt,g}}{\sum_g y_{jt,g} + \sum_{ng} y_{jt,ng}} \tag{1.19}$$

where $y_{jt,g}$ and $y_{jt,ng}$ denote the production values of green and non-green products, respectively, in industry j and year t. We assign each patent a corresponding green market share based on the industries to which it is linked. Since patents may span multiple industries, we compute a weighted average using industry weights ω_{ij}:

$$\text{Green Share Patent}_i = \sum_j \omega_{ij} \cdot \frac{\text{Green Share}_{jt} + \text{Green Share}_{jt-1}}{2} \tag{1.20}$$

where t is the filing year. This characterizes each patent by the green intensity of its market environment in the year of filing and the prior year. Results are robust to using a 1-year or 3-year window instead (Appendix 4.2.11). We then split the sample of green patents by whether they fall above or below the median green market share and estimate regressions separately for each group. Results are presented in columns (3) and (4) of Table 1.7.

We find that market size plays a significant role in shaping inventors' responses. Inventors in industries with larger green product markets respond more strongly to disaster exposure than those in less developed green markets. The difference is statistically significant (Wald test p-value = 0.039). These results align with the predictions of the directed technical change framework. Larger green markets amplify the expected returns to green innovation following disaster exposure. Since market size and inventor expectations are complementary, the effects are strongest where both are aligned.

Together, these findings underscore the central role of market incentives in shaping

innovation responses to climate change, and the importance of well-functioning, competitive markets in ensuring that inventors act on heightened climate salience.

1.7 Conclusion

This paper demonstrates that personal experiences affect inventors' research choices. We bridge the literatures on experience effects and the drivers of innovation and show that personally experiencing natural disasters significantly increases the invention of green technologies aimed at mitigation. This effect is highly localized and depends on direct personal experience. The effect is stronger in competitive markets and sectors with larger green demand.

Our empirical findings and theoretical framework point to a central mechanism, where the personal experience of natural disasters alters inventors' higher-order beliefs about consumer beliefs, increasing expectations of future green demand and regulatory tightening. This raises the perceived profitability of green R&D. The effect is strongest in competitive markets, where firms are more responsive to shifts in expected demand. While intrinsic motivation may also matter, our results highlight the central role of profit expectations in shaping innovation responses. These findings extend the experience-effects literature to the production side of the economy. Prior work has shown how personal experiences shape household expectations; we show they also influence high-stakes investment decisions with global implications.

We document an endogenous channel through which climate shocks affect the direction of technological change. Climate change is among the defining challenges of the twenty-first century. Its projected impacts, including more frequent floods (Hirabayashi et al. 2013; Roudier et al. 2016), deteriorating environmental conditions (Intergovernmental Panel on Climate Change (IPCC) 2023), and sea-level rise between 30 and 240 centimeters, are severe and global in scope. The economic literature highlights technological innovation, particularly in mitigation, as essential to addressing these risks (Acemoglu et al. 2012; Acemoglu et al. 2016). Yet, how innovation itself responds to climate change remains insufficiently understood. Our results suggest that forward-looking technological change may be more adaptive to rising climate risk than is typically assumed in integrated assessment models.

Beyond its theoretical contributions, our study has practical implications for policy. A well-functioning market and consumers' belief in anthropogenic climate change are crucial to ensuring that inventors act on increases in climate change salience. Our findings emphasize the role of private-sector incentives in shaping climate change-mitigating innovation.

Chapter 2

Welfare Weights and the Social Cost of Carbon

Joint with Felix Bierbrauer and Niklas Taft

Abstract

We use a simple model to study the interdependence of two policy design problems, an income tax problem and the design of a corrective tax. We clarify conditions under which corrective taxes vary along the Pareto-frontier. Under these conditions one needs to take stance on distributive issues to determine the social cost of carbon and the corresponding Pigouvian correction. We look at the problem both from a static and a dynamic perspective. There is a close correspondence between Pareto-efficient corrective taxes in the static model and the Pareto-efficient speed of the transition to a Carbon-neutral economy.

Acknowledgements. Felix Bierbrauer: University of Cologne & ECONtribute,
Marten Ritterrath: University of Cologne, Supported by the Joachim Herz Foundation.
Niklas Taft: University of Cologne

2.1 Introduction

Cost-benefit analysis of climate policy relies on the concept of the social cost of carbon, the present value of future income losses from climate change. Prior work examines how this measure depends on the discount rate and on how we account for uncertainty about future damages (Stern 2007; Tol 2009; Nordhaus 2017; Känzig 2023). We show that the social cost of carbon also depends on society's distributive preferences between rich and poor.

Public finance establishes separability results that generalize the results of Atkinson and Stiglitz (1976). Under separable preferences, good policy should address redistribution through income taxation alone. Other policies, commodity taxes, public goods provision, corrective taxes, should follow first-best principles. This logic suggests climate policy should use a first-best Pigouvian tax that makes individuals internalize the social cost of carbon when consuming carbon-intensive goods.

In this paper, we revisit the separable case and make one point: computing the social cost of carbon requires taking a stance on welfare weights. The first-best rule "set the corrective tax equal to the social cost of carbon" applies at every point along the Pareto frontier, but the corrective tax itself varies as we change welfare weights and move along the frontier. Even within the separable case, where corrective and redistributive motives are cleanly separated, the corrective tax is not uniquely pinned down.

We use the simplest possible setup. Two types of individuals differ in productive ability, as in Mirrleesian income tax models. These individuals consume a green good and a brown good. Aggregate consumption of the brown good creates an atmospheric externality. A separability assumption ensures that, conditional on disposable income, all individuals have the same marginal rate of substitution between green and brown goods. We then characterize the Pareto frontier using a mechanism design approach. We can reach any point on the frontier with a policy package combining nonlinear income taxation and a Pigouvian tax, although our setup is stated in terms of allocations alone.

The natural reference point for this exercise is the line of results that follows Atkinson and Stiglitz (1976). Under the separability conditions of that literature, redistribution should be carried out through the income tax alone, and commodity taxes should follow first-best principles independent of distributional concerns. Applied to a climate setting, the prescription is that the Pigouvian correction equals the social cost of carbon and is invariant to welfare weights (Cremer et al. 1998; Kaplow 2006). We take this prescription as our starting point and ask under which conditions it holds across the entire Pareto frontier.

We show that the invariance of the optimal carbon tax to welfare weights rests on a specific combination of assumptions. Under homogeneous environmental preferences, linear damages, and a CES consumption aggregator, the optimal Pigouvian tax is con-

stant along the entire Pareto frontier. As soon as one of these assumptions is relaxed, the location on the Pareto frontier matters for the level of the corrective tax.

Relaxing the linear-damages assumption introduces an equilibrium brown consumption channel. Even when both types value clean air identically, the marginal damage from brown consumption depends on the equilibrium level of aggregate brown consumption, which itself depends on the planner's distributive choice. A Rawlsian planner who favors the low type compresses labor supply and consumption to redistribute, which produces a lower aggregate level of brown consumption. Under convex damages, the marginal damage at this Rawlsian optimum is correspondingly lower, and the optimal carbon tax with it. A planner who favors the high type accepts a wider gap between the two types' allocations, sustains a higher aggregate brown consumption, and faces a higher marginal damage. The carbon tax inherits the distributive choice not through a redistributive motive, but through the equilibrium feedback of aggregate brown consumption on marginal damage. Relaxing the homogeneity of environmental preferences activates a separate, corrective channel. When the two types differ in how strongly they value clean air, the Pigouvian correction embeds a welfare-weighted average of the two damage valuations, and the planner's welfare weights determine whose valuation receives more weight. A Rawlsian planner who places more weight on the type with weaker environmental concern implements a lower corrective tax. A planner who places more weight on the type with stronger environmental concern implements a higher one. The corrective tax thereby acquires a redistributive role even in the separable Atkinson-Stiglitz setting, where the standard result says it should not. The two channels are logically distinct. The first operates through the level of brown consumption in equilibrium, the second through the welfare weighting of heterogeneous damage valuations. They reinforce one another when damages are convex and environmental preferences are heterogeneous.

We then extend the analysis to a dynamic setting and show that welfare weights shape not only how much society taxes carbon, but when it abandons it altogether and to what long-run damage level. In a model where the economy transitions from a brown to a green production technology at an endogenous switching time, the optimal timing of the green transition varies along the Pareto frontier in precisely the same way as the static carbon tax. An effective damage weight, analogous to the coefficient that governs the static Pigouvian tax, determines how the planner values future damages relative to the per-period cost of green production. This weight is lowest when the planner puts more weight on individuals with a weaker taste for environmental quality (in our setting the low type), and highest when it puts more weight on those with a stronger one. A Rawlsian planner who favors the environmentally less-concerned low type therefore delays the green transition, while a planner who favors the high type transitions earlier. The same distributional logic governs a second object, the terminal pollution stock at which accumulated emissions stabilize once the economy switches. A planner who delays the

transition also accepts a higher long-run stock, while a planner who transitions early stabilizes at a lower one. Timing and the long-run damage level are jointly distributive choices.

Contribution to the Literature: Our results primarily speak to two literatures. The social cost of carbon estimation literature (e.g. Stern 2007; Tol 2009; Nordhaus 2017; Känzig 2023) varies the discount rate and damage uncertainty but does not take a stance on welfare weights. Our result implies that existing SCC estimates are implicitly conditioned on a distributional choice that is never made explicit.

Our results suggest that introducing heterogeneity in productive abilities, in incomes, and in exposure to climate damages is a challenge not only for the positive analysis of how economic outcomes are affected by climate change and vice versa, but also for the normative analysis of climate policy. How to adequately price externalities is one of the core questions in public finance and traces back to Pigou (1920). Bovenberg and Goulder (1996), Cremer et al. (1998), Laroque (2005), and Kaplow (2006) investigate the optimal taxation of externalities in models that feature both income and commodity taxation. In the spirit of Atkinson and Stiglitz (1976), they show that the optimal commodity tax depends only on an objective damage criterion. More recent contributions such as Barrage (2020) bridge the gap between climate-economy models and this public finance literature by combining a DICE model with a fully fledged income tax schedule. A large part of this work is set in representative-agent environments, where the question of welfare weights does not arise and unweighted utilitarianism is the default by construction. Contributions that do allow for heterogeneity typically take a particular social welfare function as given and derive the corrective tax from it. Our result shows that the policy prescriptions of this literature change depending on where on the Pareto frontier one stands, and that the choice of welfare weights is unavoidable once heterogeneity in environmental preferences is admitted, even under Atkinson-Stiglitz separability.

Our finding that the optimal timing of the green transition varies along the Pareto frontier contributes to the literature on integrated assessment and optimal dynamic carbon taxation, which derives optimal transition paths within calibrated climate-economy models (Nordhaus 1992; Golosov et al. 2014; Rezai and Van der Ploeg 2016; Ploeg and Rezai 2019; Barrage 2020), and to a more recent strand that introduces household heterogeneity into these frameworks (Douenne et al. 2024a; Fried et al. 2024). Both take the social welfare function as given, as do earlier contributions in the Hotelling–backstop and directed-technical-change traditions (Van Der Ploeg and Withagen 2012; Acemoglu et al. 2012). We extend these literatures by showing that, the switching date varies systematically with welfare weights once environmental preferences are heterogeneous. The question of when to go green therefore embeds a distributional judgment that existing estimates of the optimal transition date leave implicit.

Deviations from the separability assumptions of Atkinson and Stiglitz (1976) have recently attracted interest. Foundational contributions include Stiglitz (1982) and Saez (2002). More recent work in this direction includes Ferey et al. (2024) and, in the climate policy context, Douenne et al. (2024b) and Bierbrauer (2024). We deliberately shut down this channel, which is known to provide a rationale for differential and non-linear commodity taxation. In our setting, every Pareto-efficient allocation is such that the marginal rate of substitution between green and brown consumption is equalized across all individuals. The Pigouvian correction is therefore the same for all.

Here we add a problem of corrective commodity taxation. The advantage of the simple two type economy that we study is that we can provide a complete, analytical and numerical, characterization of the whole Pareto-frontier. This is without precedence in the previous literature. An expositional advantage is that we can show with simple graphs how the Pigouvian correction varies in the set of Pareto-efficient allocations. We view this as an intuitive complement to the approach that needs to be taken in richer models, namely deriving an optimal tax formula that solves a Pareto- or a welfare-maximization problem and then providing comparative statics results with respect to the welfare weights that enter these formulas.

We add to the literature by explicitly characterizing the Pareto frontier for commodity taxation. Methodologically, this positions our paper closest to a literature on Pareto-efficient income taxation, including Stiglitz (1982), Werning (2007), Bierbrauer and Boyer (2014), and Bierbrauer et al. (2023). We derive our results under the typical setting of Atkinson and Stiglitz (1976), separability of utility and labor, and show that even in this setting welfare judgments enter the design of optimal commodity taxation.

We extend the theoretical literature by admitting heterogeneity in environmental preferences. We synthesize findings from the empirical literature on environmental preferences and policy support (e.g. Andre et al. 2024; Dechezleprêtre et al. 2025a; Stoetzer and Zimmermann 2024), which are so far absent from theoretical public finance. When we admit these heterogeneities, the social cost of carbon varies with the applied welfare criterion.

2.2 Model

We consider a world in which there is a continuum of agents, who maximize utility (2.1).

$$U(g_i, b_i, y_i) = \alpha(g_i, b_i) - \theta_i \, \phi(\pi_i b_i + \pi_j b_j) - k(y_i, \omega_i) \tag{2.1}$$

Each agent i's ($i \in \{l, h\}$) preferences consist of three parts. First, she derives direct consumption utility $\alpha(g_i, b_i)$ from the consumption of two goods. Consumption of a green good g produces utility only, while consumption of a brown product b is also associated

with a negative externality. The direct utility function α is uniform for all agents. Second, the agent suffers directly from an atmospheric externality. The impact of the externality $\phi(b_l, b_h)$ is monotonically increasing in the overall amount of brown consumption in the economy. While the accumulated amount of pollution is the same across our economy, its impact on individual utility is not. How much individuals care about the externality, e.g. clean air, depends on their preference type θ_i. Third, individuals have to exert effort when providing labor, which reduces overall utility. In this, agents differ with respect to their ability to generate income via their respective labor effort function $k(y_i, \omega_i)$. The effort cost function $k(y, \omega)$ is twice continuously differentiable, monotonically increasing in y_i and monotonically decreasing in ω_i (like in a classical model of Mirrleesian income taxation), satisfying $k_y > 0$, $k_{yy} > 0$, $k_\omega < 0$, $k_{y\omega} < 0$, and $k_y(y, \omega_h)/k_y(y, \omega_l)$ non-increasing in y.[1]

We specifically choose our preferences to be separable with respect to consumption utility, disutility from the externality and labor effort. In this, we are in line with the literature in the tradition of Atkinson and Stiglitz (1976). We use the simplest setup, considering only two types. The types (low and high ability) differ with regard to ability ω and preference on the externality θ_i.

To examine how the optimal carbon price behaves in response to how much weight the social planner puts on either type, we derive the Pareto frontier of utility. We let our social planner maximize low type utility, while keeping the high type's utility above a certain level. We incorporate this residual utility requirement (2.2) as an additional constraint in the optimization problem of the social planner.

$$U^h(g_h, b_h, y_h) = \alpha(g_h, b_h) - \theta_h\, \phi(\pi_l b_l + \pi_h b_h) - k(y_h, \omega_h) \geq \overline{V}_h \tag{2.2}$$

Since agents' type is private information, the optimization problem also features an incentive constraint for both types (with κ_l and κ_h being the respective Lagrange multipliers).

$$\begin{aligned} U^i(g_i, b_i, y_i) &= \alpha(g_i, b_i) - \theta_i\, \phi(\pi_i b_i + \pi_j b_j) - k(y_i, \omega_i) \\ &\geq U^i(g_j, b_j, y_j) = \alpha(g_j, b_j) - \theta_i\, \phi(\pi_i b_i + \pi_j b_j) - k(y_j, \omega_i) \end{aligned} \tag{2.3}$$

Finally, there is an aggregate resource constraint determined by income and an exogenous revenue requirement $\overline{R} \geq 0$, which denotes a per-period public-revenue requirement that the planner must finance through the income tax. The full Lagrangian and first-order conditions are given in Section 5.1.1 of the appendix.

We work throughout with the standard decentralization in which the planner deploys two instruments, a Pigouvian commodity tax $\tau_t \geq 0$ on brown consumption, and a nonlin-

[1] The last part implies that as effort rises, the productivity advantage of the high type (in terms of marginal effort cost) does not diminish.

ear Mirrlees income tax $T(y)$. In the brown phase the consumer faces an effective relative price of one unit of brown consumption equal to $1+\tau$, and in the green phase the technological cost is q. The planner internalizes the externality through τ and redistributes through $T(\cdot)$.

Two structural features of the model make the relevant separation results applicable. First, utility ((2.1) and for the dynamic extension(2.22)) is quasi-linear in consumption. Second, the externality enters utility additively as $-\theta_i\,\phi()$, where the environmental damage depends only on aggregate brown consumption and not on individual labor supply. Under these conditions the optimal-tax literature establishes a clean separation between the corrective and redistributive instruments. The result, due in various forms to Cremer et al. (1998), Pirttilä and Tuomala (2002), and Kaplow (2012), is that the income tax is the standard Mirrlees schedule designed as if no externality were present, while the commodity tax is set at the (welfare-weighted) marginal social damage. The two instruments do not interact at the optimum.

For the analysis that follows, the operative consequence of additivity is that the income-tax schedule preserves the standard Mirrlees properties on the consumer's labor-leisure margin. In the high-incentive-constraint (HIC) region, the high type faces no marginal income-tax distortion, so $k_y(y_h^B, \omega_h) = 1$ on the consumer's optimization problem. Symmetrically, in LIC the low type faces no marginal income-tax distortion, $k_y(y_l^B, \omega_l) = 1$. The Pigouvian wedge is borne entirely by the price of brown consumption, not by the labor margin.

2.2.1 The optimal carbon tax

The optimal carbon price is the marginal rate of substitution (MRS) between green and brown consumption, net of the marginal rate of transformation (MRT). By assumption, the MRT is constant and equal to 1. Dividing the FOC for b_l by the FOC for g_l, the coefficient $(1+\kappa_l-\kappa_h)$ cancels, giving

$$\frac{\alpha_b(g_l,b_l)}{\alpha_g(g_l,b_l)} = 1 + \frac{(\theta_l+\lambda\,\theta_h)\,\phi'\big(\pi_l b_l+\pi_h b_h\big)}{\psi}. \tag{2.4}$$

Dividing the FOC for b_h by the FOC for g_h yields the same expression for type h. Both types therefore face the same MRS at any point of the Pareto frontier. The optimal carbon tax takes the universal Samuelson form

$$\tau = \frac{(\theta_l+\lambda\,\theta_h)\,\phi'\big(\pi_l b_l+\pi_h b_h\big)}{\psi} \tag{2.5}$$

at every point on the Pareto frontier. The incentive multipliers κ_l and κ_h do not appear explicitly in this characterization. They enter only when we decompose the welfare weight

λ and the marginal value of public funds ψ through the consumption first-order conditions. The resulting decomposition is useful for the comparative statics that follow. Eliminating λ and ψ via the consumption FOCs (see Section 5.1.5 for the full derivation), we can express the carbon tax as

$$\tau = \phi'\big(\pi_l b_l + \pi_h b_h\big) \left[\frac{\left[\theta_l + (\kappa_l - \kappa_h)\,\theta_h\right]\pi_l}{(1 + \kappa_l - \kappa_h)\,\alpha_g(g_l, b_l)} + \frac{\theta_h\,\pi_h}{\alpha_g(g_h, b_h)} \right]. \tag{2.6}$$

Lemma 2.1. *At any point of the Pareto frontier, at most a single incentive constraint is binding.*

See Section 5.1.2 in the appendix for the proof of Lemma 2.1. From Lemma 2.1, there are three distinct regions along the Pareto frontier. In the high-type IC binding region ($\kappa_h > 0$, $\kappa_l = 0$), in the first-best region where no IC binds ($\kappa_l = \kappa_h = 0$), and in the low-type IC binding region ($\kappa_l > 0$, $\kappa_h = 0$). From (2.6), the MRS in the three regions follows by substitution. In the first-best region ($\kappa_l = \kappa_h = 0$),

$$\frac{\alpha_b(g_i, b_i)}{\alpha_g(g_i, b_i)} \;=\; 1 \;+\; \phi'\big(\pi_l b_l + \pi_h b_h\big) \left[\theta_l\,\frac{\pi_l}{\alpha_g(g_l, b_l)} + \theta_h\,\frac{\pi_h}{\alpha_g(g_h, b_h)} \right]. \tag{2.7}$$

When the low-type IC binds ($\kappa_l > 0$, $\kappa_h = 0$),

$$\frac{\alpha_b(g_i, b_i)}{\alpha_g(g_i, b_i)} \;=\; 1 \;+\; \phi'\big(\pi_l b_l + \pi_h b_h\big) \left[\frac{(\theta_l + \kappa_l\,\theta_h)\,\pi_l}{(1 + \kappa_l)\,\alpha_g(g_l, b_l)} \;+\; \frac{\theta_h\,\pi_h}{\alpha_g(g_h, b_h)} \right]. \tag{2.8}$$

When the high-type IC binds ($\kappa_h > 0$, $\kappa_l = 0$),

$$\frac{\alpha_b(g_i, b_i)}{\alpha_g(g_i, b_i)} \;=\; 1 \;+\; \phi'\big(\pi_l b_l + \pi_h b_h\big) \left[\frac{(\theta_l - \kappa_h\,\theta_h)\,\pi_l}{(1 - \kappa_h)\,\alpha_g(g_l, b_l)} \;+\; \frac{\theta_h\,\pi_h}{\alpha_g(g_h, b_h)} \right]. \tag{2.9}$$

At any point of the Pareto frontier, the optimal corrective tax is a uniform commodity tax across types. This follows directly from the derivation above. The FOC ratios for both types yield the same MRS, so the wedge between MRS and MRT, which the planner implements as a commodity tax on brown consumption, is identical for both types. The optimal corrective tax equates the damages from the externality weighted at a certain point of the Pareto frontier with the loss of welfare that results from a deviation from the MRT. If there are no damages ($\phi' = 0$), the wedge disappears altogether.

2.3 Main Result

The integrated assessment and climate-economy literature treats the social cost of carbon as a single number to be estimated. Estimates vary with the discount rate (Stern 2007; Nordhaus 2017), with assumptions about damages (Tol 2009), and with the treatment of

uncertainty (Känzig 2023). The underlying object is taken to be well-defined once these inputs are fixed. We show that this is not so. The social cost of carbon, and the optimal Pigouvian correction associated with it, depends on the planner's welfare weights. The dependence does not require departures from the separability conditions of the public finance literature. It survives the Atkinson–Stiglitz benchmark in which redistribution and correction are formally distinct.

The dependence operates through two channels. The first is an equilibrium channel. The optimal carbon tax tracks the marginal damage, the marginal damage depends on the equilibrium level of aggregate brown consumption, and aggregate brown consumption depends on the planner's distributive choice. The second is a corrective channel. When types differ in their environmental preferences, the corrective tax embeds a welfare-weighted average of the affected agents' damages, and the weighting shifts along the Pareto frontier. The two channels are logically independent. They operate jointly in the empirically relevant case in which environmental preferences are heterogeneous and damages are convex, and they reinforce one another when the planner favors the environmentally more concerned type.

We organize the section as a sequence of perturbations away from an invariance benchmark. The benchmark requires three structural conditions to hold at once. Environmental preferences must be homogeneous across types. Consumption must aggregate through a CES technology and damages must be linear in aggregate brown consumption. Utility must satisfy the Atkinson–Stiglitz separability conditions, which is built into the utility specification (2.1). Under all three, the optimal carbon tax takes a single numerical value that is invariant to the planner's distributional stance. Relaxing the linear-damages assumption activates the equilibrium channel. Relaxing the homogeneity assumption activates the corrective channel. The benchmark is knife-edge and the social cost of carbon is generically a function of welfare weights.

2.3.1 The Invariance Benchmark

We begin with the case in which the social cost of carbon is genuinely invariant to welfare weights. This is the reference point against which all subsequent results are stated. The benchmark requires structural assumptions on damages and on consumption.

Assumption 2.1. *Damages are linear, $\phi(B) = B$, so $\phi' \equiv 1$.*

Assumption 2.2. *Direct utility over consumption is CES,*

$$\alpha(g_i, b_i) \;=\; \left(a\, g_i^{\rho} + (1-a)\, b_i^{\rho}\right)^{1/\rho}, \tag{2.10}$$

with $a \in (0,1)$ and $\rho \in (-\infty, 1) \setminus \{0\}$. Let $\sigma \equiv 1/(1-\rho) \in (0,\infty) \setminus \{1\}$ denote the elasticity of substitution between green and brown goods.

The Cobb–Douglas case $\alpha = g^a b^{1-a}$ obtains as the limit $\rho \to 0$ ($\sigma \to 1$), and the analysis below extends to it by continuity. The CES aggregator is linearly homogeneous, so α_g is homogeneous of degree zero in (g, b) and depends only on the common ratio g/b. As we show below, this property is what allows the endogenous consumption-side denominators in the carbon-tax formula to collapse to a function of τ alone.

Proposition 2.1. *Under Assumptions 2.1–2.2, if preferences are homogeneous, $\theta_l = \theta_h = \theta$, the optimal carbon tax does not vary along the Pareto frontier. It is the unique solution of*

$$H(\tau) \;\equiv\; \tau \left[\, a^\sigma + (1-a)^\sigma (1+\tau)^{1-\sigma} \,\right]^{1/(\sigma-1)} \;=\; \theta. \tag{2.11}$$

The tax is determined entirely by the common damage parameter θ and the consumption-side primitives (a, σ). It does not depend on welfare weights, on the incentive-constraint multipliers, on the equilibrium allocation, or on the specification of labor effort costs. A planner who knew θ, a, and σ would report a single social cost of carbon without taking any stance on whose welfare to weight more heavily. The proof is given in Section 5.1.6 of the appendix.

The first observation is that homogeneous environmental preferences, on their own, do not deliver invariance. They eliminate the corrective channel but leave the equilibrium channel open. The next proposition states the result without imposing CES consumption or linear damages.

Proposition 2.2. *If preferences are homogeneous, $\theta_l = \theta_h = \theta$, the optimal carbon tax along the Pareto frontier is given by*

$$\tau \;=\; \theta\, \phi'\big(\pi_l b_l + \pi_h b_h\big) \left[\frac{\pi_l}{\alpha_g(g_l, b_l)} + \frac{\pi_h}{\alpha_g(g_h, b_h)} \right], \tag{2.12}$$

independent of which incentive constraint is binding. The Pigouvian correction depends only on the common damage parameter θ, on the technology of pollution, and on the equilibrium consumption allocation. It does not depend on the redistributive structure of the optimum.

The proof is in Section 5.1.3 of the appendix. Under homogeneous preferences the externality enters both types' utilities symmetrically. Aggregate pollution is the same regardless of which bundle an agent claims, so the externality terms cancel within both incentive constraints. The corrective tax is therefore independent of the incentive-constraint multipliers, and its formula (2.12) is the same across all three regions of the Pareto frontier.

Invariance of the formula is not invariance of the tax. The Pigouvian correction in (2.12) still depends on the equilibrium consumption allocation (g_l, b_l, g_h, b_h) through the consumption-side denominators, and on the marginal damage $\phi'(\pi_l b_l + \pi_h b_h)$ evaluated

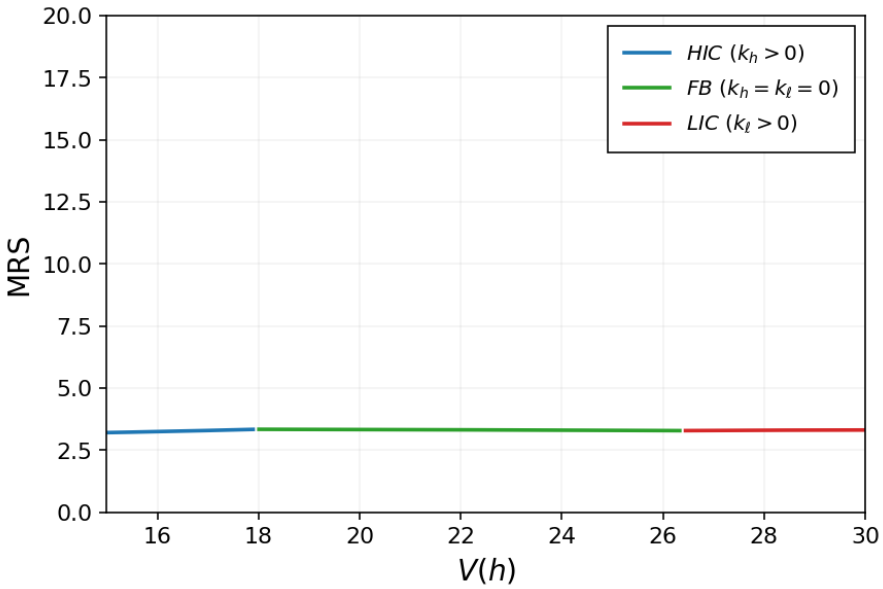


Figure 2.1: **Homogeneous preferences: MRS by V_h**

The figure depicts the MRS between green and brown consumption as a function of the welfare weight of the high type $V(h)$. Consumption utility is $\alpha(g, b) = 2g^\alpha b^\alpha$ with $\alpha = 0.4$, effort costs are $k(y; \omega) = (y/\omega)^2$ with $\omega_\ell = 21$, $\omega_h = 35$. Damages from brown consumption are $\theta_i \, \phi(\pi_l b_l + \pi_h b_h)$ with ϕ being an identity function, with $\theta_h = 2$ and $\theta_l = 0.5$ and equal population weights.

at the equilibrium level of aggregate brown consumption. Both objects respond to the planner's distributive choice. The CES and linear-damages assumptions of the benchmark eliminate this dependence. Without them, homogeneity of environmental preferences is not sufficient for invariance.

Figure 2.1 illustrates this point. It shows the MRS along the pareto frontier. The MRS and thus the optimal carbon tax does not vary with the reservation utility V_h.

2.3.2 Equilibrium Brown Consumption

We now look at how equilibrium brown consumption affects the corrective tax. We keep environmental preferences homogeneous and CES consumption in place, and we replace linear damages with the simplest strictly convex specification. The Pigouvian correction is no longer pinned down by primitives. It depends on the planner's distributive choice through the equilibrium level of aggregate brown consumption.

Under linear damages, $\phi'(B)$ is a constant, and CES consumption ensures the consumption-side denominators in (2.12) reduce to a function of τ alone. The carbon tax is fully pinned down by primitives. Under strictly convex damages, $\phi'(B)$ is sensitive to the level of aggregate brown consumption. Aggregate brown consumption in turn depends on the equilibrium allocation, which itself responds to the planner's distributive choice. A redistributive optimum that distorts labor supply substantially produces less aggregate consumption, and therefore less aggregate brown consumption, than an optimum

that distorts little. Under convex damages, the marginal damage at the redistributive optimum is correspondingly smaller, and the optimal Pigouvian correction is smaller too. The corrective tax inherits the distributive choice not through who values the externality more, but through how much is consumed in equilibrium.

We formalize the argument for the Rawlsian direction of the Pareto frontier, where the high type's incentive constraint binds and the planner favors the low type. To deliver closed-form expressions, we strengthen the consumption-side assumption of Proposition 2.1 with two additional structural requirements.

Assumption 2.3. *Damages are strictly convex of the power form*

$$\phi(B) \;=\; \frac{d}{1+\gamma}\, B^{1+\gamma}, \qquad \phi'(B) \;=\; d\, B^{\gamma}, \qquad d>0,\ \gamma>0. \tag{2.13}$$

Assumption 2.4. *Effort costs are iso-elastic in effort $\ell_i = y_i/\omega_i$,*

$$k(y,\omega) \;=\; \frac{(y/\omega)^{1+1/\varepsilon}}{1+1/\varepsilon}, \qquad k_y(y,\omega) \;=\; \frac{1}{\omega}\Big(\frac{y}{\omega}\Big)^{1/\varepsilon}, \qquad \varepsilon>0. \tag{2.14}$$

Assumption 2.3 replaces the linear-damages benchmark of Assumption 2.1 with the simplest strictly convex specification. The parameter $\gamma > 0$ governs the curvature of damages. Setting $\gamma = 0$ recovers linear damages and the invariance of Proposition 2.1. Assumption 2.4 is the standard iso-elastic effort specification used in Mirrlees-type models. It delivers tractable closed forms for the equilibrium labor allocations under a binding incentive constraint, which is what we need to sign the comparative static below.

Under Assumption 2.2, both types share a common ratio g/b at any point on the Pareto frontier (see the proof of Proposition 2.5), so $\alpha_g(g_l, b_l) = \alpha_g(g_h, b_h) \equiv \alpha_g^*(\tau)$ is a function of τ alone. Substituting into the carbon-tax formula of Proposition 2.2 under homogeneous preferences,

$$H(\tau) \;\equiv\; \tau\, \alpha_g^*(\tau) \;=\; \theta\, \phi'\big(B(\lambda)\big) \;=\; \theta\, d\, B(\lambda)^{\gamma}, \tag{2.15}$$

where $B(\lambda) = \pi_l b_l + \pi_h b_h$ denotes aggregate brown consumption at the Pareto-efficient allocation indexed by the planner's welfare weight λ on the high type. The left-hand side is the strictly increasing map H used in the CES characterization of Proposition 2.5. The right-hand side, in contrast to the linear-damages case, depends on the planner's distributive choice through $B(\lambda)$ whenever $\gamma > 0$. Whenever $B(\lambda)$ varies with λ, the carbon tax varies with it.

The next lemma signs the comparative static of $B(\lambda)$ on the Rawlsian half of the frontier.

Lemma 2.2. *Under Assumption 2.4, aggregate brown consumption $B(\lambda)$ is strictly increasing in λ on the HIC region of the Pareto frontier.*

The argument proceeds in the standard Mirrleesian way. In the HIC region the high type's labor margin is undistorted, so y_h is constant at its laissez-faire level. The low type's labor margin is distorted downward to relax the high type's incentive constraint, and the size of the downward distortion is governed by κ_h. Moving toward the first-best, equivalently raising λ from the Rawlsian limit toward π_h/π_l, shrinks κ_h toward zero and relaxes the distortion. The low type's income y_l rises strictly, aggregate income rises with it, and under CES consumption a strictly larger share of the additional income flows into brown consumption. Aggregate brown consumption therefore rises strictly. The full proof is in Section 5.1.4.

Combining the lemma with the implicit equation (2.15) delivers the result.

Proposition 2.3. *Suppose preferences are homogeneous, $\theta_l = \theta_h = \theta$, and Assumptions 2.2, 2.3, and 2.4 hold. Then the optimal carbon tax is strictly increasing in λ on the HIC region of the Pareto frontier. In particular, the carbon tax varies with welfare weights along the frontier even though environmental preferences are common across types.*

Proof. By Proposition 2.2 and the CES collapse, the optimal carbon tax under homogeneous preferences satisfies the implicit equation (2.15). The map $H(\tau) = \tau\,\alpha_g^*(\tau)$ is strictly increasing on $(-1, \infty)$ by Proposition 2.5. The right-hand side, $\theta\, d\, B(\lambda)^\gamma$, is strictly increasing in λ on the HIC region by Lemma 2.2 together with $\gamma > 0$. Applying H^{-1} to a strictly increasing function of λ yields a function that is itself strictly increasing in λ, so $\tau(\lambda)$ is strictly increasing on the HIC region. □

The economic content of the proposition is that the Pigouvian principle "set τ equal to marginal damage" is silent on which marginal damage. In a heterogeneous-agent economy with convex damages, the marginal damage at a Rawlsian optimum and the marginal damage at a Paretian optimum favoring the high type are not the same object. The Rawlsian planner achieves redistribution by compressing labor supply and consumption, including brown consumption, across the population. Under convex damages, the resulting low aggregate B generates a low $\phi'(B)$, and the appropriate Pigouvian correction is correspondingly small. A planner who places more weight on the high type accepts a wider gap between the high and low type's allocations, with the high type producing and consuming more in equilibrium. Aggregate B is then larger, $\phi'(B)$ is larger, and the appropriate correction is larger. The carbon tax tracks the equilibrium level of damages, and the equilibrium level of damages tracks the distributive choice.

Two features of Proposition 2.3 are worth emphasizing. First, the variation in τ along the frontier here is qualitatively distinct from the variation established under heterogeneous preferences in Section 2.3.3 below. There, the corrective component of the tax shifts because the planner places different weight on the environmental concerns of different types. Here, the corrective component is the common parameter θ, and what shifts is

the marginal damage $\phi'(B)$. The mechanism is therefore not a redistributive distortion of the corrective motive, but an equilibrium feedback from the level of brown consumption. The two mechanisms can operate in opposite directions or reinforce one another. Under heterogeneous θ with $\theta_h > \theta_l$ and convex damages, they reinforce one another on the HIC region of the frontier.

Second, the result clarifies the role of linear damages in Proposition 2.1. The full invariance of τ in the benchmark is not a consequence of homogeneous environmental preferences alone, nor of the CES consumption aggregator alone. It additionally requires that the marginal damage be insensitive to the equilibrium level of brown consumption, which is exactly what linear damages deliver. Whenever damages are strictly convex, the carbon tax in a homogeneous-preference economy is no longer a primitive. It depends on the planner's distributive choice through the equilibrium allocation, even when the corrective and redistributive motives are formally separable in the Atkinson–Stiglitz sense.

2.3.3 Heterogeneous Preferences

We now turn to the corrective channel. We restore linear damages and let environmental preferences differ across types. Heterogeneity in θ activates a separate margin along which welfare weights enter the carbon tax. The mechanism is independent of the convexity of damages and operates even when the equilibrium channel is shut down.

There are two ways to interpret heterogeneity in θ. One is variation in impact, with the high type suffering more from a given level of aggregate pollution. The other is variation in perception. When the damages of an externality are difficult to estimate, beliefs and expectations can differ across the population. The empirical literature on environmental attitudes and policy support documents both (Andre et al. 2024; Gill et al. 2023; Dechezleprêtre et al. 2025a; Stoetzer and Zimmermann 2024). We take $\theta_h > \theta_l$ throughout this subsection without further interpretive commitment.

Proposition 2.4. *Suppose preferences on the externality differ across types, $\theta_h \neq \theta_l$. Then the optimal carbon tax varies with welfare weights along the Pareto frontier.*

Proof. From (2.6), the optimal carbon tax can be written as

$$\tau \;=\; \phi'\big(\pi_l b_l + \pi_h b_h\big)\left[\, f(\kappa_l,\kappa_h)\,\frac{\pi_l}{\alpha_g(g_l,b_l)} + \frac{\theta_h\,\pi_h}{\alpha_g(g_h,b_h)}\right], \tag{2.16}$$

where

$$f(\kappa_l,\kappa_h) \;\equiv\; \frac{\theta_l + (\kappa_l-\kappa_h)\,\theta_h}{1+\kappa_l-\kappa_h}. \tag{2.17}$$

At the FB–LIC boundary, $\kappa_l = 0$ and the equilibrium allocations are continuous. A

marginal increase in κ_l above zero changes f by

$$\left.\frac{\partial f}{\partial \kappa_l}\right|_{\kappa_l=\kappa_h=0} \;=\; \theta_h - \theta_l \;\neq\; 0, \tag{2.18}$$

so f moves strictly away from θ_l as the LIC region is entered. By continuity of the equilibrium consumption allocations at the boundary, the bracketed term in τ changes by a nonzero amount, and hence τ changes. The same argument applies at the FB–HIC boundary, where $\partial f/\partial \kappa_h\big|_{\kappa_l=\kappa_h=0} = -(\theta_h - \theta_l) \neq 0$. Therefore τ is not constant along the Pareto frontier whenever $\theta_h \neq \theta_l$. The complete proof is in Section 5.1.5. □

Proposition 2.4 establishes that the carbon tax varies across the three regions of the Pareto frontier and that the changes at the FB–LIC and FB–HIC boundaries have opposite signs. To order τ across regions in closed form we restore the CES consumption-side assumption of the benchmark and keep linear damages so that the equilibrium channel of Section 2.3.2 is shut down.

Proposition 2.5. *Under Assumptions 2.1–2.2 the optimal carbon tax along the Pareto frontier satisfies the implicit equation*

$$H(\tau) \;\equiv\; \tau\,\alpha_g^*(\tau) \;=\; f(\kappa_l,\kappa_h)\,\pi_l + \theta_h\,\pi_h \;=\; \theta_h - \frac{\pi_l(\theta_h-\theta_l)}{1+\kappa_l-\kappa_h}, \tag{2.19}$$

where

$$\alpha_g^*(\tau) \;=\; \left[\, a^\sigma + (1-a)^\sigma (1+\tau)^{1-\sigma} \,\right]^{1/(\sigma-1)}. \tag{2.20}$$

The map $H : (-1,\infty) \to \mathbb{R}$ is strictly increasing. Whenever $\theta_h > \theta_l$, the carbon tax is non-decreasing as Pareto weight shifts to the high type. It is strictly increasing within HIC and within LIC, and constant within FB:

$$\tau^{HIC} \;<\; \tau^{FB} \;<\; \tau^{LIC}, \tag{2.21}$$

with τ^{FB} the unique solution of $H(\tau) = \theta_l\pi_l + \theta_h\pi_h$.

Figure 2.2 illustrates Proposition 2.5 for a parameterized example. The MRS rises monotonically as the welfare weight on the high type increases. The region left of the first kink corresponds to a binding high-type incentive constraint, and the region right of the second kink to a binding low-type incentive constraint. A planner who attributes higher weight to the high type accommodates the larger distaste for the externality by imposing a higher Pigouvian tax. The tax is uniform across types at every point on the frontier and distorts brown consumption downward for both.

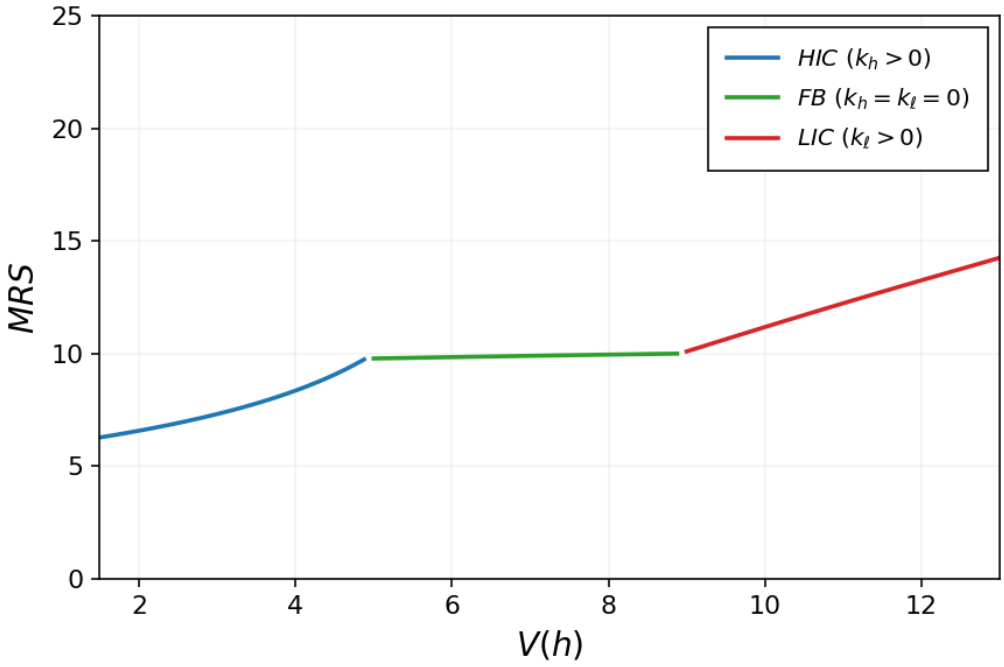


Figure 2.2: **Heterogeneous preferences: MRS by V_h**

The figure depicts the MRS between green and brown consumption as a function of the welfare weight of the high type $V(h)$. Consumption utility is $\alpha(g, b) = 2g^{\alpha}b^{\alpha}$ with $\alpha = 0.4$, effort costs are $k(y; \omega) = (y/\omega)^2$ with $\omega_\ell = 21$, $\omega_h = 35$. Damages from brown consumption are $\theta_i\, \phi(\pi_l b_l + \pi_h b_h)$ with ϕ being an identity function, with $\theta_h = 2$ and $\theta_l = 0.5$ and equal population weights.

2.4 Dynamic Extension: When Does the Economy Become Green?

The static model shows that the level of the optimal carbon tax varies along the Pareto frontier when environmental preferences are heterogeneous. We now extend the framework to a dynamic setting. The question becomes: when does the economy switch from brown to green production? We show that the optimal switching time varies along the Pareto frontier in an analogous way. Welfare weights determine not only how much to tax carbon, but when to abandon it.

2.4.1 Setup

Time is continuous, $t \in [0, \infty)$. As before, there are two types $i \in \{l, h\}$ with population shares π_l and π_h. Types differ in ability ω_i and in their valuation of environmental damages θ_i.

At each instant, each agent consumes a single consumption good produced by one of two technologies. Brown technology converts one unit of income into one unit of consumption but adds one unit to the pollution stock. Green technology converts one unit of income into $1/q$ units of consumption ($q > 1$) and generates no pollution. Both

technologies deliver the same consumption good, so preferences over consumption are:

$$U_i = \int_0^\infty e^{-\rho t}\Big[c_{it} - \theta_i\,\phi(S_t) - k(y_{it},\omega_i)\Big]\,dt, \tag{2.22}$$

where c_{it} is consumption, y_{it} is pre-tax income, $\rho > 0$ is the common discount rate, and ϕ is a convex damage function ($\phi' > 0$, $\phi'' > 0$, $\phi(0) = 0$). The pollution stock accumulates without decay:

$$\dot{S}_t = \pi_l\, b_{lt} + \pi_h\, b_{ht}, \qquad S_0 = 0, \tag{2.23}$$

where b_{it} denotes brown consumption of type i at time t.

We adopt the following functional forms throughout. The damage function is the power form

$$\phi(S) \;=\; \frac{d}{1+\gamma}\,S^{1+\gamma}, \qquad \phi'(S) \;=\; d\,S^{\gamma}, \qquad d > 0,\ \gamma > 0, \tag{2.24}$$

and the labor effort cost is iso-elastic in effort $\ell_i = y_i/\omega_i$:

$$k(y_i,\omega_i) \;=\; \frac{(y_i/\omega_i)^{1+1/\varepsilon}}{1+1/\varepsilon}, \qquad k_y(y_i,\omega_i) \;=\; \frac{1}{\omega_i}\left(\frac{y_i}{\omega_i}\right)^{1/\varepsilon}, \qquad \varepsilon > 0. \tag{2.25}$$

Since green and brown consumption are perfect substitutes in utility, the planner faces a bang-bang choice at each instant: produce all consumption using brown technology or all using green technology. We index this choice by a switching time T^*. For $t < T^*$, all consumption is brown ($b_{it} = c_{it}$). For $t \geq T^*$, all consumption is green ($b_{it} = 0$).

Within each phase, the per-period allocation problem is time-independent. The planner chooses stationary allocations (c_i^B, y_i^B) during the brown phase and (c_i^G, y_i^G) during the green phase. The per-period resource constraints are

$$\text{Brown phase:} \quad \pi_l\, y_l^B + \pi_h\, y_h^B \;\geq\; \pi_l\, c_l^B + \pi_h\, c_h^B + \bar{R}, \tag{2.26}$$

$$\text{Green phase:} \quad \pi_l\, y_l^G + \pi_h\, y_h^G \;\geq\; q\big(\pi_l\, c_l^G + \pi_h\, c_h^G\big) + \bar{R}, \tag{2.27}$$

where $\bar{R} \geq 0$ denotes an exogenous per-period public-revenue requirement that the planner must finance through the income tax. Aggregate pre-tax income and aggregate brown consumption in the brown phase are

$$Y_B \;\equiv\; \pi_l\, y_l^B + \pi_h\, y_h^B, \qquad B \;\equiv\; \pi_l\, c_l^B + \pi_h\, c_h^B. \tag{2.28}$$

With (2.26) binding, $B = Y_B - \bar{R}$. Both objects play a central role below. The pollution stock at the switching time is

$$S^* \;\equiv\; S(T^*) \;=\; B\,T^*. \tag{2.29}$$

For $t \geq T^*$, the stock remains at S^*.

2.4.2 Incentive Compatibility

Types are private information. At $t = 0$, each agent reports a type, and the planner assigns a lifetime allocation stream based on this report. The incentive constraint for type i requires that truthful reporting yields weakly higher lifetime utility than mimicking type j.

The externality term $\theta_i \, \phi(S_t)$ is determined by aggregate brown consumption and does not depend on individual reports. It therefore cancels from the incentive constraints, exactly as in the static model. The lifetime incentive constraint for type i reduces to:

$$D_B\Big[c_i^B - k(y_i^B, \omega_i)\Big] + D_G\Big[c_i^G - k(y_i^G, \omega_i)\Big] \;\geq\; D_B\Big[c_j^B - k(y_j^B, \omega_i)\Big] + D_G\Big[c_j^G - k(y_j^G, \omega_i)\Big], \tag{2.30}$$

where $D_B = (1 - e^{-\rho T^*})/\rho$ and $D_G = e^{-\rho T^*}/\rho$ are the discounted durations of the brown and green phases.

2.4.3 The Planner's Problem

The planner maximizes the low type's lifetime welfare subject to a reservation utility $\bar{V}_h$ for the high type:

$$\max_{T^*,\,(c_i^B, y_i^B),\,(c_i^G, y_i^G)} \; W_l + \lambda\big(W_h - \bar{V}_h\big) + \text{IC and RC terms}, \tag{2.31}$$

where the lifetime welfare of type i is:

$$W_i = D_B \, v_i^B + D_G \, v_i^G - \theta_i \, \Gamma(T^*), \tag{2.32}$$

with per-period flow utility $v_i^B = c_i^B - k(y_i^B, \omega_i)$, $v_i^G = c_i^G - k(y_i^G, \omega_i)$, and cumulative discounted damages:

$$\Gamma(T^*) = \int_0^{T^*} e^{-\rho t} \, \phi(B\,t) \, dt \;+\; \frac{e^{-\rho T^*}}{\rho} \, \phi(S^*). \tag{2.33}$$

The first term captures rising damages during the brown phase. The second term captures the permanent damage burden after the switch. The full Lagrangian and first-order conditions are given in Section 5.2.1 of the appendix.

2.4.4 Optimal Switching Time

We characterize the optimal switching time using a Lagrangian approach. Differentiating the Lagrangian with respect to T^* yields a first-order condition that equates the welfare

cost of extending the brown phase by one period with the present value of the additional future damages that period creates. The full derivation from the Lagrangian, including the κ-terms that arise from the incentive constraints, is given in Section 5.2.2. Eliminating λ via the within-phase consumption FOCs (see Section 5.2.2 for the derivation), the complete first-order condition can be written as

$$\begin{aligned} &\pi_l\big(v_l^B - v_l^G\big) + \big(\pi_h + \kappa_l - \kappa_h\big)\big(v_h^B - v_h^G\big) + \pi_l\big(\kappa_l\,\Delta_l + \kappa_h\,\Delta_h\big) \\ &\quad = \ \big[\pi_l\,\theta_l + \big(\pi_h + \kappa_l - \kappa_h\big)\,\theta_h\big]\,\frac{B\,\phi'(S^*)}{\rho}, \end{aligned} \tag{2.34}$$

where $\Delta_i \equiv (v_i^B - \tilde{v}_{ij}^B) - (v_i^G - \tilde{v}_{ij}^G)$ is the change in incentive slack of type i between the brown and green phases. The κ-terms arise because the lifetime incentive constraints involve the discounted phase durations D_B and D_G, which depend on T^*. By Lemma 1, at most one of these terms is nonzero at any point of the Pareto frontier. The only Lagrange multipliers that enter the switching condition are the incentive constraint multipliers κ_l and κ_h.

In the first-best region ($\kappa_l = \kappa_h = 0$), the switching condition simplifies to

$$\underbrace{\pi_l\big(v_l^B - v_l^G\big) + \pi_h\big(v_h^B - v_h^G\big)}_{\text{population-weighted cost of going green}} \ = \ \underbrace{\big(\pi_l\,\theta_l + \pi_h\,\theta_h\big)\,\frac{B\,\phi'(S^*)}{\rho}}_{\text{PV of avoided future damages}}. \tag{2.35}$$

The left-hand side is the population-weighted per-period cost of consuming green rather than brown. Green production is less efficient ($q > 1$), so the green-phase allocations deliver lower flow utility, with $v_i^G < v_i^B$. The right-hand side is the present value of avoided future marginal damages from one additional period of brown consumption. The damage valuation is weighted by $\pi_l\,\theta_l + \pi_h\,\theta_h$, the population-weighted environmental preference. The planner switches when these are equal.

2.4.5 The switching condition in λ-Form

The high type's welfare weight λ is tied to the incentive multipliers by

$$\lambda \ = \ \frac{\pi_h + \kappa_l - \kappa_h}{\pi_l}, \qquad 1 + \kappa_l - \kappa_h \ = \ \pi_l(1 + \lambda). \tag{2.36}$$

The three regions of the Pareto frontier partition cleanly in λ-space,

$$\text{HIC: } \lambda < \tfrac{\pi_h}{\pi_l}, \qquad \text{FB: } \lambda = \tfrac{\pi_h}{\pi_l}, \qquad \text{LIC: } \lambda > \tfrac{\pi_h}{\pi_l}. \tag{2.37}$$

Substituting $\pi_h + \kappa_l - \kappa_h = \pi_l \lambda$ on both sides of (2.34) and dividing through by π_l produces the master form of the switching FOC,

$$\left(v_l^B - v_l^G\right) + \lambda \left(v_h^B - v_h^G\right) + \kappa_l \Delta_l + \kappa_h \Delta_h = \left[\theta_l + \lambda \theta_h\right] \frac{B \phi'(S^*)}{\rho}. \tag{2.38}$$

With the power damage form (2.24) and $S^* = BT^*$, the right-hand side becomes $\left[\theta_l + \lambda\theta_h\right] d\, B^{1+\gamma}(T^*)^\gamma/\rho$. Define the effective damage valuation

$$\hat{\theta}(\lambda) \equiv \frac{\theta_l + \lambda\, \theta_h}{1+\lambda}, \qquad \frac{d\hat{\theta}}{d\lambda} = \frac{\theta_h - \theta_l}{(1+\lambda)^2}. \tag{2.39}$$

It is convenient to repackage the entire left-hand side as a single per-capita object. Define the *aggregate switching cost*

$$\Phi(\lambda) \equiv \left(v_l^B - v_l^G\right) + \lambda \left(v_h^B - v_h^G\right) + \kappa_l \Delta_l + \kappa_h \Delta_h \tag{2.40}$$

and its per-capita counterpart

$$\overline{\Delta v}(\lambda) \equiv \frac{\Phi(\lambda)}{1+\lambda}. \tag{2.41}$$

The switching condition (2.38) then takes the compact form

$$\overline{\Delta v}(\lambda) = \frac{d}{\rho}\, \hat{\theta}(\lambda)\, B(\lambda)^{1+\gamma}\, T^*(\lambda)^\gamma. \tag{2.42}$$

2.4.6 Structural identity and within-phase channel

The compact FOC (2.42) packages the entire comparative-statics problem into the joint behavior of three objects along the frontier: the per-capita switching cost $\overline{\Delta v}(\lambda)$, aggregate brown consumption $B(\lambda)$, and the effective damage valuation $\hat{\theta}(\lambda)$. This subsection establishes the structural properties of the first two that we draw on throughout the remainder of the section. The effective damage valuation is governed by primitives alone and is treated separately in the homogeneous and heterogeneous subsections that follow.

To make the analysis transparent, we exploit the iso-elastic specification (2.25) together with one mild restriction on the within-phase allocations.

Assumption 2.5. *The per-phase incentive constraints bind in their respective regions:* $\Delta_h = 0$ *in HIC,* $\Delta_l = 0$ *in LIC.*

Under quasi-linear preferences, the planner is indifferent across allocations that satisfy the binding lifetime incentive constraint with the same per-phase totals. Assumption 2.5 selects the natural representative of this set of optima, in which both per-phase incentive constraints separately bind. It therefore comes at no welfare loss. It is a normalization rather than a substantive restriction.

The next result is the structural identity that drives every comparative static below.

Proposition 2.6 (Structural identity)**.** *Under Assumption 2.5 and the iso-elastic specifications* (2.24)–(2.25)*, the per-capita switching cost is affine in pre-tax aggregate brown income in both HIC and LIC:*

$$\overline{\Delta v}(\lambda) \;=\; \frac{\eta}{1+\varepsilon}\, Y_B(\lambda) \;-\; K \;=\; \frac{\eta}{1+\varepsilon}\, B(\lambda) \;-\; K_0, \tag{2.43}$$

where

$$\eta \;\equiv\; 1-q^{-(1+\varepsilon)}, \qquad K \;\equiv\; \bar{R}\,(1-1/q), \qquad K_0 \;\equiv\; K-\frac{\eta\,\bar{R}}{1+\varepsilon} \;=\; \bar{R}\left[(1-1/q)-\frac{\eta}{1+\varepsilon}\right]. \tag{2.44}$$

The constant K_0 is non-negative, and strictly positive whenever $\bar{R}>0$ and $q>1$.

The proof, given in Section 5.2.3, exploits two facts. First, combining the consumption and income FOCs in the green phase yields $y_i^G = y_i^B/q^\varepsilon$ for both types in all regions, so green allocations are simply scaled-down brown allocations. Second, the iso-elastic relation $k_y \cdot y = (1+1/\varepsilon)\,k$ together with the binding per-phase ICs allows the consumption and labor-cost components of $\overline{\Delta v}$ to be combined into a single expression in Y_B. The non-negativity of K_0 follows from the convexity of $1-x^{-(1+\varepsilon)}$ in $x=q$ around $x=1$.

We also note one regularity property that we use implicitly below.

Lemma 2.3 (Regularity of $\overline{\Delta v}$)**.** *Under Assumption 2.5 and* (2.43)*, $\overline{\Delta v}(\lambda)>0$ on the admissible range whenever $B(\lambda)>(1+\varepsilon)K_0/\eta$. In particular, $\overline{\Delta v}>0$ throughout the admissible range whenever $\bar{R}=0$ (since $K_0=0$) and, more generally, whenever aggregate brown income is not dominated by the public-revenue requirement.*

Economically, $\overline{\Delta v}>0$ is the requirement that the green technology is per-period costly. Switching to green production strictly reduces the welfare-weighted flow utility of consumption net of labor cost. We treat this as a maintained regularity condition. It is satisfied automatically whenever $\bar{R}=0$ and, under the parameter restriction (A*) introduced below, throughout the admissible range whenever the inequality is strict.

Lemma 2.4. *Under* (2.25)*, $B(\lambda)$ is strictly increasing in λ on the admissible range. Its infimum on the admissible range is the limit value as $\lambda\to 0$ (deepest HIC), given by*

$$Y_B^{\min} \;=\; \pi_l\,\omega_l^{1+\varepsilon}\left(\frac{\pi_l}{1-\pi_h\,r^{1+1/\varepsilon}}\right)^{\varepsilon} + \pi_h\,\omega_h^{1+\varepsilon}, \qquad B^{\min}=Y_B^{\min}-\bar{R}, \tag{2.45}$$

where $r=\omega_l/\omega_h$. The limit is attained on the admissible range only if $\lambda=0$ lies in it; otherwise $Y_B^{\min}$ serves as a strict lower bound.

Proof. In HIC, $\kappa_h(\lambda)=\pi_h-\pi_l\lambda$ decreases in λ, so $A_B(\lambda)\equiv(1-\kappa_h)/(1-\kappa_h r^{1+1/\varepsilon})$ increases, $y_l^B=\omega_l^{1+\varepsilon}A_B^\varepsilon$ increases, and $y_h^B=\omega_h^{1+\varepsilon}$ is constant. In LIC, $\kappa_l(\lambda)=\pi_l\lambda-\pi_h$

increases, so $A_h(\lambda) \equiv (\lambda - \kappa_l)/(\lambda - \kappa_l r^{-(1+1/\varepsilon)})$ increases, $y_h^B = \omega_h^{1+\varepsilon} A_h^{\varepsilon}$ increases, and $y_l^B = \omega_l^{1+\varepsilon}$ is constant. The minimum in HIC is at $\lambda \to 0$, where $\kappa_h \to \pi_h$ and $A_B \to \pi_l/(1 - \pi_h r^{1+1/\varepsilon})$. □

The identity (2.43) immediately delivers clean log-decompositions for the two natural objects of interest along the frontier, the optimal switching time $T^*(\lambda)$ and the optimal terminal pollution stock $S^*(\lambda) = B(\lambda)\, T^*(\lambda)$. Define the public-good ratio

$$\zeta(\lambda) \;\equiv\; \frac{K_0}{\overline{\Delta v}(\lambda)} \;\geq\; 0. \tag{2.46}$$

Differentiating (2.43) logarithmically yields the within-phase identity

$$\frac{d\log \overline{\Delta v}}{d\lambda} - (1+\gamma)\frac{d\log B}{d\lambda} \;=\; \big(\zeta(\lambda) - \gamma\big)\,\frac{d\log B}{d\lambda}, \tag{2.47}$$

which collapses the joint behavior of $\overline{\Delta v}$ and B into the single scalar $\zeta(\lambda) - \gamma$ multiplied by the elasticity of B with respect to λ.

Taking logs in the compact FOC (2.42) and differentiating with respect to λ, then substituting (2.47), gives the log decomposition of the switching time,

$$\gamma\,\frac{d\log T^*}{d\lambda} \;=\; \underbrace{\frac{d\log \overline{\Delta v}}{d\lambda} - (1+\gamma)\,\frac{d\log B}{d\lambda}}_{\text{within-phase channel}} \;-\; \underbrace{\frac{d\log \hat{\theta}}{d\lambda}}_{\text{corrective channel}}. \tag{2.48}$$

Substituting $T^* = S^*/B$ into (2.42) and rearranging produces an implicit form for the terminal stock,

$$S^*(\lambda)^{\gamma} \;=\; \frac{\rho}{d\,\hat{\theta}(\lambda)}\left[\frac{\eta}{1+\varepsilon} - \frac{K_0}{B(\lambda)}\right]. \tag{2.49}$$

Taking logs and differentiating delivers the companion decomposition for S^*,

$$\gamma\,\frac{d\log S^*}{d\lambda} \;=\; \underbrace{\zeta(\lambda)\,\frac{d\log B}{d\lambda}}_{\text{within-phase channel}} \;-\; \underbrace{\frac{d\log \hat{\theta}}{d\lambda}}_{\text{corrective channel}}. \tag{2.50}$$

The within-phase channel enters (2.48) with coefficient $\zeta(\lambda) - \gamma$ and (2.50) with coefficient $\zeta(\lambda)$. The difference is exactly $\gamma\, d\log B/d\lambda$, which captures the fact that a larger B shortens the brown phase needed to accumulate any given stock. The optimal switching time is therefore more responsive to changes in B than the terminal stock is, and the two objects can move in opposite directions along the frontier.

Equations (2.48) and (2.50) are the two decompositions on which the remainder of the section draws. Each comparative-statics result reduces to signing the corrective channel $d\log \hat{\theta}/d\lambda$ and the within-phase channel through $\zeta(\lambda)$ and $d\log B/d\lambda$. Lemma 2.4 signs the latter unambiguously. The behavior of the corrective channel depends on whether

preferences are homogeneous, which is the question to which we turn next.

2.4.7 Homogeneous Preferences and the Invariance Benchmark

Homogeneous preferences mean $\theta_l = \theta_h = \theta$. The corrective motive that governs the timing of the green transition then carries no redistributive component. The next proposition states this formally.

Proposition 2.7. *If $\theta_l = \theta_h = \theta$, the effective damage valuation $\hat{\theta}(\lambda)$ in the switching condition* (2.42) *is constant along the Pareto frontier and equal to θ.*

Proof. Set $\theta_l = \theta_h = \theta$ in (2.39). Then $\hat{\theta}(\lambda) = (\theta + \lambda\theta)/(1 + \lambda) = \theta$ for every admissible λ. The right-hand side of (2.42) reduces to $\theta\, d\, B^{1+\gamma}(T^*)^{\gamma}/\rho$, with the only remaining λ-dependence operating through the within-phase allocations (B, T^*). □

Proposition 2.7 is the dynamic counterpart of Proposition 2.2. It shuts down the corrective channel in equations (2.48) and (2.50), but it does not by itself imply that T^* or S^* are invariant along the frontier. The within-phase channel still operates through the dependence of $B(\lambda)$ on the incentive-binding region. Two additional structural assumptions close this channel and deliver a dynamic counterpart to the static benchmark of Proposition 2.1.

Assumption 2.6. *Per-period consumption utility is linear in c_{it}.*

Assumption 2.7. *The public-revenue requirement satisfies $\bar{R} = 0$.*

Assumption 2.6 is the dynamic counterpart of the CES restriction in Proposition 2.1. In the static model, CES forces both types to share a common ratio g/b and collapses the consumption-side denominators $\alpha_g(g_l, b_l)$ and $\alpha_g(g_h, b_h)$ into a single function of τ. In the dynamic model, where green and brown enter consumption utility as perfect substitutes within each phase, linear consumption utility plays the same role. It eliminates any dependence of the switching FOC on the cross-sectional shape of the consumption allocation. Assumption 2.7 eliminates the fixed-cost term K_0 in the structural identity (2.43), so that $\overline{\Delta v}(\lambda) = \eta B(\lambda)/(1 + \varepsilon)$ is proportional to aggregate brown consumption. Both assumptions are restrictions on primitives, in parallel to the linear-damages and CES assumptions of the static benchmark.

