DISS. ETH NO. 28241

Essays in Public Finance and Climate Economics: Carbon Pricing and Capital Taxation

A thesis submitted to attain the degree of
Doctor of Sciences of ETH ZURICH
(Dr. sc. ETH Zurich)

presented by
Oliver Kalsbach
M.Sc. Industrial Engineering
born on 26.05.1992
citizen of the Federal Republic of Germany

accepted on the recommendation of
Prof. Dr. Sebastian Rausch
Prof. Dr. Lucas Bretschger

2022

Summary

This dissertation studies the role of public economic policy to mitigate climate change and economic inequalities. The thesis comprises three essays. Essay 1 and 2 take up on climate change mitigation. Essay 1 investigates optimal carbon pricing rules when private economic agents evaluate time differently than a benevolent planner. Essay 2 looks at optimal carbon pricing in the presence of capital income tax distortions. Essay 3 is concerned with how the taxation of labor and capital income can mitigate economic inequalities. More precisely, the essay shows how the societies' inequality aversion governs the optimal taxation of income.

Essay 1 theoretically and numerically characterizes optimal carbon prices in a dynamic, multi-sector, general equilibrium economy in which a benevolent planner values the wellbeing of future generations more than private agents. Motivation for such setting is that reaching climate neutrality requires large-scale transformations, for instance the economy must substitute away from fossil fuels to "green" capital in production. This makes thinking about the social valuation of capital accumulation in the transformation necessary. How private agents accumulate capital depends on their discount rate, however, society may place a higher value on the well-being of future generations than private agents do. Private agents then discount the future too much, which is why essay 1 allows the discount rates of a benevolent planner and private actors to differ.

Essay 1 emphasizes two main insights: first, the general wisdom in the field to price carbon uniformly across all emitting entities is only efficient under the restrictive assumptions that firms produce with identical technologies or that the household's and the planner's discount rates perfectly align. If, however, sectors produce with heterogeneous technologies and the planner places a higher welfare weight on future generations above the current generation's private altruism towards the future, non-uniform carbon prices are optimal. This arises because private agents discount the future too much and, thus, under-accumulate capital. Optimal non-uniform carbon prices then trigger the firms' capital demand and spur capital accumulation which benefits future generations, just as favoured by the planner's preferences with greater weights on future generations. Second, numerical results reveal that depending on sectoral technology heterogeneity, optimal carbon prices differ widely across sectors and yield substantial welfare gains relative to uniform pricing.

Essay 2 extends the decentralized equilibrium of essay 1 for a capital income tax that finances public consumption. The public economics literature shows that a capital income tax is distortive in a setting where a capital income tax raises the rental price for capital which lowers the firms' capital demand and eventually leads to capital under-accumulation. Contrarily to this theoretical insight, real world economies tax, however, capital incomes. Essay 2 thus asks how carbon emissions should be taxed in the presence of a capital income tax distortion.
The essay's main finding runs counter to the consensus in the field to price carbon emissions uniformly across all sectors. Optimal carbon prices are indeed non-uniform if capital income is taxed. The intuition is similar as in essay 1. Optimally differentiated sectoral carbon prices address the environmental externality and the under-accumulation of capital (due to the tax distortion). The sectoral carbon prices stimulate capital demand, i.e. they stimulate capital accumulation, and are thus beneficial to the economy. The degree of carbon price differentiation is driven by heterogeneous sectoral production technologies, in particular heterogeneity in

the elasticity of substitution between capital and carbon emissions. Heterogeneity in substitution elasticities between emissions and labor or on the final good level play less of role. Numerically, relative to a uniform price, sectoral carbon price differentiation is substantial, spurs the capital accumulation and leads to significant welfare gains.

Essay 3 examines optimal redistributive income taxes under different principles of justice. The essay builds on an exogenously growing Ramsey economy with two types of households: savers are intertemporal more elastic than workers, savers demand thus lower returns to capital and own all capital. Workers in contrast consume their entire labor income. The framework also accounts for various social preferences structures reflecting libertarian, egalitarian and utilitarian principles of justice.

The essay's insights are threefold: first, taxing capital income is optimal when the planner values the future with the workers' intertemporal elasticity of substitution (IES) instead of the savers' IES. The intuition is that a capital income tax disincentivizes the savers' capital investments and makes capital accumulate as preferred by the workers (who prefer more consumption smoothing than savers). Taxing capital income (or not) is thus a normative matter about the society's intertemporal preferences. This result opposes the common view that capital income should not be taxed as put forward by the public economics literature which (mostly) assumes a utilitarian perspective for the evaluation of household wellbeing. Assuming a utilitarian planner is thus equivalent to implicitly assuming a zero-capital income tax. Second, how much income the government redistributes from rich households to poor households depends on the planner's inequality aversion. For instance, a more libertarian planner (with less intratemporal inequality aversion relative to a utilitarian social welfare function) is less concerned with the distribution of income across households—and the planner redistributes less income in optimum. Third, capital income tax revenues should not be used to enhance the living conditions of poor households, instead optimality requires to return more capital tax revenues back to the savers. Empirically, savers supply labor less elastic, making the subsidy to their labor income less distortionary.

Kurzfassung

Diese Dissertation untersucht die Rolle von öffentlicher Wirtschaftspolitik um den Klimawandel abzuwenden und ökonomische Ungleichheiten zu mildern. Die Thesis umfasst drei Aufsätze. Aufsatz 1 untersucht optimale Regeln zur Bepreisung von Kohlenstoffdioxidemissionen (CO_2-Emissionen) wenn private, ökonomische Agenten die Zukunft unterschiedlich bewerten als die Gesellschaft. Aufsatz 2 analysiert optimale Regeln zur Bepreisung von CO_2 wenn eine Kapitaleinkommensteuer die Ökonomie verzerrt. Aufsatz 3 beleuchtet wie Arbeits- und Kapitaleinkommenssteuern ökonomische Ungleichheiten mildern können. Der Aufsatz zeigt auf wie soziale Normen bezüglich ökonomischer Ungleichheit die optimalen Einkommenssteuersätze beeinflussen.

Aufsatz 1 untersucht die optimale Ausgestaltung von sektor-übergreifenden CO_2-Preisen mit theoretischen und numerischen Methoden. Die Analyse beruht auf einem dynamischen, mehr-sektoralen, allgemeinen Gleichgewichtsmodel in welcher der Planer das Wohlergehen zukünftiger Generationen mehr gewichtet als der Haushalt. Hintergrund dafür ist folgender: Um Klimaneutralität zu erreichen, müssen Ökonomien von Grund auf transformiert werden. Firmen müssen weg von fossilen Brennstoffen hinzu "grünem" Kapital substituieren, was es nötig macht darüber nachzudenken wie Kapital in einer Volkswirtschaft akkumuliert werden sollte. Festzuhalten ist, dass die Kapitalakkumulation durch die Diskontrate der privaten Haushalte gesteuert wird, jedoch vermag die Gesellschaft die Zukunft unter-

schiedlich diskontieren, i.e. die Gesellschaft gewichtet das Wohlbefinden zukünftiger Generationen mehr als private Haushalte. Um der Logik gerecht zu werden, dass private Haushalte die Zukunft zu stark diskontieren, unterscheidet dieser Aufsatz deshalb zwischen einer privaten Diskontrate der Haushalte und einer sozialen Diskontrate der Gesellschaft.
Aufsatz 1 betont zwei Ergebnisse: Erstens, das verbreitete Verständnis CO_2 Emissionen mit einem Einheitspreis zu regulieren ist nur unter restriktiven Annahmen effizient, nämlich dass Firmen mit identischen Technologien produzieren und dass die Diskontraten des Haushalts und des Regulierers übereinstimmen. Uneinheitliche CO_2-Preise sind jedoch optimal wenn Firmen mit heterogenen Technologien produzieren und der Planer das Wohlergehen zukünftiger Generationen stärker gewichtet als der Haushalt. Der Haushalt akkumuliert jedoch zu wenig Kapital wenn der Planer die Zukunft schwächer diskontiert als der Haushalt. Es ist dann optimal von dem Einheitspreisprinzip abzuweichen. Uneinheitliche CO_2-Preise sollen die Kapitalnachfrage der Firmen und somit die Kapitalakkumulation der Haushalte anreizen, ganz so wie es die Präferenzen des Planers mit einer stärkeren Gewichtung für künftige Generationen vorsehen. Zweitens, die numerischen Ergebnisse zeigen, dass—abhängig von der sektoralen Technologieheterogenität—sich die optimalen CO_2-Preise stark von Sektor zu Sektor unterscheiden und zu signifikanten Wohlfahrtsgewinnen gegenüber einem Einheitspreis führen.

In **Aufsatz 2** führe ich eine Kapitaleinkommensteuer in das dezentrale Gleichgewicht aus Aufsatz 1 ein. Die Kapitaleinkommensteuer dient der Finanzierung von Staatsausgaben. Die finanzwissenschaftliche Literatur zeigt, dass eine Kapitaleinkommenssteuer in diesem Modellrahmen verzerrend wirkt weil sie den Kapitalpreis in die Höhe treibt, so die Kapitalnachfrage senkt und schließlich zu einer zu geringen Kapitalakkumulation führt. Im Widerspruch zu diesem theoretischen Verständnis besteuern reale Volkswirtschaften jedoch Kapitaleinkünfte. In Aufsatz 2 stelle ich deshalb die Frage, wie CO_2-Emissionen bepreist werden sollten wenn eine Kapitaleinkommenssteuer die Ökonomie verzerrt.

Das zentrale Ergebnis widerspricht dem verbreiteten Verständnis, CO_2-Emissionen einheitlich zu bepreisen. Ich zeige, dass optimale CO_2-Preise uneinheitlich sind wenn Kapitaleinkommen besteuert werden. Die Intuition ist ähnlich wie in Aufsatz 1. Optimal differenzierte CO_2-Preise internalisieren die externen Umweltkosten und adressieren gleichzeitig die zu geringe Kapitalakkumulation (aufgrund der Steuerverzerrung). Die sektoralen CO_2-Preise stimulieren die Kapitalnachfrage, d.h. sie regen Investitionen in den Kapitalstock an und sind somit für die Ökonomie von Nutzen. Hauptfaktor für das Ausmaß der CO_2-Preisdifferenzierung sind heterogene, sektorale Produktionstechnologien, insbesondere die Heterogenität hinsichtlich der Substitutionselastizität zwischen Kapital und CO_2-Emissionen. Heterogenität in den sektoralen Substitutionselastizitäten zwischen Emissionen und Arbeit oder in der Aggregationen von sektoralen Output spielen eine untergeordnete Rolle. Aus numerischer Sicht ist die sektorale Differenzierung des CO_2-Preises signifikant, sie stimuliert die Kapitalakkumulation und führt im Vergleich zu einem Einheitspreis zu beträchtlichen Wohlfahrtsgewinnen.

Aufsatz 3 untersucht die optimale Umverteilung von Einkommen mittels Steuern unter Berücksichtigung verschiedener Gerechtigkeitsprinzipien. Der Aufsatz basiert auf einer exogen wachsenden Ramsey-Ökonomie mit zwei Haushaltstypen. Sparer sind intertemporal elastischer als Arbeiter, Sparer verlangen daher geringere Kapitalrenditen und besitzt somit das gesamte Kapital. Arbeiter hingegen konsumieren ihr gesamtes Arbeitseinkommen. Der Ansatz berücksichtigt auch verschiedene soziale Präferenzen, nämlich libertäre, egalitäre und utilitaristische Prinzipien der sozialen Gerechtigkeit.

Der Aufsatz liefert drei wesentliche Erkenntnisse. Erstens ist eine Kapitaleinkommensteuer optimal wenn der Planer die Zukunft mit der intertemporalen Substitutionselastizität (EIS) der Arbeiter und nicht mit der EIS der Sparer bewertet. Kapitaleinkommen zu besteuern ist somit eine normative Frage über die intertemporalen

Präferenzen der Gesellschaft. Die Intuition hierfür ist, dass eine Kapitaleinkommenssteuer die Investitionen der Sparer hemmt und somit zu einer geringen Kapitalakkumulation führt, eben wie es die Arbeiter mit ihrer Präferenz für eine stärker Konsumglättung bevorzugen. Dieses Ergebnis steht im Kontrast zu der verbreiteten Ansicht, dass Kapitaleinkommen nicht besteuert werden sollten—beruhend auf der Annahme das der Planner das Wohlbefinden der Haushalte aus einer utilitaristischen Perspektive beurteilt. Die Hypothese eines utilitaristischen Planers ist daher gleichbedeutend mit der impliziten Annahme Kapitaleinkommen nicht zu besteuern. Zweitens hängt die Frage wie viel Einkommen der Staat von wohlhabenden Haushalten an mittellose Haushalte umverteilt von den Gerechtigkeitsvorstellungen des Planers ab. Ein eher libertärer Planer (mit einer schwächeren intratemporalen Ungleichheitsaversion als ein utilitaristischer Planer) ist beispielsweise weniger an der Einkommensumverteilungen interessiert—und der Planer verteilt weniger Arbeitseinkommen um. Drittens sollten die Kapitalertragssteuereinnahmen nicht dazu verwendet werden die Lebensbedingungen ärmerer Haushalte zu verbessern; stattdessen sollte ein grösserer Teil der Kapitalsteuereinnahmen an die Sparer zurückfließen. Empirisch gesehen ist das Arbeitsangebot der Sparer weniger elastisch, sodass das Arbeitseinkommen der Sparer zu "subventionieren" die Ökonomie weniger verzerrt.

Acknowledgments

First, I would like to thank my primary thesis advisor, Sebastian Rausch, for giving me the opportunity to pursue my doctorate under his supervision. Thank you, Sebastian, for teaching me the economic mindset, for taking your time to advance me as a person and my work, and, of course, for your endless efforts to improve my writing.

Next, I want to thank my co-advisor, Lucas Bretschger. Thank you, Lucas, for always being open to discuss my work and for welcoming me with open arms into your group which has offered me a great working environment.

I want to thank Massimo Filippini. Thank you, Massimo, for your support throughout the last two years and for chairing my PhD examination committee.

A special thanks goes to Jan Abrell and Florian Landis. Thank you, Jan and Florian, for always challenging my thoughts critically and for making an important contribution to teach an ignorant PhD student the economic mindset.

Very important throughout my PhD have been Adrian, Gustav, Noe, Claudio and Fabio. I am glad to have made good friends with you here at ETH. At this point, I also want to thank my friends outside of ETH.

Next, I thank all my colleagues inside and outside of ETH. I would like to mention a few who helped me in particular along the way or made my time more comfortable. Thank you Mirjam, Clemens, Renger, Rina, Tobias, Aleksei, Clement, Christos and Alena.

My very special thanks, however, goes to my parents Ellen and Wolfgang. Thank you for your continuous support and trust, words can not express my gratitude. I want to thank my siblings Andreas, Leonard, Luisa & family and Tobias, too. Thank you for being you. Of course, all of this also holds for my partner Miriam. This dissertation is devoted to you.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Combating climate change and mitigating economic inequalities are two important economic challenges (Blanchard and Tirole 2021). Climate change is problematic for various reasons—it will lead to a rising sea level, extreme weather events and a greater spread of diseases. Also, it will complicate food production when crops and vegetation are unable to adapt to the new environmental conditions (Halsnaes et al. 2007). Economic inequalities may be problematic because a concentration of income at a small number of households reduces the possible market size which is detrimental for profitable innovations. Less innovation, however, lead to slower growth and thus less economic prosperity in the future (Schmookler 1966).

Climate change and economic inequalities are pressuring issues. The negative consequences when not addressing these two economic challenges are expected to be severe. For instance, climate change accelerates if all sectors continue to emit carbon dioxide into the atmosphere. An equally undesirable development is that differences in the households' wealth are likely to continue growing, given the current uneven distribution of capital and labor incomes. To overcome these issues, policy countermeasures are urgently needed.

There is a consensus among economists that carbon pricing is a cost-effective instrument to combat climate change. Putting a price on carbon emissions makes the use of fossil fuels more expensive, firms then substitute from ”dirty” to ”green” inputs in production and emit thus less carbon. But how shall carbon prices be designed when ”green” capital, a substitute for fossil fuels, accumulates insufficiently because households discount the future too much or because capital income taxes distort the capital formation? Does an insufficient capital accumulation alter the dogma of pricing carbon emissions uniformly across the economy? Equally important is to mitigate economic inequalities with adequate instruments. For instance, how should a policymaker make use of redistributive income taxes to shape the distribution of income according to the social preferences? And, are capital income taxes really undesirable as put forward by a strand of literature in the field of public economics? If capital income taxes are indeed desirable, should the tax revenues be used to make low-income households better-off?

This dissertation looks at optimal carbon pricing to combat anthropogenically caused climate change and at redistributive income taxes to mitigate economic inequalities. Essay 1 studies optimal carbon pricing when the economic activity rests on different sectors, capital is a key input for the transformation towards a climate-neutral economy and a benevolent planner puts a greater weight on future generations than private actors do. Essay 2 asks if carbon emissions should really be priced uniformly across all sectors in the presence of a capital income tax distortion. Essay 3 focuses on how to mitigate economic inequalities. It asks how intra- and inter-generational equity concerns change the optimal re-distributive (capital and labor) income taxation scheme.

This dissertation's results can be summarized as follows. Essay 1 and 2 provide alternative views in light of the conventional wisdom to price carbon emissions uniformly across the economy and find that optimal carbon prices are indeed non-uniform. The intuition is that households or markets under-accumulate capital when they discount the future too much as argued in essay 1 or when a capital income tax distorts the households' savings behavior as assumed in essay 2. The planner uses thus optimally differentiated carbon prices to stimulate the firms' demand for capital, and households invest more resources in the capital stock. The degree of carbon price differentiation is driven by sectoral technology heterogeneity. Sectors that substitute well between capital and emissions receive a higher carbon price, sectors that treat capital and emissions more complementary in production receive a lower carbon price. This price deviation from uniform increases the aggregated capital demand, incentivises capital investments and leads to intertemporal efficiency gains. Essay 3 provides a novel motivation for the taxation of capital incomes. Essay 3 builds on an economy in which capital ownership is determined by the households' heterogeneous preferences for intertemporal substitution. The paper argues that a zero capital income tax is only optimal when the planner evaluates the future with the intertemporal preference of households with capital holdings. Capital income should, however, be taxed when the planner evaluates the future with the intertemporal preference of households without capital holdings. The intuition is that the capital income tax makes capital accumulate as preferred by households without capital holdings. Also, capital income tax revenues should not be used to make low-income households better-off. Empirically, these households supply labor more elastic, so allocating the major share of tax revenue to them creates a too great distortion.

Scientific contributions

This thesis contributes to the public and climate economics literature economics literature on topics of optimal carbon pricing and optimal re-distributive income taxation.

Essay 1

Economists have long argued that a uniform carbon price on carbon dioxide (CO_2) is the most cost-efficient carbon abatement instrument. For instance, if a heavily coal-based electricity sector has better opportunities for substituting fossil energy for capital than private transport or buildings, the electricity sector should reduce emissions more than these other sectors. This is exactly the idea behind a uniform carbon price under which the electricity sector abates carbon emissions up to the point where it reaches the other sectors' marginal costs of carbon abatement. The challenge to de-carbonize entire economies, however, is enormous. To reach carbon neutrality, carbon pricing must stimulate all economic sectors to substitute away from fossil fuels to "green" inputs. A central role in this context is how private households build up "green" capital over time. Optimal carbon pricing in a multi-sector economy is thus intricately linked to how private and public decision-makers value the costs and benefits of climate policy over time. Intuitively, if society or a benevolent social planner places a higher value on the well-being of future generations than private agents do, to which I refer as differential social discounting, private agents discount the future too much and the government should promote future oriented policies that increase the capital stock. Of course, I am not the first to see the relevance of differential social discounting for the design of environmental policies (Kaplow et al. 2010, Goulder and Williams 2012, van der Ploeg and Rezai 2019, von Below 2012, Barrage 2018). It is, however, all the more surprising that the economic discipline has so far overlooked to ask if a uniform carbon price across all sectors is really desirable if private and social discount rates differ. This essay[1] contributes to the literature by attempting to fill this gap.

First, I employ a dynamic, multi-sector, general equilibrium economy in which

[1] This article is joint work together with Sebastian Rausch (SR). Oliver Kalsbach (OK) is the sole contributor to developing and applying the numerical model and producing quantitative results. OK and SR have equally contributed to the analysis of quantitative results and to the writing and exposition of the paper. "I" should thus be read as "we".

private agents and the social planner discount utility differently. I show theoretically and numerically that optimal carbon prices are non-uniform in such setting. The intuition is that private agents discount the future too much and thus under-accumulate capital. Optimal non-uniform carbon prices aim at stimulating the firms' capital demand which then triggers the household's capital accumulation and makes future generations better off—just a preferred by the more patient social planner.

Second, technology heterogeneity matters for the degree of carbon price differentiation. If all sectors produce with identical technologies, certainly an oversimplified representation of the production side, optimal carbon prices are uniform. If sectors produce, however, with heterogeneous technologies, optimal carbon prices are non-uniform. I find that sectors that substitute well between capital and emissions receive a higher carbon price whereas sectors that combine capital and emissions more complementary in production receive a lower carbon price. The intuition is that allocating more emissions to sectors with a complementary technology significantly increases the sectors' capital use. Abating these additional emissions in sectors with more substitutable technologies—to keep total emissions constant—decreases these sectors' capital use. Yet, the capital use in sectors with more complementary technologies increases more than the capital use in sectors with more substitutable technologies decreases, and the total capital use thus increases. Households accumulate more capital which benefits future generations.

Third, the results are also significant from a quantitative perspective. I calibrate the multi-sector Ramsey economy to data of the EU-28 and find substantial price differentiation and welfare gains from optimally differentiated carbon prices, relative to uniform. For example, for a 40% economy-wide emissions reduction in the European economy, optimal sectoral prices range from €42.3 to €116.9 per ton of CO2. By comparison, achieving the same emissions target would require a price of €91.3/ton. The magnitude of welfare gains largely depends on technology heterogeneity. For intermediate cases, based on or close to the empirical estimates of sector-specific capital-energy substitutability, I calculate that the welfare cost of climate policy to achieve a 20% emissions reduction reduces by 14–22% per period.

Essay 2

Implementing environmental policies to combat climate change and raising tax revenues to finance public consumption are duties of the government. When considering each problem separately, the public economics literature makes clear recommendations: (1) A uniform Pigouvian carbon tax should regulate the environmental externality from carbon emissions; (2) Goods ought to be taxed according to their price elasticities to induce minimal distortions when raising tax revenues. From (2) it follows that capital income should not be taxed because it is infinitely elastic in the longrun of a Ramsey economy. In reality, however, governments do tax capital incomes. This observation motivates Essay 2, in which I ask how carbon emissions should be priced in the presence of a capital income tax distortion. Previous work has shown that optimal carbon prices deviate from the Pigouvian tax rate (Sandmo 1975, Bovenberg and de Mooij 1994, 1997, Bovenberg and Goulder 1996, Bovenberg and van der Ploeg 1994, Parry et al. 1999, Schwartz and Repetto 2000, Kaplow 2012, Barrage 2020), or are non-uniform (Landis et al. 2018, Boeters 2014), each in the presence of distortionary taxes. Yet, no study has elaborated on the optimality of non-uniform carbon prices when a capital income tax distorts the economy. This essay is an attempt to fill this gap by showing theoretically and numerically the optimality of non-uniform carbon prices, and by providing an intuition for the direction of the sectoral price deviation as governed by sectoral technology heterogeneity.

First, I base the analysis on a dynamic, multi-sector, general equilibrium economy in which a capital income tax finances public consumption. I follow Barrage

(2018) and argue that the regulator cannot dissolve the capital income tax distortions because fiscal and environmental policy decisions are made independently. I thus think of a regulator who is restricted to only set carbon prices in each sector. I combine sectoral capital and sectoral emissions in one bundle which then produces, together with sectoral labor input, sectoral output. The intuition for the optimality of non-uniform carbon prices is simple: a capital income tax increases the capital market clearing price, and capital (as an input to production) becomes more expensive. Firms demand less capital and the household under-accumulates capital. If sectors produce with heterogeneous technologies, the regulator can stimulate the firms' capital use (and thus the household's accumulation of capital which is too low to due the capital income tax) through optimally differentiated sectoral carbon prices.

Second, the degree of carbon price differentiation is in particular driven by (1) the substitutability between capital and emissions and by (2) the substitutability between labor and the capital-emissions bundle. I find (1) that sectors that do not substitute well between CO_2 and capital receive a lower carbon price than sectors which substitute well between CO_2 and capital. Decreasing the carbon price in sectors that combine emissions and capital as ”complements” in production triggers the sectoral capital use which is greater than the capital decrease in sectors with higher carbon prices (that have a better substitutability). Overall, the total capital stock and welfare increases, relative to a uniform carbon price. I find (2) that optimal carbon prices also consider the sectoral substitutability between labor and the capital-emissions bundle. The general idea is to re-allocate the sectoral use of labor to sectors that combine labor and the capital-emissions bundle as ”complements” in production, i.e. that these ”more complement” sectors use more capital when given more sectoral labor inputs. Accordingly, sectors that substitute well between labor and the capital-emissions bundle receive a lower carbon price to incentivise these sectors to substitute away from labor towards the capital-emissions bundle.

The ”now free” labor is taken up by the ”more complement” sectors which triggers their sectoral capital use. Capital use in all sectors increase, firms produce with more resources and the household is better-off, relative to a uniform carbon price.

Third, carbon price differentiation is quantitatively significant. When I build the numerical analysis on realistic, empirically estimated values on the sector-specific substitution elasticities, I find that the steady-state optimal prices for carbon emissions differ on average by 11.07%. The optimal carbon prices increase the longrun capital stock by 0.26% and increase the lifetime wellbeing of households by 4.3%, relative to a uniform carbon price. I find that heterogeneity in the sectoral substitution elasticities between capital and emissions drives the degree of carbon price differentiation because these parameters govern the sectoral emission’s direct impact on capital accumulation. The other elasticities on the production side (on the final good level or between labor and the capital-emissions bundle) have less impact on the capital use and play thus less of a role. Also, sectoral carbon prices deviate stronger from uniform when sectors produce with more heterogeneous technologies, i.e. when the sectoral elasticities of input factor substitution cover a wider range.

Essay 3

Countries use the tax system for the redistribution of income to varying extents. For instance, the reduction in income inequality in Sweden after redistribution through taxes and transfers is twice as large as in the United States, and Sweden achieves a level of inequality in disposable income that is about half that of the United States (OECD 2015). This evidences that the targeting of income equality through the tax system may play a larger role in Sweden than in the United States—and Sweden has potentially more egalitarian social preferences than the libertarian United States. Against this background, this essay ask the following questions. What if social preferences are not utilitarian as assumed in most economic models but more egalitarian (as in Sweden) or more libertarian (as in the

United States)? How does the optimal tax system for redistributing income change? Should capital income be taxed? And if so, should the revenue from the capital income tax be used to make poor households better-off? This essay uses theoretical and numerical methods to answer how a (non-) utilitarian policymaker of a growing economy should redistribute income. The essay elucidates how different social views about justice/fairness impact the design of income taxes. Surprisingly, to the best of my knowledge, the public economics literature has so far overlooked to investigate the meaning of different principles of justice for the optimal taxation of (capital) income in a Ramsey-type economy.

The essay builds on an exogenously growing Ramsey economy with two types of households: savers are intertemporal more elastic than workers, savers demand thus lower returns to capital and own all capital. Savers also earn income from labor. Workers consume their entire labor income which denotes their only income source. The essay conveys three main insights.

First, I provide a novel motive for the taxation of capital income. I show that the zero-capital income tax result as put forward by Judd (1985) and Chamley (1986) crucially depends on the assumption of a utilitarian social welfare function (SWF). If, however, a benevolent planner values the future with the workers' lower intertemporal elasticity of substitution (IES)—and not with the savers' higher IES as implicitly induced by a utilitiarian SWF like in Chamely and Judd—capital income should indeed be taxed. The intuition is that workers value the future less than savers, so a capital income tax disincentivizes capital investments and implements a capital accumulation path as preferred by the workers. If capital income should be or should not be taxed is thus a normative matter about the society's intertemporal preferences. The public economics literature has identified other motives for positive capital income taxes rates. For instance, capital taxation is desirable when capital over-accumulates as in overlapping generation models (Diamond 1965, Conesa et al. 2009) or in the presence of uncertainties (Aiyagari 1994). Other examples include heterogeneous preferences for wealth (Saez and Stantcheva 2018), heterogeneous returns to capital (Gahvari and Micheletto 2016, Kristjansson 2016, Jacobs et al. 2020), heterogeneity in household time preferences (Saez 2002, Diamond and Spinnewijn 2011, Golosov et al. 2013), heterogeneity in household and social planner time preferences (Acemoglu et al. 2011, von Below 2012, Belfiori 2017, Barrage 2018), or when labor is not a flow variable but education enables human capital accumulation (Jacobs and Bovenberg 2010, Stantcheva 2017).

Second, I find that workers receive more transfers from savers under a more egalitarian social objective. Alternatively, laissez-faire policies represent an extreme case in which zero income redistribution is optimal. This is well in line with findings in the literature where deviations from a utilitarian SWF lead to different patterns of income redistribution: a more egalitarian objective increases the marginal welfare from consumption of low-income households which makes more income redistribution from high-income households to low-income households socially desirable (Boskin and Sheshinski 1978, Fair 1971, Atkinson 1973).

Third, most of the revenues from the capital income tax are returned back to the high-income households. The reason is that, empirically, high-income households supply labor less elastically than low-income workers, and face thus less costs from supplying labor when receiving a dollar of capital income tax revenue. In other words, returning the tax revenue back to the savers is less distortive than allocating it to the workers.

Policy implications

In addition to the scientific contributions, the work presented in this PhD thesis bears out several policy and implications:

- The findings of Essay 1 have important implications for assessing the cost of climate policy based on carbon pricing and for designing carbon prices in real-world policy making. First, model-based evaluations of climate policy which (unknowingly) assume that social and private discount rates are equal and ignore technology heterogeneity amount to assuming that a uniform carbon price is optimal. To the extent that these assumptions are not warranted, they overlook better carbon pricing policies and overstate the costs of climate policies. Second, simplifying intertemporal decisions making as in static, multi-sector models or when assuming myopic agents that do not form expectations about the future implicitly imposes that a uniform carbon price is optimal, and hence overlooks (potentially) better carbon pricing policies. Third, partitioned emissions regulation through separate ETSs within one jurisdiction, as is expected for EU climate policy, does not necessarily lead to higher costs. I show that two separate ETSs may be superior to a single, comprehensive ETS if sectors are assigned to each ETS such that sectors with relatively low and high capital-energy substitutability are clustered separately. These considerations are relevant to the practical design of ETSs, especially since decarbonization efforts in many countries relies on one or more markets for tradable emissions permits as the cornerstone of climate policy.

- Essay 2 shows that the presence of a capital income tax distortion has severe implications for the optimal design of price-based climate policies. Assuming that decisions on fiscal and environmental policies are made separately and that a positive tax on capital income is distortive, it is insufficient to create a single carbon market as, for instance, envisioned in the EU's longrun environmental policy design. Instead, optimality requires to price carbon emissions non-uniformly across sectors, i.e. firms should pay (or receive) a tax (or subsidy) for carbon emissions in addition to the market clearing price. More precisely, sectors that have good substitution possibilities between "dirty" emissions and ”clean” capital (perhaps electricity generating firms) should be priced above the carbon market clearing price. In contrast, sectors that have worse substitution possibilities (perhaps cement producing firms) should receive a subsidy on carbon emissions to lower their (sectoral) price for carbon emissions. This pricing scheme triggers capital demand, and households accumulate more capital. Essay 1 argues for the same directional, sectoral carbon price derivations from a uniform carbon price, given different private and social discount rates.

- Essay 3 argues that a utilitarian SWF is only one of many possible perspectives on social justice. Under different perspectives on the SWF—all of which are perfectly justifiable—other taxation schemes become optimal. First, I show that the optimal taxation of capital income is related to how a planner substitutes consumption over time. A utilitarian SWF implies that the planner values the future with the IES of the capital-owning household, which effectively amounts to assuming that a zero capital income tax in the steady-state is optimal. However, using the IES of the household without capital holdings is equally justifiable, and capital taxation becomes optimal. Second, a utilitarian SWF also determines the degree of income redistribution among households. Less (or more) income redistribution is also equally well justifiable because the planner may have more libertarian (or more egalitarian) social preferences. A laissez-faire policy may then even be optimal from a social perspective. A final important lesson for the policymaker is that capital income tax revenues should not be used to improve the living conditions of poor households. To enhance the living conditions of poor households, the government should tax high labor incomes to finance a subsidy for low labor incomes (or a lower tax on low labor incomes)—in line with the idea to tax labor incomes progressive

Chapter 2

Pricing carbon in a multi-sector economy with social discounting*

Abstract

Economists tend to view a uniform emissions price as the most cost-effective approach to reducing greenhouse gas emissions. This paper offers a different view, focusing on economies where society values the well-being of future generations more than private actors. Employing analytical and numerical general equilibrium models, we show that a uniform carbon price is efficient only under restrictive assumptions about technology homogeneity and intertemporal decision-making. Non-uniform pricing spurs capital accumulation and benefits future generations. Depending on sectoral heterogeneity in the substitutability between capital and energy inputs, we find that optimal carbon prices differ widely across sectors and yield substantial welfare gains relative to uniform pricing.

*This chapter represents joint work together with Prof. Sebastian Rausch (ZEW – Leibniz Centre for European Economic Research and Heidelberg University).

2.1 Introduction

Climate change is a long-term problem—the costs of avoiding greenhouse gas emissions must be justified by the benefits of avoided impacts well into the future. At the same time, efficient climate change mitigation requires avoiding emissions where it is cheapest. In most economies, however, the marginal costs of reducing emissions vary widely across sectors and technologies. Economists have long argued that a uniform price on carbon dioxiode (CO_2) emissions minimizes the welfare costs of achieving an economy-wide emissions target as it incentivizes abatement up to the point where marginal abatement costs are equalized (Cropper and Oates 1992, Goulder and Parry 2008, Metcalf 2009).

This has strong implications for the contribution that each sector should make in a country's decarbonization effort. For example, if the marginal cost of reducing emissions in a heavily coal-based electricity sector is lower than the cost of reducing emissions in private transport or buildings, where the opportunities for substituting fossil energy for capital are more expensive and limited, the electricity sector should reduce emissions more than these sectors. This, of course, is precisely the idea behind the standard climate policy recommendation to "put a price on carbon," which manifests itself in numerous proposals for efficient climate policy—for example, the idea of a global, comprehensive carbon price as once envisioned under the Kyoto Protocol, expanding the scope of major emissions trading schemes in Europe, the United States, and China, or linking regional carbon pricing regimes (Nordhaus 2015).

As economies around the world pursue increasingly ambitious decarbonization goals for the whole economy, the question of how to price CO_2 emissions in different sectors is of great importance. It is all the more surprising that the economic discipline has so far overlooked the idea that optimal carbon pricing in a multi-sector economy is intricately linked to how private and public decision-makers value the costs and benefits of climate policy over time. Intuitively, if society or a benevolent social planner places a higher value on the well-being of future generations than private agents do, agents discount the future too much and the government should promote future oriented policies, i.e. policies which increase the capital stock. The transformation to a green economy requires a massive substitution of capital for fossil energy, but the substitutability of capital and fossil energy varies widely across sectors, depending on technology. Thus, while it appears desirable from a partial equilibrium perspective to price carbon uniformly, efficient multi-sector carbon pricing that enables "large" transformations requires considering the heterogeneous general equilibrium effects on capital accumulation and their social valuation.

In this paper, we examine optimal carbon pricing in a multi-sector economy when private agents and the social planner discount utility differently. We do not propose a new model but base our analysis on a standard neoclassical growth model where capital substitutes for CO_2 emissions that are a by-product of fossil fuel combustion in sectoral production. We focus on cost-effectiveness and ask how carbon should be priced to meet an exogenously specified economy-wide emissions budget at the lowest cost.[1] We question the generality of the established wisdom that efficient carbon pricing entails a single, uniform price across different sectors. We show that this view rests on restrictive and unrealistic assumptions once differential social discounting, i.e. social and private (market) discount rates differ, is taken into account: either different economic sectors are "identical" in terms of substitutability of capital and energy inputs, or the capital stock is exogenous. We show that technology heterogeneity causes optimal CO_2 prices to differ across sectors, and quantitatively explore the welfare gains relative to uniform carbon pricing for the European economy. We explore the implications of optimally differentiated CO_2 prices for the design of emissions trading systems as a major market-based instrument for climate policy.

Before summarizing our results in greater detail, it is useful to set the stage

[1] We abstract from the benefits of averted damages from climate change and endogenous environmental quality which are subject of study in integrated assessment models (see, for example, Nordhaus 2000, Tol 2009).

for social discounting. Evaluation over time is a recurring and highly controversial topic in economics, especially when the costs and benefits are evaluated over very long time scales, as is the case with climate change. For many philosophers (Sidgwick 1874, Broome 1994) and economists (Ramsey 1928, Pigou 1932, Solow 1974) intergenerational discounting is ethically indefensible: basic fairness, i.e. impartiality, non-discrimination, and equal treatment, rule out discounting.[2] Zero or near-zero discount rates, however, do not align with households' inter-temporal preferences as revealed through their savings behavior (Nordhaus 2007, Dasgupta 2008). Social discounting reconciles both perspectives, allowing the planner to place a higher welfare weight on future generations above the current generation's private altruism towards the future (Bernheim 1989, Fahri and Werning 2007, Kaplow et al. 2010, Goulder and Williams 2012). Accepting the view that society places more importance on future generations than private actors is, of course, controversial and debatable.[3] Our goal is not to enter this debate, but to analyze the extent to which the established wisdom that uniform carbon pricing is efficient can be upheld when differential social discounting is indeed accepted as a reasonable premise.

We start from the result previously shown by Barrage (2018) that under differential social discounting, a planner in a first-best setting uses a capital income subsidy to equalize social and private marginal returns to savings. We show that if capital income subsidies can be optimally chosen, sectoral CO_2 prices are uniform. In reality, however, capital income as a whole is taxed and not subsidized. Moreover, climate policy is usually pursued separately from fiscal policy; in particular, it

[2] Robert Solows argumentation against intergenerational discounting of utility is in line with Pigou and Ramsey, stating that "[in] social decision making, however, there is no excuse for treating generations unequally, and the time-horizon is, or should be, very long" (Solow 1974, p. 9).

[3] When individuals discount both the past and the future, Caplin and Leahy (2004) show that policy makers should be more patient than private citizens, whose choices define the most short-sighted Pareto optimum. From a non-academic perspective, there is also evidence of such a view. Recently, the Federal Constitutional Court (2021) has ruled that the "German Climate Law" in its current form violates the civil liberties of future generations. In particular, it is not acceptable to postpone necessary CO_2 emissions reductions largely into the future in order to spare the present. This can be interpreted as support for the view that society should not only explicitly consider the well-being of future generations in today's policy decisions, but give it more weight.

is not directly linked to policy decisions on capital taxation. As Barrage (2018), we thus argue that the relevant perspective for carbon pricing policy in a second-best world is to assume that capital income as a whole cannot be subsidized.

Given the two premises of differential social discounting and no capital income subsidies, we show that uniform carbon pricing is only efficient, if social and private "market" discount rates coincide. When discount rates differ, a uniform carbon price is only efficient if climate policy either ignores the effects of a carbon price on investment and capital accumulation or assumes that sectoral production technologies are identical. Both are clearly unreasonable.

Technological heterogeneity across sectors is key to our main result: efficient multi-sector carbon pricing differentiates CO_2 prices across sectors. The economic intuition we provide for this result is based on the heterogeneous sectoral responses to a carbon price in adjusting capital and fossil energy (CO_2 emissions) inputs. Sectors in which CO_2 emissions are not easily substitutable with "clean" capital should receive a lower carbon price than sectors where these two inputs are better substitutes. When capital in a given sector is a "poor" substitute for emissions, pricing carbon destroys more capital as compared to reducing the same amount of CO_2 emissions in a sector with a higher substitutability. If the social valuation of capital and future wealth were based on the "market" discount rate, the sectoral allocations of CO_2 emissions and capital stemming from the equilibrium decisions of firms and households would also be efficient from a social perspective. With differential social discounting, however, economic agents discount the future too much and differentiating sectoral CO_2 prices avoid households failing to invest sufficiently in the economy's capital stocks: it directs capital to where it is socially most valuable, increasing capital accumulation and benefiting future generations.

Multi-sector carbon pricing also constitutes a first-order deviation from uniform carbon pricing from a quantitative perspective. To explore the relevance of our theoretical results in an empirical context, we use a steady-state model which is calibrated to the European economy and resolves major economic sectors. Incorporating empirical estimates on sector-specific capital-energy substitutabilities, we find that the CO_2 price differentiation across sectors is significant. For example, for a 40% economy-wide emissions reduction in the European economy, optimal sectoral prices range from €42.3 to €116.9 per ton of CO_2, with an emissions-weighted mean CO_2 price of €74.4/ton. By comparison, achieving the same emissions target would require a price of €91.3/ton. The variation in optimal CO_2 prices is not much affected by considering lower or higher economy-wide emissions reduction targets: the coefficient of variation of optimal sectoral CO_2 prices is roughly constant at 40% for targets of up to 80%.

Compared to uniform carbon pricing, the failure to differentiate CO_2 prices by sector forgoes potentially large welfare gains. The magnitude of welfare gains largely depends on technology heterogeneity. For intermediate cases, based on or close to the empirical estimates of sector-specific capital-energy substitutability, we calculate that the welfare cost of climate policy to achieve a 20% emissions reduction reduces by 14–22% per period. If the economy is composed of sectors in which some sectors have strong capital-energy substitutability while others have strong capital-energy complementarity, the reduction in welfare costs can be as high as 50% per period. In contrast, if capital and energy are highly substitutable in all sectors, the welfare gains from differentiating CO_2 prices are negligible. In general, welfare gains decrease as emissions reductions increase.

Our findings have important implications for the design of emissions trading systems (ETSs): a single, comprehensive ETS is not optimal when private and social discount rates differ and sectoral production technologies differ. This runs counter to the established policy recommendation to broaden the scope of an ETS or integrate the carbon markets of separate ETSs (Böhringer et al. 2006, Abrell and Rausch 2017). We show that when CO_2 emissions are regulated by two ETSs within a jurisdiction, price differences between the systems need not be costly. In fact, two separate ETSs may be superior to a single, comprehensive ETS if sectors are assigned to each ETS in such a way that sectors with relatively low and high capital-energy substitutability are clustered separately. These considerations are relevant to the practical design of ETSs, especially since decarbonization efforts in many countries relies on one or more markets for tradable emission permits as the cornerstone of climate policy.

The results of our analysis are also relevant for model-based evaluations of climate policy. First, model-based evaluations of climate policy which macroeconomic models which adopt a single-output framework (for example, DICE-type models by Nordhaus 2007, Barrage 2018), assume that social and private discount rates are equal, and ignore technology heterogeneity amount to assuming that a uniform carbon price is optimal. To the extent that these assumptions are not warranted, they overlook better approaches to carbon pricing and thus overstate the costs of climate policy. Second, models which adopt a richer multi-sector perspective, on the other hand, are often static (Böhringer et al. 2016, Landis et al. 2018) or assume myopic economic agents ruling out that expectations affect the future capital stock—for example, models used for climate policy assessments by the European Commission (Capros et al. 2013) and in the academic community (Paltsev et al. 2005, Fawcett et al. 2014, Thompson et al. 2014). Simplifying the intertemporal decision rules in such ways also effectively assumes that a uniform carbon price is optimal. Third, the partial equilibrium result of a uniform cost-effective pricing of emissions can only be transferred to a general equilibrium under restrictive assumptions. Partial equilibrium evaluations should thus be used with caution.

More generally, our analysis contributes to the large body of literature on discounting in the context of climate change mitigation (Stern 2006, Nordhaus 2007, Weitzman 2007). Previous work has pointed out that a clearer distinction should be made between the concepts of social and private "market" discount rates when evaluating climate policies (Kaplow et al. 2010, Goulder and Williams 2012, van der Ploeg and Rezai 2019). von Below (2012) is the first to operationalize both concepts in one framework for climate policy analysis. Most closely related to our analysis, Barrage (2018) shows that if the government cannot subsidize capital income, the constrained-optimal carbon tax may be up to 50% below the present value of marginal damages (the social cost of carbon) due to the general equilibrium effects of climate policy on household savings. We contribute by providing the first analysis which examines optimal multi-sector carbon pricing when social and private discount rates differ. Non-uniform pollution prices have been found to be optimal in economies that feature either imperfectly competitive markets (Sandmo 1975, Markusen 1975, Krutilla 1991, Rauscher 1994), settings where social equity concerns over heterogeneous households are present (Landis et al. 2018, Abrell et al. 2018), or when border adjustment tariffs on carbon leakage are not possible (Hoel 1996). The literature has also identified non-uniform carbon prices to be optimal in settings with heterogeneous countries. When it is not possible to equalize marginal utilities of consumptions across countries (e.g., when lump-sum transfers are not possible), poor countries should receive a lower carbon price given that they have a higher marginal utility from consumption, and rich countries should receive a higher carbon price (Chichilnisky and Heal 1994, Sheeran 2006).[4] This paper adds an important new motive for non-uniform pollution prices.

Section 2.2 describes the economic environment and sets up the social planning problem. Section 2.3 presents our qualitative results. Section 2.4 describes the calibration of the numerical model. In Section 2.5 we explain the computational experiments, with results presented in Section 2.6. Section 2.7 provides additional sensitivity analysis. Section 3.8 concludes. Proofs are contained in the Appendix.

[4]Uniform carbon prices are, however, optimal if welfare weights are integrated in the social welfare function such that equalizing the "weighted" marginal utilities of consumption across countries is possible. This is the case when the welfare weights correspond to the inverse of the countries' marginal utilities of consumption—as, for example, guaranteed by time-varying Negishi weights (Nordhaus and Yang 2006, Nordhaus 2011).

2.2 The economic environment

2.2.1 Private and social discounting

Following recent literature (see, for example, Barrage 2018, Belfiori 2017, Fahri and Werning 2007), we assume that the social planner places positive weight on future generations' welfare above and beyond the current living population's private altruism. Social discounting thus means that there is a disagreement in utility discount factors between a dynastic household (ζ) and the social planner (ζ_S) and that $\zeta > \zeta_S \geq 0$. Time is discrete and extends to infinity $t = 0, \ldots, \infty$. Dynastic households maximize:

$$U = \sum_{t=0}^{\infty} \frac{1}{(1+\zeta)^t} u(C_t) \tag{2.1}$$

where $u(C_t)$ is a standard concave period utility function which allows for general CRRA preferences: $u(C_t) = C_t^{1-\sigma}/(1-\sigma)$. $1/\sigma$ denotes the elasticity of inter temporal substitution, C_t is consumption of a final good at time t, and ζ is the private discount rate of households. The social planner maximizes the social welfare function:[5]

$$W = \sum_{t=0}^{\infty} \frac{1}{(1+\zeta_S)^t} u(C_t)\,. \tag{2.2}$$

2.2.2 Decentralized economy

HOUSEHOLDS.—Household earn income from supplying labor at the market rate w_t, return R_t (net of depreciation) minus capital income tax Ξ_t on capital holdings K_t, profits from sectoral production Π_t, and lump-sum transfers Λ_t from the government which comprise revenues from taxing carbon emissions and capital income. The households' budget constraint is given by:

$$C_t + \bar{K}_{t+1} \leq w_t \bar{L} + [1 + R_t(1-\Xi_t)]\bar{K}_t + \Pi_t + \Lambda_t\,, \tag{2.3}$$

[5] Bernheim (1989) and Barrage (2018) show that the social welfare function corresponds to a time-consistent social planner's problem for appropriately chosen welfare weights.

where $\bar{L}$ denotes the perfectly inelastic aggregated labor supply and $\bar{K}_t$ denotes the aggregated capital supply at time t.

Households maximize (3.2) subject to (3.1), yielding the following savings optimality condition, which relates the costs of forgone consumption today to the discounted value of future capital income:

$$U_{Ct} = \frac{1}{1+\zeta} U_{Ct+1} \left(1 + R_{t+1}(1 - \Xi_{t+1})\right) , \tag{2.4}$$

where $U_{Ct} := \partial u(C_t)/\partial C_t$.

PRODUCTION.—The final good at time t, $\hat{Y}_t$, is produced by profit-maximizing firms using contemporaneous output of sector $j, k = 1, \ldots, J$, Y_{jt}, with a linearly homogeneous production function:

$$\hat{Y}_t = \prod_{j=1}^{J} Y_{jt}^{\gamma_j} \tag{2.5}$$

where γ_j, with $\sum_j \gamma_j = 1$, is a share parameter.

Sectoral goods Y_{jt} are produced using labor, capital, and CO_2 emissions inputs according to a constant-elasticity-of-substitution (CES) production function:

$$Y_{jt} = L_{jt}^{\alpha_j} \left[\beta_{Kj}(H_{Kj}K_{jt})^{\rho_j} + \beta_{Ej}(H_{Ej}E_{jt})^{\rho_j}\right]^{\frac{1-\alpha_j}{\rho_j}} \tag{2.6}$$

where α_j and βs are share parameters, and $\beta_{Kj} + \beta_{Ej} = 1$. H_{Kj} and H_{Ej} denote the input- and sector-specific productivity factors of capital and emissions. $-\infty < \rho_j < 1$ denotes the elasticity parameter and is related to the elasticity of substitution $\sigma_j = 1/(1 - \rho_j)$. Throughout, we say that inputs in production are substitutes when $\rho_j > 0$ and complements when $\rho_j < 0$. $\rho_j = 0$ indicates the Cobb-Douglas case. L_{jt}, K_{jt}, and E_{jt} denote the amount of labor, capital, and CO_2 emissions used in sector j at time t, respectively.

Firms maximize profits taking prices as given on perfectly competitive output and factor markets. The optimal input choices K_{jt}, L_{jt}, E_{jt}, and Y_{jt} are therefore determined by:

$$r_{jt} = p_{jt}\frac{\partial Y_{jt}}{\partial K_{jt}}, \quad w_t = p_{jt}\frac{\partial Y_{jt}}{\partial L_{jt}}, \quad \tau_{jt} = p_{jt}\frac{\partial Y_{jt}}{\partial E_{jt}}, \quad p_{jt} = \hat{p}_t\frac{\partial \hat{Y}_t}{\partial Y_{jt}}\,, \tag{2.7}$$

where p_{jt} and $\hat{p}_t$ denote the prices for sectoral output Y_{jt} and final output $\hat{Y}_t$, respectively. We choose the price of final output as the numeraire ($\hat{p}_t = 1$).

We focus specifically on multi-sector economies with technology heterogeneity. Sectors are said to be *heterogeneous* if share parameters β_{Kj}, substitution parameters $\rho_j \neq 0$, input factor-specific productivities (H_{Kj}, H_{Ej}), or sector-specific depreciation rates δ_j or a combination of these parameters, differ across sectors. Two sectors are *identical* if these parameters take on the same respective value or if $\rho_j = 0$ across the two sectors.

Market clearing conditions of labor and capital at time t are:

$$\sum_{j=1}^{J} L_{jt} = \bar{L}, \quad \forall j : K_{jt} = \bar{K}_{jt}\,, \tag{2.8}$$

such that $\bar{L}$ is perfectly mobile across sectors, and $\bar{K}_{jt}$ denotes the capital supplied to sector j at time t. Aggregated capital, $\bar{K}_t = \sum_j \bar{K}_{jt}$, accumulates over time according to:

$$\sum_{j=1}^{J} \bar{K}_{jt+1} = \sum_{j=1}^{J}(1-\delta_j)\bar{K}_{jt} + I_t\,, \tag{2.9}$$

where δ_j is the sector-specific, periodic depreciation rate and I_t denotes investments. Capital is homogeneous and mobile across sectors such that the return net of sector-specific depreciation is equal across the economy, i.e. households base their savings decision on $R_t = r_{jt} - \delta_j$, $\forall j$. We include sector-specific capital rental rates to incorporate an additional aspect of sectoral heterogeneity.

The final good can be used for investment and consumption purposes:

$$C_t + I_t = \hat{Y}_t\,. \tag{2.10}$$

CARBON EMISSIONS.—CO_2 emissions are the by-product created by the combustion of fossil fuels inputs in production. To simplify, we abstract from explicitly including energy (from fossil fuels) and instead directly represent CO_2 emissions as an input in production.[6]

At the center of our analysis is the investigation of sectorally differentiated carbon prices. We thus include J separate emission markets at each point in time which determine sectoral emissions prices, τ_{jt}:

$$E_{jt} = \bar{E}_{jt} \,. \tag{2.11}$$

$\bar{E}_{jt}$ denotes the sector-j emissions budget, where in addition the sum of sectoral emissions is constrained by an exogenously given economy-wide emissions budget, $\bar{E}_t$, according to:

$$\sum_{j=1}^{J} \bar{E}_{jt} = \bar{E}_t \,. \tag{2.12}$$

By setting $\bar{E}$, we represent economy-wide emission reduction targets.

EQUILIBRIUM.—The definition of the equilibrium for the decentralized economy is standard and provided in Appendix 2.9.1.

2.2.3 Planner's problem: first-best policy

The planner's problem is to maximize social welfare (2.2) subject to final good production (3.4), sectoral production (3.5), market clearing conditions for labor and capital (3.7), aggregate capital accumulation (3.8), the resource constraint (3.10), and attaining the economy-wide emissions constraint $\sum_j E_{jt} = \bar{E}_t$. The planner's first-order conditions (FOCs) with respect to the use of sectoral emissions (E_{jt}) require that the social marginal costs and benefits of using carbon emissions be

[6] In terms of our model, we thus use "carbon emissions" and "fossil energy" interchangeably.

equated for every j (see Appendix 2.9.2):

$$\underbrace{\mu_{jt}}_{\substack{\text{marginal value}\\ \text{of sectoral output}}} \underbrace{\frac{\partial Y_{jt}}{\partial E_{jt}}}_{\substack{\text{marginal product of}\\ \text{sectoral emissions}}} = \underbrace{\lambda_t^E}_{\substack{\text{marginal social cost of}\\ \text{economy-wide emissions}}} . \tag{2.13}$$

It is thus straightforward to see that in the social optimum sectoral carbon prices are uniform.

To see that in a decentralized market economy uniform sectoral carbon prices can achieve the socially optimal outcome, combine the planner's FOCs with respect to consumption C_t, capital K_{t+1}, and investment I_t to obtain the planner's Euler equation:

$$U_{Ct} = \frac{1}{1+\zeta_S} U_{Ct+1} \left(1 + R_{t+1}\right) . \tag{2.14}$$

Comparing the planner's optimality conditions with those governing the behavior of households (3.3) and firms (3.6) in decentralized equilibrium, it is straightforward to show that the first-best allocation can be decentralized by a combination of uniform sectoral carbon prices according to:

$$\underbrace{p_{jt}}_{\substack{\text{price of}\\ \text{sectoral output}}} \underbrace{\frac{\partial Y_{jt}}{\partial E_{jt}}}_{\substack{\text{marginal product of}\\ \text{sectoral emissions}}} = \tau_{jt} = \underbrace{\tau_t}_{\substack{\text{marginal social cost of}\\ \text{economy-wide emissions}}} \quad \forall\, j , \tag{2.15}$$

where $\tau_t = \lambda_t^E / U_{ct}$, and a capital income subsidy to equate the social and private marginal returns on savings:

$$\Xi_{t+1} = \frac{\zeta_s - \zeta}{1+\zeta_S} \frac{1 + R_{t+1}}{R_{t+1}} . \tag{2.16}$$

The intuition behind this result is that private households are too impatient from the planner's perspective, and consequently subsidies on capital are desirable to increase returns to savings to avoid households failing to invest sufficiently in the economy's capital stocks. This result has been shown previously by von Below

(2012) and Barrage (2018). Importantly, in a first-best setting in which capital income subsidies can be optimally chosen, optimal sectoral carbon prices are uniform.

2.3 Qualitative results

2.3.1 Constrained-optimal carbon prices

In reality, however, capital income as a whole is taxed and not subsidized. Moreover, climate policy is usually pursued separately from fiscal policy; in particular, it is not directly linked to policy decisions on capital taxation. We thus argue that the relevant perspective for carbon pricing policy is to assume that capital income as a whole cannot be subsidized. For the remainder of the analysis, we therefore assume that capital income subsidies are not available and focus on a climate policy which is concerned with choosing constrained-optimal carbon prices.

To analyze carbon pricing in a setting which rules out capital income subsidies, we can re-state the no-subsidy constraints $\Xi_t \geq 0$ in terms of the households' intertemporal optimality conditions (3.3) as:

$$\frac{U_{Ct}(1+\zeta)}{U_{Ct+1}} \leq (1 + R_{t+1})) \,, \quad R_{t+1} = MPK_{jt+1} - \delta_j, \quad \forall j, \text{ and } t > 0 \,. \tag{2.17}$$

MPK_{jt} denotes the marginal product of capital in sector j which corresponds to the price for sectoral capital (r_{jt}) in the decentralized equilibrium. The constrained policy problem is given by adding (2.17) to the social planner's problem (see Appendix 2.9.3). Comparing the constrained planner's optimality conditions with those governing the behavior of households and firms in the decentralized equilibrium, it is straightforward to show that:

Proposition 1 *In a second-best setting when capital income subsidies are not feasible ($\Xi_t = 0$), the constrained-optimal allocation can be decentralized by sector-specific carbon taxes which are implicitly defined by equating the marginal benefits* *of emissions use with the marginal social cost of emissions which comprise a Pigouvian and a social discounting externality-correcting term:*

$$\underbrace{p_{jt}}_{\substack{\text{price of}\\ \text{sectoral output}}} \quad \underbrace{\frac{\partial Y_{jt}}{\partial E_{jt}}}_{\substack{\text{marginal product of}\\ \text{sectoral emissions}}} = \underbrace{\tau_t}_{\substack{\text{social cost of}\\ \text{carbon emissions}}} - \underbrace{(1+\zeta_S)\tilde{\phi}^K_{jt}\frac{\partial MPK_{jt}}{\partial E_{jt}}}_{\substack{\text{marginal sectoral emissions' impact}\\ \text{on sectoral marginal product of capital}}} \tag{2.18}$$

where $\tilde{\phi}^K_{jt} = \phi^K_{jt}/U_{Ct}$, $\phi^K_{jt} > 0$, *denote the social costs of constrained capital prices that are governed by the private Euler equation.*

PROOF: See Appendix 2.9.3. □

Proposition 1 implies that in an economy which endorses differential social discounting and cannot subsidize capital income may want to impose sectoral CO_2 prices which deviate from the Pigouvian carbon tax, where the latter is given by the marginal social cost of economy-wide carbon emissions ($\tau_t = \lambda^E_t/U_{ct}$). Intuitively, this is because the planner cares about the overall level of assets given to future generations and consequently takes into account the general equilibrium effects of sectoral carbon taxes on households' savings behavior. Without a capital income subsidy, private savings are not optimal. Consequently, climate policy's impacts on private investment can now have first-order welfare effects. Mathematically, if $\zeta > \zeta_s$, the no-subsidy constraint is binding ($\phi^K_{jt} > 0$), the Pigouvian carbon tax is corrected by a term which reflects how reductions in sectoral emissions affect the sectoral marginal product of capital ($\partial MPK_{jt}/\partial E_{jt}$). Intuitively, this is because the sectoral CO_2 prices are used to bring savings closer to the socially optimal path in the presence of the social discounting externality.[7] $\partial MPK_{jt}/\partial E_{jt}$ reflects the substitutability (complementarity) between capital and emissions inputs in a given sector. For heterogeneous sectoral technologies, the $\partial MPK_{jt}/\partial E_{jt}$ terms vary across sectors, implying that contrained-optimal CO_2 prices differ across sectors.

[7] Without exploring the implications for differentiated CO_2 prices, Barrage (2018) and Belfiori (2017) show in a one-sector integrated assessment model that constrained-optimal carbon price reflects the discounted sum of marginal climate damages, corresponding to λ^E_t in our framework, and the general equilibrium effects on the households' savings decisions.

2.3.2 When is a uniform CO_2 price optimal?

Before we further investigate the role of technology heterogeneity for multi-sector carbon pricing, it is useful to illustrate the conditions that lead to a uniform carbon price. Given the importance of the established wisdom in environmental economics that a uniform carbon price across different emission sources is efficient, this helps to reveal the implicit assumptions behind this result.

The following result follows straightforwardly from Proposition 1:

Proposition 2 *Constrained-optimal sectoral carbon prices are uniform if the economy displays at least one of the following characteristics:*

(i) sectoral production technologies are identical,

(ii) the capital stock is exogenously given and fixed, or

(iii) there is no differential social discounting, i.e. social and private market discount rates coincide ($\zeta_S = \zeta$).

Proposition 2 highlights the three fundamental premises which underlie the standard perspective often inherent in economic assessments (and models) to study price-based climate change policies. It has important implications for the model-based evaluations of climate policy in the literature.

Integrated assessment models (Nordhaus 2007, Barrage 2018) primarily adopt a single-output framework.[8] This, however, overlooks relevant heterogeneity at the sectoral level in terms of differences in carbon emission intensity and substitutability between capital and energy inputs in production. As emissions (reductions) and capital are bound together, the optimal use of carbon pricing needs to take into account how a carbon price affects the marginal product of capital in different sectors (see $\partial MPK_{jt}/\partial E_{jt}$ in equation (2.18)). When differential social discounting plays a role, the single-sector perspective is only inconsequential for the optimal

[8] The general model of Golosov et al. (2014) considers multiple energy-producing technologies, but no further sectoral detail below the final goods sector.

carbon pricing rule when technology heterogeneity is "negligible", i.e. when sectors are largely identical. In fact, by model construction, such models assume that the optimal CO_2 is uniform. To the extent better carbon pricing policies are overlooked, they overstate the costs of climate policy.

Economic models which adopt a richer multi-sector perspective are often static (Böhringer et al. 2016, Landis et al. 2018) or assume myopic economic agents ruling out that expectations affect the future capital stock—for example, models used for climate policy assessments by the European Commission (Capros et al. 2013) and in the academic community (Paltsev et al. 2005, Fawcett et al. 2014, Thompson et al. 2014). This amounts to holding fixed the capital stock or simplifying intertemporal decision rules in a way which also boils down to assuming that a uniform carbon price is optimal. Then, the rationale to differentiate carbon prices according to Proposition 1 vanishes and the optimal carbon price for each sector is equal to the (uniform) Pigouvian level. Proposition 2 *(ii)* also implies that the partial equilibrium intuition of uniform cost-effective emissions pricing, which by construction ignore effects on capital income, does not carry over to a general equilibrium setting.

Finally, if private agents discount with the same rate as the social planner, there is no need to incentivize households' savings beyond the market remuneration. Hence, $\phi_{jt}^K = 0$ in (2.18) and carbon emissions in all sectors are priced uniformly according to the Pigouvian principle ($\tau_{jt} = \tau_t = \lambda_t^E / U_{ct}$, $\forall j$), regardless of whether sectors are heterogeneous or not.

In summary, accepting the differential social discounting perspective and the (empirical) fact that capital income is not subsidized, the assumptions leading to the conventional climate policy recommendation of a uniform CO_2 price are quite stark: either all sectoral outputs in an economy must be produced with identical technologies or intertemporal economic choices must be ignored.

2.3.3 Optimal CO_2 prices: the role of technology heterogeneity

This section investigates multi-sector carbon pricing when sectoral production technologies are heterogeneous.[9]

Based on Proposition 1 and steady-state FOCs, the rule for efficient pricing of carbon emissions in sector j can then be expressed in terms of technology parameters (see Appendix 2.9.4 for derivation):

$$\tau_j = \underbrace{\tau}_{\substack{\text{social cost of}\\ \text{carbon emissions}}} \times \underbrace{\frac{1}{1 - \bar{\phi}_j \rho_j \theta_j^K}}_{\substack{\Gamma_j:=\text{sector-specific correction factor}\\ \text{to account for social discounting externality}}} \tag{2.19}$$

where $\bar{\phi}_j = (1 + \zeta_S)\tilde{\phi}_j^K {K_j}^{-1}$ incorporates the additional cost imposed by deviating from the socially optimal path of capital accumulation as a result of the no-capital income subsidy constraint, or, expressed equivalently, due to requiring that capital prices are governed by the private, but not the social, Euler equation. θ_j^K denotes the value share of capital in the capital-emissions bundle in the production of sectoral output j according to (3.5).

Proposition 3 *If sectoral production technologies are heterogeneous, the constrained optimal carbon prices differ across sectors: $\tau_j \neq \tau_k$, $\forall j, k$. In particular, $\tau_j > \tau_k$ if ceteris paribus:*

(i) capital is a better substitute for emissions in sector j relative to sector k ($\rho_j > \rho_k$),

(ii) the capital share is higher in sector j relative to sector k ($\beta_{Kj} > \beta_{Kk}$) if both sectors are substitutes ($\rho_j = \rho_k > 0$) and vice versa if both sectors are complements ($\rho_j = \rho_k < 0$),

[9] We assume throughout that the social planner equally weighs current and future (unborn) generations, implying that the social discount rate is zero. $\zeta_S = 0$ also enables us to focus on steady-state equilibrium and to derive closed-form solutions.

(iii) capital is more productive ($H_{Kj} > H_{Kk}$) or emissions are less productive ($H_{Ej} < H_{Ek}$) in sector j relative to sector k, if both sectors are complements ($\rho_j = \rho_k < 0$), and vice versa if both sectors are substitutes ($\rho_j = \rho_k > 0$), or

(iv) the capital depreciation rate is lower in sector j relative to sector k ($\delta_j < \delta_k$) if both sectors are substitutes ($\rho_j = \rho_k > 0$) and vice versa if both sectors are complements ($\rho_j = \rho_k < 0$).

PROOF: See Appendix 2.9.4. □

The important insight from Proposition 3 is that—in an environment where the planner adopts a differential social discounting perspective and capital income subsidies are not feasible—technology heterogeneity provides an economic rationale for differentiating carbon prices across sectors. Sectoral carbon pricing is a means to influence household savings and move the economy closer to the social optimum. Intuitively, the efficient pattern of sectoral carbon taxes depends on how the amount of capital employed in each sector—and thus the marginal productivity of capital—reacts to the pricing of emissions. This, in turn, hinges on the production technology in each sector, i.e. the way in which profit-maximizing inputs of capital and emissions are combined to produce final goods for consumption.

If sectoral production technologies are identical, the quantity of capital employed in each sector is equally affected by a carbon price. Hence, a uniform carbon price is sufficient to implement the constrained social optimum—simply restating Proposition 2. Also, technology heterogeneity does not matter if there is no differential social discounting as it is then efficient to price emissions in each sector with the uniform social cost of carbon τ.[10]

One important aspect of technology heterogeneity relates to the substitutability or complementarity between capital and emissions in sectoral production. Proposition 3 (i) shows that, everything else being equal, it is optimal to price carbon emissions at a higher rate in sectors which can more easily substitute between cap-

[10] This can also be inferred from the sectoral carbon pricing rule in (2.19). Without differential social discounting, $\bar{\phi}_j = 0$, $\forall j$, implies $\Gamma_j = 1$, $\forall j$.

ital and emissions. Consider the extreme case in which capital and emissions are perfect complements in one sector and highly substitutable in another sector. A uniform carbon price can then not be optimal and lowering the carbon tax in the "complement" sector increases capital demand in this sector. To compensate for higher emissions (given the fixed economy-wide emissions target), the carbon price in the "substitute" sector has to increase. To the extent that capital and emissions are substitutes, this triggers a substitution away from emissions towards capital in the "substitute" sector. As output in the "substitute" sector, which now faces a higher carbon tax, falls, the amount of capital used in the "substitute" sector may fall. When capital in this sector is a strong enough substitute for emissions, the reduction in capital use in the "substitute" sector is smaller than the increase in capital use in the "complement" sector. Overall, the differentiation of sectoral carbon taxes leads to an increased use of capital in the economy, thus pushing the economy on a path with higher capital accumulation relative to the case with uniform carbon pricing. Put differently: with uniform carbon pricing, households invest too little and the (steady-state) consumption level is sub-optimally low. This is because savings decisions only take into account private discounting. It is thus optimal for climate policy to tax carbon in a way which boost households' savings by implicitly subsidizing capital, or, steering the relative price of capital to CO_2 emissions. The efficient pattern of implicit capital subsidies is inversely related to the elasticity of capital use in each sector: subsidize capital less (tax carbon more) in sectors where capital and emissions are "good" substitutes and subsidize capital more (tax carbon less) in sectors where capital and emissions are "bad" complements.

A similar logic applies if sectoral heterogeneity encompasses other aspects of technology which imply that the adjustment of capital (and emissions) inputs with respect to climate policy differs across sectors. Even if capital and emissions inputs are identical in terms of their degree of substitutability (same ρ's across sectors), a sector with a high capital share (β_K) should receive a higher sectoral carbon tax as compared to a sector with a low share of capital if both sectors are substitutes (see Proposition 3 (*ii*)). The reason is that a sector with a higher share of capital can more easily substitute between capital and emissions (provided that $\rho > 0$). This is also the case if the productivity of capital in a sector is higher (H_K) or if capital in a sector depreciates at a lower rate (δ)—see Proposition 3 (*iii*) and (*iv*). In all cases, it is efficient to tax carbon emissions in this sector at a higher rate as compared other to sectors with a lower β_K or H_K, or a higher δ, as such a carbon tax differentiation increases investments and economy-wide capital stock, and in turn future consumption and welfare while achieving the same level of carbon emissions.

2.4 Quantitative model and calibration

To explore the relevance of our theoretical results in an empirical context, we use a steady-state version of the model described in Section 2.2 calibrated to the current EU economy. Appendix 3.9.7 provides the steady-state conditions.

To capture technological heterogeneity, we decompose the main economic sectors and integrate empirical estimates regarding the sector-specific substitutability of capital and energy. The sectoral disaggregation is driven by the following considerations: we want to identify sectors which are responsible for the majority of CO_2 emissions in a typical industrialized economy as the EU, exhibit differences in capital emissions intensities, and are either subject to carbon pricing policies, for example, the EU ETS, or not. We distinguish the following sectors $j \in J = \{E, I, T, S\}$: electricity (E), energy-intensive industries (I), transportation (T), and services (S).[11]

Given policy choices on $\Xi = 0$ and $\tau_j \geq 0$, we need to pin down the following parameters to calibrate a steady-state model: (α_j, γ_j, δ_j, ζ, ρ_j, β_{Kj}, H_{Kj}, H_{Ej}, $\bar{L}$).

[11] We do not explicitly include agriculture as a separate sector as it only accounts for a small share of CO_2 emissions. It is beyond the scope of this paper to include emissions from non-CO_2 greenhouse gases.

Table 2.1. Overview of key parameter values for model calibration

Parameter	Model sector j			
	Electricity	*Industry*	*Transport*	*Services*
Capital rental rate (r_j)	0.14	0.20	0.10	0.10
Baseline carbon price (τ_j) [€ per ton of CO_2]	30	30	≈0	≈0
Share share of sectoral output in final output (γ_j)	0.04	0.21	0.05	0.70
Share share of L versus K-E aggregate (α_j)	0.29	0.65	0.63	0.61
Share share of K in K-E aggregate (β_{Kj})	0.76	0.83	0.99	0.99
Substitution elasticity between K-E in Y_j (ρ_j)				
All complements (central case)	-0.50	-5.00	-1.00	-4.00
All substitutes	0.40	0.10	0.30	0.20
Substitutes & complements	0.40	-5.00	0.15	-1.00
Strong substitutes & complements	0.50	-8.00	-1.00	-6.00

Notes: Parameter values are selected based on the following information. r_j: based on the *World Input-Output Database* (WIOD, Timmer et al. 2015). τ_j: chosen values intend to portray an average carbon price over the period 2010-2020 in the EU ETS. To enable the calculation of a positive value shares for carbon for the transport and services sectors, which have so far not been subject to (explicit) carbon pricing under the EU ETS, we choose a very small but positive value for τ_j for these sectors. γ_j and α_j: value share parameters based on WIOD. ρ_j: parameters describing the substitution between capital and energy are taken from the literature (for references, see text).

We select these parameters to capture the structure of the aggregated EU economy with respect to (i) the sectoral composition of output, (ii) the mix of capital, labor, and emissions input, and (iii) the observed savings rate. All parameter values are found from data targets without the need to simulate the model. Table 3.1 summarizes the parametrization for the central case of the model.

2.4.1 External parameters

To parametrize γ_j and α_j, we use two data sources. First, we obtain the value of sector-specific inputs of employment and capital from the WIOD (*World Input-Output Database*, Timmer et al. 2015). WIOD contains information at the country level, and we aggregate the data to the EU-28. We use data for the most recent available year 2014. We use the Standard Industrial Classification, Revision 4, to map WIOD categories to the four model sectors according to the mapping shown in Appendix 2.10. For European countries outside of the Euro zone, we apply exchange rate data from OECD (2020). Second, we obtain information on CO_2

emissions at the sectoral level from the European Commission's EDGAR database (*Emission Database for Global Atmospheric Research*, Crippa et al. 2019) and link emissions from fossil-fuel combustion as well as process emissions to the sectors using the mapping shown in Appendix 2.10. In order to compute value shares, we need to complement emissions data with an assumption about the costs of carbon emissions at the sectoral level. For the P and I sectors, which are subject to EU ETS regulation, we assume a carbon price of $\tau_P = \tau_I = 30$ € per ton of CO_2. The T and S sectors are not subject to EU ETS regulation, and we assume a carbon price of zero. We set $H_{Kj} = H_{Ej} = 1$, $\forall j$, as we cannot separately identify H and β from the data. Using the information on the value of inputs for K, L, and E for each sector, we can infer γ_j and α_j.

We survey the literature (Koesler and Schymura 2015, Okagawa and Ban 2008, van der Werf 2008, Costantini et al. 2019, Dissou et al. 2015, Papageorgiou et al. 2017) to pin down reasonable values for ρ_j. First, complementarity is the highest in the industry and services sectors (with lower estimates ranging from -6.96 to -2.33 and higher estimates ranging from -1.56 to 0). Second, the power and transportation sectors seem to exhibit a small degree of complementarity (with lower estimates ranging from -1.70 to -1.22 and upper estimates ranging from -1.17 to 0.21).

Given the uncertainty of empirical estimates and the central role of sectoral technology heterogeneity for our analysis, we consider four different cases representing different assumptions about the substitutability between capital and energy (emissions) input in sectoral production. *"All complements"* assumes that $\rho_j < 0$ for all sectors, reflecting a case where the substitutability between capital and energy inputs is limited. *"All substitutes"* represents a case with high substitutability. The other two cases (*"Substitutes & complements"* and *"Strong substitutes & complements"*) represent intermediate cases where capital and energy inputs are (strong) complements in some sectors and (strong) substitutes in others.

2.4.2 Calibrated parameters

Using the steady-state no-arbitrage condition (i.e., $\delta_j = r_j - \zeta$) and data on r_j from WIOD (taking the value for 2014), we determine ζ and δ_j. To obtain r_j, we divide the total sectoral capital compensation by the sectoral nominal capital stock employed. We calibrate the private discount rate of ζ by targeting an aggregate savings rate of 22 percent for the EU-28 (World Bank 2020). The savings rate is related to capital depreciation according to $\sum_j \delta_j K_j / \hat{Y}$. We obtain $\zeta = 0.0475$. Given ζ, we infer sector-specific depreciation rates δ_j. As labor is exogenous and enters in a Cobb-Douglas manner in sectoral production, we can normalize $\bar{L} = 1$.

Given ρ_j and τ_j, and using data on emissions from EDGAR and on the value of capital inputs and rental prices by sector from WIOD, we calibrate β_{Kj} based on combining the FOCs for profit-maximizing inputs of capital and emissions in each sector.

2.5 The computational experiment

To obtain insights into the nature of optimal multi-sector carbon pricing and associated welfare gains, we interact three dimensions in our simulations: the structure of sectoral CO_2 pricing ($\mathcal{S}$), policy stringency ($\mathcal{P}$), and technology heterogeneity ($\mathcal{H}$).

STRUCTURE OF CO_2 PRICING.—We consider three cases that differ in terms of the permissible structure of sectoral CO_2 prices: $s \in \mathcal{S} = \{$*Optimal*, *Uniform*, *Partitioned Pricing*$\}$. These cases cover the theoretical case of "optimal" pricing, where carbon prices τ_j are optimally differentiated across sectors (corresponding to a numerical evaluation of (2.18) in Proposition 1), and the case of uniform pricing which reflects "conventional" policy recommendation. An intermediate case represents a situation where sectoral CO_2 prices can only be partially differentiated because they are constrained by the presence of disintegrated carbon markets or multiple ETSs within a jurisdiction.

POLICY STRINGENCY.—We vary climate policy stringency to examine how the multi-sector carbon pricing depends on the economy-wide emissions budget.[12] We consider economy-wide emissions reductions of up to 80% (relative to a 2014 baseline emissions level): $p \in \mathcal{P} = \{0, 20, 40, 60, 80\}$.[13]

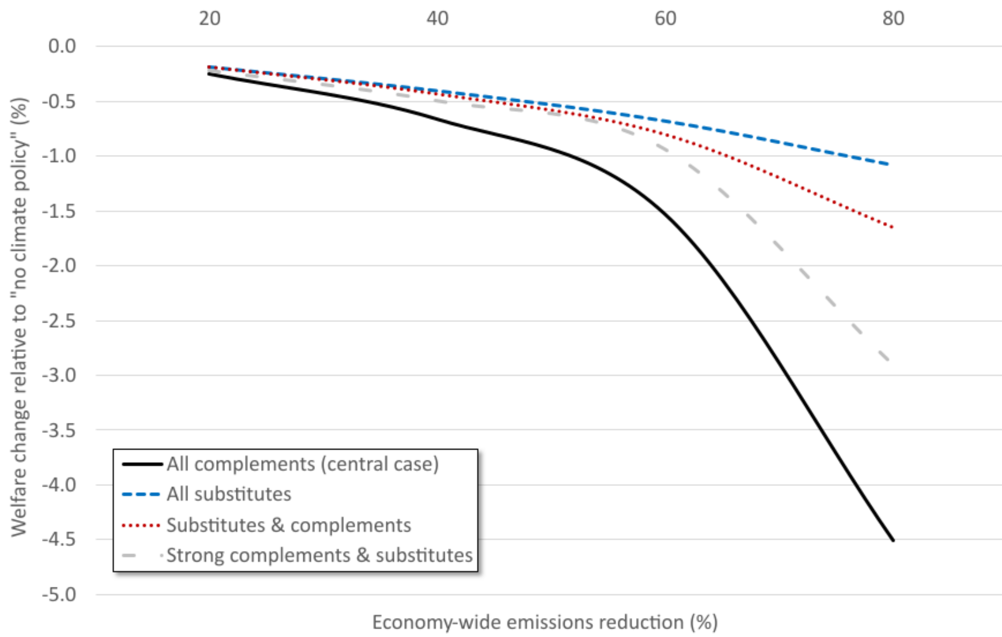
HETEROGENEOUS CAPITAL-ENERGY SUBSTITUTABILITY.—Technology heterogeneity in terms of the substitutability (complementarity) between capital K and energy (emissions) E inputs in sectoral production is the key driver of our main result that a uniform carbon price is not optimal (see Propositions 2 and 3). An important aspect of our computational experiment is examine the role of technology heterogeneity for multi-sector carbon pricing. We consider four cases ranging from "poor" to "high" K–E substitutability: $h \in \mathcal{H} = \{$*All complements*, *All substitutes*, *Substitutes & complements*, *Strong substitutes & complements*$\}$ as defined in Section 2.4 and Table 3.1.

MEASURING WELFARE GAINS.—We measure the economic cost of climate policy as:

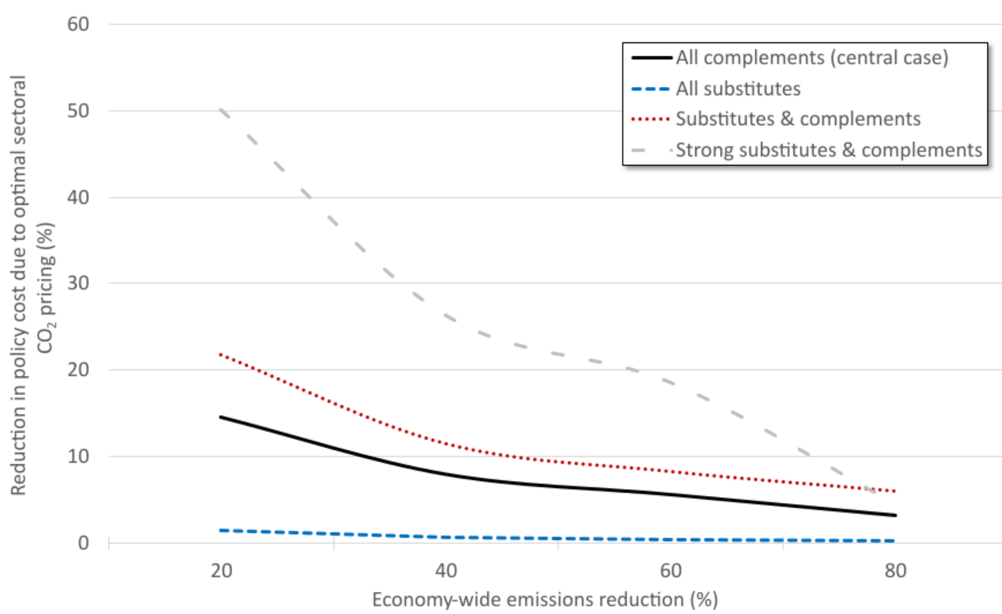
$$\Gamma^s_{h,p} := \frac{C^s_{h,p}}{C^s_{h,0}} - 1 \tag{2.20}$$

where $C^s_{h,p}$ is the steady-state aggregate consumption level. We calculate the reduction in the welfare cost of climate policy due to *"Optimal"* relative to *"Uniform"* carbon pricing as:

$$\Psi_{h,p} := \frac{C^{Optimal}_{h,p} - C^{Uniform}_{h,p}}{C^{Uniform}_{h,p} - C^{Uniform}_{h,0}} \, . \tag{2.21}$$



(a) Cost of climate policy with uniform carbon pricing ($\Gamma_{h,p}^{Uniform}$)



(b) Reduction in cost of climate policy under optimal sectoral relative to uniform carbon pricing ($\Psi_{h,p}$)

FIGURE 2.1. Welfare cost of climate policy with uniform and optimal CO_2 pricing by policy stringency $\mathcal{P}$ and for alternative assumptions about technology heterogeneity $\mathcal{H}$.

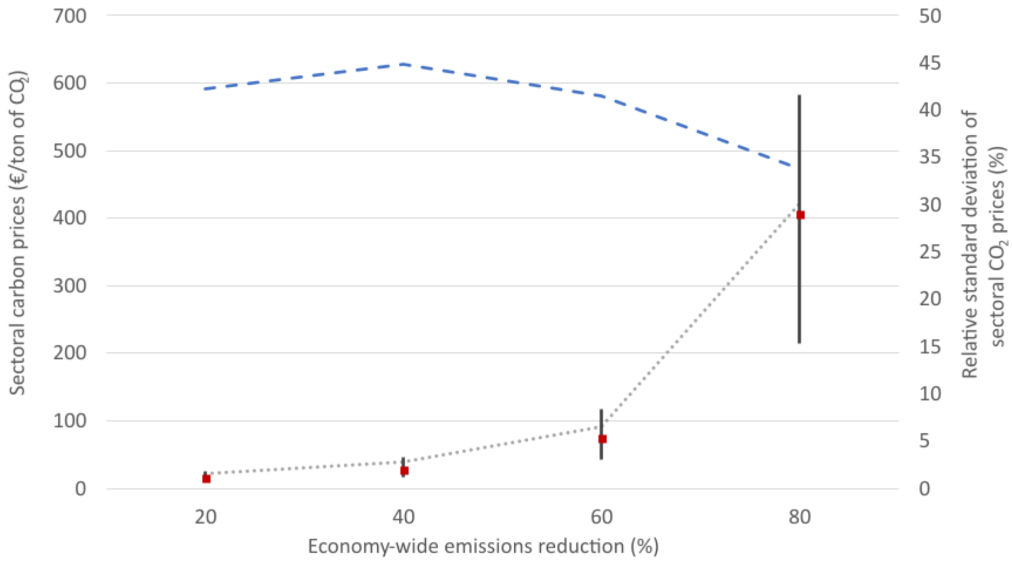


FIGURE 2.2. Dispersion of constrained-optimal sectoral CO_2 prices (τ_j), average CO_2 price, and uniform carbon price for alternative levels of policy stringency ($\mathcal{P}$).