The main result of the benchmark is an invariance result for the terminal pollution stock rather than for the switching time itself.

Proposition 2.8. *Suppose $\theta_l = \theta_h = \theta$ and that Assumptions 2.5, 2.6, and 2.7 hold. Then the optimal terminal pollution stock is invariant along the Pareto frontier,*

$$S^*(\lambda) \;\equiv\; B(\lambda)\, T^*(\lambda) \;=\; \left[\frac{\eta\,\rho}{(1+\varepsilon)\, d\, \theta}\right]^{1/\gamma}, \tag{2.51}$$

where $\eta \equiv 1 - q^{-(1+\varepsilon)}$. *The optimal switching time varies inversely with aggregate brown consumption,* $T^*(\lambda) = S^*/B(\lambda)$.

Proof. Under Assumptions 2.6 and 2.7, the structural identity (2.43) reduces to $\overline{\Delta v}(\lambda) = \eta B(\lambda)/(1+\varepsilon)$ since $K_0 = 0$. Under homogeneous preferences, $\hat{\theta}(\lambda) = \theta$ by Proposition 2.7. Substituting both into the compact switching FOC (2.42),

$$\frac{\eta\, B(\lambda)}{1+\varepsilon} \;=\; \frac{d}{\rho}\,\theta\, B(\lambda)^{1+\gamma}\, T^*(\lambda)^{\gamma}. \tag{2.52}$$

Cancelling one factor of $B(\lambda) > 0$ on each side and rearranging,

$$\big[B(\lambda)\, T^*(\lambda)\big]^{\gamma} \;=\; \frac{\eta\,\rho}{(1+\varepsilon)\, d\,\theta}, \tag{2.53}$$

which delivers (2.51) after taking the γ-th root. The right-hand side contains only primitives, so S^* is invariant in λ. The relation $T^*(\lambda) = S^*/B(\lambda)$ follows immediately from the definition of S^*. On the first-best region, the planner faces no binding incentive constraint, B takes its unconstrained value independent of λ, and hence T^* is constant on FB. □

The invariant object in the dynamic benchmark is the terminal pollution stock, not the switching time. Under the benchmark, the planner targets a fixed long-run level of accumulated pollution, and the timing of the green transition is whatever delivers that target given the rate at which the brown phase produces emissions. Off the first-best region, $B(\lambda)$ varies through the labor allocation, and $T^*(\lambda)$ varies inversely. The cross-sectional response is mechanical. A planner who distorts labor downward to redistribute also consumes less brown per period, so a longer brown phase is needed to reach the same target stock. This is the dynamic counterpart of the static result that under Proposition 2.1 the optimal carbon tax is pinned down by primitives alone. In the static model the invariant object is a price, the wedge between MRS and MRT, and the cross-section adjusts the consumption allocation to that wedge. In the dynamic model the invariant object is a stock, the terminal level of accumulated emissions, and the cross-section adjusts the switching date to that stock.

The same logic clarifies a subtle point about the homogeneous case absent the full benchmark. Even when only Proposition 2.7 applies, with $\bar{R} > 0$ or non-linear consumption utility, the headline invariance does not extend to either object across the full frontier. Setting $d\log\hat{\theta}/d\lambda = 0$ in (2.50), the within-phase channel survives,

$$\gamma\,\frac{d\log S^*}{d\lambda} \;=\; \zeta(\lambda)\,\frac{d\log B}{d\lambda} \;\geq\; 0,$$

with equality whenever $\bar{R} = 0$. Homogeneous preferences are therefore sufficient for the

corrective motive to be invariant, but not for the terminal stock or the switching time to be invariant along the full frontier. The first-best region is the cleanest case in either direction. There B is constant, both channels vanish, and S^* and T^* are both constant.

The benchmark is knife-edge in three respects, in parallel to the static case. Relaxing homogeneous preferences reactivates the corrective channel through $\hat{\theta}(\lambda)$, which is the subject of the next subsection. Relaxing $\bar{R} = 0$ introduces the fixed-cost term $K_0 > 0$ in $\overline{\Delta v}$, so that the switching FOC no longer factors cleanly into a stock-target condition. Relaxing linear consumption utility reintroduces a cross-sectional channel through the consumption-side denominators, analogous to the one shut down by CES in the static model.

2.4.8 Heterogeneous Preferences

When the two types differ in their valuation of environmental damages, $\theta_l \neq \theta_h$, the corrective motive itself becomes redistributive, and $\hat{\theta}(\lambda)$ moves with λ.

Proposition 2.9. *Suppose $\theta_l \neq \theta_h$. Then the effective damage valuation $\hat{\theta}(\lambda) = (\theta_l + \lambda\,\theta_h)/(1+\lambda)$ governing the switching condition (2.42) is strictly monotone in λ, with*

$$\frac{d\hat{\theta}}{d\lambda} \;=\; \frac{\theta_h - \theta_l}{(1+\lambda)^2} \;\neq\; 0.$$

Hence the corrective motive for the timing of the green transition acquires a redistributive component. The present-value valuation of avoided future damages varies along the Pareto frontier.

Proof. Immediate from (2.39). Whenever $\theta_h \neq \theta_l$, $d\hat{\theta}/d\lambda$ is nonzero for every admissible λ, so $\hat{\theta}$ is not constant along the Pareto frontier. The right-hand side of (2.42) therefore depends on λ through $\hat{\theta}$, in addition to any dependence through the within-phase allocations (B, T^*). □

As in the static model, heterogeneity in environmental preferences is sufficient but not necessary for the optimal switching time to depend on welfare weights. The corrective channel $d\log\hat{\theta}/d\lambda$ vanishes when $\theta_l = \theta_h$. The within-phase channel in (2.48) does not, since aggregate brown consumption $B(\lambda)$ varies along the frontier through the equilibrium allocation.

The next step is to sign the within-phase channel. By (2.47), the sign of the within-phase channel in (2.48) is exactly the sign of $\zeta(\lambda) - \gamma$ multiplied by the (strictly positive) elasticity of B. We can guarantee that the within-phase channel does not work *against* the corrective channel by imposing a primitive-level bound on the public-revenue requirement.

Assumption 2.8. *The level of non-environmental public spending satisfies*

$$\bar{R} \;\leq\; \frac{\gamma\,\eta\,Y_B^{\min}}{(1+\varepsilon)(1-1/q)(1+\gamma)-\eta}, \tag{A*}$$

with $Y_B^{\min}$ *as defined in* (2.45) *and* $\eta = 1 - q^{-(1+\varepsilon)}$.

The right-hand side of (A*) is strictly positive for $\gamma > 0$ and $q > 1$, so the condition is satisfied trivially when $\bar{R} = 0$ and, more generally, whenever $\bar{R}$ is not too large relative to aggregate productive capacity $Y_B^{\min}$. The condition is satisfied for any plausible parameterization in which public spending is small relative to aggregate income.

Proposition 2.10. *Suppose* $\theta_h > \theta_l$*, and that Assumptions 2.5 and 2.8 hold. Then the optimal switching time satisfies*

$$\frac{dT^*}{d\lambda} \;<\; 0 \tag{2.54}$$

on the admissible range of the Pareto frontier. In particular, $T^{*,\,HIC} > T^{*,\,FB} > T^{*,\,LIC}$.

Proof. By Proposition 2.6 and (2.47), the within-phase channel in (2.48) equals $(\zeta(\lambda) - \gamma)\, d\log B/d\lambda$.

Sign of $\zeta(\lambda) - \gamma$. Since $\zeta(\lambda) = K_0/\overline{\Delta v}(\lambda)$ and $\overline{\Delta v}$ is increasing in B by (2.43), ζ is decreasing in B. By Lemma 2.4, B attains its minimum $B^{\min} = Y_B^{\min} - \bar{R}$ at $\lambda \to 0$, so ζ attains its maximum there:

$$\zeta_{\max} \;=\; \frac{K_0}{\eta B^{\min}/(1+\varepsilon) - K_0}.$$

The condition $\zeta_{\max} \leq \gamma$ rearranges to

$$K_0\,(1+\gamma) \;\leq\; \frac{\gamma\,\eta\,B^{\min}}{1+\varepsilon},$$

and substituting $K_0 = \bar{R}[(1-1/q) - \eta/(1+\varepsilon)]$ and $B^{\min} = Y_B^{\min} - \bar{R}$ produces (A*). Hence under Assumption 2.8, $\zeta(\lambda) \leq \gamma$ for all admissible λ, so the within-phase channel $(\zeta(\lambda) - \gamma)\, d\log B/d\lambda \leq 0$, with equality only at $\bar{R} = 0$ and $\lambda \to 0$.

Sign of $d\log B/d\lambda$. Strictly positive on the admissible range by Lemma 2.4.

Sign of $d\log\hat{\theta}/d\lambda$. From (2.39), $d\log\hat{\theta}/d\lambda = (\theta_h - \theta_l)/[(1+\lambda)(\theta_l + \lambda\theta_h)] > 0$ whenever $\theta_h > \theta_l$.

Combining these three facts in (2.48): the first term on the right-hand side is non-positive, the second is strictly negative, so $d\log T^*/d\lambda < 0$ on the admissible range. Applying the partition (2.37) yields $T^{*,HIC} > T^{*,FB} > T^{*,LIC}$. □

Under (A*), the within-phase channel has the *same sign* as the corrective channel. Both push toward earlier switching as λ rises. This is the dynamic counterpart to the

static result that the level of the carbon tax varies along the Pareto frontier under heterogeneous environmental preferences. Welfare weights determine not only how much to tax carbon, but when to abandon it.

The terminal pollution stock

The switching time is one of two natural objects to track along the frontier. The other is the terminal pollution stock $S^*(\lambda) = B(\lambda)\,T^*(\lambda)$, the level at which accumulated emissions are stabilized once the economy switches to the green technology. The two objects answer different questions. The switching time describes when the economy abandons brown production. The terminal stock describes the long-run level of accumulated damages that the planner is willing to accept. A planner who delays the green transition does not necessarily accept a higher long-run damage burden, because a longer brown phase at a lower per-period rate of brown consumption can deliver the same terminal stock as a shorter brown phase at a higher rate. Whether the two move together or apart along the Pareto frontier is therefore a separate question.

The implicit form (2.49) and its log decomposition (2.50) provide the answer. The terminal stock is determined by two objects, the effective damage valuation $\hat{\theta}(\lambda)$ and aggregate brown consumption $B(\lambda)$, with the latter entering only when $K_0 > 0$. The decomposition (2.50) shows that the corrective and within-phase channels enter S^* with opposite signs whenever $\theta_h > \theta_l$ and $K_0 > 0$.

Proposition 2.11. *Suppose $\theta_h > \theta_l$ and Assumption 2.5 holds. Then the optimal terminal stock satisfies $dS^*/d\lambda < 0$ on the admissible range whenever $\bar{R} = 0$. More generally, $dS^*/d\lambda < 0$ whenever*

$$\zeta(\lambda)\,\frac{d\log B}{d\lambda} \;<\; \frac{d\log\hat{\theta}}{d\lambda}. \tag{2.55}$$

In particular, $S^{*,HIC} > S^{*,FB} > S^{*,LIC}$ whenever the inequality holds throughout the admissible range.

Proof. The right-hand side of (2.49) is strictly positive under Lemma 2.3, so $S^*(\lambda)$ is well-defined. When $\bar{R} = 0$, $K_0 = 0$ and $\zeta(\lambda) \equiv 0$, so the within-phase channel in (2.50) vanishes and $\gamma\, d\log S^*/d\lambda = -d\log\hat{\theta}/d\lambda < 0$ by (2.39) and $\theta_h > \theta_l$. When $\bar{R} > 0$, the within-phase channel is non-negative because $\zeta \geq 0$ and $d\log B/d\lambda > 0$ by Lemma 2.4, and the corrective channel is strictly positive. The sign of $d\log S^*/d\lambda$ is therefore determined by which channel dominates, with (2.55) the exact condition under which the corrective channel wins. □

Two channels work against one another on S^* whenever $K_0 > 0$, in contrast to the case of T^* where they reinforce under (A*). The corrective channel pulls the target stock down as λ rises. The within-phase channel pushes it up, because a higher λ raises B and

a higher B raises the per-period flow utility cost of switching, which the planner offsets by accepting a larger terminal stock. The two effects can be of either relative size when $K_0 > 0$.

The economic content of Proposition 2.11 is that heterogeneous environmental preferences shift not only the timing of the green transition but also the long-run level of pollution that the planner targets. A Rawlsian planner who favors the environmentally less-concerned type tolerates both a later transition and a higher stabilized stock of accumulated emissions. A planner who favors the environmentally more-concerned type transitions earlier and stabilizes at a lower stock. When public-revenue requirements are absent, the second result is unconditional. When public-revenue requirements are present, the two channels work against one another on the stock, and the comparative static depends on whether the within-phase response of B to redistribution is small enough not to overturn the corrective effect.

A useful limiting case is the homogeneous-preference benchmark of Proposition 2.7. When $\theta_l = \theta_h$, the corrective channel in (2.50) vanishes and $dS^*/d\lambda = (\zeta/\gamma)\, d\log B/d\lambda \geq 0$, with equality whenever $\bar{R} = 0$. The terminal stock is then constant along the frontier whenever public-revenue requirements are absent, and weakly increasing otherwise. The first-best region is the cleanest case. There B takes its unconstrained value independent of λ, so both channels in (2.50) vanish and S^* is constant on FB regardless of $\bar{R}$. The variation of T^* within FB inherited from changes in $\hat{\theta}$ at the FB allocation translates one-for-one into compensating changes in the brown-phase duration, leaving the terminal stock fixed at its FB level.

Figure 2.3 illustrates Proposition 2.10 numerically. The optimal switching time is decreasing in the reservation utility $V(h)$. In the HIC region, where the planner favours the low type, T^* is high. The economy delays the green transition because the effective damage weight $\hat{\theta}$ is low. In the LIC region, where the planner favours the high type, T^* is low. The economy transitions earlier because $\hat{\theta}$ is high. The first-best region connects the two with a flat segment, confirming that T^* is constant when no incentive constraint binds, as implied by Proposition 2.7 under homogeneous preferences and by the invariance of the corrective motive at the first-best more generally.

2.5 Conclusion

In this paper, we limit ourselves to a simple setup to illustrate how welfare judgments are inherent in determining optimal climate policy. We deliberately choose a two-type world in order to characterize a clear theoretical contribution. Within the Atkinson–Stiglitz Assumptions, we show that a unique optimal carbon tax can be pinned down only under a knife-edge combination of structural conditions: homogeneous environmental preferences, linear damages, and a CES consumption aggregator must all hold simultaneously.

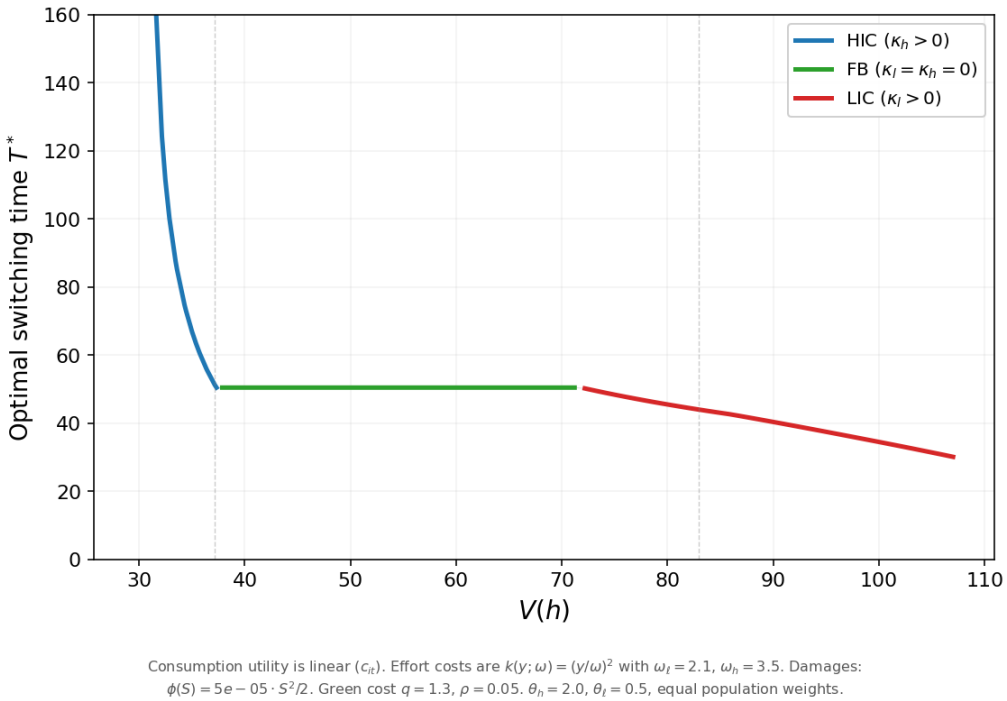


Figure 2.3: **Optimal switching time T^***

The figure depicts the optimal switching time T^* as a function of the reservation utility of the high type $V(h)$. The three regions correspond to binding high-type IC (blue, left), first-best (green, centre), and binding low-type IC (red, right). Consumption utility is linear (c_{it}). Effort costs are $k(y;\omega) = (y/\omega)^2$ with $\omega_\ell = 2.1$, $\omega_h = 3.5$. Damages are $\phi(S) = 5 \times 10^{-5}\, S^2/2$. Green cost $q = 1.3$, $\rho = 0.05$, $\theta_h = 2$, $\theta_l = 0.5$, equal population weights. See Section 5.2.4 for details.

Any small perturbation along any of these dimensions — convexity in damages, even modest heterogeneity in environmental preferences, or departures from CES — breaks the invariance and introduces a distributional component to the optimal tax. In such a world, finding the optimal carbon tax is no longer a question of damage assessment or discount factors alone, but requires taking a stance on distribution.

We extend this insight to a dynamic setting, showing that distributional preferences govern not only the level of the carbon tax but also the timing of the green transition and the long-run pollution stock at which emissions stabilize. Both objects are, jointly, distributional choices.

Chapter 3

Morals and the Political Economy of Corrective Taxes

Joint with Felix Bierbrauer, Mattias Polborn, and Georg Weizsäcker

Abstract

We study the political economy of corrective taxes in a model with *neoclassical consumers* who take prices and country-wide emissions as given and *socially responsible consumers* who internalize the social cost of carbon even without corrective taxes. They are modeled as rule-utilitarian ethical agents who take the behavior of neoclassical consumers as given. We characterize political equilibrium taxes and analyze how such taxes change when moral arguments become central to the debate. Finally, we report on a survey that investigates how preferences over carbon taxes correlate with a measure of how socially responsible individuals are.

Acknowledgements. The authors gratefully acknowledge support by the Deutsche Forschungsgemeinschaft (German Research Foundation) via Germany's Excellence Strategy – EXC2126/1 – 390838866 – and CRC TRR 190 (Project number 280092119). We benefitted from the comments of seminar audiences at WZB Berlin, TU Dortmund, University of Vienna and the Central European University and from conversations with Marc Kaufmann and Sebastian Siegloch.

Felix Bierbrauer: Department of Economics, University of Cologne, Germany
Mattias Polborn: Vanderbilt University, Nashville, USA
Marten Ritterrath: University of Cologne, Germany
Georg Weizsäcker: Humboldt-Universität zu Berlin, Germany

3.1 Introduction

An important theme in welfare economics are corrective taxes that counter market failures due to externalities. Applied to climate policy, cost-benefit analyses yield an estimate of the social cost of carbon, on which corrective taxes should be based. In this paper, by contrast, we take a political economy perspective on corrective taxes.

A traditional approach in the political economy of taxation is to analyze how the preferences of self-interested voters shape political equilibrium outcomes. The contribution of this paper is to enrich this approach with an additional dimension of heterogeneity, the extent to which individuals, as consumers or as citizens, are motivated by morals. Specifically, we consider three dimensions of heterogeneity: Individuals differ their demand of dirty goods and hence in individual carbon footprints. They also differ in their valuation of climate damages, or, equivalently, in their willingness to pay for a public good, *"the reduction of country-wide carbon emissions"* Finally, they are divided on whether or not individuals have a moral obligation to decrease their carbon footprint beyond material self-interest.[1]

Our interest on the interplay of morals with more traditional political economy forces stems from the acrimony of climate policy debates, with slogans such as *"There is no Planet B"* or *"Drill, Baby, Drill!"*. Proponents of ambitious climate policy often appeal to consequences for future generations or regions that will suffer from sea-level rises, floods and draughts. By contrast, a widespread critique of climate policy at the national level appeals to self-interest.[2]

This paper has a theory part and an empirical part. We first present a theoretical analysis of an economy that has two groups of consumers: Neoclassical consumers (henceforth, *N-agents*) take prices and overall carbon emissions as given and maximize individual consumer surplus. Socially responsible consumers (henceforth, *S-agents*), by contrast, internalize, at least in part, the damages associated with their carbon footprints. We characterize political equilibrium taxes in such an economy, and also how moral activism affects political equilibrium taxes. In the empirical part of the paper, we report on a survey on how carbon tax preferences relate to consumption habits, views on the severity of climate damages and attitudes towards moral activism.

The theory part involves various steps that are, to the best of our knowledge, novel and of stand-alone interest. We propose a microfoundation for preferences over corrective taxes, both for neoclassical and for socially-responsible consumers. We derive comparative statics predictions on how these preferences and also equilibrium taxes change when, say, the population share of socially responsible agents increases; In particular, we show that more S-agents may lead to *lower* taxes and *higher* emissions. We show, moreover, that moral activism may give rise to equilibrium multiplicity; there may exist both an equilibrium with moral pressure by

[1] Violating such an obligation, may, for instance, cause "flight shame" a term that has been coined to characterize air travelers who feel bad about of the climate-damaging consequences of their journey, see Doran et al. 2022 or Laurent et al. 2026.

[2] According to this critique, national climate policy comes with higher energy costs that cause a loss of production and employment, with little effects on global warming as the sellers of oil and gas will find new customers while firms relocate to pollution havens.

socially responsible consumers and high taxes and an alternative low-tax-equilibrium in which this moral pressure is countered by rebellious neoclassical consumers.[3]

Finally, we are interested in a welfare comparison of equilibria with and without moral activism. Who wins and who loses when contentious moral obligations play a more central role in the discourse about climate policy? We show that an equilibrium with moral pressure and moral rebellion can be Pareto-dominated by an equilibrium with no moral activism at all.

Many of our formal propositions rest on the premise that N-agents tend to have larger carbon footprints and perceive smaller social costs of carbon than S-agents. The survey shows that this is indeed the case. Individuals who are more similar to our model's S-agents tend to have a greener life-style, view climate change as a more serious problem, want carbon taxes to be larger and are more supportive of moral pressure on consumers who disregard social costs. In the remainder of the introduction, we give a more detailed overview of how our analysis proceeds and then discuss related literature.

Preferences over carbon taxes. We present a theory of how N- and S-agents form preferences over corrective taxes. For N-agents, we follow an approach that has been used in the literature on the political economy of redistributive taxation and adapt it to a setting with corrective taxes.[4] Specifically, N-agents take damages due to carbon emissions into account and value the mitigative effect of a corrective "Pigouvian" tax.[5] In addition, there is a distributive effect as the tax revenue is rebated. When forming preferences over taxes, an N-agent balances the mitigation and the redistribution effects with her loss of consumer surplus when taxes go up. The interplay of the Pigouvian and the distributive component typically implies that N-agents want to deviate from a pure Pigouvian benchmark (*"tax rate equal to marginal damage"*). For instance, an N-agent with a below-average carbon footprint wants taxes that are higher than this benchmark.

We model S-agents as ethical agents. In an environment composed entirely of such agents, they adhere to a Kantian prescription: *Behave such that an optimal social outcome would result if everybody behaved this way.* However, in an environment that also includes neoclassical consumers, the prescription is modified: *Take as given that N-agents maximize consumer surplus. Behave such that an optimal social outcome results for those who follow the rule.* This modified rule nests both individual payoff maximization and the adherence to the Kantian prescription as special cases. The shifter between these extreme cases is the population share of socially responsible consumers. If there is only one such consumer, her behavior is the same as the

[3]The U.S. moved in and out of the Paris agreement with the recent changes of its presidency. The *green new deal* has been the key initiative when a new European Commission took office in December 2019, after a year that was marked by the protests of the Fridays for Future movement. As of 2026, the climate policy debate in the EU is about moderation. Through the lens of our political economy model, these ups and downs in climate policy can be interpreted as shifts between a low-tax and a high-tax-equilibrium.

[4]Seminal references are Roberts 1977 and Meltzer and Richard 1981. Bierbrauer et al. 2022 and Berliant and Boyer 2024 discuss more recent contributions to the political economy of taxation.

[5]We assume that agents differ in their willingness to pay taxes for damage reduction. Our theory allows for the willingness to pay to be zero, which is plausible for a small country with no significant impact on global emissions. The survey provides an empirical perspective for the US and the UK.

behavior of the N-agents surrounding her. By contrast, if everybody else also is an S-agent, then the modified rule stipulates Kantian behavior.[6] The modified rule involves an antagonism: There is an in-group of socially responsible comrades whose welfare is taken into account and an out-group of selfish payoff maximizers whose welfare is ignored.[7]

This group-rule-utilitarian calculus internalizes some of the social costs of carbon: Differently from an individual, the whole group of S-agents has a noticeable impact on country-wide emissions. Therefore, they show a behavioral response not only to changes of (tax-inclusive) consumer prices, but also reduce their consumption when the social cost of carbon goes up. This distinguishes them from neoclassical consumers.

The payoff that S-agents reach as a group depends on the corrective tax. We infer their tax preferences from the value function of their group-rule utilitarian optimization problem. Its derivative gives the marginal benefit of higher taxes. Again, there is a distributive and a Pigouvian component, but they differ from those that shape N-agents tax preferences. The Pigouvian motive is entirely due to the fact that the latter show a behavioral response to taxation. The S-agents' behavior is optimal and hence needs no correction. Whether the distributive component calls for under- or overtaxation relative to the pure Pigouvian benchmark depends on a group comparison: When the average S-agent has a smaller carbon footprint than the average N-agent (as the latter flies more often, drives a bigger car...), then S-agents want high taxes to take advantage of this materialism, beyond the pure Pigouvian correction.

We look at stylized special cases of a more general setup to obtain comparative statics predictions for political equilibrium taxes. One such special case is as follows: All N-agents have the same consumption preferences and damage assessments. The same is true for all S-agents, so that there is a representative agent for each group,[8] and we assume that the S-agent has a weakly lower carbon footprint, and a weakly higher damage estimate than the N-agent, with one of these inequalities strict. We finally assume that N-agents outnumber S-agents. When the population share of S-agents increases (while staying a minority), the politically decisive N-agent prefers lower taxes than before for distributive reasons. Thus, while the new S-agents consume less than the N-agents they replace, the tax decrease leads to a higher consumption, especially among continuing N-agents.

Moral pressure and moral rebellion. We enrich the above framework by allowing S-agents to exert moral pressure on N-agents to reduce their carbon footprints. The latter then

[6]There is related literature on moral behavior that analyzes how Kantian prescriptions are traded off against other objectives, including material incentives. Brekke et al. 2003 take a similar approach to ours in their translation from outcome utilities to a Kantian prescription, embedding it into policy regimes (but not political equilibria like ours). Ellingsen and Mohlin 2026 and Alger 2026 examine circumstances under which Kantian motives can dominate other motives.

[7]In part 6.1.8 of the Appendix, we explore universalism as an alternative to group-rule-utilitarianism. Specifically, we characterize the socially responsible consumers' demand for dirty goods and their preferences over taxes on the assumption that their choices are governed by a universalist calculus that incorporates also the welfare implications for neoclassical consumers. Our choice to give more emphasis to group-rule-utilitarianism is in line with theories such as Alger 2026 that consider social comparisons like *aversion to behindness* as separate motives that are distinct from universalism.

[8]Our general setup has a joint distribution of preferences for the dirty good and views of the social cost of carbon among N-agents and a separate distribution among S-agents.

face a choice of whether to comply with this moral pressure or to reject it, a case to which we refer as moral rebellion. We present an analysis of how moral pressure and moral rebellion affect carbon tax preferences and ultimately the political equilibrium outcome.

Specifically, we assume that exerting moral pressure has a cost, e.g. to assume an unpleasant role in private interactions and to remind others of their moral obligations. There also is a benefit: neoclassical consumers don't want to be frowned at, so they reduce their consumption of the dirty good when they fear critical reactions. We assume that S-agents weigh the costs and benefits to determine the intensity of their moral pressure using, again, a group-rule utilitarian calculation.

We show that moral pressure is a substitute for taxes: The lower the tax rate and the higher the consumption of N-agents, the larger is the optimal intensity of moral pressure. N-agents adjust their tax preferences accordingly. They understand that there is now an additional benefit from higher taxes, as they crowd out moral pressure. When the decisive voter is an N-agent, political equilibrium taxes increase and carbon emissions decrease, relative to a baseline without moral pressure. Thus, moral pressure pays off for S-agents, not only directly, but also politically.

We then allow N-agents to reject the moral pressure, modeled as an act of moral rebellion that requires a costly mental effort but generates the benefit of unhindered consumer surplus maximization. We assume a strategic complementarity: rejecting moral pressure is less costly when more people are doing it. As in other models, complementarities may give rise to equilibrium multiplicity and we analyze conditions under which an equilibrium in which N-agents simply accept moral pressure exists alongside an equilibrium with widespread moral rebellion. In the latter equilibrium, moral rebellion is more attractive for high-intensity consumers. Hence, the equilibrium comes with a partition of N-agents into a set of heavy consumers who are rebellious and a set of light consumers who can still be reached by moral pressure.

We also find that the set of light consumers who are receptive to moral pressure expands with higher taxes. Therefore, S-agents now perceive moral pressure and taxes as complementary. They intensify moral pressure when taxes are high, which creates an additional cost of higher taxes: taxes crowd in moral pressure. If the decisive voter is an N-agent, the political equilibrium taxes may therefore be lower and carbon emissions higher, relative to the baseline with no moral pressure and no moral rebellion. Thus, moral rebellion pays off for N-agents, not only directly, but also politically.

We construct a numerical example in which, in the equilibrium with moral pressure, S-agents are better off while most N-agents (except those with small carbon footprints) are worse off than in the baseline equilibrium. Then, the equilibrium that also allows for moral rebellion makes S-agents worse off while most N-agents are better off. However, in comparison to the initial equilibrium without any moral activism, everyone is worse off. Thus, it can be bad for all if politics is turned into a moral battle.

Survey. Finally, we present a survey in which we measure climate concern, the importance of carbon-intensive consumption, socially responsible behavior, and carbon tax preferences. In

particular, the survey sheds light on how close people are to the stylized categories of N- and S-agents of our theory.[9] These two stylized types are distinct in their consumption behavior, notably their behavioral responses to changes in consumer prices and emission damages. N-agents respond to the former but not to the latter, whereas S-agents respond to both. Moreover, some of our formal propositions focus on an environment in which S-agents want higher taxes than N-agents, due to either their smaller average carbon footprints or their larger damage estimates. Our survey allows us to see whether this is an empirically relevant environment, by examining how the respondents' stated consumption patterns correlate with tax preferences, damage assessments, and attitudes towards moral pressure and moral rebellion.

We run an online survey on Prolific with representative samples of US and UK respondents, stratified by political affiliation. Our main measures of socially responsible behavior are the *Damage Response*, the stated adjustment of car usage to an increase in the social cost of carbon at fixed prices, and a domain-general measure of the self-perceived trade-off between consumption value and social costs, *Price vs. Harm*. We also elicit proxies for the importance of carbon intensive consumption and climate concern, preferences over carbon taxes, and attitudes towards moral pressure.

We document three main findings, each consistent with our theory. First, the behavioral distinction shows up in the data: respondents with a stronger Damage Response are more concerned about climate change, place less weight on carbon-intensive consumption, and weigh harm more heavily relative to price. Second, more socially responsible behavior (as measured by a larger Damage Response) predicts higher preferred carbon taxes in both countries, and does so conditional on consumption importance and climate concern. Third, more socially responsible respondents are more likely to approve of moral pressure and less likely to view it as intrusive, while respondents for whom carbon-intensive consumption is important show the opposite pattern. Approval of moral pressure, in turn, correlates with higher preferred taxes.

Related Literature. A rich literature on the political economy of redistributive taxes focuses on how changes in income inequality transmit via the political process into changes in redistributive policies. A well-known result is the Meltzer-Richard (1981) hypothesis that an increase in inequality, measured by the gap between average and median income, leads to a larger and more redistributive government.[10] A contribution of this paper is to derive a version of the Meltzer-Richard hypothesis for corrective taxes: The gap between the median and the average carbon footprint explains whether there is over- or undertaxation - relative to a pure Pigouvian tax.

We also move beyond this extension in our analysis of how moral pressure, moral rebellion and carbon taxes jointly emerge as endogenous political equilibrium outcomes. The work of Bénabou and Tirole 2006 and Alesina and Angeletos 2005 and this paper have in common that

[9]There is an emerging literature that uses survey methods to explore views on climate change and on policies addressing it, see in particular Dechezleprêtre et al. 2025b and the references therein. To this literature we add a survey that is guided by our political economy model.

[10]The Meltzer-Richard hypothesis often fails empirically, and this has motivated alternative positive theories of the welfare state, e.g. by Bénabou and Tirole 2006 or Alesina and Angeletos 2005.

"opinions" – views on the origins of economic success in their case, views on moral obligations to reduce carbon footprints in ours – matter for equilibrium taxes. Our work is related to Bénabou and Tirole 2006 in that there is an explicit attempt of one group to influence the views of another group.[11] We move beyond their approach in that we give those who are being influenced a choice to accept or to reject this influence.

We model socially responsible consumers as group-rule utilitarians.[12]. Group-rule utilitarianism has previously been used in political economy models with endogenous turnout, as a way around the paradox of voting, see Coate and Conlin 2004 and Feddersen and Sandroni 2006. To the best of our knowledge, there is no previous work that uses group-rule utilitarianism to formalize moral behavior in markets. Instead, the literature has explored Kantian morality and other-regarding preferences.

The Kantian approach uses a counterfactual as a normative benchmark: Morality requires to follow a rule that would yield the socially optimal outcome if all individuals complied. Models of Kantian morality assume that it is costly for individuals to violate this norm, see Alger and Weibull 2013 and Roemer 2015 for foundational work and Leeuwen and Alger 2024 or Capraro and Rand 2018 for lab experiments with a focus on Kantian morality. Implications of Kantian morality for climate policy are analyzed in Eichner and Pethig 2021.

There is a rich literature on other-regarding preferences such as altruism or inequity aversion, and whether such preferences show in laboratory experiments; see Fehr and Schmidt 1999, Bolton and Ockenfels 2000 or Charness and Rabin 2002 for important early references, or Fehr and Charness 2025 for a recent review. In many experiments, players affect other individuals' payoffs, giving other-regarding preferences a chance to play out. By contrast, individuals with other-regarding preferences are indistinguishable from selfish ones in Dufwenberg et al. 2011 where players interact via the price mechanism in large markets. The same is true in our analysis. Individuals who do not belong to the group of socially responsible consumers may still have other-regarding preferences that may shape their tax preferences, but these other-regarding preferences do not show in the behavior as a consumer.[13] An alternative approach for the analysis of large markets has recently been developed by Kaufmann et al. 2024 where an individual's impact on the equilibrium quantity is non-negligible even when there are many

[11] In Bénabou and Tirole 2006, these are parents who seek to influence their children's beliefs about the returns to educational effort. In our case, socially responsible consumers seek to influence the views of selfish payoff maximizers on the acceptability of large carbon footprints. In doing so, socially responsible consumers distinguish an in-group of like-minded individuals and an out-group of individuals whose behavior warrants a correction. Thereby, our analysis also speaks to the large literature on political polarization (Fiorina and Abrams 2008; Abramowitz and Webster 2016; Krasa and Polborn 2014; Callander and Carbajal 2022; Gennaioli and Tabellini 2025).

[12] A related specification is *Homo Moralis* in Alger and Weibull 2013 who is also an intermediate version of a Kantian and a neoclassical agent. The difference to group-rule utilitarianism is that the ethical component of the Homo Moralis preferences is *universalist*.

[13] Models of social preferences have the property that individuals are selfish in the absence of externalities, see Bierbrauer and Netzer 2016: When individuals have a choice between A and B and when A and B have the same consequences for everybody else, they choose the alternative that yields the larger payoff for themselves. When the participant in a market system with many participants takes aggregate outcomes, i.e. prices and quantities, as given, then the individual's choices maximize consumer surplus, irrespective of whether or not the individual has other-regarding preferences.

market participants. Therefore, individuals who care about others internalize some of the social damages from carbon emissions even in the absence of corrective taxes.

We view group-rule utilitarianism as an approach that is complementary to the literatures on Kantian morality and on other-regarding preferences. Our ethical agents care for the welfare of other ethical agents and, like in the Kantian literature, their behavior is in line with the solution of a collective action problem. However, our ethical agents consider only the welfare of their own group of ethical agents. Their normative standard thus changes with the composition of the economy. For instance, their response to changes in the social cost of carbon depends on the population share of neoclassical consumers. This malleability of moral standards is absent from the Kantian treatments, but broadly consistent with the evidence on social preferences in lab experiments. In particular, group utilitarianism has a flavor of conditional cooperation (Fischbacher and Gächter 2010; Thöni and Volk 2018) in that it formulates a standard for moral behavior by referencing an in-group of moral people, while selfish payoff maximizers form the out-group. Their behavior warrants a costly corrective activity that takes the form of moral pressure. In lab experiments, we often observe a willingness to sacrifice own payoff for the purpose of punishing defectors (Fehr and Gächter 2000).

Formally, we treat moral pressure as a behavioral intervention or "nudge" that makes neoclassical consumers behave as if they were facing a higher price, without actually paying it, see Farhi and Gabaix 2020 for an analysis of nudges and behavioral biases in the context of optimal tax models. List et al. 2023 discuss "nudges as moral taxes" and "guilt per unit of consumption induced by a (...) nudge."[14]

Outlook. The next section introduces our baseline model of an economy with neoclassical and socially responsible consumers. In this baseline, individuals consume and vote. Their respective preferences over taxes are characterized in Section 3, while political equilibrium taxes are analyzed in Section 4. Sections 5 and 6 extend the basic setup. Exerting moral pressure becomes possible in Section 5, rejecting it in Section 6. We report on a survey in Section 7. The last section contains concluding remarks. Formal proofs are relegated to the Appendix.

3.2 The Model

There is an atmospheric externality; CO2 emissions from the production of a dirty good contribute to global warming. We consider a country that is inhabited by a continuum of individuals who differ in their demand for the dirty good and, possibly, also in their views on the desirability to reduce the country's aggregate emissions. Moreover, two subgroups of the population differ in how they behave as consumers of the dirty good. *Neoclassical consumers* (henceforth, N-types) take both prices and aggregate consumption as given when acting as consumers. They equalize their marginal utility from the dirty good with its tax-inclusive price. As we explain in the next

[14]To be clear, this uses the term "nudge" for every non-price intervention. "Moral pressure" is not an example of a nudge in the sense of Thaler and Sunstein 2009, a mere change of the choice architecture that helps biased agents to make better choices without having a negative effect on others.

section, in their role as voters, N-types also have preferences over carbon taxes. However, their political preferences are not revealed by their behavior as consumers. The second group consists of *socially responsible consumers* (henceforth, S-types). They differ from N-types because they, motivated by ethical concerns, internalize environmental damages. Below, we formalize their consumption choices as resulting from a rule-utilitarian calculus. Subsequently, we turn to their political preferences.

We let θ parameterize the marginal utility of consuming the dirty good. More formally, let x be a generic consumption level of the dirty good and denote consumption utility by $u(x, \theta)$. We denote the marginal utility of consumption by $u_x(\cdot)$ and assume that it is increasing in θ, i.e. $u_{x\theta}(\cdot) \geq 0$. For tractability, we sometimes impose the assumption that the function u is iso-elastic; i.e. $u(x, \theta) = \theta^{\frac{1}{\varepsilon}} \frac{1}{1-\frac{1}{\varepsilon}} x^{1-\frac{1}{\varepsilon}}$, for $\varepsilon > 1$. An individual's attitude towards the country's emissions is given by a damage function $\bar{x} \mapsto D(\bar{x}, \delta)$, where $\bar{x}$ denotes aggregate or per-capita consumption of the dirty good and δ is a parameter. We let $D_{\bar{x}\delta}(\cdot) > 0$ so that a higher value of δ implies larger perceived marginal damages.[15] We also let $\delta = 0$ imply that $D(\bar{x}, \delta) = 0$ for all $\bar{x}$. For tractability, we sometimes assume that the damage function D is linear; i.e. $D(\bar{x}, \delta) = \delta\, \bar{x}$. Below, when we have fully specified preferences over corrective taxes, we expand on the interpretation of δ.

There is a joint distribution of θ and δ among N-types, and one among S-types. Our model allows both for the possibility that these distributions differ or are the same. We will refer to stylized special cases to illustrate the theory's implications.

Corrective taxes increase the consumer price of the dirty good, given by $q(t) = k + t$, where k is the marginal cost of production and t is the tax on the dirty good. Tax revenue is rebated lump sum.

3.2.1 Neoclassical consumers

N-agents benefit from the tax revenue that is redistributed in a lump-sum fashion and they are harmed by the damages that are associated with the country's emissions. However, they take damages and tax revenue as exogenously given and choose x to maximize $u(x, \theta) - qx$. The solution, denoted by $x^*(t, \theta)$, satisfies the first order condition

$$u_x(x^*(t, \theta), \theta) \quad = \quad q(t). \tag{3.1}$$

(3.1) has implications for the comparative statics of consumption: N-types respond to a change in taxes, but not to changes of marginal damages δ. This is akin to an assumption of price-taking behavior. A single individual cannot move the aggregate variable $\bar{x}$ and therefore takes $D(\bar{x}, \delta)$ as exogenously given when deciding on consumption. We denote indirect utility by $v(q(t), \theta)$. For later reference, note that by Roy's identity $v_t(q(t), \theta) = -x^*(q(t), \theta)$, where the subscript t indicates a partial derivative with respect to the tax rate. The per-capita consumption of

[15] The extent of the atmospheric externality depends also on the volume of emissions in the rest of the world X. We treat X as fixed in the analysis that follows and, for ease of notation, do not include it as an argument of the damage function.

N-types is henceforth denoted by

$$\bar{x}_N(t) = \mathbf{E}_N[x^*(q(t), \theta)] \,,$$

where the expectations operator $\mathbf{E}_N$ indicates an aggregation in the set of N-types.

3.2.2 Socially responsible consumers

A Kantian approach to socially responsible consumption could be captures by a rule that say: *Consume according to the rule that generates the socially optimal outcome if followed by every one.* However, our S-agents understand that N-agents will not behave according to this prescription. Therefore, they take the behavior of N-agents as given and focus on the welfare implications of rule-based behavior for ethical agents: *Consume according to the rule that maximizes welfare for S-agents if followed by all S-agents.*[16] This amounts to group rule-utilitarianism, where the rule then takes the form of a function $\theta \mapsto x_S(t, \theta)$ that is chosen to maximize

$$R(t) \;+\; \mathbf{E}_S[u(x_S(t,\theta),\theta) - q\; x_S(t,\theta) - D(\bar{x}(t),\delta)] \,.$$

Here, $R(t)$ is tax revenue and the operator $\mathbf{E}_S$ indicates an aggregation over the set of S-agents. Per-capita tax revenue can also be written as $R(t) = t\; \bar{x}(t)$ where

$$\bar{x}(t) = \phi_S\; \bar{x}_S(t) + \phi_N\; \bar{x}_N(t)$$

is per-capita consumption, $\phi_S = 1 - \phi_N$ is the population share of S-agents, $\bar{x}_S(t)$ is per-capita consumption among S-agents and $\bar{x}_N(t)$ is per-capita consumption among N-agents. The S-agents' problem can therefore also be written as: Choose $\theta \mapsto x_S(t, \theta)$ to maximize

$$t\; \phi_N\; \bar{x}_N(t) \;\; + \;\; \mathbf{E}_S[u(x_S(t,\theta),\theta)] - \big(k + \phi_N\; t\big)\; \bar{x}_S(t) - \mathbf{E}_S[D(\bar{x}(t),\delta)] \,. \tag{3.2}$$

Optimal choices satisfy the first order condition

$$u_x(x_S(t,\theta),\theta) \quad = \quad k \;+\; \phi_N\; t \;+ \phi_S\; \mathbf{E}_S[D_x(\bar{x}(t),\delta)] \,. \tag{3.3}$$

Thus, S-agents internalize the damages they impose on other S-agents and they are responsive to taxes. When almost everybody is an S-agent ($\phi_S \approx 1$), then S-agents hardly respond to taxes and are very responsive to damages. In contrast, if they are mostly surrounded by N-agents, S-agents behave almost like N-agents.

For later reference, note that, with a linear damage function, the consumption-choices of S-agents are as if they were N-agents facing a consumer price of

$$q_S(t) \quad = \quad k \;+\; \phi_N\; t \;+ \phi_S\; \bar{\delta}_S \quad \text{where} \quad \bar{\delta}_S := \mathbf{E}_S[\delta] \,.$$

[16] In the Appendix, we also discuss *universalism*. U-agents feel socially responsible for every one, so they also care for the well-being of N-agents.

In this case, the consumption of an S-agent with type θ is given by $x^*(q_S(t), \theta)$.

3.3 Preferences over taxes

3.3.1 Neoclassical consumers

The policy preference of a (θ, δ) N-type is represented by the utility function

$$U(t, \theta, \delta) \quad = \quad R(t) + v(q(t), \theta) - D(\bar{x}(t), \delta) \ ,$$

which describes her preferences in her role as a citizen who evaluates carbon taxes. Such taxes affect her via three channels: (i) welfare derived as a consumer of the dirty good, as reflected by $v(q(t), \theta)$, (ii) tax revenue, captured by $R(t)$, and (iii) the damages $D(\bar{x}(t), \delta)$ caused by the per-capita consumption of the dirty good $\bar{x}(t)$. The ideal tax rate for the agent solves $U_t(t, \theta, \delta) = 0$, where

$$U_t(t, \theta, \delta) \quad = \quad \bar{x}(t) - x^*(q(t), \theta) \ + \ \Big(t - D_{\bar{x}}(\bar{x}(t), \delta)\Big) \ \bar{x}'(t) \ .$$

The first term, $\bar{x}(t) - x^*(q(t), \theta)$, captures the distributive effect: A tax increase mechanically increases tax revenue (which is redistributed lump sum) proportional to $\bar{x}(t)$. This benefit is counterbalanced by a reduction in consumption utility; by Roy's identity, this loss equals $-x^*(q(t), \theta)$. The net effect is positive for individuals with below-average consumption of the dirty good and vice versa.

Behavioral responses to a tax increase affect $U_t(t, \theta, \delta)$ via the second term, $\big(t - D_{\bar{x}}(\bar{x}(t), \delta)\big) \ \bar{x}'(t)$. Again, there is a netting of two effects. Consumption decreases, which leads to a revenue loss given by $t \ \bar{x}'(t) < 0$, and a reduction in damages given by $-D_{\bar{x}}(\bar{x}(t), \delta) \ \bar{x}'(t) > 0$, which adds to the benefits of a tax increase. Note that $U_{t\theta}(\cdot) < 0$ and $U_{t\delta}(\cdot) > 0$, so that N-types' ideal taxes are decreasing in θ and increasing in δ.

For N-agents, there is a clear distinction between private choices and collective choices. Even if an N-type demands substantial taxes in the political process, this does not affect her consumption choice. More formally, δ is a shifter for an N-agent's marginal rate of substitution between her own numeraire good consumption and the per-capita consumption of the dirty good in the country; i.e. it is a measure of the willingness to pay for a reduction of the country's overall emissions. If this willingness to pay is high, the N-agent wants a high corrective tax, while consuming as much as another N-agent with a lower δ (who therefore wants a lower corrective tax) but has the same marginal utility of consumption. Observing N-agents private choices therefore does not reveal their political preferences.

An individual who thinks that domestic emissions have no measurable impact on global emissions (either because the own country is small, or because he believes that any domestic emission reduction will, through international general equilibrium effects, be completely offset by other countries consuming more) and who also has purely instrumental preferences (i.e., only global emissions matter for him), will have $\delta = 0$. An agent with $\delta > 0$ cares about the country's

carbon footprint for other reasons, unrelated to the country's marginal contribution to global warming. All these possibilities are consistent with our framework. The survey in Section 3.7 documents that, for most individuals in the US and the UK, δ is bounded away from zero.

Single-peaked preferences. The following Proposition 3.1 states sufficient conditions for preferences over taxes to be single-peaked. Single-peakedness implies, in particular, that the ideal tax of an N-agent with type (θ, δ) is well-defined.

Proposition 3.1. *Let* $u(x,\theta) = \theta^{\frac{1}{\varepsilon}} \frac{1}{1-\frac{1}{\varepsilon}} x^{1-\frac{1}{\varepsilon}}$, *for* $\varepsilon > 1$, *and let the damage function be linear. N agents' preferences are single-peaked if one the following statements holds:*

1. *There are only neoclassical consumers:* $\phi_N = 1$.
2. *The cost parameter* k *exceeds a critical value* $\hat{k}$, *the population share* ϕ_S *does not exceed a critical value* $\hat{\phi}_S$, $\theta \in \Theta = [0,1]$, *and* $\bar{\theta}_N > \frac{1}{2}$, *where* $\bar{\theta}_N$ *is the average value of* θ *among* N*-agents.*

Finding sufficient conditions for single-peakedness is more demanding when there is a mix of N- and S-agents. As we show in the proof of Proposition 3.1, an N-agent has an ideal tax for the subpopulation of N-agents and another ideal tax for the subpopulation of S-agents. The ideal tax for the mixed population lies between the two, with the implication that there is a range of taxes where type (θ, δ) is confronted with a trade-off between overtaxing one group and undertaxing the other one. The conditions in Proposition 3.1 are needed to ensure that, for any given pair (θ, δ), the function $t \mapsto U(t, \theta, \delta)$ is concave over this range.

How do N-agents' political preferences depend on the population share of S-agents? The function $t \mapsto U(t, \theta, \delta)$ depends parametrically on the population share of S-agents, ϕ_S. To make this explicit we expand our notation and write $t \mapsto U(t, \phi_S, \theta, \delta)$.

Proposition 3.2. *Suppose that damage functions are linear. Consider an* N*-agent with type* (θ, δ). *Let* $t \leq \delta$ *and* $t < \bar{\delta}_S$. *Suppose that for all such* t, $\bar{x}_N(t) \leq \bar{x}_S(q_S(t))$.[17] *Then there is* $\hat{\phi}_S$ *so that, for* $\phi_S < \hat{\phi}_S$, $U_{t\phi_S}(t, \phi_S, \theta, \delta) \quad < \quad 0$.

Proposition 3.2 considers the political preferences of a neoclassical consumer of type (θ, δ). It starts from an initial situation in which taxes are low: They neither exceed the marginal damages δ that are relevant to this individual, nor the damages $\bar{\delta}_S$ that are relevant to socially responsible consumers. It then shows that the marginal benefits of raising taxes are decreasing in the population share of S-agents. This is driven by two forces: S-agents consume less than N-agents and they are less responsive to tax changes. Thus, higher taxes bring less revenue and they are less effective in mitigating damages when more S-agents are around. Proposition 3.2 can be interpreted as a crowding-out result: A larger share of S-agents lowers the N-agents' ideal taxes.

[17] This is ensured if $\bar{\theta}_N \geq \bar{\theta}_S$ and $u(x,\theta) = \theta^{\frac{1}{\varepsilon}} \frac{1}{1-\frac{1}{\varepsilon}} x^{1-\frac{1}{\varepsilon}}$, for $\varepsilon > 1$.

3.3.2 Socially responsible consumers

Denote by $W_S(t)$ the value function associated with the group rule-utilitarian consumption problem in (3.2). The welfare realized by S-agents depends on the tax level t. The derivative $W_S'(t)$ gives the S-agents' marginal benefit from a tax increase. An application of the Envelope Theorem yields

$$W_S'(t) \quad = \quad \phi_N(\bar{x}_N(t) - \bar{x}_S(t)) \; + \phi_N \left(t - \mathbf{E}_S[D_x(\bar{x}(t), \delta)]\right) \bar{x}_N'(t) \; .$$

The first term in this sum captures the S-agents' distributive benefit; S-agents benefit if they consume less than N-agents, and this effect is the more important the larger is ϕ_N. The second term captures the corrective effect of taxation; when taxes rise, N-agents consume less ($\bar{x}_N'(t) < 0$), and if the tax rate falls short of the marginal damages perceived by the S-agents ($t < \mathbf{E}_S[D_x(\bar{x}(t), \delta)]$), then the marginal corrective taxation benefits are positive. Observe that S-agents want to correct only the N-agents' demand for dirty goods, while their own demand needs no correction.

Single-peaked preferences. Suppose for a moment that the distribution of θ is identical among N- and S-agents. If $t = \mathbf{E}_S[D_x(\bar{x}(t), \delta)]$ then per-capita consumption is the same in both groups, $\bar{x}_N(t) = \bar{x}_S(t)$. Hence $W_S'(t) = 0$. By contrast, $t < \mathbf{E}_S[D_x(\bar{x}(t), \delta)]$ implies $\bar{x}_N(t) > \bar{x}_S(t)$ and hence $W_S'(t) > 0$; whereas $t > \mathbf{E}_S[D_x(\bar{x}(t), \delta)]$ implies $\bar{x}_N(t) < \bar{x}_S(t)$ and hence $W_S'(t) < 0$. Thus, the function W is single-peaked with a peak at $t = \mathbf{E}_S[D_x(\bar{x}(t), \delta)]$. Proposition 3.3 extends this observation to a case in which N-agents exceed the consumption of S-agents even when $t = \mathbf{E}_S[D_x(\bar{x}(t), \delta)]$.

Proposition 3.3. *Suppose that $\bar{\theta}_N \geq \bar{\theta}_S$. Let $u(x, \theta) = \theta^{\frac{1}{\varepsilon}} \, \frac{1}{1-\frac{1}{\varepsilon}} \, x^{1-\frac{1}{\varepsilon}}$, for $\varepsilon > 1$. Suppose that damage functions are linear. Then S-agents have single-peaked preferences. The peak t_W^* equals $\bar{\delta}_S$ if $\bar{\theta}_N = \bar{\theta}_S$ and exceeds $\bar{\delta}_S$ if $\bar{\theta}_N > \bar{\theta}_S$.*

When $\bar{\theta}_N > \bar{\theta}_S$, the S-agents' ideal tax involves overtaxation relative to the Pigouvian benchmark with $t = \bar{\delta}_S$. The S-agents then want to take advantage of the N-agents' larger demand for dirty goods.

How do the political preferences of S-agents depend on the population share of S agents? The function $t \mapsto W_S(t)$ which captures the political preferences of S-agents depends in a parametric way on the population share of S-agents, ϕ_S. Again, to make this explicit we expand our notation and write $t \mapsto W_S(t, \phi_S)$.

Proposition 3.4. *Let $t \leq \bar{\delta}_S$. Suppose that $\bar{\theta}_N \geq \bar{\theta}_S$ and $u(x, \theta) = \theta^{\frac{1}{\varepsilon}} \, \frac{1}{1-\frac{1}{\varepsilon}} \, x^{1-\frac{1}{\varepsilon}}$, for $\varepsilon > 1$. Then, $\frac{\partial^2 W_S(t, \phi_S)}{\partial t \partial \phi_S} < 0$.*

From the perspective of S-agents, the benefits from higher taxes are smaller when their own population share is larger. Both the corrective and the redistributive benefits of higher taxes are due to N-agents. If there are fewer of them, it becomes less attractive to tax. Note that

N-agents share this view. According to Proposition 3.2, an N-agent's marginal benefit from taxes also decreases when there are more S-agents.

3.4 Political equilibrium

We do not adopt a specific model of the political process, but take the shortcut of equating the median voter's preferred tax policy with the political equilibrium outcome.[18] In the following, we characterize political equilibrium taxes for various special cases of our general setup: The purpose is to clarify how political equilibrium taxes are shaped by (i) heterogeneity in the demand for dirty goods, (ii) heterogeneity in the assessment of damages and (iii) the mix between neoclassical and socially responsible consumers.

3.4.1 Country A: Neoclassical consumers only

Heterogeneity only in the demand for dirty goods. Consider a country that is inhabited only by N-types. Suppose, moreover, that these agents have a common understanding of the relevant damages, i.e., $\delta = \delta_N$ for all N-agents. The ideal tax rate for a type θ agent then solves $U_t(t, \theta, \delta_N) = 0$, where

$$U_t(t, \theta, \delta) \quad = \quad \bar{x}(t) - x^*(q(t), \theta) \quad + \quad \Big(t - D_{\bar{x}}(\bar{x}(t), \delta_N)\Big)\, \bar{x}'(t) \; .$$

The decisive voter is the person with the median value of θ, denoted by θ_M. If the distribution of consumption levels is symmetric so that median and average consumption coincide ($\bar{x}(t) = x^*(q(t), \theta)$), then $t = D_{\bar{x}}(\bar{x}(t), \delta_N)$, i.e., the median voter's preferred tax rate is the standard "Pigouvian tax."

If the distribution is asymmetric, then distributive effects must be considered in addition to the Pigouvian motive. If the median voter consumes less of the dirty good than the average consumption level, then he distributionally benefits from higher taxes, and the political equilibrium involves overtaxation relative to the Pigouvian benchmark, and vice versa. Formally, if $x^*(q(t), \theta) < \bar{x}(t)$, then $U_t(t, \theta_M, \delta_N) = 0$, implies $t > D_{\bar{x}}(\bar{x}(t), \delta_N)$, and vice versa. This result is related to the well-known Meltzer-Richard hypothesis in the political economy of income taxation where equilibrium taxes are predicted to be increasing the gap between mean and median income. Here, an analogous force is embedded into an analysis of corrective taxation.

Heterogeneity in the demand for dirty goods and in damage assessments. Suppose now that there is a non-degenerate joint distribution of θ and δ among N-types. Since

[18] There are alternative specifications that could be used. When preferences are single-peaked, the median voter's preferred policy is a Condorcet winner, i.e., it wins a majority against any other alternative. One implication is that the median voter's preferred policy is the equilibrium outcome in a Downsian model of two party competition. Another implication is that the median voters' ideal policy is the (only) stable outcome when citizens can propose an alternative to the status quo and initiate a referendum. By Black's (1948) theorem, the majority is rule transitive when preferences are single-peaked. Therefore, a sequence of pairwise votes "converges" to the median voter's ideal policy.

preferred tax rates are decreasing in θ and increasing in δ, those with the lowest θ and the highest δ, prefer the highest tax rate, and those with the highest θ and the lowest δ prefer the lowest tax rate.

For example, let $u(x,\theta) = \theta^{\frac{1}{\varepsilon}} \frac{1}{1-\frac{1}{\varepsilon}} x^{1-\frac{1}{\varepsilon}}$, for $\varepsilon > 1$, and let $\overline{\theta}_N := \mathbf{E}_N[\theta]$. Suppose that damage functions are linear. The preferred tax rate of an N-agent of type (θ, δ) is given by

$$t^*(\theta,\delta) \quad = \quad \frac{\frac{1}{\varepsilon}\left(1-\frac{\theta}{\overline{\theta}_N}\right)k + \delta}{1-\frac{1}{\varepsilon}\left(1-\frac{\theta}{\overline{\theta}_N}\right)} \,. \tag{3.4}$$

The optimal tax is decreasing in θ and increasing in δ. Formula (3.4) also displays the inverse elasticities logic that is familiar from optimal tax theory: By how much an agent with type $\theta < \bar{\theta}_N$ wants to raise taxes above δ is inversely related to the elasticity ε.

3.4.2 Country B: Socially responsible consumers only

Consider a country composed only of S-agents. It follows from the discussion In Section 3.3.2 that the welfare realized by ethical agents does not depend on the tax rate when everyone is an ethical agent.[19] The consumption of S-agents then follows from

$$u_x(x_S(\theta),\theta) \quad = \quad k \; + \mathbf{E}_S[D_x(\bar{x}(t),\delta)] \,,$$

i.e. everyone fully internalizes the social costs of the dirty good.

Comparing countries A and B. For ease of comparison, suppose that N- and S-agents agree on the damages (i.e., $\delta = \bar{\delta}$ for all N- and S-agents). Also suppose that the distribution of θ among N-agents and among S-agents is the same.

The outcome in country B with only ethical agents is a normative benchmark and may be considered the best outcome one may hope for. Remarkably, if mean and median consumption coincide, Country A with only selfish agents reaches the same outcome via a Pigouvian tax. If mean and median consumption diverge, then Country A will have higher or lower taxes and, correspondingly, a somewhat different allocation.

3.4.3 Country C: A mix of N- and S-agents

We now turn to country C with a mixed population consisting of N- and S-agents. For tractability, we make the following assumptions: Let the damage function be linear and suppose that there is common group-wide understanding among N- and S-agents, respectively, of the relevant damages, i.e. $\delta = \delta_N$ for all N-agents, and $\delta = \delta_S$ for all S-agents. Furthermore, the distributions of consumption preferences are degenerate: For all N-agents, $\theta = \theta_N$, and for all S-agents, $\theta = \theta_S$. Assume that $\theta_N \geq \theta_S$ and $\delta_N \leq \delta_S$.

[19] Formally, for all t, $W'_S(t) = 0$ if $\phi_S = 1$.

These assumptions ensure that, when $t \leq \delta_S$, the representative N-agent consumes more of the dirty good than the representative S-agent, $\bar{x}_N(t) \geq \bar{x}_S(t)$, with equality only if $\theta_N = \theta_S$ and $t = \delta_S$. Hence, carbon taxes are redistributing from the representative N-agent to the representative S-agent. We summarize the implications in the following Corollary that we state without proof.

Corollary 3.1. *Let $0 < \phi_N < 1$, $\theta_N \geq \theta_S$ and $\delta_N \leq \delta_S$.*

1. *The representative N-agent's preferred tax t_N^* is weakly smaller than δ_N.*
2. *The representative S-agent's preferred tax t_S^* is weakly larger than δ_S.*
3. *Their preferred taxes coincide if and only if $\theta_N = \theta_S$ and $\delta_N = \delta_S$.*

Suppose that $\theta_N > \theta_S$ so that the distributive consequences of corrective taxes go against N-agents and are favorable to S-agents. Then, relative to their respective Pigouvian benchmarks, N-agents want undertaxation and S-agents want overtaxation because of the presence of the other group.

If the median voter is an N-agent, then there is undertaxation in equilibrium, both relative to the Pigouvian benchmarks of N-agents, $t_N^P = \delta_N$, and of S-agents, $t_S^P = \delta_S$. The third observation in the corollary is that a mix of N- and S-agents is not a problem per se. Ethical diversity does not spur political conflict when the representative N- and S-agent have the same consumption benefits and damage assessment.

A political economy waterbed effect. Suppose that at least one of the inequalities $\theta_N \geq \theta_S$ and $\delta_N \leq \delta_S$ holds strictly so that there is some political disagreement, and consider what happens if some N-agents are converted into S-agents. Clearly, for a given tax, the demand for the dirty good goes down. But the tax is not given; by Proposition 3.2, the political equilibrium tax goes down (assuming that N-agents remain in the majority). This countervailing effect pushes dirty good consumption upwards.

To see this formally, we write the equilibrium tax rate as a function of the population share of S-agents, $t^{eq}(\phi_S)$, and denote its derivative by $t^{eq\prime}(\phi_S) < 0$. Equilibrium consumption of the dirty good is then given by

$$\bar{x}^{eq}(\phi_S) \;:=\; \phi_S\, \bar{x}_S(q_S(t^{eq}(\phi_S), \delta_S)) + (1-\phi_S)\, \bar{x}_N(q(t^{eq}(\phi_S)))\,,$$

and its derivative equals

$$\begin{aligned} \bar{x}^{eq'}(\phi_S) \;=\; & -\Big(\bar{x}_N(q(t^{eq}(\phi_S))) - \bar{x}_S(q_S(t^{eq}(\phi_S)), \delta_S)\Big) \\ & + t^{eq\prime}(\phi_S)\, \Big(\phi_S\, \bar{x}_{S,q_S}(\cdot) + (1-\phi_S)\, \bar{x}_{N,q}(\cdot)\Big)\,, \end{aligned}$$

where the first line contains the mechanical effect when N-agents are replaced by S-agents, and

the second line contains the waterbed effect due to the tax change.[20] The partial derivatives $\bar{x}_{S,q_S}(\cdot)$ and $\bar{x}_{N,q}(\cdot)$ are both negative so that the second line is positive. This decomposition raises the question which of these effects is dominant. Figure 3.1 illustrates a setting where the second line dominates: Having more socially responsible agents then increases equilibrium emissions.