Notes: Results show the dispersion of sectoral CO_2 prices (assuming the policy case "*Optimal*", central-case technology heterogeneity) is illustrated using the following metrics: **Red box** = emissions-weighted mean of sectoral prices, **Black vertical line** = indicates the range of minimum and maximum sectoral CO_2 prices; **Grey dotted line** = uniform CO_2 price under the policy case "*Uniform*"; **Blue dashed line** (on secondary vertical axis) = coefficient of variation, i.e. the emissions-weighted standard deviation divided by the emissions-weighted mean × 100.

2.6 Quantitative results

2.6.1 Multi-sector vs. uniform carbon pricing: Welfare gains

How large are the welfare gains from optimally differentiating carbon prices across sectors relative to uniform carbon pricing? Given the insights from Propositions 1–3, the answer to this question largely depends on the degree of technology heterogeneity in terms of capital-emissions (K–E) substitutability between sectors.

It is well known that the K–E substitutability influences the macroeconomic

[12]When comparing welfare, we always hold economy-wide CO_2 emissions $\bar{E}_t$ or, equivalently, emission reductions $\Delta\bar{E}_t$, fixed.

[13]The higher targets are in line with the stated ambitions of the European Commission (2020) to reduce greenhouse gas emissions in the EU economy by 55% in 2030, relative to 1990 levels, and to achieve climate neutrality by mid-century.

cost of climate change mitigation policies (see, for example, Weyant et al. 1996, Lu and Stern 2016). Thus, to provide context for our analysis, we first report the welfare cost of climate policy with uniform and optimal CO_2 pricing by policy stringency $\mathcal{P}$ and for alternative assumptions about technology heterogeneity $\mathcal{H}$ (Figure 2.1). Panel (a) shows $\Gamma_{h,p}^{Uniform}$, the welfare cost of achieving a given economy-wide emissions reduction target with uniform carbon pricing for the different cases of technology heterogeneity. Not surprisingly, the cost of climate policy increase more than proportionally with policy stringency, reflecting the increasing cost to substitute emissions for capital—given the CES production technologies in (3.5). Also, welfare cost is substantially higher if the economy exhibits a "poor" substitutability between capital and energy in production, in particular when the economy-wide emissions reduction target is high.

The main insight from Panel (b) in Figure 2.1 is that uniform carbon pricing gives away substantial efficiency gains by failing to (optimally) differentiate sectoral CO_2 prices. The reduction in the cost of climate policy due to *"Optimal"* carbon pricing ($\Psi_{h,p}$) are up to 50% relative to *"Uniform"* pricing. These gains, however, decrease with policy stringency as higher emissions reductions imply that eventually all sectors need to decarbonize substantially, in turn diminishing the scope for efficiency gains by shifting sectoral abatement patterns. Moreover, these gains depend crucially on the degree of technology heterogeneity between sectors. If all sectors exhibit a high substitutability between capital and energy, the welfare gain from differentiating sectoral CO_2 prices is low. In contrast, if the economy is composed of sectors with complementary relationships between K and E inputs and, in addition, features sectors where K and E are substitutes, the reductions in climate policy cost can be substantial, even for high levels of policy stringency.

2.6.2 Technology heterogeneity matters

CO_2 PRICE DIFFERENTIATION.—We find that the differentiation of CO_2 prices across sectors under optimal policy is quantitatively significant.

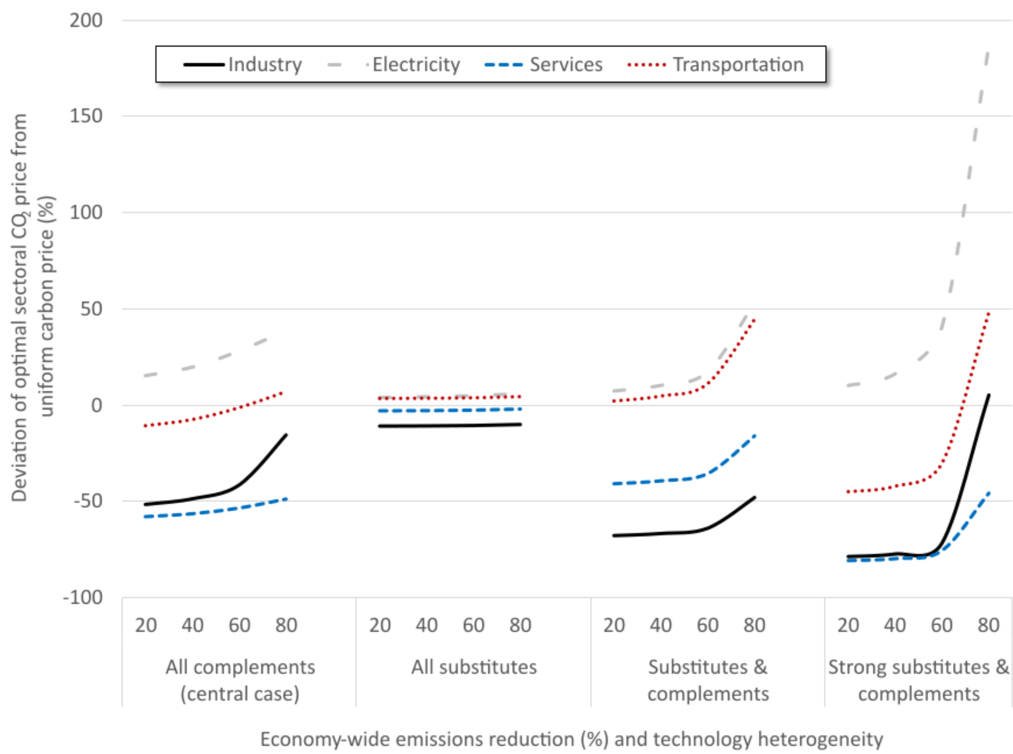


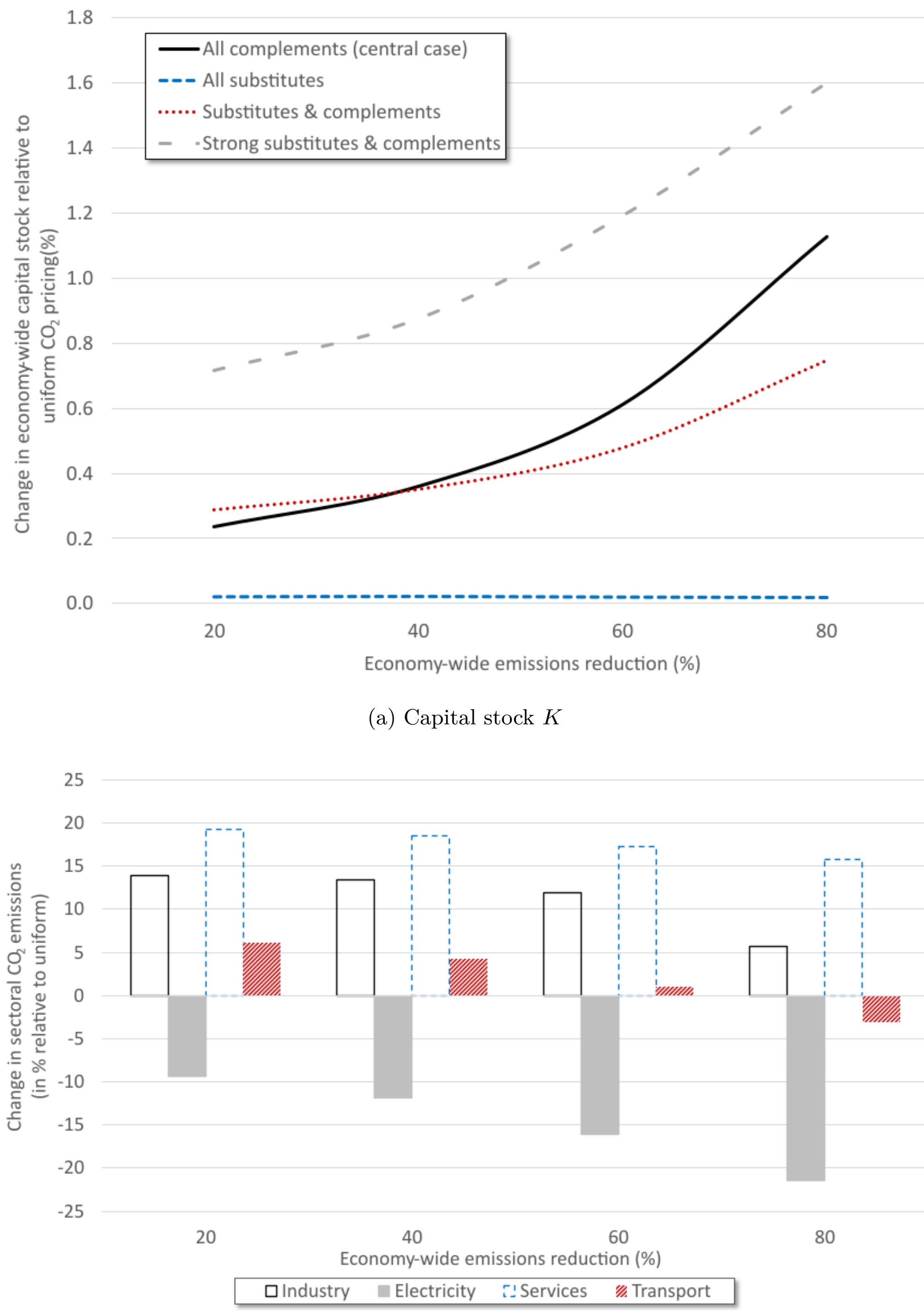
Figure 2.3. Deviation of constrained-optimal sectoral CO_2 prices from uniform carbon pricing by policy stringency by technology heterogeneity.

Table 2.2. CO_2 prices (€/ton of CO_2) under uniform and constrained-optimal sectoral carbon pricing.

Policy target (%)	*"Optimal"* sectoral CO_2 pricing $\tau^*_{Industry}$	$\tau^*_{Electricity}$	$\tau^*_{Services}$	$\tau^*_{Transport}$	$\overline{\tau^*}^a$	*"Uniform"* CO_2 pricing $\tau = \tau_j, \forall j$
20	10.7	25.6	9.3	19.8	16.2	22.2
40	20.0	46.8	17.0	36.1	28.7	39.1
60	53.3	116.9	42.3	89.9	74.4	91.3
80	355.1	583.5	214.7	449.6	406.4	420.7

Notes: Results shown assume the central-case for technology heterogeneity. [a]: Emissions-weighted mean of optimal sectoral CO_2 prices τ^*_j.

Figure 2.2 shows the dispersion of optimal sectoral CO_2 prices, the respective mean, and the corresponding uniform carbon price to achieve a given emission reduction target. For example, for a 40% economy-wide emissions reduction in the



(a) Capital stock K

(b) Sectoral CO_2 emissions E_j

Figure 2.4. Impacts on steady-state capital stock and sectoral emissions under constrained-optimal sectoral CO_2 prices relative to uniform carbon pricing.

European economy, optimal sectoral prices range from €42.3 to €116.9 per ton of CO_2, with an emissions-weighted mean CO_2 price of €74.4/ton. By comparison, achieving the same emissions target would require a price of €91.3/ton. We also find that the variation in optimal CO_2 prices is robust with respect to the economy-wide emission reduction target: the coefficient of variation of optimal sectoral CO_2 prices is roughly constant at 40% for targets of up to 80%.

Figure 2.3 reports the deviation of optimal sectoral CO_2 prices from uniform carbon pricing. The differentiation of sectoral CO_2 prices hinges on (1) the heterogeneity in the K–E substitutability between sectors and (2) the complementarity between K–E for at least some sectors. First, if K–E are "substitutes" in all sectors of the economy, the deviations from a uniform carbon price are negligible. Then, sectoral CO_2 pricing has a small effect on capital demand in each sector compared to uniform pricing, implying that differentiated CO_2 prices are not very effective in creating an incentive for capital accumulation beyond what is already implied by economic agents' saving decisions based on private discounting. Accordingly, the impact on the aggregate capital stock is small under *"All Substitutes"* (see Figure 2.4), consistent with the negligible welfare gains from optimal sectoral CO_2 pricing (see Panel (b) in Figure 2.1). Thus, an overall high K–E substitutability in the economy, as reflected by the case of $\rho_j > 0$ in all sectors, almost entirely dampens the mechanism of using sectoral emission prices to implicitly subsidize capital to address the capital externality resulting from differential social discounting.

Second, if all sectors are "complements" and differ with respect to the degree of complementarity, differentiating sectoral CO_2 provides an effective way of increasing economy-wide capital demand and accumulation. Table 2.2 complements Figure 2.3 by showing the absolute level of sectoral CO_2 prices in €/ton of CO_2. Carbon is priced at a substantially lower rate in sectors for which the K–E substitutability is relatively poor (i.e., $\tau^*_{Industry}$ and $\tau^*_{Services}$ are about half of the CO_2 price under *"Uniform"* emissions pricing and about one third lower than the mean of optimal sectoral carbon prices $\overline{\tau^*}$). Accordingly, sectoral CO_2 prices in "Electricity" and "Transport", which exhibit a higher substitutability between capital and emissions relative to the "Industry" and "Services", exceed the uniform carbon price and $\overline{\tau^*}$ considerably.

Third, for a given emissions reduction target, the optimal CO_2 pricing policy yields greater welfare gains, the more heterogeneous is the substitutability between capital and emissions among sectors. *"Strong substitutes & complements"* represents the largest degree of technology heterogeneity among the four cases and is associated with the largest gains in aggregate capital and welfare. Here, the optimal pattern of sectoral CO_2 price differentiation is such that the complementary sectors "Industry" and "Services" each receive a significantly lower price and decrease less relative to the substitution sectors "Electricity" and "Transport". This large heterogeneity in the substitutability between K–E implies quite heterogeneous sectoral responses in the sectoral use of capital. Optimal carbon pricing can then exploit this mechanism to incentivize economy-wide capital accumulation with positive welfare effects for future generations.

IMPACTS ON CAPITAL STOCK AND SECTORAL EMISSIONS.—Panel (a) in Figure 2.4 shows that the capital stock increases when sectoral CO_2 prices are optimally differentiation. The more stringent the climate policy, the greater the increase. Intuitively, steering the sectoral use of emissions and indirectly of capital through appropriate sectoral CO_2 prices has greater leverage when the economy-wide emissions budget is small. However, the incremental gains from capital accumulation at higher reduction targets are not large enough to overcompensate for the increasing economy-wide costs resulting from the limited substitutability between capital and emissions at the economy-wide level. Regardless of the gains from capital accumulation, therefore, the reduction in the costs of climate policy fall with policy stringency, as shown in Panel (b) in Figure 2.1.

Consistent with the pattern of sectoral carbon prices, Panel (b) in Figure 2.4 shows that, more (less) CO_2 emissions are abated in the relatively flexible (inflexible) sectors, relative to uniform carbon pricing. This is in line with Proposition

3: allocating more (less) of the economy's aggregate emissions budget to the less (more) flexible sectors spurs additional capital accumulation in the economy consistent with a social perspective that places a higher value on the welfare of future generations than private discounting implies.

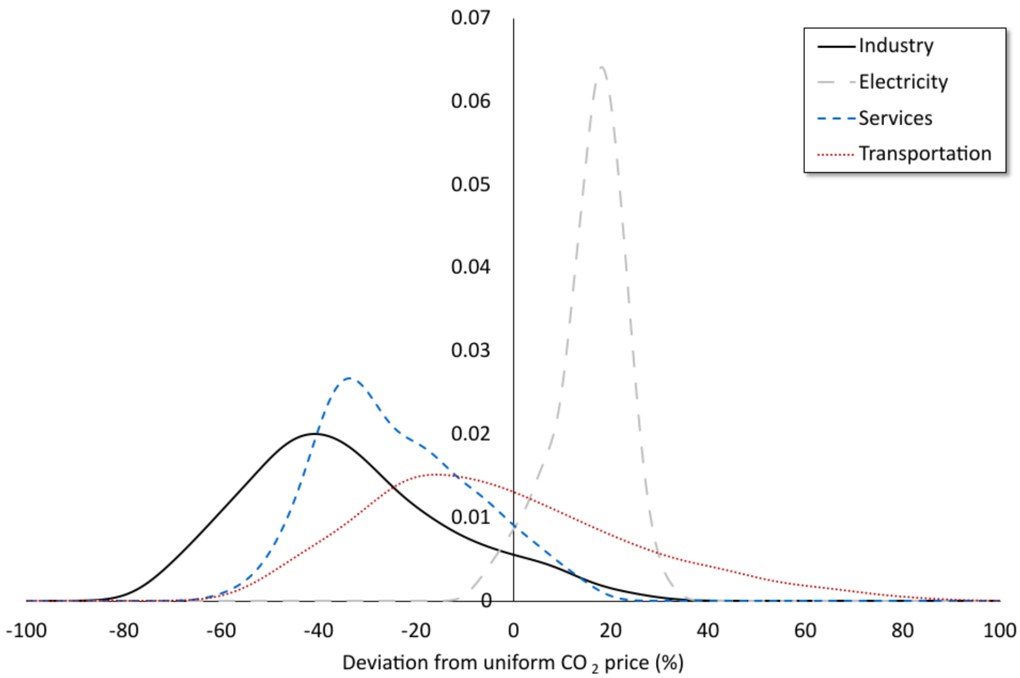
2.6.3 Non-uniform optimal carbon prices: How convincing?

While the four cases of technology heterogeneity considered so far are useful to develop an intuition for the conditions under which optimal sectoral CO_2 pricing yields quantitatively significant welfare gains (relative to uniform carbon pricing), we next conduct a systematic assessment of the impact of technology heterogeneity.

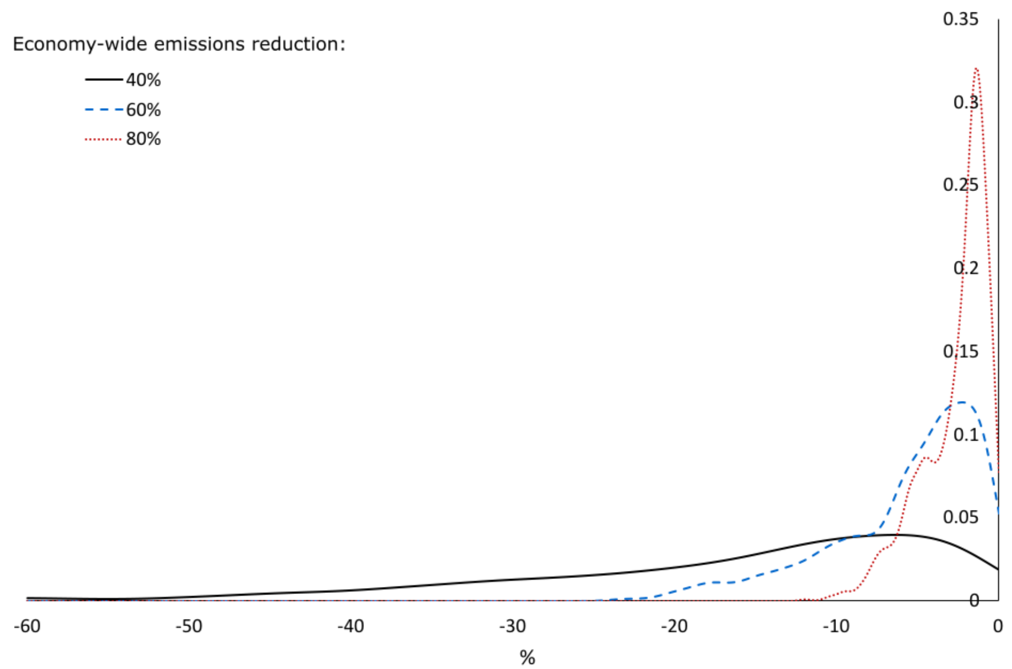
SYSTEMATIC SENSITIVITY ANALYSIS.—Given the lack of empirical estimates that would characterize the (joint) distribution of the capital-energy substitutability in sectoral production (ρ_j), we adopt the following approach to sample technology heterogeneity. We assume that $\rho_j,\ j \in \{E, I, T, S\}$ are independently and uniformly distributed with support $[a_j, b_j]$. We draw a sample of 10'000 sets of parameter values using our survey of the literature on "piecemeal" estimates for ρ_j (see Section 2.4.1).[14] To decompose the impact of policy stringency, we conduct Monte-Carlo simulations for economy-wide emissions reduction targets of 40, 60, and 80 percent, respectively.

ROBUSTNESS OF NON-UNIFORM CO_2 PRICES AND WELFARE GAINS.—Figure 2.5 shows Kernel density estimations of PDFs for the deviation in optimal sectoral CO_2 prices relative to uniform carbon pricing in Panel (a) and the reduction in policy cost from optimal sectoral relative to uniform carbon pricing for alternative economy-wide emissions reduction targets in Panel (b). The key insight, which supports our previous results, is that both the welfare gains and the extent of CO_2

[14] Specifically, we assume that $[a_E, b_E]$=$[-1.52, 0.65]$, $[a_I, b_I] = [-6.69, -0.041]$, $[a_T, b_T]$=$[-1.70, 0.21]$, and $[a_S, b_S]$=$[-2.70, -0.47]$. We note that the "empirically-informed" support for our Monte-Carlo analysis emphasizes the complementarity between K–E at the sectoral level. This reflects the difficulty of replacing fossil fuels in *all* sectors of the economy in the absence of major break-through technologies. In the long-run, ρ may well be very large. Our analysis should thus be viewed as representing a time horizon of several decades at most.



(a) Optimal sectoral carbon prices (τ_j^*) vs. uniform carbon price (τ) for a 40% reduction target



(b) Reduction in policy cost due to optimal sectoral relative to uniform carbon pricing ($\Psi_{h,p}$) for alternative levels of emissions reduction

Figure 2.5. Kernel estimation of probability density functions (PDFs) for systematic variation of technology heterogeneity.

Notes: Based on Monte-Carlo simulations with 10'000 draws systematically varying the sector-specific capital-energy (emissions) substitutability. Sample mean and standard deviation for PDFs in Panel (a) by sector: "Electricity": (15.7%, 7.3%); "Industry": (-34.1%, 20.9%); "Transportation": (-2.9%, 27.1%); "Services": (-24.0%, 15.6%). Sample mean and standard deviation for PDFs in Panel (b) by emissions reduction target: 40%: (-17.4%, 0.15%); 60%: (-5.7%, 0.10%); 80%: (-2.6%, 0.02%).

price differentiation significantly differ comparing constrained-optimal to uniform carbon pricing. First, for the given variation in technology heterogeneity, there is a large probability for optimal sectoral CO_2 to deviate from the respective uniform carbon price. The two sectors "Industry" and "Services," where K–E substitutability is drawn from distributions with relatively large negative (small positive) values for the lower (upper) bounds, show the highest downward deviation with a mean of 34.1% and 24.0% and a standard deviation of 20.9% and 15.6%, respectively. This is in line with the intuition developed so far that "complement" sectors abate less (i.e., face a lower CO_2 price) under optimally differentiated carbon pricing. CO_2 emissions in the "Electricity" sector, which is characterized by a relatively large degree of K–E substitutability, are in almost all cases priced at a higher rate than under uniform pricing, with a mean and standard deviation of 15.7% and 7.3%, respectively. The optimal sectoral CO_2 price in "Transportation" is lower or higher than under uniform pricing, with a mean and standard deviation of -2.9% and 27.1%. Second, the mean reduction in the policy cost from optimal sectoral CO_2 pricing is 17.4% for a 40% emission reduction target, but drops to 5.7% and 2.6% for 60% and 80% reduction targets, respectively. While focusing on the mean impact reflects the intuition behind Panel (b) in Figure 2.1, the PDFs of relative welfare gains show a large dispersion, in particular for smaller emissions reduction targets. This underscores again the impact of technology heterogeneity.[15]

2.6.4 Implications for designing emissions trading

So far, we have essentially adopted a carbon tax perspective and assumed that a social planner can set τ_j^* directly and without constraints. In real-world climate policy, a price on carbon is often established through the "twin" of the tax-based approach: an emissions trading system (ETS). This section explores the implications of our main finding that CO_2 prices should not be uniform but should differ

[15] For example, for an emissions reduction target of 40%, optimal sectoral CO_2 reduces the policy cost by at least one third in about 25% of the cases.

across sectors for the design of an ETS.

POLICY CONTEXT.—Our reasoning is motivated by the policy context for the application of emissions trading systems, which applies to virtually all countries that apply an ETS. Major examples of ETSs are the EU ETS, national trading systems in China and Canada, and regional carbon markets in the U.S. (for an overview of carbon pricing policies, see, for example, World Bank 2021). In virtually all of these jurisdictions, however, the ETS covers only a subset of sectors and, hence, emissions. Emissions in the remaining parts of the economy are subject to separate environmental regulation. A prominent example of such partitioned regulation is the EU's climate policy, where the overall emissions target is split between sources that fall within and outside the EU ETS. In fact, the European Commission (2021c) plans to introduce a second ETS in Europe starting in 2026 that would regulate most of the emissions sources that are currently outside of EU ETS. Hence, two ETSs would exist in parallel.

Against this background, we analyze three important questions that policymakers face in designing cost-effective emissions trading systems: (1) is a single, comprehensive ETS that overcomes partitioned emissions trading desirable, (2) how should the economy-wide CO_2 budget be allocated among ETSs, or, equivalently, what should be the emissions cap in each ETS, and (3) how should economic sectors be assigned to ETSs?

A SINGLE COMPREHENSIVE OR MULTIPLE NARROW ETSs?—Motivated by the EU climate policy context, we focus on the case of two ETSs, labeled *"ETS-1"* and *"ETS-2"*, and examine how the assignment of economic sectors and the allocation of the economy-wide emissions budget to the two ETSs impacts economic costs. In terms of assigning sectors to ETSs, we consider all possible permutations obtained by assigning the four sectors resolved in our quantitative model to two ETSs (see Figure 2.6). The first case represents in a stylized way the current EU climate policy framework: *"EU ETS"* contains the *"Electricity"* and the *"Industry"* sectors while *"Transportation"* and *"Services"* are regulated outside of the EU ETS.

For the emissions trading context, our main finding that optimal CO_2 prices differ across sectors leads to the following conclusion:

Corollary 1 *When social and private discount rates are different and sectoral production technologies are heterogeneous, a single comprehensive ETS is not optimal.*

This follows directly from Proposition 3, assuming that prices (carbon taxes) and quantities (emissions trading) are equivalent.[16] If so, our analysis up to this point can be viewed as an analysis of the regulation of aggregate CO_2 emissions by J partitions or ETSs, where the emissions cap of each ETS is given by $\bar{E}_j$. Notably, Corollary 1 stands in sharp contrast to the widely-held view among economists that the scope of an ETS should be as broad as possible and that several parallel ETSs should be integrated into a single ETS.

This insight is also supported by our quantitative analysis. Figure 2.6 shows the reduction in policy cost ($\Psi_{h,p}$) from partitioned emissions trading relative to a single ETS, or, equivalently, uniform carbon pricing. It displays the three alternative assignments of economic sectors to the two ETSs for different emissions budgets in the *"ETS-1"*. It is straightforward to see that the allocation of the economy-wide emissions budget across the two partitions significantly affects costs. A key main insight borne out by Figure 2.6 is that partitioned emissions trading can outperform a single ETS that covers all sectors and thus prices carbon uniformly across the economy. The reason is that in an economy with differential social discounting, it is not optimal to price carbon uniformly. Following the intuition developed by analysis in Sections 2.3 and 2.6), differentiating sectoral CO_2 prices exploits sectoral differences in substitutability between "dirty" fossil energy and "clean" capital, and help create incentives for investment and capital accumulation beyond is implied by households' savings based on private discounting.[17]

[16] For example, by abstracting from uncertainties in technology abatement costs (Weitzman 1974) or market power considerations (Hahn 1984).

[17] Our result of differentiated CO_2 prices could also be implemented in a integrated ETS with a single market where exchange rates for permits reflect the heterogeneous substitutability of sectors. Compared to two separate ETSs, such a setting would likely lead to an improved outcome, as both

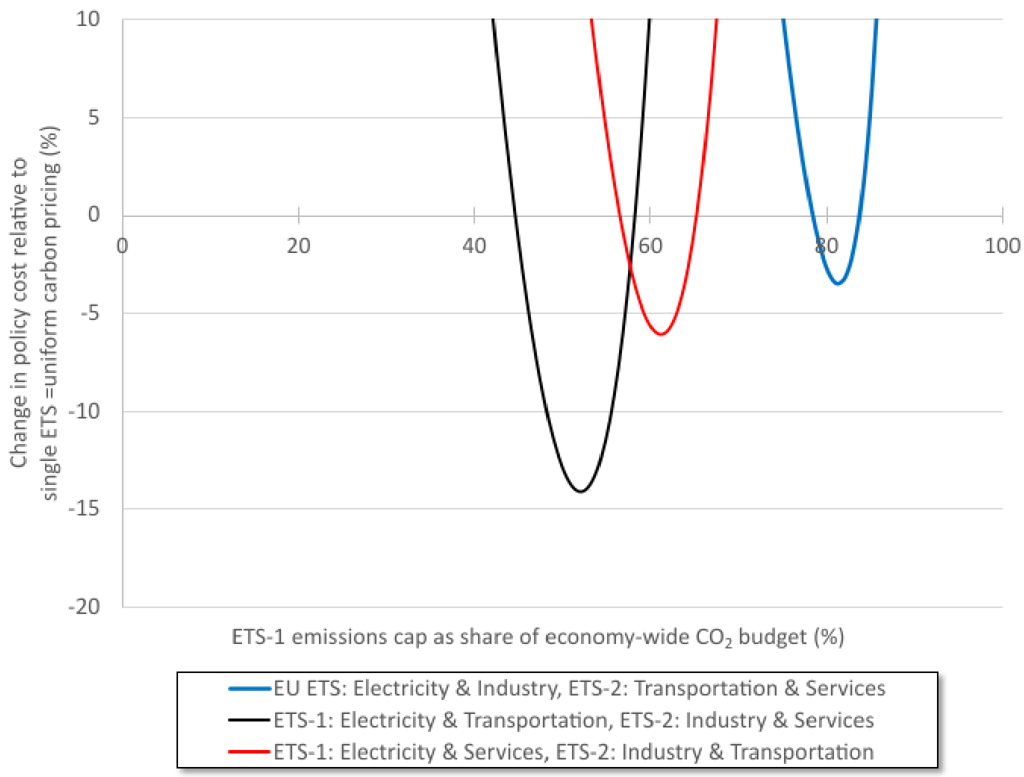


FIGURE 2.6. Reduction in policy cost ($\Psi_{h,p}$) for 40% economy-wide emissions reduction under partitioned ETSs relative to uniform carbon pricing for (1) different assignments of sectors to partitions and (2) different allocations of the economy-wide CO_2 budget (assuming central-case technology heterogeneity).

KEY POLICY CHOICES: ASSIGNING SECTORS AND CAPS TO DIFFERENT ETSs.— If a single, comprehensive ETS is not desirable, how should partitioned emissions trading be designed? Unlike in the case of carbon taxes, it is not possible to directly set a sectoral CO_2 price. Whether and to what extent partitioned emissions trading outperforms uniform carbon pricing depends on how sectors are assigned to partitions and how the emissions budget is allocated among partitions. First, if too little or too much emissions budget is allocated to one of the two ETSs, costs rise sharply and the efficiency gains from differentiating CO_2 across partitions cannot

approaches need approximately the same information and the larger market would trade more often and exploit greater heterogeneity.

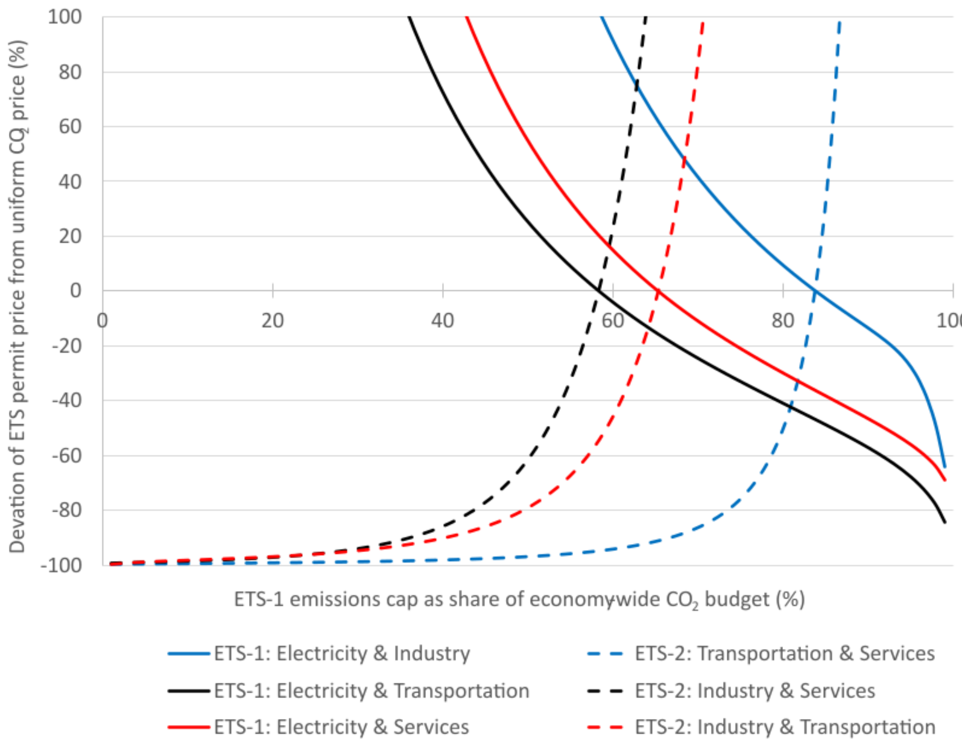


FIGURE 2.7. Deviation of ETS permit prices under partitioned trading from uniform carbon pricing for (1) different assignments of sectors to partitions and (2) different allocations of the economy-wide CO_2 budget (assuming central-case technology heterogeneity).

compensate for this. Intuitively, relying largely on an ETS which only covers a subset of emissions source forgoes efficiency gains from exploiting "where-flexibility" across sectors in emissions abatement. Second, the assignment of sectors to partitions plays an important role, too. Based on Proposition 3 and the quantitative results in Section 2.6.2, sectors with a low (high) capital-energy substitutability should receive a relatively high (low) CO_2.

The implication of this result is that one should assign sectors to ETSs based on their relative capital-energy substitutability: sectors with a relatively poor substitutability (*"Industry"* and *"Services"*) should be clustered in one partition and sectors with a relatively high substitutability (*"Electricity"* and *"Transportation"*) in the other partition. Under such a case, represented by the black line in Figure 2.6, partitioned emissions trading yields the largest efficiency gains relative to uniform carbon pricing, i.e. a single ETS. The assignment of sectors, on the other hand, as is done in the current EU climate policy, represented by the blue line, leads only to minor efficiency gains.

Since the CO_2 permit price cannot differentiate between sectors within a partition, the assignment of sectors to ETSs is important to provide the ability to appropriately differentiate CO_2 prices in the first place. When the assignment of sectors is chosen such that sectors with low and high capital-energy substitutability are bundled together in one partition, the permit prices between *"ETS-1"* and *"ETS-2"* have to differ substantially from one another in order to create reductions in the policy cost relative to uniform carbon pricing (see Figure 2.7). For example, policy cost are reduced by nearly 15% if the ETS-2 permit price deviates downward by about 60% and the ETS-1 permit price deviates upward by about 5% from the uniform CO_2 price (black line). With the sectoral assignment reflective of the current situation in EU climate policy (blue line), the permit price between ETSs do not differ much at the point where the reduction in policy cost is maximal. This is because the sectoral assignment does not allow to differentiate the underlying sectoral CO_2 prices in an appropriate way: differentiating permit prices across the ETSs would effectively imply high CO_2 prices for sectors with a high and low capital-energy substitutability which is not efficient.

2.7 Robustness

FINAL GOOD AGGREGATION.—We have emphasized so far the role technology heterogeneity at the sectoral level. At the level of the final good $\hat{Y}$, we have assumed that sectoral goods Y_j are aggregated in a Cobb-Douglas fashion. While this assumption is not uncommon in some of major energy-environment-economy models (Goulder et al. 2016, for example,), sectoral goods may in fact be imperfect substitutes, suggesting an elasticity of substitution (EOS) below 1 (for example, Chen et al. 2015, Capros et al. 2013). Using a CES aggregator with an EOS<1, we find that the optimal sectoral CO_2 prices are differentiated to an even higher degree. Everything else equal, we find that the more complementary sectoral goods in final-good aggregation are, the larger (lower) is the optimal CO_2 price for sectors for which capital and energy are relatively poor (good) substitutes. The intuition is that when sectoral goods cannot be easily substituted, climate policy has a larger leverage on capital accumulation through differentiating sectoral CO_2 prices. To the extent that the central model overestimates substitutability between sectoral goods, the case for optimal CO_2 prices that differ across sectors is even stronger.

LIMITED SUBSTITUTABILITY BETWEEN LABOR AND CAPITAL-ENERGY.—The central model assumes a unitary EOS between labor (L) and the capital-energy (K-E) bundle in sectoral production Y_j. As substitutability between L and K-E decreases in sectors where K and L are relatively highly complementary, the extent of sectoral CO_2 price differentiation decreases. In the limiting case, when L and K-E are perfect complements, we find that uniform carbon pricing is approximately optimal. On the other hand, decreasing the substitutability between L and K-E in the sectors where K and L are relatively good substitutes, increases the extent of sectoral CO_2 price differentiation diminishes. When L is a complement relative to K-E, differentiating CO_2 prices does not only affect capital but also distorts to allocation labor across sectors. These costs have to weighed against the benefits from incentivizing the accumulation capital in line with the social discounting motive.

While there is some evidence in support of the capital-skill complementarity hypothesis at the aggregate production level, the evidence is not very strong (Duffy et al. 2004). If we assume that the EOS between L and $K - E$ is 0.5, i.e. taking a midpoint between our central case (EOS=1) and the case of perfect complements (EOS=0), we still find that optimal CO_2 prices are significantly differentiated between sectors and that welfare gains are 85% of the gains in the central-case parametrization.

NEAR-ZERO SOCIAL DISCOUNTING.—While we have assumed throughout the analysis that $\zeta_S = 0$, our main results are robust to using a social discount rate close to zero (Stern 2006), provided the distance from the private discount rate (ζ) is sufficiently large. As both discount rates converge, the case for differentiating sectoral CO_2 prices diminishes. If $\zeta_S = \zeta$, uniform carbon pricing is optimal (see Proposition 2).

PRE-EXISTING, DISTORTIONARY INCOME TAXES.—Our analysis abstracts from pre-existing income taxes that are prevalent in today's economies. To the extent that such fiscal instruments are distortionary, optimal carbon taxes have been shown to be smaller compared to using lump-sum taxes to raise government revenues (Bovenberg and Goulder 1996, Barrage 2019). If we were to include, for example, positive capital income taxes $\Xi > 0$ in our model, we would add another distortion to the capital market in addition to differential social discounting. To see this, note that from the private Euler equation (3.3) the steady-state capital rental rate in the decentralized equilibrium is given by $r_j = \zeta/(1 - \Xi) + \delta_j$. The optimal capital rental rate, given the social discount rate and using the Chamley (1986)–Judd (1985) result that capital income should not be taxed, is given by: $r_j^* = \zeta_s + \delta_j$. It is thus straightforward to see that a positive capital income tax, like the social discounting externality ($\zeta > \zeta_S$), means that $r_j > r_j^*$. If capital income taxes cannot be adjusted by policy, the argument for CO_2 price differentiation would become even stronger, since the τ_j^* would then have to correct for two simultaneous capital market distortions. When capital income taxes could be partially adjusted to reduce the distortion from income taxation, but could not be completely reduced to zero (for example, because an exogenous level of government spending must be financed), the extent of CO_2 price differentiation would still be greater than in our central-case model.[18]

[18] We admit that an in-depth analysis of the role of pre-existing and distortionary income taxation for multi-sectoral carbon pricing is beyond the scope of this paper, which focuses on differential social discounting.

2.8 Conclusion

This study revisited the well-known result that a uniform carbon price minimizes the welfare costs of achieving a given economy-wide emissions target. Our analysis revealed the implicit assumptions behind this result. We showed that a uniform carbon price is optimal only when the social and private discount rates are equal. When discount rates differ, strong assumptions are required: it must be possible to subsidize aggregate capital income or the various sectors in an economy must be identical in terms of their substitutability of "dirty" fossil energy with "clean" capital. Otherwise, the result only survives if one assumes that a carbon price has no effect on investments and capital accumulation. These assumptions are not plausible.

When these assumptions are not met, this paper showed that optimal multi-sector carbon pricing differentiates CO_2 prices across sectors. Technological heterogeneity is key to our finding: sectors in which "dirty" fossil energy (i.e., CO_2 emissions) are not easily substitutable with "clean" capital should receive a lower carbon price than sectors where these two inputs are better substitutes. When capital in a given sector is a "poor" substitute for emissions, pricing carbon destroys more capital as compared to reducing the same amount of CO_2 emissions in a sector with a higher substitutability. With differential social discounting, economic agents discount the future too much and differentiating sectoral CO_2 prices avoid households failing to invest sufficiently in the economy's capital stocks, boosting capital accumulation and benefiting future generations.

We showed that the differentiation of CO_2 prices across sectors has a first-order effect on welfare. For a 40% economy-wide emissions reduction in the EU economy, we estimated that optimal sectoral prices range from €42.3 to €116.9 per ton of CO_2, with a mean CO_2 price of €74.4/ton. To achieve the same environmental target, a much higher uniform carbon price of €91.3/ton is required. Depending on technological heterogeneity and policy stringency, welfare gains can be as much as half the cost of climate policy under uniform carbon pricing.

We argued that our findings have important implications. First, model-based evaluations of climate policy which (unknowingly) assume that social and private discount rates are equal and ignore technology heterogeneity amount to assuming that a uniform carbon price is optimal. To the extent that these assumptions are not warranted, they overlook better carbon pricing policies and overstate the costs of climate policies. Second, partitioned emissions regulation through separate ETSs within one jurisdiction, as is expected for EU climate policy, does not necessarily lead to higher costs. We showed that two separate ETSs may be superior to a single, comprehensive ETS if sectors are assigned to each ETS such that sectors with relatively low and high capital-energy substitutability are clustered separately. These considerations are relevant to the practical design of ETSs, especially since decarbonization efforts in many countries relies on one or more markets for tradable emission permits as the cornerstone of climate policy.

In this paper, we expounded the assumptions in general equilibrium which underlie the established view that uniform carbon pricing is optimal. Different models offer different policy recommendations and we should settle the mapping from models to policy recommendations, on the one hand, and discuss the applicability of one model versus another, on the other hand. The scope of this paper has been concerned with the former, not the latter.

2.9 Appendix A: Theoretical derivations and proofs

2.9.1 Definition of decentralized equilibrium

Given the supply of sector emissions $\bar{E}_{jt}$, $\forall jt$, an equilibrium is given by the sequence of prices and quantities comprising consumption (C_t), capital and labor supply $(\bar{K}_{jt}, \bar{L})$, final good and sectoral outputs $(\hat{Y}_t, Y_{jt})$, demands for capital, labor, and emissions (K_{jt}, L_{jt}, E_{jt}), investments I_t, wage and capital rental rates (w_t, r_{jt}), prices for final and sectoral goods $(\hat{p}_t = 1, p_{jt})$, and sectoral prices for CO_2 emissions (τ_{jt}) such that: (i) (C_t) maximizes lifetime utility of households; (ii) $(\hat{Y}_t, Y_{jt})$ maximize profits of the final good producer; (iii) $(Y_{jt}, K_{jt}, L_{jt}, E_{jt})$ maximize profits of the sectoral goods producers; (iv) the wage and capital rental rates (w_t, r_{jt}) and prices for final and sectoral goods $(\hat{p}_t = 1, p_{jt})$ clear respective goods markets, (v) sectoral carbon prices (τ_{jt}) clear sectoral emission markets, and (vi) the evolution of the capital stock is governed by (3.8).

2.9.2 First-best social planner's problem

In the first-best setting, the social planner solves the following problem:

$$
\begin{aligned}
&\max_{\{K_{jt},\bar{K}_{jt+1},Y_{jt},\hat{Y}_t,C_t,L_{jt},\bar{E}_{jt},E_{jt},I_t\}_{j=1}^{J}} \sum_{t=0}^{\infty}\left(\frac{1}{1+\zeta_S}\right)^t \Bigg[u(C_t) \qquad (2.22)\\
&+\lambda_t^E\Big(\bar{E}_t-\sum_{j=1}^{J}E_{jt}\Big)+\lambda_t^L\Big(\bar{L}-\sum_{j=1}^{J}L_{jt}\Big)\\
&+\sum_{j=1}^{J}\lambda_{jt}^{K_j}\Big(\bar{K}_{jt}-K_{jt}\Big)+\lambda_t^K\Big(-\sum_{j=1}^{J}\bar{K}_{jt+1}+\sum_{j=1}^{J}(1-\delta_j)\bar{K}_{jt}+I_t\Big)\\
&+\sum_{j=1}^{J}\mu_{jt}\Big(Y_{jt}(L_{jt},K_{jt},E_{jt})-Y_{jt}\Big)+\hat{\mu}_t\Big(\hat{Y}_t(Y_{1t},\ldots,Y_{jt},\ldots,Y_{Jt})-\hat{Y}_t\Big)\\
&+\mu_t\Big(\hat{Y}_t-C_t-I_t\Big)\Bigg],
\end{aligned}
$$

where $\lambda_{jt}^{E_j}, \lambda_{jt}^{K_j}, \lambda_t^L$ denote the shadow prices of input choices in sector j for emissions, capital, and labor, respectively. λ_t^E is the shadow price of economy-wide emissions, λ_t^K the shadow price of aggregate capital, μ_{jt} and $\hat{\mu}_{jt}$ are the shadow prices of sector j and final output, and μ_t is the shadow cost of consumption. The FOCs are given by:

$$\begin{aligned}
U_{Ct} &= \mu_t \\
\mu_{jt} \frac{\partial Y_{jt}(L_{jt}, K_{jt}, E_{jt})}{\partial L_{jt}} &= \lambda_t^L \\
\mu_{jt} \frac{\partial Y_{jt}(L_{jt}, K_{jt}, E_{jt})}{\partial E_{jt}} &= \lambda_t^E \\
\mu_{jt} \frac{\partial Y_{jt}(L_{jt}, K_{jt}, E_{jt})}{\partial K_{jt}} &= \lambda_{jt}^{K_j} \\
\hat{\mu}_t \frac{\partial \hat{Y}_t(Y_{1t}, \ldots, Y_{jt}, \ldots, Y_{Jt})}{\partial Y_{jt}} &= \mu_{jt} \\
\hat{\mu}_t &= \mu_t \\
\lambda_t^K &= \mu_t \\
\left(\frac{1}{1+\zeta_S}\right)^{t+1} \lambda_{jt+1}^{K_j} - \left(\frac{1}{1+\zeta_S}\right)^{t} \lambda_t^K + \left(\frac{1}{1+\zeta_S}\right)^{t+1} \lambda_{t+1}^K (1-\delta_j) &= 0\,.
\end{aligned}$$

Using the conditions for optimal household and firm behavior ((3.3) and (3.6)), the decentralized equilibrium coincides with the social optimum:

$$\begin{aligned}
\mu_{jt} =& p_{jt} U_{Ct} \\
\lambda_t^L =& w_t U_{Ct} \\
\lambda_t^E =& \tau_t U_{Ct} \\
\lambda_{jt}^{K_j} =& r_{jt} U_{Ct} \\
\lambda_t^K =& \mu_t = \hat{\mu}_t = U_{Ct}\,.
\end{aligned}$$

From the conditions above, it is evident that the social optimum can be decentralized by a carbon tax which is uniform across all j sectors—which shows (2.15):

$$p_{jt}\frac{\partial Y_{jt}}{\partial E_{jt}} = \tau_{jt} = \lambda_t^E / U_{Ct}\,, \quad \forall j$$

and a capital income subsidy (or tax) Ξ_t^* which is chosen such that the social and private Euler equations coincide for each t, respectively:

$$U_{Ct} = \frac{1}{1+\zeta_S}U_{Ct+1}\left(1+R_{t+1}\right) \quad \text{and} \quad U_{Ct} = \frac{1}{1+\zeta}U_{Ct+1}\left(1+R_{t+1}[1-\Xi_{t+1}]\right)\,.$$

If $\zeta > \zeta_S$, the social optimum entails a subsidy on capital income, i.e. $\Xi_{t+1} < 0$, given by the following expression—which shows (2.16):

$$\Xi_{t+1}^* = \frac{\zeta_S - \zeta}{1+\zeta_S}\frac{(1+R_{t+1})}{R_{t+1}}\,.$$

2.9.3 Constrained-optimal policy problem

The constrained planner's problem is identical to the one in 2.9.2 with the no-subsidy constraint (2.17) ($\Xi_t \geq 0 \Longrightarrow U_{Ct}(1+\zeta)/U_{Ct+1} \leq 1 + R_{t+1}$). Using the firms' optimality conditions $\partial\hat{Y}_t/\partial K_{jt} = MPK_{jt} = r_{jt}$ and adding the private Euler equation $U_{Ct}(1+\zeta)/U_{Ct+1} = 1 + R_{t+1}$ with $R_{t+1} = MPK_{jt+1} - \delta_j$, we can write the social planner's problem as:

$$
\begin{aligned}
&\max_{\{K_{jt},\bar{K}_{jt+1},Y_{jt},C_t,\hat{Y}_t,L_{jt},E_{jt},I_t\}_{j=1}^{J}} \sum_{t=0}^{\infty}\left(\frac{1}{1+\zeta_S}\right)^t \Big[u(C_t) \\
&+\lambda_t^E\Big(\bar{E}_t-\sum_{j=1}^{J}E_{jt}\Big)+\lambda_t^L\Big(\bar{L}-\sum_{j=1}^{J}L_{jt}\Big)+\sum_{j=1}^{J}\lambda_{jt}^{K_j}\Big(\bar{K}_{jt}-K_{jt}\Big) \\
&+\lambda_t^K\Big(-\sum_{j=1}^{J}\bar{K}_{jt+1}+\sum_{j=1}^{J}(1-\delta_j)\bar{K}_{jt}+I_t\Big) \\
&+\sum_{j=1}^{J}\mu_{jt}\Big(Y_{jt}(L_{jt},K_{jt},E_{jt})-Y_{jt}\Big)+\hat{\mu}_t\Big(\hat{Y}_t(Y_{1t},\ldots,Y_{jt},\ldots,Y_{Jt})-\hat{Y}_t\Big) \\
&+\mu_t\Big(\hat{Y}_t-C_t-I_t\Big)+\hat{\phi}_{t+1}\Big(-\frac{U_{Ct}(1+\zeta)}{U_{Ct+1}}+(1+R_{t+1})\Big) \\
&+\sum_{k=1}^{J}\phi_{kt+1}^K\left(-R_{t+1}+MPK_{kt+1}-\delta_k\right)\Big].
\end{aligned}
\tag{2.23}
$$

$\hat{\phi}_{t+1}$ denotes the shadow costs of the no-subsidy constraint and ϕ_{kt+1}^K are the shadow costs of the constrained capital prices. Both constraints are present from period $t > 0$ onwards. MPK_{kt} is the marginal product of capital in sector k at time t and defined as: $MPK_{kt} = \gamma_k(1-\alpha_k)\theta_{kt}^K\hat{Y}_t K_{kt}^{-1}$, where γ_k is the value share of the sectoral output relative to aggregated output, $1-\alpha_k$ is the value share of the emissions-capital bundle in sectoral output and θ_{kt}^K is the value share of capital in the emissions-capital bundle.

The FOCs for $t > 0$ read:

$$
\begin{aligned}
C_t: &\quad U_{ct} - \mu_t - \hat{\phi}_{t+1} \frac{U_{cct}(1+\zeta)}{U_{ct+1}} + (1+\zeta_S)\hat{\phi}_t \frac{U_{ct-1}(1+\zeta)}{U_{ct}} \frac{U_{cct}}{U_{ct}} = 0 \\
\hat{Y}_t: &\quad \mu_t - \hat{\mu}_t + (1+\zeta_S) \sum_{k=1}^{J} \phi_{kt}^K MPK_{kt} \hat{Y}_t^{-1} = 0 \\
I_t: &\quad \lambda_t^K - \mu_t = 0 \\
\bar{K}_{jt+1}: &- \lambda_t^K + \frac{1}{1+\zeta_S} \lambda_{t+1}^K (1-\delta_j) + \frac{1}{1+\zeta_S} \lambda_{jt+1}^{K_j} = 0 \\
Y_{jt}: &\quad \hat{\mu}_t \frac{\partial \hat{Y}_t}{\partial Y_{jt}} - \mu_{jt} = 0 \\
K_{jt}: &\quad \mu_{jt} \frac{\partial Y_{jt}}{\partial K_{jt}} - \lambda_{jt}^{K_j} + (1+\zeta_S)\phi_{jt}^K \frac{\partial MPK_{jt}}{\partial K_{jt}} = 0
\end{aligned}
$$

$$
\begin{aligned}
L_{jt}: &\quad \mu_{jt} \frac{\partial Y_{jt}}{\partial L_{jt}} - \lambda_t^L = 0 \\
E_{jt}: &\quad \mu_{jt} \frac{\partial Y_{jt}}{\partial E_{jt}} - \lambda_t^E + (1+\zeta_S)\phi_{jt}^K \frac{\partial MPK_{jt}}{\partial E_{jt}} = 0\,.
\end{aligned}
$$

The Lagrangian multiplier $\lambda_{jt+1}^{K_j}$ is thus given by:

$$\lambda_{jt+1}^{K_j} = \lambda_t^K (1+\zeta_S) - \lambda_{t+1}^K (1-\delta_j)\,,$$

with:

$$\lambda_{jt}^K = \mu_t = \hat{\mu}_t - (1+\zeta_S) \sum_{k=1}^{J} \phi_{kt}^K MPK_{kt} \hat{Y}_t^{-1}\,.$$

Using the conditions for optimal household and firm behavior (3.3) and (3.6), respectively, the decentralized equilibrium coincides with the social optimum if:

$$\begin{aligned}
\mu_{jt} =& p_{jt}U_{Ct} \\
\lambda_t^L =& w_t U_{Ct} \\
\lambda_{jt}^{K_j} - (1+\zeta_S)\phi_{jt}^K \frac{\partial MPK_{jt}}{\partial K_{jt}} =& r_{jt}U_{Ct} \\
\lambda_t^E - (1+\zeta_S)\phi_{jt}^K \frac{\partial MPK_{jt}}{\partial E_{jt}} =& \tau_{jt}U_{Ct} \\
\hat{\mu}_t =& U_{ct} \, .
\end{aligned}$$

The constrained-optimal pricing rule for carbon emissions in sector j in the decentralized economy is thus given by:

$$\underbrace{\mu_{jt}}_{=U_{ct}p_{jt}} \frac{\partial Y_{jt}}{\partial E_{jt}} = \underbrace{\lambda_t^E}_{=U_{ct}\tau_t} - (1+\zeta_S)\phi_{jt}^K \frac{\partial MPK_{jt}}{\partial E_{jt}} \, , \tag{2.24}$$

which shows (2.18) in Proposition 1.

2.9.4 Constrained-optimal carbon pricing in the steady-state equilibrium

The steady-state equilibrium conditions are given by:

$$\begin{aligned}
\lambda^K &= U_c - (1+\zeta_S)\sum_{k=1}^{J} \phi_k^K MPK_k \hat{Y}^{-1} \\
\lambda_j^{K_j} &= \lambda^K(\zeta_S + \delta_j) \\
U_c r_j &= \lambda_j^{K_j} - (1+\zeta_S)\phi_j^K \frac{\partial MPK_j}{\partial K_j} \\
r_j &= \zeta + \delta_j \, ,
\end{aligned}$$

where MPK_k denotes the marginal product of capital in sector k. Using all FOCs with respect to K_j yields an expression for ϕ_j^K:

$$\phi_j^K = -U_c\left(\frac{\zeta-\zeta_S}{1+\zeta_S}\right)\frac{1+\sum_{k=1}^J(\delta_k-\delta_j)(\frac{\partial MPK_k}{\partial K_k})^{-1}MPK_k\hat{Y}^{-1}}{1+\sum_{k=1}^J(\zeta_S+\delta_k)(\frac{\partial MPK_k}{\partial K_k})^{-1}MPK_k\hat{Y}^{-1}}\left(\frac{\partial MPK_j}{\partial K_j}\right)^{-1}.$$

Using the steady-state analogue of (2.18) for efficient carbon prices under the constrained policy:

$$\underbrace{U_c\tau_j}_{=U_{ct}p_{jt}\partial Y_{jt}/\partial E_{jt}} = \underbrace{U_c\tau}_{=\lambda^E} - (1+\zeta_S)\phi_j^K\left(\frac{\partial MPK_j}{\partial E_j}\right),$$

we can write for all j:

$$\tau_j = \tau + (\zeta-\zeta_S)\left[\frac{1+\sum_{k=1}^J(\delta_k-\delta_j)(\frac{\partial MPK_k}{\partial K_k})^{-1}MPK_k\hat{Y}^{-1}}{1+\sum_{k=1}^J(\zeta_S+\delta_k)(\frac{\partial MPK_k}{\partial K_k})^{-1}MPK_k\hat{Y}^{-1}}\right]\frac{\left(\frac{\partial MPK_j}{\partial E_j}\right)}{\left(\frac{\partial MPK_j}{\partial K_j}\right)}. \quad (2.25)$$

The MPK and the marginal product of emission (MPE) are given, respectively, by:

$$MPK_j = \gamma_j(1-\alpha_j)\theta_j^K\hat{Y}K_j^{-1} \stackrel{!}{=} r_j, \qquad MPE_j = \gamma_j(1-\alpha_j)\theta_j^E\hat{Y}E_j^{-1} \stackrel{!}{=} \tau_j$$

where θ_j^K is the value share of capital within the capital-emissions bundle and $\theta_j^E = 1-\theta_j^K$ the value share of emissions. θ_j^K can be expressed in terms of technology parameters and equilibrium prices:

$$\theta_j^K = \frac{\beta_{Kj}H_{Kj}^{\rho_j}K_j^{\rho_j}}{\beta_{Kj}H_{Kj}^{\rho_j}K_j^{\rho_j}+\beta_{Ej}H_{Ej}^{\rho_j}E_j^{\rho_j}} = \frac{\beta_{Kj}H_{Kj}^{\rho_j}\left(\frac{r_j}{\beta_{Kj}H_{Kj}^{\rho_j}}\right)^{\frac{\rho_j}{\rho_j-1}}}{\beta_{Kj}H_{Kj}^{\rho_j}\left(\frac{r_j}{\beta_{Kj}H_{Kj}^{\rho_j}}\right)^{\frac{\rho_j}{\rho_j-1}}+\beta_{Ej}H_{Ej}^{\rho_j}\left(\frac{\tau_j}{\beta_{Ej}H_{Ej}^{\rho_j}}\right)^{\frac{\rho_j}{\rho_j-1}}}.$$

From this it follows that:

$$\frac{\partial MPK_j}{\partial K_j} = [(1-\theta_j^K)\rho_j - 1]K_j^{-1}MPK_j$$

$$\frac{\partial MPK_j}{\partial E_j} = -\rho_j \theta_j^K K_j^{-1} MPE_j \,,$$

and we can re-write (2.25) to obtain the constrained-optimal sectoral emissions pricing rule (2.19) as:

$$\tau_j = \tau + (\zeta - \zeta_S)\left(\frac{1 + \sum_{k=1}^{J}(\delta_k - \delta_j)\Psi_k}{1 + \sum_{k=1}^{J}(\zeta_S + \delta_k)\Psi_k}\right)\left(\frac{-\rho_j \theta_j^K \tau_j}{((1-\theta_j^K)\rho_j - 1)r_j}\right) \,, \quad \forall\, j$$

where $\Psi_k = [(1-\theta_k^K)\rho_k - 1]^{-1} K_k Y^{-1}$ and $K_k Y^{-1} = \frac{\gamma_k \alpha_k \theta_k^K}{\zeta + \delta_k}$.

To prove Proposition 3, we proceed in two parts. Note first that $\bar{\phi}_j \neq \bar{\phi}_k$ if either $\delta_j \neq \delta_k$, $\beta_{Kj} \neq \beta_{Kk}$, $H_{Kj} \neq H_{Kk}$, $H_{Ej} \neq H_{Ek}$, or $\rho_j \neq \rho_k$. It is then straightforward to see that technology heterogeneity implies that $\Gamma_j \neq \Gamma_k$, as the denominators take on different values depending on the sector-specific technology parameters, and $\tau_j \neq \tau_k$. The second part of the proposition follows readily from inspecting the expressions for $\bar{\phi}_j$ and ρ_j and θ_j^K in (2.19). First, τ_j increases with ρ_j because $\partial(\rho_j \theta_j^K)/\partial \rho_j > 0$ and thus $\tau_j > \tau_k$ whenever $\rho_j > \rho_k$, ceteris paribus. Second, $\beta_{Kj}, H_{Kj}, H_{Ej}$ and δ_j impact on sectoral carbon prices depends on the elasticity parameter ρ_j. Whenever $\rho_j = \rho_k > 0$, a higher capital value share ($\theta_j^K > \theta_k^K$) leads to higher carbon prices ($\tau_j > \tau_k$). We thus investigate how θ_j^K changes with the respective parameters:

- $\left.\frac{\partial \theta_j^K}{\partial \beta_{Kj}}\right|_{\rho_j > 0} > 0$ and τ_j increases with β_{Kj} when $\rho_j > 0$.

- $\left.\frac{\partial \theta_j^K}{\partial H_{Kj}}\right|_{\rho_j > 0} > 0$ and τ_j increases with H_{Kj} when $\rho_j > 0$.

- $\left.\frac{\partial \theta_j^K}{\partial H_{Ej}}\right|_{\rho_j > 0} < 0$ and τ_j decreases with H_{Ej} when $\rho_j > 0$.

- $\left.\frac{\partial \theta_j^K}{\partial \delta_j}\right|_{\rho_j > 0} < 0$ and τ_j decreases with δ_j when $\rho_j > 0$.

Whenever $\rho_j = \rho_k < 0$, a higher capital value share ($\theta_j^K > \theta_k^K$) leads to lower carbon prices ($\tau_j < \tau_k$). We thus investigate how θ_j^K changes with the respective parameters:

- $\left.\frac{\partial \theta_j^K}{\partial \beta_{Kj}}\right|_{\rho_j<0} < 0$ and τ_j decreases with β_{Kj} when $\rho_j < 0$.

- $\left.\frac{\partial \theta_j^K}{\partial H_{Kj}}\right|_{\rho_j<0} < 0$ and τ_j decreases with H_{Kj} when $\rho_j < 0$.

- $\left.\frac{\partial \theta_j^K}{\partial H_{Ej}}\right|_{\rho_j<0} > 0$ and τ_j increases with H_{Ej} when $\rho_j < 0$.

- $\left.\frac{\partial \theta_j^K}{\partial \delta_j}\right|_{\rho_j<0} > 0$ and τ_j increases with δ_j when $\rho_j < 0$.

Proposition 3 summarizes these findings.

2.9.5 Steady-state conditions

The household's Euler equation with $\Xi = 0$ and capital investments in steady-state reveal

$$r_j = \zeta + \delta_j\,, \quad I = \sum_{j=1}^{J} \delta_j K_j\,.$$

The final good sector's FOCs for all j are

$$\gamma_j \frac{\prod_{k=1}^{J} Y_k^{\gamma_k}}{Y_j} = p_j\,.$$

The optimality conditions for sectoral output are

$$w = \alpha_j p_j \frac{Y_j}{L_j}, \quad r_j = p_j \frac{\partial Y_j}{\partial K_j}, \quad \tau_j = p_j \frac{\partial Y_j}{\partial E_j}\,,$$

where

$$\frac{\partial Y_j}{\partial K_j} = (1-\alpha_j)\beta_{Kj}(H_{Kj}K_j)^{\rho_j}/K_j \times L_j^{\alpha_j}\left[\beta_{Kj}(H_{Kj}K_j)^{\rho_j} + \beta_{Ej}(H_{Ej}E_j)^{\rho_j}\right]^{\frac{1-\alpha_j}{\rho_j}-1}$$
$$\frac{\partial Y_j}{\partial E_j} = (1-\alpha_j)\beta_{Ej}(H_{Ej}E_j)^{\rho_j}/E_j \times L_j^{\alpha_j}\left[\beta_{Kj}(H_{Kj}K_j)^{\rho_j} + \beta_{Ej}(H_{Ej}E_j)^{\rho_j}\right]^{\frac{1-\alpha_j}{\rho_j}-1}.$$

Labor supply is given by $\sum_j L_j = \bar{L}$ and total emission is given by $\sum_j E_j = \bar{E}$. Final output and sectoral output are given by

$$\hat{Y} = \prod_{j=1}^{J} Y_j^{\gamma_j}, \quad Y_j = L_j^{\alpha_j}\left[\beta_{Kj}(H_{Kj}K_j)^{\rho_j} + \beta_{Ej}(H_{Ej}E_j)^{\rho_j}\right]^{\frac{1-\alpha_j}{\rho_j}}.$$

Consumption is thus $C = \hat{Y} - I$.

2.10 Appendix B: Mappings of sectors and emissions data to model sectors

We map the sectors in the WIOD dataset, as identified by alphabetic categories according to the *International Standard Industrial Classification of All Economic Activities (ISIC), Revision 4*, to the four sectoral aggregates in our model. The many-to-one sectoral mapping is as follows: sectors B, D, E are aggregated in *"Electricity"*, sectors C, F in *"Industry"*, sectors H in *"Transportation"*, and sectors G, I-U in *"Services"*.

We use information on CO_2 emissions at the sectoral level from the European Commission's EDGAR database (*Emission Database for Global Atmospheric Research*, Crippa et al. 2019) and link emissions from fossil-fuel combustion as well as process emissions to the sectors using the mapping shown in Table 3.5.

Table 2.3. Mapping of emissions data categories to model sectors.

Category in the data	Model sector
Biological Treatment of Solid Waste	Electricity
Emissions from biomass burning	Electricity
Incineration and Open Burning of Waste	Electricity
Main Activity Electricity and Heat Production	Electricity
Oil and Natural Gas	Electricity
Petroleum Refining - Manufacture of Solid Fuels and Other Energy Industries	Electricity
Solid Fuels	Electricity
Solid Waste Disposal	Electricity
Wastewater Treatment and Discharge	Electricity
Chemical Industry	Industry
Manufacturing Industries and Construction	Industry
Metal Industry	Industry
Cement Production	Industry
Lime Production	Industry
Glass Production	Industry
Other Process Uses of Carbonates	Industry
Non-Energy Products from Fuels and Solvent Use	Industry
Liming	Industry
Urea application	Industry
Civil Aviation	Transportation
Other Transportation	Transportation
Railways	Transportation
Road Transportation	Transportation
Water-borne Navigation	Transportation
Non-Specified	Services
Other Sectors	Services

Chapter 3

Should carbon emissions really be priced uniformly? The case of capital income tax distortions

Abstract

How should carbon emissions be taxed when a capital income tax finances public consumption? When accounting for capital income taxes, this paper argues that carbon emissions should be priced non-uniformly across sectors which runs counter to the general wisdom in the field to price carbon uniformly. This paper shows theoretically and numerically how optimally differentiated carbon prices address the under-accumulation of capital arising from the income tax distortion. Sectoral carbon price differentiation is substantial and welfare gains are significant, relative to a uniform price. The degree of sectoral carbon price differentiation is driven by heterogeneous sectoral production technologies, i.e. how differently firms' substitute carbon emissions with other input goods.

3.1 Introduction

Governments have various duties—for instance they raise tax revenues that finance public consumption and they implement carbon mitigation policies to combat climate change. Many studies in the environmental economics literature, however, abstract from the need to raise tax revenues and address exclusively the externality from carbon emissions (Nordhaus 2007, 2017, Golosov et al. 2014). A common insight from these studies is that a uniform carbon price denotes the most cost-efficient carbon abatement instrument—resulting in the call for a global carbon price as in the Kyoto Protocol, scope expansion of emissions trading schemes in Europe, the United States and China, and linkage of different carbon pricing regimes (Nordhaus 2015). The presence of fiscal policies in the real-world alters, however, the theoretical optimal pricing scheme because environmental and fiscal policies interact. For instance, previous studies show how distortionary taxes change the level of an uniform carbon price in a dynamic economy (Bovenberg and Goulder 1996, Barrage 2020), or how distortionary taxes impact optimal non-uniform carbon prices in a static economy (Landis et al. 2018). But how does a fiscal policy instrument that finances public consumption—for instance a capital income tax—impact the optimal carbon pricing scheme? Surprisingly, the literature has so far overseen to ask if carbon prices should really be uniform when capital income taxes (partly) finance public consumption.

In this paper, I examine how the economic rationale for a uniform carbon price is altered by the presence of distortionary capital income taxes that finance public consumption. By this, I acknowledge that real world taxation schemes diverge from theoretical recommendations, i.e. that governments raise revenues with capital income taxes although tax theory suggest not to (Judd 1985, Chamley 1986, Chari et al. 2020). I argue that the regulator cannot dissolve the capital income tax distortions because fiscal and environmental policy decisions are made independently (Barrage 2018). My finding that carbon emissions should then be priced non-uniformly runs counter to the established policy recommendation to price carbon emissions with a single price. The main intuition for the result is that a capital income tax is ”maximal” distortive and leads to capital under-accumulation. Optimal, non-uniform carbon prices spur, however, capital demand (and capital accumulation) and are hence desirable from a social welfare perspective. The degree of carbon price differentiation depends on the sectoral technological heterogeneity that governs how *differently* firms substitute between emissions and capital and between labor and the capital-emissions bundle. Also, a higher capital income tax rate induces a greater tax distortion, and carbon price differentiation increases.

Before summarizing my results in greater detail, it is useful to introduce my setting. I build the analysis around the Ramsey growth model. I analyze a multi-sector economy which produces with capital, labor and carbon emissions (CO_2). I combine sectoral capital and sectoral emissions in one bundle which then produces, together with sectoral labor input, sectoral output. I impose an exogenous, economy-wide carbon emissions budget in every period. The government needs to finance an exogenous sequence of public consumption with tax revenues from carbon emissions and capital income. I impose a time-invariant capital income tax rate because fiscal policies are independent of environmental policies. The government operates a balanced budget in every period, I thereby abstract from government debt. Tax revenues from capital income and carbon emission—that are not used to finance public consumption—are returned to the household. The regulator maximizes social welfare using sectoral carbon prices subject to equilibrium conditions of the de-centralized economy. I explore theoretically how the presence of the capital income tax distortion, and given technology heterogeneity, determines the degree of carbon price differentiation. I calibrate the model to data of the EU-28 economy to explore numerically the role of technology heterogeneity for non-uniform sectoral carbon pricing and the welfare implications, relative to a uniform carbon pricing scheme.

Given a capital income tax distortion, I obtain the following results. First, uniform carbon prices are only optimal when all sectors produce with identical technologies—certainly an unrealistic assumption. Second, and given sectoral technology heterogeneity, optimal carbon prices are non-uniform. The intuition is that a capital income tax yields an under-accumulation of capital, so non-uniform carbon prices incentivise the household savings to push capital accumulation upwards. This improves intertemporal efficiency, i.e. firms produce with more resources and households consume more in the longrun.

The degree of carbon price differentiation is driven by the sectoral technology heterogeneity, in particular by (1) the substitutability between capital and emissions and by (2) the substitutability between labor and the capital-emissions bundle. I find (1) that sectors that do not substitute well between CO_2 and capital receive a lower carbon price than sectors which substitute well between CO_2 and capital. Decreasing the carbon price in sectors that combine emissions and capital as "complements" in production triggers the sectoral capital use which is greater than the capital decrease in sectors with higher carbon prices (that have a better substitutability). Overall, the total capital stock and welfare increases, relative to a uniform carbon price. I find (2) that optimal carbon prices also consider the sectoral substitutability between labor and the capital-emissions bundle. The general idea is to re-allocate the sectoral use of labor to sectors that combine labor and the capital-emissions bundle as "complements" in production, i.e. that these "more complement" sectors use more capital when given more sectoral labor inputs. Accordingly, sectors that substitute well between labor and the capital-emissions bundle receive a lower carbon price to incentivise these sectors to substitute *away* from labor towards the capital-emissions bundle. The "now free" labor is taken up by the "more complement" sectors which triggers their sectoral capital use. Capital use increases, firms produce with more resources and the household is better-off, relative to a uniform carbon price.

An important third insight is that price deviations from uniform are numerically significant: when I build the numerical analysis on realistic, empirically estimated values on the sector-specific substitution elasticities, I find that the steady-state optimal prices for carbon emissions differ on average by 11.07%. The optimal carbon prices increase the longrun capital stock by 0.26% and increase the lifetime wellbeing of households by 4.3%, relative to today under a uniform carbon price. I find that heterogeneity in the sectoral substitution elasticities between capital and emissions drives the degree of carbon price differentiation because these parameters govern the sectoral emission's direct impact on capital accumulation. The other elasticities on the production side (on the final good level or between labor and the capital-emissions bundle) have less impact on the capital use and play thus less of a role. Also, the degree of technology heterogeneity matters a lot. I deviate from the empirical estimates of sector-specific substitutabilities and find that the degree of carbon price differentiation increases to 18% when assuming more extreme substitution elasticities. In contrast, almost eliminating sectoral heterogeneity decreases the coefficient of variation in carbon prices down to 0.6%. Both sets represent rather unrealistic cases of sectoral heterogeneity but show that from a conceptual perspective, optimal carbon price differentiation may take extreme values or may be approximately uniform, depending on the sectoral technology heterogeneity. From a practical perspective—with reasonable elasticity values—the degree of carbon price differentiation is, however, substantial and leads to significant welfare gains relative to a uniform price.

The paper's findings have important implications for real-world policymaking. I show that the general wisdom to price carbon emissions uniformly across all sectors does no longer hold if the government diverges from the theoretical recommendations not to tax capital income. It is then not sufficient to create a single carbon market, as for instance envisioned in the EU's longrun environmental policy design. Instead, firms covered by the European Emissions Trading Scheme should also pay (or receive) an additional tax (or subsidy) on their carbon emissions. In this paper, I show how these additional sectoral taxes or subsidies depend on the sectoral substitution possibilities between "dirty" emissions and "clean" capital.

RELATED LITERATURE.—My paper contributes to a large body of literature that evaluates optimal carbon pricing policies. The first strand of literature builds on general equilibrium growth models with capital accumulation that abstract from distortionary taxes (van der Ploeg and Withagen 1991, 2014, Bovenberg and Smulders 1996, Hassler and Krusell 2012, Golosov et al. 2014, Nordhaus 2007). In these frameworks—without further (tax) distortions—a Pigouvian tax that equalizes the social costs of carbon emissions with the emission's benefits is optimal (See Pigou (1932)).

Optimal carbon prices deviate, however, from the Pigouvian principle when considering distortionary taxes (Sandmo 1975, Bovenberg and de Mooij 1994, 1997, Bovenberg and Goulder 1996, Bovenberg and van der Ploeg 1994, Parry et al. 1999, Schwartz and Repetto 2000, Kaplow 2012, Barrage 2020). The general idea is that carbon pricing generates tax revenues but may also decrease the revenues generated by other taxes because the carbon tax decreases the respective tax base. A carbon price equal to the Pigouvian tax rate may then not be optimal. For instance, Bovenberg and Goulder (1996) advocate to price carbon emissions below the Pigouvian rates because the carbon tax's fiscal interactions likely exceeds their non-environmental revenue benefits. Barrage (2020)'s integrated assessment framework adds an additional motive, namely that tax distortions decrease the size of the economy and hence the value of marginal damages. She finds that the optimal carbon carbon price decreases by -3% to -36% due to tax distortions.

These papers exclusively focus on a uniform carbon pricing scheme, however, non-uniform carbon prices are optimal in multi-sector general equilibrium economies when distortionary taxes are present—see for instance Landis et al. (2018) or Boeters (2014). The rational for non-uniform carbon prices is the presence of intermediate input taxes on refined fuels. These taxes are economically indistinguishable from carbon prices, so sectoral carbon prices aim at compensating any differences in the intermediate input taxes.

The literature has identified other motives for non-uniform carbon prices. For instance non-uniform carbon prices are optimal when markets are not perfectly competitive (Markusen 1975, Krutilla 1991, Rauscher 1994), when carbon leakage cannot be addressed with border adjustment tariffs (Hoel 1996, Böhringer et al. 2014), or when social equity concerns are at play (Abrell et al. 2018, Kalsbach and Rausch 2021).

With my paper, I aim to provide another motive for non-uniform carbon prices, namely the presence of capital income taxes to finance public consumption. I contribute to the literature in multiple ways. I am the first to show theoretically the optimality of non-uniform carbon prices in a *dynamic* setting with capital income taxes. Second, I provide an intuition for the direction of the sectoral price deviation as governed by the sectoral technology heterogeneity. Third, I show the relevance of my finding by providing numerical estimates on the degree of price differentiation and the welfare implication, relative to a uniform carbon price.

The paper is organized as follows. I introduce the competitive economy in Section 3.2. In Section 3.3, I derive the optimal carbon prices in a setting where the capital income tax distortion cannot be dissolved. I show then in Section 3.4 that optimal carbon prices are non-uniform. Section 3.5 calibrates the model to the EU-28 economy. I investigate the numerical relevance of non-uniform carbon prices in Section 3.6 and perform sensitivity analysis in Section 3.7. Section 3.8 concludes.

3.2 The competitive economy

I extend the standard Ramsey-model in three ways: (i) I impose a multi-sector economy, (ii) I consider carbon emissions as an input to sectoral production, and (iii) I introduce a capital income tax that finance public consumption.

HOUSEHOLD.—In every period $t = \{0, ..., \infty\}$, the household supplies labor $\bar{L}$ inelastically at the market rate w_t, the household supplies capital $\bar{K}_t$ at the market rate r_t and the household pays a time-invariant capital income tax $\bar{\tau}^K$ on capital holdings. The after-tax capital return is $R_t = r_t(1 - \bar{\tau}^K)$. Further income sources are profits from sectoral production Π_t and government lump-sum transfers Λ_t^G. The household's budget constraint is:

$$C_t + \bar{K}_{t+1} = w_t \bar{L} + R_t \bar{K}_t + (1 - \delta)\bar{K}_t + \Pi_t + \Lambda_t^G \,, \tag{3.1}$$

where δ denotes the capital depreciation rate. The household's objective is to maximize the discounted sum of household's utility $u(C_t)$:

$$U = \sum_{t=0}^{\infty} \left(\frac{1}{1+\zeta}\right)^t u(C_t) \,, \tag{3.2}$$

where ζ denotes the household's discount rate and C_t is consumption. Utility is of CRRA-type: $u(C_t) = C_t^{1-\sigma}/(1-\sigma)$ where $1/\sigma$ is the elasticity of intertemporal substitution. Maximizing (3.2) subject to (3.1) yields the intertemporal optimality condition:

$$U_{Ct} = \frac{1}{1+\zeta} U_{Ct+1} \left(1 + R_{t+1} - \delta\right) \,, \tag{3.3}$$

with $U_{Ct} := \partial u(C_t)/\partial C_t$.

PRODUCTION.—The final good, $\hat{Y}_t$, is produced with sectoral output Y_{jt} where $j, i, k = 1, \ldots, J$:

$$\hat{Y}_t = \hat{A} \left(\sum_{j=1}^{J} \beta_{Y_j} Y_{jt}^{\hat{\rho}} \right)^{\frac{1}{\hat{\rho}}} \,. \tag{3.4}$$

$\hat{A}$ denotes the total factor productivity. β_{Y_j} are share parameters with $\sum_j \beta_{Y_j} = 1$. The elasticity parameter $\hat{\rho}$ relates to the elasticity of substitution according to $\hat{\sigma} = 1/(1-\hat{\rho})$. Inputs in production are substitutes when $0 < \hat{\rho} < 1$ and complements when $-\infty < \hat{\rho} < 0$. $\hat{\rho} = 0$ indicates the Cobb-Douglas case.

Sectoral output Y_{jt} is produced with sectoral labor L_{jt}, capital K_{jt}, and CO_2 emissions E_{jt} as inputs to a nested constant-elasticity-of-substitution (CES) production function:

$$Y_{jt} = A_j \left(\beta_{L_j} \left(L_{jt}\right)^{\rho_j} + \beta_{KE_j} \left(\underbrace{\left[\beta_{E_j} E_{jt}^{\rho_j^{KE}} + \beta_{K_j} K_{jt}^{\rho_j^{KE}} \right]^{\frac{1}{\rho_j^{KE}}}}_{:=Y_{jt}^{KE}} \right)^{\rho_j} \right)^{\frac{1}{\rho_j}}. \tag{3.5}$$

I do not explicitly include energy from fossil fuels but instead represent carbon emissions (from burning fossil fuels) as an input in production. A_j denotes the sectoral productivity factor. $\beta_{L_j} + \beta_{KE_j} = 1$ and $\beta_{E_j} + \beta_{K_j} = 1$ are share parameters. Sectoral production requires a capital-emissions bundle, Y_{jt}^{KE}, on the lower level. Together with labor input, the capital-emissions bundle is used for the production of sectoral output Y_{jt}. ρ_j^{KE} and ρ_j are the elasticity parameter between capital and emissions and the elasticity parameter between labor and the capital-emission-bundle in sector j, respectively.

I focus specifically on multi-sector economies with technology heterogeneity.

Definition 1 *(**Heterogeneous sectoral technologies**) Sector j and k, with $j \neq k$, are heterogeneous if j and k*

- *have different elasticities of substitution on either level ($\rho_j \neq \rho_k$ or $\rho_j^{KE} \neq \rho_k^{KE}$), or*
- *if the elasticities of substitution coincide but are not Cobb-Douglas ($\rho_j = \rho_k \neq 0$ and $\rho_j^{KE} = \rho_k^{KE} \neq 0$) and the share parameters are different ($\beta_{L_j} \neq \beta_{L_k}$ or $\beta_{E_j} \neq \beta_{E_k}$).*

Otherwise sector j and k are identical.

Firms take prices and taxes given. The FOCs for K_{jt}, L_{jt}, E_{jt}, and Y_{jt} read:

$$r_t = p_{jt} \frac{\partial Y_{jt}}{\partial K_{jt}}, \quad w_t = p_{jt} \frac{\partial Y_{jt}}{\partial L_{jt}}, \quad p_t^E (1 + \tau_{jt}^{E_j}) = p_{jt} \frac{\partial Y_{jt}}{\partial E_{jt}}, \quad p_{jt} = \frac{\partial \hat{Y}_t}{\partial Y_{jt}}, \tag{3.6}$$

where p_{jt} and $p_t^E(1 + \tau_{jt}^{E_j})$ denote the prices for sectoral output Y_{jt} and sectoral emissions E_{jt}, respectively. The price of final output is the numeraire.

Market clearing conditions of labor and capital are:

$$\sum_{j=1}^{J} L_{jt} = \bar{L}, \quad \sum_{j=1}^{J} K_{jt} = \bar{K}_t\,, \tag{3.7}$$

and the law of capital accumulation is given by:

$$\bar{K}_{t+1} = (1-\delta)\bar{K}_t + I_t\,, \tag{3.8}$$

where I_t denotes investments.

GOVERNMENT.— The government finances public consumption G with capital income tax revenues, $\bar{\tau}^K r_t \bar{K}_t$, and carbon emissions tax revenues, $\sum_j p_t^E(1+\tau_{jt}^{E_j})E_{jt}$. p_t^E denotes the corresponding clearing price of the total emissions supply. A set of sector-specific taxes and subsidies on sectoral carbon emission, $\{\tau_{jt}^{E_j}\}_{j=1}^J$, implements a (potentially) non-uniform carbon pricing scheme. The government's budget constraint is:

$$G + \Lambda_t^G = \bar{\tau}^K r_t \bar{K}_t + \sum_{j=1}^{J} p_t^E(1+\tau_{jt}^{E_j})E_{jt}\,, \tag{3.9}$$

with $\Lambda_t^G \geq 0$ to rule out lump-sum taxes that finance public consumption. Throughout, I assume that tax revenues are sufficiently large to finance G.