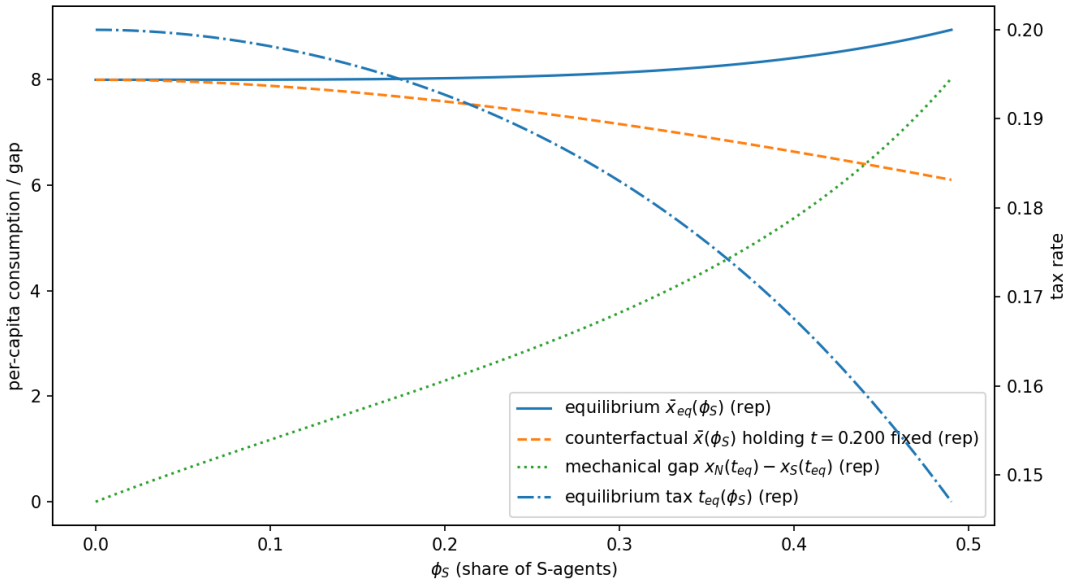


Figure 3.1: Strong Waterbed Effect

Notes: Equilibrium per-capita consumption of the dirty good as a function of ϕ_S. The solid line is equilibrium aggregate consumption $\bar{x}^{eq}(\phi_S)$; the dashed line is the counterfactual $\bar{x}(\varphi_S)$ holding the tax fixed at $t^{eq}(\phi_S{=}0)$; the dotted line is the mechanical consumption gap $x_N(t^{eq}) - x_S(t^{eq})$. The dash-dotted line shows the equilibrium tax $t^{eq}(\phi_S)$, measured on the right-hand vertical axis. Parameters: $\varepsilon = 2$, $k = 0.05$, $\delta_S = 0.40$, $\delta_N = 0.20$, $\theta_N = \theta_S = 0.5$.

It is also possible, however, that the mechanical effect dominates the waterbed effect so that pollution decreases when more S-agents are around. This case is illustrated by Figures 3.2a and 3.2b, which are not based on a specification with a representative S-agent and a representative N-agent. Instead, δ_N is dispersed across the N-population, while θ is equally distributed across the two populations, with $\bar{\theta}_N = \bar{\theta}_S$.[21]

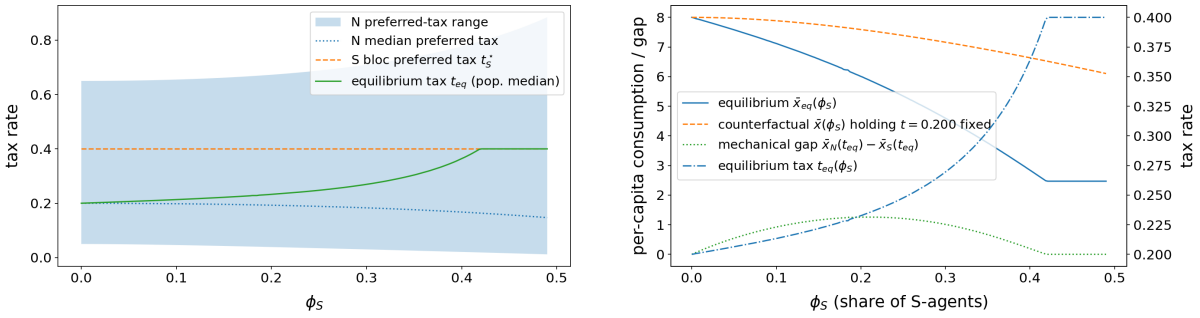
Panel (a) of Figure 3.2a shows the political equilibrium tax as a function of ϕ_S, together with the S-bloc's preferred tax and the within-N median preferred tax. The S-bloc's preferred tax (orange dashed) is at $\bar{\delta}_S = 0.40$, as Proposition 3.3 implies under $\bar{\theta}_N = \bar{\theta}_S$. The within-$N$ median preferred tax (blue dotted) starts near $\bar{\delta}_N = 0.20$ and drifts downward as ϕ_S grows, consistent

[20] In Kaufmann et al. 2024, a waterbed effect arises for a different reason, namely an upward sloping supply function. Therefore, if socially responsible consumers lower their demand, this lowers prices and leads to demand adjustments by others that partially, but not fully, offset the initial reduction. Here, by contrast, the dirty good is produced with constant returns to scale so that the supply curve is horizontal. Changes of consumer prices are therefore entirely due to tax changes.

[21] Under $\bar{\theta}_N > \bar{\theta}_S$, Proposition 3.3 instead implies $t^*_S > \bar{\delta}_S$: S-agents prefer a tax strictly above their own damage assessment, leaning on the N-bloc's larger demand for the dirty good. Consistent with the logic of Proposition 3.4, this t^*_S then falls as ϕ_S grows, since the N tax base on which the S-bloc leans shrinks. Appendix 6.1.10 reproduces Figure 3.2 for this case, and the numerical exercise makes both forces directly visible: t^*_S sits above $\bar{\delta}_S$ throughout and slopes downward in ϕ_S.

with Proposition 3.2. The equilibrium tax (green solid) is the median of the pooled electorate, and two forces shape it. The first is the individual-level crowding-out of Proposition 3.2, which pulls the preferred tax of the majority of N-agents down. The second is a composition effect that operates on the position of the decisive voter: as ϕ_S grows, the S-bloc adds voting mass at 0.40 and the decisive voter walks rightward through the N-agent distribution of preferred taxes. With heterogeneous δ_N, the second force dominates, and the equilibrium tax rises smoothly from $\bar{\delta}_N$ toward $\bar{\delta}_S$. At $\phi_S \approx 0.42$, the cumulative weight of the S-bloc plus the N-agents whose preferred tax is above $\bar{\delta}_S$ crosses one-half, and the decisive voter switches to the S-bloc itself. The equilibrium tax then pins at $\bar{\delta}_S$ for all larger ϕ_S. In contrast to the homogeneous-δ_N benchmark, where Proposition 3.2 alone pulls the equilibrium tax down with ϕ_S and induces a compensating N-agent consumption response, here the median-composition effect pushes the equilibrium tax up.[22]

Figure 3.2b shows the consumption consequences. Equilibrium per-capita consumption (blue solid) falls. The counterfactual (orange dashed) freezes the tax at $t^{eq}(\phi_S{=}0) = 0.20$ and so isolates the composition channel: replacing N-agents with S-agents, who face $q_S > q_N$, mechanically reduces aggregate demand. The remaining decline is the policy channel, and reflects the fact that the rising equilibrium tax depresses both groups' consumption further.



(a) Political-Equilibrium Tax

(b) Aggregate Consumption

Figure 3.2: Weak Waterbed Effect

Notes: Panel (a): The shaded band gives the cross-sectional range of N-agents' preferred taxes over the support of (θ, δ_N). The dotted blue line is the within-N median preferred tax. The orange dashed line is t_S^*, invariant at $\bar{\delta}_S = 0.40$ (Proposition 3.3). The green solid line is the political equilibrium tax, which is the median preferred tax over the entire population. The discontinuity at $\phi_S \approx 0.42$ marks the switch of the decisive voter from an N-agent to the S-bloc. Panel (b): The blue solid line is equilibrium per-capita consumption $\bar{x}^{eq}(\phi_S) = (1-\phi_S)\bar{x}_N(t^{eq}) + \phi_S \bar{x}_S(q_S(t^{eq}, \phi_S))$. The orange dashed line holds the tax fixed at $t^{eq}(\phi_S{=}0) = 0.20$, isolating the composition channel. The green dotted line is the consumption gap $\bar{x}_N(t^{eq}) - \bar{x}_S(t^{eq})$. The blue dash-dotted line (right axis) is $t^{eq}(\phi_S)$ from panel (a). Parameters: $\varepsilon = 2$, $k = 0.05$, $\theta \sim \text{Uniform}[0,1]$ in both groups, $\delta_N \sim \text{Uniform}[0.1, 0.3]$, $\delta_S \sim \text{Uniform}[0.3, 0.5]$.

[22]The band of N-agent preferred taxes extends well above the within-N median. Its upper edge corresponds to low-θ, high-δ_N types, whose objective collapses to maximising net revenue $(t - \delta_N)\bar{x}(t)$ and whose preferred tax is therefore $t^* = \delta_N + \bar{x}/|\bar{x}'|$. As ϕ_S rises, the inelastic demand of S-agents makes the tax base less tax-responsive, the Laffer-style markup $\bar{x}/|\bar{x}'|$ grows, and the corner type's preferred tax rises with it.

3.5 Moral pressure

As shown in the previous section, the equilibrium tax may be insufficient from the perspective of S-agents. We now extend the model and give S-agents the opportunity to express their discontent with the carbon footprints of N-Agents. We are particularly interested in how such behavior affects tax preferences and, ultimately, the political equilibrium. Our approach is motivated both by climate protests in the streets, such as the Fridays-for-Future movement, but also by concerns raised in private conversations, e.g. on car use or short-distance flights. In such conversations, those who drive or fly may be frowned at, especially when less carbon-intense alternatives are available.

We will model an attempt of S-agents to establish and uphold a social norm to "worry about the climate and to reduce the own carbon footprint." Our approach is inspired by a literature that uses group-rule utilitarianism to model turnout in elections. In this literature, it is assumed that voting is costly and that ethical voters who are politically like-minded determine a cutoff level of voting costs. Those with costs below the cutoff are then morally obliged to vote, whereas those with higher costs are excused. The cutoff is a measure of how hard the group is fighting to win the election. If the stakes are high, then the cutoff is also high and almost everybody should turn out. Otherwise, there is more tolerance of "indifference-based abstention" see Coate and Conlin 2004, Feddersen and Sandroni 2006 or Bierbrauer et al. 2022.

We assume that upholding the social norm is costly. It requires to participate in protests or to take an unpleasant role in private conversations, namely to criticize others for their lifestyle. S-agents are heterogeneous in how costly it is for them to engage in such activities. We assume that S-agents apply a rule-utilitarian calculus not only to their consumption of dirty goods, but also to how confrontational they are in their fight for the social norm. Whether such norm enforcement pays off depends on how N-agents respond to it. In the following, we will assume that N-agents reduce their consumption of the dirty good when they get more moral pressure from S-agents. From the perspective of S-agents, this is the benefit from fighting harder. We now introduce a formalism that captures these considerations.

Denote by σ the fraction of S-agents who participate in upholding the social norm. We assume that S-agents have an idiosyncratic cost c of doing this. This cost is taken to be an *iid* draw from a uniform distribution on the unit interval. Those with a cost below σ call out norm violations, those with a cost above σ feign indifference. For S-agents as a group, the total cost associated with a participation rate of σ is

$$C(\sigma) \quad = \quad \int_0^\sigma c \, dc \quad = \quad \frac{1}{2}\,\sigma^2 \,.$$

S-agents create a "shaming" function $x \mapsto S(x, \phi_S\,\sigma)$, where x is the consumption of the dirty good and $\phi_S\,\sigma$ is the population share of agents who "sanction" norm violations. We assume that this function is increasing in the first argument and has a positive cross-derivative; that is, the shame associated with, say, an extra flight is increasing when more people show their disapproval. For convenience, we impose a linear functional form in what follows and let

$S(x, \phi_S\, \sigma) = x\ \phi_S\, \sigma$. N-agents choose x to maximize

$$u(x, \theta) - q\ x - S(x, \phi_S\, \sigma) \tag{3.5}$$

The first order condition characterizing their behavior then is

$$u_x(\cdot) \quad = \quad q(t) + \phi_S\, \sigma\ .$$

The solution to the consumer choice problem is $x^*(q(t) + \phi_S\, \sigma, \theta)$. Thus, N-agents now behave as if they were facing a higher price than the tax-inclusive price $q(t)$.[23] S-agents understand that N-agents respond to moral pressure in this way. They choose σ to minimize their total costs from damages and social norm enforcement. Assuming a linear damage function, this objective can be written, with $\bar{\delta}_S \equiv E_S(\delta)$, as

$$\bar{\delta}_S\ \phi_N\ \bar{x}_N(q(t) + \phi_S\, \sigma) \quad + \quad C(\sigma)\ . \tag{3.6}$$

It is easy to verify that (3.6) is decreasing in σ for σ close to 0. Thus, the optimum involves $\sigma > 0$. Moreover, with iso-elastic consumption utility, the objective is globally convex. Thus, either the S-agents go all in and choose $\sigma = 1$, or they choose an interior σ implicitly defined by the rearranged first-order condition

$$\sigma \quad = - \quad \bar{\delta}_S\ \phi_N\ \phi_S\ \bar{x}_{N,q}\big(q(t) + \phi_S\, \sigma\big)\ . \tag{3.7}$$

Lemma 3.1 below, which we state without proof, shows that the S-agents' pressure goes down if taxes are higher.[24] Intuitively, if N-agents already reduce their demand because high taxes, there is less need for moral pressure. Lemma 3.1 also shows that, while σ falls when t rises, the tax-and-norm-inclusive-price $q(t) + \phi_S\ \sigma_t^*(t)$ increases in t. Thus, while S-agents become more tolerant when taxes go up, the net effect of a higher tax is still a reduction of the dirty goods demand of N-agents.

Lemma 3.1. *Let* $u(x, \theta) = \theta^{\frac{1}{\varepsilon}}\ \frac{1}{1-\frac{1}{\varepsilon}}\ x^{1-\frac{1}{\varepsilon}}$ *, for* $\varepsilon > 1$*. Let* $\sigma^*(t)$ *be the value of* σ *that solves (3.7). Then* $\sigma_t^*(t) < 0$ *and* $1 + \phi_S\, \sigma_t^*(t) > 0$*.*

We now revisit the N-agents' tax preferences. In particular, we seek to clarify how they are affected by the S-agents' moral pressure.

Proposition 3.5. *Let* $u(x, \theta) = \theta^{\frac{1}{\varepsilon}}\ \frac{1}{1-\frac{1}{\varepsilon}}\ x^{1-\frac{1}{\varepsilon}}$ *, for* $\varepsilon > 1$*. Consider an* N*-agent of type* (θ, δ) *and an ideal tax of* $t^*(\theta, \delta) \leq \delta$ *when* $\sigma = 0$*. Denote by* $t_N^*(\theta, \delta)$ *this agents' ideal tax with endogenous moral pressure. There is* $\hat{\phi}_S$ *so that* $\phi_S < \hat{\phi}_S$ *implies* $t_N^*(\theta, \delta) > t^*(\theta, \delta)$*.*

Given the S-agents' pressure, N-agents evaluate tax policy using the function

$$U(t, \theta, \delta) \quad = \quad (t - \delta)\ \bar{x}\big(t, \phi_S\, \sigma^*(t)\big)\ +\ \max_x \Big\{ u(x, \theta) - \big(q(t) + \phi_S\, \sigma^*(t)\big)\ x \Big\}\ ,$$

[23] We follow Farhi and Gabaix 2020 in that behavioral interventions or "nudges" – here: the moral pressure exerted by S-agents – make people behave as if they were confronted with a higher tax. A nudge differs from a tax, however, in that it does not generate tax revenue.

[24] The Lemma follows from a straightforward application of the implicit function theorem.

where $\bar{x}\big(t, \phi_S\, \sigma^*(t)\big) \;=\; \phi_N\, \mathbf{E}_N[x^*(q(t)+\phi_S\, \sigma^*(t)), \theta] \,+\, \phi_S \mathbf{E}_N[x^*(q_S(t), \theta]$. By the envelope theorem, the marginal benefit of higher taxes is given by

$$U_t(t, \theta, \delta) \;=\; \bar{x}(\cdot) \;-\; x^*(\theta, \cdot)\Big(1 + \hat{\sigma}_t^*(t)\Big) \;+\; \Big(t - \delta\Big)\, \bar{x}_t(\cdot)\;.$$

Higher taxes decrease consumer surplus. When $\hat{\sigma}_t^*(t) < 0$ the loss from higher taxes is dampened because moral pressure goes down at the same time. The preferred tax rates of neoclassical consumers are therefore larger in the presence of moral pressure.[25]

The premise of Proposition 3.5 is $t^*(\theta, \delta) \leq \delta$ when $\sigma = 0$.[26] As shown in Section 3.4.3, in a country with representative N- and S-agents, $\theta_N \geq \theta_S$ and $\delta_N \leq \delta_S$, the representative N-agent's preferred tax is bounded from above by δ_N, so that the premise of Proposition 3.5 is satisfied so that the N-agents' preferred tax increases. If the representative N-agent is, moreover, the decisive median voter, then the political equilibrium tax goes up. Importantly, this upward revision does not arise because the median voter now internalizes a social norm to reduce the own carbon footprint. She is not turned into an S-agent, but rather understands that the S-agents' moral pressure is costly for her, and that these costs decrease when taxes are higher.

3.6 Moral rebellion

Right-wing populism often rejects "moralizing" in politics and argues that policy-making should be driven more by national self-interest and be less constrained by (leftist) ethical principles. In the following, we present an attempt to introduce such a "rebellion against moralizing" into the model of Section 3.5. While that model links moral pressure by ethical agents to political equilibrium outcomes, it does so in a simplistic manner. In particular, the pressured agents can respond only by adjusting their consumption of the dirty good and by recalibrating their preferences over taxes.

We now give them the opportunity to reject the demands of ethical agents. Specifically, we assume that, when taking their consumption decisions, N-agents make a binary choice $\gamma \in \{0, 1\}$ whether to accept ($\gamma = 0$) or reject ($\gamma = 1$) the social norm. We assume that rejecting the social norm comes with a (non-monetary) cost, but less so if this is very common among N-agents. Formally, consider an N-agent who expects that (i) σ is the share of S-agents upholding the social norm, (ii) the share of N-agents rejecting the social norm equals $\bar{\gamma}$. Given this expectation, the N-agent chooses $\gamma \in \{0, 1\}$ and the consumption of the dirty good x to maximize

$$u(x, \theta) \;-\; q(t)\, x \;-\; \phi_S\, \sigma\, (1 - \gamma)\, x \;-\; \frac{1}{\bar{\gamma}}\, \gamma\;.$$

[25] This force interacts, however, with the fact that $\bar{x}(\cdot)$, $x^*(\cdot)$ and $\bar{x}_t(\cdot)$ depend on $\sigma^*(t)$. Under the premises of the Proposition, and as formally shown in the Appendix, these effects reinforce each other so that an N-agent's preferred tax with the social norm is larger than without.

[26] Proposition 3.5 states only a sufficient condition, so preferred taxes may be higher also when its premises do not apply.

Moral rebellion is more attractive for high θ types. An N-agent rejects the social norm if the gain in consumption utility outweighs the cost; this is the case when

$$v(q(t),\theta) - v(q(t) + \phi_S\, \sigma, \theta) \quad \geq \quad \frac{1}{\bar{\gamma}}\,. \tag{3.8}$$

Using Roy's identity, the left-hand side can be written as $\int_0^\sigma x^*(q + \phi_S\, s, \theta)ds$, an expression that is increasing in θ. Therefore, there is a cutoff type $\hat{\theta}(\bar{\gamma}, \sigma, q)$ so that all N-agents with $\theta > \hat{\theta}(\cdot)$ choose $\gamma = 1$ and all N-agents with $\theta < \hat{\theta}(\cdot)$ choose $\gamma = 0$.

Monotone comparative statics arguments can be applied to the value $\hat{\theta}$ that satisfies (3.8) with equality. This yields the following observations (i) $\hat{\theta}_{\bar{\gamma}}(\cdot) < 0$, revealing a complementarity amongst the choices of N-agents: if the share of N-agents rejecting the social norm goes up, the cutoff $\hat{\theta}$ goes down so that rejecting pays off for more N-agents; (ii) $\hat{\theta}_\sigma(\cdot) < 0$, more pressure from S-agents implies that more N-agents find it attractive to reject the social norm; and (iii) $\hat{\theta}_q > 0$; higher taxes reduce the consumption of the dirty good and fewer N-agents reject the social norm.

Equilibrium in the coordination game played by N-agents. Given σ, N-agents play a coordination game in which, in equilibrium, for all N-agents have correct expectations and their choice of γ is individually optimal, given their expectations about $\bar{\gamma}$. Given σ and q, this requires that $\bar{\gamma}$ solves the following fixed point equation

$$\bar{\gamma} \quad = \quad 1 - F_N(\hat{\theta}(\bar{\gamma}, \sigma, q))\,, \tag{3.9}$$

where F_N is a cumulative distribution function, with density denoted by f_N, that describes the distribution of θ among N-agents. Thus, equation (3.9) says that $\bar{\gamma}$ is equal to the mass of N-agents with θ above the cutoff $\hat{\theta}(\cdot)$. There may be multiple solutions of (3.9) in which case there are multiple equilibria in the coordination game among N-agents (see Figure 3.3 for an illustration). Note first that $\bar{\gamma} = 0$ "solves" the fixed point equation in (3.9).[27] This implies, in particular, that the political equilibrium that was characterized in Section 3.5 continues to be a political equilibrium here: There is an equilibrium in which all N-agents accept the moral pressure and adjust their consumption accordingly.

To see that there is also scope for an equilibrium with moral rebellion note that, at $\bar{\gamma} = 1$, the left hand side of (3.9) exceeds the right-hand side as soon as F_N has positive mass to the left of $\hat{\theta}(1, \sigma, q)$.[28] If, moreover the slope of the right-hand side exceeds 1 at $\bar{\gamma} = 0$, there have to be multiple equilibria. Finally, if the right-hand side is a concave function of $\bar{\gamma}$, then there is exactly one further stable equilibrium.[29]

[27] More formally, equation (3.8) defines $\hat{\theta}$ only for $\bar{\gamma} > 0$. Note that $\bar{\gamma} \to 0$ implies that $\hat{\theta} \to \infty$ and hence that $1 - F_N(\hat{\theta}) \to 0$.

[28] In Words: When everybody expects that all N-agents reject the moral claims of S-agents, then there are some N-agents for whom rejecting does not pay off. This contradicts the existence of an equilibrium in which all N-agents reject.

[29] An equilibrium is called stable when the slope of the right-hand side is smaller than 1, with the implication that $1 + f_N(\cdot)\, \hat{\theta}_{\bar{\gamma}}(\cdot) \quad > \quad 0\,.$

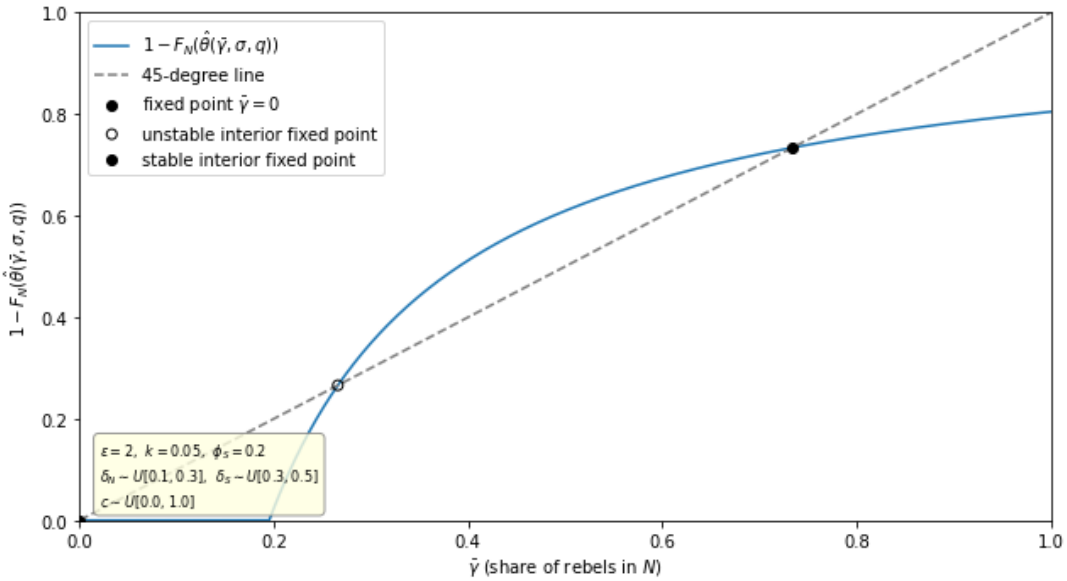


Figure 3.3: Rebellion in Response to Moral Pressure

Notes: The figure shows the solutions of equation (3.9), with moral pressure intensity fixed at its equilibrium level σ^{eq}. The curve plots $1 - F_N(\hat{\theta}(\bar{\gamma}, \sigma^{eq}, q))$; intersections with the 45-degree line are equilibria. The filled marker denotes the stable interior fixed point and the open marker the unstable one. Parameters: $\varepsilon = 2$, $k = 0.05$, $t = 0.05$, $\varphi_S = 0.20$; $\theta \sim$ Uniform$[0, 1]$ in both groups; $\delta_N \sim$ Uniform$[0.1, 0.3]$, $\delta_S \sim$ Uniform$[0.3, 0.5]$.

The following Lemma that we state without proof states that comparative statics properties of a stable fixed point.[30]

Lemma 3.2. *Given σ and q, let $\bar{\gamma}^*(\sigma, q)$ be a stable, interior fixed point. Then $\hat{\gamma}^*_\sigma(\sigma, q) > 0$ and $\hat{\gamma}^*_q(\sigma, q) < 0$.*

According to Lemma 3.2, in a stable equilibrium with moral rebellion by some N-agents, the intensity of rebellion increases in σ, the moral pressure of S-agents, and decreases in the tax rate which drives the consumer price q up.

S-agents' best response: Adjusting the moral pressure to the fraction of non-rebel N-agents. S-agents will understand that their moral pressure now only affects the behavior of those N-agents with a type below $\hat{\theta}$, so they minimize

$$\frac{1}{2}\,\sigma^2 \quad + \quad \delta_S\,\phi_N \int_0^{\hat{\theta}} x^*(q + \phi_S\,\sigma, \theta)\; f_N(\theta)\; d\theta\;. \tag{3.10}$$

Again, assume an interior solution $\sigma^*(q, \hat{\theta})$ that is characterized by a first order condition

$$\sigma^* \quad = \quad -\delta_S\,\phi_N\,\phi_S \int_0^{\hat{\theta}} x^*_q(q + \phi_S\,\sigma^*, \theta)\; f_N(\theta)\; d\theta\;. \tag{3.11}$$

[30]The Lemma follows from a straightforward application of the implicit function theorem, upon using the comparative statics properties of $\hat{\theta}$ and the requirement of stability.

It is easy to see that σ^* is increasing in $\hat{\theta}$: the more people respond to pressure, the larger is the incentive for S-agents to exert it. As already discussed in Section 3.5, σ^* is decreasing in q; when consumer prices are high, there is less of a need for a social norm.

Moral pressure and moral rebellion as a Nash-equilibrium. For a given a price q, we define a Nash equilibrium as a pair $(\sigma^{eq}, \hat{\theta}^{eq})$ with the following properties: First, given that $\hat{\theta} = \hat{\theta}^{eq}$, σ^{eq} is a best response for S-agents, i.e. σ^{eq} is a minimizer of the costs in (3.10). Second, given that $\sigma = \sigma^{eq}$, $\hat{\theta}^{eq}$ is an equilibrium outcome in the coordination game amongst N-agents; formally, there is a solution $\bar{\gamma}^*(\sigma^{eq}, q)$ to the fixed-point equation in (3.9) so that

$$v(q, \hat{\theta}^{eq}) - v(q + \phi_S \, \sigma^{eq}, \hat{\theta}^{eq}) \quad = \quad \frac{1}{\bar{\gamma}^*(\sigma^{eq}, q)} \, .$$

Using the function $(\bar{\gamma}, \sigma, q) \mapsto \hat{\theta}(\bar{\gamma}, \sigma, q)$ introduced above, this condition can also be stated as

$$\hat{\theta}^{eq} \quad = \quad \hat{\theta}(\bar{\gamma}^*(\sigma^{eq}, q), \sigma^{eq}, q) \, .$$

An equilibrium can also be described graphically as the intersection of two functions (see Figure 3.4). Here, the "best-response function" of N-agents is $\sigma \mapsto \hat{\theta}(\bar{\gamma}^*(\sigma, q), \sigma, q)$ The arguments above imply that this function is downward sloping since

$$\frac{\hat{\theta}(\cdot)}{d\sigma} \quad = \quad \hat{\theta}_{\bar{\gamma}}(\cdot) \, \bar{\gamma}^*_{\sigma}(\cdot) \; + \; \hat{\theta}_{\sigma}(\cdot) \quad < \quad 0 \, .$$

The S-agents' best-response function is given by the function $\hat{\theta} \mapsto \sigma^*(q, \hat{\theta}, q)$ which solves (3.11) and is upward sloping. A Nash equilibrium is a pair $(\sigma^{eq}, \hat{\theta}^{eq})$ at which these functions intersect.

How does the Nash equilibrium change when taxes go up? When the tax t, or, equivalently, the price q increases, both best response functions shift. Specifically, the N-agents' best response function shifts upwards since

$$\frac{\hat{\theta}(\cdot)}{dq} \quad = \quad \hat{\theta}_{\bar{\gamma}}(\cdot) \, \bar{\gamma}^*_{q}(\cdot) \; + \; \hat{\theta}_{q}(\cdot) \quad > \quad 0 \, .$$

S-agents' best-response function shifts downwards: for given $\hat{\theta}$, the optimal σ is lower when taxes are higher. Consequently, the function $t \mapsto \hat{\theta}^{eq}(t)$ is increasing. The sign of $t \mapsto \hat{\sigma}^{eq}(t)$ is not a priori clear. There are two countervailing effects: If $\hat{\theta}$ moves up, more people can be reached and this increases the S-agents' incentives to exert moral pressure. On the other hand, if taxes go up, people consume less and this reduces the need for moral pressure. With an additional assumption, iso-elastic consumption utility, we can determine which effect dominates (see Figure 3.4 for an illustration).

Lemma 3.3. *Let* $u(x, \theta) = \theta^{\frac{1}{\varepsilon}} \, \frac{1}{1 - \frac{1}{\varepsilon}} \, x^{1 - \frac{1}{\varepsilon}}$ *, for* $\varepsilon > 1$*. Consider a Nash equilibrium* $(\sigma^{eq}, \hat{\theta}^{eq}))$ *with* $0 < \sigma^{eq} < 1$*. Then,* $t \mapsto \hat{\sigma}^{eq}(t)$ *is increasing.*

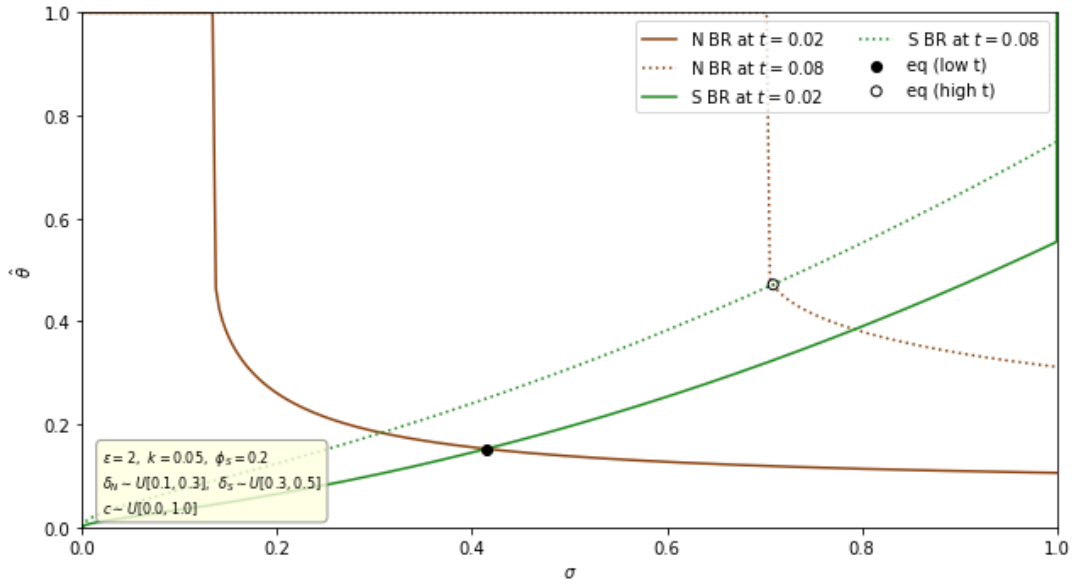


Figure 3.4: How Taxes Affect Moral Pressure and Moral Rebellion in Equilibrium

Notes: The figure illustrates Lemma 3.3 by depicting the best-response functions of N-agents (brown) and S-agents (green); their intersection is a Nash equilibrium. Solid lines correspond to a low tax rate ($t = 0.02$) and dotted lines to a high tax rate ($t = 0.08$). The filled marker is the Nash equilibrium at the low tax rate, the open marker that at the high tax rate. Parameters: $\varepsilon = 2$, $k = 0.05$, $\varphi_S = 0.20$; $\theta \sim$ Uniform$[0, 1]$ in both groups; $\delta_N \sim$ Uniform$[0.1, 0.3]$, $\delta_S \sim$ Uniform$[0.3, 0.5]$.

Implications for tax preferences. The possibility that $t \mapsto \sigma^{eq}(t)$ may be increasing contrasts with the results of Section 3.5. When there is no rebellion this function is decreasing by Lemma 3.1. This has implications for preferences over taxes. With rebellion, N-agents may revise their ideal taxes downwards, rather than upwards.

Proposition 3.6. *Let* $u(x, \theta) = \theta^{\frac{1}{\varepsilon}} \frac{1}{1-\frac{1}{\varepsilon}} x^{1-\frac{1}{\varepsilon}}$, *for* $\varepsilon > 1$. *Consider an* N*-agent of type* (θ, δ) *and an ideal tax of* $t^*(\theta, \delta) \simeq \delta$ *in the basic model with* $\sigma(t) = 0$, *for all* t. *Denote by* $t_N^*(\theta, \delta)$ *this agents' ideal tax in an equilibrium without rebellion. Suppose that there is an equilibrium with rebellion so that* $\theta < \hat{\theta}^{eq}$. *Denote by* $t_R^*(\theta, \delta)$ *this agent's ideal tax in this equilibrium. There is* $\hat{\phi}_S$ *so that* $\phi_S < \hat{\phi}_S$ *implies*

$$t_N^*(\theta, \delta) \quad > \quad t^*(\theta, \delta) \quad > \quad t_R^*(\theta, \delta)\ .$$

The premise of Proposition 3.6 applies for an interesting special case of the general setup: When the share of S-agents is small, and among N-agents, $\theta/\bar{\theta}_N$ is sufficiently close to 1 for all θ, (see equation (3.4)) then the decisive voter is an N-agent with $t^*(\theta, \delta)$ close to δ in the basic model. According to the Proposition, this decisive voter wants higher taxes in response to moral pressure, provided that the moral pressure is universally accepted. By contrast, the decisive voter wants lower taxes when the moral pressure is challenged by people with large carbon footprints.

Numerical illustration: preferred taxes. We plot voters' tax preferences and equilibrium payoffs by each agent type and their respective consumption importance θ and climate

concern δ. For ease of presentation, we begin by fixing $\delta_S = 0.4$ and $\delta_N = 0.2$ to be uniform within group but relax this assumption later on. Figure 3.5 illustrates how preferred tax rates change across the three regimes considered: the basic model, moral pressure without rebellion, and moral pressure with rebellion.[31] For N-agents, the preferred tax $t^*(\theta)$ is a decreasing function of θ, as agents with a larger taste for the dirty good prefer lower taxes. In the basic model (blue), the median voter at $\theta_M = 0.444$ chooses a tax of 0.213, above the common damage parameter $\delta_N = 0.20$. This is due to their lower consumption, relative to the average.

The rebellion regime (green) exhibits three distinct regions. For low θ (up to about 0.34), agents have a weak enough taste for the dirty good and therefore find the costs of rebellion unappealing. These agents prefer taxes high enough that no interior rebellion equilibrium can be sustained—in this calibration, taxes above 0.193 eliminate rebellion entirely, consistent with Lemma 3.1. Since the unique continuation is $\bar{\gamma} = 0$, the rebellion and moral-pressure regimes deliver the same preferred tax, and the green and orange curves coincide. At the discontinuous drop, agents with a somewhat stronger taste for the dirty good prefer to lower the tax to the point where rebellion can just be sustained. They are willing to accept a lower tax because the presence of rebellion reduces the moral pressure they face. The gain from weaker moral pressure outweighs the cost of a lower corrective tax. Their preferred tax is thus pinned at the upper boundary of the range that supports an interior rebellion equilibrium. For still higher θ, agents are sufficiently heavy consumers that their preferred tax is dictated entirely by their type θ and is no longer constrained from above by the requirement that rebellion be sustained. The preferred tax curve then declines smoothly, well below the basic-model level, consistent with Proposition 3.6.

Under moral pressure without rebellion (orange), the pattern is reversed. Agents with an intermediate taste for the dirty good prefer a high tax that reduces the intensity of moral pressure. Moving rightward to agents with a stronger taste, the preferred tax drops discontinuously to a low level that entails high moral pressure. These higher-θ agents accept the greater moral pressure because the tax savings outweigh the cost of social disapproval. Conversely, the lower-θ agents to the left of the discontinuity accept a higher tax because, in the absence of rebellion, moral pressure decreases with the tax rate (Lemma 3.1), so the cost relief from lower moral pressure compensates for the higher tax burden.

The median voter's preferred tax rises from 0.213 in the basic model to 0.334 under moral pressure without rebellion, consistent with Proposition 3.5, but falls to only 0.193 under moral pressure with rebellion, below even the basic-model level, consistent with Proposition 3.6. The S-agents' preferred taxes, shown as horizontal dotted lines, are constant in θ: as group rule-utilitarians, all S-agents maximize the same objective $W_S(t)$. When they can exert moral

[31] In order to allow for rebellion in the calibration of Figures 3.5 and 3.6, we introduce two cost-scaling parameters, α and β. The parameter α rescales the S-agents' cost of exerting moral pressure, so that the cost function becomes $\frac{\alpha}{2}\sigma^2$ rather than $\frac{1}{2}\sigma^2$; the interior first-order condition then reads $\alpha\,\sigma^* = -\delta_S\phi_N\phi_S\int_0^{\hat{\theta}} x_q^*(q+\phi_S\sigma^*,\theta)\,f_N(\theta)\,d\theta$, so that $\alpha < 1$ makes activism cheaper and raises equilibrium σ^*. The parameter β rescales the non-monetary cost of rebellion borne by N-agents, replacing $1/\bar{\gamma}$ with $\beta/\bar{\gamma}$ in the rebellion payoff. This shifts the existence threshold for interior rebellion equilibria from $K(q,\sigma) \geq 4$ to $K(q,\sigma) \geq 4\beta$, with interior roots $\bar{\gamma} = \frac{1}{2}\big(1 \pm \sqrt{1-4\beta/K}\big)$; lower β thus makes rebellion easier to sustain and enlarges the rebellious region.

pressure, they prefer a slightly higher tax (0.433) than in the basic model (0.400).

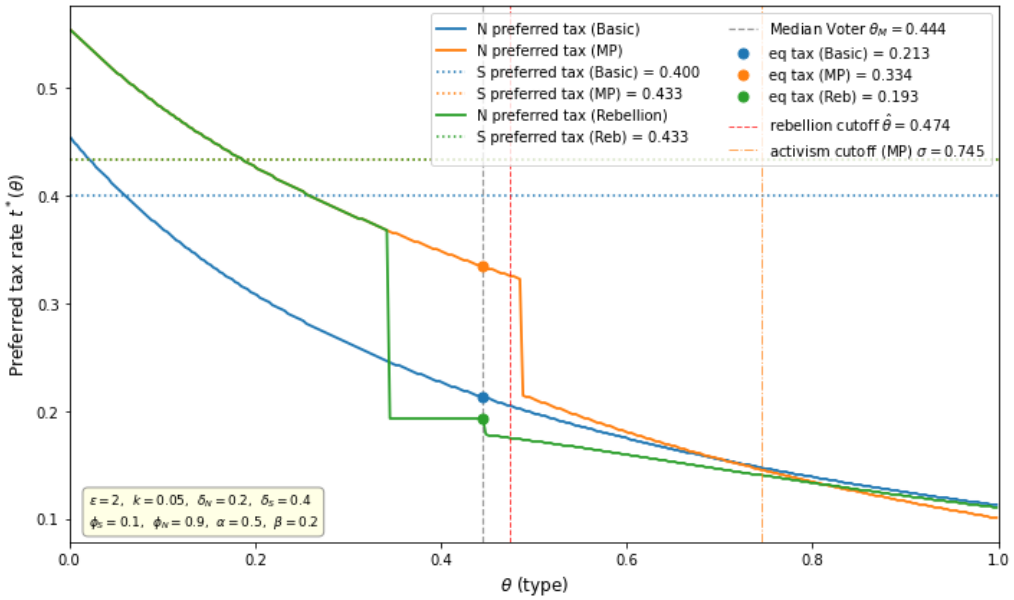


Figure 3.5: Preferred Tax Rates by Agent Type across Regimes

Notes: Preferred tax rate $t^*(\theta)$ as a function of the consumption-taste parameter θ for N-agents (solid lines) and S-agents (dotted horizontal lines) across three regimes: basic model (blue), moral pressure without rebellion (orange), and moral pressure with rebellion (green). Filled markers indicate the political-equilibrium tax in each regime, determined by the median voter $\theta_M = 0.444$ (dashed grey line). The dashed red line marks the rebellion cutoff $\hat{\theta} = 0.474$; the dashed orange line marks the activism cutoff with $c \sim \text{Uniform}[0, 1]$ on the same scale. S-agents' preferred taxes are constant in θ because they are group rule-utilitarians who all maximize the same objective $W_S(t)$. Parameters: $\varepsilon = 2$, $k = 0.05$, $\varphi_S = 0.1$, $\delta_S = 0.40$, $\delta_N = 0.20$, $\alpha = 0.5$, $\beta = 0.2$; $\theta \sim \text{Uniform}[0, 1]$ in both groups; activism costs $c \sim \text{Uniform}[0, 1]$.

Figures 3.5 and 3.6 and the comparison in Proposition 3.6 presume that an equilibrium with moral rebellion coexists with the equilibrium without rebellion. It is however important to note that an interior rebellion equilibrium exists only when the gain from rebelling is sufficiently large relative to its cost, i.e. when $K(q, \sigma^{eq}) \geq 4\beta$. Whether this condition holds depends on the underlying parameters, in particular on the damage parameters, the elasticity of demand, and the cost of rebellion. The illustrations in Figures 3.5 and 3.6 fix parameters inside the coexistence region.

Equilibrium payoffs. Figure 3.6 shows the corresponding equilibrium payoffs. The equilibrium with rebellion may be Pareto-dominated by the one in the basic model: once moral dispute becomes the norm, all types can end up worse off than in an equilibrium in which morality plays no role at all.

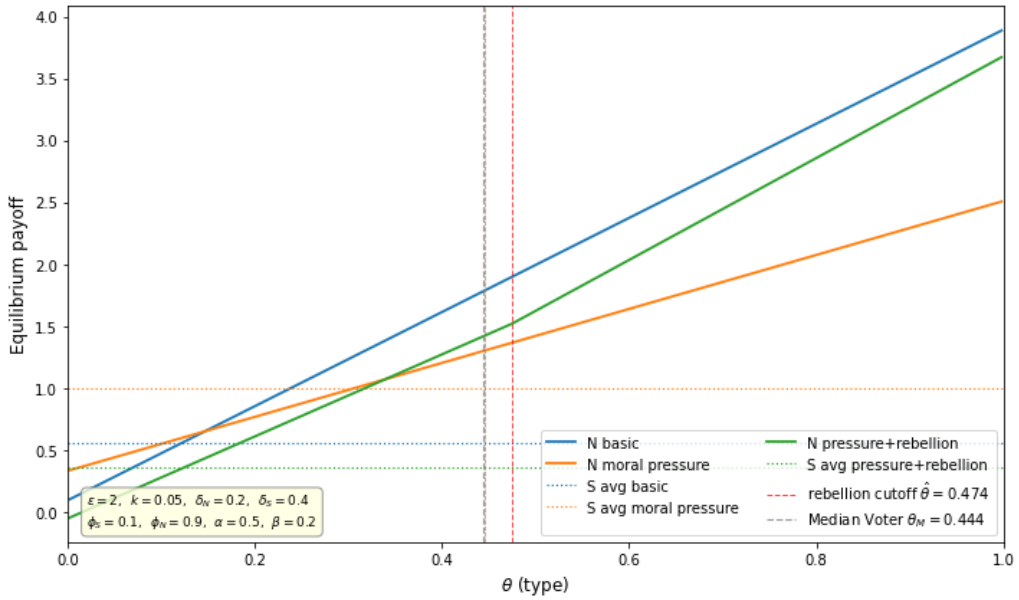


Figure 3.6: Political-Equilibrium Payoffs

Notes: Moral pressure and rebellion reshape N-agents' equilibrium payoffs relative to the basic model. Solid lines plot the equilibrium payoff $U(\theta)$ for N-agents; dashed horizontal lines plot the average S-agent payoff. Three regimes: basic model (blue), moral pressure only (orange), and moral pressure with rebellion (green). The vertical dashed grey line marks the median voter θ_M; the red vertical line marks the rebellion cutoff $\hat{\theta}$. Parameters: $\varepsilon = 2$, $k = 0.05$, $\varphi_S = 0.10$, $\delta_S = 0.40$, $\delta_N = 0.20$, $\alpha = 0.5$, $\beta = 0.2$; $\theta \sim \text{Uniform}[0, 1]$ in both groups.

The baseline equilibrium is based on a setting where people only consume and vote. Figure 3.6 shows how payoffs change when moral activism enters the scene. The baseline equilibrium and the equilibrium with moral pressure (and no rebellion) cannot be Pareto-ranked. In going from the former to the latter, S-agents and N-agents with small carbon footprints win, while all other N-agents lose. Neither is there a Pareto-ranking for the two equilibria with moral pressure: When rebellion emerges, N-agents with large carbon footprints are better off, while everyone else is worse off. The sequence, however, leads to an outcome that is Pareto-dominated by the initial situation. All moral activists and also all other citizens would be better off if all they could do was to consume and to vote. Broadly, the following logic is at play: S-agents exert moral pressure, this is costly for them. Then, N-agents neutralize this pressure. Again, this is costly. The consequence is a departure from the status quo that is too small to justify the costs of this back and forth.

Two-dimensional heterogeneity. The numerical examples above (Figures 3.5 and 3.6) hold the damage parameters fixed at δ_N and δ_S within the N-group and S-group and trace preferred taxes and payoffs along the consumption-taste dimension θ only. We now allow for a non-degenerate joint distribution of (θ, δ) within each group. The preferred tax of an N-agent is decreasing in θ and increasing in δ, and the median voter is the agent with the median position in the resulting two-dimensional ordering of ideal taxes (see Section 3.4). We replicate the three-regime comparison under such two-dimensional heterogeneity. The qualitative pattern is unchanged. The political equilibrium tax rises from the basic model to the moral-pressure

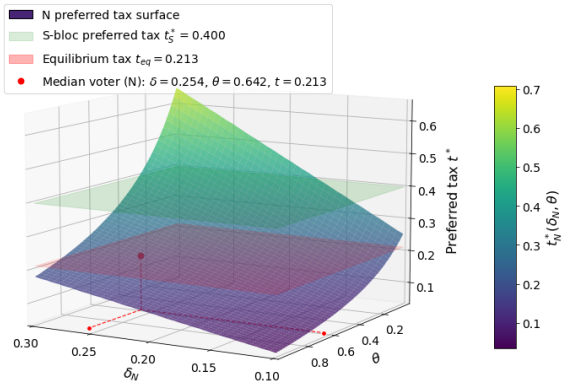
regime and falls back below the basic-model level once rebellion is sustained.

We let $\delta_N \sim \text{Uniform}[0.1, 0.3]$ and $\delta_S \sim \text{Uniform}[0.3, 0.5]$, with $\theta \sim \text{Uniform}[0, 1]$ in both groups and (θ, δ) independent within each group. Other parameters are as before ($\varepsilon = 2$, $k = 0.05$, $\varphi_S = 0.1$, $\alpha = 0.5$, $\beta = 0.2$).

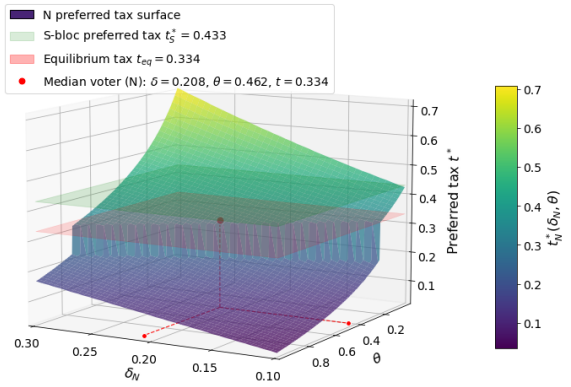
Preferred tax surfaces (Figure 3.7a–3.7c). Each of the three panels plots the N-agents' preferred tax $t_N^*(\delta_N, \theta)$ as a (colored) surface over the (δ_N, θ)-plane, the constant S-bloc preferred tax t_S^* as a green plane, and the political equilibrium tax (the median voter's ideal tax in the joint N–S population) as a red plane. The red dot marks the median voter's parameters (δ_M, θ_M) together with her preferred tax. Across the three regimes, the figures show the same qualitative pattern as in the one-dimensional numerical exercise above. Under moral pressure without rebellion, both the N-surface and the equilibrium plane shift upward relative to the basic model. Once a rebellion action is added, the N-surface kinks at the rebellion cutoff $\hat{\theta}^{eq}$, where N-agents with $\theta_N < \hat{\theta}^{eq}$ prefer no rebellion and thus higher taxes, while those with $\theta_N > \hat{\theta}^{eq}$ prefer rebellion and therefore lower taxes. In this illustration the median voter prefers rebellion and the political equilibrium tax falls below the basic-model level.

Equilibrium payoffs (Figure 3.7d). Figure 3.7d shows the analogous extension of Figure 3.6, where N-agents' equilibrium payoffs are plotted as surfaces over (δ, θ), evaluated at the median voter's preferred tax in each regime, alongside the average S-payoff in each regime as a horizontal plane. The comparison across regimes mirrors Figure 3.6. The moral-pressure surface lies above the basic-model surface for small θ but below it for high-θ types, while the rebellion surface is below the basic-model surface for the bulk of types. We recover the Pareto-inferior outcome in the rebellion equilibrium highlighted above.

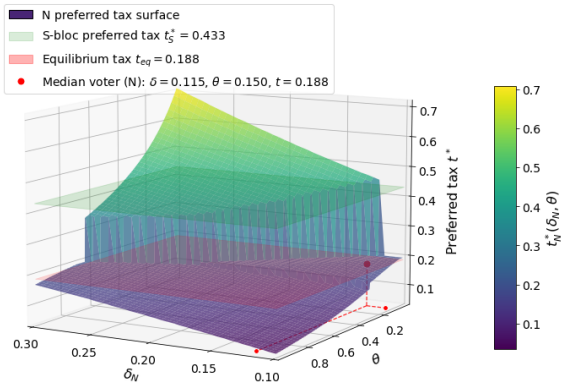
The comparative statics of moral pressure and moral rebellion that drive the main propositions and the one-dimensional numerical illustration carry over.



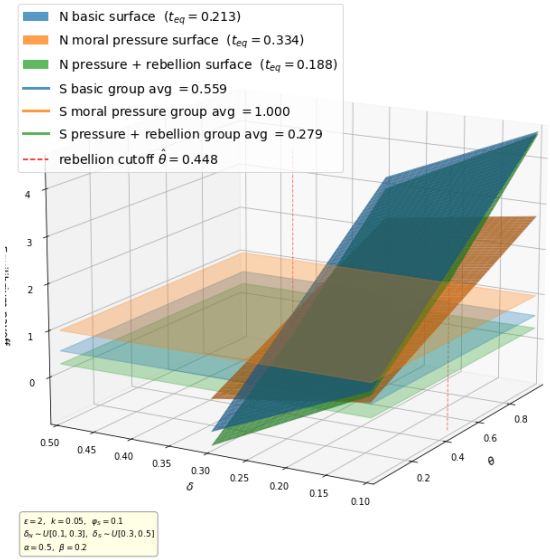
(a) Preferred Taxes: Basic Model



(b) Preferred Taxes: Moral Pressure



(c) Preferred Taxes: Moral Rebellion



(d) Political-Equilibrium Payoffs

Figure 3.7: Two-Dimensional Heterogeneity: Preferred Taxes and Equilibrium Payoffs

Notes: Panels (a)–(c) show N-agents' preferred tax $t_N^*(\delta_N, \theta)$ (colored surface), the S-bloc's preferred tax t_S^* (green plane), and the political-equilibrium tax t_{eq} (red plane) across the three regimes; t_{eq} equals the preferred tax of the median voter in the joint N–S population (red dot). Panel (d) plots equilibrium N-payoffs (colored surfaces) and average S-payoffs (horizontal planes); the vertical red dashed line marks $\hat{\theta}^{eq}$. Panel (d) is the two-dimensional analogue of Figure 3.6. Parameters: $\varepsilon = 2$, $k = 0.05$, $\varphi_S = 0.1$, $\alpha = 0.5$, $\beta = 0.2$; $\theta \sim \text{Uniform}[0, 1]$, $\delta_N \sim \text{Uniform}[0.1, 0.3]$, $\delta_S \sim \text{Uniform}[0.3, 0.5]$, with θ and δ drawn independently within each group.

3.7 Survey

Two samples of respondents, for the US and the UK, were collected through Prolific, using a representative sampling approach with emphasis on political affiliation. The survey instrument collects demographics, stated behavioral responses to price and damage changes, preferences over carbon taxes, and attitudes towards moral pressure and social responsibility. Respondents answer simple comprehension questions on the social cost of carbon and the mechanics of the carbon tax. Those who fail these checks are excluded. We also drop respondents who indicate at

the end of the survey that their answers should not be used for research. After these filters, the final samples comprise 901 respondents for the US and 894 for the UK. Our survey instructions are reported in Section 6.1.12 of the Appendix.

When reporting our survey results, we report but do not extensively discuss magnitudes and instead focus on the signed correlations. This is both for compactness of the exposition and due to the fact that our survey results are based on hypothetical behavior. For comparability, all tables use untransformed dependent variables and z-standardized independent variables. Our coefficients of interest are therefore interpreted as the change in the raw dependent variable associated with a one-standard-deviation increase in the independent variable.

3.7.1 Primitives

Our primary way to identify socially responsible behavior in the domain of climate change is the *Damage Response*: the self-reported adjustment of car usage in response to new information about the social cost of carbon, holding prices fixed.[32] In Table 3.1 we check how the model's consumption and damage primitives correlate with this measure. Our hypothesis is that higher climate concern (δ) is associated with a stronger (more negative) damage response and that higher consumption importance (θ) is associated with a weaker response.[33]

Columns (2) and (5) replaces these primitives with the *Price vs. Harm* measure, which directly elicits how respondents weigh price against harm when making consumption decisions.[34] Unlike *Damage Response*, which is specific to car usage and carbon emissions, *Price vs. Harm* is domain-general. It asks respondents how they trade off affordability against the harm that their consumption causes to others, including but not limited to environmental damages. Both variables proxy for socially responsible orientation and we expect *Price vs. Harm* to correlate negatively with *Damage Response*.

While not our primary focus about socially responsible behavior, columns (3) and (6) include the *Oil vs Tax Gap*, the difference in stated car-use reduction between a carbon-tax-induced and an oil-price-induced gasoline price increase.[35] A neoclassical consumer responds only to prices and should not distinguish between the two sources (see equation (3.1)). A stronger oil price response relative to the tax response is indicative of S-type behavior (see equation (3.3)).

The results in Table 3.1 confirm these hypotheses in both countries. Climate Concern (δ) is strongly negatively associated with *Damage Response*, while Consumption Importance (θ) is

[32] *Damage Response* is the respondent's self-reported percentage change in car usage (slider from -100 to $+100$) in response to: "Suppose scientists now estimate that the social cost of carbon from driving gasoline-powered cars has increased by 30%, while prices remain the same. How would you adjust the use of your car?"

[33] *Climate Concern* (δ): "How serious of a problem do you think climate change is?" (1 = Not a problem at all, 5 = Extremely serious). *Consumption Importance* (θ): "How important are carbon-intensive goods and activities (such as driving, air travel, or meat consumption) to your lifestyle?" (1 = Not at all important, 5 = Essential).

[34] *Price vs. Harm*: "When you consider buying something, how do you weigh price/value (0) against whether your consumption might cause harm to others (10)?"

[35] *Oil vs Tax Gap*: difference between stated car-use reduction for a 30% gasoline price increase due to a carbon tax and one due to higher oil extraction costs (both on -100 to $+100$ sliders, so positive numbers mean a larger reduction due to an oil price price increase).

Table 3.1: Foundations — Correlations with Socially Responsible Behavior

	Dependent variable: Damage Response					
	UK	UK	UK	US	US	US
	(1)	(2)	(3)	(4)	(5)	(6)
Climate Concern (δ)	-5.806^{***}		-5.793^{***}	-8.449^{***}		-8.478^{***}
	(0.777)		(0.784)	(0.876)		(0.883)
Consumption Importance (θ)	2.313^{***}		2.328^{***}	1.557^{*}		1.518^{*}
	(0.811)		(0.809)	(0.876)		(0.879)
Price vs. Harm		-5.330^{***}			-4.066^{***}	
		(1.015)			(0.972)	
Oil vs Tax Gap			-0.934			-2.152^{**}
			(1.305)			(0.892)
Constant	-16.742^{***}	-16.647^{***}	-16.742^{***}	-16.633^{***}	-16.644^{***}	-16.633^{***}
	(0.820)	(0.827)	(0.820)	(0.805)	(0.842)	(0.802)
Mean of DV	-16.740	-16.647	-16.740	-16.633	-16.644	-16.633
Observations	893	890	893	900	900	900
R^2	0.060	0.045	0.062	0.110	0.025	0.117
Adjusted R^2	0.058	0.044	0.058	0.108	0.024	0.114

Notes: Robust (HC1) standard errors in parentheses. All independent variables are standardized (z-scored) so coefficients represent the change in the dependent variable per one-standard-deviation increase in the regressor. Mean of DV gives the mean of the dependent variable. The mean can differ across columns due to the different samples, and few respondents not answering all relevant questions. *p<0.1; **p<0.05; ***p<0.01.

positively associated with it. *Price vs. Harm* enters with the expected negative sign, and the *Oil vs. Tax Gap* is significant in the US but not in the UK.

3.7.2 Tax Preferences

We elicit respondents' *Preferred Carbon Tax* by asking, after a brief explainer of a revenue-rebated carbon tax, how high they would like the carbon tax to be in their country, with an additional display of the implied increase in gasoline prices.[36]

Our theory states that the N-agent's preferred tax is decreasing in θ ($U_{t\theta} < 0$) and increasing in δ ($U_{t\delta} > 0$) (see section 3.3.1). Moreover, *Damage Response*, as a proxy for socially responsible orientation, should be associated with higher preferred taxes (negative coefficient, since *Damage Response* itself is negative for S-types).[37] Table 3.2 presents five specifications per country: the primitives θ and δ alone; *Damage Response* alone; both together; *Price vs. Harm* alone; and θ, δ together with *Price vs. Harm*.

Results support our hypotheses. Consumption importance θ and climate concern δ predict tax preferences with the expected signs. Moreover, the proxies for socially responsible behavior

[36] *Preferred Carbon Tax* is the respondent's self-reported tax rate (slider in £/ton in the UK and $/ton in the US) in response to: How high would you like the carbon tax to be in your country? (At the bottom we show you the estimated increase in gasoline prices.)" Respondents are first reminded that under the proposed carbon tax, everyone pays more for high-emission goods, but all tax revenue is returned equally to the public. Low emitters get back more than they paid."

[37] See section 3.3.1 for our discussion on N-agent's preferred tax.

(*Damage Response* and *Price vs. Harm*) are robustly correlated with a higher preferred carbon tax. In the UK, the association between social responsibility and higher carbon tax preferences persists even after controlling for θ and δ (Columns (3) and (5)). This suggests that socially responsible individuals prefer higher taxes above and beyond what is explained by their lower consumption preferences and higher damage assessments alone.

Table 3.2: Tax Preferences — Level

	Dependent variable: Preferred Carbon Tax									
	UK	UK	UK	UK	UK	US	US	US	US	US
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Climate Concern (δ)	15.832***		14.881***		11.780***	15.788***		15.813***		11.384***
	(1.393)		(1.413)		(1.482)	(1.177)		(1.330)		(1.307)
Consumption Importance (θ)	−4.254***		−3.875***		−3.768***	−0.908		−0.913		−0.717
	(1.450)		(1.451)		(1.407)	(1.263)		(1.258)		(1.215)
Damage Response		−7.844***	−4.139***				−5.117***	0.075		
		(1.637)	(1.572)				(1.847)	(1.797)		
Price vs. Harm				14.654***	10.103***				15.241***	10.369***
				(1.484)	(1.553)				(1.321)	(1.464)
Constant	53.574***	53.589***	53.571***	53.527***	53.498***	40.221***	40.271***	40.221***	40.289***	40.262***
	(1.426)	(1.502)	(1.421)	(1.440)	(1.390)	(1.270)	(1.362)	(1.270)	(1.277)	(1.233)
Mean of DV	53.570	53.589	53.570	53.527	53.507	40.221	40.271	40.221	40.289	40.239
Observations	893	894	893	890	889	900	901	900	900	899
R^2	0.128	0.030	0.136	0.104	0.168	0.146	0.015	0.146	0.137	0.198
Adjusted R^2	0.126	0.029	0.133	0.103	0.165	0.144	0.014	0.144	0.136	0.196

Notes: Robust (HC1) standard errors in parentheses. All independent variables are standardized (z-scored). *p<0.1; **p<0.05; ***p<0.01.

Additional results on how respondents adjust their preferred tax in response to hypothetical local and global damage increases are reported in Appendix Tables 6.1 and 6.2. The patterns there are broadly consistent with the findings in Table 3.2: δ is a strong positive predictor and *Price vs. Harm* is significant in all specifications, while θ is weakly negative. One noteworthy difference is that, in the US, *Damage Response* is positively signed (though insignificant) conditional on θ and δ in the tax-adjustment regressions, whereas it is essentially zero in the level regression.

3.7.3 Moral Pressure

We measure attitudes toward moral pressure with two items capturing, respectively, normative endorsement and normative rejection of confronting others about their carbon-intensive consumption (*MP Approval* and *MP Opposition*) together with a behavioral measure of own susceptibility to such pressure, *MP Responsiveness*.[38]

In the model, S-agents exert moral pressure on N-agents to reduce their carbon footprints. Our hypothesis is that individuals with higher δ should approve of moral pressure more, and oppose it less, as reducing others' consumption is more valuable for those highly concerned

[38] *MP Approval* is the respondent's agreement (1,=,Strongly disagree, 5,=,Strongly agree) with: "Applying moral pressure by confronting individuals with the carbon emissions of their consumption choices is justified." *MP Opposition* is the respondent's agreement on the same five-point scale with: "Moral pressure about individuals' carbon-intensive consumption is intrusive and should be avoided." *MP Responsiveness* is the respondent's answer (1,=,Never, 5,=,Very often) to: "How often do you change what you buy or consume to avoid feeling moral pressure about carbon emissions?"

about the externality. Higher θ, in turn, should be associated with less approval and more opposition, since moral pressure raises the effective price of consumption for high-θ individuals (equation (3.5)). Likewise, we hypothesize that rebellion increases in θ (see section 3.6). Our proxies for socially responsible behavior (a more negative *Damage Response*, and larger *Price vs. Harm*) should correlate with greater approval and less opposition to moral pressure.

Table 3.3: Support and Opposition to Moral Pressure — UK

	Dependent variable:							
	MP Approval				MP Opposition			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Climate Concern (δ)			0.519***	0.434***			−0.486***	−0.423***
			(0.031)	(0.033)			(0.035)	(0.036)
Consumption Importance (θ)			−0.082**	−0.084***			0.136***	0.141***
			(0.034)	(0.032)			(0.035)	(0.034)
Damage Response	−0.210***		−0.085***		0.210***		0.089**	
	(0.037)		(0.032)		(0.037)		(0.035)	
Price vs. Harm		0.446***		0.279***		−0.381***		−0.217***
		(0.033)		(0.032)		(0.036)		(0.036)
Constant	3.203***	3.207***	3.203***	3.206***	2.936***	2.934***	2.937***	2.936***
	(0.036)	(0.033)	(0.032)	(0.030)	(0.037)	(0.036)	(0.034)	(0.033)
Mean of DV	3.203	3.205	3.202	3.204	2.936	2.935	2.937	2.936
Observations	892	888	891	887	893	889	892	888
R^2	0.037	0.167	0.256	0.308	0.035	0.113	0.220	0.246
Adjusted R^2	0.036	0.166	0.253	0.306	0.033	0.112	0.217	0.243

Notes: Robust (HC1) standard errors in parentheses. All independent variables are standardized (z-scored). *p<0.1; **p<0.05; ***p<0.01.

Table 3.3 presents results for the UK.[39] Our proxies for socially responsible behavior correlate with greater approval and lower opposition to moral pressure, and climate concern δ and consumption importance θ enter with the expected signs. Consistent with our theory, individuals with a higher damage assessment approve of moral pressure more strongly, while those who place greater value on consumption tend to oppose it.

Corresponding US results are in Table 3.4. The US findings largely corroborate those for the UK. Two differences are worth noting. First, *Damage Response* loses significance for MP Approval in the US once θ and δ are controlled for (column 3), whereas it remains significant in the UK. Second, θ is not significant for MP Opposition in the US, whereas it is in the UK.

[39]Since *MP Approval* and *MP Opposition* are measured on five-point Likert scales, we report OLS estimates in the main text for interpretability and comparability with our other regressions, and show in Appendix 6.1.11 that all results are robust to estimating ordered probit models instead.