CLOSED ECONOMY.— The household, firms and the government live in a closed-economy in which the final good is used for investments, and private and public consumption:

$$I_t + C_t + G_t = \hat{Y}_t\,. \tag{3.10}$$

CARBON EMISSIONS.—I impose an economy-wide, exogenous emissions budget in every period, $\bar{E}_t$, according to

$$\sum_{j=1}^{J} E_{jt} = \bar{E}_t \,. \tag{3.11}$$

EQUILIBRIUM— The competitive equilibrium is defined as:

Definition 2 *Given prices* $\{w_t, r_t, p_t^E, \{p_{jt}\}_{j=1}^J\}$ *and taxes* $\{\bar{\tau}^K,\ \{\tau_{jt}^{E_j}\}_{j=1}^J\}$*, the allocations* $\{C_t, \hat{Y}_t, I_t,\ \bar{L}, \bar{K}_t, \bar{E}_t, \{L_{jt}, K_{jt}, E_{jt}\}_{j=1}^J\}$ *yield the competitive equilibrium, where it holds that*

(i) the allocations solve the household's and firms' problem with given prices and taxes,

(ii) public consumption G_t *is financed with tax revenues from capital income and carbon emission,*

(iii) tax revenues beyond the desired public consumption level are returned lump-sum to the household, and

(iv) all markets clear.

3.3 Carbon pricing with capital income tax distortions

3.3.1 How do capital income taxes distort the economy?

Capital income should not be taxed to finance public consumption (Judd 1985, Chamley 1986, Chari et al. 2020). One intuition for this insight is that goods ought to be taxed according to their price elasticities. Capital supply is, however, very elastic in the Ramsey-framework (even infinitely elastic in the longrun) and a capital income tax would thus induce the "strongest" possible tax distortion. Instead, public consumption should be financed with taxes on private consumption or labor because these instrument induces weaker responses, i.e. smaller distortions. In reality, however, governments do tax capital income with the outcome that households save too little, i.e. under-accumulate capital from the normative perspective of a Ramsey economy. But what drives this outcome? A capital income tax increases the capital market clearing price, and capital (as an input to production) becomes more expensive. Firms demand less capital and the household under-accumulates capital. Policies that increase the aggregated capital stock are then socially desirable: benefits from (environmental) policies stem from averted climate change damages through carbon emissions abatement *and* from the policies' impact on the household's accumulation of capital.

3.3.2 Can environmental policy stimulate capital accumulation?

The intertemporal optimality conditions governs the household's saving behavior: if future returns from capital exceed the steady-state level, i.e. $r_{t+1} > (\zeta+\delta)/(1-\tau^{\bar{K}})$, the household wants to substitute consumption today for consumption tomorrow. In other words, the household wants to invest more resources in tomorrow's capital stock (which has been too low due to the capital income tax and when given a uniform price on carbon emissions). The regulator can increase tomorrow's return on capital investments for the household, r_{t+1}, indirectly through tomorrow's environmental policy because the firms' marginal product of capital depends on the sectoral allocation of carbon emissions. With an adequate sectoral allocation of carbon emission, the economy converges then to an intertemporal more efficient steady-state equilibrium with more capital investments. The remainder of this paper thus discusses what the presence of a capital income tax imply for environmental policies, i.e. how carbon prices impact the firms' capital demand which in return stimulates the household's capital accumulation.

3.3.3 The regulator's problem

I have in mind a regulator who uses sectoral carbon prices to maximize the discounted sum of the household's utility (3.2) subject to the equilibrium conditions of the decentralized economy. These conditions are given by the household's intertemporal optimality conditions (3.3), the production technologies (3.4 and 3.5), the firms' optimal input choices for production (3.6), the resource constraints for labor and capital (3.7), the law of capital accumulation (3.8), the government's budget constraint (3.9), the household's budget constraint—or the closed-economy condition—(3.10), the economy-wide periodic emissions constraint (3.11), and given the initial capital stock $\bar{K}_0$. I impose an exogenously given capital income tax $\bar{\tau}^K$ and that revenues from taxing capital income and carbon emissions are sufficiently large to finance public consumption ($\Lambda_t^G > 0$). Mathematically, I formulate the problem in primal form, so the regulator can be thought of as using quantities (and not prices) as control (See Chari and Kehoe (1999)). Comparing the solution of the centralized problems to the solution of the de-centralized problem yields Proposition 4.

Proposition 4 *The optimal allocation of emissions across sectors equalizes marginal cost to marginal benefits from sectoral carbon emissions.*

PROOF: I derive in Appendix 3.9.1 the complete solution to the regulator's problem. The first-order condition with respect to E_{jt} reads (3.12):

$$\lambda_t^{\bar{E}} = \underbrace{\hat{\mu}_t \frac{\partial \hat{Y}_t}{\partial E_{jt}}}_{\substack{\text{marginal benefit}\\ \text{from production}}} + \underbrace{\sum_{i=1}^{J} \mu_t^{K_i} (1-\bar{\tau}^K) \frac{\partial MPK_{it}}{\partial E_{jt}}}_{\substack{=E_{jt}\text{'s impact on the}\\ \text{aggregated capital demand}}} + \underbrace{\sum_{i=1}^{J} \mu_t^{L_i} \left(-\frac{\partial w_t}{\partial E_{jt}} + \frac{\partial MPL_{it}}{\partial E_{jt}} \right)}_{\substack{=E_{jt}\text{'s impact on the}\\ \text{sectoral allocation of labor}}}, \forall j \,. \tag{3.12}$$

□

The LHS in (3.12) shows the Lagrangian multiplier, $\lambda_t^{\bar{E}}$, from the economy-wide emissions constraint (3.11) and captures the policy induced scarcity of CO_2.

The RHS in (3.12) gathers the benefits from sectoral carbon emissions. Sectoral carbon emissions from fossil fuel are an input to production and yield the utility-weighted marginal benefit from emission, $\hat{\mu}_t \frac{\partial \hat{Y}_t}{\partial E_{jt}}$, as captured by the first term on RHS. $\mu_t^{K_i}$ are the Lagrangian multipliers from firm i's optimal capital use:

$R_t = (1 - \bar{\tau}^K)MPK_{it}$, where MPK_{it} denotes the marginal product of capital in sector i. More emissions in sector j increases the firm j's marginal product of capital and thus the firm j's capital demand (while it (may) also decreases the other j sectors' capital demand) which spurs capital accumulation. The second term on the RHS thus denotes the benefits from carbon emissions as a vehicle to address the under-accumulation of capital. $\mu_t^{L_j}$ are the Lagrangian multipliers from firm j's optimal labor use: $w_t = MPL_{jt}$, where MPL_{jt} denotes the marginal product of labor in sector j. The intuition is that sectoral carbon emissions also impact the allocation of labor across sectors and thus indirectly change the firms' demand for capital. The third terms on the RHS thus adjusts E_{jt}'s benefits for indirect effects on capital accumulation through the labor allocation channel.

It is important to say that the first-best is a situation where capital is not taxed and government revenues are collected using a non-distortionary instruments (e.g., a lump-sum tax). If the capital income tax cannot be set to zero, the regulator would prefer to address the capital income tax distortion directly by subsidizing the firm's capital input (or by subsidizing the firms' output and taxing labor input) which I rule, however, out by restricting the regulator's instrument set to carbon prices only. The intuition is that the regulator prefers fiscal instruments—and not environmental instruments—to offset the capital income tax distortion on the household's side. If an optimal fiscal policy was available, capital would then accumulate like the "socially optimal" path as in a setting without tax distortion. The regulator would also implement a uniform carbon pricing scheme to guarantee efficient carbon emissions reduction—the standard result in the literature. I derive the first-best instruments (dropping the no lump-sum tax condition) in Appendix 3.9.2.

These settings require, however, rather stark assumptions which contrast the common policy practice, namely that climate policy is usually pursued separately from fiscal policy and is not directly linked to policy decisions on output, labor input or capital input taxes.[1] I thus focus on the more realistic setup in which the available instrument set comprises only carbon prices, and I investigate next the underlying mechanism that drives the optimality of non-uniform carbon prices through spurring the capital accumulation.

3.4 When are optimal carbon prices (non-)uniform?

From (3.12) follows that optimal carbon prices depend on the sectoral substitutability (complementarity) between capital and labor with emissions inputs, i.e. they depend on the terms $\partial\left(MPK_{it}\right)/\partial E_{jt}$ and $\partial\left(-w_t + MPL_{it}\right)/\partial E_{jt}$. These terms vary across j when sectors produce with heterogeneous technologies. Closed-form expressions for MPK_{it}, MPL_{it} and w_t are given in Appendix 3.9.6. I first investigate in Section 3.4.1 how carbon prices depend on the sectoral technology heterogeneity. Second, I investigate in Section 3.4.2 how the capital income tax rate impacts the sectoral carbon prices.

3.4.1 The sectoral heterogeneities' role for (non-)uniform carbon prices

Proposition 5 *Given a positive capital income tax, optimal carbon prices are uniform if sectors produce with identical technologies.*

PROOF: See Appendix 3.9.3. □

In the presence of a capital income tax and identical production technologies, all sectors substitute identically between input goods in response to a higher sectoral carbon price. Carbon prices cannot stimulate the total firms' capital demand and

[1] In particular, subsidizing the firms' used capital quantities translates into subsidizing corporate profits in the real-world, which is, of course, not possible. Also, governments do not subsidize output but impose positive VATs (value-added taxes). I thus argue that a more relevant perspective for carbon pricing policy (in a setting with capital income taxes to finance public consumption) is to assume that subsidies on corporate profits and final output, and optimal taxes on labor input are not available to a policymaker when designing environmental policies. Instead, I focus on optimal carbon prices when decisions on fiscal and environmental polices are made separably from each other.

deviating from a uniform carbon price is not optimal. I argue, however, that this representation of the firms' production technology clearly oversimplifies the production side.[2]

Proposition 6 *Given a positive capital income tax, optimal carbon prices are non-uniform if sectors produce with heterogeneous technologies. Ceteris paribus, sector* j *receives a higher carbon price than sector* k $(p_t^E(1+\tau_{jt}^{E_j}) > p_t^E(1+\tau_{kt}^{E_k}))$,

(i) if sector j *substitutes more easily between capital and emissions than sector* k $(\rho_j^{KE} > \rho_k^{KE})$,

(ii) or if sector j *substitutes less easily between labor and the capital-emissions bundle than sector* k $(\rho_j < \rho_k)$.

PROOF: See Appendix 3.9.3. □

If sectors produce with *heterogeneous* production technologies, optimal carbon prices are non-uniform. The general intuition is that optimally differentiated carbon prices aim to steer the capital accumulation upwards which has been too low due to the capital income tax. Increasing the capital stock yields intertemporal efficiency improvements which benefits the economy and leads to welfare gains. Proposition 6 highlights that two heterogeneity aspects are particularly relevant for optimal price differentiation to spur capital accumulation. I investigate the role of the substitutability between capital and emissions and the role of the substitutability between labor and the capital-emissions bundle for carbon pricing.

SUBSTITUTABILITY BETWEEN CAPITAL AND EMISSIONS.— Proposition 6 (i) shows that sectors which substitute easily between capital and emissions receive a higher carbon price that sectors that combine capital and emissions as complements in production receive a lower carbon price. The main intuition is that a more complementary sectoral technology stimulates a "relatively" greater sectoral

[2] The environmental economics literature often (implicitly) assumes identical technologies when using a single-sector economy framework (Golosov et al. 2014, Nordhaus 2007, Barrage 2018, 2020). A uniform carbon price is then, of course, optimal when there is only one emitting entity.

capital use when the sector receives more emissions. Accordingly, these sectors receive lower carbon prices (i.e. more emission) to spur capital accumulation. To comprehend the underlying mechanism fully, consider a two sector economy where the ”complement” sector combines capital and emissions as complements in production and the ”substitute” sector combines capital and emissions as substitutes in production. Also assume that both sectors combine labor and the capital-emissions bundle using identical technologies. A uniform carbon price is not optimal for a simple reason. Lowering the carbon price in the ”complement” sector (and allowing more emission) increases the ”complement” sector's capital demand. Given the fixed economy-wide emissions budget, the ”substitute” sector must emit less CO2 and the ”substitute” sector receives a higher carbon price. A higher carbon price in the ”substitute” sector increases the capital-emissions bundle's price to the extend that the ”substitute” sector uses less capital in production. The strong substitutability between capital and emissions in the ”substitute” sector means, however, a weaker leverage from emissions on capital than in the ”complement” sector, so the decrease in capital use in the ”substitute” sector is smaller than the increase in capital use in the ”complement” sector. Overall, optimally differentiated carbon prices trigger the firms' capital use and spur capital accumulation (which has been too low due to the capital income tax), relative to the case with uniform carbon pricing.

SUBSTITUTABILITY BETWEEN LABOR AND THE CAPITAL-EMISSIONS BUNDLE.— Proposition 6 (ii) reveals that sectors that substitute easily between labor (L) and the capital-emissions bundle (K-E) receive a lower carbon price while sectors that substitute less easily between L and K-E receive a higher carbon price. The main intuition here is to allocate labor to sectors with complementary substitution technologies between L and K-E: more L in these sectors triggers a greater sectoral use of K-E which spurs the capital accumulation due to the greater leverage from L on K. To comprehend the underlying mechanism fully, consider again a two sector economy where the ”complement” sector combines L and K-E as complements in production and the ”substitute” sector combines L and K-E as substitutes in production. For simplicity, also assume that firms combine capital and emissions using the same technology. But why is a uniform carbon price not optimal in this setting? First, decrease the carbon price in the ”substitute” sector to allow for more carbon emissions which also stimulates the capital use in the ”substitute” sector. The K-E bundle becomes relatively cheaper and the ”substitute” sector substitutes away from L towards K-E. Second, labor is perfectly mobile across sectors so the ”now free” labor quantities are taken up by the ”complement” sector. Given the strong sectoral complementarity between L and K-E, additional labor quantities in production trigger a greater demand for the K-E bundle in the ”complement” sector, i.e. the ”complement” sector uses more capital and emissions. The market clearing price for emission, p_t^E, adjusts for the emissions budget to hold. In summary, the regulator allocates more CO2 (and less labor) to the ”substitute” sector but less CO2 (and more labor) to the ”complement” sector, relative to a uniform carbon price, and the total capital use in the economy increases. In other words, optimal carbon prices make use of the heterogeneous sectoral elasticity of substitution between the capital-emissions bundle and labor to incentivise capital accumulation.

THE SPECIAL CASE OF COBB-DOUGLAS TECHNOLOGIES.—Lastly, Cobb-Douglas production technologies are a special case which eliminates the role of technology heterogeneity for carbon price differentiation.

Proposition 7 *Given a positive capital income tax and $\rho_j^{KE} = \rho_j = 0, \forall j$, i.e. Cobb-Douglas production technologies in all sectors, and a final good aggregation of Cobb-Douglas type ($\hat{\rho} = 0$), optimal sectoral carbon prices are uniform.*

PROOF: See Appendix 3.9.4. □

Cobb-Douglas technologies imply constant value shares for each input good. All firms then substitute equally between capital and emission, and between labor and the capital-emissions bundle, and the motive for differentiated carbon prices (to increase the overall firms' capital demand and to incentivise capital accumulation) vanishes. Mathematically, the marginal products of capital and labor do not change with sectoral emissions ($\partial MPK_{it}/\partial E_{jt} = \partial MPL_{it}/\partial E_{jt} = 0, \forall j, i$), so deviating from a uniform carbon price cannot increase the firms' capital demand to spur capital accumulation.

3.4.2 The capital income taxes' role for (non-)uniform carbon prices

Optimal carbon prices are uniform if capital income is not taxed, also when sectors produce with heterogeneous technologies (See Appendix 3.9.5 for a proof of this claim). The intuition is clear. Only in the absence of a capital income tax, the household demands the ”optimal” capital quantities, so the motive for non-uniform carbon prices to spur capital accumulation vanishes. I conclude that increasing the capital income tax rate, $\bar{\tau}^K$, induces a stronger motive for carbon price differentiation. The intuition is that the household demands higher capital returns for the capital supply, however, firms use less capital in production when the price for capital increases relative to the other input goods. As a result the household accumulates less capital. The regulator then deviates stronger from the uniform carbon price to increase the firm's capital demand, i.e. to incentive more capital investment.

3.5 The quantitative analysis

Throughout the numerical analysis I focus on the steady-state, as given in Appendix 3.9.7. In absence of transitional dynamics however, I need to re-formulate the regulator's problem from Section 3.3 to be able to make meaningful welfare comparisons. Section 3.5.1 thus shows how a steady-state setup changes the regulator's objective function. Section 3.5.2 presents our calibration strategy to bring the model to data of the EU-28 economy. Section 3.5.3 introduces different policy setups that I investigate thereafter.

3.5.1 The steady-state of the regulator's problem

Steady-state capital investments are governed by the household's intertemporal preference which must be taken into account also when abstracting from transitional dynamics: only if $\zeta = 0$ (and when abstracting from transitional dynamics), maximizing (3.2) yields the intertemporal efficient allocation of resources (See Phelps (1961)). If $\zeta > 0$ (and when abstracting from transitional dynamics), I need, however, to adjust the social welfare function to be consistent with the household's intertemporal optimality condition (3.3):

$$U = \sum_{t=0}^{\infty} \left(\frac{1}{1+\zeta} \right)^t \left(C_t - \zeta \bar{K}_t \right) . \tag{3.13}$$

Note two major differences between (3.13) and (3.2). First, because capital supply in the Ramsey-framework is infinitely elastic in the steady-state, I include a "linear disutility from capital holdings" term, $\zeta \bar{K}_t$, in (3.13). The parameter ζ guarantees that the optimal choice of $\bar{K}_t$ yields the steady-state version of the intertemporal optimality condition (3.3), i.e. $r_t = (\zeta + \delta)/(1 - \bar{\tau}^K)$. Second, because I abstract from growth, the intertemporal elasticity of substitution $1/\sigma$ does not affect the households effective discount rate and thus the equilibrium, and I impose linear utility in consumption in (3.13). I further impose steady-state capital investments ($I_t = \delta \bar{K}_t$). The complete "steady-state of the regulator's problem" is given in Appendix 3.9.8. The solution to the regulator's problem as presented in this section and the steady-state solution of the regulator's problem as presented in Section 3.3 are identical, i.e. the regulator allocates carbon emissions across sectors equally in both problems.

3.5.2 Calibration

I choose parameters to capture the structure of the aggregated EU-28 economy with respect to (i) the sectoral composition of output, (ii) the mix of capital, labor, and emissions input, and (iii) the observed environmental and capital tax revenues.

I implement the steady-state version of the model in Matlab. To obtain optimal carbon prices, a numerical solver solves the system of equation comprising of the regulator's optimality conditions with respect to L_{jt}, K_{jt} and E_{jt}.

HOUSEHOLD.— Gollier and Hammitt (2014) conclude from the literature that the private sector applies low rates to discount risk-free projects. Accordingly, I set $\zeta = 0.01$—an estimate close to the identified lower bound. Capital depreciates annually with 10%. σ does not impact the steady-state equilibrium, which is why I leave it uncalibrated.

SECTORAL PRODUCTION.—I make use of data from the Global Trade Analysis Project (GTAP) 10 from the year 2014 to calibrate the sectoral production technologies (Aguiar et al. 2019). I distinguish between the following sectors $j \in J = \{AGR, ENE, EIT, MAC, TRN, SER\}$: agriculture ($AGR$), energy ($ENE$), energy-intensive industry (EIT), macro (MAC), transport (TRN) and services (SER). ENE comprises the economic activity from gas, coal, petroleum coal products and crude oil and from the electricity sector. Appendix 3.10 shows the aggregation of the other sectors. I interpret sectoral output as wealth created by the economy and thus focus on *value added*, i.e. the market value of goods and services produced using the primary input factors capital, labor and emissions. The data reveals that the service sector creates the most value and that agriculture creates the least value. The energy sector emits the most emission, followed by the transport sector and the energy-intensive industry.

I survey the literature to find estimates for the elasticity parameters $(\hat{\rho}, \rho_j, \rho_j^{KE})$: Hobijn and Nechio (2019) estimate $\hat{\rho}$ for Europe and confirm the assumption made in many macroeconomic frameworks that impose a unitary elasticity for the final good at a high level of sectoral aggregation. Accordingly, I set $\hat{\rho} = 0$. Okagawa and Ban (2008), van der Werf (2008), Costantini et al. (2019), Dissou et al. (2015), Papageorgiou et al. (2017), Baccianti (2013), Henningsen et al. (2019), Kemfert and Welsch (2000) estimate the substitution elasticity between L and K-E ($\sigma_j = 1/(1 - \rho_j)$) and the substitution elasticity between K and E ($\sigma_j^{KE} =$

$1/(1-\rho_j^{KE})$). The empirical estimates for the elasticity between L and K-E (ρ_j) vary from lower estimates of $(-2.70; -13.3; -24.0; -32.3; -4.91; -4.00)$ to upper estimates of $(-0.85; -0.72; -0.06;\ 0.23;\ -0.15; -2.45)$ in the sectors (AGR, ENE, EIT, MAC, TRN, SER). Complementarity between K and E (ρ_j^{KE}) is the highest in AGR, EIT and MAC—with lower estimates of $(-32.3;\ -24;\ -24)$ and higher estimates of $(-0.11;\ 0;\ 0.03)$ respectively. ENE, TRN and SER substitute more easily between K and E—with lower estimates of $(-1.5;\ -2.2;\ -2.7)$ and upper estimates of $(0.49;\ 0.96;\ 0)$ respectively. For the baseline calibration, I average all estimated σ_j's and σ_j^{KE}'s from the studies on the real-world sectors, given that the real-world sectors fall in the same model sector j. I obtain ρ_j's and ρ_j^{KE}'s via $\rho_j = (\bar{\sigma}_j - 1)/\bar{\sigma}_j$ and $\rho_j^{KE} = (\bar{\sigma}_j^{KE} - 1)/\bar{\sigma}_j^{KE}$, where $\bar{\sigma}_j$ and $\bar{\sigma}_j^{KE}$ denote the averaged estimated elasticity parameter values for sector j from the available studies. I deviate from this strategy solely for ρ_{SER} due to an unreasonable strong complementarity. Instead, I follow Duffy et al. (2004) who find evidence in support of the capital-skill complementarity hypothesis at the aggregate production level—which is a reasonable assumption for the service sector in our case because it accounts for ca. 68.8% of GDP—and I set $\rho_{SER} = -0.05$.

Capital income tax and carbon emissions taxes.— Total capital tax revenues in the EU-28 amounted to €1390 billion in 2019 which is equivalent to 8.4% of GDP (European Commission 2021a). Accordingly, I set the baseline capital income tax rate to 17.5% to obtain the observed capital tax revenues. The steady-state capital income tax revenues determine G, i.e. $G = \tau^K rK$. Total environmental tax revenues in the EU-28 amounted to €389.4 billion, representing 2.4% of EU-28 GDP, in 2019 (European Commission 2021b). My baseline calibration effort aims to reflect this policy design: I impose a uniform carbon price ($p_t^E = 120$, $\tau_{jt}^{E_j} = 0, \forall j$) that raises environmental (and energy) tax revenues of ca. 2.4% of EU-28 GDP in the baseline calibration.

I set the price for labor inputs and sectoral outputs to unity: $w = p_j = 1, \forall j$. The price for capital is given by the steady-state version of the household's in-

TABLE 3.1. Overview of key production values for the baseline calibration.

Parameter	Model sector j					
	AGR	*ENE*	*EIT*	*MAC*	*TRN*	*SER*
Capital expenditure [in bill. €]	87.55	225.03	539.51	1176.34	335.63	5500.81
Labor expenditure [in bill. €]	135.91	81.66	590.96	1475.60	351.45	5515.64
carbon emissions [in Gt CO_2]	0.050	1.767	0.248	0.094	0.999	0.151
Substitution elasticity: L and K-E (ρ_j)	-1.45	-1.84	-1.04	-0.85	-1.24	-0.05
Substitution elasticity: K and E (ρ_j^{KE})	-1.15	0.19	-1.1	-0.92	0.74	-0.34

Notes: I use the primary factors labor and capital from the GTAP 10 database. Capital and labor expenditure denote the amount spent by each sector on the respective input (including taxes and subsidies on inputs). I also treat carbon emissions as a primary factor. carbon emissions stem from intermediate inputs, i.e. gas, coal and crude oil, to each sector which is also provided in the GTAP 10 database.

tertemporal optimality condition (3.3): $r = (\zeta + \delta)/(1 - \bar{\tau}^K)$. Given values and prices of input factors, I obtain the quantities L_j, K_j, Y_j and $\hat{Y}$ in the baseline calibration. Table 3.1 shows the primary factors used in sectoral production and the corresponding *baseline* substitution elasticities. Using all quantities, prices, tax rates and elasticity parameters, I derive $\hat{A}, A_j$ and $\beta_{Y_j}, \beta_{L_j}, \beta_{KE_j}, \beta_{E_j}, \beta_{K_j}, \forall j$ which I list in Appendix 3.9.9.

3.5.3 Different policy setups

I vary the policy setups across three dimensions: the carbon pricing scheme (S), the periodic carbon emissions budget (B) and sectoral technology (Γ).

THE CARBON PRICES.—I distinguish between the policies $s \in S = \{Uniform, Optimal\}$. $Uniform$ imposes a single uniform carbon price. $Optimal$ sectoral carbon prices solve the constrained regulator's problem and are given by (3.12). Both policies implement the carbon emissions budget.

THE CARBON EMISSIONS BUDGET.—I distinguish between different carbon emissions budgets $b \in B = \{0\%, 20\%, 40\%, 60\%, 80\%\}$. In the 0% scenario, the economy does not reduce emissions and emits as much carbon as in the baseline. This scenario serves as a baseline to highlight the economic mechanism of the carbon pricing policies without quantity effects from carbon abatement. I decrease the carbon emissions budget by 20%, 40%, 60% and 80%, given $b \in \{20\%, 40\%, 60\%, 80\%\}$ respectively.

SECTORAL TECHNOLOGY—I vary the sectoral production technologies to investigate which technology heterogeneity aspects drives the degree of carbon prices stronger, i.e. whether heterogeneity in ρ_j or ρ_j^{KE} drive a greater wedge between optimal and uniform carbon prices. The different technology sets are given by $T \in \Gamma =$ $\{Baseline$, *Heterogeneity only between L and K-E*, *Heterogeneity only between K and E*$\}$. In *Baseline*, sectors produce with heterogeneous technologies with respect to ρ_j and ρ_j^{KE} as given by the baseline calibration. In *Heterogeneity only between L and K-E*, I set $\rho_j^{KE} = 0$ but leave ρ_j unchanged. In *Heterogeneity only between K and E*, I set $\rho_j = 0$ but leave ρ_j^{KE} unchanged. For each set, I also re-calibrate the share parameters to be consistent with the baseline.

WELFARE GAINS.— Compensating variation is given by $\sum_{t=0}^{\infty}\left(\frac{1}{1+\zeta}\right)^t((C_{t,b,T}^{Optimal} - \zeta\bar{K}_{t,b,T}^{Optimal}) - (C_{t,b,T}^{Uniform} - \zeta\bar{K}_{t,b,T}^{Uniform}))$, where $C_{t,b,T}^s$ and $\bar{K}_{t,b,T}^s$ denote consumption and the capital stock in period t under budget b, policy s and given Technology T. I express lifetime welfare gains, $\psi_{b,T}$, from optimal carbon prices relative to units of today (assuming uniform carbon prices):

$$\psi_{b,T} = \sum_{t=0}^{\infty}\left(\frac{1}{1+\zeta}\right)^t \left(\frac{\left(C_{t,b,T}^{Optimal} - \zeta\bar{K}_{t,b,T}^{Optimal}\right) - \left(C_{t,b,T}^{Uniform} - \zeta\bar{K}_{t,b,T}^{Uniform}\right)}{\left(C_{0,b,T}^{Uniform} - \zeta\bar{K}_{0,b,T}^{Uniform}\right)}\right). \tag{3.14}$$

$\psi_{20,Baseline}$ thus measures how much the household is better off over his or her lifetime, relative to today, when applying optimal instead of uniform carbon prices, reducing emissions by 20% and given the technology from the baseline calibration.

3.6 The computational experiment

This section assesses the quantitative importance of non-uniform carbon prices in the presence of capital income taxation by evaluating (1) the degree of carbon price differentiation, (2) the increases in the aggregated capital stock and (3) welfare gains.

3.6.1 Optimal vs. uniform carbon prices

I first keep periodic emissions at the baseline level to highlight the economic mechanism of the carbon pricing policies without quantity effects from carbon emissions abatement. I find that optimal sectoral carbon prices differ significantly from uniform.

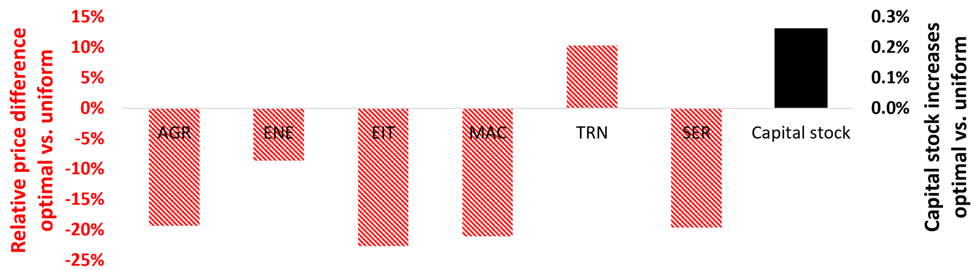


FIGURE 3.1. Optimal sectoral carbon price differentiation and aggregated capital stock increases, relative to a uniform carbon price in %, and given $b = 0\%$.

Notes: Sectoral price deviations from *Optimal* to *Uniform* are calculated via: $p_t^{E,optimal}(1+\tau_{jt}^{E_j})/p^{E,uniform}-1$, where $p^{E,optimal}(1+\tau_{jt}^{E_j})$ are the sectoral carbon prices in *Optimal* and $p_t^{E,uniform}$ is the carbon price in *Uniform*, each in steady-state.

Figure 3.1 shows the deviation from optimal carbon prices relative to the uniform carbon price, under the baseline emissions budget. The differentiation of sectoral CO2 prices depends on (1) the heterogeneity in the K-E substitutability across sectors and on (2) the heterogeneity in the substitutability between L and the K-E bundle across sectors. Sectors with high ρ_j^{KE}-values and low ρ_j-values receive higher prices. This is, for instance, true for the transport sector which receives the highest carbon price that is 10.3% above the uniform carbon price. In contrast, the energy-intensive industry receives the lowest carbon price that is 22.7% below the uniform carbon price. This is because (1) *EIT* substitutes poorly between E and K (given a "low" ρ_{EIT}^{KE}), so a lower carbon price allocates more emissions to *EIT* which triggers the sectoral capital use. Also, (2) *EIT* substitutes relatively well between L and the K-E bundle (given a "high" ρ_{EIT}), so a low sectoral carbon price "frees up" labor resources when EIT substitutes away from labor towards the capital emissions bundle. Other sectors that combine labor and K-E with a "more complementary" technology use these labor resources which triggers their capital use. On average, sectoral carbon prices deviate by 11.1%.[3] Under optimally differentiated carbon prices, the household accumulates up to 0.26% more capital than under a uniform carbon price. The economy produces 0.23% more output and consumes up to 0.07% more in every "steady-state" period. Welfare gains ($\psi_{0,Baseline}$) amount to 4.3% , i.e. the household is 4.3% better off over its lifetime under *Optimal* compared to $Uniform$, relative to today.

3.6.2 Which technology heterogeneity aspect matters the most?

Next, I investigate which heterogeneity aspect in the sectoral production technology matters the most, i.e. whether heterogeneity in ρ_j or heterogeneity in ρ_j^{KE} drive the degree in carbon price differentiation and thus welfare gains. Figure 3.2 shows how carbon prices depend on these sectoral production technology heterogeneity aspects. Relative to the optimal carbon price differentiation in the baseline without carbon emissions reduction (middle red bars), the carbon price wedges decrease when there is *Heterogeneity only between L and K-E* (but not between K and E given $\rho_j^{KE} = 0, \forall j$). See for instance that the left grey bars deviate less from zero than the red bars. In fact, when we observe *Heterogeneity only between L and K-E*, all sectoral carbon prices differ by less than 10% from uniform and the capital stock increases only by 0.01%. The carbon price wedges increase, however, when there is *Heterogeneity only between K and E* (but not between L and K-E because $\rho_j = 0, \forall j$). To see this note that the right blue bars differ more from zero than

[3] I measure the degree of carbon price differentiation with the coefficient of variation in carbon prices $\frac{\sigma_t}{\mu_t}$, where $\sigma_t = \left[\sum_j \frac{E_{jt}}{E_t}\left(p_t^E(1+\tau_{jt}^{E_j}) - \mu_t\right)^2\right]^{0.5}$ is the quantity-weighted standard deviation in sectoral carbon prices and $\mu_t = \sum_j \frac{E_{jt}}{E_t} p_t^E(1+\tau_{jt}^{E_j})$ is the quantity-weighted average carbon price.

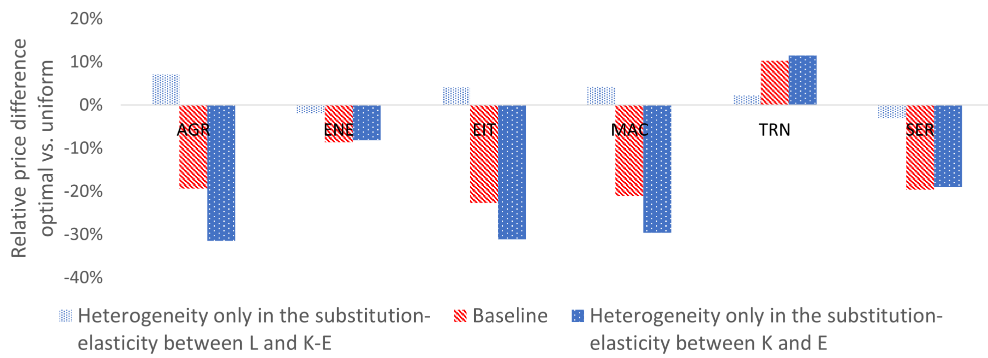


FIGURE 3.2. Relative price differences by sector between *Optimal* vs. *Uniform*, for different technology heterogeneity.

Notes: In Baseline, sectors produce with heterogeneous production technologies on the L-KE level and on the K-E level. Heterogeneity only in the substitution-elasticity between L and K-E assumes $\rho_j^{KE} = 0, \forall j$, and carbon price differentiation decreases relative to the Baseline. Heterogeneity only in the substitution-elasticity between K and E assumes $\rho_j = 0, \forall j$, and carbon price differentiation increases relative to the Baseline.

the middle red bars, and the capital stock increases by 0.38%. But why does heterogeneity in the substitution elasticity between K and E (ρ_j^{KE}) play a more important role to spur capital accumulation using non-uniform carbon prices than heterogeneity in the substitution elasticity between L and K-E (ρ_j)? Note that setting $\rho_j^{KE} = 0$ as in T = *Heterogeneity only between L and K-E* shuts down the direct impact of carbon emissions on the firms' capital demand (and thus capital accumulation) and allows only for indirect effects from CO_2 on capital through sector-specific labor demand responses. In contrast, setting $\rho_j = 0$ as in T = *Heterogeneity only between K and E* allows for direct impacts of sectoral CO_2 on capital. Intuitively, allowing for a direct impact of CO_2 on capital offers a greater leverage to spur capital accumulation and optimal carbon prices differentiation is greater.

3.6.3 The carbon emissions budget

I investigate next how the carbon emissions budget impact these results. I compare how optimal vs. uniform carbon prices change the capital stock and welfare gains.

Figure 3.3 visualizes that carbon price differentiation decreases from 11.1% to 7.5%,

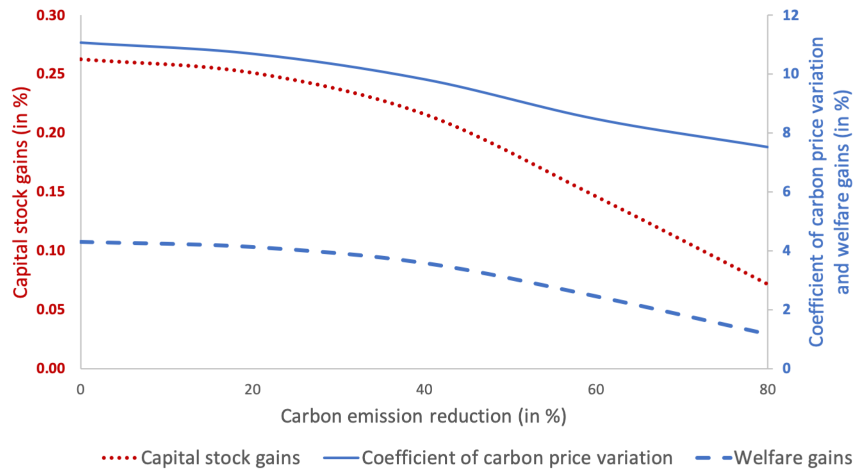


FIGURE 3.3. Capital stock gains, Variation of carbon price coefficient and welfare gains from *Optimal* to $Uniform$ for different carbon emissions reduction targets.

Notes: Capital stock gains decrease, welfare gains decrease and the level of carbon price differentiation decreases with a higher emissions reduction target.

the capital stock gains decrease from 0.26% to 0.07%, and welfare gains decrease from $\psi_{0,Baseline} = 4.3\%$ to $\psi_{80,Baseline} = 1.2\%$, each when reducing emissions by 80%. The intuition is that all sectors need to abate more CO_2, and we observe that carbon prices in all sectors increase by similar magnitudes when decreasing the carbon emissions budget. The coefficient of carbon price variation thus decreases because the standard deviation of carbon prices increases ”less” than the ”mean” carbon price. With a higher reduction target, it is becoming increasingly difficult to stimulate the aggregate capital demand using the sectoral allocation of CO_2. Differences in the capital stock between *Optimal* and $Uniform$ decrease and welfare gains from *Optimal* shrink.

3.7 Sensitivity analysis

3.7.1 Carbon price differentiation depends on the degree of technology heterogeneity

Heterogeneity in the sectoral elasticities of substitution determines the degree of carbon price differentiation. Yet, the literature review as based on (Okagawa and Ban 2008, van der Werf 2008, Costantini et al. 2019, Dissou et al. 2015, Papageorgiou et al. 2017, Baccianti 2013, Henningsen et al. 2019, Kemfert and Welsch 2000) reveals that the empirical estimates on the sectoral substitutability are very uncertain. I thus vary the baseline's elasticity parameters to assess how sensitive the degree of carbon price differentiation is to the sectoral elasticities of substitution ρ_j and ρ_j^{KE}.

DIFFERENT ELASTICITY PARAMETERS.—I vary the elasticity parameters across three scenarios, $\Gamma = \{Baseline, Homogeneous, Heterogeneous\}$. First, the *Baseline* scenario uses the elasticity parameters from the baseline calibration. Second, the elasticity parameters in the *Heterogeneous* scenario cover a wider range than in the baseline. The goal is to show how more heterogeneity in the elasticities of substitution translates into a greater degree of carbon price differentiation. The sectors ENE and TRN have the highest carbon prices in the baseline calibration. Following Proposition 6, I thus decrease ρ_j and I increase ρ_j^{KE} for $j \in \{ENE, TRN\}$ to obtain even higher carbon prices in these sectors. The sectors EIT and MAC have the lowest carbon prices in the baseline calibration. Accordingly, I increase ρ_j and I decrease ρ_j^{KE} for $j \in \{EIT, MAC\}$ to obtain lower carbon prices in these sectors.[4] Third, the *Homogeneous* scenario intends to show a minimal degree of carbon price differentiation. Following Proposition 7, I reduce the degree of carbon price differentiation by setting $\rho_j = 0$ and $\rho_j^{KE} = 0$ whenever "zero" falls within the lower and upper bounds of the estimated values of our literature review. If not,

[4] I adjust each parameter by a factor of two if in line with the estimated bounds from the literature, otherwise I choose the respective bound as identified in the literature. I leave the elasticity parameters of AGR and SER unchanged.

I choose the value that is the closest to zero. Also, I apply $\rho_{SER} = -0.05$ as in the baseline calibration.

Table 3.4 in Appendix 3.9.9 summarizes the elasticity parameters of all scenario. Prices, quantities and taxes remain unchanged from the baseline calibration. Again, using all quantities, prices, tax rates and elasticity parameters, I determine $\hat{A}, A_j$ and $\beta_{Y_j}, \beta_{L_j}, \beta_{KE_j}, \beta_{E_j}, \beta_{K_j}, \forall j$.

THE ECONOMY UNDER 80% CARBON EMISSIONS REDUCTION.— Table 3.2 compares the differences in the capital stock, the degree of carbon price differentiation and the welfare from optimal, relative to uniform. *Optimal* carbon prices spur

TABLE 3.2. Welfare gains ($\psi_{80,\Gamma}$), capital increases (ΔK), coefficients of variation and carbon prices deviation from $Uniform$, all in % (and in the steady-state).

Γ	$\psi_{80,\Gamma}$	ΔK	Coef. of Var.
Heterogeneous	3.6	0.2	18
Baseline	1.2	0.07	8
Homogeneous	0.01	0.0005	0.6

capital accumulation and yields welfare gains relative to an $uniform$ carbon price, but the degree of carbon price differentiation depends crucially on the sectoral elasticity of substitution technology parameters. For instance, carbon prices vary the most in the *Heterogeneous* scenario, on average by 18%, and vary the least in the *Homogeneous* scenario, on average by 0.6%. Driver for the price differentiation is how well sectoral carbon prices can spur the total capital stock accumulation. Optimal carbon prices increase the steady-state capital stock in the *Heterogeneous* scenario by 0.2%, however, optimal carbon prices increase the steady-state capital stock only by 0.0005% in the *Homogeneous* scenario. Welfare gains increase when optimal sectoral carbon prices differ stronger from uniform, and the household ”consumes” 3.6% more over its lifetime in the *Heterogeneous* scenario but welfare gains decrease to negligible 0.01% in the *Homogeneous* scenario.

3.7.2 The capital income tax rate matters

I have so far assumed a capital income tax rate of 17.5%. But what if the government would want to raise more tax revenues from capital income taxes? What are the consequences for the optimal environmental policy? To answer this, I increase and decrease the capital income tax distortion. Welfare gains ($\psi_{0\%,Baseline}$) decreases from 4.3% with $\bar{\tau}^K = 17.5\%$ down to 1.42% when $\bar{\tau}^K = 10\%$ and increase to 8.7% when $\bar{\tau}^K = 25\%$. The intuition is that a greater capital income tax distortion yields a "greater" under-accumulation of capital and carbon prices deviate more from uniform to address this under-accumulation of capital. Accordingly, welfare gains increase.

3.7.3 The elasticity of substitution at the final good level barely matters

I vary the final good aggregation elasticity parameter in Figure 3.4 to assess $\hat{\rho}$'s impact on optimal carbon pricing and the economy. To match the baseline, I also re-calibrate β_{Y_i}. All other parameters remain unchanged. I find $\hat{\rho}$ to have a

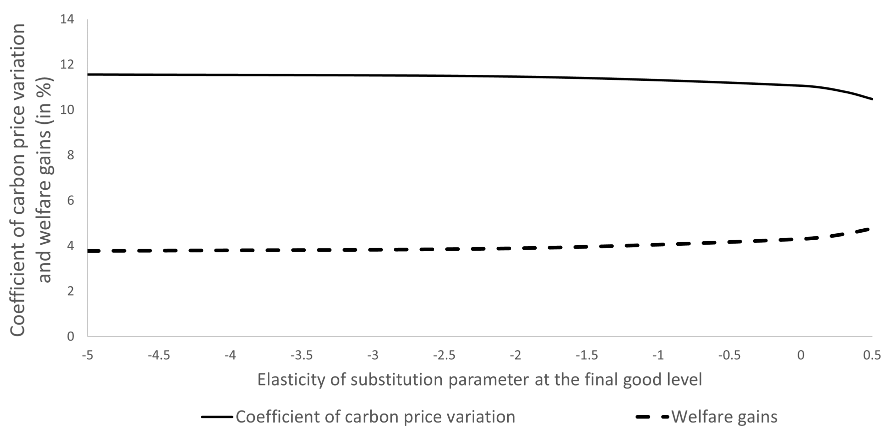


FIGURE 3.4. Variation of carbon price coefficient and welfare gains from *Optimal* to *Uniform* for different $\hat{\rho}$, and given $b = 0\%$.

negligible impact on the degree of carbon price differentiation and welfare, relative to uniform, for a broad range of reasonable $\hat{\rho} \in [-5, 0.5]$. The degree of carbon price differentiation and welfare gains from optimal, relative to uniform, remain on comparable levels.

3.8 Conclusion

This paper asks if carbon emissions should really be priced uniformly when capital income is taxed. I show that a uniform carbon price is optimal only under very restrictive assumptions: either, governments set the capital income tax rate to zero or all sectors must produce with identical technologies. None of these assumptions is plausible.

In a more realistic setting in which the capital income tax distortion prevails and sectors produce with heterogeneous technologies, optimal carbon prices are non-uniform. The intuition is that the capital income tax increases the rental price for capital, so firms use too little in production and households under-accumulate capital. The regulator can then use non-uniform carbon prices to trigger the firms' capital demand which spurs capital accumulation. I show that the sectoral price deviation from uniform depends crucially on the sectoral carbon emissions' elasticity of substitution with other input goods. Generally, sectors that substitute poorly between capital and emissions (or substitute well between labor and the capital-emissions bundle) receive a lower carbon price. The intuition is that sectors receive lower carbon prices, and are thus permitted more emission, the better sectoral emissions can trigger sectoral capital use. The idea is to stimulate the households' capital accumulation (which has been too low due to the capital income tax) through a greater capital demand.

Kalsbach and Rausch (2021) show the same directional carbon price deviations from uniform when the regulator discounts the future less than private actors (and when abstracting from a capital income tax). In such setting, households under-accumulate capital due to a too high private discount rate, and sectoral carbon price deviations aim at triggering the firms' capital use to stimulate the households capital investments. Kalsbach and Rausch (2021) and the here presented work have in common the mechanics how non-uniform carbon prices address the household's under-accumulation of capital, yet, the respective capital market imperfection—different private and social discount rates versus a capital income tax distortion—which drives the result to price carbon emissions non-uniformly are very different in nature.

I show for a reasonable calibration of the EU-28 economy that the sectoral carbon price differentiation is substantial and yields significant welfare gains. For instance, keeping total emissions constant at the baseline level, optimal sectoral carbon prices are up to 23% below or 10% above the uniform carbon price and the household is 4.3% better off over it's lifetime than under a uniform carbon price. The main driver to price carbon non-uniformly is the heterogeneity in the substitution elasticity between capital and emissions because it directly governs how sectoral emissions impact the capital accumulation. The substitution elasticity on the final good level and between labor and the capital-emissions bundle play less of a role. Also, more extreme sectoral elasticity parameters increase the degree of carbon price differentiation and yield greater welfare gains. Imposing almost unitary elasticities in all sectors makes optimal sectoral carbon prices, however, approximately uniform with negligible welfare gains.

Of course, my paper also has its limitations. First, the optimality of non-uniform carbon prices is bound to the distortive characteristics of a capital income tax. But to what extend is the result that capital income should not be taxed really generalizable? In fact, the public economics literature has identified several motives under which capital income taxes are optimal. For instance, capital income taxation is desirable if households over-accumulate capital due to uncertainties (Aiyagari 1994), if capital taxes act as an implicit subsidy for human capital accumulation (Jacobs and Bovenberg 2010), or if returns on saving are heterogenous across households and should be taxed for redistributive reasons (Saez 2002, Gahvari and Micheletto 2016, Kristjansson 2016, Jacobs et al. 2020). When leaving the standard Ramsey framework as done by these authors, it is not clear that my paper's result to price carbon emissions non-uniformly across sectors survives. Second, the directional deviations from a uniform carbon price depend on the nesting structure of production, i.e. on the sectoral elasticity of substitution between capital and emissions and between labor and the capital-emissions bundle. My paper misses, however, to elucidate how the directional price deviations change when imposing a different nesting structure of production or when accounting for intermediate sectoral inputs. Third, I would like to mention that a policymaker needs comprehensive information on sectoral technologies to stimulate the sectors' demand for capital with sectoral carbon emissions. Yet, it remains to be clarified whether a policymaker can really collect and update these information.

In this paper, I show the importance of fiscal tax distortions for the environmental policy design. More precisely, I argue that the presence of a capital income tax that finance public consumption alters the general wisdom in the field to price carbon uniformly. I show that optimally differentiated, sectoral carbon prices can yield large benefits because they spur capital accumulation which has been to low due to the capital income tax distortion.

3.9 Appendix A: Theoretical derivations and proofs

3.9.1 Constrained-optimal policy problem

The constrained regulator's problem includes the resource constraints for carbon emission, labor and capital, the law of capital accumulation, the resource constraint for final good use and the final good production, the household's intertemporal optimality condition and the no lump-sum tax constraint. I also add $2 \times J$ constraints that govern the firms' demand for capital and labor: First, I add $R_t = (1 - \bar{\tau}^K) MPK_{jt}, \forall j$, where $r_t = MPK_{jt} = (1 - \alpha_{jt}) \theta^K_{jt} \gamma_{jt} \hat{Y}_t K^{-1}_{jt}$ denotes the marginal product of capital in sector j. These constraints guarantee that the regulator can not tax (or subsidize) the firms' capital input or final output. Second, I add $w_t = MPL_{jt}, \forall j$, where $w_t = \sum_{k=1}^{J} \alpha_{kt} \gamma_{kt} \hat{Y}_t / \bar{L}$ denotes the labor market clearing price and $MPL_{jt} = \alpha_{jt} \gamma_{jt} \hat{Y}_t L^{-1}_{jt}$ denotes the marginal product of labor in sector j. These constraints prevent the regulator from taxing or subsidizing the firms' labor input, i.e. each firm pays the uniform price w_t for labor input.

Further, γ_{jt} is the value share of the sectoral output relative to aggregated output, $1 - \alpha_{jt}$ is the value share of the emissions-capital bundle in sectoral output and θ^K_{jt} is the value share of capital in the emissions-capital bundle. The value shares given in 3.9.6. The constrained regulator's problem reads:

$$
\begin{aligned}
&\max_{\{\bar{K}_{t+1},C_t,\hat{Y}_t,I_t,\{K_{jt},Y_{jt},L_{jt},E_{jt}\}_{j=1}^{J}\}} \sum_{t=0}^{\infty}\left(\frac{1}{1+\zeta}\right)^t u(C_t) \\
&+\lambda_t^{\bar{E}}\Big(\bar{E}_t-\sum_{j=1}^{J}E_{jt}\Big)+\lambda_t^{\bar{L}}\Big(\bar{L}-\sum_{j=1}^{J}L_{jt}\Big)+\lambda_t^{\bar{K}}\Big(\bar{K}_t-\sum_{j=1}^{J}K_{jt}\Big) \\
&+\lambda_t^{K}\Big(-\bar{K}_{t+1}+(1-\delta)\bar{K}_t+I_t\Big)+\mu_t\Big(\hat{Y}_t-C_t-I_t-G_t\Big) \\
&+\hat{\mu}_t\Big(\hat{Y}_t(Y_{1t}(L_{1t},K_{1t},E_{1t}),\ldots,Y_{jt}(L_{jt},K_{jt},E_{jt})),\ldots,Y_{Jt}(L_{Jt},K_{Jt},E_{Jt}))-\hat{Y}_t\Big) \\
&+\sum_{j=1}^{J}\mu_t^{K_j}\Big(-\underbrace{R_t}_{=r_t(1-\bar{\tau}^K)}+(1-\bar{\tau}^K)\underbrace{(1-\alpha_{jt})\theta_{jt}^{K}\gamma_{jt}\hat{Y}_tK_{jt}^{-1}}_{:=MPK_{jt}}\Big) \\
&+\sum_{j=1}^{J}\mu_t^{L_j}\Big(-\underbrace{\left(\sum_{k=1}^{J}\frac{\alpha_{kt}\gamma_{kt}\hat{Y}_t}{\bar{L}}\right)}_{:=w_t}+\underbrace{\alpha_{jt}\gamma_{jt}\hat{Y}_tL_{jt}^{-1}}_{:=MPL_{jt}}\Big) \\
&+\frac{1}{1+\zeta}\phi_{t+1}\Big(-\Big(1+\underbrace{R_{t+1}}_{=r_{t+1}(1-\bar{\tau}^K)}-\delta\Big)^{\frac{1}{\sigma}}C_t+(1+\zeta)^{\frac{1}{\sigma}}C_{t+1}\Big) \\
&+\Omega_t\underbrace{\Big(-\bar{\tau}^K\sum_{j}^{J}\underbrace{(1-\alpha_{jt})\theta_{jt}^{K}\gamma_{jt}\hat{Y}_t}_{:=MPK_{jt}K_{jt}}-\sum_{j}^{J}\underbrace{(1-\alpha_{jt})(1-\theta_{jt}^{K})\gamma_{jt}\hat{Y}_t}_{:=MPE_{jt}E_{jt}}+G\Big)}_{:=\Lambda_t^G>0}.
\end{aligned}
\tag{3.15}
$$

$\mu_t^{K_j}, \mu_t^{L_j}$ denote the shadow costs of the no-instrument constraint on sectoral capital, labor and output. ϕ_{jt+1} is the Lagrangian multiplier from the household's Euler equation. I assume that tax revenues from taxing capital income and carbon emissions are sufficiently large to finance public consumption ($\Lambda_t^G = \bar{\tau}^K r_t\bar{K}_t + \sum_{j=1}^{J} p_t^E(1+\tau_{jt}^{E_j})E_{jt} - G > 0$). This makes the Lagrangian multiplier on non-negativity constraint of lump-sum taxes, Ω_t, zero, which is why I drop the constraint in the FOCs. The FOCs then read:

$$
\begin{aligned}
C_0: \quad & C_0^{-\sigma} - \mu_0 - \left(\frac{1}{1+\zeta}\right)\left(1 + r_1((1-\bar{\tau}^K) - \delta\right)^{\frac{1}{\sigma}} \phi_1 = 0 \\
C_{t+1}: \quad & \frac{1}{1+\zeta} C_{t+1}^{-\sigma} + \frac{1}{1+\zeta}\phi_{t+1}(1+\zeta)^{\frac{1}{\sigma}} - \left(\frac{1}{1+\zeta}\right)^2 \left(1 + r_{t+1}((1-\bar{\tau}^K) - \delta\right)^{\frac{1}{\sigma}} \phi_{t+2} \\
& - \frac{1}{1+\zeta}\mu_{t+1} = 0 \\
& \Leftrightarrow \left(C_{t+1}^{-\sigma} - \mu_{t+1}\right)(1+\zeta) + (1+\zeta)^{\frac{1}{\sigma}} \left(\phi_{t+1}(1+\zeta) - \phi_{t+2}\frac{C_{t+2}}{C_{t+1}}\right) = 0 \\
I_t: \quad & -\mu_t + \lambda_t^K = 0
\end{aligned}
$$

The optimality conditions for the capital supply, final output and capital income taxes are:

$$
\begin{aligned}
\bar{K}_{t+1}: \quad & \mu_t - \frac{1}{1+\zeta}\mu_{t+1}(1-\delta) - \frac{1}{1+\zeta}\lambda_{t+1}^{\bar{K}} = 0 \\
& \Rightarrow \lambda_{t+1}^K = \mu_{t+1} R_{t+1} \\
\hat{Y}_t: \quad & \hat{\mu}_t - \mu_t - \sum_{j=1}^{J} \frac{\mu_{jt}^{K_j}}{K_{jt}}(1-\bar{\tau}_t^K)(1-\alpha_{jt})\theta_{jt}^K \gamma_{jt} = 0
\end{aligned}
$$

Optimal choice for sectoral emissions is:

$$
\begin{aligned}
E_{jt}: \lambda_t^{\bar{E}} = MPE_{jt}[\hat{\mu}_t &+ \frac{\mu_t^{K_j}}{K_{jt}}(1-\tau^K)\left(\rho_j \alpha_{jt}\theta_j^K - \rho_j^{KE}\theta_t^K + \hat{\rho}(1-\alpha_{jt})\theta_j^K\right) \\
&+ \sum_{i=1}^{J} \frac{\mu_t^{K_i}}{K_{it}}(1-\tau^K)\left((-\hat{\rho})(1-\alpha_{it})\theta_i^K \gamma_{it}\right) \\
&+ \frac{\mu_t^{L_j}}{L_{jt}}\left(-\rho_j \alpha_{jt} + \hat{\rho}\alpha_{jt}\right) + \sum_{l=1}^{J} \frac{\mu_t^{L_l}}{L_{lt}} \frac{\alpha_{lt}\gamma_{lt}}{\sum_{k=1}^{J}\alpha_{kt}\gamma_{kt}}\left(\rho_i \alpha_{it} - \hat{\rho}\alpha_{it}\right)]
\end{aligned}
$$

Optimal choice for sectoral labor is:

$$
\begin{aligned}
L_{jt}: \quad \lambda_t^{\bar{L}} = {} & \hat{\mu}_t \frac{\partial \hat{Y}_t}{\partial L_{jt}} + \frac{\mu_t^{K_j}}{K_{jt}}(1-\tau^K)\left(-\rho_j(1-\alpha_{jt})\theta_j^K + \hat{\rho}(1-\alpha_{jt})\theta_j^K\right) MPL_{jt} \\
& + \sum_{i=1}^{J} \frac{\mu_t^{K_i}}{K_{it}}(1-\tau^K)\left(-\hat{\rho}(1-\alpha_{it})\theta_i^K\gamma_{it}\right) MPL_{jt} \\
& + \frac{\mu_t^{L_j}}{L_{jt}}\left(-1+\rho_j(1-\alpha_{jt})+\hat{\rho}\alpha_{jt}\right) MPL_{jt} \\
& + \sum_{l=1}^{J} \frac{\mu_t^{L_l}}{L_{lt}} \frac{\alpha_{lt}\gamma_{lt}}{\sum_{k=1}^{J}\alpha_{kt}\gamma_{kt}}\left(-\rho_i(1-\alpha_{it})-\hat{\rho}\alpha_{it}\right) MPL_{jt}
\end{aligned}
$$

Optimal choice for sectoral capital is:

$$
\begin{aligned}
K_{jt}: \quad \lambda_t^{\bar{K}} = {} & \hat{\mu}_t \frac{\partial \hat{Y}_t}{\partial K_{jt}} + \frac{\mu_t^{K_j}}{K_{jt}}(1-\tau^K)\left(-1+\rho_j\alpha_{jt}\theta_j^K + \rho_j^{KE}\theta_t^E + \hat{\rho}(1-\alpha_{jt})\theta_j^K\right) MPK_{jt} \\
& + \sum_{i=1}^{J} \frac{\mu_t^{K_i}}{K_{it}}(1-\tau^K)\left(-\hat{\rho}(1-\alpha_{it})\theta_i^K\gamma_{it}\right) MPK_{jt} \\
& + \frac{\mu_t^{L_j}}{L_{jt}}\left(-\rho_j\alpha_{jt}+\hat{\rho}\alpha_{jt}\right) MPK_{jt} \\
& + \sum_{l=1}^{J} \frac{\mu_t^{L_l}}{L_{lt}} \frac{\alpha_{lt}\gamma_{lt}}{\sum_{k=1}^{J}\alpha_{kt}\gamma_{kt}}\left(\rho_i\alpha_{it}-\hat{\rho}\alpha_{it}\right) MPK_{jt}
\end{aligned}
$$

I use $\lambda_t^{\bar{L}} = \mu_t \frac{\partial \hat{Y}_t}{\partial L_{jt}}$ and obtain the following set of equations that fully determine the sectoral carbon prices:

$$
\begin{aligned}
0 =& \frac{\mu_t^{K_j}}{K_{jt}}(1-\tau^K)\left(-\rho_j(1-\alpha_{jt})\theta_j^K + \hat{\rho}(1-\alpha_{jt})\theta_j^K\right) \\
&+ \sum_{i=1}^{J}\frac{\mu_t^{K_i}}{K_{it}}(1-\tau^K)\left((1-\hat{\rho})(1-\alpha_{it})\theta_i^K\gamma_{it}\right) + \frac{\mu_t^{L_j}}{L_{jt}}\left(-1+\rho_j(1-\alpha_{jt})+\hat{\rho}\alpha_{jt}\right) \\
&+ \sum_{l=1}^{J}\frac{\mu_t^{L_l}}{L_{lt}}\frac{\alpha_{lt}\gamma_{lt}}{\sum_{k=1}^{J}\alpha_{kt}\gamma_{kt}}\left(-\rho_i(1-\alpha_{it})-\hat{\rho}\alpha_{it}\right) \\
0 =& \mu_t\tau^K + \frac{\mu_t^{K_j}}{K_{jt}}(1-\tau^K)\left(-1+\rho_j\alpha_{jt}\theta_j^K + \rho_j^{KE}\theta_t^E + \hat{\rho}(1-\alpha_{jt})\theta_j^K\right) \\
&+ \sum_{i=1}^{J}\frac{\mu_t^{K_i}}{K_{it}}(1-\tau^K)\left((1-\hat{\rho})(1-\alpha_{it})\theta_i^K\gamma_{it}\right) \\
&+ \frac{\mu_t^{L_j}}{L_{jt}}\left(-\rho_j\alpha_{jt}+\hat{\rho}\alpha_{jt}\right) + \sum_{l=1}^{J}\frac{\mu_t^{L_l}}{L_{lt}}\frac{\alpha_{lt}\gamma_{lt}}{\sum_{k=1}^{J}\alpha_{kt}\gamma_{kt}}\left(\rho_i\alpha_{it}-\hat{\rho}\alpha_{it}\right) \\
\lambda_t^{\bar{E}} =& MPE_{jt}[\mu_t + \frac{\mu_t^{K_j}}{K_{jt}}(1-\tau^K)\left(\rho_j\alpha_{jt}\theta_j^K - \rho_j^{KE}\theta_t^K + \hat{\rho}(1-\alpha_{jt})\theta_j^K\right) \\
&+ \sum_{i=1}^{J}\frac{\mu_t^{K_i}}{K_{it}}(1-\tau^K)\left((1-\hat{\rho})(1-\alpha_{it})\theta_i^K\gamma_{it}\right) \\
&+ \frac{\mu_t^{L_j}}{L_{jt}}\left(-\rho_j\alpha_{jt}+\hat{\rho}\alpha_{jt}\right) + \sum_{l=1}^{J}\frac{\mu_t^{L_l}}{L_{lt}}\frac{\alpha_{lt}\gamma_{lt}}{\sum_{k=1}^{J}\alpha_{kt}\gamma_{kt}}\left(\rho_i\alpha_{it}-\hat{\rho}\alpha_{it}\right)]
\end{aligned}
$$

I further re-arrange and obtain the optimality condition for L_{jt}:

$$
\begin{aligned}
0 =& -\rho_j(1-\alpha_{jt})\left(\frac{\mu_t^{K_j}}{K_{jt}}(1-\tau^K)\theta_{jt}^K - \frac{\mu_t^{L_j}}{L_{jt}} + \sum_{l=1}^{J}\frac{\mu_t^{L_l}}{L_{lt}}\frac{\alpha_{lt}\gamma_{lt}}{\sum_{k=1}^{J}\alpha_{kt}\gamma_{kt}}\right) \\
&- \hat{\rho}\left(-(1-\alpha_{jt})\frac{\mu_t^{K_j}}{K_{jt}}(1-\tau^K) - \alpha_{jt}\frac{\mu_t^{L_j}}{L_{jt}} + \alpha_{jt}\sum_{l=1}^{J}\frac{\mu_t^{L_l}}{L_{lt}}\frac{\alpha_{lt}\gamma_{lt}}{\sum_{k=1}^{J}\alpha_{kt}\gamma_{kt}}\right) - \frac{\mu_t^{L_j}}{L_{jt}} \\
&+ \sum_{i=1}^{J}\frac{\mu_t^{K_i}}{K_{it}}(1-\tau^K)\left((1-\hat{\rho})(1-\alpha_{it})\theta_i^K\gamma_{it}\right),
\end{aligned}
\tag{3.16}
$$

the optimality condition for K_{jt}:

$$
\begin{aligned}
0 =\ & \rho_j \alpha_{jt} \left(\frac{\mu_t^{K_j}}{K_{jt}} (1-\tau^K)\theta_{jt}^K - \frac{\mu_t^{L_j}}{L_{jt}} + \sum_{l=1}^{J} \frac{\mu_t^{L_l}}{L_{lt}} \frac{\alpha_{lt}\gamma_{lt}}{\sum_{k=1}^J \alpha_{kt}\gamma_{kt}} \right) \\
& - \hat{\rho} \left(-(1-\alpha_{jt}) \frac{\mu_t^{K_j}}{K_{jt}} (1-\tau^K) - \alpha_{jt} \frac{\mu_t^{L_j}}{L_{jt}} + \alpha_{jt} \sum_{l=1}^{J} \frac{\mu_t^{L_l}}{L_{lt}} \frac{\alpha_{lt}\gamma_{lt}}{\sum_{k=1}^J \alpha_{kt}\gamma_{kt}} \right) + \mu_t \tau^K \\
& + \frac{\mu_t^{K_j}}{K_{jt}} (1-\tau^K)(-1 + \rho_j^{KE}\theta_{jt}^E) \\
& + \sum_{i=1}^{J} \frac{\mu_t^{K_i}}{K_{it}} (1-\tau^K) \Big((1-\hat{\rho})(1-\alpha_{it})\theta_i^K \gamma_{it} \Big) \ ,
\end{aligned}
$$

and the optimality condition for E_{jt}:

$$
\begin{aligned}
\frac{\lambda_t^{\bar{E}}}{MPE_{jt}} =\ & -\rho_j^{KE} \frac{\mu_t^{K_j}}{K_{jt}} (1-\tau^K)\theta_{jt}^K + \rho_j \alpha_{jt} \left(\frac{\mu_t^{K_j}}{K_{jt}} (1-\tau^K)\theta_{jt}^K - \frac{\mu_t^{L_j}}{L_{jt}} + \sum_{l=1}^{J} \frac{\mu_t^{L_l}}{L_{lt}} \frac{\alpha_{lt}\gamma_{lt}}{\sum_{k=1}^J \alpha_{kt}\gamma_{kt}} \right) \\
& - \hat{\rho} \left(-(1-\alpha_{jt}) \frac{\mu_t^{K_j}}{K_{jt}} (1-\tau^K) - \alpha_{jt} \frac{\mu_t^{L_j}}{L_{jt}} + \alpha_{jt} \sum_{l=1}^{J} \frac{\mu_t^{L_l}}{L_{lt}} \frac{\alpha_{lt}\gamma_{lt}}{\sum_{k=1}^J \alpha_{kt}\gamma_{kt}} \right) + \mu_t \\
& + \sum_{i=1}^{J} \frac{\mu_t^{K_i}}{K_{it}} (1-\tau^K) \Big((1-\hat{\rho})(1-\alpha_{it})\theta_i^K \gamma_{it} \Big) \ .
\end{aligned}
$$

3.9.2 Optimal policy with an unrestricted policy instrument set

If the policy instrument set is unrestricted, the social regulator's problem includes the resource constraints for carbon emission, labor and capital, the law of capital accumulation, the resource constraint for final good use, the final good production, the household's intertemporal optimality condition and the no lump-sum tax constraint.

$$
\begin{aligned}
&\max_{\{K_{jt},\bar{K}_{t+1},Y_{jt},C_t,\hat{Y}_t,L_{jt},E_{jt},I_t\}_{j=1}^{J}} \sum_{t=0}^{\infty}\Big(\frac{1}{1+\zeta}\Big)^t \Big[u(C_t) \\
&+\lambda_t^{\bar{E}}\Big(\bar{E}_t-\sum_{j=1}^{J}E_{jt}\Big)+\lambda_t^{\bar{L}}\Big(\bar{L}-\sum_{j=1}^{J}L_{jt}\Big)+\lambda_t^{\bar{K}}\Big(\bar{K}_t-\sum_{j=1}^{J}K_{jt}\Big) \\
&+\lambda_t^{K}\Big(-\bar{K}_{t+1}+(1-\delta)\bar{K}_t+I_t\Big)+\mu_t\Big(\hat{Y}_t-C_t-I_t-G_t\Big) \\
&+\hat{\mu}_t\Big(\hat{Y}_t(Y_{1t}(L_{1t},K_{1t},E_{1t}),\ldots,Y_{jt}(L_{jt},K_{jt},E_{jt})),\ldots,Y_{Jt}(L_{Jt},K_{Jt},E_{Jt}))-\hat{Y}_t\Big) \\
&+\frac{1}{1+\zeta}\phi_{t+1}\Big(-\Big(1+\underbrace{R_{t+1}}_{=r_{t+1}(1-\bar{\tau}^K)}-\delta\Big)^{\frac{1}{\sigma}}C_t+(1+\zeta)^{\frac{1}{\sigma}}C_{t+1}\Big) \\
&+\Omega_t\underbrace{\Big(-\bar{\tau}^K\sum_j^J\underbrace{(1-\alpha_{jt})\theta_{jt}^K\gamma_{jt}\hat{Y}_t}_{:=MPK_{jt}K_{jt}}-\sum_j^J\underbrace{(1-\alpha_{jt})(1-\theta_{jt}^K)\gamma_{jt}\hat{Y}_t}_{:=MPE_{jt}E_{jt}}+G\Big)}_{:=\Lambda_t^G>0}.
\end{aligned}
\tag{3.17}
$$

where $\lambda_t^{\bar{E}},\lambda_t^{\bar{K}},\lambda_t^{\bar{L}}$ denote the shadow prices of economy-wide emissions, capital and labor, respectively. $\hat{\mu}_{jt}$ is the shadow price of final output, and μ_t is the shadow cost of consumption. ϕ_{t+1} is the Lagrangian multiplier from the household's intertemporal optimality condition. I assume that tax revenues from taxing capital income and carbon emissions are sufficiently large to finance public consumption ($\Lambda_t^G = \bar{\tau}^K r_t\bar{K}_t+\sum_{j=1}^{J}p_t^E(1+\tau_{jt}^{E_j})E_{jt}-G>0$). This makes the Lagrangian multiplier on non-negativity constraint of lump-sum taxes, Ω_t, zero, which is why I drop the constraint in the FOCs.

The FOCs are given by:

$$
\begin{aligned}
U_{C0} - \mu_0 - \left(\frac{1}{1+\zeta}\right)\left(1 + r_1((1-\bar{\tau}^K) - \delta\right)^{\frac{1}{\sigma}} \phi_1 = 0 \\
\frac{1}{1+\zeta} U_{Ct+1} + \frac{1}{1+\zeta} \phi_{t+1} \left(1+\zeta\right)^{\frac{1}{\sigma}} =... \\
= \left(\frac{1}{1+\zeta}\right)^2 \left(1 + r_{t+1}((1-\bar{\tau}^K) - \delta\right)^{\frac{1}{\sigma}} \phi_{t+2} + \frac{1}{1+\zeta}\mu_{t+1} \\
\hat{\mu}_t \frac{\partial \hat{Y}_t(...)}{\partial L_{jt}} = \lambda_t^{\bar{L}} \\
\hat{\mu}_t \frac{\partial \hat{Y}_t(...)}{\partial E_{jt}} = \lambda_t^{\bar{E}} \\
\hat{\mu}_t \frac{\partial \hat{Y}_t(...)}{\partial K_{jt}} = \lambda_t^{\bar{K}} \\
\hat{\mu}_t = \mu_t \\
\lambda_t^K = \mu_t \\
\left(\frac{1}{1+\zeta}\right)^{t+1} \lambda_{t+1}^{\bar{K}} - \left(\frac{1}{1+\zeta}\right)^{t} \lambda_t^K + \left(\frac{1}{1+\zeta}\right)^{t+1} \lambda_{t+1}^K (1-\delta) = 0 \,.
\end{aligned}
$$

With $\lambda_{t+1}^{\bar{K}} = U_{Ct} r_{t+1}(1-\bar{\tau}^K)$, the regulator's intertemporal optimality condition coincides with the household's intertemporal optimality conditions. Using the conditions for optimal household and firm behavior ((3.3) and (3.6)), the decentralized equilibrium coincides with the social optimum:

$$
\begin{aligned}
\phi_t =& 0 \\
\mu_t =& U_{Ct} \\
\mu_{jt} =& p_{jt} U_{Ct} \\
\lambda_t^{\bar{L}} =& w_t U_{Ct} \\
\lambda_t^{\bar{E}} =& p_t^E U_{Ct} \\
\lambda_t^{\bar{K}} =& r_t(1-\tau^K) U_{Ct} \\
\lambda_t^K =& \mu_t = \hat{\mu}_t = U_{Ct} \,.
\end{aligned}
$$

From the conditions above, it is evident that the social optimum can be decentralized by a carbon tax which is uniform across all j sectors:

$$p_{jt}\frac{\partial Y_{jt}}{\partial E_{jt}} = p_t^E = \lambda_t^E / U_{Ct}\,, \quad \forall j$$

and a capital input subsidy on the firm side, τ_t^K, which is equivalent to the capital income tax on the household side ($\bar{\tau}^K = \tau^K$):

$$r_t(1-\tau_t^K)U_{Ct} = \lambda_t^{\bar{K}} = \hat{\mu}_t \frac{\partial \hat{Y}_t(Y_{1t},\ldots,Y_{jt},\ldots,Y_{Jt})}{\partial Y_{jt}} \frac{\partial Y_{jt}(L_{jt},K_{jt},E_{jt})}{\partial K_{jt}} = U_{Ct}MPK_{jt},$$

where MPK_{jt} denotes the marginal product of capital in sector j.

1) If the regulator has an unrestricted policy set available, the optimal allocation requires a uniform carbon price and each firm receives a capital input subsidy that is equal to the household's capital income tax rate ($\tau^K = \bar{\tau}^K$).
PROOF: I know the social regulator's optimality conditions:

$$\underbrace{\hat{\mu}_t \frac{\partial \hat{Y}_t(Y_{1t},\ldots,Y_{jt},\ldots,Y_{Jt})}{\partial Y_{jt}}}_{:=U_{Ct}p_{it}} \frac{\partial Y_{jt}(L_{jt},K_{jt},E_{jt})}{\partial E_{jt}} = \underbrace{\lambda_t^{\bar{E}}}_{:=U_{Ct}p_t^E} \tag{3.18a}$$

$$\underbrace{\hat{\mu}_t \frac{\partial \hat{Y}_t(Y_{1t},\ldots,Y_{jt},\ldots,Y_{Jt})}{\partial Y_{jt}}}_{:=U_{Ct}p_{it}} \frac{\partial Y_{jt}(L_{jt},K_{jt},E_{jt})}{\partial K_{jt}} = \underbrace{\lambda_t^{\bar{K}}}_{:=U_{Ct}r_t(1-\tau^K)}, \tag{3.18b}$$

where (3.18a) is the optimality condition w.r.t. E_{jt} and (3.18b) is the optimality condition w.r.t. K_{jt}. (3.18a) and (3.18b) show the marginal benefit of emissions and capital use in sector j on the LHS. The marginal costs are given by the Langragian multipliers $\lambda_t^{\bar{E}}$ and $\lambda_t^{\bar{K}}$ from the respective resource constraints on the RHS. Under the curly brackets are the corresponding prices (and taxes) of the decentralised economy. Decentralizing the social optimum is possible by pricing carbon emissions uniformly with p_t^E and by subsidizing the firms' capital use with a rate equal to the capital income tax rate $\bar{\tau}^K = \tau^K$. $\square$

The equivalent allocation can also be implemented with a different instrument set:

2) If capital input cannot be subsidized, i.e. $\tau^K = 0$*, the optimal allocation from 1) is also implemented with a subsidy on final output, a tax on labor input and a tax on emissions input.*

PROOF: The complete optimality conditions from Appendix 3.9.2 that govern the firms' behavior are given by:

$$\frac{\partial \hat{Y}_t}{\partial Y_{it}} = p_{it}, \quad p_{it}\frac{\partial Y_{it}}{\partial K_{it}} = r_t(1-\tau^K), \quad p_{it}\frac{\partial Y_{it}}{\partial E_{it}} = p_t^E, \quad p_{it}\frac{\partial Y_{it}}{\partial L_{it}} = w_t\,, \tag{3.19}$$

with $\tau^K = \bar{\tau}^K$. The *identical* allocation can be implemented with a zero-tax on capital inputs ($\tau^K = 0$), an output subsidy, $\tau^{\hat{Y}}$, sectoral carbon emissions taxes, $\tau_{it}^{E_i}$, and labor input taxes, τ^L:

$$(1+\tau^{\hat{Y}})\frac{\partial \hat{Y}_t}{\partial Y_{it}} = p^*_{it}, \quad p^*_{it}\frac{\partial Y_{it}}{\partial K_{it}} = r_t, \quad p^*_{it}\frac{\partial Y_{it}}{\partial E_{it}} = p_t^E(1+\tau_{it}^{E_i}), \quad p^*_{it}\frac{\partial Y_{it}}{\partial L_{it}} = w_t(1+\tau^L)\,, \tag{3.20}$$

where p^*_{it} are sectoral output prices. The allocations in (3.19) and (3.20) coincide—that is when all quantities are identical—if

$$\tau^{\hat{Y}} = \frac{\bar{\tau}^K}{1-\bar{\tau}^K}, \quad \tau_{it}^{E_i} = \frac{\bar{\tau}^K}{1-\bar{\tau}^K}, \quad \tau^L = \frac{\bar{\tau}^K}{1-\bar{\tau}^K}\,.\square$$

3.9.3 Proof to Proposition 5 and 6

First, I show that optimal carbon prices are uniform when firms produce with identical production technologies as in Proposition 5. Second, I show that optimal carbon prices are non-uniform when sectors produce with heterogenous technologies. Third, I show how sectoral heterogeneity in ρ_j and ρ_j^{KE} drive the degree of carbon price differentiation. The second and the third part combined proof Proposition 6.

1) Optimal carbon prices are uniform when all sectors produce with identical production technologies.

PROOF: Look at the optimality condition w.r.t. E_{jt}:

$$\begin{aligned}\frac{\lambda_t^{\bar{E}}}{MPE_{jt}} &= -\rho_j^{KE}\frac{\mu_t^{K_j}}{K_{jt}}(1-\tau^K)\theta_{jt}^K \\ &+ \rho_j\alpha_{jt}\left(\frac{\mu_t^{K_j}}{K_{jt}}(1-\tau^K)\theta_{jt}^K - \frac{\mu_t^{L_j}}{L_{jt}} + \sum_{l=1}^{J}\frac{\mu_t^{L_l}}{L_{lt}}\frac{\alpha_{lt}\gamma_{lt}}{\sum_{k=1}^{J}\alpha_{kt}\gamma_{kt}}\right) + \mu_t \\ &+ \sum_{i=1}^{J}\frac{\mu_t^{K_i}}{K_{it}}(1-\tau^K)\left((1-\alpha_{it})\theta_i^K\gamma_{it}\right) .\end{aligned}$$

Note that all terms on the RHS are identical when firms produce with identical technologies (including the Lagrangian multipliers $\mu_t^{K_j}$ and $\mu_t^{L_j}$), and optimal carbon prices must be uniform ($MPE_{jt} = MPE_{it}, \forall i, j$). $\square$

2) Optimal carbon prices are non-uniform when sectors produce with heterogeneous production technologies.

PROOF: Look at the terms on the RHS that vary across sectors when firms produce with heterogeneous technologies (different $\rho_j, \rho_j^{KE}, \theta_{jt}^K, \alpha_{jt}$). Intuitively, optimal carbon prices are thus non-uniform, i.e. $MPE_{jt} \neq MPE_{it}$ for at least one combination of i, j.$\square$

3) Ceteris paribus, sector j receives a higher carbon price than sector i when $\rho_j^{KE} > \rho_i^{KE}$, sector j receives a lower carbon price than sector i when $\rho_j < \rho_i$, $j \neq i$.

PROOF: Combine the optimality conditions w.r.t. L_{jt} and K_{jt}:

$$\begin{aligned}
&-\mu_t \tau^K + \frac{\mu_t^{K_j}}{K_{jt}}(1-\tau^K)(1-\rho_j^{KE}\theta_{jt}^E) = \\
&\rho_j \left(\frac{\mu_t^{K_j}}{K_{jt}}(1-\tau^K)\theta_{jt}^K - \frac{\mu_t^{L_j}}{L_{jt}} + \sum_{l=1}^{J} \frac{\mu_t^{L_l}}{L_{lt}} \frac{\alpha_{lt}\gamma_{lt}}{\sum_{k=1}^{J}\alpha_{kt}\gamma_{kt}} \right) + \frac{\mu_t^{L_j}}{L_{jt}} \, ,
\end{aligned} \tag{3.21}$$

and use it in the optimality conditions w.r.t. E_{jt} to obtain:

$$\begin{aligned}
\frac{\lambda_t^{\bar{E}}}{MPE_{jt}} = &\frac{\mu_t^{K_j}}{K_{jt}}(1-\tau^K)\left(\rho_j - \rho_j^{KE}\right)\theta_{jt}^K \\
&+ \tau^K \mu_t + \frac{\mu_t^{L_j}}{L_{jt}}(1-\rho_j) + \rho_j \sum_{l=1}^{J} \frac{\mu_t^{L_l}}{L_{lt}} \frac{\alpha_{lt}\gamma_{lt}}{\sum_{k=1}^{J}\alpha_{kt}\gamma_{kt}} \, .
\end{aligned} \tag{3.22}$$

The directional sectoral price derivations from uniform depend on ρ_j and ρ_j^{KE}, i.e. sector j's carbon price decreases with ρ_j and increases with ρ_j^{KE}.