Table 3.4: Support and Opposition to Moral Pressure — US

	Dependent variable:							
	MP Approval				MP Opposition			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Damage Response	−0.174***		0.033		0.292***		0.105**	
	(0.045)		(0.038)		(0.044)		(0.045)	
Price vs. Harm		0.454***		0.236***		−0.386***		−0.158***
		(0.038)		(0.041)		(0.039)		(0.040)
Consumption Importance (θ)			−0.077**	−0.072**			0.052	0.057
			(0.034)	(0.032)			(0.036)	(0.036)
Climate Concern (δ)			0.624***	0.512***			−0.565***	−0.532***
			(0.033)	(0.038)			(0.037)	(0.038)
Constant	3.007***	3.006***	3.006***	3.005***	3.111***	3.112***	3.113***	3.113***
	(0.038)	(0.035)	(0.033)	(0.032)	(0.039)	(0.038)	(0.034)	(0.034)
Mean of DV	3.007	3.006	3.006	3.004	3.111	3.112	3.114	3.115
Observations	901	900	900	899	899	898	898	897
R^2	0.023	0.154	0.281	0.314	0.060	0.105	0.259	0.267
Adjusted R^2	0.021	0.153	0.279	0.312	0.059	0.104	0.257	0.264

Notes: Robust (HC1) standard errors in parentheses. All independent variables are standardized (z-scored). All variables are defined as in Tables 3.1 and 3.3. *p<0.1; **p<0.05; ***p<0.01.

Moral Pressure and Tax Preferences

Our theory hypothesizes that individuals who approve of moral pressure should prefer higher carbon taxes, while those who oppose it should prefer lower taxes. Additionally, the model's rebellion mechanism implies that individuals who oppose moral pressure but nonetheless respond to it behaviorally (i.e., do not rebel) should prefer higher taxes than those who oppose and also resist it.

Table 3.5 tests these hypotheses. Columns (1) and (4) include MP Approval and MP Opposition alongside θ and δ. Columns (2)–(3) and (5)–(6) introduce MP Responsiveness and its interaction with MP Opposition to test the rebellion channel.

The results are consistent with the first hypothesis in both countries. MP Approval is positively and significantly associated with preferred tax levels, and MP Opposition enters negatively. Climate concern δ remains a strong positive correlate, and consumption importance θ is significantly negative in the UK but not in the US, consistent with the patterns in Table 3.2.

The rebellion results differ across countries. In the US, the interaction between MP Opposition and MP Responsiveness is positive and significant (columns 5–6), indicating that among those who oppose moral pressure, individuals who nonetheless adjust their behavior are associated with higher tax preferences, consistent with non-rebellious N-types supporting corrective taxation. In the UK, by contrast, the interaction is not significant, suggesting that the rebellion channel is less present in this sample.

Table 3.5: Moral Pressure Attitudes and Tax Preferences

	Dependent variable: Preferred Carbon Tax					
	UK	UK	UK	US	US	US
	(1)	(2)	(3)	(4)	(5)	(6)
Climate Concern (δ)	9.485***		9.647***	9.178***		7.113***
	(1.671)		(1.646)	(1.401)		(1.487)
Consumption Importance (θ)	−3.069**		−2.915**	0.012		0.019
	(1.424)		(1.448)	(1.229)		(1.202)
MP Approval	7.442***			10.911***		
	(2.378)			(2.078)		
MP Opposition	−5.606**	−13.198***	−9.639***	−1.572	−9.030***	−6.572***
	(2.319)	(1.690)	(1.831)	(2.076)	(1.461)	(1.527)
MP Responsiveness		7.565***	5.179***		11.866***	9.475***
		(1.672)	(1.689)		(1.542)	(1.598)
MP Opposition × MP Responsiveness		0.232	−0.843		4.266***	3.238**
		(1.519)	(1.518)		(1.295)	(1.301)
Constant	53.442***	53.681***	53.259***	40.090***	42.044***	41.559***
	(1.379)	(1.540)	(1.516)	(1.215)	(1.442)	(1.420)
Mean of DV	53.410	53.594	53.574	40.061	40.111	40.061
Observations	891	893	892	898	899	898
R^2	0.181	0.148	0.183	0.211	0.208	0.227
Adjusted R^2	0.177	0.145	0.178	0.207	0.206	0.222

Notes: Robust (HC1) standard errors in parentheses. All independent variables are standardized (z-scored); the interaction *MP Opposition × MP Responsiveness* is the product of the two standardized variables. The interaction *MP Opposition × MP Responsiveness* tests whether individuals who oppose moral pressure but are nonetheless behaviorally responsive to it prefer different tax levels. *p<0.1; **p<0.05; ***p<0.01.

3.8 Concluding Remarks

This paper analyzes preferences over carbon taxes und political equilibrium outcomes in an economy consisting of neoclassical consumers and socially responsible consumers.

In a hypothetical country A inhabited only by neoclassical consumers, individuals' choose independently how much of a dirty good to consume. However, the resulting atmospheric externality is taken into account through the political process that determines the corrective tax. The political equilibrium tax is the median voter's preferred tax and reflects his views on damage severity are, as well as a distributive consideration: It is increasing in the gap between the average and the median carbon footprint. A first best Pigouvian tax so that the *"tax rate equals the marginal damage"* is reached in country A only if the average and the median footprint coincide.

In a hypothetical country B that is inhabited only by socially responsible consumers, individuals internalize the social damages due to their consumption of a dirty good so that there is no need for a corrective tax. It is easy to specify stylized environments that result in the same allocation in countries A and B. For instance, if the distribution of carbon footprints is

symmetric in country A, and if all individuals have the same views on what the social costs of carbon are, then country A reaches a first-best outcome via a Pigouvian tax, and country B reaches the same outcome because its inhabitants internalize the social costs of carbon even in the absence of a corrective tax.

Our analysis mainly deals with country C, a country with a mixed population consisting of neoclassical and socially responsible consumers – with focus on settings where socially responsible consumers have a lower willingness to pay for dirty goods and larger estimates of the social costs of carbon. We present extensive survey evidence that this is the empirically plausible case. Our main theoretical results characterize the tensions in country C.[40] In particular, we show that an equilibrium with moral pressure by socially responsible consumers and moral rebellion by a subset of neoclassical consumers, may give rise to a Pareto-dominated equilibrium in which carbon taxes are very low.

While this paper focuses on carbon taxes, we believe that a similar approach could be fruitful for an analysis of political economy forces also in other domains. At a more abstract level, agents in our economy take individual consumption decisions, but also have preferences over economy-wide consumption patterns. For example, individuals may wish that there are fewer imports from countries with no effective ban of child labor, or, they may wish that there is more demand for local products and more business for local dealers. A stereotypical neoclassical consumer articulates such preferences only in her role as a citizen in the political process, while stereotypical socially responsible consumers would not need any public policy to achieve their preferred consumption pattern, if they were alone. Again, heterogeneity will lead to interesting interactions.

Our survey shows there is significant heterogeneity in how socially-responsible-consumer-like individuals are. It also shows that the distinction of the polar cases *neoclassical* and *socially responsible* is useful in that it captures attitudes towards both policy and the consumption habits of others. Individuals who are more-socially-responsible-consumer-like prefer higher corrective taxes and are more supportive of moral pressure on consumers who disregard social costs.

[40] There are alternative specifications of the models' primitives, so that the same allocation results in countries A, B and C. Thus, having a mix of neoclassical and socially responsible consumers is not per se a source of tensions.

Bibliography

Abramowitz, A. I. and S. Webster (2016). "The rise of negative partisanship and the nationalization of US elections in the 21st century". In: *Electoral Studies* 41, pp. 12–22.

Acemoglu, D. (2002). "Directed Technical Change". In: *The Review of Economic Studies* 69.4, pp. 781–809.

— (2007). "Equilibrium Bias of Technology". In: *Econometrica* 75.5, pp. 1371–1409.

Acemoglu, D., P. Aghion, L. Bursztyn, and D. Hemous (2012). "The Environment and Directed Technical Change". In: *American Economic Review* 102.1, pp. 131–166.

Acemoglu, D., U. Akcigit, D. Hanley, and W. Kerr (2016). "Transition to Clean Technology". In: *Journal of Political Economy* 124.1, pp. 52–104.

Acemoglu, D. and J. Linn (2004). "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry*". In: *The Quarterly Journal of Economics* 119.3, pp. 1049–1090.

Ackerman, F., S. J. DeCanio, R. B. Howarth, and K. Sheeran (2009). "Limitations of integrated assessment models of climate change". In: *Climatic Change* 95.3, pp. 297–315.

Agatz, N., Y. Fan, and D. Stam (2021). "The Impact of Green Labels on Time Slot Choice and Operational Sustainability". In: *Production and Operations Management* 30.7, pp. 2285–2303.

Aghion, P., U. Akcigit, A. Hyytinen, and O. Toivanen (2017). *The Social Origins of Inventors*. Tech. rep. w24110. National Bureau of Economic Research.

Aghion, P., R. Bénabou, R. Martin, and A. Roulet (2023). "Environmental Preferences and Technological Choices: Is Market Competition Clean or Dirty?" In: *American Economic Review: Insights* 5.1, pp. 1–20.

Aghion, P., A. Bergeaud, M. Lequien, and M. J. Melitz (2024). "The Heterogeneous Impact of Market Size on Innovation: Evidence from French Firm-Level Exports". In: *The Review of Economics and Statistics* 106.3, pp. 608–626.

Aghion, P., N. Bloom, R. Blundell, R. Griffith, and P. Howitt (2005). "Competition and Innovation: An Inverted-U Relationship". In: *The Quarterly Journal of Economics* 120.2, pp. 701–728.

Aghion, P., A. Dechezleprêtre, D. Hémous, R. Martin, and J. Van Reenen (2016). "Carbon Taxes, Path Dependency, and Directed Technical Change: Evidence from the Auto Industry". In: *Journal of Political Economy* 124.1, pp. 1–51.

Aghion, P. and P. Howitt (1992). "A Model of Growth Through Creative Destruction". In: *Econometrica* 60.2, pp. 323–351.

Akcigit, U., J. Grigsby, and T. Nicholas (2017). *The Rise of American Ingenuity: Innovation and Inventors of the Golden Age*. Tech. rep. w23047. National Bureau of Economic Research.

Alesina, A. and G.-M. Angeletos (2005). "Fairness and Redistribution". In: *American Economic Review* 95.4, pp. 960–980.

Alger, I. (2026). "Norms and norm change - driven by social Kantian preferences". In: *Toulouse School of Economics Working Papers* 1605.

Alger, I. and J. W. Weibull (2013). "Homo Moralis—Preference Evolution Under Incomplete Information and Assortative Matching". In: *Econometrica* 81.6, pp. 2269–2302.

Andre, P., T. Boneva, F. Chopra, and A. Falk (2024). "Globally representative evidence on the actual and perceived support for climate action". In: *Nature Climate Change* 14.3, pp. 253–259.

Angeletos, G.-M. and Z. Huo (2021). "Myopia and Anchoring". In: *American Economic Review* 111.4, pp. 1166–1200.

Atkinson, A. and J. Stiglitz (1976). "The design of tax structure: Direct versus indirect taxation". In: *Journal of Public Economics* 6.1-2, pp. 55–75.

Azoulay, P., J. S. Graff Zivin, D. Li, and B. N. Sampat (2019). "Public R&D Investments and Private-sector Patenting: Evidence from NIH Funding Rules". In: *The Review of Economic Studies* 86.1, pp. 117–152.

Banerjee, S., R. Kaniel, and I. Kremer (2009). "Price Drift as an Outcome of Differences in Higher-Order Beliefs". In: *The Review of Financial Studies* 22.9.

Barrage, L. (2020). "Optimal Dynamic Carbon Taxes in a Climate–Economy Model with Distortionary Fiscal Policy". In: *The Review of Economic Studies* 87.1, pp. 1–39.

Barreca, A., K. Clay, O. Deschenes, M. Greenstone, and J. S. Shapiro (2016). "Adapting to Climate Change: The Remarkable Decline in the US Temperature-Mortality Relationship over the Twentieth Century". In: *Journal of Political Economy* 124.1, pp. 105–159.

Bastone, K. (2018). *Behind the Curtain at BioLite*. `https://www.outsideonline.com/business-journal/brands/biolite-behind-curtain/`.

Bell, A., R. Chetty, X. Jaravel, N. Petkova, and J. Van Reenen (2019). "Who Becomes an Inventor in America? The Importance of Exposure to Innovation*". In: *The Quarterly Journal of Economics* 134.2, pp. 647–713.

Bénabou, R. and J. Tirole (2006). "Belief in a just world and redistributive politics". In: *Quarterly Journal of Economics* 121.2, pp. 699–746.

Berliant, M. and P. C. Boyer (2024). "Politics and income taxes: Progress and progressivity". In: *Journal of Public Economic Theory* 26.4, e12704.

Bierbrauer, F., A. Tsyvinski, and N. Werquin (2022). "Taxes and Turnout: When the Decisive Voter Stays at Home". In: *American Economic Review* 112.2, pp. 689–719.

Bierbrauer, F. and N. Netzer (2016). "Mechanism design and intentions". In: *Journal of Economic Theory* 163, pp. 557–603.

Bierbrauer, F. J. (2024). *A Uniform Price for Carbon: Optimal or Not?* Working Paper. University of Cologne.

Bierbrauer, F. J. and P. C. Boyer (2014). "The Pareto-Frontier in a Simple Mirrleesian Model of Income Taxation". In: *Annals of Economics and Statistics* 113/114, pp. 185–206.

Bierbrauer, F. J., P. C. Boyer, and E. Hansen (2023). "Pareto-improving tax reforms and the Earned Income Tax Credit". In: *Econometrica* 91.3, pp. 1077–1103.

Black, D. (1948). "On the Rationale of Group Decision-making". In: *Journal of Political Economy* 56.1, pp. 23–34.

Bloom, N., M. Schankerman, and J. Van Reenen (2013). "Identifying Technology Spillovers and Product Market Rivalry". In: *Econometrica* 81.4, pp. 1347–1393.

Bloom, N. (2007). "Uncertainty and the Dynamics of R&D". In: *American Economic Review* 97.2, pp. 250–255.

Bolton, G. E. and A. Ockenfels (2000). "ERC: A Theory of Equity, Reciprocity, and Competition". In: *The American Economic Review* 90.1, pp. 166–193.

Bontadini, F. and F. Vona (2023). "Anatomy of Green Specialisation: Evidence from EU Production Data, 1995–2015". In: *Environmental and Resource Economics* 85.3, pp. 707–740.

Bordalo, P., N. Gennaioli, and A. Shleifer (2012). "Salience Theory of Choice Under Risk". In: *The Quarterly Journal of Economics* 127.3, pp. 1243–1285.

— (2022). "Salience". In: *Annual Review of Economics* 14.1, pp. 521–544.

Boustan, L. P., M. E. Kahn, P. W. Rhode, and M. L. Yanguas (2020). "The effect of natural disasters on economic activity in US counties: A century of data". In: *Journal of Urban Economics* 118, p. 103257.

Bovenberg, A. L. and L. H. Goulder (1996). "Optimal Environmental Taxation in the Presence of Other Taxes: General- Equilibrium Analyses". In: *The American Economic Review* 86.4, pp. 985–1000.

Brekke, K. A., S. Kverndokk, and K. Nyborg (2003). "A n economic model of moral motivation". In: *Journal of Public Economics* 87, pp. 1967–1983.

Cai, Y. (2020). *The Role of Uncertainty in Controlling Climate Change.*

Calel, R. and A. Dechezleprêtre (2016). "Environmental Policy and Directed Technological Change: Evidence from the European Carbon Market". In: *Review of Economics and Statistics* 98.1, pp. 173–191.

Callander, S. and J. C. Carbajal (2022). "Cause and effect in political polarization: A dynamic analysis". In: *Journal of Political Economy* 130.4, pp. 825–880.

Capraro, V. and D. G. Rand (2018). "Do the Right Thing: Experimental evidence that preferences for moral behavior, rather than equity or efficiency per se, drive human prosociality". In: *Judgement and Decision Making* 13.1, pp. 99–111.

CBS News (2006). *The Brain Behind Netflix*. `https://www.cbsnews.com/news/the-brain-behind-netflix/`.

Chae, R. L., R. Siddiqui, and Y. Xu (2025). "Vulnerability to natural disasters and sustainable consumption: Unraveling political and regional differences". In: *Proceedings of the National Academy of Sciences* 122.5, e2409851122.

Chaisemartin, C. de and X. D'Haultfœuille (2024a). *Difference-in-Differences Estimators of Intertemporal Treatment Effects*. SSRN Scholarly Paper. Rochester, NY.

Chaisemartin, C. de and X. D'Haultfœuille (2023). "Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: a survey". In: *The Econometrics Journal* 26.3, pp. C1–C30.

Charness, G. and M. Rabin (2002). "Understanding Social Preferences with Simple Tests". In: *The Quarterly Journal of Economics* 117.3, pp. 817–869.

Coate, S. and M. Conlin (2004). "A Group Rule-Utilitarian Approach to Voter Turnout: Theory and Evidence". In: *The American Economic Review* 94.5, pp. 1476–1504.

Coibion, O., Y. Gorodnichenko, S. Kumar, and J. Ryngaert (2021). "Do You Know that I Know that You Know...? Higher-Order Beliefs in Survey Data*". In: *The Quarterly Journal of Economics* 136.3, pp. 1387–1446.

CompNet (2022). *CompNet 2022, User Guide for the 9th Vintage of the CompNet Dataset*. Tech. rep.

Cremer, H., F. Gahvari, and N. Ladoux (1998). "Externalities and optimal taxation". In: *Journal of Public Economics* 70.3, pp. 343–364.

De Rassenfosse, G., J. Kozak, and F. Seliger (2019). "Geocoding of worldwide patent data". In: *Scientific Data* 6.

Dechezleprêtre, A., A. Fabre, T. Kruse, B. Planterose, A. Sanchez Chico, and S. Stantcheva (2025a). "Fighting Climate Change: International Attitudes toward Climate Policies". In: *American Economic Review* 115.4, pp. 1258–1300.

— (2025b). "Fighting Climate Change: International Attitudes toward Climate Policies". In: *American Economic Review* 115.4, pp. 1258–1300.

Dechezleprêtre, A., Y. Ménière, and M. Mohnen (2017). "International patent families: from application strategies to statistical indicators". In: *Scientometrics* 111.2, pp. 793–828.

Dechezleprêtre, A., C. B. Muckley, and P. Neelakantan (2021). "Is firm-level clean or dirty innovation valued more?" In: *The European Journal of Finance* 27.1-2, pp. 31–61.

Demski, C., S. Capstick, N. Pidgeon, R. G. Sposato, and A. Spence (2017). "Experience of extreme weather affects climate change mitigation and adaptation responses". In: *Climatic Change* 140.2, pp. 149–164.

Dernis, H. and M. Khan (2004). *Triadic Patent Families Methodology*. OECD Science, Technology and Industry Working Papers 2004/02.

Djourelova, M., R. Durante, E. Motte, and E. Patacchini (2024). "Experience, Narratives, and Climate Change Beliefs". In: *SSRN Electronic Journal*.

Doran, R., S. Pallesen, G. Böhm, and C. A. Ogunbode (2022). "When and why do people experience flight shame?" In: *Annals of Tourism Research* 92, p. 103254.

Douenne, T., A. J. Hummel, and M. Pedroni (2024a). "DP19151 Optimal Fiscal Policy in a Climate-Economy Model with Heterogeneous Households". In.

— (2024b). *Optimal Fiscal Policy under Heterogeneous Environmental Preferences*. CESifo Working Paper 19151. CESifo.

Duckworth, J. J., M. Randle, L. S. McGale, A. Jones, B. Doherty, J. C. G. Halford, and P. Christiansen (2022). "Do front-of-pack 'green labels' increase sustainable food choice and willingness-to-pay in U.K. consumers?" In: *Journal of Cleaner Production* 371, p. 133466.

Dufwenberg, M., P. Heidhues, G. Kirchsteiger, F. Riedel, and J. Sobel (2011). "Other-Regarding Preferences in General Equilibrium". In: *The Review of Economic Studies* 78.2, pp. 613–639.

Eichner, T. and R. Pethig (2021). "Climate Policy and Moral Consumers". In: *The Scandinavian Journal of Economics* 123.4, pp. 1190–1226.

Einiö, E., J. Feng, and X. Jaravel (2025). "Social push and the direction of innovation". In: *American Economic Review*.

Eisensee, T. and D. Strömberg (2007). "News Droughts, News Floods, and U. S. Disaster Relief*". In: *The Quarterly Journal of Economics* 122.2, pp. 693–728.

Ellingsen, T. and E. Mohlin (2026). "A Model of Social Duties". In: *Journal of Political Economy* 134.5, pp. 1307–1346.

EUROSTAT (2025). *PRODCOM*.

Farhi, E. and X. Gabaix (2020). "Optimal Taxation with Behavioral Agents". In: *American Economic Review* 110.1, pp. 298–336.

Feddersen, T. and A. Sandroni (2006). "A Theory of Participation in Elections". In: *The American Economic Review* 96.4, pp. 1271–1282.

Fehr, E. and G. Charness (2025). "Social Preferences: Fundamental Characteristics and Economic Consequences". In: *Journal of Economic Literature* 63.X, pp. X–X.

Fehr, E. and S. Gächter (2000). "Cooperation and Punishment in Public Goods Experiments". In: *American Economic Review* 90.4, pp. 980–994.

Fehr, E. and K. M. Schmidt (1999). "A Theory of Fairness, Competition, and Cooperation". In: *The Quarterly Journal of Economics* 114.3, pp. 817–868.

Ferey, A., B. B. Lockwood, and D. Taubinsky (2024). "Sufficient Statistics for Nonlinear Tax Systems with Preference Heterogeneity". In: *American Economic Review* 114.10, pp. 3206–3249.

Fiorina, M. P. and S. J. Abrams (2008). "Political polarization in the American public". In: *Annu. Rev. Polit. Sci.* 11.1, pp. 563–588.

Fischbacher, U. and S. Gächter (2010). "Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods Experiments". In: *American Economic Review* 100.1, pp. 541–556.

Fried, S., K. Novan, and W. B. Peterman (2024). "Understanding the Inequality and Welfare Impacts of Carbon Tax Policies — Journal of the Association of Environmental and Resource Economists: Vol 11, No S1". In: *Journal of the Association of Environmental and Resource Economists*.

Gallagher, J. (2014). "Learning about an Infrequent Event: Evidence from Flood Insurance Take-Up in the United States". In: *American Economic Journal: Applied Economics* 6.3, pp. 206–233.

Gennaioli, N. and G. Tabellini (2025). "Presidential Address: Identity Politics". In: *Econometrica* 93.6, pp. 1937–1967.

Gill, A., M. Heinz, H. Schumacher, and M. Sutter (2023). "Social Preferences of Young Professionals and the Financial Industry". In: *Management Science* 69.7, pp. 3905–3919.

Giuliano, P. and A. Spilimbergo (2025). "Aggregate Shocks and the Formation of Preferences and Beliefs". In: *Journal of Economic Literature* 63.2, pp. 542–597.

Golosov, M., J. Hassler, P. Krusell, and A. Tsyvinski (2014). "Optimal Taxes on Fossil Fuel in General Equilibrium". In: *Econometrica* 82.1, pp. 41–88.

Guha-Sapir, D., R. Below, and P. Hoyois (2022). *EM-DAT — The international disasters database*.

Hall, B. H., A. Jaffe, and M. Trajtenberg (2005). "Market Value and Patent Citations". In: *The RAND Journal of Economics* 36.1, pp. 16–38.

Harhoff, D., F. Narin, F. M. Scherer, and K. Vopel (1999). "Citation Frequency and the Value of Patented Inventions". In: *The Review of Economics and Statistics* 81.3, pp. 511–515.

Hassler, J., P. Krusell, and C. Olovsson (2012). *Energy-Saving Technical Change*. Tech. rep. w18456. Cambridge, MA: National Bureau of Economic Research.

Herrnstadt, E. and E. Muehlegger (2014). "Weather, salience of climate change and congressional voting". In: *Journal of Environmental Economics and Management* 68.3, pp. 435–448.

Hirabayashi, Y., R. Mahendran, S. Koirala, L. Konoshima, D. Yamazaki, S. Watanabe, H. Kim, and S. Kanae (2013). "Global flood risk under climate change". In: *Nature Climate Change* 3.9, pp. 816–821.

Hulme, M., E. M. Barrow, N. W. Arnell, P. A. Harrison, T. C. Johns, and T. E. Downing (1999). "Relative impacts of human-induced climate change and natural climate variability". In: *Nature* 397.6721, pp. 688–691.

Huo, Z. and N. Takayama (2025). "Rational Expectations Models with Higher-Order Beliefs". In: *The Review of Economic Studies* 92.5, pp. 3138–3173.

Intergovernmental Panel on Climate Change (IPCC) (2023). *Climate Change 2022 – Impacts, Adaptation and Vulnerability: Working Group II Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press.

Jaffe, A. B. and G. de Rassenfosse (2017). "Patent citation data in social science research: Overview and best practices". In: *Journal of the Association for Information Science and Technology* 68.6, pp. 1360–1374.

Jones, B. F. (2021). "Science and innovation: The under-fueled engine of prosperity". In: *Rebuilding the Post-Pandemic Economy, ed. Melissa S. Kearney and Amy Ganz (Washington DC: Aspen Institute Press, 2021)*.

Jones, B. F. and L. H. Summers (2020). *A Calculation of the Social Returns to Innovation*. Working Paper.

Kalatzi Pantera, D., T. Böhmelt, and Z. Bakaki (2023). "The transnational influence of natural disasters on environmental attitudes". In: *European Journal of Political Research* 62.3, pp. 761–780.

Känzig, D. R. (2023). "The Unequal Economic Consequences of Carbon Pricing". In: *Working Paper*.

Kaplow, L. (2006). "On the undesirability of commodity taxation even when income taxation is not optimal". In: *Journal of Public Economics* 90.6-7, pp. 1235–1250.

— (2012). "Optimal Control of Externalities in the Presence of Income Taxation". In: *International Economic Review* 53.2, pp. 487–509.

Kaufmann, M., P. Andre, and B. Koszegi (2024). "Understanding Markets with Socially Responsible Consumers*". In: *The Quarterly Journal of Economics* 139.3, pp. 1989–2035.

Keding, L. and M. C. Ritterrath (2025). *Inventors' Personal Experience of Natural Disasters and Green Innovation*. Discussion Paper 380. ECONtribute.

Krasa, S. and M. K. Polborn (2014). "Social Ideology and Taxes in a Differentiated Candidates Framework". In: *American Economic Review* 104.1, pp. 308–322.

Kuchler, T. and B. Zafar (2019). "Personal Experiences and Expectations about Aggregate Outcomes". In: *The Journal of Finance* 74.5, pp. 2491–2542.

Laroque, G. R. (2005). "Indirect taxation is superfluous under separability and taste homogeneity: a simple proof". In: *Economics Letters* 87.1, pp. 141–144.

Laudenbach, C., A. Weber, R. Weber, and J. Wohlfart (2024). "Beliefs about the Stock Market and Investment Choices: Evidence from a Survey and a Field Experiment". In: *The Review of Financial Studies*, hhae063.

Laurent, S., P. Chiambaretto, and C. Bildstein (2026). "How should we measure flight shame and what are its main drivers?" In: *Transportation Research Part A: Policy and Practice* 204, p. 104802.

Le, H., T. Nguyen, A. Gregoriou, and J. Healy (2024). "Natural disasters and corporate innovation". In: *The European Journal of Finance* 30.2, pp. 144–172.

Leeuwen, B. van and I. Alger (2024). "Estimating Social Preferences and Kantian Morality in Strategic Interactions". In: *Journal of Political Economy Microeconomics* 2.4, pp. 665–706.

List, J. A., M. Rodemeier, S. Roy, and G. Sun (2023). "Judging Nudging: Understanding the welfar effects of nudges versus taxes". In: *Unpublished Manuscript*.

Lorenzoni, G. (2009). "A Theory of Demand Shocks". In: *American Economic Review* 99.5, pp. 2050–2084.

Malmendier, U. and S. Nagel (2011). "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?*". In: *The Quarterly Journal of Economics* 126.1, pp. 373–416.

— (2016). "Learning from Inflation Experiences *". In: *The Quarterly Journal of Economics* 131.1, pp. 53–87.

Manso, G. (2011). "Motivating Innovation". In: *The Journal of Finance* 66.5, pp. 1823–1860.

McCrary, J. (2007). "The Effect of Court-Ordered Hiring Quotas on the Composition and Quality of Police". In: *American Economic Review* 97.1, pp. 318–353.

Meltzer, A. and S. Richard (1981). "A Rational Theory of the Size of Government". In: *Journal of Political Economy* 89, pp. 914–927.

Miao, Q. and D. Popp (2014). "Necessity as the mother of invention: Innovative responses to natural disasters". In: *Journal of Environmental Economics and Management* 68.2, pp. 280–295.

Moscona, J. and K. A. Sastry (2023). "Does Directed Innovation Mitigate Climate Damage? Evidence from U.S. Agriculture*". In: *The Quarterly Journal of Economics* 138.2, pp. 637–701.

Nordhaus, W. D. (1992). "An Optimal Transition Path for Controlling Greenhouse Gases". In: *Science* 258.5086, pp. 1315–1319.

— (2017). "Revisiting the Social Cost of Carbon". In: *Proceedings of the National Academy of Sciences* 114.7, pp. 1518–1523.

Osberghaus, D. and C. Fugger (2022). "Natural disasters and climate change beliefs: The role of distance and prior beliefs". In: *Global Environmental Change* 74, p. 102515.

Owen, A. L., E. Conover, J. Videras, and S. Wu (2012). "Heat Waves, Droughts, and Preferences for Environmental Policy". In: *Journal of Policy Analysis and Management* 31.3, pp. 556–577.

Peters, V., W. S. Langford, M. Sanders, and M. P. Feldman (2024). *Local Finance and Economic Resilience During Extreme Weather Events*. SSRN Scholarly Paper. Rochester, NY.

Pigou, A. (1920). *The Economics of Welfare*. Macmillan.

Pirttilä, J. and M. Tuomala (2002). "Publicly provided private goods and redistribution: A general equilibrium analysis". In: *Scandinavian Journal of Economics* 104.1, pp. 173–188.

Ploeg, F. van der and A. Rezai (2019). "Simple Rules for Climate Policy and Integrated Assessment". In: *Environmental and Resource Economics* 72.1, pp. 77–108.

Popp, D. (2002). "Induced Innovation and Energy Prices". In: *American Economic Review* 92.1, pp. 160–180.

Rassenfosse, G. de and B. van Pottelsberghe de la Potterie (2009). "A policy insight into the R&D–patent relationship". In: *Research Policy* 38.5, pp. 779–792.

Rezai, A. and F. Van der Ploeg (2016). "Intergenerational Inequality Aversion, Growth, and the Role of Damages: Occam's Rule for the Global Carbon Tax — Journal of the Association of Environmental and Resource Economists: Vol 3, No 2". In: *Journal of the Association of Environmental and Resource Economists*.

Roberts, K. (1977). "Voting over Income Tax Schedules". In: *Journal of Public Economics* 8.

Roemer, J. E. (2015). "Kantian optimization: A microfoundation for cooperation". In: *Journal of Public Economics* 127, pp. 45–57.

Rosenberg, N. (1998). "Uncertainty and Technological Change". In: *The Economic Impact of Knowledge*. Routledge. ISBN: 978-0-08-050502-2.

Rosvold, E. L. and H. Buhaug (2021). "GDIS, a global dataset of geocoded disaster locations". In: *Scientific Data* 8.1, p. 61.

Roudier, P., J. C. M. Andersson, C. Donnelly, L. Feyen, W. Greuell, and F. Ludwig (2016). "Projections of future floods and hydrological droughts in Europe under a +2°C global warming". In: *Climatic Change* 135.2, pp. 341–355.

Saez, E. (2002). "The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes". In: *Journal of Public Economics*.

Schmidheiny, K. and S. Siegloch (2023). "On event studies and distributed-lags in two-way fixed effects models: Identification, equivalence, and generalization". In: *Journal of Applied Econometrics* 38.5, pp. 695–713.

Stern, N. (2007). *The Economics of Climate Change: The Stern Review*. Tech. rep. Cambridge University Press.

Stern, S. (2004). "Do Scientists Pay to Be Scientists?" In: *Management Science* 50.6, pp. 835–853.

Stiglitz, J. E. (1982). "Self-selection and Pareto efficient taxation". In: *Journal of Public Economics* 17.2, pp. 213–240.

Stoetzer, L. S. and F. Zimmermann (2024). "A representative survey experiment of motivated climate change denial". In: *Nature Climate Change* 14.2, pp. 198–204.

Sun, L. and S. Abraham (2021). "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects". In: *Journal of Econometrics* 225.2, pp. 175–199.

Thaler, R. H. and C. R. Sunstein (2009). *Nudge: Improving Decisions About Health, Wealth, and Happiness*. New York, NY: Penguin.

Thöni, C. and S. Volk (2018). "Conditional Cooperation: Review and Refinement". In: *Economics Letters* 171, pp. 37–40.

Tol, R. S. J. (2009). "The Economic Effects of Climate Change". In: *Journal of Economic Perspectives* 23.2, pp. 29–51.

Trajtenberg, M. (1990). "A Penny for Your Quotes: Patent Citations and the Value of Innovations". In: *The RAND Journal of Economics* 21.1, pp. 172–187.

Van Der Ploeg, F. and C. Withagen (2012). "Is there really a green paradox?" In: *Journal of Environmental Economics and Management* 64.3, pp. 342–363.

Werning, I. (2007). *Pareto Efficient Income Taxation*. Working Paper. Massachusetts Institute of Technology.

ZEW - Leibniz Centre for European Economic Research (n.d.). *Mannheim Innovation Panel (MIP)*.

Chapter 4

Appendix for Chapter 1

4.1 Appendix - Theory

This appendix formalizes the consumer demand structure, characterizes equilibrium outcomes for monopoly producers, and offers our proofs. It also includes welfare analysis and explores alternative ways of modeling market size effects on innovation.

4.1.1 Demand & Market-Clearing

We consider a consumer with utility

$$U = \int_0^1 \ln \bar{y}_j \, dj, \quad \text{where} \quad \bar{y}_j = \int_{\mathcal{F}_j} y_{j,f} \left(q_{j,f}\right)^{\delta} df,$$

subject to the overall budget constraint

$$\int_0^1 \int_{\mathcal{F}_j} p_{j,f} \, y_{j,f} \, df \, dj = M.$$

Intra-Good Optimization

For each good j, assume the consumer allocates an expenditure m_j (with $\int_0^1 m_j \, dj = M$) across varieties $f \in \mathcal{F}_j$ by solving

$$\max_{\{y_{j,f}\}} \ln \left(\int_{\mathcal{F}_j} y_{j,f} \left(q_{j,f}\right)^{\delta} df \right)$$

subject to

$$\int_{\mathcal{F}_j} p_{j,f} \, y_{j,f} \, df = m_j.$$

Defining

$$\bar{y}_j = \int_{\mathcal{F}_j} y_{j,f} \left(q_{j,f}\right)^{\delta} df,$$

the Lagrangian for this subproblem is

$$\mathcal{L}_j = \ln(\bar{y}_j) - \lambda_j \left(\int_{\mathcal{F}_j} p_{j,f}\, y_{j,f}\, df - m_j \right).$$

Taking the first-order condition with respect to $y_{j,f}$ gives

$$\frac{(q_{j,f})^\delta}{\bar{y}_j} - \lambda_j\, p_{j,f} = 0 \quad \Longrightarrow \quad \frac{(q_{j,f})^\delta}{p_{j,f}} = \lambda_j\, \bar{y}_j \quad \forall f \text{ with } y_{j,f} > 0.$$

Thus, only those varieties that maximize the ratio

$$\frac{(q_{j,f})^\delta}{p_{j,f}}$$

can receive positive demand. Define

$$f^*(j) = \arg\max_{f \in \mathcal{F}_j} \frac{(q_{j,f})^\delta}{p_{j,f}}.$$

Then all expenditure in good j is allocated to variety $f^*(j)$. The budget within good j therefore satisfies

$$p_{j,f^*(j)}\, y_{j,f^*(j)} = m_j,$$

so that

$$y_{j,f^*(j)} = \frac{m_j}{p_{j,f^*(j)}}$$

and

$$\bar{y}_j = \frac{m_j}{p_{j,f^*(j)}} (q_{j,f^*(j)})^\delta.$$

Allocation Across Goods

Substituting the expression for $\bar{y}_j$ into the overall utility yields

$$U = \int_0^1 \ln \left[\frac{m_j}{p_{j,f^*(j)}} \left(q_{j,f^*(j)}\right)^\delta \right] dj = \int_0^1 \Big\{ \ln m_j - \ln p_{j,f^*(j)} + \delta\, \ln(q_{j,f^*(j)}) \Big\} dj.$$

The allocation $\{m_j\}$ is chosen subject to

$$\int_0^1 m_j\, dj = M.$$

Form the Lagrangian for the allocation across goods:

$$\mathcal{L} = \int_0^1 \ln m_j\, dj - \int_0^1 \ln p_{j,f^*(j)}\, dj + \int_0^1 \delta\, \ln(q_{j,f^*(j)})\, dj + \theta \left(M - \int_0^1 m_j\, dj \right).$$

Taking the derivative with respect to m_j for each j gives

$$\frac{1}{m_j} - \theta = 0 \quad \Longrightarrow \quad m_j = \frac{1}{\theta}, \quad \forall\, j.$$

Then $\int_0^1 m_j \, dj = M$ implies

$$\frac{1}{\theta} = M \quad \Longrightarrow \quad \theta = \frac{1}{M}.$$

Hence, the optimal allocation is

$$m_j = M \quad \text{for all } j \in [0,1],$$

so that each good j receives an equal share of the total budget.

Characterization of Expected Monopoly Outcomes

Due to the structure of consumer demand, a successful inventor who upgrades quality by a factor γ faces the unsuccessful producer whose quality remains one "step" behind, $q_{\text{comp}} = q_M/\gamma$, and who sells at marginal cost c. Indifference between the innovator's variety and that of her competition requires

$$p_{Mj}\,(q_M)^{\delta} \;=\; p_{\text{comp}}\left(q_M/\gamma\right)^{\delta} \;\Longrightarrow\; p_{Mj} \;=\; \gamma^{\delta}\, c.$$

With total expenditure on good j normalized to one, this implies

$$y_{Mj} \;=\; \frac{1}{p_{Mj}} \;=\; \frac{1}{\gamma^{\delta}\, c}, \qquad \pi_{Mj} \;=\; 1 - \frac{c}{p_{Mj}} \;=\; 1 - \gamma^{-\delta}.$$

Taking local expectations $E_l[\cdot]$ over δ immediately yields

$$E_l[y_{Mj}] = \frac{1}{E_l[p_{Mj}]} = \frac{1}{\gamma^{\,E_l[\delta]}\, c}, \qquad E_l[\pi_{Mj}] = 1 - \frac{1}{\gamma^{\,E_l[\delta]}},$$

as in equation (1.11). Thus, prospective inventors anticipate these output and profits whenever their R&D succeeds.

4.1.2 Proof of Proposition 1

In this Section we prove Proposition 1 of Section 1.6. We know that $\gamma > 1$ and that for each j we have $E[\delta] \geq 0$. Let

$$\Phi\left(E[\delta]\right) \equiv 1 - \frac{1}{\gamma^{E[\delta]}}.$$

A straightforward calculation shows that

$$\Phi'(x) = \gamma^{-x} \ln \gamma \quad \text{and} \quad \Phi''(x) = -\gamma^{-x} (\ln \gamma)^2.$$

Since $\gamma^{-x} > 0$ for all $x \geq 0$ and $\ln \gamma > 0$, it follows that $\Phi''(x) < 0 \quad$ for all $x \geq 0$. Thus, Φ is strictly concave on $[0, \infty)$. Let the average environmental expectation across regions be $\overline{E[\delta]} = \int_l E_l[\delta]\, f(l)\, dl$. Then, by Jensen's inequality for the strictly concave function Φ, we have

$$\Phi\Big(\overline{E_l[\delta]}\Big) > \int_l \Phi\Big(E_l[\delta]\Big) f(l)\, dl, \tag{4.1}$$

provided that the $E_l[\delta]$ are not degenerate. Explicitly, this is

$$1 - \frac{1}{\gamma^{\overline{E[\delta]}}} > \int_l \left(1 - \frac{1}{\gamma^{E_l[\delta]}}\right) f(l)\, dl.$$

Multiplying both sides of equation (4.1) by the positive constant Δ_j/κ yields

$$\frac{\Delta_j}{\kappa}\left(1 - \frac{1}{\gamma^{\overline{E[\delta]}}}\right) > \frac{\Delta_j}{\kappa}\int_l \left(1 - \frac{1}{\gamma^{E_l[\delta]}}\right) f(l)\, dl.$$

Now, suppose the intrinsic motivation is heterogeneous (i.e., it may vary with l). The average of intrinsic motivation μ_l is defined as: $\overline{\mu} = \int_l \mu_l\, f(l)\, dl$. Since the intrinsic motivation term enters linearly, we have

$$\frac{\overline{\mu}}{\alpha\kappa} = \int_l \frac{\mu_l}{\alpha\kappa}\, f(l)\, dl.$$

Therefore, adding these linear terms to both sides we obtain

$$\frac{\Delta_j}{\kappa}\left(1 - \frac{1}{\gamma^{\overline{E[\delta]}}}\right) + \frac{\overline{\mu}}{\alpha\kappa} > \int_l \left[\frac{\Delta_j}{\kappa}\left(1 - \frac{1}{\gamma^{E_l[\delta]}}\right) + \frac{\mu_l}{\alpha\kappa}\right] f(l)\, dl \quad \forall j. \tag{4.2}$$

Inequality (4.2) completes the proof. Notice that the nonlinearity (and hence the strict inequality) originates solely from the concave function Φ; the linear intrinsic motivation term averages exactly.

4.1.3 Welfare Derivation and Proof of Proposition 2

Societal welfare depends on quality-adjusted consumption (equation (1.9)) and an emission externality. With consumption expenditure normalized to 1, aggregate emissions are:

$$X = \int_j (1 - z_j) y(\Delta_j) + z_{lj} y_{M_j}/\gamma dj \tag{4.3}$$

The social planner forms expectations using the population-average belief, so that

$$\hat{E}[y_{Mj}] = \frac{1}{E[p_{Mj}]} = \frac{1}{\gamma^{\overline{E[\delta]}_C}}, \quad \hat{E}[\pi_{Mj}] = 1 - \frac{1}{\gamma^{\overline{E[\delta]}}}. \tag{4.4}$$

The social planner chooses the research rate in every sector j to maximize welfare

$$\begin{aligned}\max_{z_j} W = &\int_j (1-z_j)\ln \hat{E}[y(\Delta_j)] + z_j \ln\left(\gamma^{\overline{E[\delta]}}\hat{E}[y_{M_j}]\right)\\ &-\psi\left[(1-z_j)\hat{E}[y(\Delta_j)] + z_j\hat{E}[y_{M_j}]/\gamma\right]\\ &.+\lambda\left[(1-z_j)(1-\Delta_j)\hat{E}[\pi_{M_j}] + z_j\hat{E}[\pi_{M_j}] - K(\eta_j)z_j^2/2 + \frac{\overline{\mu}}{\alpha\kappa}\right]dj\end{aligned} \tag{4.5}$$

subject to inventor utility being non-negative. The Lagrange multiplier λ denotes the value the social planner attaches to innovators' profits/utility; when $\lambda \to \infty$, the social planner simply maximizes inventor utility.

Proof of Proposition 2. Define the average private research rate and the average-expectations research rate as

$$\underline{z_j}(\Delta_j, E_l[\delta], \eta_j) \equiv \int_l \left(\frac{\Delta_j}{K(\eta_j)}\left(1 - \frac{1}{\gamma^{E_l[\delta]}}\right) + \frac{\mu_l}{\alpha\kappa}\right) f(l)dl,$$

$$\overline{z_j}(\Delta_j, \overline{E[\delta]}, \eta_j) \equiv \frac{\Delta_j}{K(\eta_j)}\left(1 - \frac{1}{\gamma^{\overline{E[\delta]}}}\right) + \frac{\overline{\mu}}{\alpha\kappa}.$$

Proposition 1 (equation 1.17) extends straightforwardly to heterogeneous market size, giving $\forall j$

$$\frac{\Delta_j}{K(\eta_j)}\left(1 - \frac{1}{\gamma^{\overline{E[\delta]}}}\right) + \frac{\overline{\mu}}{\alpha\kappa} > \int_l \left(\frac{\Delta_j}{K(\eta_j)}\left(1 - \frac{1}{\gamma^{E_l[\delta]}}\right) + \frac{\mu_l}{\alpha\kappa}\right) f(l)dl \tag{4.6}$$

so that $\overline{z_j} > \underline{z_j}$. From the first-order condition with respect to z_j we obtain

$$z_j^*(\Delta_j, \overline{E[\delta]}, \eta_j) = \underbrace{\overline{z_j}(\Delta_j, \overline{E[\delta]}, \eta_j)}_{\text{average expectations}} + \frac{1}{\lambda K(\eta_j)}\left[\underbrace{\ln\left[\frac{\gamma^{\overline{E[\delta]}}\hat{E}[y_{M_j}]}{\hat{E}[y(\Delta_j)]}\right]}_{\geq 0} + \underbrace{\psi\left[\hat{E}[y(\Delta_j)] - \hat{E}[y_{M_j}]/\gamma\right]}_{>0}\right] \tag{4.7}$$

The first correction term is non-negative: from (4.4) we have $\hat{E}[y_{M_j}]\gamma^{\overline{E[\delta]}} = 1/c$, which together with (1.14) implies $\hat{E}[y(\Delta_j)] \leq 1/c = \gamma^{\overline{E[\delta]}}\hat{E}[y_{M_j}]$. The second term is strictly positive since $y_j(\Delta_j) \geq y_{M_j}$ for all $\Delta_j \in [1/2, 1]$ (from (1.14)) and $\gamma > 1, \psi > 0$. Combined with $\overline{z_j} > \underline{z_j}$, this establishes $z_j^* > \underline{z_j}$.

4.1.4 Parameter Calibration - Comparative Statics

Figure 4.1 plots our comparative static with respect to market size.

This section provides the parameters with which we calibrate our model in order to plot Figures 1.7 and 4.1.

For the purpose of our comparative statics analysis, we calibrate the model with the following parameters. In the baseline model without market size, we set the scaling constant of costs to

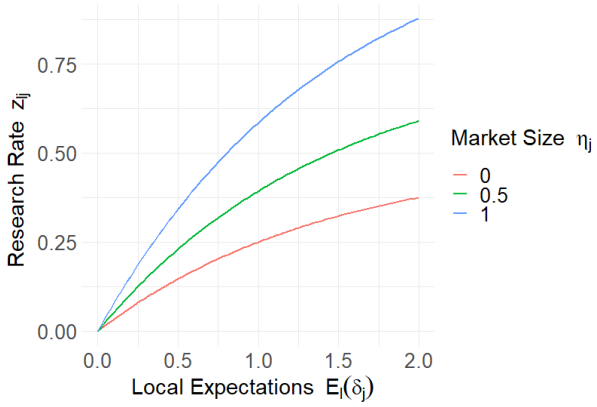


Figure 4.1: Innovation Responses dependent on market size η_j

Note: The figure plots inventors' innovation response dependent on their expectations $E_l[\delta]$ and market size η_j.

$\kappa = 1$ and chose $\gamma = 2$. Innovation intensity, Δ_j, is varied across the representative values 0, 0.5, and 1, while local expectations $E_l[\delta]$ are sampled continuously over the interval $[0, 1]$. The intrinsic utility parameter μ_l is evaluated on the interval $[0, 1]$.

In the extended model that incorporates market size, we fix the innovation intensity at $\Delta_j = 0.5$ and maintain $\gamma = 2$. The market size parameter η_j is evaluated at 0, 0.5, and 1. The function $K(\eta_j)$ is calibrated as

$$K(\eta_j) = \Big(1 - 0.5 \log(1 + \eta_j)\Big)^2,$$

which provides a decreasing and concave relationship in η_j consistent with the assumption of diminishing marginal effects.

4.2 Appendix - Empirical Analysis

This appendix provides empirical robustness checks and supplementary analyses for the main paper. It includes results verifying the robustness of findings to different technology trend controls, examines heterogeneity by technology class and subclass, explores spillover effects across neighboring regions, and tests robustness to alternative definitions of green innovation and disaster exposure. Additionally, it assesses heterogeneity based on inventor experience, disaster severity, disaster type, competition intensity, and green market size.

4.2.1 Alternative Estimator

To address concerns about heterogeneous treatment effects in difference-in-differences designs, we re-estimate our main green innovation results using the estimator proposed by Chaisemartin and D'Haultfœuille (2023) and Chaisemartin and D'Haultfœuille (2024a). This estimator remains valid under treatment effect heterogeneity and serves as a robustness check for our baseline specification in Equation (1.4).

In contrast to our preferred approach, which uses all untreated regions as controls regardless

of treatment history, the alternative estimator restricts comparisons to regions with identical treatment trajectories up to period $t-1$. For example, it compares regions with two prior disasters that receive a third at time t, to regions with the same prior exposure that are not treated at t. Once a control region becomes treated, it drops out of the control group. As a result, the control group diminishes over time as more units receive treatment.

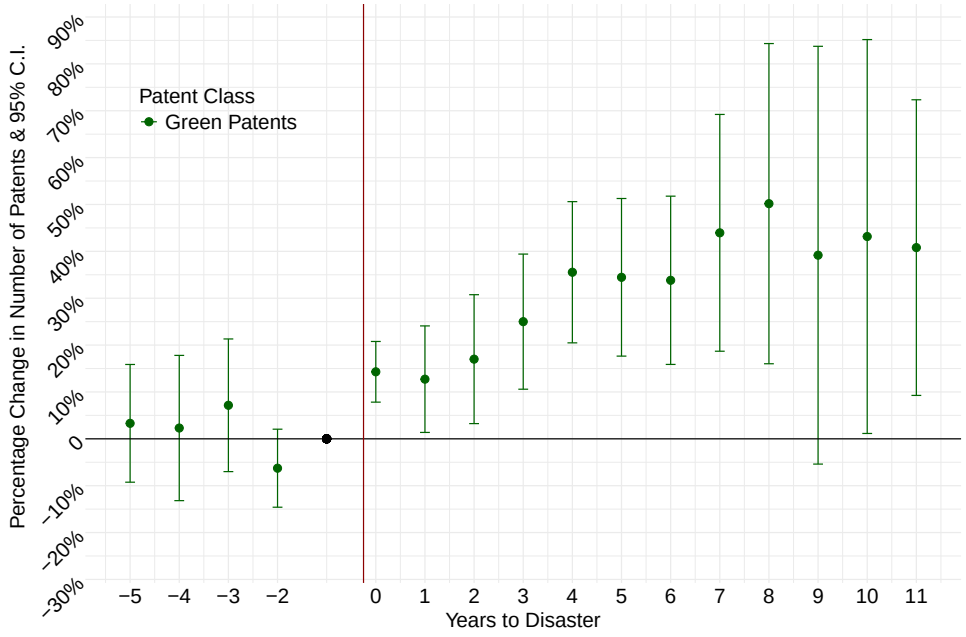


Figure 4.2: Patenting following Exposure to a Natural Disaster — Estimator of Chaisemartin and D'Haultfœuille 2023; Chaisemartin and D'Haultfœuille 2024a

Note: This figure depicts the results of our alternative event-study using the estimator of Chaisemartin and D'Haultfœuille (2023) and Chaisemartin and D'Haultfœuille (2024a). Standard errors are clustered at the region level, and confidence intervals are drawn for the 95% interval.

The resulting estimates (Figure 4.2) are somewhat larger and more persistent than those from our baseline model (Figure 1.3). This is encouraging, as it suggests that our findings are not an artifact of bias introduced by treatment effect heterogeneity. However, the alternative estimator also introduces trade-offs. The shrinking control group reduces statistical power in later event periods, and the restriction to regions with identical treatment histories may limit generalizability.

While it is not entirely clear which approach is more suitable in our context, the qualitative similarity across both estimators increases confidence in the robustness of our results. The alternative estimator offers protection against heterogeneity bias, while our preferred specification leverages a broader and more interpretable sample. Taken together, the results from both approaches point to a consistent and significant relationship between disaster exposure and green innovation. Given that the TWFE specification produces more conservative estimates and facilitates clearer interpretation, we retain it as our primary approach in the main analysis.

4.2.2 Patenting Results with Additional Controls

Our baseline specification does not include controls for regional technology composition and regional GDP per capita, since both variables may themselves be affected by disaster exposure

or be outcomes of the additional green patents induced by disasters, rendering them potentially endogenous ("bad") controls. In this section, we show that our results are robust to including these controls.

We first add technology trend controls. For every region, we calculate the percentage of patents falling into each broad CPC class at time t.[1] These controls account for different time trends in a region's patenting industry composition, for instance, the impact of a large pharmaceutical company leaving a region, which would affect patenting in class C (Chemistry; Metallurgy). Figure 4.3 reports the event-study estimates with these controls included. The results are near identical to our baseline specification, both in magnitude and timing.

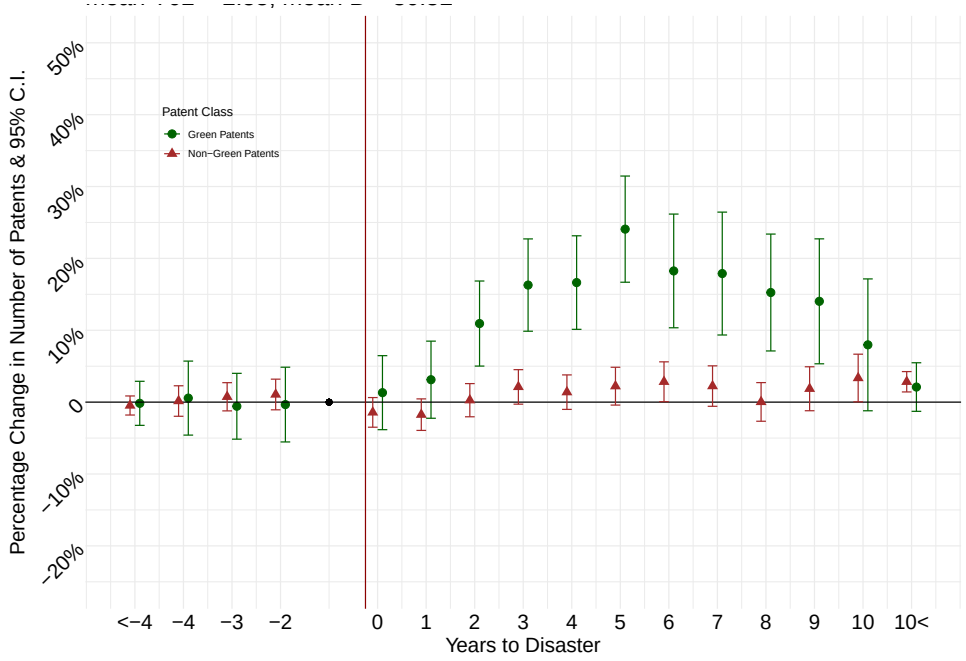


Figure 4.3: Patenting following exposure with technology trend controls

Note: This figure depicts the results for our baseline event-study specification, additionally including technology trend controls (regional CPC class shares). We plot one regression for green and one for non-green patents. Standard errors are clustered on the region level, and confidence intervals are drawn for the 95% interval.

Second, we control for regional GDP per capita. Regional GDP per capita data for Germany is obtained from the INKAR database (*Indikatoren und Karten zur Raum- und Stadtentwicklung*), maintained by the Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR).[2] INKAR provides socioeconomic indicators at the Kreis level, which is the same administrative unit at which our analysis is conducted (403 Kreise in our sample). The GDP control therefore matches the granularity of our treatment and outcome variables precisely. For France, regional GDP per capita data is sourced from the Institut National de la Statistique et des Études Économiques (INSEE). However, INSEE does not publish GDP figures at the arrondissement or département level; the finest available granularity is the *région*,

[1]These classes are: A - "Human Necessities - Agriculture", B - "Performing Operations; Transporting", C - "Chemistry; Metallurgy", D - "Textiles; Paper", E - "Fixed Constructions", F - "Mechanical Engineering; Lighting; Heating; Weapons; Blasting", G - "Physics", H - "Electricity". For example, in 2007, 33% of all patented inventions by inventors in Dunkerque had the CPC class C (Chemistry; Metallurgy).

[2]See `https://www.inkar.de/`.

of which there are 13 in metropolitan France. Since our French sample contains 350 arrondissements, the GDP control assigns the same value to all arrondissements within a given *région*, absorbing only broad regional economic trends rather than local variation. Despite this difference in granularity, the asymmetry is inconsequential for identification: country-by-year fixed effects ensure that treated and control regions are compared only within the same country and year, so the different levels of measurement across countries cannot introduce bias.

Figure 4.4 shows that including GDP per capita controls leaves the event-study pattern broadly unchanged, with estimates slightly smaller but consistent with the baseline results.

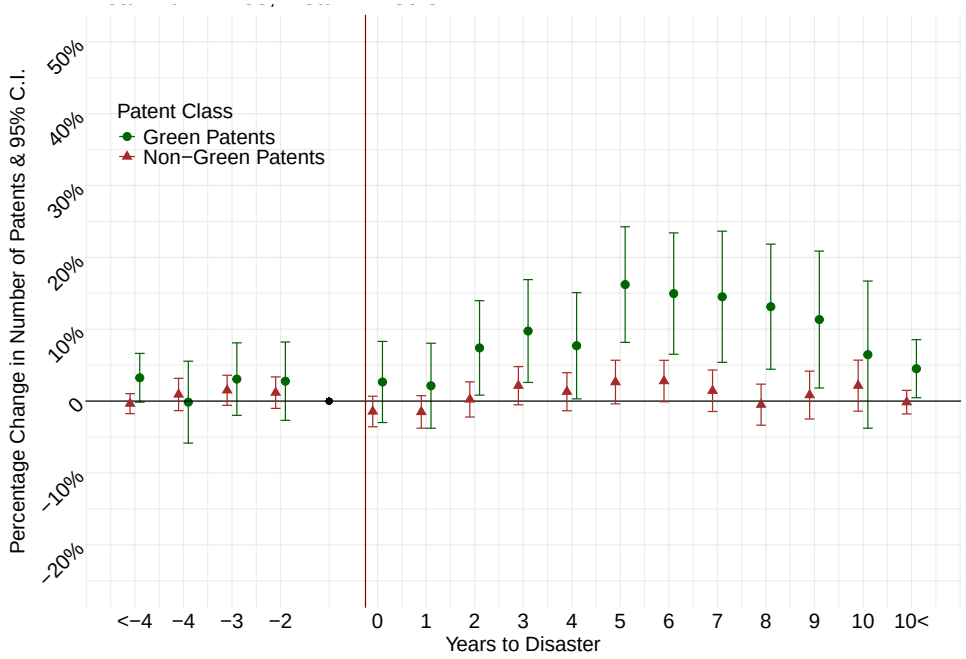


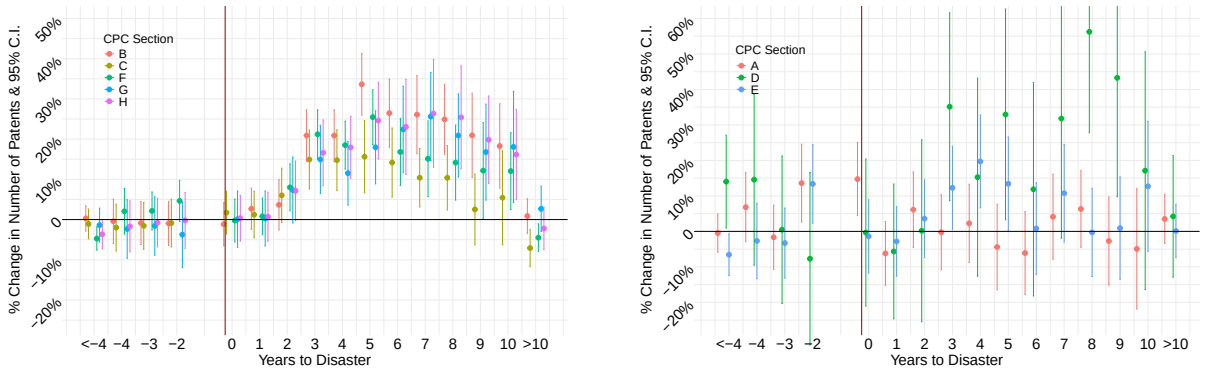
Figure 4.4: Patenting following exposure with GDP per capita controls

Note: This figure depicts the results for our baseline event-study specification, additionally including regional GDP per capita controls (Kreis level for Germany, *région* level for France). We plot one regression for green and one for non-green patents. Standard errors are clustered on the region level, and confidence intervals are drawn for the 95% interval.

4.2.3 Green innovation by technology class

In this section of the appendix, we examine whether the baseline effect of natural disasters on green innovation is concentrated in particular technological domains or reflects a broader pattern across fields. To this end, we replicate our baseline green innovation event study specification separately by broad technology/CPC class. Panel A in Figure 4.5 shows results for classes B (Performing Operations; Transporting), C (Chemistry; Metallurgy), F (Mechanical Engineering; Lighting; Heating), G (Physics), and H (Electricity). These classes closely mirror the baseline response in both timing and magnitude, exhibiting a smooth and persistent increase in green patenting following disasters. This consistency suggests that the average effect is not driven by any single technological area.

Panel B in Figure 4.5 presents results for classes A (Human Necessities), D (Textiles; Paper), and E (Fixed Constructions), where greater heterogeneity is observed. Class D shows particularly large and volatile effects, while class E remains broadly aligned with the baseline



(a) Effects for CPC classes B,C,F,G,H

(b) Effects for CPC classes A,D,E

Figure 4.5: Green innovation response to natural disaster by technology class

Note: This Figure depicts the effects of natural disaster exposure on green patenting for patents from different technology classes (1-digit CPC classes). The mean number of green patents in CPC classes is as follows: A: 0.114, B: 0.451, C: 0.477, D: 0.054, E: 0.102, F: 0.766, G: 0.221, H: 0.550. Standard errors are clustered on the region level, and confidence intervals are drawn for the 95% interval.

pattern, albeit with more noise. Class A is the most distinct, possibly reflecting the different dynamics of green innovation in agriculture and related fields.

Taken together, these results indicate that the positive innovation response to disasters is not confined to a narrow set of technologies, but is instead distributed across a wide range of broad CPC classes.

4.2.4 Subclasses of Y02

Our results suggest that natural disasters are a robust and broad-based driver of green innovation. This has important implications for understanding the determinants of environmental technological change, as it indicates that external shocks can meaningfully stimulate patenting activity across a diverse set of green technology fields.

We can split our sample further by a patents respective "green subclass". See Table 4.1 for these subclasses.

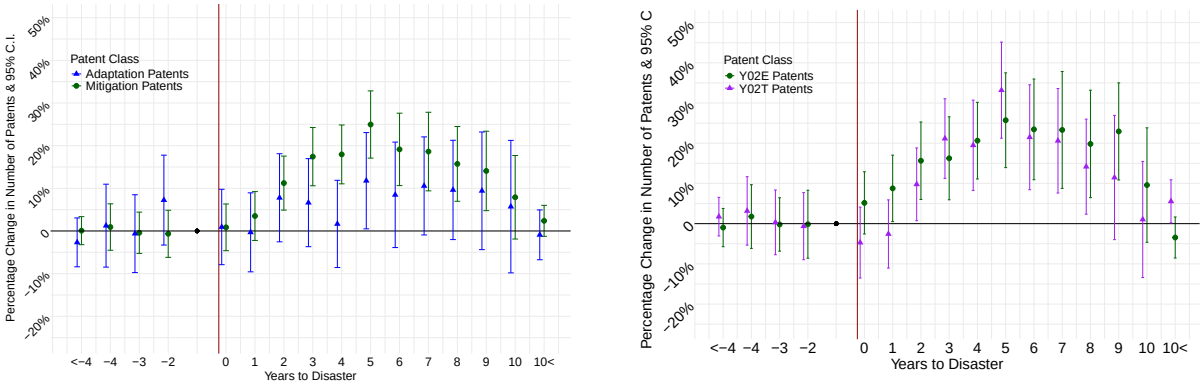
In Table 1.2 and Figure 4.6a we pool subclasses Y02B-Y02W into mitigation technologies. Figure 4.6a depicts our results when estimating (1.4) on the split sample of adaptation and mitigation technologies.

In Table 4.2, we report estimates of the effect of natural disaster exposure on different Y02 green patent subclasses. The outcome variables are subclass-specific indicators for green patent filings, and the main regressor is the cumulative count of natural disasters over the past five years. Each column represents a separate regression. All models include country-year fixed effects, and region fixed effects.

We find consistently positive and statistically significant effects across all Y02 subclasses, significant at the 1% level in all cases. The largest effects are observed in Y02D (energy generation), Y02C (carbon capture and storage), and Y02B (building technologies), indicating particularly strong responsiveness in domains most directly tied to energy systems and emissions mitigation. Transport (Y02T) and energy conservation (Y02E) technologies, while not associated with the largest effects, are notable for their relatively high baseline levels of patenting

Class	Description
Y02A	Climate change mitigation technologies related to agriculture, forestry, or land use
Y02B	Climate change mitigation technologies related to buildings, e.g., housing, appliances, or related end-user applications
Y02C	Capture, storage, sequestration, or disposal of greenhouse gases
Y02D	Climate change mitigation technologies in information and communication technologies (ICT), aiming at reducing ICT-related energy use
Y02E	Reduction of greenhouse gas (GHG) emissions related to energy generation, transmission, or distribution
Y02P	Climate change mitigation technologies in the production or processing of goods
Y02T	Climate change mitigation technologies related to transportation
Y02W	Climate change mitigation technologies related to wastewater treatment or waste management

Table 4.1: Y02 subclasses of climate change mitigation technologies



(a) Patenting in Mitigation and Adaptation

(b) Patenting in Y02E and Y02T subclasses

Figure 4.6: Green innovation patterns in Germany and France by technology focus

Note: **Left panel:** Each time we split the sample to only contain technologies from the respective CPC class(es). Therefore the figure depicts 2 separate regressions. The sample average of mitigation patents is 2.323, while the sample average of adaptation patents is 0.2256. **Right panel:** This figure plots our baseline specification, when only looking at the Y02E and Y02T subclasses. The sample average of Y02E patents is 0.885, while the sample average of Y02T patents is 0.933. Standard errors are clustered on the region level and confidence intervals are drawn for the 95% interval.

activity and their relevance to energy-saving policy goals. Given their substantive importance and distinct temporal dynamics, we present separate event-study plots for these subclasses in Appendix Figure 4.6b.

Taken together, these results demonstrate that the innovation response to natural disasters is not confined to a narrow subset of green technologies, but rather operates across a wide range of sectors, with particularly strong effects in energy-related and emissions-reducing fields. This underscores the role of natural disasters as a catalyst for environmentally beneficial technological change.

Table 4.2: Regression by Y02 subclass

	Dependent variable:							
	Y02A	Y02B	Y02C	Y02D	Y02E	Y02P	Y02T	Y02W
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cumulative Count	0.044***	0.135***	0.178***	0.263***	0.083***	0.078***	0.086***	0.043***
	(0.013)	(0.016)	(0.047)	(0.044)	(0.014)	(0.016)	(0.012)	(0.015)
Country-Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample Means	0.226	0.251	0.036	0.075	0.885	0.425	0.933	0.152
Observations	15,813	15,813	15,813	15,813	15,813	15,813	15,813	15,813
R^2	0.513	0.483	0.362	0.461	0.627	0.640	0.609	0.416
Adj. R^2	0.487	0.456	0.329	0.432	0.607	0.621	0.589	0.385

Note: This table reports the results of our baseline difference-in-differences regression for different subclasses of green patents. We report results for all subclasses of Y02. Sample Means gives the respective sample means per region-year for these subclasses. Standard errors are clustered on the region level and are reported in parentheses. P-values are as follows: *p<0.1; **p<0.05; ***p<0.01

4.2.5 Most Severe-Disaster Severity

We use EM-DAT data to identify each region's most severe natural disaster, defined as the event with the highest recorded number of deaths. We then construct a treatment indicator that equals one in all years following the region's most deadly disaster. This specification restricts treatment to occur at most once per region and allows us to implement the staggered adoption event-study estimator of Sun and Abraham (2021), which is designed for settings where treatment is irreversible and varies in timing.

Figure 4.7 shows the resulting event-study estimates. The identifying variation in this setting comes from comparing regions that have already experienced their most severe disaster to those that have not yet done so. Because treatment is limited to a single (and extreme) event per region, we reduce concerns about overlapping or serially correlated shocks. At the same time, this comes at the cost of substantially reduced variation, which leads to increased standard errors and greater noisiness in the estimated effects.

Despite the reduced statistical precision, we find that the most severe disasters are associated with stronger innovation responses than in our baseline specification. However, the effect is more delayed, with increases in green patenting becoming visible only several years after the event.

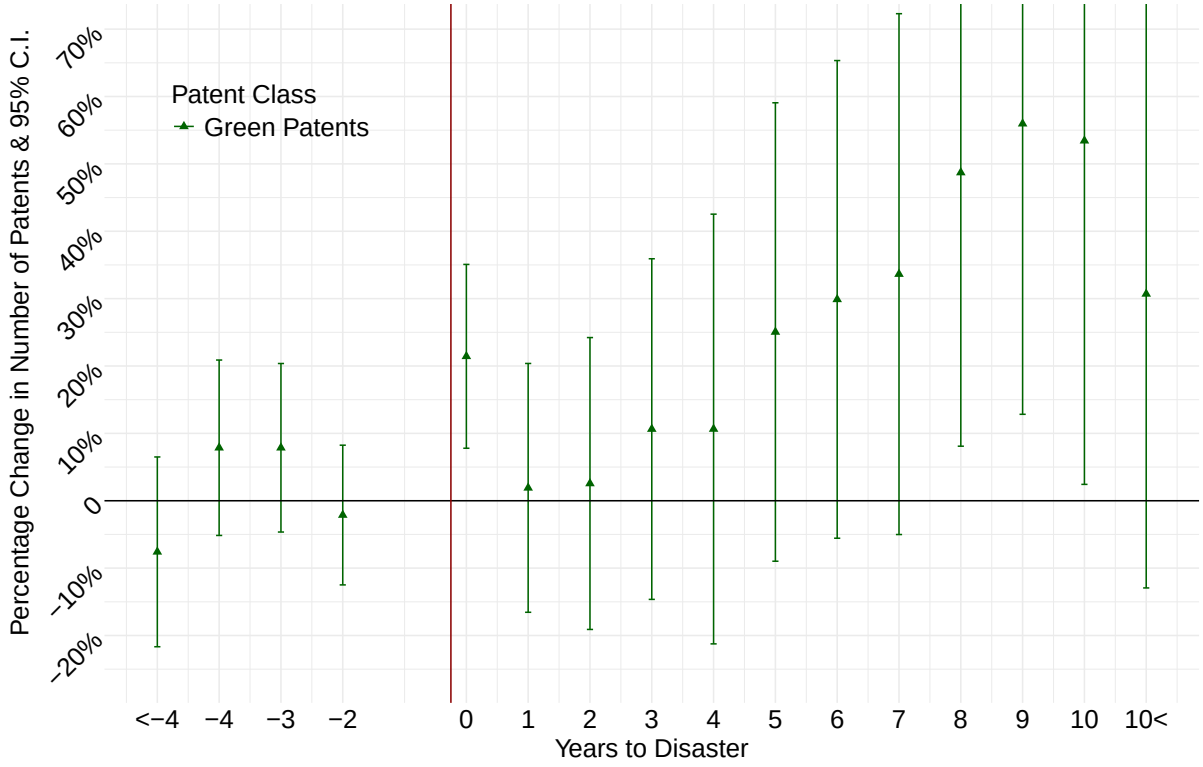


Figure 4.7: Patenting following the Exposure to the Most Severe Natural Disaster

Note: This figure plots green patenting after the most severe disaster, as measured by deaths, in our sample. We estimate effects using the estimator of Sun and Abraham (2021). Standard errors are clustered at the region level, and 95% confidence intervals are shown.

4.2.6 Disaster Type

Figure 4.8 plots our event-study estimates separately for each disaster type: extreme temperature, flood, and storm. This breakdown allows us to explore potential heterogeneity in the innovation response depending on the nature of the shock.

While the estimates are noisier due to the smaller number of observations for each disaster category, the overall patterns are consistent with our main findings in Figure 1.3. In all three cases, we observe a notable increase in green patenting activity in the years following disaster exposure. This supports our interpretation that natural disasters act as catalysts for environmentally oriented innovation, regardless of the specific type of event.

These results suggest that the mechanism linking disaster exposure to innovation is not driven by one specific type of shock, but appears to hold more generally across a range of climate-related events.

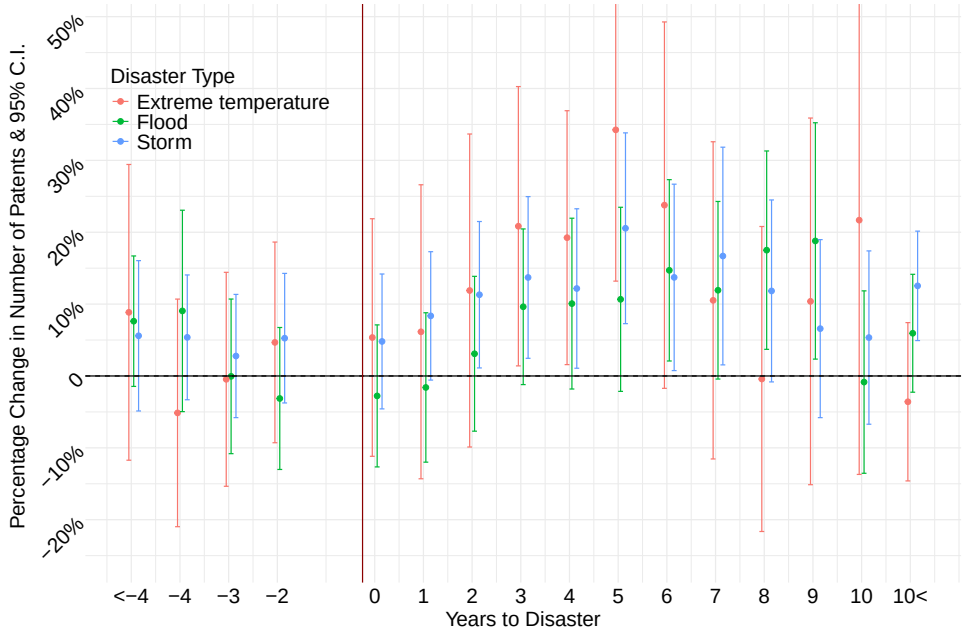


Figure 4.8: Green Patenting Following Different Disaster Types

Note: This figure presents event-study estimates for the impact of three different disaster types (extreme temperature, flood, and storm) on green patenting. We omit droughts, as there are not enough droughts (3) in our sample to plot meaningful estimates. Confidence intervals correspond to the 95% level, and standard errors are clustered at the region level.

4.2.7 Spillovers - Alternative Modeling

Table 4.3 depicts alternative sample restrictions compared to Table 1.3 in the main text. In columns (1)–(3) we exclude regions affected in the past 3 years, (4)–(6) in the past 4 years, (7)–(9) in the past 6 years, and (10)–(12) in the past 7 years. The more regions we remove, the smaller our sample size comes. We do so as to not have affected regions in our control group. Overall the results for spillovers from regions that are 100km or 150km remain largely the same. However, when we remove all regions that were affected in the past 6 or 7 years, we find somewhat larger effects of natural disasters in regions closer than 50km away.

Table 4.4 presents regression results examining the spillover effects of neighboring disasters using an unrestricted sample that includes regions directly affected by disasters. The estimates are disaggregated by distance thresholds of 50km, 100km, and 150km. At the 50km threshold (column 1), the effect of cumulative neighboring disasters is positive and statistically significant, with a coefficient of 0.0115. For the 100km and 150km thresholds (columns 2 and 3), the coefficients are negative. This is due to the control group now containing regions that are affected by natural disasters. We essentially invert treatment and control group.

Table 4.3: Spillovers of Neighboring Disasters - Alternative Windows

	Dependent variable: $P(Y02_{lt})$											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Cumulative Count (50km)	0.0102***			0.0099***			0.0144***			0.0272**		
Neighboring Disasters	(0.0032)			(0.0033)			(0.0051)			(0.0107)		
Cumulative Count (100km)		0.00005			−0.0003			0.0005			0.0006	
Neighboring Disasters		(0.0005)			(0.0005)			(0.0006)			(0.0006)	
Cumulative Count (150km)			−0.0003			−0.0006*			−0.0002			−0.0001
Neighboring Disasters			(0.0004)			(0.0003)			(0.0004)			(0.0003)
Country-Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,464	5,464	5,464	4,680	4,680	4,680	3,704	3,704	3,704	3,309	3,309	3,309
R^2	0.8416	0.8409	0.8409	0.8542	0.8536	0.8537	0.8456	0.8445	0.8444	0.8511	0.8497	0.8496
Adjusted R^2	0.8149	0.8141	0.8142	0.8249	0.8242	0.8244	0.8119	0.8105	0.8105	0.8175	0.8158	0.8157

Note: This table gives the results of our spillover analysis for different samples dependent on regions past natural disaster exposure. We exclude all regions with past disaster exposure for different time windows. In columns (1)–(3) we exclude regions affected in the past 3 years, (4)–(6) in the past 4 years, (7)–(9) in the past 6 years, and (10)–(12) in the past 7 years. The more regions we remove, the smaller our sample size becomes. "Cumulative Count Neighboring Disasters" is the count of past natural disasters in neighboring regions. Which regions are considered as "neighbors" depends on the distance threshold as shown in Figure 1.4. Standard errors are clustered on the region level and are reported in parentheses. P-values: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 4.4: Spillovers of Neighboring Disasters - Unrestricted Sample

	Dependent variable:		
	$P(Y02_{lt})$		
	(1)	(2)	(3)
Cumulative Count (50km) Neighboring Disasters	0.0115^{***}		
	(0.0026)		
Cumulative Count (100km) Neighboring Disasters		-0.0017^{***}	
		(0.0003)	
Cumulative Count (150km) Neighboring Disasters			-0.0014^{***}
			(0.0002)
Country-Year F.E.	Yes	Yes	Yes
Region F.E.	Yes	Yes	Yes
Observations	15,813	15,813	15,813
R^2	0.7364	0.7362	0.7365
Adjusted R^2	0.7225	0.7222	0.7226

Note: This table gives the estimates of our spillover analysis when we do not constain the sample. Thus all regions, regardless of past exposure, are in the sample. Standard errors are clustered at the region level and reported in parentheses. We do not remove themselves affected regions from the sample. Significance levels: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

4.2.8 Reasons to Innovate Robustness

To assess the robustness of our baseline survey results (Table 1.6), we estimate two alternative specifications that sequentially incorporate more granular regional fixed effects. Table 4.5 includes NUTS-2 fixed effects to control for unobserved time-invariant characteristics at a broader regional level, while Table 4.6 introduces NUTS-3 fixed effects.

Across both specifications, the direction of the estimated coefficients on cumulative disaster exposure remains qualitatively consistent with the baseline, indicating that the underlying relationship is robust. However, once NUTS-3 fixed effects are included, the magnitude of estimated effects is similar, but statistical significance weakens in several cases. This reduction in significance likely reflects the limited within-region variation available across just three survey waves, which constrains identification when highly granular fixed effects are introduced. Nonetheless, the overall pattern provides reassurance that the baseline results are not driven by omitted regional heterogeneity.

French Research Funding

To explore whether changes in public research funding explain the increase in green innovation after natural disasters, we draw on ScanR, an administrative database developed by the French Ministry of Higher Education, Research, and Innovation.[3] The platform lists 121,451 publicly

[3] `https://scanr.enseignementsup-recherche.gouv.fr/`

Table 4.5: Effect of Cumulative Disaster Count on Reasons to Innovate (Robustness Check NUTS-2)

	Dependent variable:					
	Expected Regulation (1)	Expected Demand (2)	Existing Regulation (3)	Public Funding (4)	Voluntary Standard (5)	Reputation (6)
Cumulative Count	3.66***	1.72***	2.89***	1.22**	2.43***	0.263
	(0.582)	(0.497)	(0.569)	(0.499)	(0.563)	(0.745)
Firm Size F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Revenue	Yes	Yes	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Industry F.E. (2-digit)	Yes	Yes	Yes	Yes	Yes	Yes
NUTS-2 F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Y	63.33	55.64	40.66	49.07	57.95	61.47
Observations	8,787	8,754	8,862	8,776	8,782	5,423
R^2	0.554	0.492	0.610	0.422	0.547	0.670
Adj. R^2	0.548	0.485	0.605	0.414	0.541	0.663

Note: This table reports the effect of disaster exposure on the drivers of innovation as measured in the survey. One additional past disaster increases the probability of a firm mentioning these factors by β percentage points. Each column reports coefficients from separate regressions. Outcomes are: Expected Regulation = expected future environmental rules; Expected Demand = expected future market demand for green innovation; Existing Regulation = current environmental rules/charges; Public Funding = government grants/subsidies; Voluntary Standard = voluntarily joining a standard for green practices (e.g. environmental/organic label); Reputation = firm's reputation concerns. Coefficients and standard errors are scaled by 100 for interpretability as percentage point effects. "Cumulative Count" captures disaster exposure. Standard errors clustered at the region (Kreis, NUTS-3) level. Significance levels: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 4.6: Effect of Cumulative Disaster Count on Reasons to Innovate (Robustness Check NUTS-3)

	Dependent variable:					
	Expected Regulation (1)	Expected Demand (2)	Existing Regulation (3)	Public Funding (4)	Voluntary Standard (5)	Reputation (6)
Cumulative Count	2.50**	0.852	1.52	1.15	1.85	1.21
	(1.25)	(1.05)	(1.03)	(0.893)	(1.30)	(0.904)
Firm Size F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Revenue	Yes	Yes	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Industry F.E. (2-digit)	Yes	Yes	Yes	Yes	Yes	Yes
NUTS-3 F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Y	63.33	55.64	40.66	49.07	57.95	61.47
Observations	8,787	8,754	8,862	8,776	8,782	5,423
R^2	0.574	0.515	0.629	0.451	0.566	0.694
Adj. R^2	0.550	0.487	0.608	0.419	0.542	0.665

Note: This table reports the effect of disaster exposure on the drivers of innovation as measured in the survey. One additional past disaster increases the probability of a firm mentioning these factors by β percentage points. Each column reports coefficients from separate regressions. Outcomes are: Expected Regulation = expected future environmental rules; Expected Demand = expected future market demand for green innovation; Existing Regulation = current environmental rules/charges; Public Funding = government grants/subsidies; Voluntary Standard = voluntarily joining a standard for green practices (e.g. environmental/organic label); Reputation = firm's reputation concerns. Coefficients and standard errors are scaled by 100 for interpretability as percentage point effects. "Cumulative Count" captures disaster exposure. Standard errors clustered at the region (Kreis, NUTS-3) level. Significance levels: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

funded research projects initiated in France between 1999 and 2023, including funding from Horizon 2020/Horizon Europe, the French National Research Agency (ANR), and the Hubert Curien Partnership (PHC). We restrict the data to projects launched before 2019 to align with the period for which we observe disaster exposure.

We aggregate project-level information to the region-year level and construct two outcome measures: (1) the count of distinct research funding streams and (2) the total research budget per region-year. We then match this data to natural disaster exposure and estimate difference-in-differences regressions with region and year fixed effects, controlling for CPC technological structure.

Table 4.7 reports the results. Across specifications, we find no statistically significant effects of disaster exposure on either the number of funded projects or the total research budget. Both cumulative and recent disaster exposure are unrelated to regional research funding outcomes. These results rule out selective increases in public R&D support as the primary mechanism behind our main findings.

Table 4.7: French Research Funding

	Dependent variable:			
	Count	Budget	Count	Budget
	(1)	(2)	(3)	(4)
Cumulative Count	0.254	−1,224k		
	(0.385)	(1,248k)		
Disaster Count			−0.365	−165k
Last 3 Years			(0.325)	(1,228k)
Year f.e.	Yes	Yes	Yes	Yes
Region f.e.	Yes	Yes	Yes	Yes
CPC Controls	Yes	Yes	Yes	Yes
Observations	2,848	2,848	2,848	2,848
R^2	0.309	0.106	0.309	0.105
Adj. R^2	0.257	0.038	0.256	0.037

Note: This table reports results for different measures research funding. In columns (1) and (3) we report results on the number of funding allocations to a region, while in columns (2) and (4) we report results on the amount of funding allocated to a region. Cumulative count is the count of past natural disasters. "Disaster Count Last 3 Years" is the count of natural disasters in the past 3 years. We only have research funding data for France. Standard errors are clustered on the region level and are reported in parenthesis. P-values are as follows: *p<0.1; **p<0.05; ***p<0.01

4.2.9 Alternative Measures of Green Innovation Robustness

Our green product variable is derived from a survey question that asks: "During [the past two years], did your enterprise introduce new products or services with the following environmental benefits through the use of these products/services, and if yes, what was their contribution to environmental protection The survey lists the following four benefits: (a) reduced energy use, (b) reduced air, water, soil, or noise pollution, (c) improved recycling of products after use, and (d) extended product life through longer-lasting, more durable products. Respondents could answer with "Yes, significant", "Yes", insignificant, and "No" for each of the four benefits. In our analysis, the dummy variable for Green Products is assigned a value of one if a firm indicated that it has introduced a new product or service encompassing any of the four environmental benefits, regardless of whether that benefit was deemed significant or insignificant. Our within-firm green innovation indicator is based on the following survey question: "During [the past two years], did your enterprise introduce innovations that had any of the following environmental benefits, and if yes, was their contribution to environmental protection rather significant or insignificant?". The survey lists the following benefits (a) reduced energy use per unit of output, (b) reduced material use/ use of water per unit of output, (c) reduced CO2 footprint (total CO2 production), (d) reduced air pollution, (e) reduced noise pollution, (f) replaced fossil energy sourced by renewable energy sources, (g) replaced materials by less hazardous substitutes, (h) recycled waste, water, or materials for own use or sale. Firms could again indicate "Yes", significant, "Yes", insignificant, and "No" for each of the four benefits. For our analysis, the within-firm green innovation indicator equals one if a firm has introduced an innovation with any of the mentioned (significant or insignificant) benefits.

Table 4.8 provides additional robustness for our results reported in Table 1.5, which examine the effect of cumulative natural disaster exposure on survey green innovation outcomes and perceived climate affectedness. While Table 1.5 includes controls for firm size, revenue, industry, and year, it does not include region fixed effects. To provide additional robustness, Table 4.8 introduces regional fixed effects at two levels: columns (1)–(4) include NUTS-2 fixed effects, while columns (5)–(7) include NUTS-3 fixed effects. The direction and relative magnitude of the coefficients on cumulative disaster count remain qualitatively consistent across specifications. As in earlier robustness checks, the decline in statistical significance in the NUTS-3 specification is attributable to reduced statistical power due to limited within-region variation across the three available waves. Since climate affectedness is only elicited in one wave, we can not include NUTS-3 fixed effects, as this is the level of our variation.

Table 4.8: Effect of Cumulative Disaster Count on Green Innovation and Climate Affectedness (Robustness Regional Fixed Effects)

	Green Innovation Outcomes			Climate	Green Innovation Outcomes		
	Green Innovation	Within-firm	Green Products	Affectedness	Green Innovation	Within-firm	Green Products
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Cumulative Count	6.59***	5.75***	4.37***	10.7***	3.22	2.78	2.25*
	(0.702)	(0.640)	(0.532)	(0.968)	(2.08)	(1.80)	(1.18)
Firm Size F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Revenue	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	No (Single Wave)	Yes	Yes	Yes
Industry F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NUTS-2 F.E.	Yes	Yes	Yes	Yes	No	No	No
NUTS-3 F.E.	No	No	No	No	Yes	Yes	Yes
Observations	15,395	15,426	15,226	4,873	15,395	15,426	15,226
R^2	0.638	0.599	0.457	0.597	0.653	0.614	0.475
Adj. R^2	0.635	0.596	0.453	0.588	0.642	0.602	0.458

Note: This table presents robustness checks for green innovation outcomes. Coefficients and standard errors are multiplied by 100 to reflect percentage point effects. All models include firm size (based on employment), revenue controls, year fixed effects, and 2-digit NACE industry fixed effects. We either include NUTS-2 (Bezirk) or NUTS-3 (region) level fixed effects. Standard errors are clustered at the regional (Kreis) level. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

4.2.10 Alternative Competition Windows

In our baseline competition analysis (Section 1.6.2), competition intensity is measured as the average profit margin across the filing year and the year prior. To assess sensitivity to this choice, Table 4.9 re-estimates the competition split using alternative time windows: a 1-year window (the year prior to filing only) and a 3-year window (the three years prior to filing).

Qualitatively, the results remain consistent: the innovation response to natural disaster exposure is concentrated in the high-competition group, while the low-competition group shows no significant effect. The estimated coefficients in the high-competition groups (columns 1 and 3) are close to the baseline (Table 1.7, column 1), though slightly smaller in magnitude. The Wald test p-values are correspondingly larger (0.1417 for the 1-year window and 0.0526* for the 3-year window, compared to 0.0376** in the baseline), reflecting somewhat less precise estimates in the low-competition subsamples (columns 2 and 4). Overall, while the alternative competition windows yield attenuated statistical significance for the between-group difference, the direction and relative magnitude of the effects remain stable, providing additional support for the robustness of our findings.

Table 4.9: Competition Split Above/Below Median

	Dependent variable:			
	$P(Y02_{lt})$			
	1 Year - High Competition	1 Year - Low Competition	3 Year - High Competition	3 Year - Low Competition
	(1)	(2)	(3)	(4)
Cumulative Count	0.086***	0.033	0.097***	−0.001
	(0.023)	(0.031)	(0.029)	(0.029)
Country-Year F.E.	Yes	Yes	Yes	Yes
Region F.E.	Yes	Yes	Yes	Yes
P Value	0.1417		0.0526	
Sample Means	1.8046	1.382	2.0166	1.3536
Observations	9,036	9,036	7,530	7,530
R^2	0.626	0.513	0.706	0.589
Adj. R^2	0.591	0.467	0.672	0.542

Note: This table reports the results for our test of the model's comparative statics with regard to competition. Columns (1)-(2) measures competition based on the prior year, while columns (3)-(4) based on the last 3 years piror to patent filing. Cumulative count is the count of past natural disasters. The Wald-tests examine if the coefficient for cumulative disaster count significantly differs between splits. We construct a Wald-test of the form $W = \frac{(\hat{\beta}_{\text{eq1}} - \hat{\beta}_{\text{eq2}})^2}{\text{Var}(\hat{\beta}_{\text{eq1}} - \hat{\beta}_{\text{eq2}})}$, where: $\text{Var}(\hat{\beta}_{\text{eq1}} - \hat{\beta}_{\text{eq2}}) = \text{Var}(\hat{\beta}_{\text{eq1}}) + \text{Var}(\hat{\beta}_{\text{eq2}}) - 2 \cdot \text{Cov}(\hat{\beta}_{\text{eq1}}, \hat{\beta}_{\text{eq2}})$. Standard errors are clustered at the region level and reported in parentheses. Significance levels: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

4.2.11 Alternative Green Good Window

In our baseline green product analysis (Section 1.6.2), the green product share in each industry is calculated as the average across the filing year and the year prior. To assess sensitivity to this choice, Table 4.10 re-estimates the greenness split using alternative time windows: a 1-year window (the filing year alone) and a 3-year window (the filing year and the two preceding years).

The results are stable across specifications. Across all windows, inventors in industries with above-median green product shares respond more strongly to natural disaster exposure than those in less green industries. The coefficients remain positive and highly significant for both groups, with larger point estimates consistently observed in the above-median sample. Point estimates are close to the baseline (Table 1.7, columns 3 and 4) across both windows. The Wald test p-value for the 1-year window (0.0516^{*}) is slightly larger than in the baseline (0.039^{**}), while the 3-year window (0.048^{**}) remains below the 5% threshold. Overall, the results confirm that our findings are not sensitive to the specific choice of window for measuring green market size.

Table 4.10: Green Product Split by Median

	Dependent variable:			
	$P(Y02_{lt})$			
Greenness Cutoff:	1 Year Window Above Median	1 Year Window Below Median	3 Year Window Above Median	3 Year Window Below Median
	(1)	(2)	(3)	(4)
Cumulative Count	0.086^{***}	0.064^{***}	0.083^{***}	0.060^{***}
	(0.011)	(0.009)	(0.011)	(0.009)
Country-Year F.E.	Yes	Yes	Yes	Yes
Region F.E.	Yes	Yes	Yes	Yes
Wald-test p-value:	0.0516		0.048	
Sample Mean	1.2949	1.287	1.3715	1.3617
Observations	15,060	15,060	13,554	13,554
R^2	0.611	0.775	0.635	0.798
Adj. R^2	0.589	0.762	0.612	0.785

Note: This table reports the results for our test of the model's comparative statics with regard to the size of the green good market. In columns (1) and (2), for each industry, we calculate the average green product share over the present year. In columns (3) and (4), for each industry, we calculate the average green product share over the last 3 years. Cumulative Count is the count of past natural disasters. Results are for the years 1995–2014 in columns (1) and (2) and for the years 1997–2014 in columns (3) and (4). We only have PRODCOM data starting in 1995, so a 3-year window allows us to estimate results starting in 1997. We test the null hypothesis that the Disaster Count coefficient is larger for our sample of patents coming from above-median size green good markets than from patents than for our sample of patents coming from below-median size green good markets. We construct a Wald-test of the form $W = \frac{(\hat{\beta}_{\text{eq1}} - \hat{\beta}_{\text{eq2}})^2}{\text{Var}(\hat{\beta}_{\text{eq1}} - \hat{\beta}_{\text{eq2}})}$, where: $\text{Var}(\hat{\beta}_{\text{eq1}} - \hat{\beta}_{\text{eq2}}) = \text{Var}(\hat{\beta}_{\text{eq1}}) + \text{Var}(\hat{\beta}_{\text{eq2}}) - 2 \cdot \text{Cov}(\hat{\beta}_{\text{eq1}}, \hat{\beta}_{\text{eq2}})$. We can reject the Null hypothesis $H_0 : \beta_h = \beta_l$ against the alternative ($H_1 : \beta_h > \beta_l$) with the reported p value. Standard errors are clustered on the region level and are reported in parentheses. P-values are as follows: *p<0.1; **p<0.05; ***p<0.01

4.2.12 First Time & Repeat Inventors

Table 4.11 shows our results when we split the sample of patents based on an inventor having previously patented or not. Both repeat and first-time inventors react to natural disaster exposure, with the effects being significantly larger for repeat inventors. Overall, there are more inventors that have previously patented than first time inventors in our sample. We denote a patent as coming from a repeat inventor if at least one of the inventors on that patent has previously patented.

Table 4.11: Response by First Time & Repeat Inventors

	Dependent variable:	
	$P(Y02_{lt})$	
	First-Time	Repeat Inventors
	(1)	(2)
Cumulative Count	0.049***	0.095***
	(0.016)	(0.010)
Country-Year F.E.	Yes	Yes
Region F.E.	Yes	Yes
P Value	0.0119	
Sample Means	0.1682	2.0954
Observations	15,813	15,813
R^2	0.344	0.702
Adj. R^2	0.309	0.686

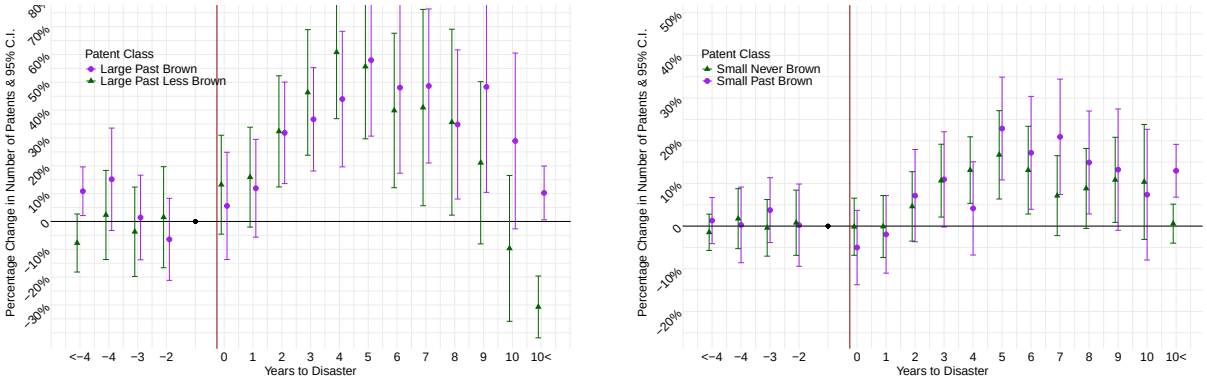
Note: This table reports our results when we split the sample of inventors into First-Time and Repeat inventors. Cumulative Count is the count of past natural disasters. We test the null hypothesis that the Disaster Count coefficient is different for our sample of first-time inventors than for our sample of repeat inventors. We construct a Wald-test of the form $W = \frac{(\hat{\beta}_{\text{eq1}} - \hat{\beta}_{\text{eq2}})^2}{\text{Var}(\hat{\beta}_{\text{eq1}} - \hat{\beta}_{\text{eq2}})}$, where: $\text{Var}(\hat{\beta}_{\text{eq1}} - \hat{\beta}_{\text{eq2}}) = \text{Var}(\hat{\beta}_{\text{eq1}}) + \text{Var}(\hat{\beta}_{\text{eq2}}) - 2 \cdot \text{Cov}(\hat{\beta}_{\text{eq1}}, \hat{\beta}_{\text{eq2}})$. We can reject the Null hypothesis $H_0 : \beta_h = \beta_l$ against the alternative $(H_1 : \beta_h \neq \beta_l)$ with the reported p value. Standard errors are clustered on the region level and are reported in parentheses. P-values are as follows: *p<0.1; **p<0.05; ***p<0.01

Our findings in Table 4.11 indicate that the innovation response to natural disasters operates along both the extensive and intensive margins. We observe significant increases in green patenting among first-time inventors as well as among those with prior patenting experience, suggesting that disasters not only induce additional activity from established innovators but also bring new inventors into the green innovation space. This broad-based response reinforces the main result that environmental shocks stimulate green technological effort across the innovation spectrum.

4.2.13 Past Patenting

In this section of the appendix, we examine whether the baseline effect of natural disasters on green innovation varies with inventors' patenting histories. To do so, we split our sample along

two dimensions. First, we classify patents based on the share of green patents previously filed by their inventors or firms. However, many patent holders in our dataset have no prior patent history, preventing us from directly calculating their past green patent shares. To address this limitation, we introduce a second classification based on the total number of patents previously filed by inventors. For inventors in the top quartile of past patent filings, we can reliably calculate their share of past green patents. For inventors below the top quartile, we instead check whether any inventor on a patent previously filed a clearly defined "brown" patent, one associated explicitly with environmentally harmful technologies. We adopt this definition of "brown" patents from Dechezleprêtre et al. (2021).



(a) Effects for Large Patent Filers

(b) Effects for Small Patent Filers

Figure 4.9: Green innovation response to natural disaster by past patenting

Note: This figure shows the effects of natural disaster exposure on green patenting, split by inventors' past patenting patterns. Panel (a) focuses on inventors in the top quartile of past patent filings. Within this group, we distinguish between those whose past portfolios contain an above-median share of brown patents ("Large Past Brown") and those with a below-median share ("Large Past Less Brown"). Both subgroups file 0.26 green patents per region on average. Panel (b) focuses on inventors below the top quartile ("Small Patent Filers"), distinguishing between those who have previously filed at least one brown patent ("Past Brown") and those who have never filed a brown patent ("Never Brown"). Sample averages are 0.67 and 0.6 green patents per region for "Never Brown" and "Past Brown," respectively. Brown patents are defined following the classification by Dechezleprêtre et al. (2021). Standard errors are clustered at the region level, and confidence intervals represent the 95% level.

The effects observed across these classifications indicate fairly similar responses for both "past brown" and "past less brown" inventor groups. However, among large patent filers (top quartile), inventors classified as "past less brown" exhibit somewhat less persistence in their response compared to their counterparts with extensive past brown patenting. In general, large past filers demonstrate stronger innovation responses following natural disaster exposure relative to smaller past filers.

4.3 Appendix - Data

4.3.1 Countries in the CompNet Dataset

Table 4.12 shows the time span for which different countries are available in the 9th vintage of the CompNet database (CompNet 2022). The information is directly taken from the CompNet website `https://www.comp-net.org/`.

Country	All firms	20e	Time Span
Belgium	**x**	**x**	2000 – 2020
Croatia	**x**	x	2002 – 2021
Czech Republic	**x**	**x**	2005 – 2020
Denmark	**x**	**x**	2001 – 2020
Finland	**x**	**x**	1999 – 2020
France	x	**x**	2003 – 2020
Germany		x	2001 – 2018
Hungary	**x**	**x**	2003 – 2020
Italy	x	x	2006 – 2020
Latvia*	**x**	**x**	2007 – 2019
Lithuania*	x	**x**	2000 – 2020
Malta	**x**	**x**	2010-2020
Netherlands	**x**	**x**	2007 – 2019
Poland		**x**	2002 – 2020
Portugal	**x**	**x**	2004 – 2020
Romania		**x**	2005 – 2020
Slovakia		**x**	2000 – 2020
Slovenia	**x**	**x**	2002 – 2021
Spain	**x**	**x**	2008 – 2020
Sweden	**x**	**x**	2003 – 2020
Switzerland	**x**	**x**	2009 – 2020
United Kingdom		x	1997 – 2019

Table 4.12: Comp Net TimpeSpans

4.3.2 Community Innovation Survey

We use data from the German part of the Community Innovation Survey, the Mannheimer Innovationspanel, administered by the ZEW – Leibniz Centre for European Economic Research (ZEW - Leibniz Centre for European Economic Research n.d.). Detailed information is available at: `https://www.zew.de/forschung/mannheimer-innovationspanel-innovationsaktivitaeten-der-deutschen-wirtschaft`.

The CIS is a biennial survey of firms in the manufacturing and service sectors, designed to monitor innovation activity across EU member states. For Germany, the survey is representative at the two-digit industry level and includes both SMEs and large firms. It does not track firms longitudinally, making it a repeated cross-section. We focus on the three survey waves that include questions on environmentally beneficial innovations: 2009, 2015, and 2021.

The dataset provides detailed information on innovation outputs, types of environmental benefits targeted, and motivations for adopting green innovation. For our analysis, we retain firms that report introducing an innovation with environmental benefits in the past two years.

To examine the motivations behind green innovation, we use responses to the following survey question:

> "During [*the last two years*], how important were the following factors in driving your enterprise's decisions to introduce innovations with environmental benefits?"

Responses are recorded on a four-point scale: "high," "medium," "low," or "not important." We construct binary indicators equal to one if the firm rated a given factor as "low," "medium," or "high," and zero if it was "not important." This inclusive definition captures all cases where the firm viewed the factor as at least somewhat relevant to its innovation decision.

Table 4.13 lists the variables and their corresponding survey items:

Table 4.13: Variable Definition - Factors driving green innovation

Variable	Corresponding survey questions
Expected demand	*Current or expected market demand for environmental innovation*
Expected regulatory changes	*Environmental regulations or taxes expected in the future*
Existing regulations	*Existing environmental regulations* OR *Existing environmental taxes, charges or fees*
Reputation	*Improving your enterprise's reputation*
Voluntary standards	*Voluntary actions or standards for environmental good practice within your sector*
Government funding	*Government grants, subsidies etc. for environmental innovations*

4.3.3 Table of Green Goods for PRODCOM

Table 4.14 gives the list of green goods we identify in PRODCOM (EUROSTAT 2025). The list is almost entirely based on Bontadini and Vona (2023), with a few minor additions. PRODCOM can be accesed on the EUROSTAT website: `https://ec.europa.eu/eurostat/web/prodcom/database`.

Table 4.14: Green Good Table

	PRODCOM Number	Label
1	24107500	Railway material (of steel)
2	25112200	Iron or steel towers and lattice masts
3	25301150	Vapour generating boilers (including hybrid boilers) (excluding central heating hot water boilers capable of producing low pressure steam, watertube boilers)
4	25301230	Auxiliary plant for use with boilers of HS 8402 or 8403
5	25301330	Parts of vapour generating boilers and super-heater water boilers
6	25991131	Sanitary ware and parts of sanitary ware of iron or steel
7	25992910	Railway or tramway track fixtures and fittings and parts thereof
8	26112220	Semiconductor light emitting diodes (LEDs)
9	26112240	Photosensitive semiconductor devices; solar cells, photodiodes, phototransistors, etc.
10	26121330	Multiple-walled insulating units of glass
11	26511200	Theodolites and tachymetres (tachometers); other surveying, hydrographic, oceanographic, hydrological, meteorological or geophysical instruments and appliances
12	26511215	Electronic rangefinders, theodolites, tacheometers and photogrammetrical instruments and appliances
13	26511235	Electronic instruments and apparatus for meteorological, hydrological and geophysical purposes (excluding compasses)
14	26511239	Other electronic instruments, n.e.c.
15	26511270	Surveying (including photogrammetrical surveying), hydrographic, oceanographic, hydrological, meteorological or geophysical instruments and appliances (excluding levels and compasses), non-electronic; rangefinders, non-electronic
16	26511280	Non electronic surveying (including photogrammatrical surveying), hydrographic, oceanographic, hydrological, meteorological or geophysical instruments and appliances (excluding rangefinders, levels and compasses),
17	26514100	Instruments and apparatus for measuring or detecting ionising radiations
18	26514200	Cathode-ray oscilloscopes and cathode-ray oscillographs
19	26514300	Instruments for measuring electrical quantities without a recording device
20	26514310	Multimeters without recording device
21	26514330	Electronic instruments and apparatus for measuring or checking voltage, current, resistance or electrical power, without recording device (excluding multimeters, and oscilloscopes and oscillographs)
22	26514355	Voltmeters without recording device

	PRODCOM Number	Label
23	26514359	Non-electronic instruments and apparatus, for measuring or checking voltage, current, resistance or power, without a recording device (excluding multimeters, voltmeters)
24	26514530	Instruments and apparatus, with a recording device, for measuring or checking electric gains (excluding gas, liquid or electricity supply or production meters)
25	26514555	Electronic instruments and apparatus, without a recording device, for measuring or checking electric gains (excluding gas, liquid or electricity supply or production meters)
26	26514559	Non-electronic instruments and apparatus, without a recording device, for measuring or checking electrical gains (excluding multimeters, voltmeters)
27	26515110	Thermometers, liquid-filled, for direct reading, not combined with other instruments (excluding clinical or veterinary thermometers)
28	26515135	Electronic thermometers and pyrometers, not combined with other instruments (excluding liquid filled)
29	26515139	Thermometers, not combined with other instruments and not liquid filled, n.e.c.
30	26515235	Electronic flow meters (excluding supply meters, hydrometric paddlewheels)
31	26515239	Electronic instruments and apparatus for measuring or checking the level of liquids
32	26515255	Non-electronic flow meters (excluding supply meters, hydrometric paddlewheels)
33	26515313	Electronic gas or smoke analysers
34	26515319	Non-electronic gas or smoke analysers
35	26515330	Spectrometers, spectrophotometers... using optical radiations
36	26515350	Instruments and apparatus using optical radiations, n.e.c.
37	26515381	Electronic ph and rh meters, other apparatus for measuring conductivity and electrochemical quantities (including use laboratory/field environment, use process monitoring/control)
38	26516350	Liquid supply or production meters (including calibrated) (excluding pumps)
39	26516370	Electricity supply or production meters (including calibrated) (excluding voltmeters, ammeters, wattmeters and the like)
40	26516500	Hydraulic or pneumatic automatic regulating or controlling instruments and apparatus
41	26516620	Test benches
42	26516650	Electronic instruments, appliances and machines for measuring or checking geometrical quantities (including comparators, coordinate measuring machines (CMMs))
43	26516683	Other instruments, appliances, for measuring or checking geometrical quantities

	PRODCOM Number	Label
44	26517015	Electronic thermostats
45	26517019	Non-electronic thermostats
46	26518200	Parts and accessories for the goods of 26.51.12, 26.51.32, 26.51.33, 26.51.4 and 26.51.5; microtomes; parts n.e.c.
47	26518550	Parts and accessories for automatic regulating or controlling instruments and apparatus
48	26702450	Other instruments and apparatus using optical radiation (UV, visible, IR)
49	26702490	Exposure meters, stroboscopes, optical instruments, appliances and machines for inspecting semiconductor wafers or devices or for inspecting photomasks or reticles used in manufacturing semiconductor devices, profile projectors and other optical instruments, appliances and machines for measuring or checking
50	27108230	Steel; iron or cast iron rails excl. current-conducting; with parts of non-ferrous metal - screws; bolts; nuts; rivets and spikes used for fixing track construction materials; assembled track
51	27108250	Iron or steel sleepers (crossties); rolled fish-plates and sole plates and check-rails (excl. screws; bolts; nuts; rivets and spikes used for fixing track construction materials)
52	27109230	Railway material (of steel)
53	27123130	Numerical control panels with built-in automatic data-processing machine for a voltage <= 1 kV
54	27123150	Programmable memory controllers for a voltage <= 1 kV
55	27123170	Other bases for electric control, distribution of electricity, voltage > 1000 V
56	27356200	Railway or tramway materials of steel or iron; not hot rolled
57	27401250	Tungsten halogen filament lamps for motorcycles and motor vehicles (excluding ultraviolet and infrared lamps)
58	27401293	Tungsten halogen filament lamps, for a voltage > 100 V (excluding ultraviolet and infra-red lamps, for motorcycles and motor vehicles)
59	27401295	Tungsten halogen filament lamps for a voltage <= 100 V (excluding ultraviolet and infrared lamps, for motorcycles and motor vehicles)
60	27401510	Fluorescent hot cathode discharge lamps, with double ended cap (excluding ultraviolet lamps)
61	27401530	Fluorescent hot cathode discharge lamps (excluding ultraviolet lamps, with double ended cap)
62	27402200	Electric table, desk, bedside, or floor-standing lamps
63	27403090	Electric lamps and lighting fittings, of plastic and other materials, of a kind used for filament lamps and tubular lamps, including lighting sets for Christmas trees
64	27403200	Lighting sets for Christmas trees

	PRODCOM Number	Label
65	27403930	Electric lamps and lighting fittings, of plastic and other materials, of a kind used for filament lamps and tubular fluorescent lamps
66	27512190	Other electromechanical appliances
67	27512690	Other electric space heaters
68	27521400	Non-electric instantaneous or storage water heaters
69	28112130	Steam turbines and other vapour turbines (excluding for electricity generation)
70	28112150	Steam turbines for electricity generation
71	28112160	Steam turbines and other vapour turbines
72	28112200	Iron or steel towers and lattice masts
73	28112400	Generating sets, wind-powered
74	28113100	Parts for steam turbines and other vapour turbines
75	28113200	Parts for hydraulic turbines and water wheels (including regulators)
76	28251130	Heat exchange units
77	28251380	Heat pumps other than air conditioning machines of HS 8415
78	28251410	Machinery and apparatus for filtering or purifying air (excluding intake filters for internal combustion engines)
79	28251420	Machinery and apparatus for filtering or purifying gases by a liquid process (excluding intake air filters for internal combustion engines, machinery and apparatus for filtering or purifying air)
80	28251430	Machinery and apparatus for filtering and purifying gases (other than air and excl. those which operate using a catalytic process, and isotope separators)
81	28251440	Machinery and apparatus for filtering or purifying gases by catalytic process (excluding intake air filters for internal combustion engines, machinery and apparatus for filtering or purifying air)
82	28251450	Machinery and apparatus for filtering and purifying gases with stainless steel housing, and with inlet and outlet tube bores with inside diameters not exceeding 1,3 cm (excluding intake filters for internal combustion engines)
83	28251470	Machinery and apparatus for filtering or purifying gases including for filtering dust from gases (excluding air filters for internal combustion engines, using liquid or catalytic process)
84	28291100	Producer gas or water gas generators; acetylene gas generators and the like; distilling or rectifying plant
85	28291230	Machinery and apparatus for filtering or purifying water
86	28291270	Machinery and apparatus for solid-liquid separation/ purification excluding for water and beverages, centrifuges and centrifugal dryers, oil/petrol filters for internal combustion engines
87	28298250	Parts for filtering and purifying machinery and apparatus, for liquids or gases (excluding for centrifuges and centrifugal dryers)

	PRODCOM Number	Label
88	28301150	Vapour generating boilers (including hybrid boilers) (excluding central heating hot water boilers capable of producing low pressure steam, watertube boilers)
89	28301230	Auxiliary plan for use with boilers of 84.02 or 84.03, used
90	28301330	Parts of vapour generating boilers and super-heater water boilers
91	28992020	Machines and apparatus used solely or principally for the manufacture of semiconductor boules or wafers
92	28992060	Machines and apparatus used solely or principally for the manufacture of flat panel displays
93	28993945	Machines and apparatus used solely or principally for a) the manufacture or repair of masks and reticles, b) assembling semiconductor devices or electronic integrated circuits, and c) lifting, handling, loading or unloading of boules, wafers, semiconductor devices, electronic integrated circuits and flat panel displays
94	29102400	Other motor vehicles for the transport of persons (excluding vehicles for transporting >=10 persons, snowmobiles, golf cars and similar vehicles)
95	29102410	Motor vehicles, with both spark-ignition or compression-ignition internal combustion reciprocating piston engine and electric motor as motors for propulsion, other than those capable of being charged by plugging to external source of electric power
96	29102430	Motor vehicles, with both spark-ignition or compression-ignition internal combustion reciprocating piston engine and electric motor as motors for propulsion, capable of being charged by plugging to external source of electric power
97	29102450	Motor vehicles, with only electric motor for propulsion
98	29102490	Other motor vehicles for the transport of persons (excluding vehicles with only electric motor for propulsion , vehicles for transporting > 10 persons, snowmobiles, golf cars and similar vehicles)
99	29105200	Motor vehicles specially designed for travelling on snow, golf cars and similar vehicles
100	29112130	Steam turbines and other vapour turbines (excl. for electricity generation)
101	29112150	Steam turbines for generation of electricity
102	29112200	Hydraulic turbines and water wheels
103	29113100	Parts for steam turbines and other vapour turbines
104	29113200	Parts of hydraulic turbines; water wheels incl. regulators
105	29231375	Absorption heat pumps
106	29231380	Heat pumps other than air conditioning machines of HS 8415
107	29231410	Machinery and apparatus for filtering or purifying air
108	29231420	Machinery and apparatus for filtering or purifying gases by a liquid process excl. intake air filters for internal combustion engines; machinery and apparatus for filtering or purifying air

	PRODCOM Number	Label
109	29231430	Machinery filtering or purifying gases; by electrostatic process
110	29231440	Machinery and apparatus for filtering/purifying gases by catalytic process excluding intake air filters for internal combustion engines, machinery and apparatus for filtering/purifying air
111	29231450	Machinery filtering or purifying gases; by thermic process
112	29231460	Machinery filtering or purifying gases; other
113	29231470	Machinery filtering or purifying gases
114	29241130	Producer gas or water gas generators, acetylene and similar water process gas generators
115	29241150	Distilling or rectifying plant
116	29241230	Machinery and apparatus for filtering/ purifying water
117	29241270	Machinery and apparatus for filtering/ purifying liquids; for chemical industry
118	29245250	Parts for filtering and purifying machinery and apparatus, for liquids or gases (excluding for centrifuges and centrifugal dryers)
119	29562582	Machines and apparatus used solely or principally for the manufacture of semiconductor boules or wafers
120	29562586	Machines and apparatus used solely or principally for the manufacture of flat panel displays
121	29562588	Machines and apparatus used solely or principally for a) the manufacture or repair of masks and reticles, b) assembling semiconductor devices or electronic integrated circuits, and c) lifting, handling, loading or unloading of boules, wafers, semiconductors.
122	29721400	Instantaneuous water heater apparatus non-electric
123	30201100	Rail locomotives powered from an external source of electricity
124	30201200	Diesel-electric locomotives
125	30201300	Other rail locomotives; locomotive tenders
126	30202000	Self-propelled railway or tramway coaches, vans and trucks, except maintenance or service vehicles
127	30203100	Railway or tramway maintenance or service vehicles (including workshops, cranes, ballast tampers, track-liners, testing coaches and track inspection vehicles)
128	30203200	Rail/tramway passenger coaches; luggage vans, post office coaches and other special purpose rail/tramway coaches excluding rail/tramway maintenance/service vehicles, self-propelled
129	30203300	Railway or tramway goods vans and wagons, not self-propelled
130	30204030	Parts of locomotives or rolling stock
131	30921000	Bicycles and other cycles (incl. delivery tricycles), non-motorized
132	30921030	Non-motorized bicycles and other cycles, without ball bearings (including delivery tricycles)

	PRODCOM Number	Label
133	30921050	Non-motorized bicycles and other cycles with ball bearings (including delivery tricycles)
134	30923010	Frames and forks, for bicycles
135	30923030	Parts of frames, front forks, brakes, coaster braking hubs, hub brakes, pedals crank-gear and free-wheel sprocket-wheels for bicycles, other non-motorized cycles and sidecars
136	30923060	Parts and accessories of bicycles and other cycles, not motorised (excl. frames and front forks).
137	30923090	Other parts and accessories of bicycles and other cycles, not motorised
138	31203150	Programmable memory controllers; voltage <= 1000 V
139	31203170	Meter mounting boards and installation panels; voltage <= 1000 V
140	31501230	Tungsten halogen filament lamps (excl. ultra-violet; infra-red): for projectors
141	31501250	Tungsten halogen filament lamps for motorcycles and motor vehicles (excl. ultraviolet and infrared lamps)
142	31501293	Tungsten halogen filament lamps; for a voltage > 100 V (excl. ultraviolet and infra-red lamps; for motorcycles and motor vehicles)
143	31501295	Other tungsten halogen lamps; <= 100 V
144	31501510	Fluorescent hot cathode discharge lamps, with double ended cap (excluding ultraviolet lamps)
145	31501530	Fluorescent hot cathode discharge lamps (excl. ultraviolet lamps, with double ended cap)
146	31502200	Electric table; desk; bedside or floor-standing lamps
147	31503430	Electric lamps and lighting fittings, of plastic and other materials, of a kind used for filament lamps and tubular fluorescent lamps
148	32105235	Semiconductor light emitting diodes (LEDs)
149	32105237	Photosensitive semiconductor devices; solar cells, photodiodes, phototransistors, etc.
150	33201215	Electronic surveying & hydrographic instr.& appliances (incl. rangefinders; levels; theodolites & tacheometers; photogrammetrical instr.& appliances; excl. compasses)
151	33201219	Non-electronic surveying, hydrographic instr. and appliances (including rangefinders, levels, theodolites and tacheometers, photogrammetrical instr. and appliances; excluding compasses)
152	33201235	Electronic instruments and apparatus for meteorological, hydrological and geophysical purposes (excl. compasses)
153	33201253	Instruments and appliances used in geodesy; topography; surveying...
154	33201255	Non-electronic meteorological; hydrological and geophysical instruments and apparatus (excl. compasses)

	PRODCOM Number	Label
155	33201257	Non-electronic surveying, hydro-, oceanographic instr./appliances (excluding rangefinders, levels, theodolites, tacheometers, photogrammetrical instr./app., compasses)
156	33203900	Installation of other special-purpose machinery n.e.c.
157	33204100	Installation of medical and surgical equipment
158	33204200	Cathode-ray oscilloscopes and cathode-ray oscillographs
159	33204330	Instruments and apparatus, for measuring or checking voltage: electronic
160	33204355	Voltmeters
161	33204359	Instruments and apparatus; for measuring or checking voltage: others
162	33205119	Other thermometers, not with other instruments, liquid, for direct reading
163	33205135	Thermometers; not combined with other instruments and not liquid filled; electronic
164	33205139	Thermometers, not combined with other instruments and not liquid filled, n.e.c.
165	33205313	Electronic gas or smoke analysers
166	33205319	Non-electronic gas or smoke analysers
167	33205330	Spectrometers, spectrophotometers using optical radiations
168	33205340	Exposure meters
169	33205350	Instruments and apparatus using optical radiations; n.e.c.
170	33205381	Electronic ph & rh meters; other apparatus for measuring conductivity & electrochemical quantities (incl. use laboratory/field environment; use process monitoring/control)
171	33205385	Viscometers, porosimeters and expansion meters
172	33205389	Other instruments and apparatus for physical and chemical analysis
173	33206350	Liquid supply or production meters (incl. calibrated) (excl. pumps)
174	33206370	Electricity supply or production meters (incl. calibrated) (excl. voltmeters; ammeters; wattmeters and the like)
175	33206550	Electronic instruments...measuring; checking geometrical quantities: 3 D
176	33206583	Other instruments, appliances, for measuring or checking geometrical quantities
177	33206589	Other instruments; appliances and machines for measuring or checking
178	33207015	Electronic thermostats
179	33207019	Non-electronic thermostats
180	33207050	Hydraulic or pneumatic automatic regulating or controlling instruments and apparatus
181	33208120	Parts and accessories for surveying, geodesy, topography, levelling, photogrammetrical, hydro-, oceanographic, hydro-, meteorological, geophysical instruments excl. compasses

	PRODCOM Number	Label
182	33208143	Parts and accessories for hydrometers and similar floating instruments, thermometers, pyrometers, barometers, hygrometers and psychrometers, recording or not, and any combination of these instruments
183	33208145	Parts and accessories of instruments and apparatus for measuring or checking the variables of liquids or gases (excl. for supply or production meters)
184	33208147	Microtomes, and parts and accessories
185	33209100	Installation of instruments and apparatus for measuring; checking; testing; navigating and other purposes
186	34102430	Vehicles with an electric motor, for the transport of persons (excl. vehicles for transporting >= 10 persons, snowmobiles, golf cars and similar vehicles)
187	34102490	Other motor vehicles for carrying people (excluding vehicles for transporting >= 10 persons, snowmobiles, golf cars and similar vehicles, electrically powered)
188	34105300	Vehicles for travelling on snow; golf cars; etc; with engines
189	35201100	Rail locomotives powered from an external source of electricity
190	35201200	Diesel-electric locomotives; =< 1000 kW power continuous rating
191	35201330	Rail locomotives powered by electric accumulators
192	35201390	Rail locomotives and locomotive tenders (excl. locomotives powered from an external source of electricity, locomotives powered by electric accumulators, diesel-electric locomotives)
193	35202030	Self-propelled railway coaches... powered by external electricity
194	35202090	Self-propelled railway or tramway coaches; vans and trucks; (diesel)
195	35203100	Railway or tramway maintenance or service vehicles (including workshops, cranes, ballast tampers, track-liners, testing coaches and track inspection vehicles)
196	35203200	Railway passenger coaches for speed =< 250 km/h; local
197	35203330	Tank wagons and the like; not self-propelled
198	35203350	Rail-or tramway goods vans & wagons; not self-propelled (incl. self-discharging and open vans & wagons) with non-removable sides; height > 60 cm; & other wagons
199	35204030	Parts of locomotives or rolling stock
200	35204055	Railway or tramway track fixtures and fittings, and mechanical or electromechanical signalling, safety or traffic control equipment
201	35204058	Parts of railway or tramway track fixtures and fittings; and for electromechanical signalling; safety or traffic control equipment
202	35204059	Mechanical (and electromechanical) signalling; safety or traffic control equipement (excluding equipment and material for track)
203	35421030	Bicycles and other cycles; not motorized; without ball bearings
204	35421050	Mountain bike

	PRODCOM Number	Label
205	35422013	Frames for bicycles, other non-motorized cycles and sidecars (excluding parts of frames)
206	35422015	Front forks for bicycles; other non-motorized cycles and sidecars (excl. parts of front forks)
207	35422019	parts of cycles
208	35422023	Wheel rims for bicycles other non-motorized cycles and sidecars
209	35422025	Wheel spokes for bicycles; other non-motorized cycles and sidecars
210	35422027	Hubs without free-wheel or braking device for bicycles, other non-motorized cycles and sidecars
211	35422033	Coaster braking hubs and hub brakes
212	35422039	Brakes for bicycles and other non-motorized cycles (excl. coaster braking hubs and hub brakes)
213	35422040	Saddles for bicycles and other non-motorized cycles
214	35422053	Pedals
215	35422055	Crank-gear
216	35422063	Handlebars
217	35422065	Luggage-carriers for bicycles and other non-motorized cycles
218	35422067	Derailleur gears for bicycles and other non-motorized cycles
219	35431200	Parts and accessories of invalid carriages
220	40301003	Heat - heating plants (heat produced by heating plants using fossil fuels; biomass or waste; sold to third parties)
221	40301005	Heat - geothermal (heat produced in geothermal fields; sold to third parties)
222	23121330	Multiple-walled insulating units of glass
226	28112160	Steam turbines and other vapour turbines (excl. for electricity generation)
227	28112200	Hydraulic turbines and water wheels
229	28113200	Parts of hydraulic turbines; water wheels incl. regulators
231	28251410	Machinery and apparatus for filtering or purifying air
232	28251441	Machinery and apparatus for filtering/purifying gases by catalytic process excluding intake air filters for internal combustion engines, machinery and apparatus for filtering/purifying air
233	28291100	Distilling or rectifying plant
234	28298251	Parts for filtering and purifying machinery and apparatus, for liquids or gases (excluding for centrifuges and centrifugal dryers)
237	27123150	Programmable memory controllers; voltage <= 1000 V
240	26516370	Voltmeters
244	26702490	Exposure meters
248	26515175	Parts and accessories for hydrometers and similar floating instruments, thermometers, pyrometers, barometers, hygrometers and psychrometers, recording or not, and any combination of these instruments

	PRODCOM Number	Label
250	30201200	Diesel-electric locomotives; =< 1000 kW power continuous rating
251	30203100	Self-propelled railway or tramway coaches; vans and trucks; (diesel)
252	30921000	Bicycles and other cycles (including delivery tricycles), non-motorised
253	30923060	Bicycles and other cycles, not motorised, with ball bearings.
254	30923010	Frames for bicycles, other non-motorized cycles and sidecars (excluding parts of frames)
255	22111200	Wheel rims for bicycles other non-motorized cycles and sidecars

Chapter 5

Appendix for Chapter 2

5.1 Basic Setup

5.1.1 Lagrangian and First-Order Conditions

The Lagrangian for the optimization problem of the social planner reads:

$$\begin{aligned}
\mathcal{L} \;=\; & \alpha(g_l, b_l) - \theta_l\,\phi(\pi_l b_l + \pi_h b_h) - k(y_l, \omega_l) \\
& + \lambda\Big(\alpha(g_h, b_h) - \theta_h\,\phi(\pi_l b_l + \pi_h b_h) - k(y_h, \omega_h) - \overline{V}_h\Big) \\
& + \psi\Big(\pi_l y_l + \pi_h y_h - \bar{R} - \pi_l(g_l + b_l) - \pi_h(g_h + b_h)\Big) \\
& + \kappa_l\Big(\alpha(g_l, b_l) - \theta_l\,\phi(\pi_l b_l + \pi_h b_h) - k(y_l, \omega_l) \\
& \qquad - \big[\alpha(g_h, b_h) - \theta_l\,\phi(\pi_l b_l + \pi_h b_h) - k(y_h, \omega_l)\big]\Big) \\
& + \kappa_h\Big(\alpha(g_h, b_h) - \theta_h\,\phi(\pi_l b_l + \pi_h b_h) - k(y_h, \omega_h) \\
& \qquad - \big[\alpha(g_l, b_l) - \theta_h\,\phi(\pi_l b_l + \pi_h b_h) - k(y_l, \omega_h)\big]\Big).
\end{aligned} \tag{5.1}$$

Note that the externality terms $\phi(\pi_l b_l + \pi_h b_h)$ cancel within each incentive constraint, since aggregate pollution is the same regardless of which bundle an agent claims. The first-order conditions are:

$$\frac{\partial \mathcal{L}}{\partial g_l} \;=\; (1 + \kappa_l - \kappa_h)\,\alpha_g(g_l, b_l) \;-\; \psi\,\pi_l \;=\; 0 \tag{5.2}$$

$$\frac{\partial \mathcal{L}}{\partial g_h} \;=\; (\lambda + \kappa_h - \kappa_l)\,\alpha_g(g_h, b_h) \;-\; \psi\,\pi_h \;=\; 0 \tag{5.3}$$

$$\frac{\partial \mathcal{L}}{\partial b_l} \;=\; (1 + \kappa_l - \kappa_h)\,\alpha_b(g_l, b_l) \;-\; (\theta_l + \lambda\,\theta_h)\,\phi'\big(\pi_l b_l + \pi_h b_h\big)\,\pi_l \;-\; \psi\,\pi_l \;=\; 0 \tag{5.4}$$

$$\frac{\partial \mathcal{L}}{\partial b_h} \;=\; (\lambda + \kappa_h - \kappa_l)\,\alpha_b(g_h, b_h) \;-\; (\theta_l + \lambda\,\theta_h)\,\phi'\big(\pi_l b_l + \pi_h b_h\big)\,\pi_h \;-\; \psi\,\pi_h \;=\; 0 \tag{5.5}$$

$$\frac{\partial \mathcal{L}}{\partial y_l} \;=\; -(1 + \kappa_l)\,k_y(y_l, \omega_l) \;+\; \kappa_h\,k_y(y_l, \omega_h) \;+\; \psi\,\pi_l \;=\; 0 \tag{5.6}$$

$$\frac{\partial \mathcal{L}}{\partial y_h} \;=\; -(\lambda + \kappa_h)\,k_y(y_h, \omega_h) \;+\; \kappa_l\,k_y(y_h, \omega_l) \;+\; \psi\,\pi_h \;=\; 0 \tag{5.7}$$

Note that the consumption FOCs (5.2)–(5.5) carry the symmetric coefficient $(1 + \kappa_l - \kappa_h)$ or $(\lambda + \kappa_h - \kappa_l)$ on both α_g and α_b. By contrast, the labor FOCs (5.6)–(5.7) involve effort costs evaluated at different ability levels, since a mimicking agent retains her own productivity. This means that $\alpha_g(g_i, b_i) = k_y(y_i, \omega_i)$ holds for both types in the first-best region ($\kappa_l = \kappa_h = 0$), for type l in the LIC region ($\kappa_l > 0$, $\kappa_h = 0$), and for type h in the HIC region ($\kappa_h > 0$, $\kappa_l = 0$), but not for the mimicked type in either incentive-binding region.

5.1.2 Proof of Lemma 1

Lemma 1: At any point of the Pareto frontier, at most a single incentive constraint is binding.

Proof. Since the externality terms cancel within each incentive constraint, the two ICs reduce to:

$$\text{IC}_l: \quad \alpha(g_l, b_l) - k(y_l, \omega_l) \;\geq\; \alpha(g_h, b_h) - k(y_h, \omega_l) \tag{5.8}$$

$$\text{IC}_h: \quad \alpha(g_h, b_h) - k(y_h, \omega_h) \;\geq\; \alpha(g_l, b_l) - k(y_l, \omega_h) \tag{5.9}$$

Suppose both bind simultaneously. Adding (5.8) and (5.9) with equality yields:

$$k(y_h, \omega_l) - k(y_l, \omega_l) + k(y_l, \omega_h) - k(y_h, \omega_h) = 0. \tag{5.10}$$

We can rewrite (5.10) as:

$$\int_{y_l}^{y_h} k_y(s, \omega_l)\,ds - \int_{y_l}^{y_h} k_y(s, \omega_h)\,ds = -\int_{y_l}^{y_h}\int_{\omega_l}^{\omega_h} k_{y\omega}(s, z)\,dz\,ds = 0. \tag{5.11}$$

By the single-crossing property ($k_{y\omega} < 0$), the integrand is strictly positive. The integral therefore equals zero only if $y_l = y_h \equiv y$.