Consider two firms h and l: firm h has a low ρ_h but a high ρ_h^{KE} and firm l has a high ρ_l but a low ρ_l^{KE}—such that $\rho_h < \rho_l$ and $\rho_h^{KE} > \rho_l^{KE}$ holds. Also, h and l have identical value shares under a uniform carbon price. Relative to a uniform carbon price, I show in the following that sector h should receive a higher carbon price and that sector l should receive a lower carbon price: $MPE_{ht} > MPE_{lt}$. Suppose for mathematical convenience that (i) $\frac{\mu_t^{K_h}}{K_{ht}} = \frac{\mu_t^{K_l}}{K_{lt}}$ and (ii) $\frac{\mu_t^{L_h}}{L_{ht}} = \frac{\mu_t^{L_l}}{L_{lt}}$. Given a uniform carbon price, $MPE_{ht} = MPE_{lt}$, and (i) and (ii), the optimality conditions (3.22) are not satisfied:

$$
\begin{aligned}
&\lambda_t^{\bar{E}} + \epsilon_{ht} = MPE_{ht} \\
&\times \left(\frac{\mu_t^{K_h}}{K_{ht}} (1-\tau^K) \left(\rho_h - \rho_h^{KE}\right) \theta_{ht}^K + \tau^K \mu_t + \frac{\mu_t^{L_h}}{L_{ht}} (1-\rho_h) + \rho_h \sum_{i=1}^{J} \frac{\mu_t^{L_i}}{L_{it}} \frac{\alpha_{it}\gamma_{it}}{\sum_{k=1}^{J} \alpha_{kt}\gamma_{kt}} \right) \\
&< MPE_{lt} \\
&\times \left(\frac{\mu_t^{K_l}}{K_{lt}} (1-\tau^K) \left(\rho_l - \rho_l^{KE}\right) \theta_{lt}^K + \tau^K \mu_t + \frac{\mu_t^{L_l}}{L_{lt}} (1-\rho_l) + \rho_l \sum_{i=1}^{J} \frac{\mu_t^{L_i}}{L_{it}} \frac{\alpha_{it}\gamma_{it}}{\sum_{k=1}^{J} \alpha_{kt}\gamma_{kt}} \right) \\
&= \lambda_t^{\bar{E}} + \epsilon_{lt} \,,
\end{aligned}
$$

where $\epsilon_{ht} < \epsilon_{lt}$ denote the error terms in the respective optimality condition. A uniform carbon price, combined with (i) and (ii) is not optimal because $\rho_h < \rho_l$ and $\rho_h^{KE} > \rho_l^{KE}$, and the optimality conditions (3.22) does thus not hold. Ceteris paribus, moving towards the correct solution is possible when increasing the carbon price in sector h (increasing MPE_{ht}) and when decreasing the carbon price in sector l (decreasing MPE_{lt}). Intuitively, the error terms ϵ_{ht} and ϵ_{lt} in the optimality conditions (3.22) shrink, and carbon prices approach the optimal level (given that ϵ_{ht} and ϵ_{lt} are zero in optimum. The optimal carbon price levels, however, also depend on the $2 \times J$ Lagrangian multipliers $\mu_t^{K_j}$ and $\mu_t^{L_j}$. The $2 \times J$ optimality conditions w.r.t. K_{jt} and L_{jt} in the regulator's problem determine how $\mu_t^{K_j}$ and $\mu_t^{L_j}$ increase or decrease, so (i) and (ii) will no longer hold in optimum. □

3.9.4 Proof to Proposition 7

Optimality requires

$$
\begin{aligned}
\frac{\lambda_t^{\bar{E}}}{MPE_{jt}} &= -\rho_j^{KE}\frac{\mu_t^{K_j}}{K_{jt}}(1-\tau^K)\theta_{jt}^K \\
&\quad + \rho_j\alpha_{jt}\left(\frac{\mu_t^{K_j}}{K_{jt}}(1-\tau^K)\theta_{jt}^K - \frac{\mu_t^{L_j}}{L_{jt}} + \sum_{l=1}^{J}\frac{\mu_t^{L_l}}{L_{lt}}\frac{\alpha_{lt}\gamma_{lt}}{\sum_{k=1}^{J}\alpha_{kt}\gamma_{kt}}\right) + \mu_t \\
&\quad + \sum_{i=1}^{J}\frac{\mu_t^{K_i}}{K_{it}}(1-\tau^K)\left((1-\alpha_{it})\theta_i^K\gamma_{it}\right) .
\end{aligned}
$$

Setting $\rho_j^{KE} = \rho_j = \hat{\rho} = 0, \forall j$, the optimal carbon pricing rule then simplifies to

$$
\lambda_t^{\bar{E}} = MPE_{jt}\left(\mu_t + \sum_{i=1}^{J}\frac{\mu_t^{K_i}}{K_{it}}(1-\tau^K)\left((1-\alpha_{it})\theta_i^K\gamma_{it}\right)\right) ,
$$

and sectoral carbon prices are uniform: $MPE_{jt} = p_t^E(1+\tau_t^E)$, i.e. $\forall j : \tau_t^E = \tau_{jt}^{E_j}$.
□

3.9.5 Proof to claim in section 3.4.2

The optimality conditions w.r.t. capital, emissions and labor read:

$$
\begin{aligned}
0 =& \frac{\mu_t^{K_j}}{K_{jt}}(1-\tau^K)\left(-\rho_j(1-\alpha_{jt})\theta_j^K + \hat{\rho}(1-\alpha_{jt})\theta_j^K\right) \\
&+ \sum_{i=1}^{J} \frac{\mu_t^{K_i}}{K_{it}}(1-\tau^K)\left((1-\hat{\rho})(1-\alpha_{it})\theta_i^K\gamma_{it}\right) + \frac{\mu_t^{L_j}}{L_{jt}}\left(-1+\rho_j(1-\alpha_{jt})+\hat{\rho}\alpha_{jt}\right) \\
&+ \sum_{l=1}^{J} \frac{\mu_t^{L_l}}{L_{lt}} \frac{\alpha_{lt}\gamma_{lt}}{\sum_{k=1}^{J}\alpha_{kt}\gamma_{kt}}\left(-\rho_i(1-\alpha_{it})-\hat{\rho}\alpha_{it}\right) \\
0 =& \mu_t\tau^K + \frac{\mu_t^{K_j}}{K_{jt}}(1-\tau^K)\left(-1+\rho_j\alpha_{jt}\theta_j^K + \rho_j^{KE}\theta_t^E + \hat{\rho}(1-\alpha_{jt})\theta_j^K\right) \\
&+ \sum_{i=1}^{J} \frac{\mu_t^{K_i}}{K_{it}}(1-\tau^K)\left((1-\hat{\rho})(1-\alpha_{it})\theta_i^K\gamma_{it}\right) \\
&+ \frac{\mu_t^{L_j}}{L_{jt}}\left(-\rho_j\alpha_{jt}+\hat{\rho}\alpha_{jt}\right) + \sum_{l=1}^{J} \frac{\mu_t^{L_l}}{L_{lt}} \frac{\alpha_{lt}\gamma_{lt}}{\sum_{k=1}^{J}\alpha_{kt}\gamma_{kt}}\left(\rho_i\alpha_{it}-\hat{\rho}\alpha_{it}\right) \\
\lambda_t^{\bar{E}} =& MPE_{jt}\Bigg[\mu_t + \frac{\mu_t^{K_j}}{K_{jt}}(1-\tau^K)\left(\rho_j\alpha_{jt}\theta_j^K - \rho_j^{KE}\theta_t^K + \hat{\rho}(1-\alpha_{jt})\theta_j^K\right) \\
&+ \sum_{i=1}^{J} \frac{\mu_t^{K_i}}{K_{it}}(1-\tau^K)\left((1-\hat{\rho})(1-\alpha_{it})\theta_i^K\gamma_{it}\right) \\
&+ \frac{\mu_t^{L_j}}{L_{jt}}\left(-\rho_j\alpha_{jt}+\hat{\rho}\alpha_{jt}\right) + \sum_{l=1}^{J} \frac{\mu_t^{L_l}}{L_{lt}} \frac{\alpha_{lt}\gamma_{lt}}{\sum_{k=1}^{J}\alpha_{kt}\gamma_{kt}}\left(\rho_i\alpha_{it}-\hat{\rho}\alpha_{it}\right)\Bigg]
\end{aligned}
$$

If $\bar{\tau}^K = 0$, the system of equation solves with $\mu_t^{K_j} = \mu_t^{L_j} = 0, \forall j$, and I obtain $MPE_{jt} = \lambda_t^{\bar{E}}$ indicating that a uniform carbon price is optimal. $\square$

3.9.6 Closed-form expressions for MPK_{it}, MPL_{it}, w_t and value shares

I note that $MPK_{it} = (1-\alpha_{it})\theta_{it}^K \gamma_{it}\hat{Y}_t/K_{it}$, $w_t = \sum_{k=1}^J \alpha_{kt}\gamma_{kt}\hat{Y}_t/\bar{L}$ and $MPL_{it} = \alpha_{it}\gamma_{it}\hat{Y}_t/L_{it}$. γ_{it} is the value share of sectoral output relative to aggregated output, $1-\alpha_{it}$ is the value share of the emissions-capital bundle in sectoral output and θ_{it}^K is the value share of capital in the emissions-capital bundle, each in sector i. The value shares are given by:

$$
\begin{aligned}
\gamma_{it} =& \frac{\beta_{Yi} Y_{it}^{\hat{\rho}}}{\sum_{j=1}^J \beta_{Yj} Y_{jt}^{\hat{\rho}}} \\
Y_{it} =& A_i \left(\beta_{Lj} L_{it}^{\rho_i} + \beta_{KEi} \left[A_j^{KE} \left(\beta_{Ei} E_{it}^{\rho_i^{KE}} + \beta_{Ki} K_{it}^{\rho_i^{KE}} \right)^{\frac{1}{\rho_i^{KE}}} \right]^{\rho_i} \right)^{\frac{1}{\rho_t}} \\
\alpha_{it} =& \frac{\beta_{Lj} L_{it}^{\rho_i}}{\beta_{Lj} L_{it}^{\rho_i} + \beta_{KEi} \left[A_j^{KE} \left(\beta_{Ei} E_{it}^{\rho_i^{KE}} + \beta_{Ki} K_{it}^{\rho_i^{KE}} \right)^{\frac{1}{\rho_i^{KE}}} \right]^{\rho_i}} \\
\theta_{it}^K =& \frac{\beta_{Ki} K_{it}^{\rho_i^{KE}}}{\beta_{Ki} K_{it}^{\rho_i^{KE}} + \beta_{Ei} E_{it}^{\rho_i^{KE}}} \,.
\end{aligned}
$$

The value shares α_{it} change with the input quantities according to:

$$
\begin{aligned}
\frac{\partial \alpha_{it}}{\partial K_{it}} =& -\rho_i \alpha_{it}(1-\alpha_{it})\theta_{it}^K K_{it}^{-1} \\
\frac{\partial \alpha_{it}}{\partial E_{it}} =& -\rho_i \alpha_{it}(1-\alpha_{it})(1-\theta_{it}^K) E_{it}^{-1} \\
\frac{\partial \alpha_{it}}{\partial L_{it}} =& \rho_i \alpha_{it}(1-\alpha_{it}) L_{it}^{-1} \,.
\end{aligned}
$$

The value shares θ_{it}^K change with the input quantities according to:

$$
\begin{aligned}
\frac{\partial \theta_{it}^K}{\partial K_{it}} =& \rho_i^{KE} \theta_{it}^K (1-\theta_{it}^K) K_{it}^{-1} \\
\frac{\partial \theta_{it}^K}{\partial E_{it}} =& -\rho_i^{KE} \theta_{it}^K (1-\theta_{it}^K) E_{it}^{-1} \,.
\end{aligned}
$$

The value shares γ_{it} change with the input quantities if $i = j$ according to:

$$\begin{aligned}
\frac{\partial \gamma_{it}}{\partial L_{it}} &= \hat{\rho}(1-\gamma_{it}) MPL_{it} \hat{Y}_t^{-1} \\
\frac{\partial \gamma_{it}}{\partial E_{it}} &= \hat{\rho}(1-\gamma_{it}) MPE_{it} \hat{Y}_t^{-1} \\
\frac{\partial \gamma_{it}}{\partial K_{it}} &= \hat{\rho}(1-\gamma_{it}) MPK_{it} \hat{Y}_t^{-1}
\end{aligned}$$

and if $i \neq j$ according to:

$$\begin{aligned}
\frac{\partial \gamma_{jt}}{\partial L_{it}} &= \hat{\rho}(-\gamma_{jt}) MPL_{it} \hat{Y}_t^{-1} \\
\frac{\partial \gamma_{jt}}{\partial E_{it}} &= \hat{\rho}(-\gamma_{jt}) MPE_{it} \hat{Y}_t^{-1} \\
\frac{\partial \gamma_{jt}}{\partial K_{it}} &= \hat{\rho}(-\gamma_{jt}) MPK_{it} \hat{Y}_t^{-1}
\end{aligned}$$

3.9.7 Steady-state conditions

The household's Euler equation and capital investments in steady-state reveal

$$r = \frac{\zeta + \delta}{1 - \bar{\tau}^K}\,, \quad I = \delta \sum_{j=1}^{J} K_j\,.$$

The final good sector's FOCs for all j are

$$p_j = \frac{\partial \hat{Y}}{\partial Y_j}\,.$$

The optimality conditions for sectoral output are

$$w = p_j \frac{\partial Y_j}{\partial L_j}, \quad r = p_j \frac{\partial Y_j}{\partial K_j}, \quad p^E(1 + \tau^{E_j}) = p_j \frac{\partial Y_j}{\partial E_j}\,.$$

Labor supply is given by $\sum_j L_j = \bar{L}$ and total emissions is given by $\sum_j E_j = \bar{E}$. Final output and sectoral output are given by

$$\hat{Y} = \hat{A}\left(\sum_{j=1}^{J} \beta_{Y_j} Y_j^{\hat{\rho}}\right)^{\frac{1}{\hat{\rho}}}$$

$$Y_j = A_j\left(\beta_{L_j}\left(L_j\right)^{\rho_j} + \beta_{KE_j}\left(\underbrace{A_j^{KE}\left[\beta_{E_j} E_j^{\rho_j^{KE}} + \beta_{K_j} K_j^{\rho_j^{KE}}\right]^{\frac{1}{\rho_j^{KE}}}}_{:=Y_j^{KE}}\right)^{\rho_j}\right)^{\frac{1}{\rho_j}}\,.$$

Consumption is $C = \hat{Y} - I - G$.

3.9.8 The steady-state equivalent regulator problem

$$
\begin{aligned}
&\max_{\{\bar{K}_t, C_t, \hat{Y}_t, \{K_{jt}, Y_{jt}, L_{jt}, E_{jt}\}_{j=1}^{J}\}} \sum_{t=0}^{\infty} \left(\frac{1}{1+\zeta}\right)^t \Big(C_t - \zeta \bar{K}_t\Big) \qquad (3.23)\\
&+ \lambda_t^{\bar{E}}\Big(\bar{E}_t - \sum_{j=1}^{J} E_{jt}\Big) + \lambda_t^{\bar{L}}\Big(\bar{L} - \sum_{j=1}^{J} L_{jt}\Big) + \lambda_t^{\bar{K}}\Big(\bar{K}_t - \sum_{j=1}^{J} K_{jt}\Big)\\
&+ \mu_t\Big(\hat{Y}_t - C_t - \delta K_t - G_t\Big)\\
&+ \hat{\mu}_t\Big(\hat{Y}_t(Y_{1t}(L_{1t}, K_{1t}, E_{1t}), \ldots, Y_{jt}(L_{jt}, K_{jt}, E_{jt})), \ldots, Y_{Jt}(L_{Jt}, K_{Jt}, E_{Jt})) - \hat{Y}_t\Big)\\
&+ \sum_{j=1}^{J} \mu_t^{K_j}\Big(- \underbrace{R_t}_{=r_t(1-\bar{\tau}^K)} + (1-\bar{\tau}^K)\underbrace{(1-\alpha_{jt})\theta_{jt}^K \gamma_{jt} \hat{Y}_t K_{jt}^{-1}}_{:=MPK_{jt}}\Big)\\
&+ \sum_{j=1}^{J} \mu_t^{L_j}\Big(- \underbrace{\left(\sum_{k=1}^{J} \frac{\alpha_{kt}\gamma_{kt}\hat{Y}_t}{\bar{L}}\right)}_{:=w_t} + \underbrace{\alpha_{jt}\gamma_{jt}\hat{Y}_t L_{jt}^{-1}}_{:=MPL_{jt}}\Big)\\
&+ \frac{1}{1+\zeta}\phi_{t+1}\Big(- \underbrace{R_{t+1}}_{=r_{t+1}(1-\bar{\tau}^K)} + \zeta + \delta\Big).
\end{aligned}
$$

3.9.9 Parameter calibration

THE BASELINE CALIBRATION.—–Production side parameters are given in Table 3.3:

TABLE 3.3. Baseline calibration parameters.

Parameter	Model sector j					
	AGR	*ENE*	*EIT*	*MAC*	*TRN*	*SER*
β_{K_j}	0.9965	0.5171	0.9975	0.9999	0.5855	0.9999
β_{KE_j}	0.8619	0.9993	0.8716	0.7849	0.9102	0.5252
A_j	0.6185	0.3299	0.5219	0.5931	0.6428	0.7161
β_{Y_j}	0.0140	0.0316	0.0707	0.1622	0.0492	0.6723
ρ_j	-1.45	-1.84	-1.04	-0.85	-1.24	-0.05
ρ_j^{KE}	-1.15	0.19	-1.1	-0.92	0.74	-0.34

Further parameters are $A = 2.9045$, $\hat{\rho} = 0$.

DIFFERENT ELASTICITY PARAMETERS FOR THE SENSITIVITY ANALYSIS.—–I vary the elasticity parameters as follows:

TABLE 3.4. Elasticity parameters in the *"Homogeneous"*, the *"Baseline"* case and in the *"Heterogeneous"* case.

Case	Parameter	Elasticity parameter value in sector j					
		AGR	*ENE*	*EIT*	*MAC*	*TRN*	*SER*
Homogeneous	ρ_j	-0.85	-0.72	-0.06	0	-0.15	-0.05
Homogeneous	ρ_j^{KE}	-0.11	0	0	0	0	0
Baseline	ρ_j	-1.45	-1.84	-1.04	-0.85	-1.24	-0.05
Baseline	ρ_j^{KE}	-1.15	0.19	-1.1	-0.92	0.74	-0.34
Heterogeneous	ρ_j	-1.45	-3.68	-0.52	-0.425	-2.48	-0.05
Heterogeneous	ρ_j^{KE}	-1.15	0.38	-2.2	-1.84	0.96	-0.34

3.10 Appendix B: Mappings of GTAP sectors to model sectors

Table 3.5. Mapping of GTAP sector data to model sectors (I/II).

Category in the data	Model sector
Paddy rice	Agriculture
Wheat	Agriculture
Cereal grains nec	Agriculture
Vegetables, fruit, nuts	Agriculture
Oil seeds	Agriculture
Plant-based fibers	Agriculture
Crops nec	Agriculture
Bovine cattle, sheep and goats, horses	Agriculture
Animal products nec	Agriculture
Raw milk	Agriculture
Wool, silk-worm cocoons	Agriculture
Forestry	Agriculture
Fishing	Agriculture
Coal	Energy
Crude Oil	Energy
Gas	Energy
Petroleum, coal products	Energy
Gas manufacture, distribution	Energy
Electricity	Energy
Other extraction	Energy-intensive Industry
Paper products, publishing	Energy-intensive Industry
Chemical products	Energy-intensive Industry
Basic pharmaceuticals	Energy-intensive Industry
Rubber and plastic products	Energy-intensive Industry
Mineral products nec	Energy-intensive Industry
Ferrous metals	Energy-intensive Industry
Metals nec	Energy-intensive Industry
Metal products	Energy-intensive Industry

Table 3.6. Mapping of GTAP sector data to model sectors (II/II).

Category in the data	Model sector
Meat products nec	Macro
Bovine meat products	Macro
Vegetable oils and fats	Macro
Dairy products	Macro
Processed rice	Macro
Sugar	Macro
Food products nec	Macro
Beverages and tobacco products	Macro
Textiles	Macro
Wearing apparel	Macro
Leather products	Macro
Wood products	Macro
Motor vehicles and parts	Macro
Transport equipment nec	Macro
Computer, electronic and optical products	Macro
Electrical equipment	Macro
Machinery and equipment nec	Macro
Manufactures nec	Macro
Construction	Macro
Land transport and transport via pipelines	Transport
Warehousing and support activities	Transport
Water transport	Transport
Air transport	Transport
Water	Services
Trade	Services
Accommodation and food service activities	Services
Communication	Services
Financial services nec	Services
Insurance	Services
Real estate activities	Services
Other business services nec	Services
Recreational and other services	Services
Public administration and defense	Services
Education	Services
Human health and social work	Services
Dwellings	Services

Chapter 4

Re-distributive income taxes in light of social equity concerns

Abstract

I examine the distributional consequences of non-utilitarian principles for the optimal taxation of income. The model considers two types of households: savers substitute consumption over time more elastically than workers. Savers demand thus lower returns to capital and own all capital. Workers consume their entire labor income. Not taxing capital income—the conventional wisdom in the field—is a special case that assumes a utilitarian perspective under which the planner (implicitly) substitutes the households' consumption over time with the savers' intertemporal elasticities of substitution. Taxation of capital income becomes, however, optimal when the planner substitutes the households' consumption over time with the workers' intertemporal elasticities of substitution. More capital tax revenues are returned back to the savers. Empirically, savers supply labor less elastic, making the subsidy to their labor income less distortionary.

4.1 Introduction

How a society perceives income inequalities is ambigious. For example, as in Bentham (1776), the utilitarian idea maximizes the sum (or average) of individual utilities. In this case, a marginal dollar is worth more to poor households than to rich households, creating room for income redistribution from the rich to the poor. In contrast, the egalitarian idea, as espoused in Rawls (1971), holds that it is society's moral obligation to maximize the utility of the worst-off and not the average individual. A third ethic—libertarianism, as presented in Nozick (1974)—sees the role of government as protecting its citizens from aggression by other individuals. Income and wealth inequality, however, are not such market failures in Nozick's view, and governments should not redistribute income.

The three mindsets base on different ethical concepts, but are all legitimate guiding principles when designing the tax system. Countries differ, however, with respect to these mindsets. For example, the reduction in income inequality in Sweden after redistribution through taxes and transfers is twice as large as in the United States, and Sweden achieves a level of inequality in disposable income that is about half that of the United States (OECD 2015). Thus, the targeting of income equality through the tax system may play a larger role in Sweden than in the United States. It appears that social norms differ between Sweden—a country perceived as more egalitarian—and the United States—a country perceived as more liberal.

Despite the observation that social preferences differ from country to country (and are not necessarily utilitarian), most economic models assume a utilitarian perspective. But what if social preferences are not utilitarian but more egalitarian as in Sweden or more libertarian as in the United States? How does the optimal tax system for redistributing income change? Should capital income be taxed? And if so, should the revenue from the capital income tax be used to make poor households better-off? The paper addresses hence the question of how a (non-) utilitarian policymaker of a growing economy should redistribute income, and shows why and to what extent tax systems differ.

In this paper, I study optimal redistributive taxes in a Ramsey framework with two households. Savers have a higher intertemporal elasticities of substitution with respect to consumption (IES) than workers, savers demand thus lower returns to capital and own all capital. Workers consume their entire labor income. I use analytical and numerical methods to examine how a variety of social welfare functions (SWFs) determine capital and labor income taxes. I focus on the following aspects. First, I challenge the common view that capital income should not be taxed in the long run. I show that this view arises when the planner applies a utilitarian SWF and thus (implicitly) substitutes consumption over time by using the savers' IES. However, capital income should be taxed if the planner substitutes consumption over time, for example, by using the IES of workers without capital. Second, I examine how a more libertarian or a more egalitarian view of social justice—compared to a utilitarian view—decreases the degree of income redistribution. Third, I find that rich households receive more capital income tax revenues and I examine how this depends on household labor supply elasticities.

Before summarizing my results in greater detail, it is useful to introduce my framework. Following Mankiw (2000), my Ramsey framework distinguishes between savers and workers.[1] The government uses taxes on capital income, savers' labor income, and workers' labor income to redistribute income. The government commits to its announced policies. Lump-sum taxes to redistribute income are not possible—the standard assumption in the literature (See Diamond and Mirrlees (1971) or Sandmo (1999) for a detailed discussion). Moreover, the tax system aims to redistribute income and does not finance government spending.

[1] This representation of household heterogeneity in wealth (or capital ownership) has found a broad application in the economics literature. For instance, to investigate the implications for monetary policy (Gali et al. 2007, Bilbiie 2008, Bilbiie and Straub 2013, Challe et al. 2017, Kaplan et al. 2018) or fiscal policy (Bilbiie et al. 2013). Moll (2014) applies the framework to show how entrepreneurial self-financing can undo frictions on the capital market. Lansing (2015) uses the concentrated capital ownership model to investigate the equity risk premium. More closely related to my paper are Judd (1985), Lansing (1999), Reinhorn (2019), Straub and Werning (2020) who investigate the problem of optimal taxation under a utilitarian planner.

But what motivates the focus on capital and labor income taxes? The public finance literature focuses on wedges that the tax system induces in marginal conditions. By introducing an intertemporal wedge (through a capital income tax) and an intratemporal wedge in each household's problem (through the respective labor income tax), I allow for a complete set of policy instruments when investigating optimal taxation of income under equity concerns.

Moreover, I extend the model to include a (non-) utilitarian SWF to capture social equity concerns across two dimensions. *Intra*temporal inequality aversion refers to the planner's perception of inequality across households within a period, i.e. how the planner values the "inequality" that one household is better-off than the other household in the same period. *Inter*temporal inequality aversion refers to the planner's perception of income inequality across periods. The planner is intertemporally inequality neutral when adopting the IES of savers. The planner is intertemporally inequality averse when, for example, adopting the IES of workers with a greater consumption smoothing motive.

The approach holds the following results. First, capital taxation is optimal in a growing economy when the social planner is intertemporally inequality averse, i.e. the planner adopts the IES of the workers without capital holdings. This contrasts with the conventional wisdom of not taxing capital income—based on the traditional Ramsey framework under a utilitarian planner who values future consumption using the IES of capital-owning households. When the planner values future consumption with the IES of workers rather than savers, the planner taxes capital income to make capital investment less attractive. Savers invest less and capital accumulates as preferred by workers. Second, deviating from a utilitarian SWF leads to different patterns of income redistribution: workers receive more transfers from savers with a more egalitarian social objective. Alternatively, laissez-faire policies represent an extreme case in which zero income redistribution is optimal. Third, the revenue from the capital income tax is not used to improve the living conditions of a poor worker—instead, the planner allocates at least twice as much of the tax revenue to savers than to workers. The reason is that, empirically, a saver supplies labor less elastic than a worker, and a saver faces fewer private costs from supplying labor when receiving a dollar of capital income tax revenue. The result is robust to a wide distribution of labor income across households and a wide range of inequality aversion parameters.

These findings are in line with the common understanding that a planner with social equity concerns trades-off efficiency with equity. The efficient (or laissez-faire) market outcome and the optimal market outcome are (potentially) two different equilibria. The efficient market outcome is implemented only by an intra- and intertemporal inequality-neutral planner who does not tax any income source. An optimal (but inefficient) market outcome is implemented by an intra- and intertemporal inequality averse planner who uses distortionary labor and capital income taxes.

The paper's findings have important implications for economic modeling and real-world policymaking. The paper argues that a utilitarian SWF is only one of many possible perspectives on social justice. Under different perspectives on the SWF—all of which are perfectly justifiable—other schemes become optimal. First, I show that the optimal taxation of capital income is related to how a planner substitutes consumption over time. A utilitarian SWF implies that the planner values the future with the IES of the capital-owning household, which effectively amounts to assuming that a zero capital income tax in the steady-state is optimal. However, using the IES of the household without capital holdings is equally justifiable, and capital taxation becomes optimal. Second, a utilitarian SWF also determines the degree of income redistribution among households. Less (or more) income redistribution is also equally well justifiable because the planner may have more libertarian (or more egalitarian) social preferences. A laissez-faire policy may then even be optimal from a social perspective. A final important lesson for the policymaker is that capital income tax revenues should not be used to improve the living conditions of poor households, even if the policymaker is strongly (intratemporal) inequality averse. Instead, intratemporal inequality aversion is better addressed by taxing high labor incomes to finance a subsidy for low labor incomes (or a lower tax on low labor incomes)—in line with the idea to tax labor incomes progressively.

RELATED LITERATURE.— The paper adds intertemporal inequality aversion to the literature as a motive for capital income taxation. However, the standard reasoning, based on the Ramsey framework, argues against capital income taxation. The intuition is that taxing capital income raises the rental price of capital, which lowers capital demand and depresses the real wage. This effect makes any tax on capital income undesirable in the long run, even from the perspective of households that do not own capital. If the government can use a *complete* set of policy instruments (i.e., taxes on labor income or consumption and capital income), neither government spending should be financed with a capital income tax nor should income be redistributed with a capital income tax. My paper thus extends the work of Chamley (1986) and Judd (1985) (and various model extensions as in Kehoe et al. (1999)) who show that capital taxation should be zero in the long run when maximizing a utilitarian SWF. My contribution is to show that capital taxation becomes, however, optimal when the planner values future consumption at, say, the lower IES of workers rather than the higher IES of savers. I derive this result under a complete set of policy instruments available to the planner.

The literature has found that taxation of capital income is optimal under an *incomplete* set of policy instruments (see Chari et al. (2020)). This body of work assumes for instance that one household derives all of its income from capital, while the other household derives all of its income from labor. First, Lansing (1999) finds that capital should indeed be taxed in the long run when the household's IES is one (log-utility). Lansing also emphasizes that his result of a non-zero capital income tax disappears when the planner employs a third instrument that addresses the intratemporal wedge, for instance a consumption tax. Second, Reinhorn (2019) shows that Lansing derives the result by relaxing the convergence hypotheses. He finds that the capital tax is zero or the IES is one to guarantee an internal steady state. The most recent influential paper on this topic is Straub and Werning (2020). They show that the long-run capital tax rate is positive and even converges to infinity when the convergence hypotheses in Chamley (1986) and Judd (1985) do not hold. This is the case when the IES is less than one.

Several other motives for a capital income tax have been proposed by extending the standard Ramsey framework or building on other models. For example, when individuals live finite lives, as in models with overlapping generations, capital taxation is desirable (Diamond 1965, Conesa et al. 2009). When capital enables households to insure themselves against shocks, capital may over-accumulate which justifies capital taxation (Aiyagari 1994). Other examples include heterogeneous preferences for wealth (Saez and Stantcheva 2018), heterogeneous returns to capital (Gahvari and Micheletto 2016, Kristjansson 2016, Jacobs et al. 2020), heterogeneity in household time preferences (Saez 2002, Diamond and Spinnewijn 2011, Golosov et al. 2013), heterogeneity in household and social planner time preferences (Acemoglu et al. 2011, von Below 2012, Belfiori 2017, Barrage 2018), or when labor is not a flow variable but education enables human capital accumulation (Jacobs and Bovenberg 2010, Stantcheva 2017).

But how should capital tax revenues be returned to households? By answering this question, the paper adds to the small literature on optimal income redistributive policies. The classic framework for studying optimal redistributive income taxes is based on a *static* Mirrlees economy.[2] Mirrlees (1971) derives the optimal tax policy under a utilitarian SWF. However, the optimal marginal tax rate on labor income is much higher (Atkinson 1973), taxes on labor income are more progressive (Boskin and Sheshinski 1978), and more income is redistributed among households (Fair 1971) under a more egalitarian objective. The results are robust

[2]A Mirrlees economy consists of a (weakly) separable consumption-leisure choice, homogeneous preferences, and heterogeneous labor productivity across households. For a more detailed discussion of optimal income taxation in a Mirrlees economy, I refer the reader to Stiglitz (2018), Bastani and Waldenström (2020), and Piketty and Saez (2013).

also in different settings (Stern 1976, Phelps 1973, Saez 2001). My work is consistent with these results. A more egalitarian objective increases the marginal welfare from consumption of low-income households. Then, more income is redistributed from high-income households to low-income households. But—by the same logic—should a planner then distribute more capital tax revenue to low-income households? Counterintuitively, I find that low-income households should receive less capital tax revenue.

The paper is organized as follows. Section 4.2 explains the model, section 4.3 introduces the planner's welfare criterion. Section 4.4 shows how the welfare criterion governs the optimal policy design. In Section 4.5, I perform computational exercises to explain the mechanisms at work that relate the social preferences to the optimal taxation scheme. The paper concludes by discussing the implications for optimal capital and labor taxation and the redistribution of capital income tax revenues under alternative social evaluation of inequality.

4.2 The economic environment

4.2.1 explains why savers own all capital. 4.2.2 presents the complete decentralized economy. 4.2.3 elaborates on further conditions that guarantee real-world applicability

4.2.1 A model of capital concentration

Following Mankiw (2000), my model captures that households possess different amounts of wealth (or capital). Driving factor is that households have different, growth adjusted discount rates as governed by their heterogenous IESs. I find evidences in the empirical literature for heterogenous IESs: for instance, the IES is increasing in household income (Blundell et al. 1994) or in wealth (Attanasio et al. 1995). The IES is larger for households with larger asset holdings (Vissing-Jorgensen 2002) and stockholders have a higher elasticity than non-stockholders

(Attanasio et al. 2002). Seeing why capital concentrates at the savers is straightforward. Assume that workers and savers live infinitely and have standard utility functions with constant IESs. Workers have a IES of $1/\sigma$ which is smaller than the savers' IES of one. On the balanced growth path, savers and workers demand then the capital returns[3] R^s and R^w:

$$R^s = \zeta + g \text{ and } R^w = \zeta + \sigma g, \text{ with } \sigma > 1.$$

$g > 0$ denotes the growth rate and ζ is the households' discount rate. Because "$\sigma > 1$", savers supply capital at a lower price than workers ($R^s < R^w$). Firms rent then capital only from the savers and savers end up owning all capital. To prevent that workers may run into debt ad infinitum by taking consumption loans from savers, I follow Blinder (1976), Campbell and Mankiw (1991), Judd (1985), Hubbard et al. (1986) and I impose a liquidity constraint arising from a limited access to loan markets. This assumption is consistent with the stylized fact that the ability to borrow against future earnings is severely limited in the real-world and that households do not have negative wealth throughout their lifetime. In other words, I impose

Assumption 1 *Savers are intertemporally more elastic with respect to consumption than workers and own thus all capital. Workers have a limited access to loan markets and can not borrow consumption loans from savers at any time.*

4.2.2 The decentralized economy

WORKERS.— There is a λ-mass of identical workers with Greenwood et al. (2011)

[3] $R^w = \log(1 + r - \delta) = \log((1 + \zeta)(1 + g)^\sigma) = \log(1 + \zeta) + \sigma \log(1 + g) \approx \zeta + \sigma g$ and $R^s = \log(1 + r - \delta) = \log((1 + \zeta)(1 + g)) = \log(1 + \zeta) + \log(1 + g) \approx \zeta + g$, where δ is the depreciation rate.

preferences (GHH)

$$U^w = \sum_{t=0}^{\infty} \left(\frac{1}{1+\zeta}\right)^t \frac{\left(C_t^w - X_t L_t^{1+\kappa}\right)^{1-\sigma} - 1}{1-\sigma} . \tag{4.1}$$

C_t^w is the individual consumption level, L_t is the labor supply of the worker and $C_t^w - X_t L_t^{1+\kappa}$ is the worker's wellbeing, each at time t. $1/\sigma < 1$ denotes the workers' IES, $1/\kappa$ is the Frisch elasticity of labor supply and governs the variation in the workers' labor supply in response to a change in their after-tax wages. An inelastic labor supply is given by $\kappa = \infty$ and labor supply is perfectly elastic if $\kappa = 0$. $X_t = X_0(1+g)^t > 0$ is a scaling variable that grows at the same rate as the economy and ensures constant working hours on the balanced growth path (Jaimovich and Rebelo 2009)[4]. The workers' budget constraint is

$$C_t^w = (1 - \tau_t^{L_w}) w_t^w L_t \,, \tag{4.2}$$

where $\tau_t^{L_w} > 0$ is the workers' labor income tax and $\tau_t^{L_w} < 0$ is a subsidy. w_t^w denotes the pre-tax wage. Maximizing the utility given in (4.1) subject to (4.2) yields the optimal labor supply decision for all t

$$(1+\kappa) X_t L_t^{\kappa} = (1 - \tau_t^{L_w}) w_t^w . \tag{4.3}$$

In optimum, workers equalize the marginal benefits of labor income with the marginal costs from supplying labor. A worker's net-benefits from labor is then

$$w_t^w (1 - \tau_t^{L_w}) L_t - X_t L_t^{1+\kappa} = \frac{\kappa}{1+\kappa} w_t^w (1 - \tau_t^{L_w}) L_t \,, \tag{4.4}$$

where $w_t^w(1 - \tau_t^{L_w})L_t$ denotes the labor income—the benefits from labor supply—

[4] King-Plosser-Rebello preferences (KPP) denote another commonly used preference structure that yields constant working hours on the balanced growth path. KPP yield, however, a constant labor supply when labor is the only income source (See 4.7.1). Workers would then supply labor independent of the wage and, I choose GHH preferences as in Saez and Stantcheva (2016).

and $-X_t L_t^{1+\kappa}$ denotes the private costs from labor supply. Ceteris paribus, the net benefits from labor, $\frac{\kappa}{1+\kappa} w_t^w (1-\tau_t^{L_w}) L_t$, increases if the household supplies labor less elastic because the household faces less private costs from supplying labor. For instance, the household faces zero private costs when supplying labor perfectly inelastic.

SAVERS.— I impose for the remainder of the paper a unity mass of savers with an IES of one with GHH-preferences

$$U^s = \sum_{t=0}^{\infty} \left(\frac{1}{1+\zeta}\right)^t \log(C_t^s - \hat{X}_t \hat{L}_t^{1+\hat{\kappa}}) \,, \tag{4.5}$$

where C_t^s is the savers' consumption level in period t. $\hat{X}_t, \hat{L}_t$ and $\hat{\kappa}$ have the corresponding interpretation as described in the workers' problem. $C_t^s - \hat{X}_t \hat{L}_t^{1+\hat{\kappa}}$ denotes the savers' wellbeing. The savers' intertemporal budget constraint is

$$K_{t+1} + C_t^s = (1-\tau_t^{L_s}) w_t^s \hat{L}_t + (1-\tau_t^K) r_t K_t + (1-\delta) K_t \,, \tag{4.6}$$

where K_t is the capital stock and r_t is the capital price and w_t^s is the savers' pre-tax wage, each in period t. $\tau_t^K, \tau_t^{L_s} > 0$ denote a capital and labor income tax and $\tau_t^K, \tau_t^{L_s} < 0$ denote a capital and labor income subsidy respectively. δ is the depreciation rate. Savers maximize (4.5) subject to (4.6). The solution to the savers' problem is given by the intertemporal budget constraint (4.6), the intertemporal optimality condition (4.7a) and the intratemporal optimality condition (4.7b)

$$\left(C_t^s - \hat{X}_t \hat{L}_t^{1+\hat{\kappa}}\right)^{-1} = \frac{1}{1+\zeta} \left(C_{t+1}^s - \hat{X}_{t+1} \hat{L}_{t+1}^{1+\hat{\kappa}}\right)^{-1} \left(r_{t+1}(1-\tau_{t+1}^K) + 1 - \delta\right) \tag{4.7a}$$

$$\hat{X}_t (1+\hat{\kappa}) \hat{L}_t^{\hat{\kappa}} = w_t^s (1-\tau_t^{L_s}) \,. \tag{4.7b}$$

(4.7a) governs the optimal consumption path over time and (4.7b) governs the optimal labor quantity supplied in each period. Using (4.7a) and (4.7b) in (4.6) yields the savers' implementability constraint

$$\sum_{t=0}^{\infty}\left(\frac{1}{1+\zeta}\right)^t \frac{C_t^s-(1+\hat{\kappa})\hat{X}_t\hat{L}_t^{1+\hat{\kappa}}}{C_t^s-\hat{X}_t\hat{L}_t^{1+\hat{\kappa}}} = \frac{1}{C_0^s-\hat{X}_0\hat{L}_0^{1+\hat{\kappa}}}(r_0(1-\tau_0^K)+1-\delta)K_0\,. \quad (4.8)$$

I solve both households' problem in Appendix 4.7.2.

TECHNOLOGY.— Output, Y_t, is produced with capital and labor using the standard neoclassical Cobb-Douglas production function

$$Y_t = A_t(K_t)^{\theta_k}(\lambda L_t)^{\theta_{l_w}}(\hat{L}_t)^{\theta_{l_s}}\,. \quad (4.9)$$

The parameters θ_k and $1-\theta_k=\theta_{l_w}+\theta_{l_s}$ are the value shares of capital and total labor. θ_{l_w} and θ_{l_s} are the value shares of workers' and savers' labor, respectively. A_t is the total factor productivity (TFP) which grows at a constant rate $A_t = (1+g)^{1-\theta_k}A_{t-1}$ and guarantees that Y_t and K_t grow with g on the balanced growth path. λL_t denotes the total labor supply of all workers. A constant-return-to-scale production function and price taking behavior yields

$$r_tK_t=\theta_kY_t, \quad w_t^w\lambda L_t=\theta_{l_w}Y_t, \quad w_t^s\hat{L}_t=\theta_{l_s}Y_t\,. \quad (4.10)$$

The technology representation with separated labor markets for savers and workers is standard (See for instance Feldstein (1973) or Lansing (2015)).

GOVERNMENT.— The government collects revenues from taxing capital and labor income and operates an intertemporal balanced budget

$$\tau_t^K r_tK_t+\tau_t^{L_w}w_t^w\lambda L_t+\tau_t^{L_s}w_t^s\hat{L}_t=0\,. \quad (4.11)$$

4.2.3 Conditions for real world applicability

To be in-line with real-world economies, I constrain parameters that govern the distribution of income and that govern the households' labor supply decisions. First, I impose $\theta_{l_s} > \theta_{l_w}/\lambda$ to capture that capital owning households have a higher labor income than households without capital holdings (See for instance European Central Bank (2020)). Second, I impose $\hat{\kappa} > \kappa$ denoting that workers respond more elastic to a wage increase than savers. This captures the stylized empirical finding as summarized in Saez (2002): lower-income earners respond at the intensive margin (working hours) and at the extensive margin (labor market participation). Higher-income earners respond mostly at the intensive margin but barely at the extensive margin. Accounting for responses at both margins, lower-income earners supply labor then more elastic than higher-income earners. Lastly, I allow that the workers' labor tax rate, $\tau_t^{L_w}$, may be different from the savers' labor tax rate, $\tau_t^{L_s}$. This is well in-line with real-world policy making where labor income is taxed progressively so households pay different marginal and average tax rates on their labor income, i.e. $\tau_t^{L_w} < \tau_t^{L_s}$.

4.3 Enriching the utilitarian approach

In 4.3.1, I re-state the utilitarian SWF that adds the households utilities. To capture different equity concept, I introduce a more flexible SWF in 4.3.2. This SWF includes the utilitarian SWF as a special case and it covers different principle of justice ranging from libertarian to egalitarian ethics.

4.3.1 The utilitarian welfare criterion

The utilitarian SWF adds the savers' and workers' utilities:

$$W = \sum_{t=0}^{\infty} \left(\frac{1}{1+\zeta}\right)^t \left[\omega_{st} \log(C_t^s - \hat{X}_t \hat{L}_t^{1+\hat{\kappa}}) + \omega_{wt} \lambda \frac{(C_t^w - X_t L_t^{1+\kappa})^{1-\sigma} - 1}{1-\sigma}\right], \tag{4.12}$$

where ω_{st}, ω_{wt} are welfare weights. The parameter λ reflects that there are "λ" more workers than savers. But is (4.12) a realistic representation of social norms? For instance, (4.12) is maximized when the marginal welfare of households' consumption[5] is equalized, or equivalently

$$(C^w - X_0 L^{1+\kappa}) = \left(\frac{1}{1+g}\right)^{t \times \frac{\sigma-1}{\sigma}} \left(\frac{\omega_{wt}\lambda}{\omega_{st}}\right)^{\frac{1}{\sigma}} \left(C^s - \hat{X}_0 \hat{L}^{1+\kappa}\right)^{\frac{1}{\sigma}}, \tag{4.13}$$

where $C^w, L, C^s, \hat{L}$ are the workers' and savers' growth adjusted consumption and labor supply on the balanced growth path. With constant welfare weights, the planner would transfer the workers' entire income to the savers in the long run ($\lim_{t\to\infty} C_t^w - X_t {L_t}^{1+\kappa} = 0$) because $\lim_{t\to\infty} (1+g)^{-t \times \frac{\sigma-1}{\sigma}} = 0$. The social desirability that workers ought to consume *nothing* due to a lower IES is, however, debatable and does (perhaps) not represent realistic social norms. To prevent that workers consume nothing in the long run, I apply time-varying welfare weights to the households' utility functions in the SWF:

$$\omega_{st} = 1, \quad \omega_{wt} = (C_t^w - X_t L_t^{1+\kappa})^{\sigma-1}. \tag{4.14}$$

The weights guarantee that the workers' IES—a parameter that foremost prevents workers to accumulate capital—no longer governs the redistribution of income.