Substituting $y_l = y_h = y$ into the binding ICs, both reduce to:

$$\alpha(g_l, b_l) - k(y, \omega_l) = \alpha(g_h, b_h) - k(y, \omega_l), \tag{5.12}$$

which gives $\alpha(g_l, b_l) = \alpha(g_h, b_h)$. From the consumption FOCs (5.2)–(5.5), both types face the same marginal rate of substitution. With strictly concave α, the same utility level and the same MRS pin down a unique consumption bundle, so $(g_l, b_l) = (g_h, b_h)$.

Now all allocations are identical across types. The g and y FOCs require:

$$(1 + \kappa_l - \kappa_h)\,\alpha_g = \psi\,\pi_l \tag{5.13}$$

$$(\lambda + \kappa_h - \kappa_l)\,\alpha_g = \psi\,\pi_h \tag{5.14}$$

$$(1 + \kappa_l)\,k_y(y, \omega_l) - \kappa_h\,k_y(y, \omega_h) = \psi\,\pi_l \tag{5.15}$$

$$(\lambda + \kappa_h)\,k_y(y, \omega_h) - \kappa_l\,k_y(y, \omega_l) = \psi\,\pi_h \tag{5.16}$$

Combining (5.13) with (5.15):

$$(1 + \kappa_l - \kappa_h)\,\alpha_g = (1 + \kappa_l)\,k_y(y, \omega_l) - \kappa_h\,k_y(y, \omega_h). \tag{5.17}$$

Combining (5.14) with (5.16):

$$(\lambda + \kappa_h - \kappa_l)\,\alpha_g = (\lambda + \kappa_h)\,k_y(y, \omega_h) - \kappa_l\,k_y(y, \omega_l). \tag{5.18}$$

Solving (5.17) for α_g:

$$\alpha_g = k_y(y, \omega_l) + \kappa_h\,\frac{k_y(y, \omega_l) - k_y(y, \omega_h)}{1 + \kappa_l - \kappa_h}. \tag{5.19}$$

Solving (5.18) for α_g:

$$\alpha_g = k_y(y, \omega_h) + \kappa_l\,\frac{k_y(y, \omega_h) - k_y(y, \omega_l)}{\lambda + \kappa_h - \kappa_l}. \tag{5.20}$$

Setting these equal and letting $D \equiv k_y(y, \omega_l) - k_y(y, \omega_h)$:

$$D + \kappa_h\,\frac{D}{1 + \kappa_l - \kappa_h} + \kappa_l\,\frac{D}{\lambda + \kappa_h - \kappa_l} = 0. \tag{5.21}$$

Since $k_{y\omega} < 0$ and $\omega_h > \omega_l$, we have $D > 0$. Dividing by D:

$$1 + \frac{\kappa_h}{1 + \kappa_l - \kappa_h} + \frac{\kappa_l}{\lambda + \kappa_h - \kappa_l} = 0. \tag{5.22}$$

All three terms are non-negative (the denominators are positive by complementary slackness and the structure of the problem), so the left-hand side is at least 1. This is a contradiction. Therefore both incentive constraints cannot bind simultaneously. □

5.1.3 Proof of Proposition 1

Proposition 1: If preferences are homogeneous, $\theta_l = \theta_h = \theta$, the optimal carbon tax is given by equation (2.12)*, independent of which incentive constraint is binding. The Pigouvian correction depends only on the common damage parameter, on the technology of pollution, and on the equilibrium consumption allocation. It does not depend on the redistributive structure of the optimum.*

Proof. From (2.4), the optimal carbon tax is $\tau = (\theta_l + \lambda\,\theta_h)\,\phi'/\psi$. From (5.2) and (5.3) we

obtain

$$\frac{1+\kappa_l-\kappa_h}{\psi}=\frac{\pi_l}{\alpha_g(g_l,b_l)},\qquad \frac{\lambda+\kappa_h-\kappa_l}{\psi}=\frac{\pi_h}{\alpha_g(g_h,b_h)}. \tag{5.23}$$

Adding these two expressions yields

$$\frac{1+\lambda}{\psi}=\frac{\pi_l}{\alpha_g(g_l,b_l)}+\frac{\pi_h}{\alpha_g(g_h,b_h)}. \tag{5.24}$$

Now set $\theta_l=\theta_h=\theta$. The carbon tax becomes

$$\tau=\frac{(1+\lambda)\,\theta\,\phi'}{\psi}=\theta\,\phi'\big(\pi_l b_l+\pi_h b_h\big)\left[\frac{\pi_l}{\alpha_g(g_l,b_l)}+\frac{\pi_h}{\alpha_g(g_h,b_h)}\right]. \tag{5.25}$$

This expression contains no incentive constraint multipliers κ_l or κ_h. The Pigouvian correction therefore takes the same form regardless of which incentive constraint binds. It equals the externality damage weighted by the marginal utilities of consumption, with no dependence on the redistributive structure of the optimum.

The intuition is as follows. Under homogeneous preferences, $\theta_l=\theta_h=\theta$, the externality term enters both individuals' utilities in the same way,

$$U_i=\tilde{U}_i(g_i,b_i,y_i)-\theta\phi(B),\qquad \tilde{U}_i(g_i,b_i,y_i):=\alpha(g_i,b_i)-k(y_i,\omega_i).$$

Aggregate brown consumption B is the same whether an agent reports truthfully or mimics another type. Hence the term $-\theta\phi(B)$ appears on both sides of each incentive constraint and cancels out. The incentive constraints depend only on the standard Mirrleesian trade-off between consumption utility and labor disutility and are independent of the externality.

Because the externality affects both types symmetrically, it constitutes a common utility component. Redistribution across types can therefore be implemented entirely through the nonlinear income tax. The corrective tax component has no redistributive role and serves only to internalize the externality. In the spirit of the Atkinson–Stiglitz result, the Pigouvian correction is determined solely by the common damage parameter $\theta\phi'(B)$ and the consumption allocation, with no dependence on the incentive constraint multipliers. □

5.1.4 Proof of Lemma 2.2

Lemma 2.2: Under Assumption 2.4, aggregate brown consumption $B(\lambda)$ is strictly increasing in λ on the HIC region of the Pareto frontier.

Proof. We show that aggregate income $Y(\lambda)\equiv\pi_l y_l+\pi_h y_h$ is strictly increasing in λ on the HIC region, and then translate this into a statement about $B(\lambda)$ using the resource constraint.

In the HIC region, $\kappa_l=0$ and $\kappa_h>0$. The labor first-order conditions (5.6) and (5.7) reduce to

$$k_y(y_l,\omega_l)-\kappa_h\,k_y(y_l,\omega_h)=\psi\,\pi_l, \tag{5.26}$$

$$(\lambda+\kappa_h)\,k_y(y_h,\omega_h)=\psi\,\pi_h. \tag{5.27}$$

The relation between λ and κ_h in HIC is pinned down by the consumption first-order conditions. Adding (5.2) and (5.3) (or, equivalently, by Walras' law on the resource constraint) yields the standard identity $\lambda = (\pi_h - \kappa_h)/\pi_l$, so

$$\kappa_h \;=\; \pi_h - \pi_l\,\lambda, \qquad 1 - \kappa_h \;=\; \pi_l(1+\lambda), \qquad \lambda + \kappa_h \;=\; \pi_h(1+\lambda). \tag{5.28}$$

The HIC region corresponds to $\kappa_h \in (0, \pi_h)$, i.e. $\lambda \in (0, \pi_h/\pi_l)$.

The high type is the unconstrained type in HIC. Dividing (5.27) by π_h and using $\lambda + \kappa_h = \pi_h(1+\lambda)$ together with the consumption FOC (5.3) evaluated under homogeneous θ yields $k_y(y_h, \omega_h) = \alpha_g(g_h, b_h)$. In our setting with CES consumption, the equilibrium income of the high type satisfies the undistorted condition $k_y(y_h, \omega_h) = 1$ on the consumer's optimization (the additivity property of Cremer et al. 1998, Pirttilä and Tuomala 2002, and Kaplow 2012; the Pigouvian wedge is borne entirely by the price of brown consumption, not the labor margin). Under Assumption 2.4 this delivers

$$y_h \;=\; \omega_h^{1+\varepsilon}, \qquad \text{constant in } \lambda \text{ on the HIC region.} \tag{5.29}$$

The low type is the constrained type. Substituting (5.28) into (5.26) and using the iso-elastic form,

$$\frac{1}{\omega_l}\left(\frac{y_l}{\omega_l}\right)^{1/\varepsilon} - \kappa_h \cdot \frac{1}{\omega_h}\left(\frac{y_l}{\omega_h}\right)^{1/\varepsilon} \;=\; \psi\,\pi_l. \tag{5.30}$$

Factoring out $(y_l)^{1/\varepsilon}$ and writing $r \equiv \omega_l/\omega_h$,

$$(y_l)^{1/\varepsilon}\,\omega_l^{-(1+1/\varepsilon)}\left[1 - \kappa_h\, r^{1+1/\varepsilon}\right] \;=\; \psi\,\pi_l. \tag{5.31}$$

Combining with the corresponding expression for ψ from the consumption FOCs ((5.2)) and using $k_y(y_l, \omega_l) = \alpha_g(g_l, b_l)\cdot A_B$ for an appropriate distortion wedge $A_B = (1-\kappa_h)/(1-\kappa_h r^{1+1/\varepsilon})$, the low type's income solves

$$y_l \;=\; \omega_l^{1+\varepsilon}\left[A_B(\lambda)\right]^{\varepsilon}, \qquad A_B(\lambda) \;\equiv\; \frac{1-\kappa_h(\lambda)}{1-\kappa_h(\lambda)\,r^{1+1/\varepsilon}}. \tag{5.32}$$

Since $r = \omega_l/\omega_h < 1$ and $1 + 1/\varepsilon > 1$, we have $r^{1+1/\varepsilon} < 1$, so the denominator of A_B is positive and strictly less than $1 - \kappa_h$ whenever $\kappa_h > 0$. Hence $A_B(\lambda) \geq 1$ on HIC, with equality only at the FB boundary where $\kappa_h = 0$.

To establish monotonicity of $y_l(\lambda)$, note that from (5.28), κ_h is strictly decreasing in λ on HIC. Differentiating A_B with respect to κ_h,

$$\frac{\partial A_B}{\partial \kappa_h} \;=\; \frac{-(1-\kappa_h\, r^{1+1/\varepsilon}) + (1-\kappa_h)\, r^{1+1/\varepsilon}}{\left(1-\kappa_h\, r^{1+1/\varepsilon}\right)^2} \;=\; \frac{r^{1+1/\varepsilon} - 1}{\left(1-\kappa_h\, r^{1+1/\varepsilon}\right)^2} \;<\; 0, \tag{5.33}$$

since $r^{1+1/\varepsilon} < 1$. Hence A_B is strictly decreasing in κ_h, and therefore strictly increasing in λ on HIC. From (5.32) and $\varepsilon > 0$, $y_l(\lambda)$ is strictly increasing in λ on HIC.

It remains to translate this into a statement about aggregate brown consumption. Aggregate

income on HIC is $Y(\lambda) = \pi_l\, y_l(\lambda) + \pi_h\, \omega_h^{1+\varepsilon}$. The first term is strictly increasing in λ by the argument above and the second is constant, so Y is strictly increasing in λ on HIC.

The aggregate resource constraint with $\bar{R} \geq 0$ implies $\pi_l(g_l + b_l) + \pi_h(g_h + b_h) = Y - \bar{R}$. Under CES consumption, both types share a common ratio $g/b = r(\tau)$, so $g_i + b_i = b_i(1 + r(\tau))$ and aggregate consumption decomposes as

$$\pi_l b_l + \pi_h b_h \;=\; \frac{Y - \bar{R}}{1 + r(\tau)} \;=\; \frac{Y(\lambda) - \bar{R}}{1 + r(\tau(\lambda))}. \tag{5.34}$$

Combining (2.15) with the strict monotonicity of Y on HIC, the mapping $\lambda \mapsto B(\lambda)$ is the composition of the strictly increasing map $\lambda \mapsto Y(\lambda)$ with a function determined jointly by $Y(\lambda)$ and $\tau(\lambda)$. We verify monotonicity directly. Suppose, toward contradiction, that $B(\lambda)$ is not strictly increasing in λ on HIC, so there exist $\lambda_1 < \lambda_2$ in HIC with $B(\lambda_1) \geq B(\lambda_2)$. By (2.15), $\tau(\lambda_1) \geq \tau(\lambda_2)$, hence $r(\tau(\lambda_1)) \geq r(\tau(\lambda_2))$ (since $r(\tau) = (a(1+\tau)/(1-a))^{\sigma}$ is strictly increasing in τ for $\sigma > 0$ and $a \in (0,1)$, as derived in Proposition 2.5). Therefore $1 + r(\tau(\lambda_1)) \geq 1 + r(\tau(\lambda_2))$, which combined with $Y(\lambda_1) < Y(\lambda_2)$ yields

$$B(\lambda_1) \;=\; \frac{Y(\lambda_1) - \bar{R}}{1 + r(\tau(\lambda_1))} \;<\; \frac{Y(\lambda_2) - \bar{R}}{1 + r(\tau(\lambda_2))} \;=\; B(\lambda_2), \tag{5.35}$$

contradicting $B(\lambda_1) \geq B(\lambda_2)$. Hence $B(\lambda)$ is strictly increasing in λ on HIC. □

Remark. The argument breaks the dependence chain by exploiting that $Y(\lambda)$ is determined directly by the labor margin and does not depend on τ. The CES consumption side then translates aggregate income into aggregate brown consumption, with the price ratio $r(\tau)$ playing an offsetting but not reversing role. The same argument does not extend mechanically to the LIC region, where the low type is undistorted and the high type's labor responds to κ_l. The direction of the comparative static in LIC depends on the relative strength of the corresponding income and substitution effects and is not pursued here.

5.1.5 Proof of Proposition 2.4

Proposition 2.4: Suppose $\theta_h \neq \theta_l$. Then the optimal carbon tax varies with welfare weights along the Pareto frontier.

Proof. We first derive the decomposition (2.6) used in the main text, then show that τ varies across regions.

Decomposition of τ. By (2.4), the optimal carbon tax in any region is $\tau = (\theta_l + \lambda\,\theta_h)\,\phi'/\psi$. From (5.3):

$$\lambda = \frac{\psi\,\pi_h}{\alpha_g(g_h, b_h)} + \kappa_l - \kappa_h. \tag{5.36}$$

Substituting into $\theta_l + \lambda\,\theta_h$ and dividing by ψ:

$$\frac{\theta_l + \lambda\,\theta_h}{\psi} = \frac{\theta_l + (\kappa_l - \kappa_h)\,\theta_h}{\psi} + \frac{\theta_h\,\pi_h}{\alpha_g(g_h, b_h)}. \tag{5.37}$$

Using (5.2) to eliminate ψ in the first term, $1/\psi = \pi_l/[(1+\kappa_l-\kappa_h)\,\alpha_g(g_l,b_l)]$, and multiplying by $\phi'(\pi_l b_l + \pi_h b_h)$ yields

$$\tau = \phi'\big(\pi_l b_l + \pi_h b_h\big)\left[\frac{[\theta_l + (\kappa_l-\kappa_h)\,\theta_h]\,\pi_l}{(1+\kappa_l-\kappa_h)\,\alpha_g(g_l,b_l)} + \frac{\theta_h\,\pi_h}{\alpha_g(g_h,b_h)}\right]. \tag{5.38}$$

Variation of τ across regions. Define

$$f(\kappa_l,\kappa_h) \;\equiv\; \frac{\theta_l + (\kappa_l-\kappa_h)\,\theta_h}{1+\kappa_l-\kappa_h}. \tag{5.39}$$

In the three regions:

$$\text{FB } (\kappa_l=\kappa_h=0): \quad f=\theta_l, \tag{5.40}$$

$$\text{LIC } (\kappa_l>0,\ \kappa_h=0): \quad f=\frac{\theta_l+\kappa_l\,\theta_h}{1+\kappa_l}, \tag{5.41}$$

$$\text{HIC } (\kappa_h>0,\ \kappa_l=0): \quad f=\frac{\theta_l-\kappa_h\,\theta_h}{1-\kappa_h}. \tag{5.42}$$

Direct computation gives:

$$f^{LIC}-f^{FB}=\frac{\kappa_l\,(\theta_h-\theta_l)}{1+\kappa_l}, \tag{5.43}$$

$$f^{FB}-f^{HIC}=\frac{\kappa_h\,(\theta_h-\theta_l)}{1-\kappa_h}. \tag{5.44}$$

Both differences are nonzero whenever $\theta_h \neq \theta_l$ and the corresponding multiplier is positive. They have the same sign as $\theta_h-\theta_l$.

At the FB–LIC boundary, $\kappa_l=0$, and the equilibrium allocations $(g_l,b_l,g_h,b_h,y_l,y_h)$ are continuous in $\overline{V}_h$. As κ_l becomes strictly positive, f moves discontinuously away from θ_l in the sense that $\partial f/\partial\kappa_l|_{\kappa_l=0}=\theta_h-\theta_l\neq 0$, while the allocations and hence the second summand $\theta_h\pi_h/\alpha_g(g_h,b_h)$ change continuously. Therefore the bracketed term in (5.38) changes by a nonzero amount, and τ is discontinuous in its dependence on κ_l across the boundary. The same argument applies at the FB–HIC boundary with $\partial f/\partial\kappa_h|_{\kappa_h=0}=-(\theta_h-\theta_l)\neq 0$.

Hence τ takes different values in the FB, LIC, and HIC regions whenever $\theta_h\neq\theta_l$. When $\theta_h=\theta_l$, both differences vanish and the carbon tax reduces to the invariant form of Proposition 2.2. □

Version of Proposition 2.4 with linear labor effort costs:

Proposition 5.1. *Suppose $\theta_h > \theta_l$ and labor effort costs are linear, $k(y_i, \omega_i) = y_i/\omega_i$, so that $k_y(y_i, \omega_i) = 1/\omega_i$. Then the optimal carbon tax varies along the Pareto frontier such that:*

$$\tau^{HIC} < \tau^{FB} < \tau^{LIC}$$

for $\kappa_l \geq 0$ and $1 > \kappa_h \geq 0$, and correspondingly $MRS^{HIC} < MRS^{FB} < MRS^{LIC}$.

Proof. The optimal carbon tax is given by equation (5):

$$\tau = \frac{(\theta_l + \lambda\theta_h)\phi'(B)}{\psi}. \tag{5.45}$$

We want to sign how this varies across regions of the Pareto frontier. The key challenge is that λ, ψ, and $\phi'(B)$ are all endogenous objects that vary with the equilibrium allocation. Under linear labor costs, the labor FOCs pin down ψ and λ purely in terms of multipliers and primitive parameters, eliminating the dependence on equilibrium allocations entirely. We proceed in four steps.

Step 1: Linear labor costs pin down ψ and λ.
Under $k_y(y_i, \omega_i) = 1/\omega_i$, the labor FOCs (33) and (34) become:

$$-(1 + \kappa_l)\frac{1}{\omega_l} + \kappa_h\frac{1}{\omega_h} + \psi\pi_l = 0, \tag{33'}$$

$$-(\lambda + \kappa_h)\frac{1}{\omega_h} + \kappa_l\frac{1}{\omega_l} + \psi\pi_h = 0. \tag{34'}$$

The endogenous income levels y_l and y_h have dropped out entirely. Solving (33') and (34'):

$$\psi\pi_l = \frac{1 + \kappa_l}{\omega_l} - \frac{\kappa_h}{\omega_h}, \tag{L1}$$

$$\psi\pi_h = \frac{\lambda + \kappa_h}{\omega_h} - \frac{\kappa_l}{\omega_l}. \tag{L2}$$

Dividing (L2) by (L1) and solving for λ:

$$\lambda = \frac{\pi_h}{\pi_l} \cdot \omega_h \left(\frac{1 + \kappa_l}{\omega_l} - \frac{\kappa_h}{\omega_h} \right) + \frac{\kappa_l\omega_h}{\omega_l} - \kappa_h. \tag{L3}$$

Both ψ and λ are now purely functions of primitive parameters $\{\omega_l, \omega_h, \pi_l, \pi_h\}$ and multipliers $\{\kappa_l, \kappa_h\}$. No equilibrium allocation appears.

Step 2: Compute $\Psi \equiv (\theta_l + \lambda\theta_h)/\psi$ in each region.
Define $\Psi \equiv (\theta_l + \lambda\theta_h)/\psi$, so that $\tau = \Psi \cdot \phi'(B)$. Substituting (L1) and (L3) into Ψ in each region:

First-Best ($\kappa_l = \kappa_h = 0$): $\psi^{FB} = 1/(\pi_l\omega_l)$ and $\lambda^{FB} = \pi_h\omega_h/(\pi_l\omega_l)$, giving:

$$\Psi^{FB} = \pi_l\omega_l\theta_l + \pi_h\omega_h\theta_h. \tag{FB}$$

LIC ($\kappa_l > 0$, $\kappa_h = 0$): $\psi^{LIC} = (1+\kappa_l)/(\pi_l\omega_l)$ and $\lambda^{LIC} = \pi_h\omega_h(1+\kappa_l)/(\pi_l\omega_l)+\kappa_l\omega_h/\omega_l$, giving:

$$\Psi^{LIC} = \frac{\pi_l\omega_l\theta_l + \kappa_l\pi_l\omega_h\theta_h}{1+\kappa_l} + \pi_h\omega_h\theta_h. \tag{LIC}$$

HIC ($\kappa_h > 0$, $\kappa_l = 0$): $\psi^{HIC} = 1/(\pi_l\omega_l) - \kappa_h/(\pi_l\omega_h)$ and $\lambda^{HIC} = \pi_h\omega_h/(\pi_l\omega_l) - \pi_h\kappa_h/\pi_l - \kappa_h$, giving:

$$\Psi^{HIC} = \frac{\pi_l\omega_l\theta_l + \pi_h\omega_h\theta_h - \kappa_h\omega_l\theta_h}{1-\kappa_h\omega_l/\omega_h}. \tag{HIC}$$

Step 3: Order Ψ across regions.
We invoke the following assumption:

Assumption 5.1 (D)**.** $\omega_h\theta_h > \omega_l\theta_l$.

Since $\theta_h > \theta_l$ and $\omega_h > \omega_l$ are both maintained throughout, Assumption D is automatically satisfied and requires no additional restriction beyond what is already assumed.

LIC vs FB. From (LIC) and (FB):

$$\Psi^{LIC} - \Psi^{FB} = \frac{\kappa_l\pi_l(\omega_h\theta_h - \omega_l\theta_l)}{1+\kappa_l} > 0, \tag{5.46}$$

since $\kappa_l > 0$ and $\omega_h\theta_h > \omega_l\theta_l$ by Assumption D. Therefore $\Psi^{LIC} > \Psi^{FB}$ strictly for all $\kappa_l > 0$.

HIC vs FB. Let $A \equiv \pi_l\omega_l\theta_l + \pi_h\omega_h\theta_h$ and $c \equiv \omega_l/\omega_h < 1$. From (HIC) and (FB):

$$\Psi^{FB} - \Psi^{HIC} = \frac{\kappa_h\omega_l\pi_l\,(\theta_h - \omega_l\theta_l/\omega_h)}{1-\kappa_h c}. \tag{5.47}$$

The denominator $1 - \kappa_h c > 0$ since $\kappa_h < 1$ and $c < 1$. The numerator is strictly positive if and only if $\theta_h > \omega_l\theta_l/\omega_h$, which rearranges to $\omega_h\theta_h > \omega_l\theta_l$, i.e. Assumption D. Therefore $\Psi^{FB} > \Psi^{HIC}$ strictly for all $\kappa_h > 0$.

Combining both comparisons: $\Psi^{HIC} < \Psi^{FB} < \Psi^{LIC}$.

Step 4: From Ψ to τ.
Recall $\tau = \Psi \cdot \phi'(B)$. The term $\phi'(B)$ still varies across regions since equilibrium brown consumption B differs. We invoke:

Assumption 5.2 (E)**.** *$\phi'(B)$ does not vary across regions so strongly as to reverse the ordering of Ψ. Formally, $\phi'(B^{LIC})/\phi'(B^{FB}) > \Psi^{FB}/\Psi^{LIC}$ and $\phi'(B^{HIC})/\phi'(B^{FB}) < \Psi^{FB}/\Psi^{HIC}$.*

Assumption E is satisfied trivially if ϕ is linear (ϕ' constant), which is a standard benchmark. More generally, since the LIC region corresponds to higher welfare weight on the high type who has higher environmental concern θ_h, the planner in the LIC region imposes a higher carbon tax, which reduces brown consumption $B^{LIC} \leq B^{FB}$. Under convex damages ($\phi'' > 0$), this would partially offset the higher Ψ^{LIC}. Assumption E requires that this offset is not complete. It is verified numerically at the calibrated parameter values underlying Figure 1.

Under Assumption E:

$$\tau^{HIC} < \tau^{FB} < \tau^{LIC}, \tag{5.48}$$

and since the MRS is strictly increasing in τ:

$$MRS^{HIC} < MRS^{FB} < MRS^{LIC}. \tag{5.49}$$

The transitions between regions are continuous: at the boundaries $\kappa_l = 0$ and $\kappa_h = 0$, allocations are identical on both sides and the ordering holds with equality, recovering continuity. The magnitude of the difference is increasing in $(\theta_h - \theta_l)$. When $\theta_h = \theta_l$, both differences in Ψ vanish and we recover Proposition 1.

□

Reduced version of proposition 2:

Proposition 5.2. *Suppose $\theta_h > \theta_l$. Define the corrective component of the carbon tax as:*

$$f(\kappa_l, \kappa_h) \equiv \frac{\theta_l + (\kappa_l - \kappa_h)\theta_h}{1 + \kappa_l - \kappa_h}. \tag{5.50}$$

Then f varies along the Pareto frontier such that:

$$f^{HIC} < f^{FB} < f^{LIC}$$

for $\kappa_l > 0$ and $1 > \kappa_h > 0$.

Proof. In the three regions of the Pareto frontier, f takes the following values.

First-Best ($\kappa_l = \kappa_h = 0$):

$$f^{FB} = \theta_l. \tag{5.51}$$

LIC ($\kappa_l > 0$, $\kappa_h = 0$):

$$f^{LIC} = \frac{\theta_l + \kappa_l \theta_h}{1 + \kappa_l}. \tag{5.52}$$

HIC ($\kappa_h > 0$, $\kappa_l = 0$):

$$f^{HIC} = \frac{\theta_l - \kappa_h \theta_h}{1 - \kappa_h}. \tag{5.53}$$

LIC vs FB.

$$f^{LIC} - f^{FB} = \frac{\theta_l + \kappa_l \theta_h}{1 + \kappa_l} - \theta_l = \frac{\kappa_l(\theta_h - \theta_l)}{1 + \kappa_l} > 0, \tag{5.54}$$

since $\kappa_l > 0$ and $\theta_h > \theta_l$. Therefore $f^{LIC} > f^{FB}$ strictly for all $\kappa_l > 0$.

HIC vs FB.

$$f^{FB} - f^{HIC} = \theta_l - \frac{\theta_l - \kappa_h \theta_h}{1 - \kappa_h} = \frac{\kappa_h(\theta_h - \theta_l)}{1 - \kappa_h} > 0, \tag{5.55}$$

since $1 > \kappa_h > 0$ and $\theta_h > \theta_l$. Therefore $f^{FB} > f^{HIC}$ strictly for all $1 > \kappa_h > 0$.

Combining both comparisons:

$$f^{HIC} < f^{FB} < f^{LIC}. \tag{5.56}$$

The transitions between regions are continuous: at the boundary between FB and LIC, $\kappa_l = 0$ and $f^{LIC} = f^{FB} = \theta_l$; at the boundary between FB and HIC, $\kappa_h = 0$ and $f^{HIC} = f^{FB} = \theta_l$. The magnitude of the difference is increasing in $(\theta_h - \theta_l)$ in both cases. When $\theta_h = \theta_l$, both differences vanish and f is constant along the entire Pareto frontier, consistent with Proposition 1.

Interpretation. The component $f(\kappa_l, \kappa_h)$ governs the weight placed on the low type's environmental damages in the corrective tax. From equation (6), the full carbon tax is:

$$\tau = \phi'(\pi_l b_l + \pi_h b_h) \left[\frac{f(\kappa_l, \kappa_h)\, \pi_l}{\alpha_g(g_l, b_l)} + \frac{\theta_h\, \pi_h}{\alpha_g(g_h, b_h)} \right]. \tag{5.57}$$

The ordering $f^{HIC} < f^{FB} < f^{LIC}$ therefore establishes that the corrective component of the tax is lowest when the planner favours the environmentally less-concerned low type (HIC region)

and highest when the planner favours the environmentally more-concerned high type (LIC region). Whether this ordering extends to the full carbon tax τ depends on the behaviour of $\alpha_g(g_l, b_l)$, $\alpha_g(g_h, b_h)$, and $\phi'(B)$ across regions. Under linear labor effort costs and Assumption E, Proposition 2 establishes that the ordering of f carries over to the full ordering of τ and the MRS.

□

5.1.6 Proof of Proposition 2.1

Proposition 2.1: Under Assumptions 2.1–2.2 and homogeneous preferences $\theta_l = \theta_h = \theta$, the optimal carbon tax is constant along the Pareto frontier and is the unique solution of $\tau\,\alpha_g^*(\tau) = \theta$.

Proof. By Proposition 2.2, the optimal carbon tax under homogeneous preferences is given by

$$\tau \;=\; \theta\,\phi'\big(\pi_l b_l + \pi_h b_h\big)\left[\frac{\pi_l}{\alpha_g(g_l, b_l)} + \frac{\pi_h}{\alpha_g(g_h, b_h)}\right], \tag{5.58}$$

which is independent of the incentive-constraint multipliers but still depends on the equilibrium allocations.

Under Assumption 2.2, the CES aggregator (2.10) is linearly homogeneous, so α_g is homogeneous of degree zero in (g, b) and depends only on the ratio g/b. From the MRS condition (2.4), both types face the same marginal rate of substitution at any point of the Pareto frontier. For CES this pins down a unique common ratio $r(\tau) = \big(a(1+\tau)/(1-a)\big)^{\sigma}$, and consequently $\alpha_g(g_l, b_l) = \alpha_g(g_h, b_h) \equiv \alpha_g^*(\tau)$, with the closed form (2.20) derived in the proof of Proposition 2.5. Substituting into (5.58) and using $\pi_l + \pi_h = 1$ yields

$$\tau \;=\; \frac{\theta\,\phi'\big(\pi_l b_l + \pi_h b_h\big)}{\alpha_g^*(\tau)}, \qquad \text{equivalently} \qquad \tau\,\alpha_g^*(\tau) \;=\; \theta\,\phi'\big(\pi_l b_l + \pi_h b_h\big). \tag{5.59}$$

The two type-specific denominators have collapsed into a single function of τ. The right-hand side still involves aggregate brown consumption $B = \pi_l b_l + \pi_h b_h$, which varies along the Pareto frontier as the equilibrium allocation adjusts.

Under Assumption 2.1, $\phi(B) = B$ and $\phi' \equiv 1$. The right-hand side of (5.59) reduces to the constant θ, and the implicit equation becomes

$$\tau\,\alpha_g^*(\tau) \;=\; \theta. \tag{5.60}$$

Aggregate brown consumption has dropped out. The equation contains only τ and the primitives (θ, a, σ).

It remains to verify uniqueness. As shown in the proof of Proposition 2.5, $H(\tau) \equiv \tau\,\alpha_g^*(\tau)$ is strictly increasing on $(-1, \infty)$ with $H(0) = 0$ and $H(\tau) \to \infty$ as $\tau \to \infty$. For any $\theta > 0$, the equation $H(\tau) = \theta$ therefore has exactly one positive solution. This solution is the same at every point of the Pareto frontier, in the HIC, FB, and LIC regions alike, because neither the incentive-constraint multipliers (κ_l, κ_h) nor any equilibrium object enters the determining equation.

The argument uses only the invariance of τ from Proposition 2.2, the linear homogeneity of the CES aggregator, and the constancy of ϕ'. At no point does it invoke properties of the effort cost function beyond those required for Proposition 2.2. The result therefore holds under iso-elastic costs $k(y,\omega) = (y/\omega)^{1+1/\varepsilon}/(1+1/\varepsilon)$, under linear costs $k(y,\omega) = y/\omega$, and under any other specification satisfying the regularity conditions stated in Section 2.2. □

5.2 Appendix - Dynamic Extension

5.2.1 Lagrangian and First-Order Conditions

The planner maximizes the low type's lifetime welfare subject to the participation constraint for the high type (multiplier λ), lifetime incentive constraints for both types (multipliers κ_l and κ_h), and per-phase resource constraints (multipliers ψ^B and ψ^G). The full Lagrangian is:

$$\begin{aligned}\mathcal{L} \;=\; & D_B\big(v_l^B + \lambda\, v_h^B\big) + D_G\big(v_l^G + \lambda\, v_h^G\big) - (\theta_l + \lambda\,\theta_h)\,\Gamma(T^*) \\ & + \kappa_l\Big\{D_B\big[v_l^B - \tilde{v}_{lh}^B\big] + D_G\big[v_l^G - \tilde{v}_{lh}^G\big]\Big\} \\ & + \kappa_h\Big\{D_B\big[v_h^B - \tilde{v}_{hl}^B\big] + D_G\big[v_h^G - \tilde{v}_{hl}^G\big]\Big\} \\ & + \psi^B\Big[\pi_l\, y_l^B + \pi_h\, y_h^B - \pi_l\, c_l^B - \pi_h\, c_h^B - \bar{R}\Big] \\ & + \psi^G\Big[\pi_l\, y_l^G + \pi_h\, y_h^G - q\big(\pi_l\, c_l^G + \pi_h\, c_h^G\big) - \bar{R}\Big],\end{aligned} \tag{5.61}$$

where $v_i^P = c_i^P - k(y_i^P,\omega_i)$ is the flow utility of type i in phase $P \in \{B,G\}$, and $\tilde{v}_{ij}^P = c_j^P - k(y_j^P,\omega_i)$ is the flow utility type i obtains when mimicking type j in phase P. The discounted phase durations are $D_B = (1-e^{-\rho T^*})/\rho$ and $D_G = e^{-\rho T^*}/\rho$, and cumulative discounted damages are given by (2.33).

Note that, as in the static model, the externality terms cancel within each incentive constraint. Aggregate pollution S_t is determined by the aggregate allocation and does not depend on individual type reports. A mimicking agent changes her own consumption and labor supply but does not change aggregate pollution. Therefore $\theta_i\,\Gamma(T^*)$ appears on both sides of each IC and drops out.

Within-phase consumption FOCs

Consumption utility is linear in c_i. Brown consumption enters the Lagrangian in three places: the objective's flow-utility terms via v_i^B, the incentive-constraint terms, and the damage term $-(\theta_l + \lambda\theta_h)\,\Gamma(T^*)$ through $B = \pi_l c_l^B + \pi_h c_h^B$, since

$$\frac{\partial\Gamma(T^*)}{\partial B} \;=\; \int_0^{T^*} e^{-\rho t}\,\phi'(Bt)\,t\,dt \;+\; \frac{e^{-\rho T^*}}{\rho}\,\phi'(BT^*)\,T^* \;>\; 0. \tag{5.62}$$

Define $\Gamma_B \equiv \partial\Gamma(T^*)/\partial B$ for brevity. Differentiating (5.61) with respect to c_l^B: the objective contributes $D_B \cdot 1$, the low-type IC contributes $\kappa_l \cdot D_B \cdot 1$ (type l truthfully receives c_l^B), the high-type IC contributes $-\kappa_h \cdot D_B \cdot 1$ (type h mimicking l would receive c_l^B), the damage term contributes $-(\theta_l + \lambda\,\theta_h)\,\Gamma_B\,\pi_l$, and the resource constraint contributes $-\psi^B \pi_l$. Setting to zero:

$$\frac{\partial \mathcal{L}}{\partial c_l^B} \;=\; D_B\,(1 + \kappa_l - \kappa_h) \;-\; (\theta_l + \lambda\,\theta_h)\,\Gamma_B\,\pi_l \;-\; \psi^B\,\pi_l \;=\; 0. \tag{D-FOC1}$$

By the same logic for c_h^B:

$$\frac{\partial \mathcal{L}}{\partial c_h^B} \;=\; D_B\,(\lambda + \kappa_h - \kappa_l) \;-\; (\theta_l + \lambda\,\theta_h)\,\Gamma_B\,\pi_h \;-\; \psi^B\,\pi_h \;=\; 0. \tag{D-FOC2}$$

These are the direct dynamic analogs of the static FOCs for brown consumption, which carry a Pigouvian term $(\theta_l + \lambda\theta_h)\,\phi'(\cdot)\,\pi_i$. As in the static model, the κ-terms reflect only the mimicker's deviation, so forming the ratio of (D-FOC1) to (D-FOC2) yields the population-share relation

$$\lambda \;=\; \frac{\pi_h + \kappa_l - \kappa_h}{\pi_l}, \tag{5.63}$$

since the externality term is symmetric in π_l and π_h and cancels in the ratio. Adding (D-FOC1) and (D-FOC2) (using $\pi_l + \pi_h = 1$) gives the Samuelson-type relation:

$$D_B\,(1 + \lambda) \;=\; \psi^B \;+\; (\theta_l + \lambda\,\theta_h)\,\Gamma_B. \tag{5.64}$$

The green-phase FOCs are analogous. The only difference is that green technology costs q units of income per unit of consumption, so the resource constraint contributes $-\psi^G\,q\,\pi_i$ rather than $-\psi^B\,\pi_i$:

$$\frac{\partial \mathcal{L}}{\partial c_l^G} = D_G\,(1 + \kappa_l - \kappa_h) - \psi^G\,q\,\pi_l = 0, \tag{D-FOC3}$$

$$\frac{\partial \mathcal{L}}{\partial c_h^G} = D_G\,(\lambda + \kappa_h - \kappa_l) - \psi^G\,q\,\pi_h = 0. \tag{D-FOC4}$$

Adding:

$$D_G\,(1 + \lambda) = \psi^G\,q. \tag{5.65}$$

Within-phase labor FOCs

Differentiating with respect to y_l^B: the objective contributes $-D_B \cdot k_y(y_l^B, \omega_l)$. The low-type IC contributes $-\kappa_l \cdot D_B \cdot k_y(y_l^B, \omega_l)$ (the truthful low type bears effort cost at her own ability ω_l). The high-type IC contributes $+\kappa_h \cdot D_B \cdot k_y(y_l^B, \omega_h)$ (a high type mimicking l produces y_l^B at ability ω_h). The resource constraint contributes $+\psi^B\,\pi_l$:

$$\frac{\partial \mathcal{L}}{\partial y_l^B} = -D_B\,(1 + \kappa_l)\,k_y(y_l^B, \omega_l) + D_B\,\kappa_h\,k_y(y_l^B, \omega_h) + \psi^B\,\pi_l = 0. \tag{D-FOC5}$$

Similarly for y_h^B:

$$\frac{\partial \mathcal{L}}{\partial y_h^B} = -D_B\,(\lambda + \kappa_h)\,k_y(y_h^B, \omega_h) + D_B\,\kappa_l\,k_y(y_h^B, \omega_l) + \psi^B\,\pi_h = 0. \tag{D-FOC6}$$

The green-phase labor FOCs have the same structure with superscript G replacing B and ψ^G replacing ψ^B:

$$\frac{\partial \mathcal{L}}{\partial y_l^G} = -D_G\,(1 + \kappa_l)\,k_y(y_l^G, \omega_l) + D_G\,\kappa_h\,k_y(y_l^G, \omega_h) + \psi^G\,\pi_l = 0, \tag{D-FOC7}$$

$$\frac{\partial \mathcal{L}}{\partial y_h^G} = -D_G\,(\lambda + \kappa_h)\,k_y(y_h^G, \omega_h) + D_G\,\kappa_l\,k_y(y_h^G, \omega_l) + \psi^G\,\pi_h = 0. \tag{D-FOC8}$$

These are the dynamic analogs of FOC5 and FOC6 in the static model. A mimicking agent retains her own productivity, so the effort cost is evaluated at the mimicker's ability level, not the mimicked type's.

5.2.2 Derivation of the Switching FOC

The switching time T^* enters the Lagrangian (5.61) through three channels: the discounted phase durations D_B and D_G, and the cumulative damage term $\Gamma(T^*)$. The within-phase resource constraints are per-period and do not depend on T^*.

Derivatives of D_B and D_G

Recall $D_B = (1 - e^{-\rho T^*})/\rho$ and $D_G = e^{-\rho T^*}/\rho$. Differentiating:

$$\frac{dD_B}{dT^*} = e^{-\rho T^*}, \qquad \frac{dD_G}{dT^*} = -e^{-\rho T^*}. \tag{5.66}$$

One additional period of brown production increases the discounted duration of the brown phase and decreases that of the green phase by the same amount.

Derivative of $\Gamma(T^*)$

The cumulative discounted damage is:

$$\Gamma(T^*) = \underbrace{\int_0^{T^*} e^{-\rho t}\,\phi(B\,t)\,dt}_{\text{(I): damages during brown phase}} + \underbrace{\frac{e^{-\rho T^*}}{\rho}\,\phi(B\,T^*)}_{\text{(II): post-switch permanent damages}}. \tag{5.67}$$

We differentiate each part with respect to T^*.

For part (I), by the Leibniz rule (the upper limit of integration depends on T^*):

$$\frac{d}{dT^*}\int_0^{T^*} e^{-\rho t}\,\phi(B\,t)\,dt = e^{-\rho T^*}\,\phi(B\,T^*). \tag{5.68}$$

This is the discounted damage flow at the boundary instant T^*.

For part (II), we apply the product rule. Write (II) $= \frac{1}{\rho}\, e^{-\rho T^*}\, \phi(B\,T^*)$. Then:

$$\begin{aligned} \frac{d}{dT^*}\left[\frac{e^{-\rho T^*}}{\rho}\,\phi(B\,T^*)\right] &= \frac{1}{\rho}\Big[-\rho\, e^{-\rho T^*}\,\phi(B\,T^*) + e^{-\rho T^*}\,\phi'(B\,T^*)\,B\Big] \\ &= -e^{-\rho T^*}\,\phi(B\,T^*) + \frac{e^{-\rho T^*}\,B\,\phi'(B\,T^*)}{\rho}. \end{aligned} \tag{5.69}$$

The first term is the reduction in the discounted post-switch damage burden due to the shorter green phase (D_G shrinks). The second term is the increase in permanent damages because the terminal stock $S^* = B\,T^*$ rises by B.

Adding (5.68) and (5.69):

$$\Gamma'(T^*) = e^{-\rho T^*}\,\phi(B\,T^*) - e^{-\rho T^*}\,\phi(B\,T^*) + \frac{e^{-\rho T^*}\,B\,\phi'(S^*)}{\rho} = \frac{e^{-\rho T^*}\,B\,\phi'(S^*)}{\rho}. \tag{5.70}$$

The boundary damage terms from parts (I) and (II) cancel exactly. The intuition is that at the boundary instant T^*, damages are $\phi(S^*)$ regardless of whether that instant falls in the brown phase or the green phase. The only net effect of extending the brown phase is the marginal increase in the permanent pollution stock, valued at the marginal damage $\phi'(S^*)$ and discounted over the infinite post-switch horizon (hence division by ρ).

Assembling the FOC for T^*

Differentiating the full Lagrangian (5.61) with respect to T^*, and using (5.66) and (5.70):

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial T^*} \;=\; & e^{-\rho T^*}\Big[\big(v_l^B + \lambda\, v_h^B\big) - \big(v_l^G + \lambda\, v_h^G\big)\Big] \\ & + e^{-\rho T^*}\,\kappa_l\Big[\big(v_l^B - \tilde{v}_{lh}^B\big) - \big(v_l^G - \tilde{v}_{lh}^G\big)\Big] \\ & + e^{-\rho T^*}\,\kappa_h\Big[\big(v_h^B - \tilde{v}_{hl}^B\big) - \big(v_h^G - \tilde{v}_{hl}^G\big)\Big] \\ & - (\theta_l + \lambda\,\theta_h)\,\frac{e^{-\rho T^*}\,B\,\phi'(S^*)}{\rho} \;=\; 0. \end{aligned} \tag{5.71}$$

The first line collects the welfare-weighted change in flow utilities from shifting one instant from the green phase to the brown phase (via $dD_B/dT^* = e^{-\rho T^*}$ applied to the objective and participation constraint terms). The second and third lines collect the analogous terms from the incentive constraints. The fourth line is the damage cost.

Dividing through by $e^{-\rho T^*} > 0$ and defining $\Delta_i \equiv (v_i^B - \tilde{v}_{ij}^B) - (v_i^G - \tilde{v}_{ij}^G)$:

$$\big(v_l^B + \lambda\, v_h^B\big) - \big(v_l^G + \lambda\, v_h^G\big) + \kappa_l\,\Delta_l + \kappa_h\,\Delta_h = \frac{(\theta_l + \lambda\,\theta_h)\,B\,\phi'(S^*)}{\rho}. \tag{5.72}$$

Eliminating λ

To obtain a switching condition that depends only on the incentive constraint multipliers κ_l and κ_h, we eliminate λ using the within-phase consumption FOCs. Recalling (5.63), the ratio of (D-FOC1) to (D-FOC2) yields

$$\frac{\lambda + \kappa_h - \kappa_l}{1 + \kappa_l - \kappa_h} = \frac{\pi_h}{\pi_l} \qquad \Longrightarrow \qquad \lambda = \frac{\pi_h + \kappa_l - \kappa_h}{\pi_l}, \tag{5.73}$$

since the externality term $(\theta_l + \lambda\theta_h)\,\Gamma_B\,\pi_i$ enters both FOCs symmetrically in π_i and cancels in the ratio, just as the Pigouvian term does in the static consumption FOCs. Substituting into the left-hand side of (5.72):

$$v_l^B + \lambda\, v_h^B = v_l^B + \frac{\pi_h + \kappa_l - \kappa_h}{\pi_l}\, v_h^B = \frac{\pi_l\, v_l^B + (\pi_h + \kappa_l - \kappa_h)\, v_h^B}{\pi_l}. \tag{5.74}$$

Substituting into the right-hand side:

$$\theta_l + \lambda\,\theta_h = \theta_l + \frac{\pi_h + \kappa_l - \kappa_h}{\pi_l}\,\theta_h = \frac{\pi_l\,\theta_l + (\pi_h + \kappa_l - \kappa_h)\,\theta_h}{\pi_l}. \tag{5.75}$$

Multiplying both sides of (5.72) by π_l:

$$\begin{aligned} &\pi_l\big(v_l^B - v_l^G\big) + \big(\pi_h + \kappa_l - \kappa_h\big)\big(v_h^B - v_h^G\big) + \pi_l\big(\kappa_l\,\Delta_l + \kappa_h\,\Delta_h\big) \\ &\quad = \ \big[\pi_l\,\theta_l + \big(\pi_h + \kappa_l - \kappa_h\big)\,\theta_h\big]\,\frac{B\,\phi'(S^*)}{\rho}. \end{aligned} \tag{5.76}$$

This is equation (2.34) in the main text. The multiplier λ no longer appears. The terms $\kappa_l\,\Delta_l$ and $\kappa_h\,\Delta_h$ capture how extending the brown phase changes the discounted incentive slack. In the first-best region ($\kappa_l = \kappa_h = 0$), these terms vanish and the condition reduces to (2.35).

5.2.3 Appendix: Proof of Proposition 2.6

We prove the identity $\overline{\Delta v}(\lambda) = \eta\, Y_B(\lambda)/(1+\varepsilon) - K$ in HIC and LIC. The FB case ($\kappa_l = \kappa_h = 0$) follows as the common boundary by continuity.

Framing. The proof works in the decentralization introduced earlier. A Pigouvian commodity tax internalizes the externality, and a Mirrlees nonlinear income tax handles redistribution. Under quasi-linear utility and an additive externality, Pirttilä–Tuomala–Kaplow additivity applies, and the income tax is the standard Mirrlees schedule that solves a redistributive problem in which the externality plays no direct role. Three consequences for the proof. *First*, the labor-leisure margin of the unconstrained type is undistorted by the income tax: $k_y(y_h^B, \omega_h) = 1$ in HIC and $k_y(y_l^B, \omega_l) = 1$ in LIC, evaluated on the consumer's optimization. *Second*, the constrained type's income margin satisfies the standard Mirrlees relation $k_y(\cdot) = A_B$ (HIC) or A_h (LIC), with the multipliers κ_l, κ_h reflecting the redistributive incentive constraint, exactly as in a Mirrlees economy without externalities. *Third*, the relative effective price of brown to green

consumption that enters the green-phase FOCs is q, the technological cost ratio adjusted by the Pigouvian commodity tax, which we continue to denote q for compactness.

Preliminaries. With iso-elastic costs (2.25), $k(y,\omega_l) = k(y,\omega_h)\, r^{-s}$ and $k_y(y,\omega_l) = k_y(y,\omega_h)\, r^{-s}$ where $r = \omega_l/\omega_h$ and $s = 1 + 1/\varepsilon$. The factorization $k_y(y,\omega)\cdot y = (1+1/\varepsilon)\, k(y,\omega)$ yields the useful identity

$$k(y,\omega) \;=\; \frac{\varepsilon}{1+\varepsilon}\, k_y(y,\omega)\, y. \tag{5.77}$$

Green allocations. The within-phase consumption and income FOCs in the green phase yield $k_y(y_i^G,\omega_i) = k_y(y_i^B,\omega_i)/q$ for both types in all regions. Combined with the iso-elastic form of k_y,

$$y_i^G = y_i^B/q^{\varepsilon}, \qquad k_i^G = k_i^B/q^{1+\varepsilon} \qquad \text{for } i \in \{l,h\}. \tag{5.78}$$

Aggregate consumption and per-phase ICs. Aggregate consumption in each phase is determined by the binding resource constraint. Writing $C^P \equiv \pi_l c_l^P + \pi_h c_h^P$, the brown- and green-phase RCs give $C^B = Y_B - \bar{R}$ and $C^G = (Y_G - \bar{R})/q$, respectively. Using $Y_G = Y_B/q^{\varepsilon}$ from (5.78), $C^G = Y_B/q^{1+\varepsilon} - \bar{R}/q$, hence

$$C^B - C^G \;=\; Y_B\left(1 - q^{-(1+\varepsilon)}\right) - \bar{R}\,(1-1/q) \;=\; \eta\, Y_B - K. \tag{5.79}$$

Define $\delta^P \equiv c_h^P - c_l^P$. Under Assumption 2.5, the relevant per-phase IC binds in each phase. In HIC,

$$\delta^P \;=\; k(y_h^P,\omega_h) - k(y_l^P,\omega_h) \;=\; k_h^P - r^s\, k_l^P, \tag{5.80}$$

while in LIC,

$$\delta^P \;=\; k(y_h^P,\omega_l) - k(y_l^P,\omega_l) \;=\; r^{-s}\, k_h^P - k_l^P. \tag{5.81}$$

In both cases, $\delta^G = \delta^B/q^{1+\varepsilon}$ by (5.78), hence $\delta^B - \delta^G = \eta\,\delta^B$. The resource constraints then pin down individual consumption levels:

$$c_l^P = C^P - \pi_h\,\delta^P, \qquad c_h^P = C^P + \pi_l\,\delta^P. \tag{5.82}$$

Computing $\Phi(\lambda)$. For each type,

$$\Delta v_i \;\equiv\; v_i^B - v_i^G \;=\; (c_i^B - c_i^G) - (k_i^B - k_i^G) \;=\; (c_i^B - c_i^G) - \eta\, k_i^B,$$

using (5.78). From (5.82) and (5.79),

$$\begin{aligned}
\Delta v_l &= (\eta Y_B - K) - \pi_h\,\eta\,\delta^B - \eta\, k_l^B \;=\; \eta\left[Y_B - \pi_h\,\delta^B - k_l^B\right] - K,\\
\Delta v_h &= (\eta Y_B - K) + \pi_l\,\eta\,\delta^B - \eta\, k_h^B \;=\; \eta\left[Y_B + \pi_l\,\delta^B - k_h^B\right] - K.
\end{aligned}$$

Under Assumption 2.5, $\Delta_l = \Delta_h = 0$, so the master FOC sums collapse to $\Phi(\lambda) = \Delta v_l + \lambda\,\Delta v_h$:

$$\Phi(\lambda) \;=\; -(1+\lambda)\,K + \eta\left[(1+\lambda)\,Y_B - \left(k_l^B + \lambda\, k_h^B\right) + (\pi_l\lambda - \pi_h)\,\delta^B\right]. \tag{5.83}$$

HIC simplification. In HIC, $\pi_l\lambda - \pi_h = -\kappa_h$ and $\delta^B = k_h^B - r^s\, k_l^B$ from (5.80). The bracket in (5.83) becomes

$$(1+\lambda)\, Y_B - (\lambda + \kappa_h)\, k_h^B - (1 - \kappa_h\, r^s)\, k_l^B. \tag{5.84}$$

In HIC, $y_h^B = \omega_h^{1+\varepsilon}$ implies $k_y(y_h^B, \omega_h) = 1$ and so, by (5.77), $k_h^B = \frac{\varepsilon}{1+\varepsilon} y_h^B$. The income FOC for y_l^B gives $k_y(y_l^B, \omega_l) = A_B = (1-\kappa_h)/(1-\kappa_h r^s)$, hence by (5.77), $k_l^B = \frac{\varepsilon}{1+\varepsilon} A_B\, y_l^B$. Therefore

$$(1 - \kappa_h r^s)\, k_l^B \;=\; (1 - \kappa_h r^s)\, A_B\, \frac{\varepsilon}{1+\varepsilon}\, y_l^B \;=\; (1-\kappa_h)\, \frac{\varepsilon}{1+\varepsilon}\, y_l^B.$$

Using $1 - \kappa_h = \pi_l(1+\lambda)$ and $\lambda + \kappa_h = \lambda + \pi_h - \pi_l\lambda = \pi_h(1+\lambda)$,

$$\begin{aligned}(\lambda + \kappa_h)\, k_h^B + (1 - \kappa_h r^s)\, k_l^B &= \pi_h(1+\lambda)\, \frac{\varepsilon}{1+\varepsilon}\, y_h^B + \pi_l(1+\lambda)\, \frac{\varepsilon}{1+\varepsilon}\, y_l^B \\ &= \frac{\varepsilon}{1+\varepsilon}\, (1+\lambda)\, Y_B.\end{aligned}$$

Substituting into (5.84):

$$(1+\lambda)\, Y_B \left(1 - \frac{\varepsilon}{1+\varepsilon}\right) \;=\; \frac{(1+\lambda)\, Y_B}{1+\varepsilon}.$$

Therefore $\Phi_{\text{HIC}}(\lambda) = -(1+\lambda)K + \eta\,(1+\lambda)Y_B/(1+\varepsilon) = (1+\lambda)\big[\eta\, Y_B/(1+\varepsilon) - K\big]$.

LIC simplification. In LIC, $\pi_l\lambda - \pi_h = \kappa_l$ and $\delta^B = r^{-s}\, k_h^B - k_l^B$ from (5.81). The bracket in (5.83) becomes

$$(1+\lambda)\, Y_B - (1 + \kappa_l)\, k_l^B - (\lambda - \kappa_l\, r^{-s})\, k_h^B. \tag{5.85}$$

In LIC, $y_l^B = \omega_l^{1+\varepsilon}$ gives $k_y(y_l^B, \omega_l) = 1$, so $k_l^B = \frac{\varepsilon}{1+\varepsilon} y_l^B$. The income FOC for y_h^B gives $k_y(y_h^B, \omega_h) = A_h = (\lambda - \kappa_l)/(\lambda - \kappa_l r^{-s})$, so $k_h^B = \frac{\varepsilon}{1+\varepsilon} A_h\, y_h^B$. Using $1 + \kappa_l = \pi_l(1+\lambda)$ and $\lambda - \kappa_l = \pi_h(1+\lambda)$,

$$(1 + \kappa_l)\, k_l^B \;=\; \pi_l(1+\lambda)\, \frac{\varepsilon}{1+\varepsilon}\, y_l^B,$$

$$(\lambda - \kappa_l r^{-s})\, k_h^B \;=\; (\lambda - \kappa_l r^{-s})\, A_h\, \frac{\varepsilon}{1+\varepsilon}\, y_h^B \;=\; (\lambda - \kappa_l)\, \frac{\varepsilon}{1+\varepsilon}\, y_h^B \;=\; \pi_h(1+\lambda)\, \frac{\varepsilon}{1+\varepsilon}\, y_h^B.$$

Summing,

$$(1 + \kappa_l)\, k_l^B + (\lambda - \kappa_l r^{-s})\, k_h^B \;=\; \frac{\varepsilon}{1+\varepsilon}\, (1+\lambda)\, Y_B,$$

and (5.85) reduces to $(1+\lambda)Y_B/(1+\varepsilon)$ as in HIC. Therefore $\Phi_{\text{LIC}}(\lambda) = (1+\lambda)\big[\eta\, Y_B/(1+\varepsilon) - K\big]$.

Dividing by $1+\lambda$ in either region delivers $\overline{\Delta v}(\lambda) = \eta Y_B/(1+\varepsilon) - K$, which equals $\eta B/(1+\varepsilon) - K_0$ on substituting $Y_B = B + \bar{R}$.

Non-negativity of K_0. Define $f(x) \equiv 1 - x^{-(1+\varepsilon)} - (1+\varepsilon)(1 - x^{-1})$ for $x > 0$. Then $f(1) = 0$ and

$$f'(x) \;=\; (1+\varepsilon)\, x^{-(2+\varepsilon)} - (1+\varepsilon)\, x^{-2} \;=\; (1+\varepsilon)\, x^{-2} \left(x^{-\varepsilon} - 1\right) \;\leq\; 0 \quad \text{for } x \geq 1,$$

with strict inequality for $x > 1$. Hence $f(q) \leq 0$ for $q \geq 1$, i.e. $\eta \leq (1+\varepsilon)(1-1/q)$, with strict inequality for $q > 1$. Multiplying by $\bar{R}/(1+\varepsilon) \geq 0$ yields $\eta\bar{R}/(1+\varepsilon) \leq K$, hence $K_0 = K - \eta\bar{R}/(1+\varepsilon) \geq 0$. $\square$

5.2.4 Numerical Calibration

This section describes the numerical exercise underlying Figure 2.3. The planner's problem is solved directly using constrained nonlinear optimisation (sequential quadratic programming).

Functional Forms and Parameters

Consumption utility is linear in c_{it}. Effort costs are $k(y;\omega) = (y/\omega)^2$, with $\omega_l = 2.1$ and $\omega_h = 3.5$. The damage function is quadratic, $\phi(S) = d\,S^2/2$, with $d = 5 \times 10^{-5}$. The green technology cost is $q = 1.3$, the discount rate is $\rho = 0.05$, and the exogenous revenue requirement is $\bar{R} = 0$. Environmental preference parameters are $\theta_h = 2$ and $\theta_l = 0.5$. Population shares are equal, $\pi_l = \pi_h = 0.5$.

Solution Method

The planner chooses nine objects: the switching time T^* and within-phase allocations $(c_i^B, y_i^B, c_i^G, y_i^G)$ for $i \in \{l, h\}$. The optimisation maximises the low type's lifetime welfare W_l subject to: the participation constraint $W_h \geq \bar{V}_h$, the per-phase resource constraints, and the lifetime incentive constraints for both types.

The cumulative discounted damage under the quadratic specification evaluates in closed form. With $\phi(S) = dS^2/2$,

$$\int_0^{T^*} e^{-\rho t}\,\frac{d\,B^2 t^2}{2}\,dt \;=\; \frac{d\,B^2}{2}\cdot\frac{2D_B - T^* e^{-\rho T^*}(\rho T^* + 2)}{\rho^2}, \tag{5.86}$$

using $\int_0^{T^*} t^2 e^{-\rho t}\,dt = [2D_B - T^* e^{-\rho T^*}(\rho T^* + 2)]/\rho^2$, and the post-switch term is $(d\,B^2\,T^{*2} e^{-\rho T^*})/(2\rho)$. Summing the two, the $T^{*2} e^{-\rho T^*}$ terms cancel and

$$\Gamma(T^*) \;=\; \frac{d\,B^2\left(D_B - T^*\,e^{-\rho T^*}\right)}{\rho^2}, \tag{5.87}$$

where $B = \pi_l\,c_l^B + \pi_h\,c_h^B$ is aggregate brown consumption. As a sanity check, $\Gamma'(T^*) = d\,B^2\,T^* e^{-\rho T^*}/\rho = e^{-\rho T^*} B\,\phi'(S^*)/\rho$, consistent with the general envelope result (5.70).

The figure is constructed by sweeping over a grid of reservation utility values $\bar{V}_h$. At each grid point, the problem is solved from multiple initial conditions (perturbations of the switching time around the previous solution) to guard against local optima. The sweep proceeds outward from a first-best reference point. The IC slack values are recorded to classify each solution into the HIC, FB, or LIC region.

Features of the Solution

The first-best switching time is approximately $T^* \approx 50$ periods. In the first-best region, T^* is constant across a range of $\bar{V}_h$ values, confirming that the switching condition does not depend on the incentive multipliers when no incentive constraint binds. As the planner moves into the HIC region (low $\bar{V}_h$, Rawlsian direction), T^* increases sharply, reflecting the lower effective damage weight $\bar{\theta}^{HIC}$. As the planner moves into the LIC region (high $\bar{V}_h$), T^* decreases, reflecting the higher effective damage weight $\bar{\theta}^{LIC}$. The transitions at the region boundaries are continuous, consistent with the theoretical prediction. The magnitude of the variation in T^* is substantial. Under the chosen parameterisation, the switching time ranges from approximately 30 periods in the LIC region to over 150 periods in the HIC region.

Chapter 6

Appendix for Chapter 3

6.1 Appendix

6.1.1 Proof of Proposition 3.1

Preliminaries: Implications of the functional form assumption on consumption utility

Suppose that

$$u(x,\theta) \quad = \quad \theta^{\frac{1}{\varepsilon}} \, \frac{1}{1-\frac{1}{\varepsilon}} \, x^{1-\frac{1}{\varepsilon}} \, ,$$

for $\varepsilon > 1$. Then

$$x^*(q,\theta) \quad = \quad \theta \, q^{-\varepsilon}$$

and

$$x^*_q(q,\theta) \quad = \quad -\varepsilon \, \theta \, q^{-(1+\varepsilon)} \, ,$$

and

$$v(q,\theta) \quad = \quad \theta \, \frac{1}{1-\frac{1}{\varepsilon}} \, q^{-(\varepsilon-1)} \, .$$

Statement 1: Only neoclassical consumers

Suppose that $\phi_N = 1$. Also assume that damage functions are linear. Policies are then evaluated according to

$$U(t,\theta,\delta) \quad = \quad R(t) + v(k+t,\theta) - \delta \, \bar{x}(t)$$

or, equivalently,

$$U(t,\theta,\delta) \quad = \quad v(k+t,\theta) + (t-\delta) \, \bar{x}(t) \, ,$$

where $\bar{x}(t) = \bar{\theta}_N \, q^{-\varepsilon}$ and $\bar{\theta}_N$ is the average value of θ among N-agents. We have that

$$U_t(t,\theta,\delta) \quad = \quad q^{-(1+\varepsilon)} \, p(t,\theta,\delta)$$

where

$$p(t,\theta,\delta) \quad = \quad \left\{(\bar{\theta}_N - \theta)q - \varepsilon(t-\delta)\bar{\theta}_N\right\}$$

Thus $U_t(\cdot) > 0$ if $p(\cdot) > 0$, $U_t(\cdot) = 0$ if $p(\cdot) = 0$ and $U_t(\cdot) < 0$ if $p(\cdot) < 0$. Using that $q = (k+t)$, we find that, for given θ and δ, p is an affine-linear function of t. It is now straightforward to verify the following statements:

$$\begin{aligned}
p(\cdot) \;>\; 0 \quad &\Longleftrightarrow \quad \frac{1}{\varepsilon}\left(1 - \frac{\theta}{\bar{\theta}_N}\right) \;>\; \frac{t-\delta}{k+t}\,, \\
p(\cdot) \;=\; 0 \quad &\Longleftrightarrow \quad \frac{1}{\varepsilon}\left(1 - \frac{\theta}{\bar{\theta}_N}\right) \;=\; \frac{t-\delta}{k+t}\,, \\
p(\cdot) \;<\; 0 \quad &\Longleftrightarrow \quad \frac{1}{\varepsilon}\left(1 - \frac{\theta}{\bar{\theta}_N}\right) \;<\; \frac{t-\delta}{k+t}\,,
\end{aligned}$$

and, finally, that $\frac{t-\delta}{k+t}$ is an increasing function of t. All this implies that preferences are single peaked. The ideal tax of an N agent with type (θ,δ) is given by

$$t^*(\theta,\delta) \quad = \quad \frac{\frac{1}{\varepsilon}\left(1 - \frac{\theta}{\bar{\theta}_N}\right)k + \delta}{1 - \frac{1}{\varepsilon}\left(1 - \frac{\theta}{\bar{\theta}_N}\right)}\,.$$

Statement 2: Mixing N and S agents

An N-agent evaluates taxes t, again, according to

$$U(t,\theta,\delta) \quad = \quad v(k+t,\theta) + (t-\delta)\;\bar{x}(t)\,,$$

but now

$$\bar{x}(t) \quad = \quad \phi_N\;\bar{\theta}_N\;q^{-\varepsilon} \;+\; \phi_S\;\bar{\theta}_S\;q_S^{-\varepsilon}\,,$$

where $q = k + t$, as before, and $q_S = k + \phi_N\; t + \phi_S\;\bar{\delta}_S$, where $\bar{\delta}_S$ is the average of δ among S-agents. Straightforward computations yield

$$U_t(t,\theta,\delta) \quad = \quad q^{-(1+\varepsilon)}\;p_N(t,\theta,\delta) \;+\; q_S^{-(1+\varepsilon)}\;\phi_S\;\bar{\theta}_S\;p_S(t,\delta)\,,$$

where

$$p_N(t,\theta,\delta) \quad = \quad (\phi_N\;\bar{\theta}_N - \theta)q - (t-\delta)\;\varepsilon\;\phi_N\;\bar{\theta}_N$$

and

$$p_S(t,\delta) \quad = q_S - (t-\delta)\;\varepsilon\;\phi_N\,.$$

Thus, a type (θ,δ)-agent evaluates marginal changes of t according to a weighted average of two evaluation functions with weights that, however, depend on the tax rate t. To make this explicit inn terms of notation, we let $w_N(t) = q^{-(1+\varepsilon)}$, $w_S(t) = q_S^{-(1+\varepsilon)}\phi_S\bar{\theta}_S$ and write

$$U_t(t,\theta,\delta) \quad = \quad w_N(t)\;p_N(t,\theta,\delta) \;+\; w_S(t)\;p_S(t,\delta)\,.$$

Both p_N and p_S are affine functions of t. Let

$$t_N^*(\theta,\delta) \quad = \quad \frac{\frac{1}{\varepsilon}\left(1-\frac{\theta}{\phi_N\,\bar{\theta}_N}\right)k+\delta}{1-\frac{1}{\varepsilon}\left(1-\frac{\theta}{\phi_N\,\bar{\theta}_N}\right)}\,.$$

It is easy to verify that $t \mapsto p_N(t,\theta,\delta)$ is a decreasing function of t, with $p_N(t,\theta,\delta) > 0$ for $t < t_N^*(\theta,\delta)$, $p_N(t,\theta,\delta) = 0$ for $t = t_N^*(\theta,\delta)$ and $p_N(t,\theta,\delta) < 0$ for $t > t_N^*(\theta,\delta)$. Analogously, let

$$t_S^*(\delta) \quad = \quad \frac{k+\phi_S\,\bar{\delta}_S+\varepsilon\,\phi_N\,\delta}{(\varepsilon-1)\phi_N}\,.$$

It is easy to verify that $t \mapsto p_S(t,\delta)$ is a decreasing function of t, with $p_S(t,\delta) > 0$ for $t < t_S^*(\delta)$, $p_S(t,\theta,\delta) = 0$ for $t = t_S^*(\delta)$ and $p_S(t,\delta) < 0$ for $t > t_S^*(\delta)$.

An interpretation: If we had $w_N = 1$ and $w_S = 0$, irrespective of t, then preferences would be single-peaked with a peak at $t_N^*(\theta,\delta)$. If we had $w_N = 0$ and $w_S = 1$, irrespective of t, then preferences would be single-peaked with a peak at $t_S^*(\delta)$.

All this implies that $U_t(\cdot) > 0$ when $t < t_N^*(\theta,\delta)$ and $t < t_S^*(\delta)$. Also, $U_t(\cdot) < 0$ when $t > t_N^*(\theta,\delta)$ and $t > t_S^*(\delta)$. Since U_t is continuous in t this implies that there exists some t in the intermediate range between $t_N^*(\theta,\delta)$ and $t_S^*(\delta)$ so that $U_t(t,\theta,\delta) = 0$. To establish single-peakedness we need to show that this is the only value of t so that $U_t(t,\theta,\delta) = 0$. Since U_t is itself a continuously differentiable function of t, any non-monotonicity of U over the intermediate range would be associated with another local extremum or peak, i.e. with another value of t so that $U_t(t,\theta,\delta) = 0$. In the following we provide sufficient condition for the concavity of $t \mapsto U(t,\theta,\delta)$ over the intermediate range of tax rates, i.e. we provide sufficient conditions for $U_{tt}(t,\theta,\delta) < 0$, where

$$\begin{aligned} U_{tt}(t,\theta,\delta) \quad &= \quad w_N'(t)\,p_N(t,\theta,\delta) + w_N(t)\,p_N'(t,\theta,\delta) \\ &\quad\quad + w_S'(t)\,p_S(t,\delta) + w_S(t)\,p_S'(t,\delta)\,. \end{aligned}$$

Note that we have already established above that, for all t, $w_N'(t), w_S'(t), p_N'(t,\theta,\delta)$ and $p_S'(t,\delta)$ are all negative, whereas $w_N(t)$ and $w_S(t)$ are positive.

Case 1: $t_N^*(\theta,\delta) < t_S^*(\delta)$**.** In this case, over the intermediate range of taxes we have $p_N(t,\theta,\delta) < 0$ and $p_S(t,\delta) > 0$. This implies that

$$w_S'(t)\,p_S(t,\delta) + w_S(t)\,p_S'(t,\delta) \quad < \quad 0\,.$$

A sufficient condition for $U_{tt}(\cdot) < 0$, therefore is that also

$$w_N'(t)\,p_N(t,\theta,\delta) + w_N(t)\,p_N'(t,\theta,\delta) \quad < \quad 0\,.$$

After some algebra, $w_N'(t)\,p_N(t,\theta,\delta) + w_N(t)\,p_N'(t,\theta,\delta)$ can be written as

$$q^{-(2+\varepsilon)}\,X(t,\theta,\delta)\,,$$

where

$$\begin{aligned} X(t,\theta,\delta) &= -\varepsilon\, k \left(2\, \phi_N\, \bar{\theta}_N - \theta\right) - \delta\, \phi_N\, \bar{\theta}_N\, \varepsilon\, (1+\varepsilon) \\ &\quad + t \left(\delta\, \phi_N\, \bar{\theta}_N\, \varepsilon\, (\varepsilon - 1) + \varepsilon\, \theta\right) \end{aligned}$$

The following assumptions are sufficient to ensure that $X < 0$ for all θ and δ: Values of θ are drawn from the unit interval, $\bar{\theta}_N > \frac{1}{2}$, $\bar{\theta}_N$ sufficiently close to 1 and k sufficiently large. These assumption imply that $2\ \phi_N\ \bar{\theta}_N - \theta > 0$. Recall that t is, by assumption, bounded between $t^*_N(\theta,\delta)$ and $t^*_S(\delta)$. Thus, by choosing k sufficiently large, we can always make sure that X is negative.

Case 2: $t^*_N(\theta,\delta) > t^*_S(\delta)$**.** In this case, over the intermediate range of taxes we have $p_N(t,\theta,\delta) > 0$ and $p_S(t,\delta) < 0$. This implies that

$$w'_N(t)\ p_N(t,\delta) + w_N(t)\ p'_N(t,\delta) \quad < \quad 0\ .$$

A sufficient condition for $U_{tt}(\cdot) < 0$, therefore is that

$$w'_S(t)\ p_S(t,\theta,\delta) + w_S(t)\ p'_S(t,\theta,\delta)$$

is sufficiently small. Since, this expression is linear in ϕ_S, it holds for ϕ_S sufficiently small.

6.1.2 Proof of Proposition 3.2

Recall that

$$U_t(t,\phi_S,\theta,\delta) \quad = \quad \bar{x}(t,\phi_S) - x^*(q(t),\theta) + (t-\delta)\ \bar{x}'(t,\phi_S)$$

where we have expanded the notation introduced in the main text to emphasize that both per-capita consumption $\bar{x}$ and its derivative with respect to the tax rate depend in a parametric way on the population share of S-agents, ϕ_S. To establish the Proposition, we show that, under the premises of the Proposition $\bar{x}_{\phi_S}(t,\phi_S) < 0$ and $\bar{x}'_{\phi_S}(t,\phi_S) > 0$. In words: with more S-agents the tax base shrinks, which is one force reducing the benefits from higher taxes. Also, the tax base becomes less tax sensitive, as the absolute value $\mid \bar{x}'(t,\phi_S) \mid$ goes down. Together these effects imply that $U_{t\phi_S} > 0$, when $t < \delta$. First, note that

$$\bar{x}(t,\phi_S) \quad = \quad \bar{x}_N(t) - \phi_S\Big(\bar{x}_N(t) - \bar{x}_S(q_S)\Big)$$

where $q_S = k + t + \phi_S(\bar{\delta}_S - t)$. The derivative with respect to ϕ_S equals

$$\bar{x}_{\phi_S}(t,\phi_S) \quad = \quad -\Big(\bar{x}_N(t) - \bar{x}_S(q_S)\Big)\ +\ \phi_S\ \bar{x}'_S(q_S)(\bar{\delta}_S - t)$$

which is a sum of two negative terms. Second, note that

$$\bar{x}'(t,\phi_S) \quad = \quad \bar{x}'_N(t) - \phi_S\Big(\bar{x}'_N(t) - \bar{x}'_S(q_S)(1-\phi_S)\Big)$$

The derivative with respect to ϕ_S equals

$$\bar{x}'_{\phi_S}(t,\phi_S) \quad = \quad -\Big(\bar{x}'_N(t)+\bar{x}'_S(q_S)\Big) \; - \; \phi_S(1-\phi_S)\,\bar{x}''_S(q_S)(\bar{\delta}_S - t)\;.$$

The ferst term in this sum is positive and dominates the second when ϕ_S is sufficiently small.

6.1.3 Proof of Proposition 3.3

Under the premises of the Proposition $W'_S(t)$ can be written as

$$W'_S(t) \quad = \quad \phi_s\; q^{-\varepsilon}\;\Big(LHS(t) - RHS(t)\Big)$$

where

$$LHS(t) \quad := \quad \bar{\theta}_N - \bar{\theta}_S\left(\frac{q}{q_S}\right)^{\varepsilon}$$

and

$$RHS(t) \quad := \quad \bar{\theta}_N\;\varepsilon\;\frac{t-\delta}{q}\;.$$

where $q = k+t$ and $q_S = k + \phi_N\; t + \phi_S\;\bar{\delta}_S$. Note that $t \mapsto LHS(t)$ is a decreasing function of t and $t \mapsto RHS(t)$ is an increasing function of t. Thus, there exists at most one level of t, henceforth denoted by t^*_S, so that $LHS(t) = RHS(t)$. Moreover, if such a t exists, then $W'_S(t) > 0$ for $t < t^*_S$ and $W'_S(t) < 0$ for $t < t^*_S$, which establishes single-peakedness. To show that there exists such $t^*_S >$, note that at $t = \delta$, $RHS(t) = 0$ and $LHS(t) \geq 0$, where the inequality is strict if and only if $\bar{\theta}_N > \bar{\theta}_S$. Thus, if $\bar{\theta}_N = \bar{\theta}_S$, then $t^*_S = \bar{\delta}_S$. Note that $\lim_{t\to\infty} RHS(t) = \bar{\theta}_N\;\varepsilon > \bar{\theta}_N$ and that $LHS(t)$ is bounded from above by $\bar{\theta}_N$. Thus, the functions $t \mapsto LHS(t)$ and $t \mapsto RHS(t)$ eventually cross and if $\bar{\theta}_N > \bar{\theta}_S$, then they cross at $t^*_S > \bar{\delta}_S$.

6.1.4 Proof of Proposition 3.4

Note that

$$W_t(t,\phi_S) \quad = \quad (1-\phi_S)\Big(\bar{x}_N(t) - \bar{x}_S(q_S)\Big) + (1-\phi_S)(t-\delta_S)\;\bar{x}'_N(t,\phi_S)$$

for $q_S = k + t + \phi_S(\delta_S - t)$. Straightforward computations yield

$$W_{t\phi_S}(t,\phi_S) \quad = \quad -\Big(\bar{x}_N(t) - \bar{x}_S(q_S)\Big) - (t-\delta_S)\;\Big(\bar{x}'_N(t) - (1-\phi_S)\bar{x}'_S(q_S)\Big)\;.$$

Using $t \leq \bar{\delta}_S$, $\bar{\theta}_N \geq \bar{\theta}_S$ and $u(x,\theta) = \theta^{\frac{1}{\varepsilon}}\;\frac{1}{1-\frac{1}{\varepsilon}}\;x^{1-\frac{1}{\varepsilon}}$, it is straightforward to verify that

$$\bar{x}_N(t) - \bar{x}_S(q_S) \quad > \quad 0$$

and

$$\bar{x}'_N(t,\phi_S) - (1-\phi_S)\bar{x}'_S(q_S) \quad < \quad 0\;,$$

which establishes the result.

6.1.5 Proof of Proposition 3.5

For some fixed but arbitrary $\sigma > 0$ and $\sigma_t < 0$, let $a = \phi_S\,\sigma$ and $b = \phi_S\;\sigma_t$. We use a and b to parameterize the function U that is used by an N-agent of type (θ, δ) to evaluate marginal changes of taxes,

$$U_t(t, \theta, \delta, a, b) \quad = \quad \bar{x}(t, a) - x^*(q(t) + a, \theta)(1 + b) - (t - \delta)\,\bar{x}_t(t, a)\ ,$$

where

$$\bar{x}(t, a) \quad = \quad \phi_N\,\bar{\theta}_N\,\big(q(t) + a\big)^{-\varepsilon}\ +\ \phi_S\,\bar{\theta}_S\;q_S(t)^{-\varepsilon}\ ,$$

$$x^*(q(t) + a, \theta) \quad = \quad \theta\,\big(q(t) + a\big)^{-\varepsilon}\ ,$$

and

$$\bar{x}_t(t, a) \quad = \quad -\varepsilon\Big(\phi_N\,\bar{\theta}_N\,\big(q(t) + a\big)^{-(1+\varepsilon)}\ +\ \phi_S\,\phi_N\,\bar{\theta}_S\;q_S(t)^{-(1+\varepsilon)}\Big)\ .$$

Henceforth, we denote the agent's ideal tax by $t^*(\theta, \delta, a, b)$. It solves the equation

$$U_t(t, \theta, \delta, a, b) \quad = \quad 0\ .$$

Obviously,

$$U_{tb}(t, \theta, \delta, a, b) \quad = \quad -x^*(q(t) + a, \theta) \quad < \quad 0\ .$$

Therefore, whatever a, moving from $b = 0$ to $b < 0$ increases the benefit from higher taxes and pushes the agent's ideal tax up: $t^*(\theta, \delta, a, 0) < t^*(\theta, \delta, a, b)$, when $b < 0$. Now, to establish the Proposition, we show that, for an agent with $t^*(\theta, \delta, 0, 0) \leq \delta$, and a and b sufficiently small, as implied by ϕ_S small, we have

$$U_{ta}(t^*(\theta, \delta, a, b), \theta, \delta, a, b) \quad > \quad 0\ .$$

Thus, under these conditions, moving from $a = b = 0$, to $a > 0 > b$ shifts the ideal tax up: $t^*(\theta, \delta, 0, 0) < t^*(\theta, \delta, a, b)$. To verify that, under the stated conditions, $U_{ta}(t^*(\cdot), \theta, \delta, a, b) > 0$, note that

$$U_{ta}(t, \theta, \delta, a, b) \quad = \quad -\frac{\varepsilon}{(q(t) + a)^{1+\varepsilon}}\left(\phi_N\,\bar{\theta}_N\ -\ \theta(1 + b)\ +\ (t - \delta)(1 + \varepsilon)\frac{\phi_N\,\bar{\theta}_N}{q(t) + a}\right)$$

Now, for a and b close to zero, $t^*(\theta, \delta, a, b)$ is close to $t^*(\theta, \delta, 0, 0)$. Moreover, $t = t^*(\theta, \delta, 0, 0)$ implies that

$$\phi_N\,\bar{\theta}_N\ -\ \theta\ +\ (t - \delta)\varepsilon\frac{\phi_N\,\bar{\theta}_N}{q(t)}\ .$$

Therefore, the sign of $U_{ta}(t^*(\theta,\delta,a,b),\theta,\delta,a,b)$ is equal to the sign of

$$-\frac{\varepsilon}{(q(t^*(\cdot))+a)^{1+\varepsilon}}\left(-\theta\ b\ +\ (t^*(\cdot)-\delta)\frac{\phi_N\,\bar{\theta}_N}{q(t)+a}\right)$$

or, using that $b \simeq 0$,

$$-\frac{\varepsilon}{q(t^*(\cdot))^{1+\varepsilon}}(t^*(\cdot)-\delta)\frac{\phi_N\,\bar{\theta}_N}{q(t^*(\cdot))}$$

when a and b are close to zero. This expression is positive when $t^*(\cdot) \leq \delta$.

6.1.6 Proof of Lemma 3.3

Equation (3.11) implies that, in equilibrium,

$$\hat{\sigma}^{eq}(t)(q+\phi_S\,\hat{\sigma}^{eq}(t))^{1+\varepsilon} \quad = \quad \delta_S\ \phi_N\ \phi_S \int_0^{\hat{\theta}^{eq}(t)} \theta f_N(\theta)\ d\theta\ .$$

By the arguments in the main text, $t \mapsto \hat{\theta}^{eq}(t)$ is an increasing function of t. Hence, the right hand side of the above equation is increasing in t. Therefore, the left hand side needs to be an increasing function of t too. But this requires that $t \mapsto \hat{\sigma}^{eq}(t)$ is an increasing function.

6.1.7 Proof of Proposition 3.6

The proof follows from the arguments in the proof of Proposition 3.5, upon noting that, $U_{ta}(\cdot)$ becomes small when ϕ_S and hence a and b go to zero provided that $\mid t^*(\theta,\delta)-\delta \mid$ is small. Therefore, the sign of b governs the comparative statics of the preferred tax. $b<0$ in the no-rebellion equilibrium and $b>0$ in the rebellion equilibrium then imply that $t_N^*(\theta,\delta) > t^*(\theta,\delta) > t_R(\theta,\delta)$ as had to be shown.

6.1.8 Universalists

Universalist-rule utilitarians: consumption choices. Their objective is

$$\begin{aligned}\tilde{W}_U \quad = \quad R(t) \quad &+ \quad \phi_N\,\mathbf{E}_{f_N}[u(x^*(q(t),\theta))\ -\ q\,x^*(q(t),\theta)] \\ &+ \quad \phi_S\,\mathbf{E}_{f_S}[u(x_S(t,\theta))\ -\ q\,x_S(t,\theta)] \\ &+ \quad \phi_U\,\mathbf{E}_{f_U}[u(x_U(t,\theta))\ -\ q\,x_U(t,\theta)] \quad - \quad \mathcal{D}_U(\bar{x}(t))\ ,\end{aligned}$$

where $\mathcal{D}_U(\bar{x}(t)) = \mathbf{E}_{g_U}[D(\bar{x}(t),\delta)]$.

Universalists cannot control the consumption of neoclassical consumersor group rule-utilitarians. They still internalize the utility net of damages realized by other agents. This will matter for their preferences over taxes that we will characterize in the subsequent paragraph. They are paternalist, however, in that they do not respect the view of others on what the damages are. They use the damage function $\mathcal{D}_U$ to compute damages for everyone, irrespectively of whether or nor the others agree that this is the right damage function.

The objective $\tilde{W}_U$ can equivalently be written as

$$\begin{aligned}\tilde{W}_U &= \phi_N \, \mathbf{E}_{f_N}[u(x^*(q(t),\theta)) \; - \; k \, x^*(q(t),\theta)] \\ &+ \phi_S \, \mathbf{E}_{f_S}[u(x_S(t,\theta)) \; - \; k \, x_S(t,\theta)] \\ &+ \phi_U \, \mathbf{E}_{f_U}[u(x_U(t,\theta)) \; - \; k \, x_U(t,\theta)] \quad - \quad \mathcal{D}_U(\cdot) \, ,\end{aligned}$$

Universalists choose $\theta \mapsto x_U(\theta)$ to maximize this expression. Note that their choices do no depend on the tax rate, so that we no longer need to carry t as an argument of x_U. The first order condition characterizing the optimal choice $x_U^*(\theta)$ is

$$u_x(x_U^*(\theta),\theta) \quad = \quad k \; + \; \mathcal{D}'_U(\cdot) \, .$$

This first order condition implies, in particular, that universalist rule-utilitarians show a behavioral response to changes of marginal damages, but not to taxes.

Universalist's preferences over taxes. Their preferences over taxes are given by

$$\begin{aligned}W_U(t) &= \phi_N \, \mathbf{E}_{f_N}[u(x^*(q(t),\theta)) \; - \; k \, x^*(q(t),\theta)] \\ &+ \phi_S \, \mathbf{E}_{f_S}[u(x_S(t,\theta)) \; - \; k \, x_S(t,\theta)] \\ &+ \phi_U \, \mathbf{E}_{f_U}[u(x_U^*(\theta)) \; - \; k \, x_U^*(\theta)] \quad - \quad \mathcal{D}_U((1-\phi_U)\bar{x}_{-U}(t) + \phi_U \bar{x}_U) \, ,\end{aligned}$$

where

$$\bar{x}_{-U}(t) = \frac{\phi_N}{1-\phi_U} \, \bar{x}_N(t) \; + \frac{\phi_S}{1-\phi_U} \, \bar{x}_S(t) \, .$$

Note that

$$\begin{aligned}W'_U(t) &= \phi_N \, \mathbf{E}_{f_N}[(u_x(x^*(q(t),\theta)) \; - \; k) \, x_t^*(q(t),\theta)] \\ &+ \phi_S \, \mathbf{E}_{f_S}[(u_x(x_S(t,\theta)) \; - \; k) \, x_{S,t}(t,\theta)] \\ &- \mathcal{D}'_U(\cdot) \, (1-\phi_U) \, \bar{x}'_{-U}(t) \, ,\end{aligned}$$

Using the first order conditions characterizing the consumption choices of neoclassical consumers and group rule utilitarians, this can be rewritten as

$$\begin{aligned}W'_U(t) &= \phi_N \, \Big(t - \mathcal{D}'_U(\cdot)\Big)\bar{x}'_N(t) \\ &+\phi_S\Big((1-\phi_S)t + \phi_S \, \mathcal{D}'_S(\cdot) \quad - \quad \mathcal{D}'_U(\cdot)\Big)\bar{x}'_S(t) \, .\end{aligned}$$

A special case. Suppose that $\phi_S = 0$ or that $\mathcal{D}_U = \mathcal{D}_S$. Then $t = \mathcal{D}'_U(\cdot)$ implies that $W'_U(t) = 0$; $t < \mathcal{D}'_U(\cdot)$ implies $W'_U(t) > 0$; $t > \mathcal{D}'_U(\cdot)$ implies $W'_U(t) < 0$. *Thus, the function W_U is single-peaked with a peak at $t = \mathcal{D}'_U(\cdot)$.* [Could also assume that $D_U = D_S$ so that S and U agents differ only in whom they care about/ where altruism applies.]

6.1.9 An alternative interpretation of the setup: Heterogeneity in incomes as opposed to hetereogeneity in the taste for dirty goods.

Suppose for simplicity that there are only N agents and that the damage function is linear and the same for all agents. With iso-elastic utility $u(x,\theta) = \theta^{\frac{1}{\varepsilon}} \frac{1}{1-\frac{1}{\varepsilon}}\ x^{1-\frac{1}{\varepsilon}}$, $x^*(q,\theta) = \frac{\theta}{q^\varepsilon}$. Hence, in these cases, consumption is linear in θ and the distribution of θ determines how the median relates to the mean. If $\theta_M < \bar{\theta}$, then $t^*_M > \delta$, where $\bar{\theta} := \mathbf{E}_{f_N}[\theta]$ and t^*_M is the median voter's preferred tax rate.

To get as sense of whether $\theta_M < \bar{\theta}$ is empirically a plausible case, consider a setup that is a richer than the one discussed above. Households differ in their disposable income y and in the budget share γ of their expenditures on the dirty good. The complementary fraction $1-\gamma$ is spent on an unspecific consumption good c. Households solve the following consumer choice problem: Choose x and c to max

$$\gamma \ln x + (1-\gamma)\ln c$$

subject to the constraint

$$q\ x + c \quad = \quad y\ -\ \delta\bar{x}(t)\ +\ t\ \bar{x}(t)\ ,$$

where the price of the unspecific consumption good has been normalized to 1. The demand for the dirty good x is then given by

$$x^*(\cdot) \quad = \quad \frac{\gamma\ \big(y + (t-\delta)\bar{x}(t)\big)}{q}\ .$$

Using this equation to solve for $\bar{x}(t)$ yields

$$\bar{x}(t) \quad = \quad \frac{\mathbf{E}[\gamma\ y]}{k + (1-\bar{\gamma})t + \mathbf{E}[\gamma\ \delta]}\ ,$$

where $\bar{\gamma} := \mathbf{E}[\gamma]$. Dropping irrelevant additive constants, the indirect utility function that is used for policy evaluation can be written as

$$U(t,\gamma,y,\delta) \quad = \quad \ln\Big(y + (t-\delta)\ \bar{x}(t)\Big)\ -\ \gamma\ \ln(k+t)$$

Straightforward calculations yield

$$U_t(t,\gamma,y,\delta) \quad = \quad \frac{\bar{x}(t)}{y + (t-\delta)\bar{x}(t)}\ \frac{k + \delta(1-\bar{\gamma}) + \mathbf{E}[\gamma\ \delta]}{k + t(1-\bar{\gamma}) + \mathbf{E}[\gamma\ \delta]}\ -\ \gamma\ \frac{1}{k+t}$$

It is now easy to see that $U_t(\cdot)$ is decreasing in y and γ and increasing in δ. Hence, if preferences are single-peaked, then the peaks are larger for people with high damages, low incomes and small expenditures on dirty goods. Note that the sign of $U_t(\cdot)$ is equal to the sign of $LHS(t) - RHS(t)$, where

$$LHS(t) \quad = \quad (k+t)\ \frac{k + \delta(1-\bar{\gamma}) + \mathbf{E}[\gamma\ \delta]}{k + t(1-\bar{\gamma}) + \mathbf{E}[\gamma\ \delta]}$$

and

$$RHS(t) \quad = \quad \frac{\gamma\, y}{\bar{x}(t)} + \gamma(t-\delta)\ .$$

The function $t \mapsto RHS(t)$ is increasing and the function $t \mapsto LHS(t)$ can be shown to be decreasing for k large enough. Consequently, preferences are single-peaked.

Now suppose that γ and δ are the same for everybody and that heterogeneity is entirely due to disposable incomes. Then the decisive voter is the one with the median position y_M in the cross-section distribution of y. For $y = y_M$, the condition $LHS(t) = RHS(t)$ can equivalently be written as

$$\frac{y_M}{\bar{y}} \quad = \quad 1 \ + \ \frac{q(1-\gamma)(\delta - t)}{\big[q + \gamma(\delta - t)\big]^2}$$

This equation characterizes t_M^*, the median voter's preferred tax rate. If $y_M < \bar{y}$, then it has to be the case that $q(1-\gamma)(\delta - t_M^*) < 0$, which implies that $t_M^* > \delta$.

6.1.10 Additional Numerical Illustrations

Heterogeneous mean tastes: $\bar{\theta}_N > \bar{\theta}_S$

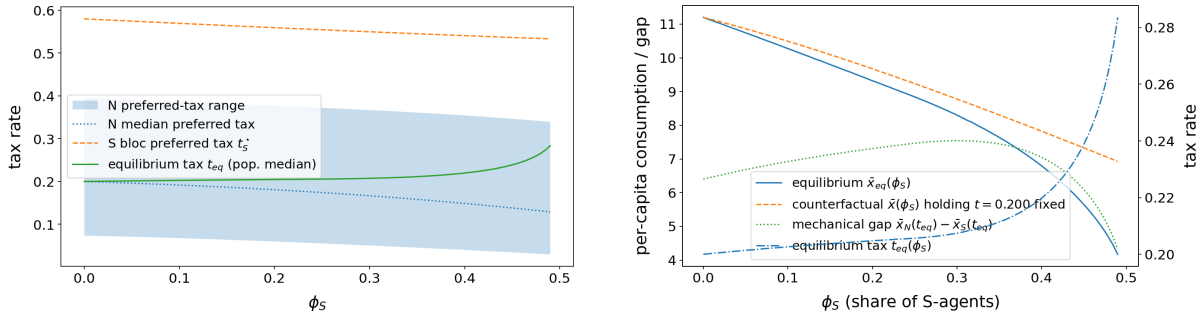
In the main text we focus on $\bar{\theta}_N = \bar{\theta}_S$, which isolates the political-economy mechanisms without bringing in the additional overtaxation motive of Proposition 3.3. Here we reproduce the political-equilibrium and consumption diagnostics under $\bar{\theta}_N > \bar{\theta}_S$, drawing $\theta_N \sim U[0.4, 1.0]$ ($\bar{\theta}_N = 0.7$) and $\theta_S \sim U[0.0, 0.6]$ ($\bar{\theta}_S = 0.3$). All other primitives are as in the main text.

Three features distinguish this case. First, the S-bloc's preferred tax (orange dashed in Panel (a) of Figure 6.1) sits strictly above $\bar{\delta}_S = 0.40$ throughout: it begins at $t_S^*(0) \approx 0.58$ and falls to ≈ 0.53 at $\phi_S = 0.5$. This is the overtaxation motive of Proposition 3.3 — when $\bar{\theta}_N > \bar{\theta}_S$, S-agents prefer a tax above their own Pigouvian benchmark $\bar{\delta}_S$ in order to lean on the N-bloc's larger demand for the dirty good. The negative slope is consistent with the logic of Proposition 3.4: as the N tax base shrinks with ϕ_S, the redistributive and corrective benefits of higher taxes that the S-bloc draws from the N-bloc both weaken, and the S-bloc's ideal tax falls.

Second, the political-equilibrium tax (green solid) rises monotonically as in the main text, but never reaches t_S^* within the plotted range. The decisive voter remains an N-agent throughout: t_S^* sits well above the upper edge of the N preferred-tax distribution, and even as the S-bloc's voting weight grows toward $\frac{1}{2}$, the cumulative-weight threshold for the decisive voter to switch to the S-bloc is not crossed for $\phi_S < \frac{1}{2}$. The visible acceleration of t^{eq} near $\phi_S = \frac{1}{2}$ reflects the population median walking through the upper end of the N preferred-tax range.

Third, in Panel (b) of Figure 6.1 the cross-sectional consumption gap $\bar{x}_N(t^{eq}) - \bar{x}_S(t^{eq})$ (green dotted) no longer starts at zero. With $\bar{\theta}_N \neq \bar{\theta}_S$, the two groups consume differently even when they face the same effective price, and at $\phi_S = 0$ the gap is already $(\bar{\theta}_N - \bar{\theta}_S)/(k + t^{eq})^{\varepsilon} \approx 6.4$. The price wedge $q_S - q_N = \phi_S(\bar{\delta}_S - t^{eq})$ still modulates the gap with ϕ_S, but its lower envelope is now bounded away from zero by the $\bar{\theta}$-difference.

The remaining elements of the picture, the composition channel through the dashed counterfactual, policy channel through the rising equilibrium tax, and reinforcing rather than offsetting behavioural responses across the two groups, are unchanged from the main-text scenario.



(a) Political-Equilibrium Tax

(b) Aggregate Consumption

Figure 6.1: Weak Waterbed Effect under $\bar{\theta}_N > \bar{\theta}_S$

Notes: Panel (a): The shaded blue band gives the cross-sectional range of N-agents' preferred taxes over the support of (θ_N, δ_N). The dotted blue line is the within-N median preferred tax. The orange dashed line is the S-bloc's preferred tax t_S^*, which exceeds $\bar{\delta}_S = 0.40$ throughout (Proposition 3.3) and falls strictly in ϕ_S (Proposition 3.4). The green solid line is the political equilibrium tax, which is the median preferred tax over the entire population. The decisive voter remains an N-agent throughout the plotted range. Panel (b): The blue solid line is equilibrium per-capita consumption $\bar{x}^{eq}(\phi_S) = (1-\phi_S)\bar{x}_N(t^{eq}) + \phi_S \bar{x}_S(q_S(t^{eq}, \phi_S))$. The orange dashed line holds the tax fixed at $t^{eq}(\phi_S{=}0) = 0.20$, isolating the composition channel. The green dotted line is the consumption gap $\bar{x}_N(t^{eq}) - \bar{x}_S(t^{eq})$, which remains strictly positive at $\phi_S = 0$ because $\bar{\theta}_N \neq \bar{\theta}_S$ even when $q_N = q_S$. The blue dash-dotted line (right axis) is $t^{eq}(\phi_S)$ from panel (a). Parameters: $\varepsilon = 2$, $k = 0.05$, $\theta_N \sim$ Uniform$[0.4, 1.0]$, $\theta_S \sim$ Uniform$[0.0, 0.6]$, $\delta_N \sim$ Uniform$[0.1, 0.3]$, $\delta_S \sim$ Uniform$[0.3, 0.5]$.

6.1.11 Empirical Appendix

Tax Response to Damage Increases

Table 3.2 examines preferred tax levels. Here we turn to how respondents adjust their preferred tax when told that the social cost of carbon rises by 30%, either only in their own country (*Local Tax Adjustment*, Table 6.1) or uniformly across all countries (*Global Tax Adjustment*, Table 6.2), holding the previously chosen tax as the reference point.[1]

The patterns are broadly consistent with the tax-level results in Table 3.2. Climate concern (δ) is strongly significant in all specifications. *Price vs. Harm* is significant everywhere, and θ is weakly negative. In both tables, the Damage Response is significant on its own but loses significance when θ and δ are included. In the US, the conditional coefficient on the Damage Response turns positive (though insignificant), unlike in the tax-level regression where it is essentially zero.

[1] *Local* (resp. *Global*) *Tax Adjustment* is the respondent's percentage change in their preferred carbon tax (slider from −100% to +100%, "−100% Abandon the tax" to "+100% Double the tax") in response to: "Now suppose that only in [your country] (resp., across all countries globally, including [your country]), the social cost of carbon ... has increased by 30%. You previously chose a carbon tax of $[X] per ton. How would you adjust your preferred carbon tax rate in [your country]?"

Table 6.1: Tax Preferences — Responsiveness to Local Damage Information

	Dependent variable: Tax Adjustment (%, Local)									
	UK	UK	UK	UK	UK	US	US	US	US	US
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Consumption Importance (θ)	−2.560*		−2.365		−2.265	−2.524*		−2.684*		−2.396
	(1.464)		(1.472)		(1.399)	(1.513)		(1.505)		(1.464)
Climate Concern (δ)	16.373***		15.883***		12.907***	21.669***		22.535***		17.432***
	(1.402)		(1.400)		(1.451)	(1.486)		(1.527)		(1.578)
Damage Response		−5.932***	−2.129				−4.842***	2.618		
		(1.380)	(1.301)				(1.824)	(1.652)		
Price vs. Harm				14.142***	9.183***				17.211***	9.937***
				(1.473)	(1.527)				(1.512)	(1.523)
Constant	−11.273***	−11.254***	−11.275***	−11.078***	−11.105***	−15.679***	−15.717***	−15.679***	−15.764***	−15.704***
	(1.292)	(1.389)	(1.291)	(1.323)	(1.261)	(1.393)	(1.559)	(1.391)	(1.459)	(1.361)
Observations	893	894	893	890	889	900	901	900	900	899
R^2	0.155	0.020	0.157	0.114	0.198	0.212	0.011	0.215	0.134	0.249
Adjusted R^2	0.153	0.019	0.155	0.113	0.195	0.210	0.010	0.212	0.133	0.246

Notes: Robust (HC1) standard errors in parentheses. All independent variables are standardized (z-scored). *p<0.1; **p<0.05; ***p<0.01.

Table 6.2: Tax Preferences — Responsiveness to Global Damage Information

	Dependent variable: Tax Adjustment (%, Global)									
	UK	UK	UK	UK	UK	US	US	US	US	US
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Consumption Importance (θ)	−2.455*		−2.359		−2.184	−3.895**		−3.960***		−3.763**
	(1.454)		(1.465)		(1.405)	(1.518)		(1.518)		(1.486)
Climate Concern (δ)	14.940***		14.701***		11.780***	20.823***		21.175***		17.225***
	(1.431)		(1.428)		(1.487)	(1.506)		(1.581)		(1.572)
Damage Response		−4.593***	−1.041				−6.002***	1.064		
		(1.438)	(1.374)				(1.824)	(1.660)		
Price vs. Harm				13.031***	8.464***				15.621***	8.453***
				(1.459)	(1.536)				(1.536)	(1.490)
Constant	−4.554***	−4.563***	−4.555***	−4.401***	−4.400***	−11.625***	−11.668***	−11.625***	−11.681***	−11.620***
	(1.266)	(1.353)	(1.267)	(1.291)	(1.240)	(1.351)	(1.506)	(1.352)	(1.429)	(1.329)
Observations	893	894	893	890	889	900	901	900	900	899
R^2	0.137	0.013	0.138	0.103	0.176	0.211	0.017	0.211	0.117	0.239
Adjusted R^2	0.135	0.012	0.135	0.102	0.174	0.209	0.016	0.209	0.116	0.236

Notes: Robust (HC1) standard errors in parentheses. All independent variables are standardized (z-scored). *p<0.1; **p<0.05; ***p<0.01.

Ordered Probit Specification for Moral Pressure Outcomes

The moral pressure outcomes *MP Approval* and *MP Opposition* are measured on five-point Likert scales (1 = Strongly disagree, 5 = Strongly agree). In the main text Tables 3.3 and 3.4, we treat these outcomes as cardinal and estimate the conditional mean by OLS with HC1 standard errors. This has the advantage of yielding coefficients that are directly interpretable as marginal effects on the response scale, and it facilitates comparison with the tax-level and tax-adjustment regressions reported elsewhere in the paper.

Because the outcomes are strictly ordinal, we also re-estimate all specifications by ordered probit, which treats the five response categories as ordered but does not impose equal spacing between them. The specifications are identical to those in Tables 3.3 and 3.4.

The ordered probit results are qualitatively identical to the OLS results (see Tables 6.3 and 6.4). Signs, significance levels, and the relative magnitudes of the coefficients are preserved across all eight specifications in both countries. We conclude that the OLS estimates reported in the main text are not driven by the cardinality assumption.

Table 6.3: Support and Opposition to Moral Pressure — UK (Ordered Probit)

	Dependent variable:							
	MP Approval				MP Opposition			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Damage Response	−0.221***		−0.108***		0.199***		0.094**	
	(0.036)		(0.038)		(0.036)		(0.037)	
Price vs. Harm		0.470***		0.321***		−0.375***		−0.235***
		(0.038)		(0.041)		(0.037)		(0.039)
Consumption Importance (θ)			−0.098***	−0.106***			0.150***	0.158***
			(0.037)	(0.037)			(0.036)	(0.036)
Climate Concern (δ)			0.582***	0.510***			−0.514***	−0.457***
			(0.040)	(0.042)			(0.039)	(0.041)
McFadden RÂ²	0.015	0.062	0.104	0.127	0.012	0.041	0.084	0.096
Observations	892	888	891	887	893	889	892	888

Notes: Ordered probit estimates; standard errors in parentheses. All independent variables are standardized (z-scored). The dependent variables are *MP Approval* and *MP Opposition*, each measured on a five-point Likert scale (1 = Strongly disagree, 5 = Strongly agree); see Table 3.3 for the question wording. All covariates are defined as in Tables 3.1 and 3.3. Cut-point estimates are suppressed for readability. *p<0.1; **p<0.05; ***p<0.01.

Table 6.4: Support and Opposition to Moral Pressure — US (Ordered Probit)

	Dependent variable:							
	MP Approval				MP Opposition			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Damage Response	−0.157***		0.044		0.263***		0.102***	
	(0.035)		(0.038)		(0.036)		(0.038)	
Price vs. Harm		0.447***		0.258***		−0.364***		−0.170***
		(0.037)		(0.040)		(0.036)		(0.039)
Consumption Importance (θ)			−0.092**	−0.089**			0.069*	0.075**
			(0.036)	(0.036)			(0.036)	(0.036)
Climate Concern (δ)			0.677***	0.569***			−0.593***	−0.560***
			(0.042)	(0.042)			(0.041)	(0.042)
McFadden RÂ²	0.007	0.055	0.11	0.126	0.02	0.038	0.101	0.105
Observations	901	900	900	899	899	898	898	897

Notes: Ordered probit estimates; standard errors in parentheses. All independent variables are standardized (z-scored). The dependent variables are *MP Approval* and *MP Opposition*, each measured on a five-point Likert scale (1 = Strongly disagree, 5 = Strongly agree); see Table 3.3 for the question wording. All covariates are defined as in Tables 3.1 and 3.3. Cut-point estimates are suppressed for readability. *p<0.1; **p<0.05; ***p<0.01.

6.1.12 Survey Instructions

This subsection documents the full survey flow shown to respondents. Screens are presented in the order they appeared. Where a screen had a separate version for U.S. respondents, the U.S. version is shown immediately after the baseline (UK) version. Where a screen has an *adjusted* version, this shows how the screen looked after the respondent moved the slider.

Introduction and Screening

Figure 6.2: Introduction screen (1/2).

Figure 6.3: Introduction screen (2/2).

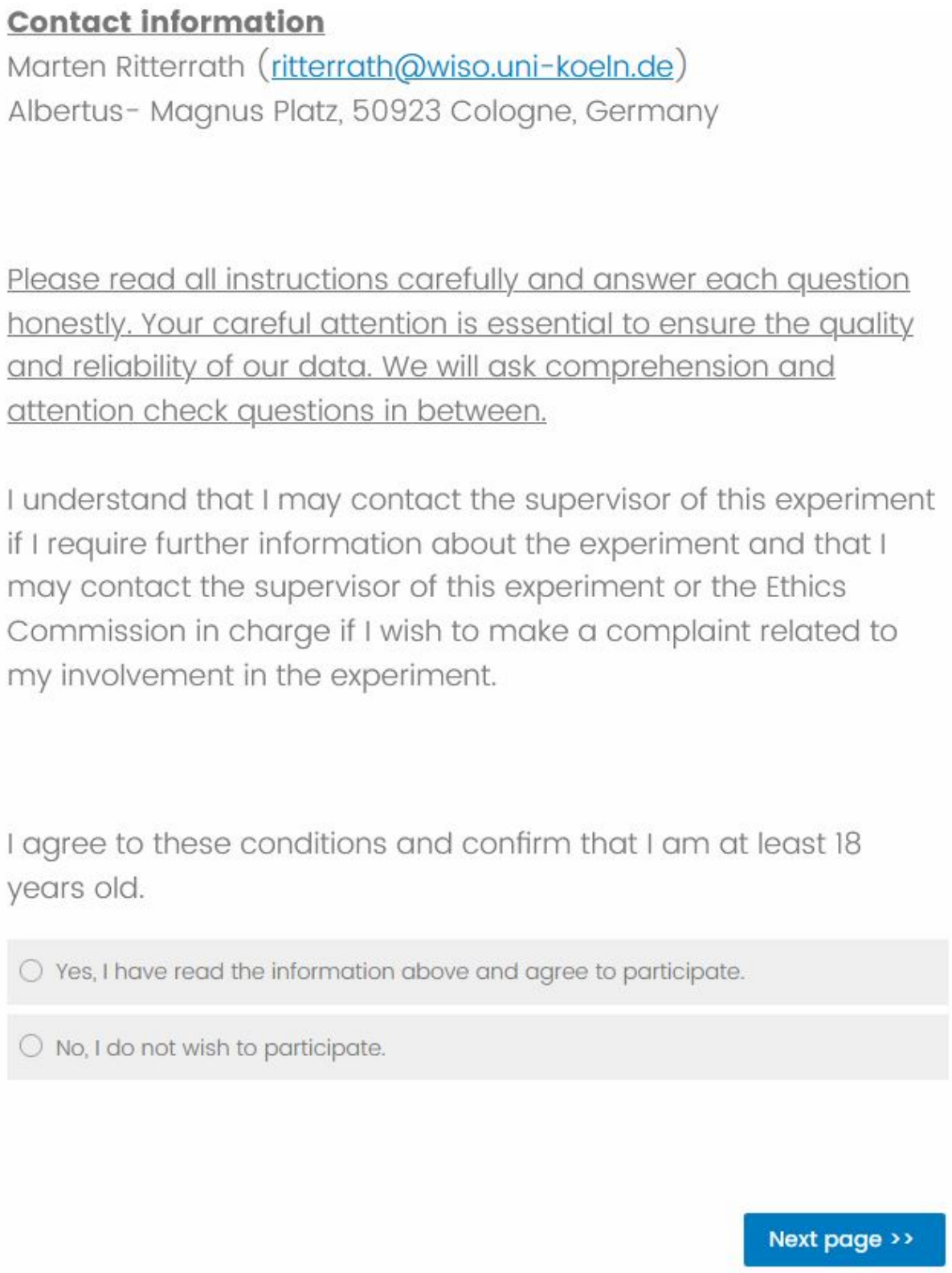


Figure 6.4: CAPTCHA screen.

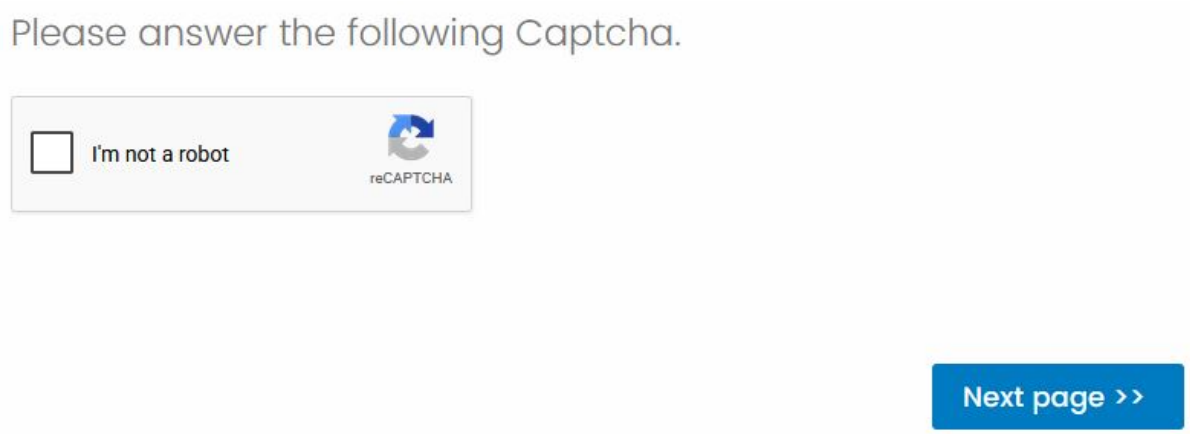


Figure 6.5: Prolific ID entry.

Demographics

Figure 6.6: Gender question.

Figure 6.7: Education question (UK version).

Figure 6.8: Education question (U.S. version). The education question was adapted to reflect the U.S. educational system.

Figure 6.9: Income question (UK version).

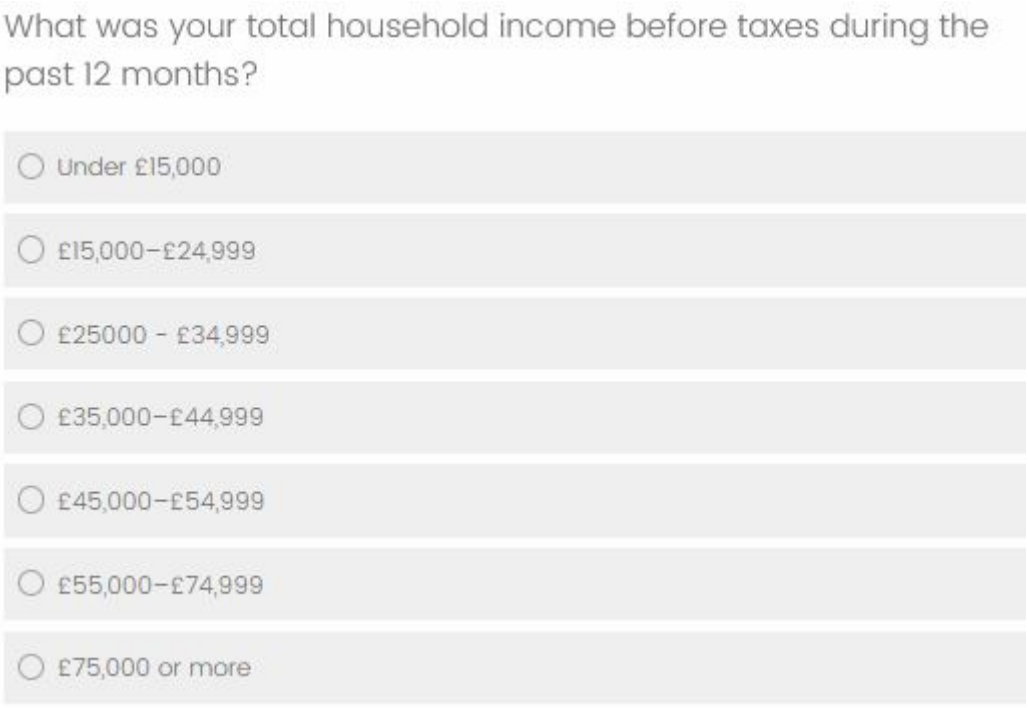
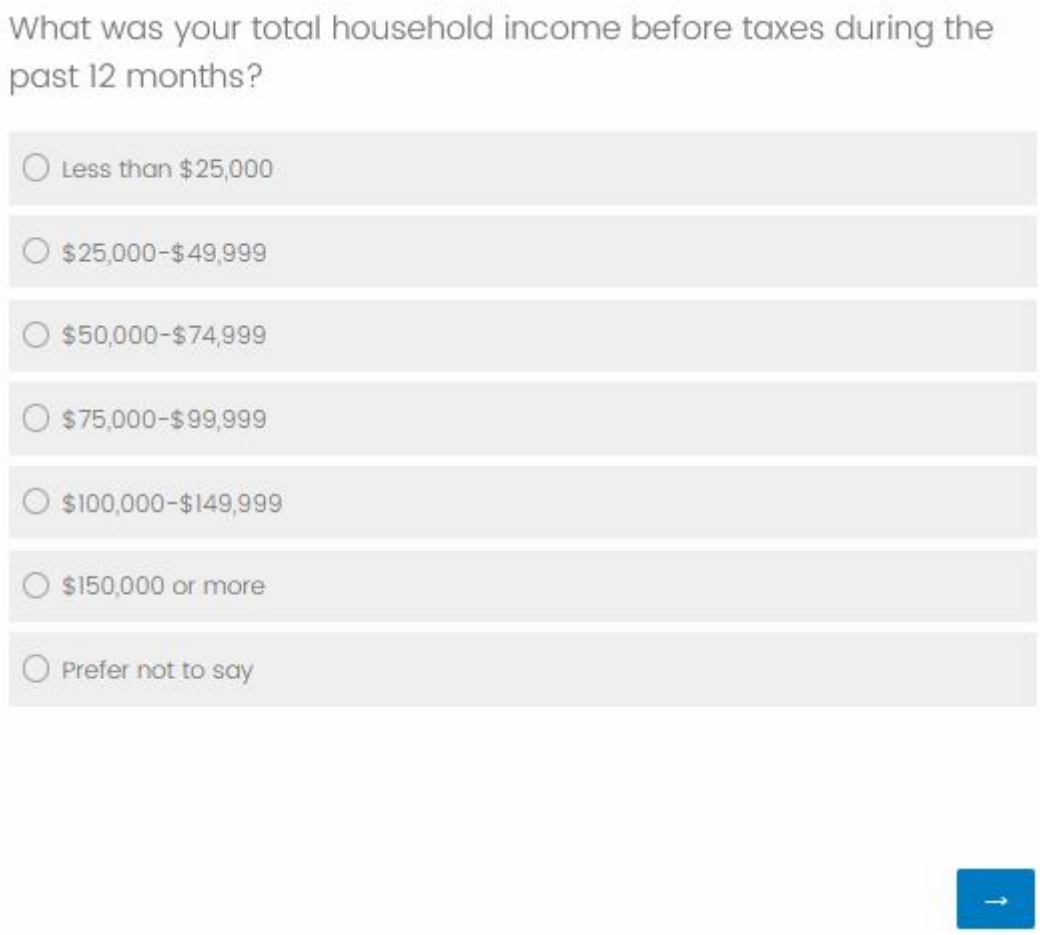


Figure 6.10: Income question (U.S. version). The income question was adapted to reflect U.S. income brackets (in U.S. dollars).



Car Ownership and Usage

Figure 6.11: Car type question.

Figure 6.12: Car usage question (UK version).

Figure 6.13: Car usage question (U.S. version). The car usage question was adapted to reflect U.S. driving patterns.

θ **and** δ

Figure 6.14: Elicitation of θ.

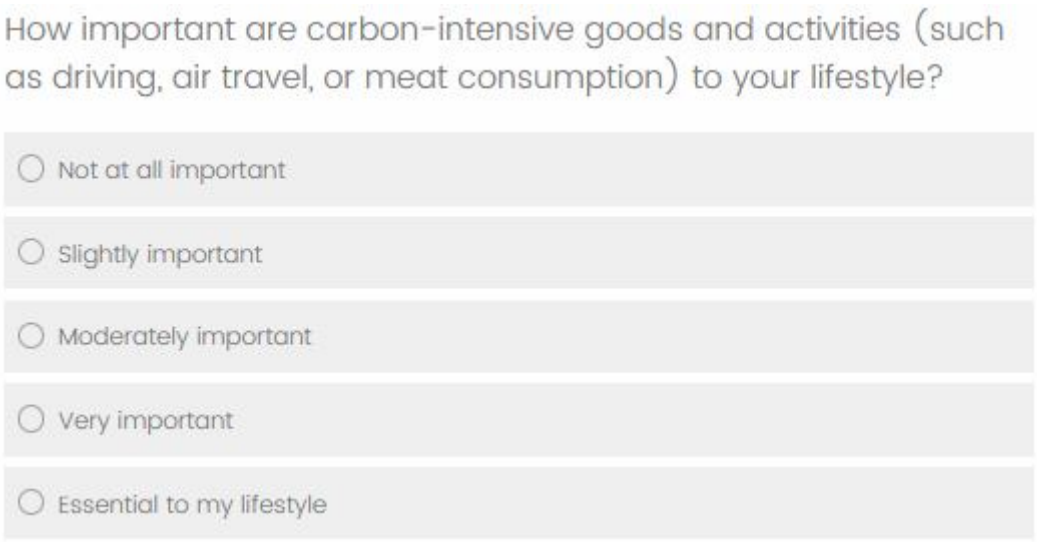
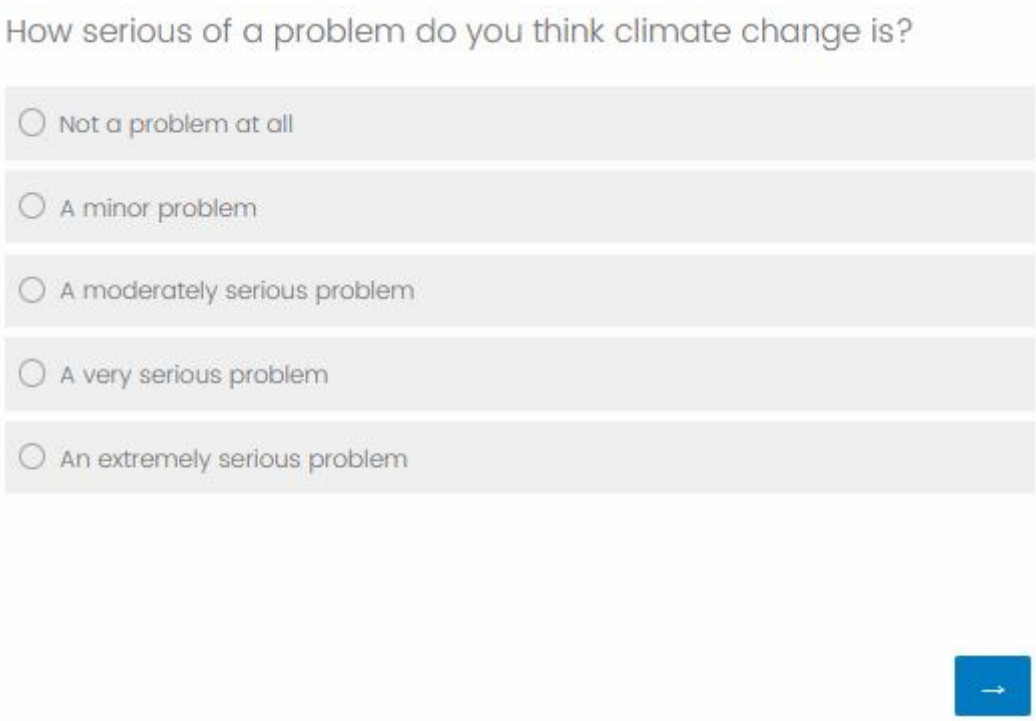


Figure 6.15: Elicitation of δ.



Carbon Tax Block

Figure 6.16: Carbon tax information screen.

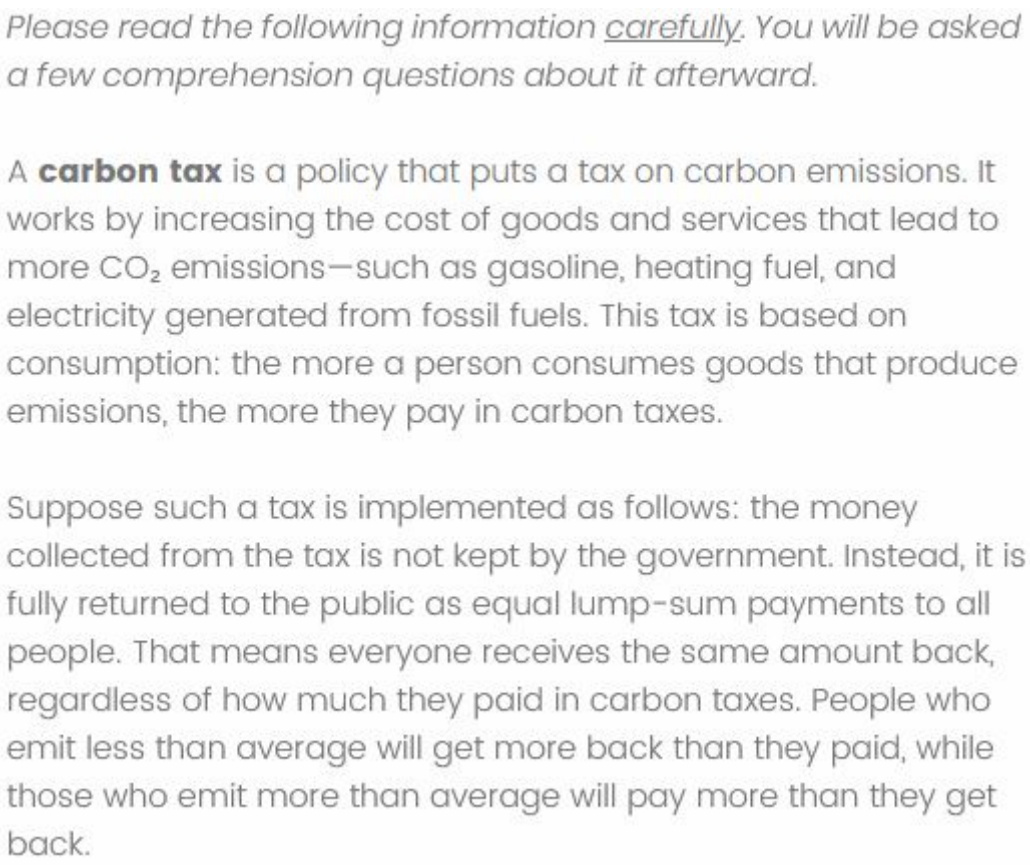


Figure 6.17: Carbon tax comprehension check.

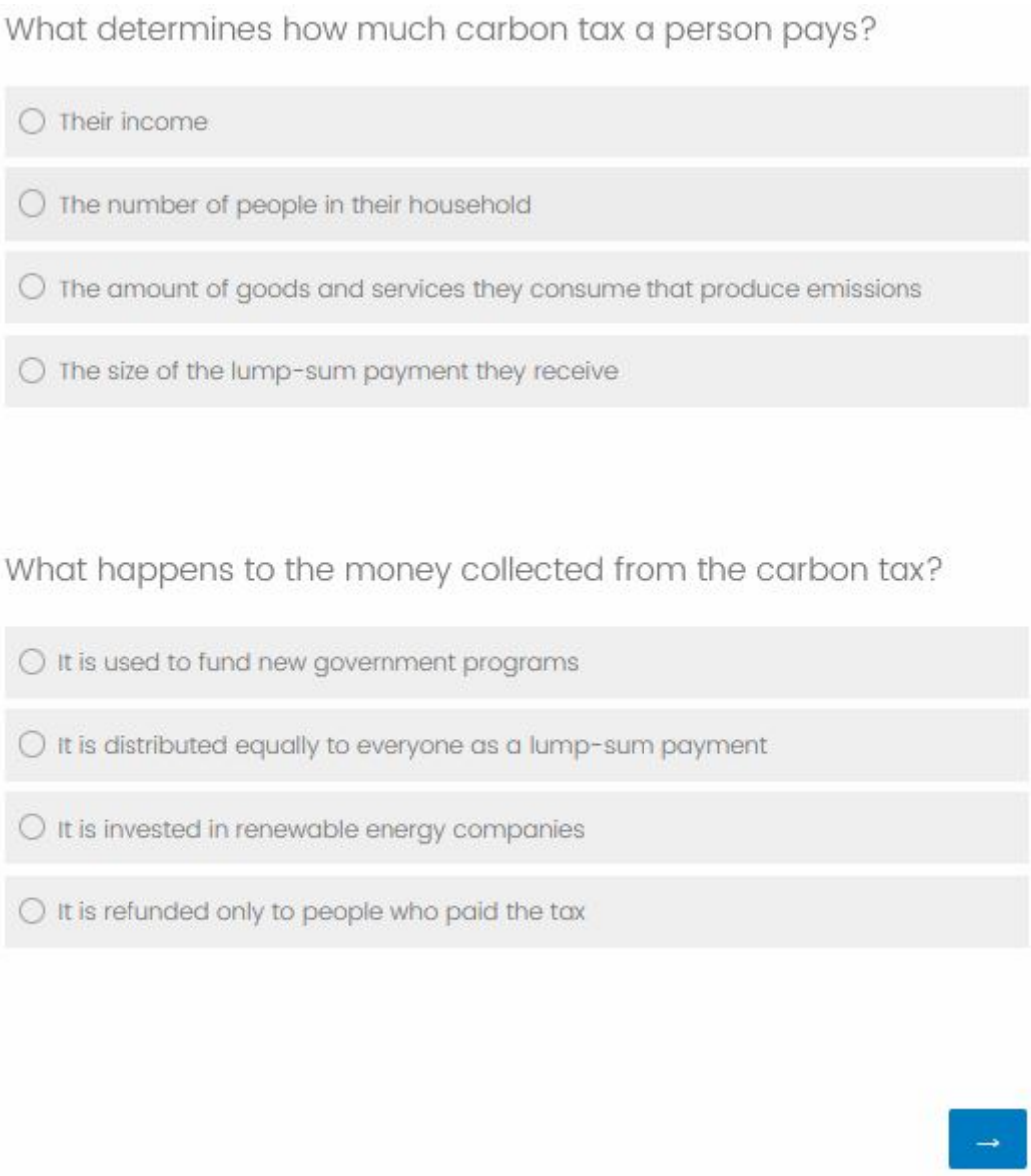


Figure 6.18: Price adjustment under carbon tax.

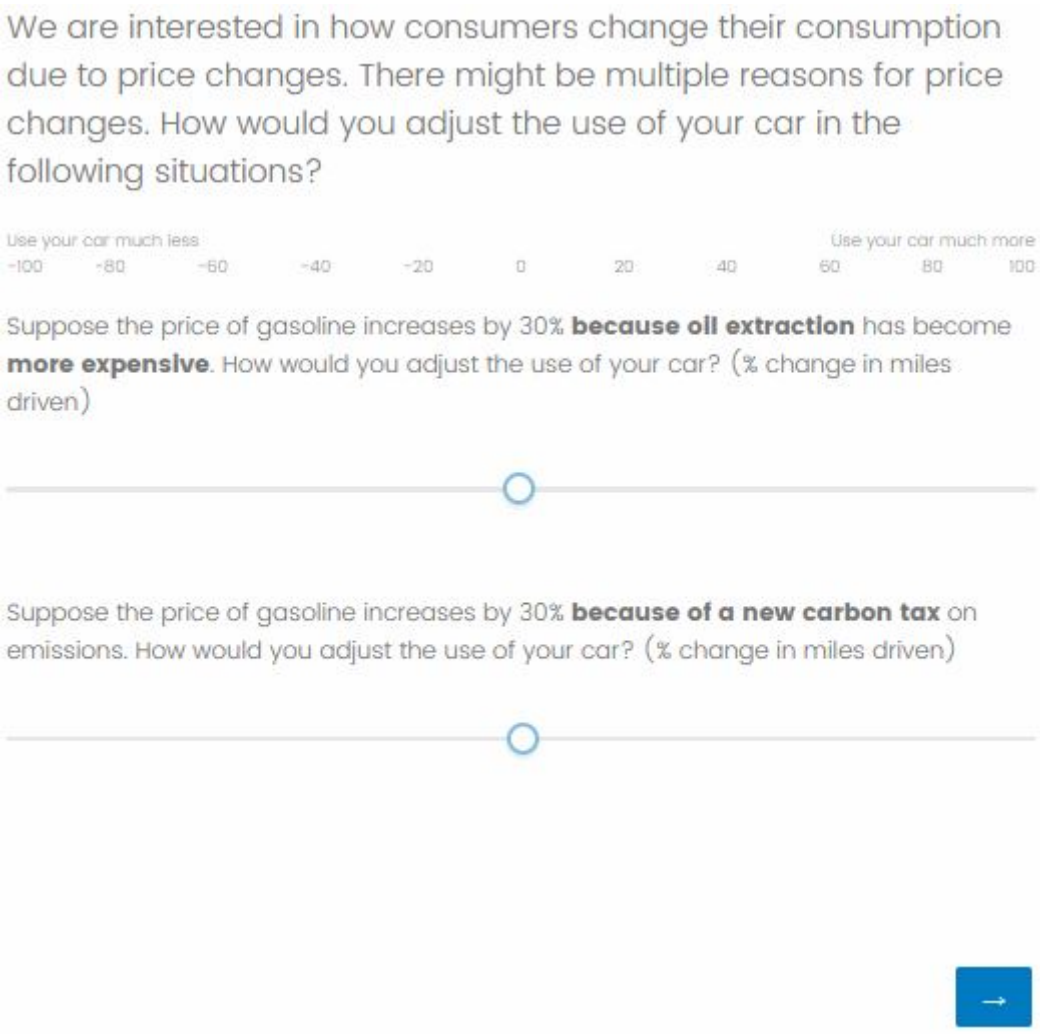
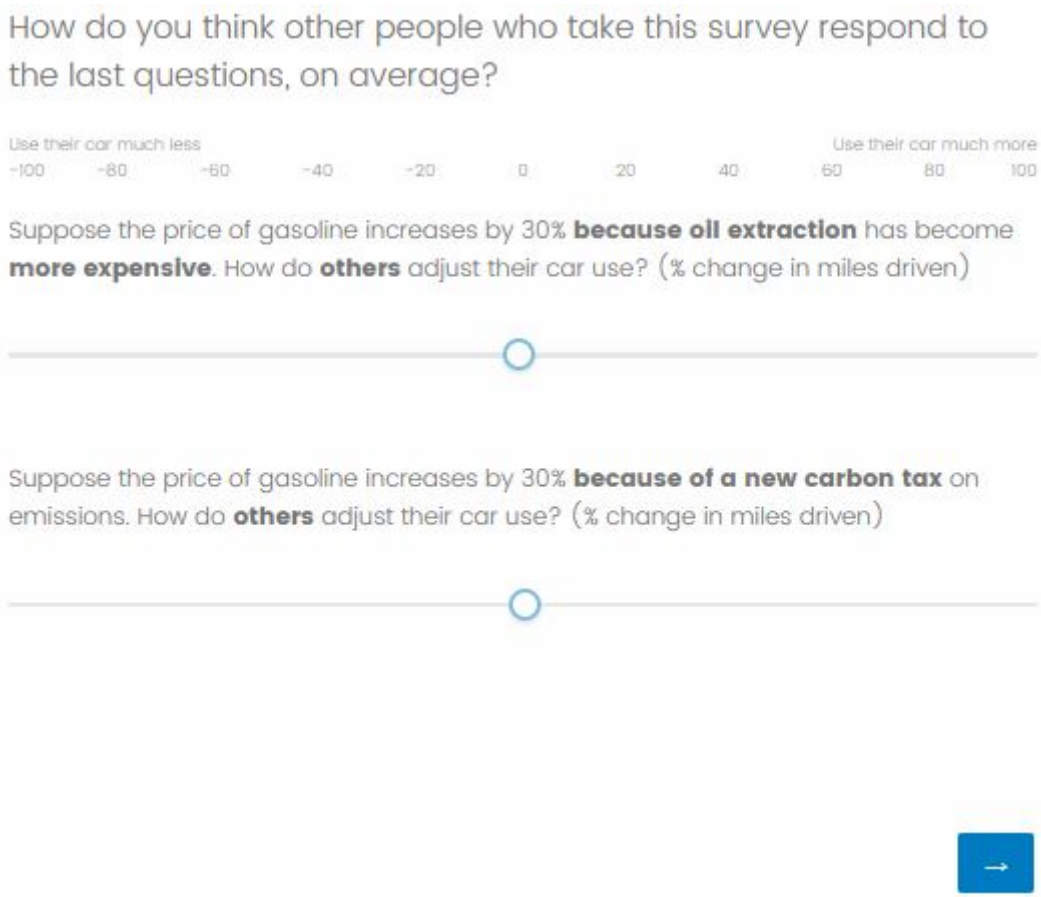


Figure 6.19: Price under carbon tax, hold-out version.



Social Cost of Carbon Block

Figure 6.20: Social cost of carbon information screen.

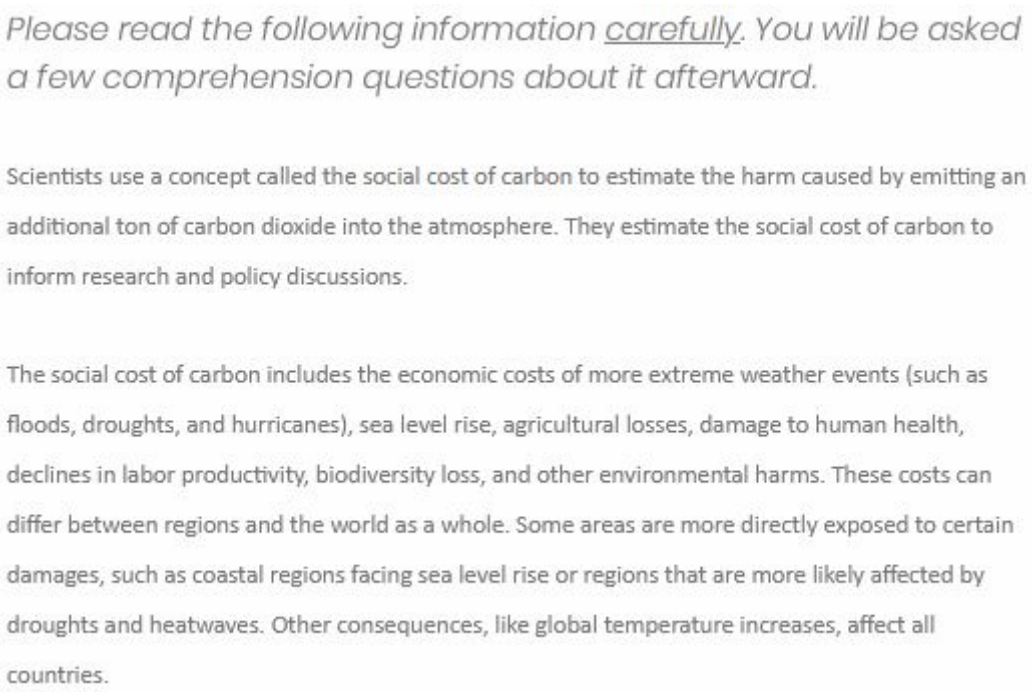


Figure 6.21: Social cost of carbon comprehension check.

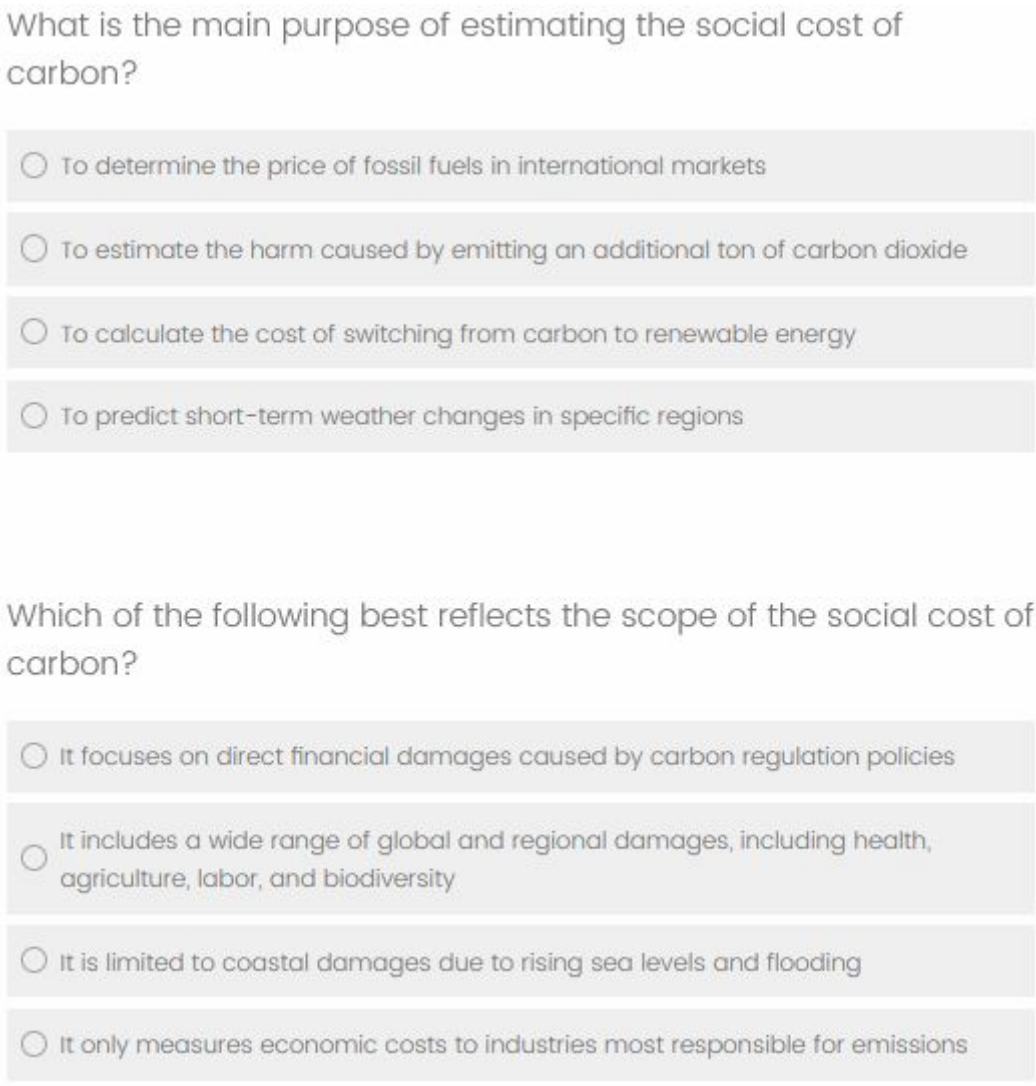


Figure 6.22: Social cost of carbon adjustment screen.

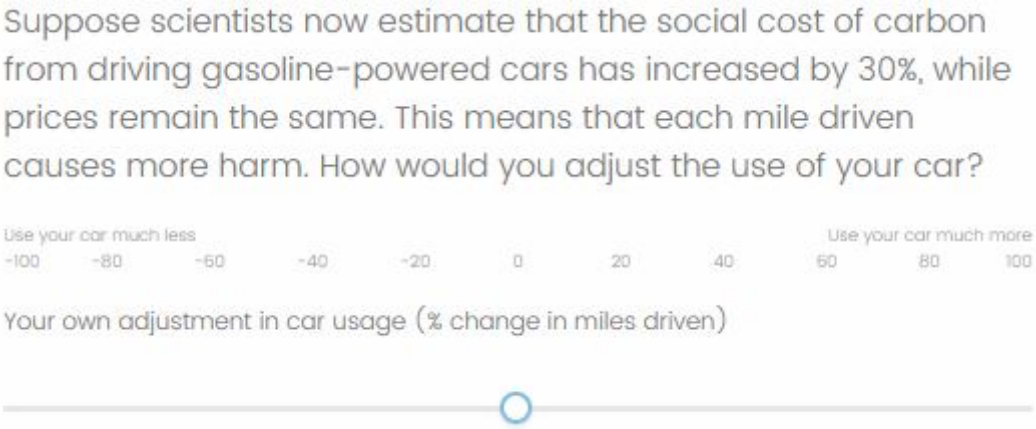
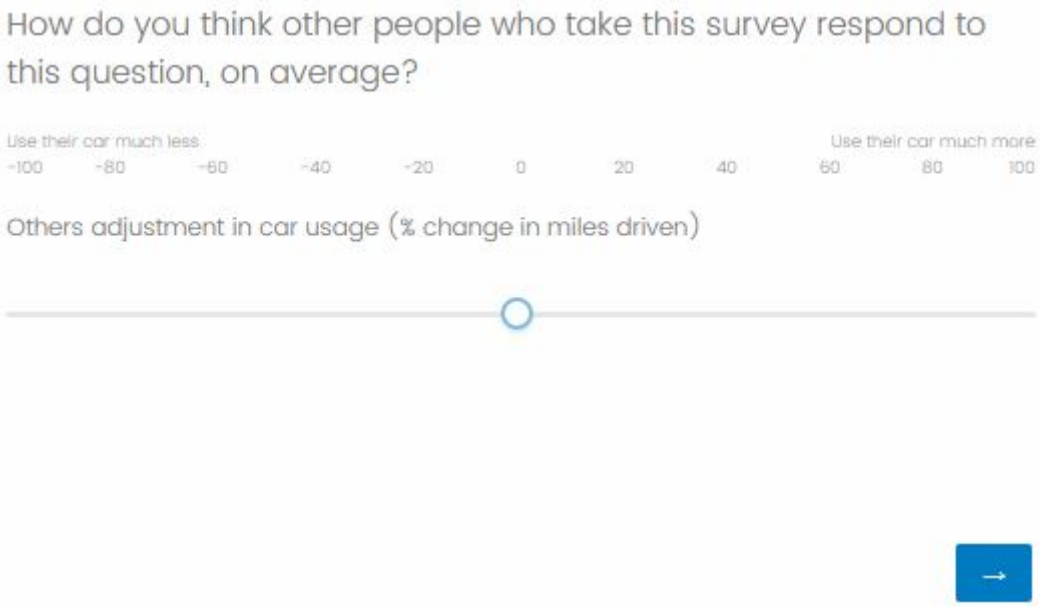


Figure 6.23: Social cost of carbon adjustment, hold-out version.



Policy Preferences: Taxation

Figure 6.24: Preferred tax level (UK baseline version).

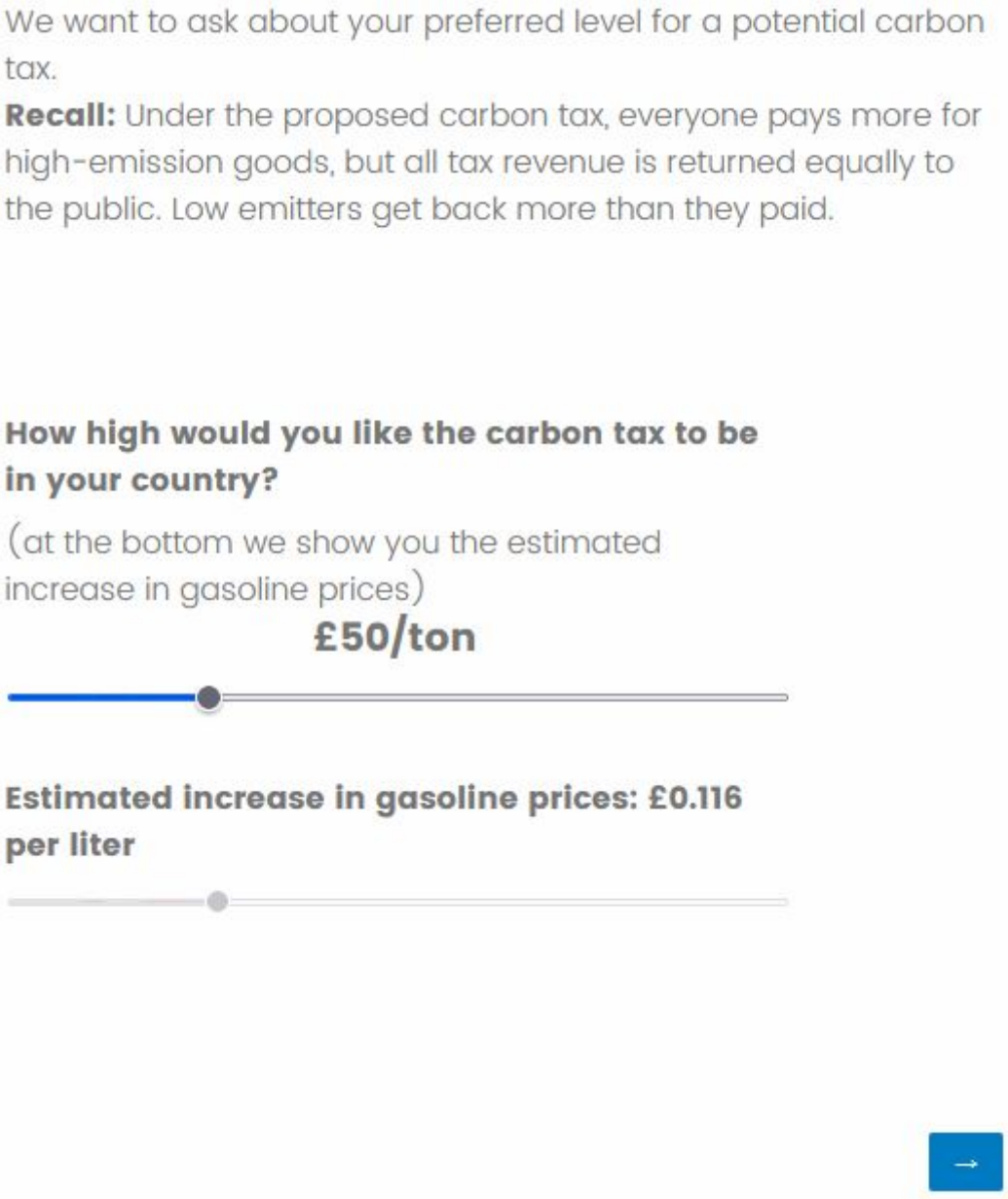


Figure 6.25: Preferred tax level, after the slider was moved.

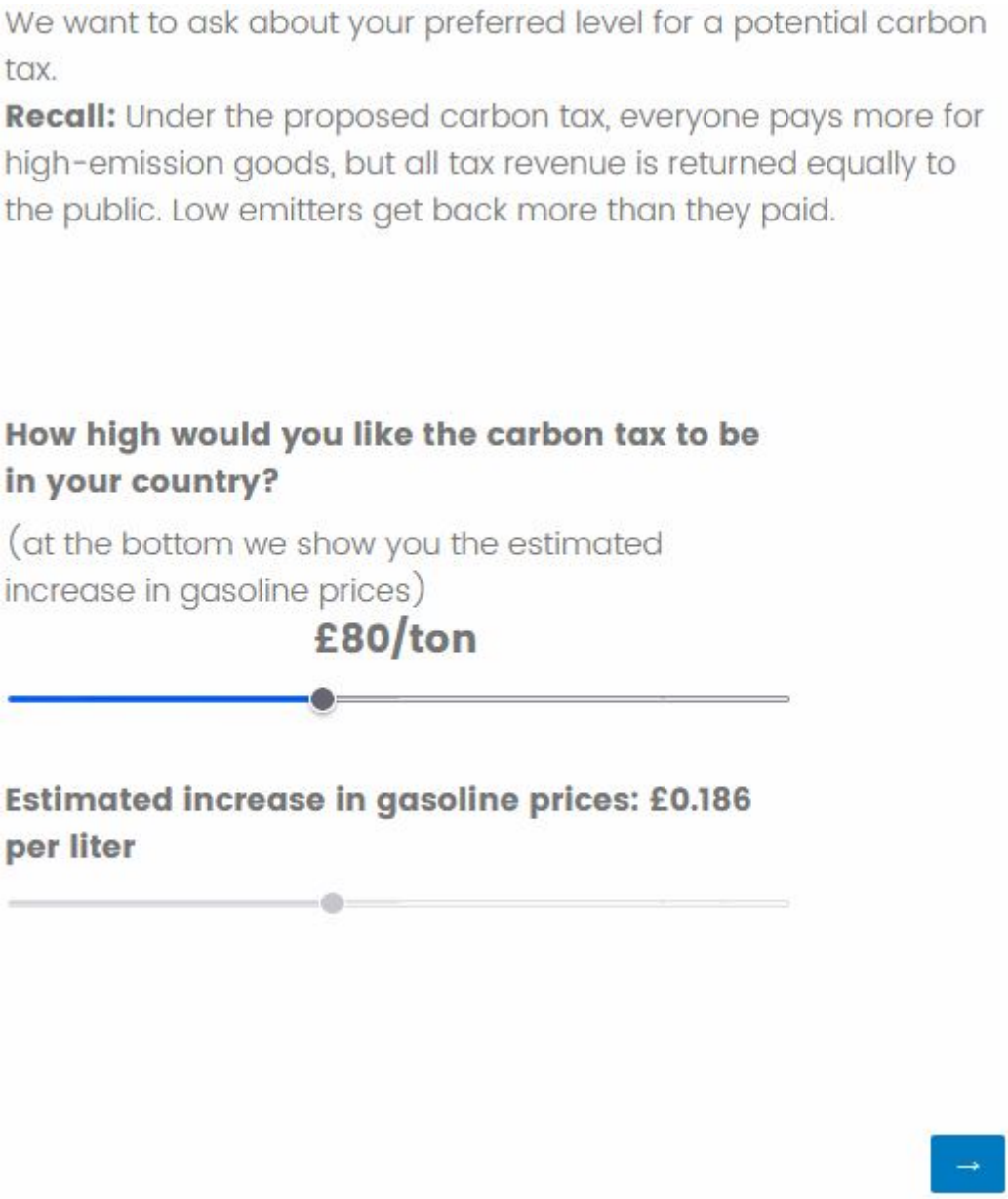


Figure 6.26: Preferred tax level (U.S. version). The preferred tax question was adapted to U.S. context (currency and reference points).

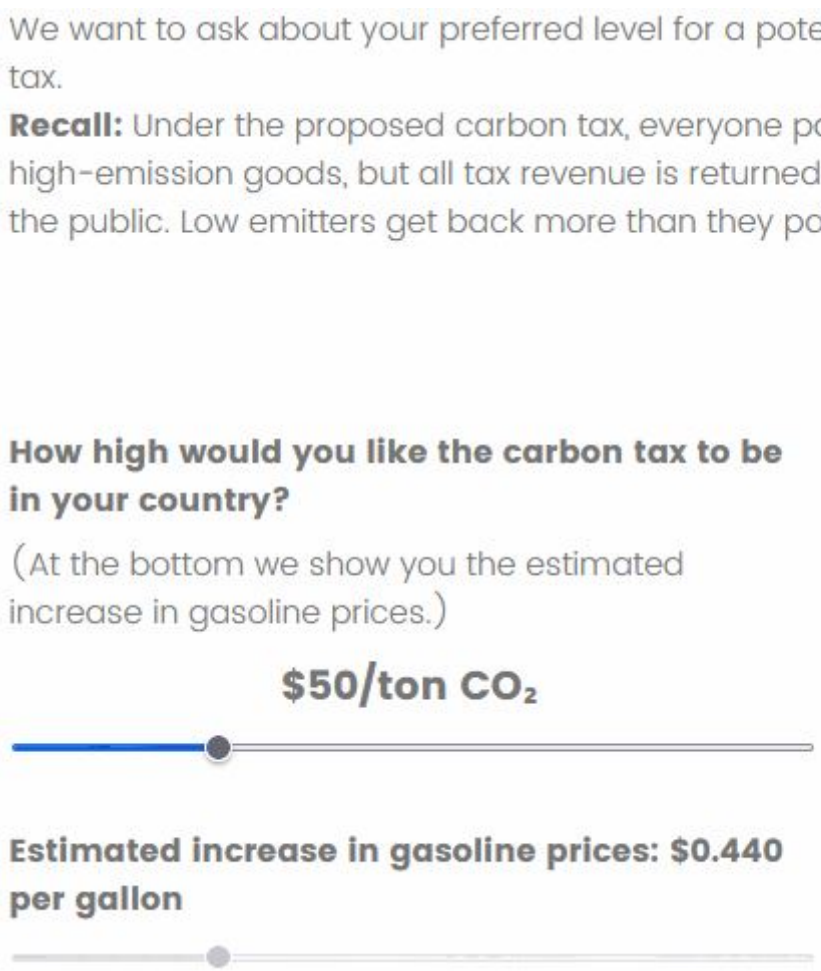


Figure 6.27: Local tax based on SCC (UK version).

Figure 6.28: Local tax based on SCC, after the slider was moved.

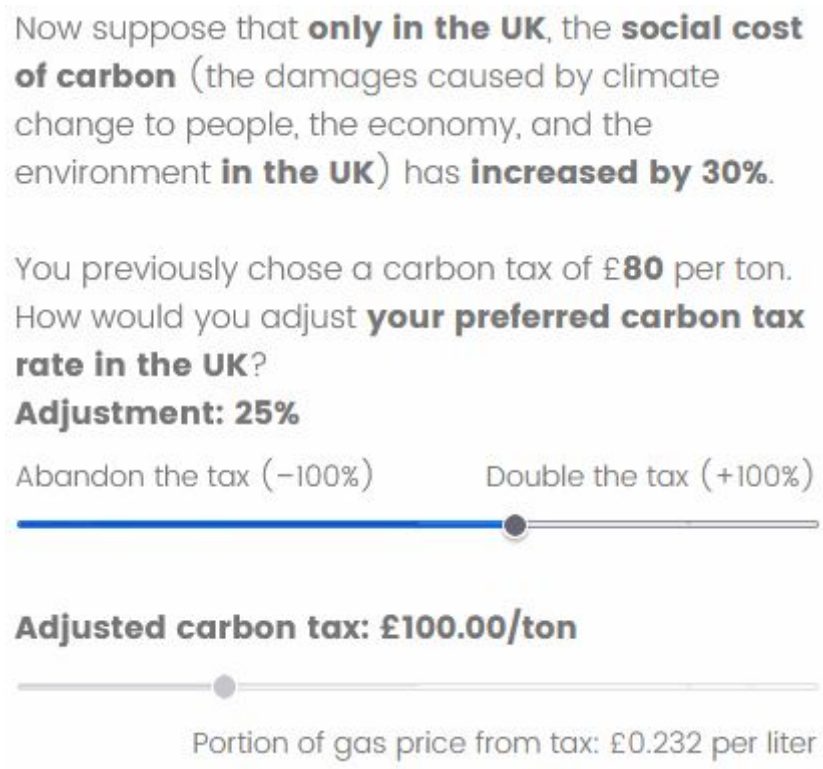


Figure 6.29: Local tax based on SCC (U.S. version). The local SCC-based tax question was adapted to U.S. context.

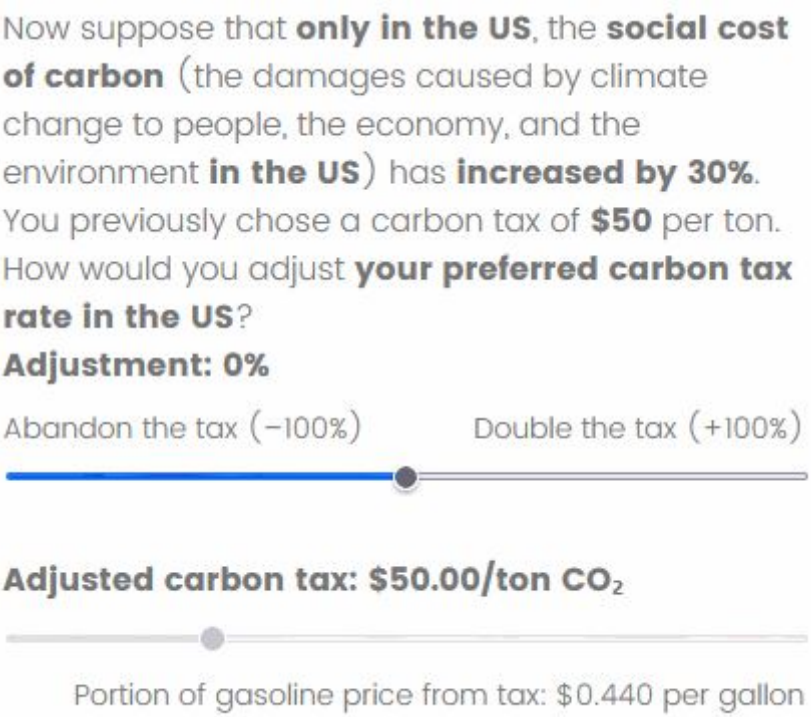


Figure 6.30: Local tax based on SCC, hold-out version.

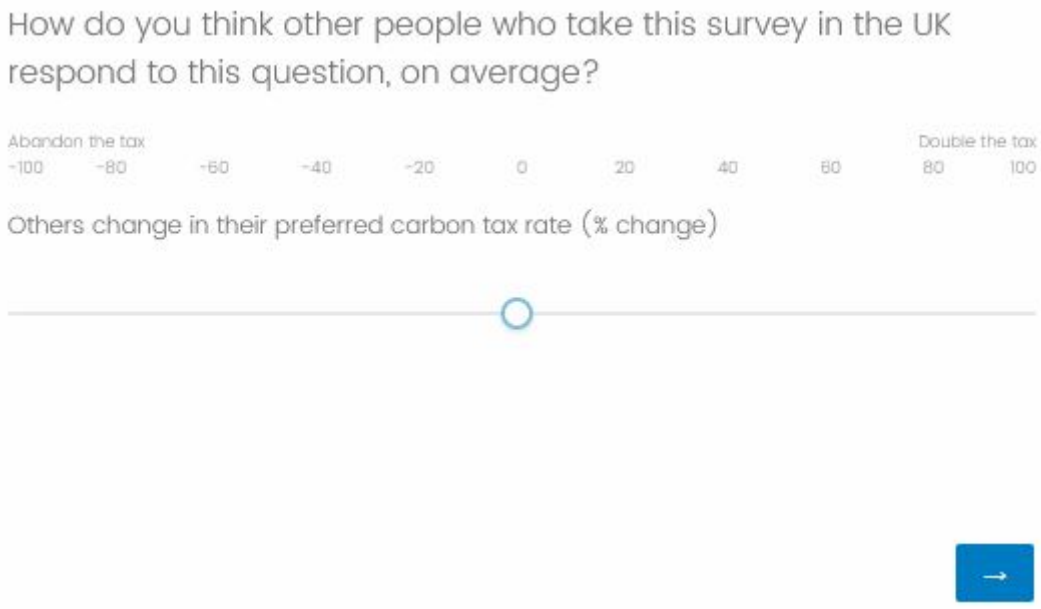


Figure 6.31: Global tax based on SCC (UK baseline version).

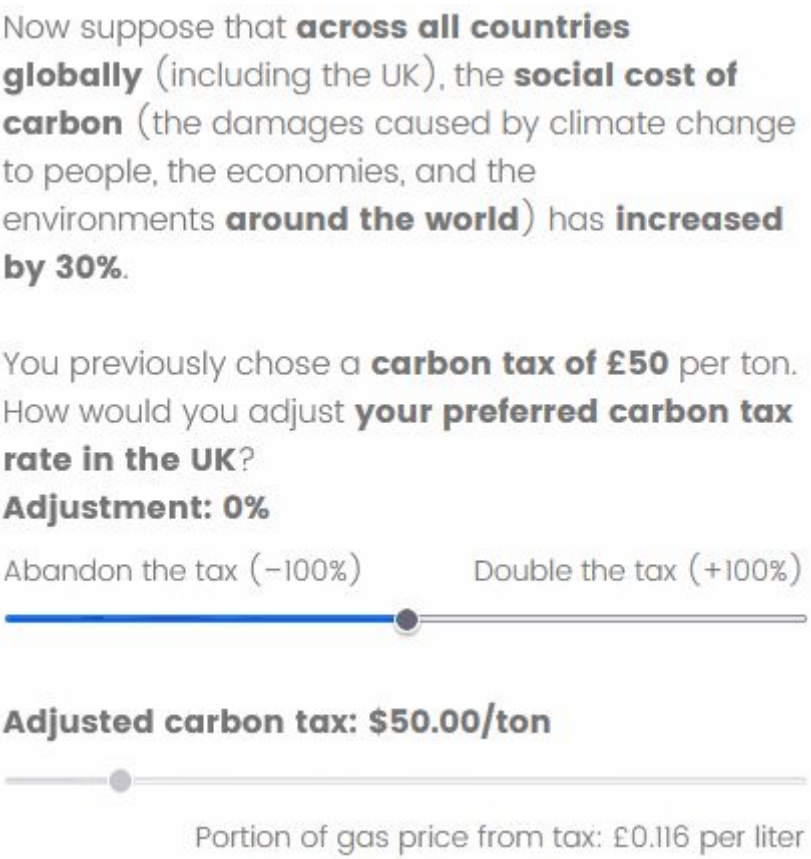


Figure 6.32: Global tax based on SCC, after the slider was moved.

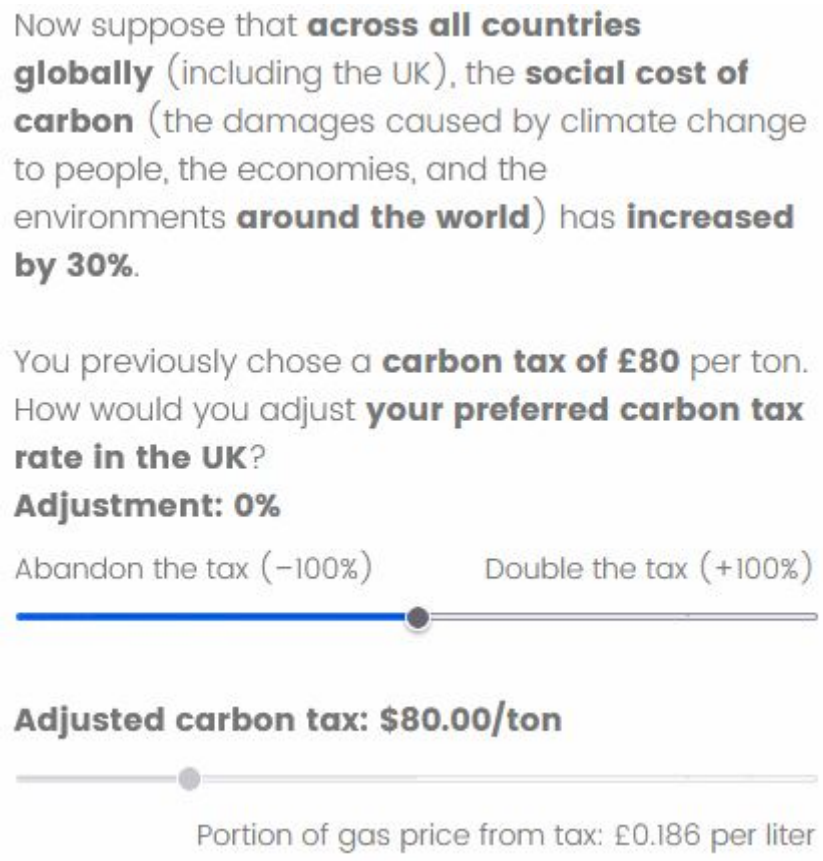
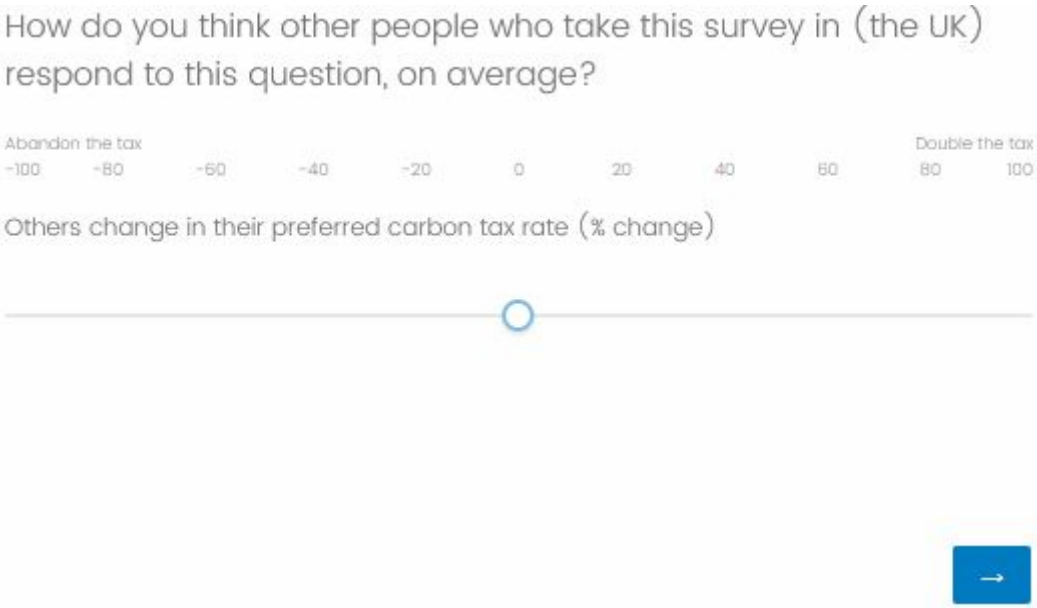


Figure 6.33: Global tax based on SCC (U.S. version). The global SCC-based tax question was adapted to U.S. context.

Figure 6.34: Global tax based on SCC, hold-out version.



Moral Pressure Block

Figure 6.35: Moral pressure, positive

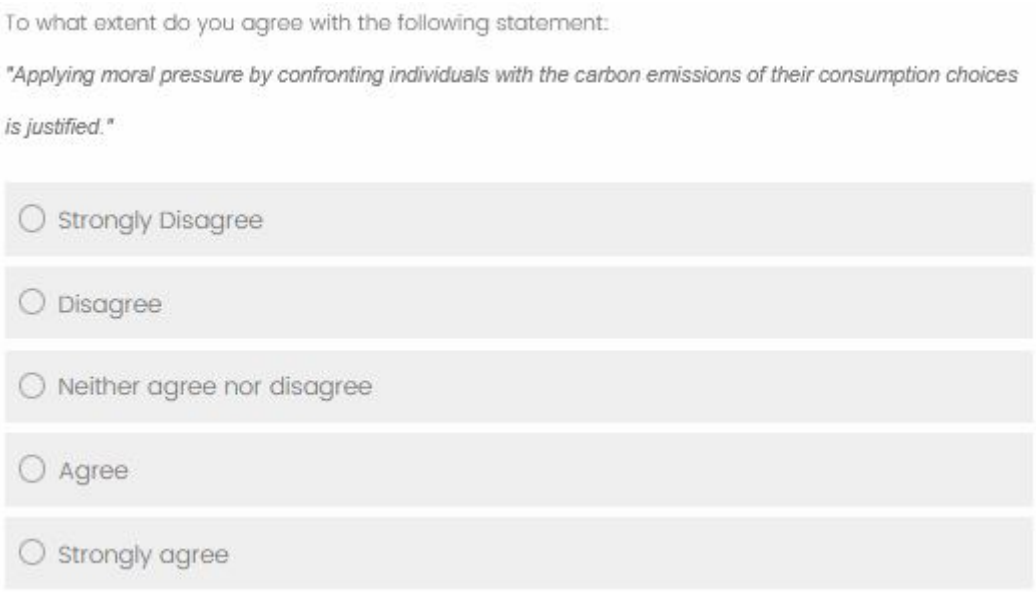


Figure 6.36: Moral pressure, negative

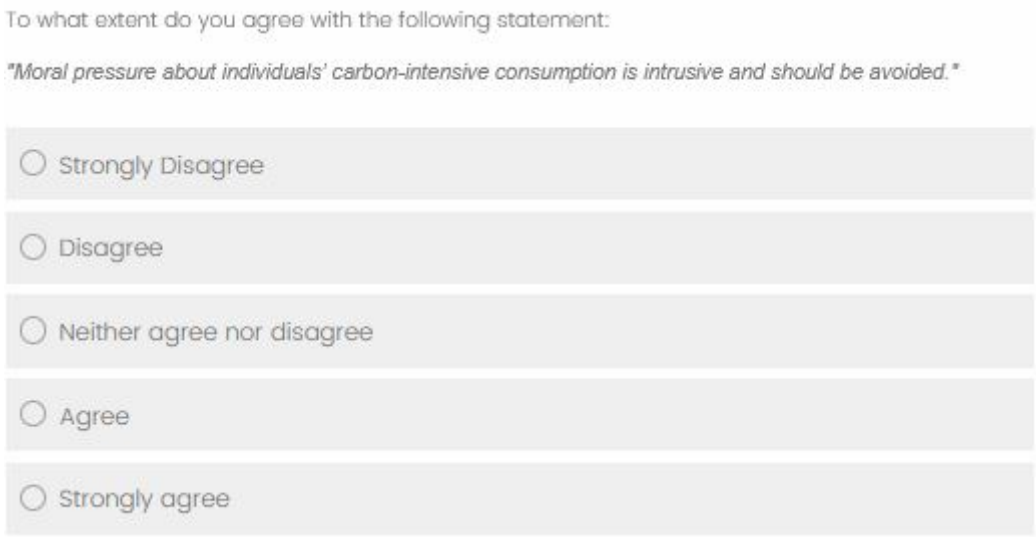


Figure 6.37: Response to Moral Pressure

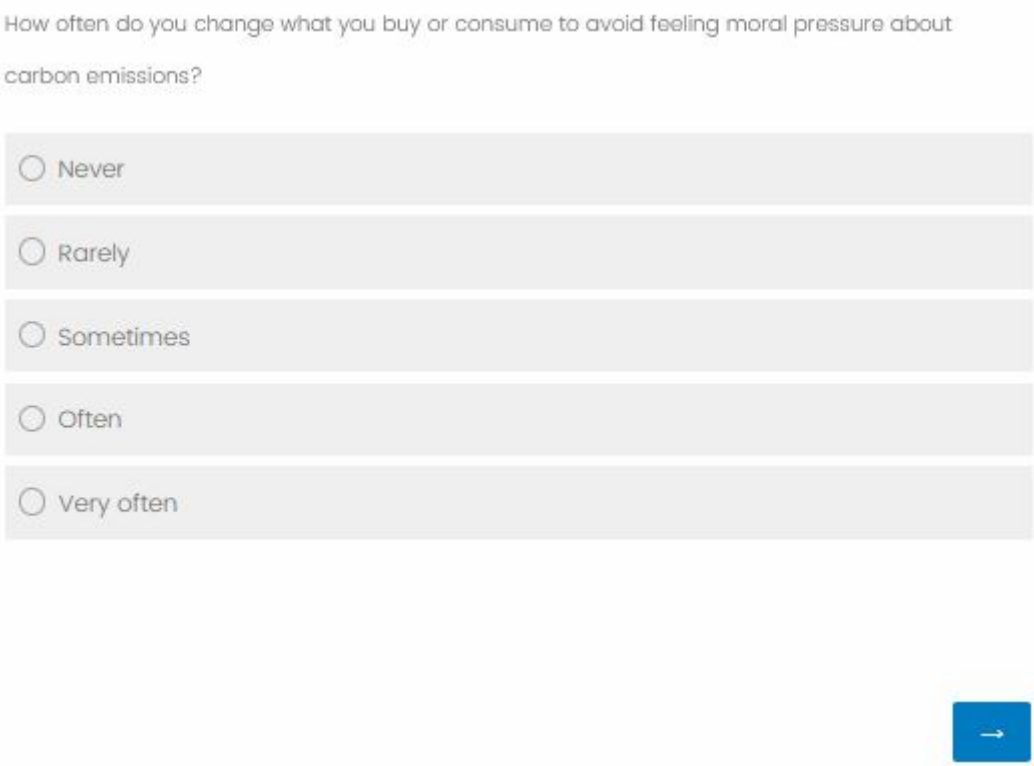


Figure 6.38: Price versus Harm

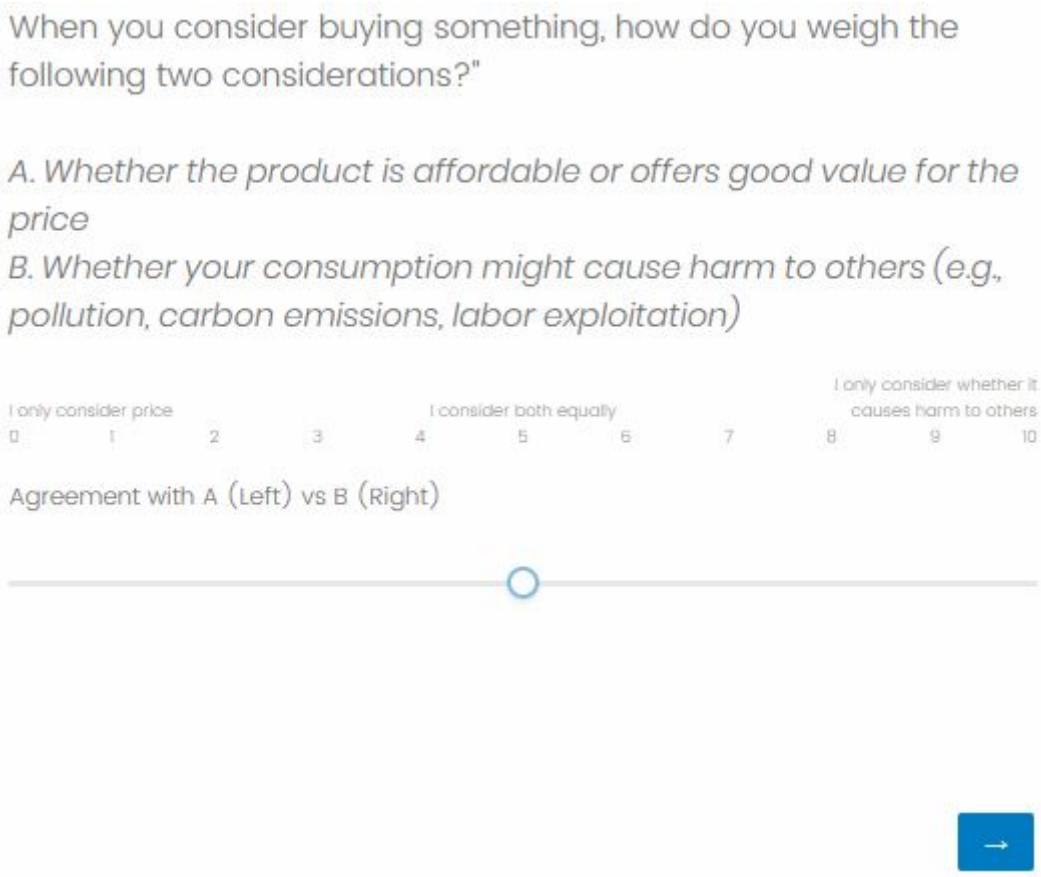
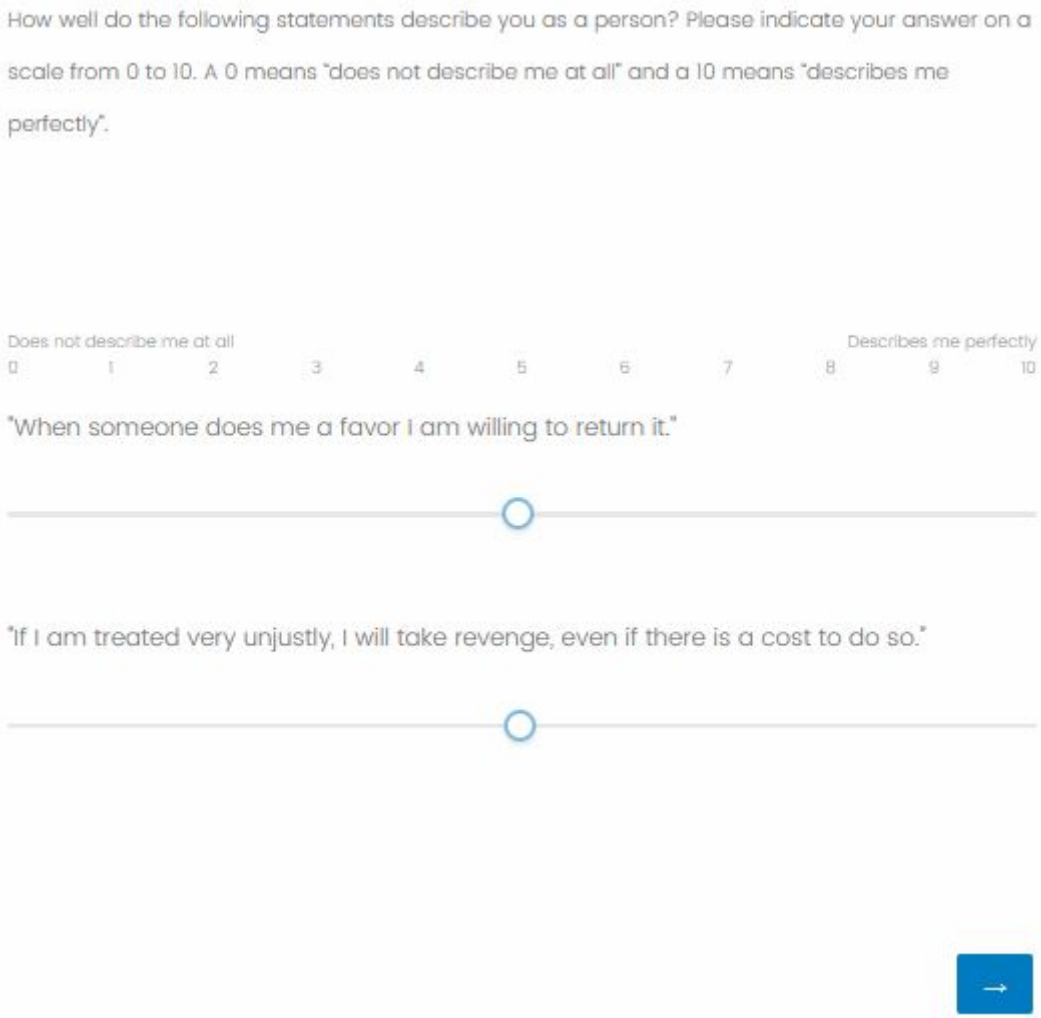


Figure 6.39: Reciprocity measure.



Clarity

Figure 6.40: Clarity check.

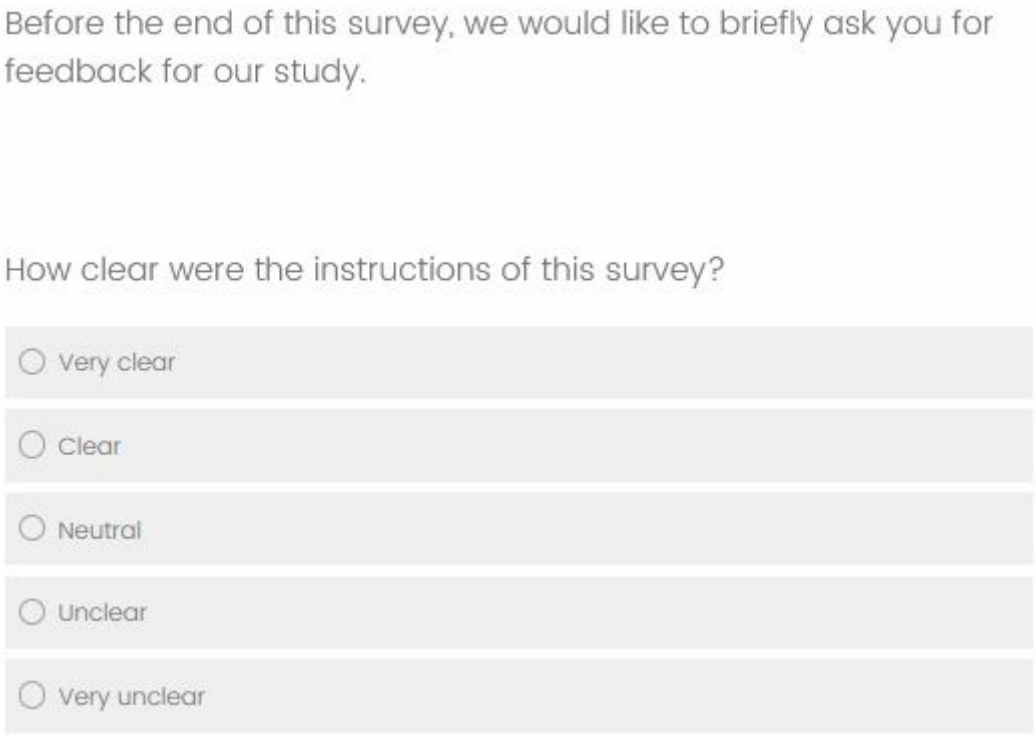
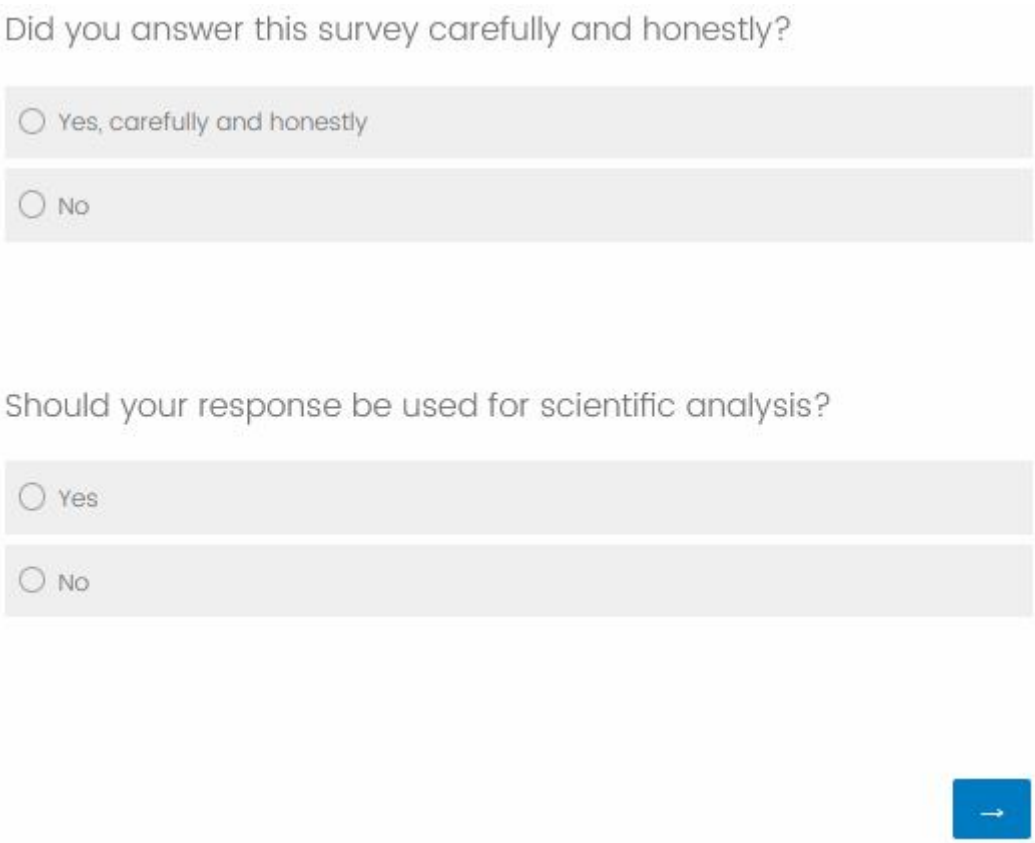


Figure 6.41: Self-assessment screen.



Chapter 7

Appendix for Software

The following specialist software was used in the preparation of this thesis:

- LaTeX (Lamport 1986)
- Overleaf (Hammersley and Lees-Miller 2025)
- R 4.1.2 (R Core Team 2021), with packages:
  - `car` (Fox and Weisberg 2019)
  - `DIDmultiplegtDYN` (de Chaisemartin and D'Haultfœuille 2024b)
  - `dplyr` (Wickham et al. 2023)
  - `fixest` (Bergé 2018)
  - `ggplot2` (Wickham 2016)
  - `lfe` (Gaure 2013)
  - `lmtest` (Zeileis and Hothorn 2002)
  - `MASS` (Venables and Ripley 2002)
  - `reshape2` (Wickham 2007)
  - `sandwich` (Zeileis 2004; Zeileis 2006; Zeileis et al. 2020)
  - `sf` (Pebesma 2018; Pebesma and Bivand 2023)
  - `stargazer` (Hlavac 2022)
  - `stringi` (Gagolewski 2022)
  - `systemfit` (Henningsen and Hamann 2007)
  - `tidyverse` (Wickham et al. 2019)
- RStudio 2022.07.2+576 "Spotted Wakerobin" (Posit Team 2022)
- Python 3.8.10, via Spyder 5.4.0 (Spyder Project Contributors 2022), with packages:
  - `matplotlib` (Hunter 2007)

  - `numpy` (Harris et al. 2020)
  - `scipy` (Virtanen et al. 2020)
- SQL Server Management Studio 15.0 (Microsoft Corporation 2022)
- Qualtrics (Qualtrics 2026)
- Prolific (Prolific 2026)
- Claude (Anthropic 2026)
- GPT-5 (OpenAI 2025)

Software Bibliography

Anthropic (2026). *Claude Opus 4.7 [Large language model]*.

Bergé, L. (2018). "Efficient Estimation of Maximum Likelihood Models with Multiple Fixed-Effects: The R Package `fixest`". In: *CREA Discussion Papers*.

de Chaisemartin, C. and X. D'Haultfœuille (2024b). "Difference-in-Differences Estimators of Intertemporal Treatment Effects". In: *Review of Economics and Statistics*.

Fox, J. and S. Weisberg (2019). *An R Companion to Applied Regression*. 3rd. Thousand Oaks, CA: Sage.

Gagolewski, M. (2022). "`stringi`: Fast and Portable Character String Processing in R". In: *Journal of Statistical Software* 103.2, pp. 1–59.

Gaure, S. (2013). "`lfe`: Linear Group Fixed Effects". In: *The R Journal* 5.2, pp. 104–117.

Hammersley, J. and J. Lees-Miller (2025). *Overleaf*.

Harris, C. R. et al. (2020). "Array Programming with NumPy". In: *Nature* 585, pp. 357–362.

Henningsen, A. and J. D. Hamann (2007). "`systemfit`: A Package for Estimating Systems of Simultaneous Equations in R". In: *Journal of Statistical Software* 23.4, pp. 1–40.

Hlavac, M. (2022). *`stargazer`: Well-Formatted Regression and Summary Statistics Tables*.

Hunter, J. D. (2007). "Matplotlib: A 2D Graphics Environment". In: *Computing in Science & Engineering* 9.3, pp. 90–95.

Lamport, L. (1986). *LaTeX: A Document Preparation System*. Reading, MA: Addison-Wesley.

Microsoft Corporation (2022). *SQL Server Management Studio*.

OpenAI (2025). *GPT-5 [Large language model]*.

Pebesma, E. (2018). "Simple Features for R: Standardized Support for Spatial Vector Data". In: *The R Journal* 10.1, pp. 439–446.

Pebesma, E. and R. Bivand (2023). *Spatial Data Science: With Applications in R*. Chapman and Hall/CRC.

Posit Team (2022). *RStudio: Integrated Development Environment for R*. Posit Software, PBC. Boston, MA.

Prolific (2026). *Prolific*.

Qualtrics (2026). *Qualtrics XM Platform*.

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria.

Spyder Project Contributors (2022). *Spyder: The Scientific Python Development Environment*.

Venables, W. N. and B. D. Ripley (2002). *Modern Applied Statistics with S*. 4th. New York: Springer.

Virtanen, P. et al. (2020). "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". In: *Nature Methods* 17, pp. 261–272.

Wickham, H. (2007). "Reshaping Data with the `reshape` Package". In: *Journal of Statistical Software* 21.12, pp. 1–20.

— (2016). *`ggplot2`: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.

Wickham, H., R. François, L. Henry, K. Müller, and D. Vaughan (2023). *`dplyr`: A Grammar of Data Manipulation*.

Wickham, H. et al. (2019). "Welcome to the `tidyverse`". In: *Journal of Open Source Software* 4.43, p. 1686.

Zeileis, A. (2004). "Econometric Computing with HC and HAC Covariance Matrix Estimators". In: *Journal of Statistical Software* 11.10, pp. 1–17.

— (2006). "Object-Oriented Computation of Sandwich Estimators". In: *Journal of Statistical Software* 16.9, pp. 1–16.

Zeileis, A. and T. Hothorn (2002). "Diagnostic Checking in Regression Relationships". In: *R News* 2.3, pp. 7–10.

Zeileis, A., S. Köll, and N. Graham (2020). "Various Versatile Variances: An Object-Oriented Implementation of Clustered Covariances in R". In: *Journal of Statistical Software* 95.1, pp. 1–36.

Curriculum Vitae

Marten Christian Ritterrath

University of Cologne – Cluster of Excellence: ECONtribute Markets & Public Policy
Albertus-Magnus-Platz, 50923 Cologne, Germany

marten.ritterrath@gmail.com | sites.google.com/view/martenritterrath

Research Fields

Environmental Economics · Innovation · Public Economics · Behavioral Economics

Current Position

VATT Institute For Economic Research *Helsinki, Finland*
PostDoctoral Researcher 2026 –

Affiliations

2024 –	DIW	Guest Researcher
2022 –	Center for Social and Economic Behavior (C-SEB)	Junior Researcher
2021 –	Cluster of Excellence ECONtribute: Markets & Public Policy	Associated Member

Education

University of Cologne *Cologne, Germany*
PhD Economics 2022 –

Central European University *Vienna, Austria*
Visiting Researcher 03.2026
- Host: Marc Kaufmann

University of California Los Angeles (UCLA) *Los Angeles, USA*
Visiting Researcher 01.2025 – 04.2025
- Host: Nico Voigtländer

CREST/École Polytechnique *Paris, France*
Visiting Researcher 10.2024 – 12.2024

- Host: Pierre Boyer

University of Cologne *Cologne, Germany*
Master of Economic Research (ranked 1/12) 2019 – 2022
- Master Thesis: *"Electricity Market Responses to Carbon Pricing in Europe: Emission Avoidance and the Relocation of Emissions"*

University of Cologne *Cologne, Germany*
BSc Economics 2014 – 2018
- Minor in Politics
- Dean's Award – *Awarded to the top 10% of students*

Hitotsubashi University *Tokyo, Japan*
Study Abroad 2017

Working Papers

Inventors' Personal Experience of Natural Disasters and Green Innovation *(with Lisa Keding)*

2024 Best Paper Award – Ivey-ARCS PhD Sustainability Academy

2025 YEP Doctoral Paper Award – ECONtribute (Universities Bonn & Cologne)

Work in Progress

Morals and the Political Economy of Carbon Taxes *(with Felix Bierbrauer, Mattias Polborn and Georg Weizsäcker)*

Risk Denial & Meritocracy *(with Melisa Kurtis)*

On the Equity-Efficiency-Tradeoff and the Social Cost of Carbon *(with Felix Bierbrauer and Niklas Taft)*

Conferences & Seminars (* scheduled, $^+$ invited)

2026 – **Central European University** (Vienna)

2025 – **UCLA Anderson** Internal Seminar (Los Angeles), **UC Berkeley** Internal Seminar$^+$ (Berkeley), **DIW** Cluster Seminar (Berlin), **Maastricht University** Interdisciplinary Brownbag Seminar of Behavioural Sciences$^+$ (Maastricht), **AERE** Summer Conference (New Mexico), **EAERE** Conference (Bergen), **Collège de France, PSE, INSEAD** Innovation Seminar$^+$ (Paris), **EEA** Conference (Bordeaux), **ESA** European Meeting (Brno)

2024 – **Royal Economic Society Conference** (Belfast), **RGS Doctoral Conference** (Essen), **EAERE** Conference (Leuven), **EAERE Summer School** on Data science in environmental economics research (Graz), **ZEW/MaCCI Conference on the Economics of**

Innovation and Patenting (Mannheim), **CREST/ENSAE** Internal Applied Seminar (Paris)

Refereeing

Journal of the European Economic Association

Grants

2025	**C-SEB**, 3000€ – Funding for a research visit at UCLA
2024	**Joachim Herz Stiftung**, 12500€ – Add-on Fellowship for Interdisciplinary Economics
2022–2025	**ECONtribute**, 8600€ – Funding for various projects

Teaching

Spring 2022–2025	**Core Module Macroeconomics**, for the BSc MES
Fall 2023 –	**Thesis Supervision**, Master 2x
Fall 2022 –	**Thesis Supervision**, Bachelor 16x
Fall 2016	**Statistics**, Teaching Assistant
2014	**English Teacher**, Suzhou Lanying (Blue Tassel) School – Suzhou, China

Professional Service & Activities

2024 – 2026	**Internal Applied Lunch Seminar**, Co-Organizer
2022 – 2024	**ECONtribute Seminar**, Co-Organizer

Languages

German (native), English (fluent), French (advanced)

Programming

Proficient in: R, Python, SQL

Eidesstattliche Erklärung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe.

Weitere Personen, neben den ggf. in der Arbeit aufgeführten Koautorinnen und Koautoren, waren an der inhaltlich-materiellen Erstellung der vorliegenden Arbeit nicht beteiligt. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten fremden Werken dem Wortlaut oder dem Sinn nach entnommen wurden, sind als solche kenntlich gemacht. Ich versichere an Eides statt, dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie – abgesehen von den angegebenen Teilpublikationen und eingebundenen Artikeln und Manuskripten – noch nicht veröffentlicht worden ist, sowie, dass ich eine Veröffentlichung der Dissertation vor Abschluss der Promotion nicht ohne Genehmigung des Promotionsausschusses vornehmen werde. Die Bestimmungen der Promotionsordnung im Promotionsprogramm Research in Management, Economics and Social Sciences der Wirtschafts- und Sozialwissenschaftlichen Fakultät der Universität zu Köln sind mir bekannt.

Darüber hinaus erkläre ich hiermit, dass ich die Leitlinien der Universität zu Köln zur Sicherung guter wissenschaftlicher Praxis gelesen und sie bei der Durchführung der der Dissertation zugrundeliegenden Arbeiten und der schriftlich verfassten Dissertation beachtet habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen. Ich versichere, dass die eingereichte elektronische Fassung der eingereichten Druckfassung vollständig entspricht. Ich versichere, dass ich nach bestem Wissen die reine Wahrheit gesagt und nichts verschwiegen habe.

Helsinki, 21.05.2026 Marten Christian Ritterrath