[5]The marginal welfare of consumption, on the balanced growth path, is

$$\frac{\partial W}{\partial C_t^w} = \left(\frac{(1+g)^{-\sigma}}{1+\zeta}\right)^t \omega_{wt}\lambda(C^w - X_0 L^{1+\kappa})^{-\sigma}, \quad \frac{\partial W}{\partial C_t^s} = \left(\frac{(1+g)^{-1}}{1+\zeta}\right)^t \omega_{st}(C^s - \hat{X}_0 \hat{L}^{1+\kappa})^{-1}.$$

(4.15) is then an equivalent representation of (4.12) with welfare weights (4.14), both SWFs yield the same terms for the marginal welfare from consumption:

$$W = \sum_{t=0}^{\infty} \left(\frac{1}{1+\zeta}\right)^t \left[\log(C_t^s - \hat{X}_{t+1}\hat{L}_{t+1}^{1+\hat{\kappa}}) + \lambda \log(C_t^w - X_t L_t^{1+\kappa})\right] . \tag{4.15}$$

Environmental economists apply the same logic in integrated assessment models that maximize global welfare (See Stanton (2011)). They apply Negishi weights to freeze the current distribution of country income. This eliminates the social desirability of transfers between countries due to heterogeneity in preferences and income. In my framework, w_{st}, w_{wt} eliminate the social desirability of transfers from low income households to high income households due to heterogeneous IESs.[6]

Utilitarianism, however, has been criticized on various grounds. First, a utilitarian SWF is insensitive to the distribution of the total sum of the individuals' wellbeing (Sen 1973). Second, additive separability rules out a number of other ethical concepts as potential candidates for the social preferences (Dasgupta et al. 1973).

4.3.2 The non-utilitarian welfare criterion

I relax the additive separability assumption in (4.15) to consider more egalitarian or more libertarian social norms. I apply the following *normative* SWF:

$$W = \sum_{t=0}^{\infty} \left(\frac{1}{1+\zeta}\right)^t \frac{\left[U_t(C_t^s, C_t^w, L_t, \hat{L}_t; X_t, \hat{X}_t, \kappa, \hat{\kappa}, \sigma_1, \lambda)\right]^{1-\sigma_2} - 1}{1-\sigma_2} \tag{4.16}$$
$$\text{with } U_t(.) = \left[\left(C_t^s - \hat{X}_t\hat{L}_t^{1+\hat{\kappa}}\right)^{1-\sigma_1} + \lambda\left(C_t^w - X_t L_t^{1+\kappa}\right)^{1-\sigma_1}\right]^{\frac{1}{1-\sigma_1}} .$$

[6] Time-varying Negishi weights distort the saving decisions (Denning and Emmerling 2017, Abbott and Fenichel 2014). Hand-to-mouth workers, however, solve a static problem, and weighting only the workers' utilities does not distort the optimal path of capital accumulation.

$U_t(.)$ denotes the instantaneous utility of generation t. Section 4.4.2 explains how (4.16) links to libertarianism, utilitarianism and egalitarianism.[7] The normative SWF embodies the following assumptions.

Assumption 2 *The social planner has CES-preferences over the savers' wellbeing* $(C_t^s - \hat{X}_t \hat{L}_t^{1+\hat{\kappa}})$ *and workers' wellbeing* $(C_t^w - X_t L_t^{1+\kappa})$*. The elasticity parameter* $\sigma_1 \geq 0$ *measures the intratemporal inequality aversion.*

The CES-preference structure captures different ethical concepts of distributive justice. In line with Atkinson (1970), σ_1 reveals the societies' preferences on inequality aversion within a period. If $\sigma_1 = 0$, the social planner is inequality neutral within a period and treats the two household types as perfect substitutes. The social planner then optimizes total consumption as if there was only one representative household. $\sigma_1 > 0$ means that the social planner is intratemporal inequality averse. If $\sigma_1 = 1$, the social planner is as intratemporal inequality averse as a utilitarian planner. If $\sigma_1 = \infty$, the indifference curves are rectangular denoting the maximal intratemporal inequality aversion of a "Rawlsian" planner.

Assumption 3 *The social planner has a constant intertemporal elasticity of substitution over* U_t*. The elasticity parameter* $\sigma_2 \in [1, \sigma]$ *governs the diminishing degree of* U_t *and measures the intertemporal inequality aversion.*

I distinguish between the households' IESs and the planner's intertemporal inequality aversion σ_2 for the following reason. Decisions on capital market only reflect the savers' but not the workers' willingness to substitute consumption over time. We can then ask if the planner should thus only consider households that participate on the capital market—or should the planner also consider households that do not participate on the capital market—to value future wellbeing? These thoughts can be applied to the SWF as follows. The planner's intertemporal inequality

[7] For instance, if $\sigma_1 = 1$, generational utility is $U_t = (C_t^s - \hat{X}_t \hat{L}_t^{1+\hat{\kappa}})(C_t^w - X_t L_t^{1+\kappa})^{\lambda}$. If $\sigma_2 = 1$, social welfare is the discounted sum of $\log(U_t)$. If $\sigma_1 = \sigma_2 = 1$, the social welfare function becomes (4.15).

aversion coincides with the savers' intertemporal elasticity when $\sigma_2 = 1$. Social preferences then only reflect the preferences of capital-owning households, and the social planner is intertemporally inequality neutral (just as a utilitarian planner). $\sigma_2 > 1$ means that social preferences reflect an (weighted) average of both households preferences. The planner uses then the workers' IES instead of savers' IES (or a weighted average) to value future consumption.

4.4 Qualitative results

We have in mind a planner that maximizes the SWF (4.16) using capital income taxes and household specific labor income taxes/subsidies. The problem is subject to the equilibrium conditions of a growing economy in which workers cannot take consumption loans. The equilibrium is given by the workers' budget constraint (4.2), a re-arranged version of the workers' labor supply constraint (4.3), the savers' intertemporal budget constraint (4.6), the savers' implementability constraint (4.8), the production technology constraint (4.9), the government's budget constraint (4.11) and using the constant-return-to-scale production FOCs (4.10) in (4.2), (4.6) and (4.11). The planner's problem is set up in primal form so that the planner can be thought as directly using the quantities as control. We obtain the optimal tax rates by comparing the solution of the planner's problem with the households' FOCs in the decentralized economy. I derive the complete analytical solution to the planner's problem in Appendix 4.7.3. For the remainder of this section, I restrict myself on the balanced growth path (and dropping the time subscription).

I first characterize how the planner's inequality aversion parameters translate into capital and labor income tax rates in 4.4.1. After I investigate how σ_1 and σ_2 relate to different ethical concepts on justice in 4.4.2. In 4.4.3 I briefly look at how capital income tax revenues should be re-distributed among the households.

4.4.1 How should income be taxed?

THE OPTIMAL TAXATION OF CAPITAL INCOME.— Proposition 8 provides the condition when capital income should and should not be taxed.

Proposition 8 *On the balanced growth path with an exogenous growth rate $g > 0$, capital income*

- *is taxed ($\tau^K > 0$) when the social planner is intertemporal inequality averse ($\sigma_2 > 1$),*
- *is not taxed ($\tau^K = 0$) when the social planner is intertemporal inequality neutral ($\sigma_2 = 1$).*

PROOF: The FOC w.r.t. capital in the planner's problem yields the optimal capital income tax rate on the balanced growth path (as derived in Appendix 4.7.3)

$$\tau^K = (1+\zeta)\frac{(1+g)^{\sigma_2} - (1+g)}{r}\,, \tag{4.17}$$

where $r = (1+g)^{\sigma_2}(1+\zeta) + \delta - 1$. Capital income is taxed if $\sigma_2 > 1$. $\sigma_2 = 1$ denotes the limiting case with $\tau^K = 0$.□

Taxing (or not taxing) capital income is thus a matter of how a planner substitutes consumption today with consumption tomorrow. Intertemporal inequality neutrality ($\sigma_2 = 1$) implies that capital income remains untaxed. The way savers accumulate capital does not raise equity concerns, and capital market interventions are then unnecessary. By not taxing capital income, the planner implicitly recognizes that capital markets perfectly reflect the principle of intertemporal equity concerns in society. In contrast, capital market participants, i.e. the savers, value future consumption too much relative to the social preferences when the planner applies the workers' IES ($\sigma_2 = \sigma$). The planner imposes a capital income tax to make investment less attractive and to incentivise consumption today, and capital accumulates as preferred by the workers.

But what is the rational to choose $\sigma_2 > 1$ and how can a planner decide on σ_2? One possibility is to let households vote on $\sigma_2 \in [1, \sigma]$. For this purpose, assume that the majority of the population has little wealth and therefore falls into the workers' category ($\lambda > 1$) and that all households know that the winning vote on σ_2 will result in a capital income tax as specified in (4.17). Under a median or majority voting system, the workers' vote would then be the winner. Each worker chooses $\sigma_2 = \sigma$ because the resulting capital income tax causes capital to accumulate in line with their IES. The planner adopts the worker's preferences and implements a capital tax rate as specified in (4.17) with $\sigma_2 = \sigma$. It evidences that deviating from savers' IES and adopting workers' IES is indeed reasonable as revealed through a democratic voting process. Capital income should then be taxed.

THE OPTIMAL TAXATION OF THE WORKERS' LABOR INCOME.— (4.18) shows how the planner's inequality aversion relates to the workers' labor income tax rate.

$$\left(\frac{\zeta(1+g)\frac{\theta_k}{r} + \frac{\hat{\kappa}}{1+\hat{\kappa}}\left(\theta_{l_s} + \tau^K\theta_k + \tau^{L_w}\theta_{l_w}\right)}{\left(\frac{\kappa}{1+\kappa}(1-\tau^{L_w})\frac{\theta_{l_w}}{\lambda}\right)}\right)^{\sigma_1} \tag{4.18}$$
$$= \frac{\left(\zeta(1+g)\frac{\theta_k}{r} + \frac{\hat{\kappa}}{1+\hat{\kappa}}\left(\theta_{l_s} + \tau^K\theta_k + \tau^{L_w}\theta_{l_w}\right)\right)\left(1 - \frac{1+\kappa}{\kappa}\tau^{L_w}\right)}{\left(\zeta(1+g)\frac{\theta_k}{r} + \frac{\hat{\kappa}}{1+\hat{\kappa}}\theta_{l_s} + \tau^K\theta_k + \tau^{L_w}\theta_{l_w}\right)(1-\tau^{L_w})}.$$

r and τ^K are given in the proof for Proposition 8 and τ^{L_s} is given by the governments budget constraint (4.11). (4.18) yields Corollary 2.

Corollary 2 *On the balanced growth path and given (i) $\infty > \hat{\kappa} > \kappa$, (ii) $\theta_{l_s} > \theta_{l_w}/\lambda$ and (iii) $\sigma_2 \geq 1$, the planner does not tax the workers' labor income ($\tau^{L_w} \ngtr 0$).*[8]

[8]The numerator in the brackets on the LHS in (4.18) denotes the savers' wellbeing and the denominator in the brackets on the LHS in (4.18) denotes the workers' wellbeing. With $\tau^{L_w} = 0$ and $(i) - (ii)$, the LHS in (4.18) is ≥ 1 denoting that savers are relatively better-off than workers: savers have a capital income, savers supply labor less elastic than workers as guaranteed by (i) and savers have a greater pre-tax labor income than workers as guaranteed by (ii). (i) and (ii) ensure a greater marginal welfare from the workers' consumption than from the savers' consumption under zero income taxes ($\zeta(1+g)\frac{\theta_k}{r}\frac{\hat{\kappa}}{1+\hat{\kappa}}\theta_{l_s}(1-\tau^{L_s}) > \frac{\kappa}{1+\kappa}\theta_{l_w}/\lambda(1-\tau^{L_w})$, with $\tau^{L_s} = \tau^{L_w} = 0$). In contrast, the RHS in (4.18) is ≤ 1 with $\tau^{L_w} = 0$ given $\tau^K \geq 0$ as guaranteed by (iii).

The intuition why the planner does not tax the workers' labor income is simple. Effectively, the planner would then transfer income from the poor workers to the rich savers, i.e. the planner would allocate resources to the household with the lower marginal welfare from consumption. So taxing the workers' labor income would widen the wedge between the LHS and RHS in (4.18) instead of closing it (the LHS would increase and savers would become even more relatively better-off than workers). The planner then moves away from the social optimum, and the planner refrains from taxing the workers' labor income.

But how do the tax rates change with the planner's intratemporal inequality aversion?

Proposition 9 *On the balanced growth path and given (i) $\infty > \hat{\kappa} > \kappa$, (ii) $\theta_{l_s} > \theta_{l_w}/\lambda$ and (iii) $\sigma_2 \geq 1$, the planner subsidizes the workers' labor income stronger if the planner is more intratemporal inequality averse ($\partial\tau^{L_w}/\partial\sigma_1 < 0$).*

PROOF: See Appendix 4.7.4 for the proof that $\partial\tau^{L_w}/\partial\sigma_1 < 0$ holds. □
(i) and (ii) imply that workers are "poorer" than workers. Corollary 2 then argues that the planner subsidizes workers' labor income. The subsidy depends, however, on the planner's inequality aversion σ_1. To fully understand the mechanism behind Proposition 9, assume that households supply labor elastically. Increasing the worker's wellbeing by one dollar requires then transfer worth more than one dollar because households incur private costs from supplying labor (see discussion after equation (4.4)). This makes the transfer *expensive* from a social perspective. A greater intratemporal inequality aversion, σ_1, mitigates, however, the planner's perception of this *extra* private cost and the planner transfers more income when σ_1 increases. It becomes more important to make poor workers better off, and the planner subsidizes the poor workers' labor income stronger.

If, however, both households supply labor perfectly inelastic, no household incurs private costs from providing labor. Taxing labor income then mimics a lump-sum transfer system that is then independent of σ_1. Perfectly inelastic labor supply is, however, a rather stark assumption because the empirical literature does find an impact of wages on the households' labor supply.

4.4.1 shows how inter- and intratemporal social justice concerns—modeled in the SWF as normative parameters to reflect ethics in a society—drive the optimal redistributive tax system. Greater intertemporal inequality aversion (σ_2) leads to a higher capital income tax, and greater intratemporal inequality aversion (σ_1) leads to a greater labor income subsidy for poor workers. To explore the extend to which σ_1 and σ_2 reflect social norms, the next section selects three exploratory sets for σ_1 and σ_2. The goal is to show the policy implications when the planner is more egalitarian or more libertarian than a utilitarian planner.

4.4.2 The policy implications of different SWFs

A UTILITARIAN PLANNER.— I first choose $\sigma_1 = \sigma_2 = 1$ to obtain a SWF that adds individual utilities—or equivalently to obtain the utilitarian SWF (4.15). This enables me to compare my findings to the literature which commonly applies a utilitarian framework. I derive the optimal capital and labor taxes

$$\tau^K = 0 \quad \text{and} \quad \tau^{L_w}(1+\lambda)\frac{\theta^{l_w}}{\lambda} = -\left(\zeta(1+g)\frac{\theta^k}{r} + \frac{\hat{\kappa}}{1+\hat{\kappa}}\theta^{l_s}\right) + \frac{\kappa}{1+\kappa}\frac{\theta^{l_w}}{\lambda}\,, \tag{4.19}$$

and the social planner balances the government's budget constraint with the saver's labor income tax τ^{L_s}.

Result 1 *The result of zero-capital income taxation bases upon the premise to maximize an intertemporal neutral—or utilitarian—SWF.*

Applying a utilitarian SWF to aggregate household wellbeing when investigating the optimal tax policy imposes (unknowingly) to value the future using the savers' IES, and a zero-capital income tax is optimal in the long-run. Capital markets then perfectly capture how the planner substitutes consumption today with consumption tomorrow. Intratemporal inequality concerns are best addressed using labor income transfers. (4.19) further shows that workers receive a labor income subsidy if there are sufficiently many poor workers (high λ or low θ_{l_w}), if workers supply labor sufficient elastically (low κ) or if savers supply labor less elastically (high $\hat{\kappa}$).

AN EGALITARIAN PLANNER.— An egalitarian welfare state corrects inequalities through redistribution policies. Egalitarianism translates to the SWF in the following way.

Result 2 *An egalitarian planner with preferences $\sigma_1 > 0, \sigma_2 > 1$ applies a capital income tax, a workers' labor income subsidy and a savers' labor income tax when assuming a sufficiently large share of poor workers.*

In contrast to the policy recommendation for a utilitarian planner, a society that follows above's egalitarian principle of justice would indeed tax capital income (See Proposition 8). Increasing the intratemporal inequality aversion beyond the level of a utilitarian planner yields also a greater workers' labor income subsidy (See Propositions 9).

A LIBERTARIAN PLANNER.— An extreme interpretation of libertarianism is that any income distribution is just leaving no room for any income redistribution.

Result 3 *A fully libertarian planner with $\sigma_1 = 0$ and $\sigma_2 = 1$ ignores distributive inequalities in a society. The planner is inter- and intratemporal inequality neutral and does not re-distributes any income ($\tau^K = \tau^{L_s} = \tau^{L_w} = 0$).*

A social planner who fully admits to libertarian principles of justice is better represented by the parameter choice $\sigma_1 = 0$ and $\sigma_2 = 1$ than by a utilitarian SWF. This type of planner ignores equity concerns and only focuses on an efficient market outcome by not distorting the economy with taxes. Result 3 shows that a *laissez-faire* policy can thus be optimal from a social perspective.

I explain next how the planner redistribute the capital tax revenues between savers and workers.

4.4.3 How should capital income tax revenues be redistributed?

Proposition 10 *On the balanced growth path and given (i) $\infty > \hat{\kappa} > \kappa$, (ii) $\theta_{l_s} > \theta_{l_w}/\lambda$ and (iii) $\sigma_2 \geq 1$, a worker receives less capital tax revenues than a saver at $\sigma_1 = 1$.*

PROOF: Appendix 4.7.5 proofs $\frac{\partial\left((1-\tau^{L_s})\theta_{l_s}\right)}{\partial\sigma_2} > \frac{\partial\left((1-\tau^{L_w})\theta_{l_w}/\lambda\right)}{\partial\sigma_2}$, given $\sigma_1 = 1.\square$
The proof to Proposition 10 shows that a saver's labor income increases more than a worker's labor income, each relative to final output—i.e. the planner allocates more tax revenue to a saver than to a worker. But what is the driving force behind this result? The intuition is that transfers to a saver are less expensive from a social perspective because allocating one dollar of the tax revenues impacts the saver's labor supply decision (and thus his private costs from labor supply) less, compared to a worker. To see this, compare the households' net-benefit from labor income (which I obtain by combining (4.4) and (4.10), net of final output and zero labor tax rates)

$$\frac{\hat{\kappa}}{1+\hat{\kappa}}\theta_{l_s} > \frac{\kappa}{1+\kappa}\frac{\theta_{l_w}}{\lambda}.$$

" $>$ " holds because (i) $\hat{\kappa}/(1+\hat{\kappa}) > \kappa/(1+\kappa)$ and because (ii) $\theta_{l_s} > \theta_{l_w}/\lambda$. (i) states that savers supply labor more elastic so the savers experience greater net-benefits from one dollar of tax revenue. (ii) guarantees that a single saver has a greater labor income than a single worker so allocating one dollar of tax revenues changes their tax rate by $\lambda\theta_{l_w}^{-1}$ and $\theta_{l_s}^{-1}$, respectively. Because the worker's after-tax wage changes relatively stronger, i.e. $\lambda\theta_{l_w}^{-1} > \theta_{l_s}^{-1}$, a worker adjust his labor supply more and faces relatively larger private costs from labor supply. Due to (i) and (ii), the saver has a greater net benefit from the tax revenue and the planner allocates more tax revenues to the saver. I show the robustness of this result for a wide range of σ_1, σ_2-values in the numerical section.

In the previous sections, I showed how the optimal re-distributive taxes depend on the social planner's inequality preferences. The following section further elaborates on this insight. It applies a numerical representation of the EU-28 (EU-27 + UK) economy to provide an in-depth analysis of re-distributive taxes.

4.5 Quantitative results

I start with the calibration strategy in 4.5.1. After, I present the numerical results to provide a more comprehensive economic intuition on the drivers behind the redistribution of household income and capital income tax revenues in 4.5.2-4.5.6.

4.5.1 Calibration

I calibrate a balanced-growth version of the theoretical model, defined by (4.2), (4.3), (4.6), (4.7), (4.9), (4.10) and (4.11), for an aggregated EU-28 economy. Periods correspond to years. I assume policy choices on $\tau^K = \tau^{L_w} = \tau^{L_s} = 0$ and choose the parameters $(\lambda, \kappa, \hat{\kappa}, X_0, \hat{X}_0, g, \theta^k, \theta^{l_w}, \theta^{l_s}, \delta, \zeta, A_0, \sigma)$. I select these parameters to reflect the real-world distribution of capital and labor income across households and how households respond to labor income changes.

CALIBRATION OF CAPITAL INCOME AND CAPITAL ACCUMULATION—For the calibration of θ_k, δ and ζ, I use the WIOT database (World Input-Output Database, Timmer et al. 2015). The data for the most recent available year 2014 provides the capital value share θ_k. I apply data from (OECD 2020) to account for currency exchange rates. I calibrate the discount rate, ζ, by targeting the average EU-28 economy-wide saving rate of 22% from 2000 onwards (World Bank 2020). The savings rate is related to capital depreciation according to $(\delta + g)K/Y$. Using the balanced-growth version of the Euler equation ($\delta = -(1+g)(1+\zeta) + r(1-\tau^K) + 1$), the average capital price ($r = 0.1182$) from 2000-2014 in the WIOT database, the average EU-28 growth rate ($g = 0.0136$) for the same period (The World Bank 2020) and $\tau^K = 0$ in the expression for the savings rate, I obtain $\zeta = 0.0489$. Capital then depreciates with $\delta = 0.055$.

CALIBRATION OF THE LABOR INCOME DISTRIBUTION.—To parameterize λ, θ^{l_s}, and θ^{l_w}, I use the Household Finance and Consumption Survey from the European Central Bank (2020). The data source shows a significant increase in the households' financial assets holding starting at the 60th percentile, and I choose $\lambda = 1.5$ to express that $\lambda/(1+\lambda) = 60\%$ of the households possess little wealth. The same data source also shows the average households' income conditional on household wealth. This, and using θ_k, enables me to derive the workers' and the savers' labor income share.[9]

CALIBRATION OF THE LABOR SUPPLY ELASTICITIES.—I calibrate next the households' Frisch elasticities $(1/\kappa, 1/\hat{\kappa})$. With one representative household, Chetty et al. (2002) recommend an economy-wide Frisch elasticity of labor supply of around 0.75, in line with various econometric studies on the Frisch elasticity (Keane 2011). Saez (2002) further disentangles this estimate in two components. First, estimates for responses at the extensive margin are large at the lower end of the income distribution (0.5 and above) but are "very likely to be small" at the upper end of the income distribution. Low income households thus enter and leave the labor market in response to wage changes while high income households barely respond. Second, there is little consensus on the intensive margin elasticity—or working hour responses—and estimates vary between 0.25 and 0.5. I use the middle value of this range ($(0.25 + 0.5)/2 = 0.3750$). I calibrate the workers' Frisch elasticity to $1/\kappa = 0.5 + 0.3750 = 0.8750$ capturing that lower income household respond at the extensive and intensive margin. I calibrate the savers' Frisch elasticity to $1/\hat{\kappa} = 0.3750$ capturing that the higher income households respond mostly at the intensive margin. The economy-wide Frisch elasticity then aggregates to 0.6750[10],

[9] The 40% richest households (savers) receive 58% of the total labor income. The economy consists of 1.5 times as many workers as savers ($\lambda = 1.5$) and the workers' value share is 72% of the savers' value share. I obtain this result by re-arranging the firm's FOCs $\frac{w^s \hat{L}}{w^w L} = \lambda \frac{\theta^{l_s} Y_t}{\theta_{l_w} Y_t}$. Data shows $\frac{w^s \hat{L}}{w^w L} = 2.064$ and $\theta^{l_w} = \frac{\lambda}{2.064}\theta^{l_s} = 0.7267\theta^{l_s}$ follows.

[10] A worker and a saver supply $1/\kappa$ and $1/\hat{\kappa}$ more units of labor, in response to increasing the after-tax wages. Weighted with the households' share of the population yields the economy-wide Fritsch elasticity: $\frac{\lambda}{1+\lambda}\frac{1}{\kappa} + \frac{1}{1+\lambda}\frac{1}{\hat{\kappa}} = 0.6750$.

in line with Chetty et al. (2002) and Keane (2011).

CALIBRATION OF SCALING PARAMETERS.—Further, I set X_0 and $\hat{X}_0$ to guarantee a unity labor supply ($L = \hat{L} = 1$) under zero taxation. This enables me to calibrate A_0 to obtain the EU-28 GDP of 2018 (ca. 18,700 billion US dollar).

In line with Blundell et al. (1994), Attanasio et al. (1995), Vissing-Jorgensen (2002), Attanasio et al. (2002), I set $\sigma = 2.5$ to ensure that a low-income worker has a lower IES than a high-income saver. Table 4.1 summarizes the calibrated

TABLE 4.1. Baseline with zero capital taxation.

Worker		Saver		Firm		Other	
κ	1.1429	$\hat{\kappa}$	2.6667	A_0	236.00	λ	1.5
IES	1/2.5	IES	1	θ_k	0.3881	g	0.0136
X_0	1491.5	$\hat{X}_0$	1805.6	θ_{l_s}	0.3549	δ	0.0550
ζ	0.0489	ζ	0.0489	θ_{l_w}	0.2570	r	0.1182

parameter choices to which I refer to as the baseline.

4.5.2 The households under different SWFs

Figure 4.1 visualizes the quantitative relevance of previous theoretical findings. The figure plots the relative differences in a single saver's and a single worker's wellbeing in steady-state, given a rather libertarian planner, a utilitarian planner and a rather egalitarian planner respectively. The saver is ca. 125% better-off than a worker under a rather libertarian planner, the utilitarian planner decreases this difference down to ca. 72%, while a more egalitarian planner decreases the difference further down to 25%. These numbers show the great dependency of the *optimal* income redistribution on the underlying SWF. Income inequalities increase under libertarian social preferences and can be mitigated under egalitarian social preferences. The figure also hindsights that the planner does not use the capital income tax revenue to make a worker better-off, relative to a saver: a saver remains ca. 25% better-off than a worker under the egalitarian planner, also when capital income is taxed.

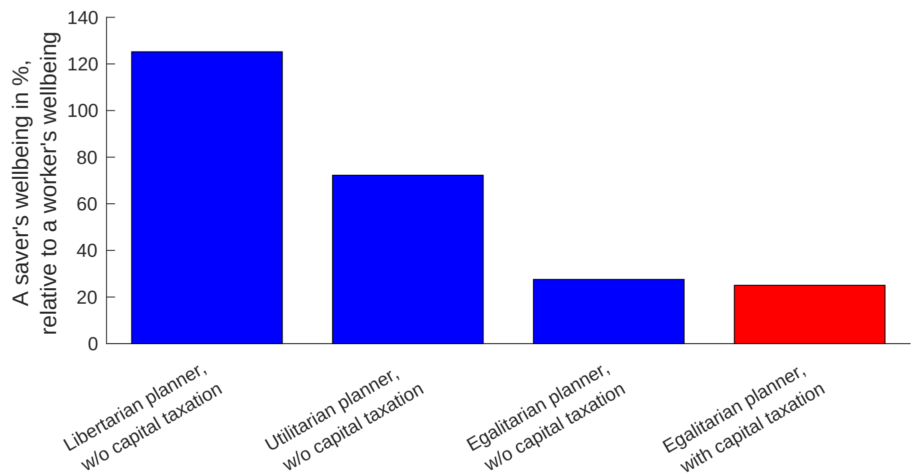


FIGURE 4.1. Differences in the households' wellbeing under different SWF.

Notes: The saver's wellbeing, in percentage points deviation from the worker's wellbeing, for a rather libertarian planner ($\sigma_1 = 0.5$), for a utilitarian planner ($\sigma_1 = 1$) and for a rather egalitarian planner ($\sigma_1 = 3$). **Blue bars**: The social planner does not tax capital income, given $\sigma_2 = 1$. **Red bar**: The social planner does tax capital income, given $\sigma_2 = 2.5$.

In the following I elaborate on the underlying mechanisms: 4.5.3 looks at how intratemporal inequality aversion σ_1 impacts re-distributive taxes. 4.5.4 investigates how intertemporal inequality aversion σ_2 impacts the economic dynamics. 4.5.5 elaborates why the capital income tax revenues should not be used to make poor households better-off. 4.5.6 performs sensitivity analyses on this issue.

4.5.3 How does intratemporal inequality aversion impact the labor income tax rate?

Figure 4.2 visualizes the workers' and savers' tax rates for different intratemporal inequality aversion parameters. The red solid line denotes the workers' labor income subsidy and the blue dashed line denotes the savers' labor income tax. I impose intertemporal inequality neutrality ($\sigma_2 = 1$) to guarantee a zero capital income tax that allows me to perform the analyses on the balanced growth path. A libertarian planner (at $\sigma_1 = 0$) transfers no income: neither the workers' labor income is subsidized, nor the savers' labor income is taxed. The workers' labor income subsidy

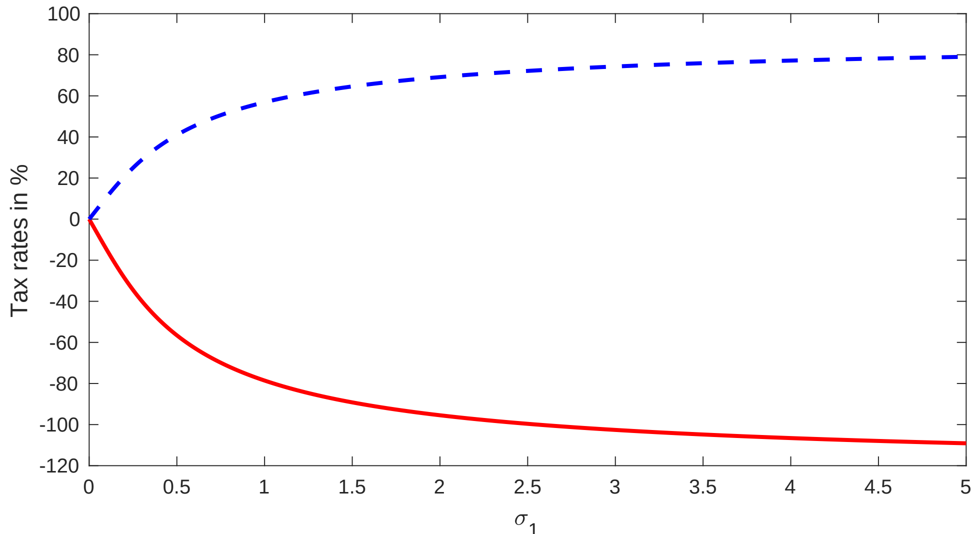


FIGURE 4.2. The households' labor income tax rates under different intratemporal inequality aversions.

Notes: **Blue dashed line**: The savers' labor income tax. **Red solid line**: The workers' labor income subsidy.

increases when the planner becomes more intratemporal inequality averse, as shown by the decreasing red solid line in Figure 4.2. The planner finances this with a greater savers' labor income tax, as shown by the increasing blue dashed line in Figure 4.2. The most extreme case is a Rawlsian planner with $\sigma_1 = \infty$ who maximizes the wellbeing of the worse-off household: the workers' labor income subsidy rate increases to ca. 110%, the savers' labor income tax rate increases to ca. 80%.

4.5.4 How does intertemporal inequality aversion impact the economic dynamics?

I compare three intertemporal inequality averse planners. The first planner—denoted by the solid lines—applies the savers' IES with $\sigma_2 = 1$, the second planner—denoted by the dotted lines—applies the workers' IES with $\sigma_2 = 2.5$ and the third planner—denoted by the dashed lines—applies the average of both IESs with $\sigma_2 = 1.75$. All planners have $\sigma_1 = 1.5$ in common. Each simulation starts with the same capital stock, putting the planner with preferences $\sigma_1 = 1.5, \sigma_2 = 1$ on the balanced growth path from period one onwards. Figure 4.3 plots the dynamics of

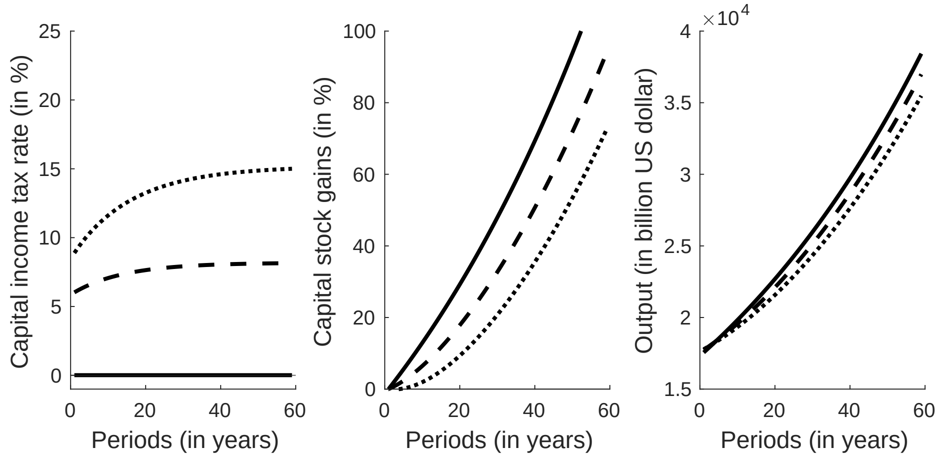


FIGURE 4.3. Dynamics under capital taxation.

Notes: Optimal capital income taxes (left image), capital stock gains (middle image) and output dynamics (right image) under different intertemporal inequality aversion parameters. **Dotted lines**: Dynamics with social preferences $\sigma_1 = 1.5, \sigma_2 = 2.5$. **Dashed lines**: Dynamics with social preferences $\sigma_1 = 1.5, \sigma_2 = 1.75$. **Solid lines**: Dynamics with social preferences $\sigma_1 = 1.5, \sigma_2 = 1$.

the optimal capital income taxes in the left image, the capital stock gains in the middle image and output in the right image, for each intertemporal inequality aversion parameters σ_2. If $\sigma_2 > 1$, the planner employs a capital income tax because the savers' investment decisions do not *correctly* reflect the intertemporal inequality aversion (See Propositions 8). The capital income tax rates converge to 8.2% and 15.2% for $\sigma_2 = 1.75$ and $\sigma_2 = 2.5$ respectively. The planner redistributes the tax revenues to the households that consume it in the same period. Fewer resources are invested and the capital stock grows at a lower rate before the balanced-growth path is reached (middle image). Less output is produced in the future (right image) leading to less consumption and a lower wellbeing of future generations. Yet, the dynamics no longer follow the savers' preferences as given by their IES of one, instead they follow the social preferences as governed by $\sigma_2 > 1$.

4.5.5 How should capital income tax revenues be redistributed?

Because a capital income tax lowers final output, I isolate the tax rate's impact and investigate only the change in the after-tax labor income, net of output and relative to the zero-capital income tax.[11] I investigate the optimal redistribution of tax revenue under the baseline calibration in Figure 4.4. Figure 4.4 visualizes how a

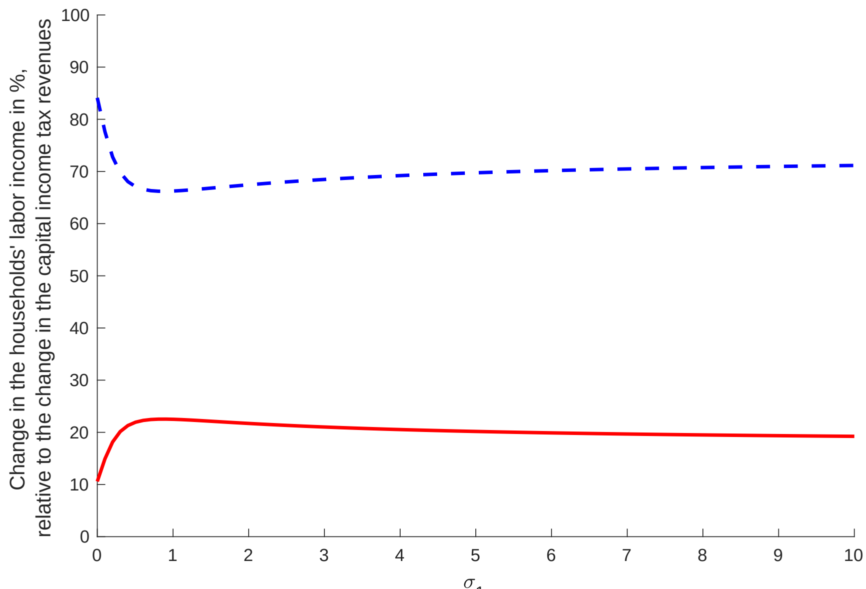


FIGURE 4.4. Change in the households' labor income due to capital taxation.

Notes: **Blue dashed line**: Decrease in the labor income tax, paid by a saver. **Red solid line**: Increase in the labor income subsidy, received by a worker.

single worker's and a single saver's labor income changes due to the capital income tax revenue, relative to a scenario without capital income taxation. The red solid line denotes by how much the subsidies' value paid to a worker increases due to the received capital income tax revenues: each worker receives a $10\% - 22.5\%$ higher labor income (after taxes). The blue dashed line denotes by how much the saver's labor income tax burden decreases: a saver has $66\% - 84\%$ greater (after taxes)

[11] The change in a saver's labor income due to capital income taxation is $(\Delta(1 - \tau^{L_s})\theta_{l_s})/(\Delta\tau^K\theta_k)$. The change in a worker's labor income due to capital income taxation is $(\Delta(1 - \tau^{L_w})\theta_{l_w}/\lambda)/(\Delta\tau^K\theta_k)$. Changes are given relative to the zero-capital income taxation equilibrium.

labor income. These changes are triggered by a ca. 10% capital income tax due to changing $\sigma_2 = 1$ to $\sigma_2 = 2$. A worker receives less tax revenue for two reasons.

A worker receives less tax revenues because (i) a worker responds more elastic to the tax rate than a saver ($\kappa < \hat{\kappa}$), so a worker faces higher private costs from adjusting his labor supply than a saver. Also, (ii) the worker's wage is lower than the savers' wage ($\theta_{l_w}/\lambda < \theta_{l_s}$). Allocating one dollar of tax revenue to the worker triggers thus greater labor supply responses (and induces thus greater private costs from labor supply) because the worker's wage rate changes more than the saver's wage rate. Due to (i) and (ii), the worker benefits less than a saver from one dollar of tax revenue—and the planner allocates more tax revenue to the saver. If the planner is very intratemporal inequality averse, for instance a "Rawlsian" planner with a high σ_1 value, another mechanism becomes relevant. A Rawlsian planner compensates the saver for his capital income losses due to the capital income tax. To make both households "equally" well-off, the planner then allocates relatively more tax revenues to a saver.

4.5.6 How robust is the capital tax revenues' re-distribution pattern?

I vary the labor supply elasticities and labor value shares in production and re-calibrate $\hat{X}_0$ (and X_0) to guarantee $\hat{L} = 1$ (and $L = 1$) in the zero-taxation baseline. Figure 4.5 plots how a single worker's labor income changes when introducing a capital income tax of ca. 10%. The figure provides two insights: (i) the re-distribution pattern of capital income tax revenues depends foremost on the households' labor supply elasticity parameters. The value shares of the households' labor supply in production play less of a role. (ii) A saver receives at least $2x$ more tax revenues than a worker for a wide range of σ_1, κ, $\hat{\kappa}$, θ_{l_s}, θ_{l_w}.

THE HOUSEHOLDS' LABOR SUPPLY ELASTICITIES MATTER A LOT.— The black solid line indicates the baseline calibration in which the saver supplies labor less

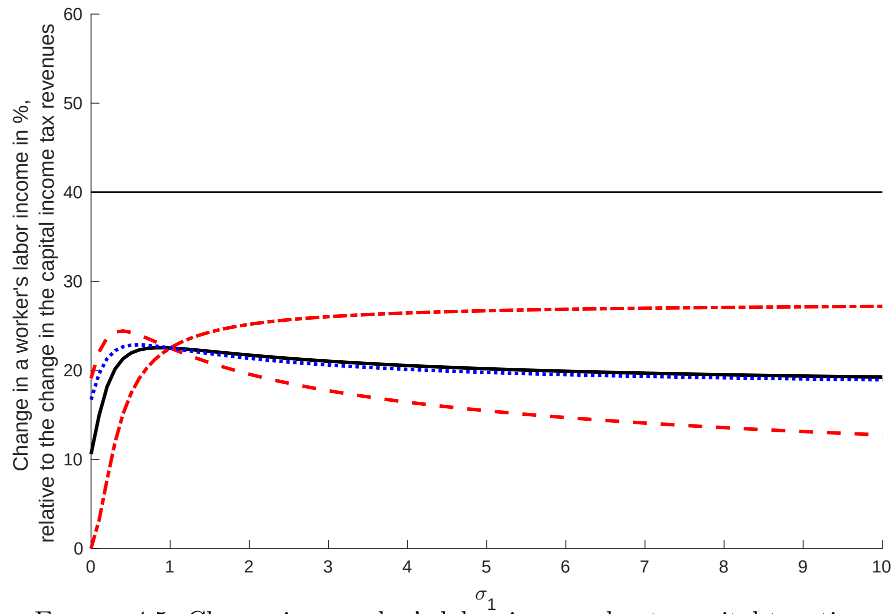


FIGURE 4.5. Change in a worker's labor income due to capital taxation

Notes: **Black horizontal line**: Threshold under which all households' receive the equal shares of capital income tax revenues. **Black solid line**: Baseline calibration, savers respond at the intensive margin ($\hat{\kappa} = 2.6667$) and workers respond at the intensive and extensive margin ($\kappa = 1.1429$). **Blue dotted line**: Basline labor supply elasticities ($\hat{\kappa} = 2.6667$, $\kappa = 1.1429$) plus a worker and a saver have the same pre-tax labor income ($\theta_l = \theta_{l_w}/\lambda = 0.2448$). **Red dashed line**: Savers and workers respond both at the intensive and extensive margin ($\hat{\kappa} = \kappa = 1.1429$). **Red dashed-dotted line**: Workers respond at the intensive and extensive margin ($\kappa = 1.1429$) and savers supply labor perfectly inelastically ($\hat{\kappa} = 1000$).

elastic than each worker ($2.6667 = \hat{\kappa} > \kappa = 1.1429$). For the red dashed-dotted line, I impose a perfectly inelastic labor supply ($1000 = \hat{\kappa}$). At $\sigma_1 = 0$, the black solid line ($2.6667 = \hat{\kappa}$) runs above the red dashed-dotted solid line ($1000 = \hat{\kappa}$). But why does an intratemporal inequality neutral planner allocate *less* tax revenue to the worker when the saver supplies labor *less elastic*? A saver faces no private costs from labor supply when supplying labor *perfectly inelastic* and allocating one dollar to a saver makes the saver better-off by exactly one dollar. An intratemporal inequality neutral planner thus redistributes all capital tax revenue back to the saver leaving nothing behind for the worker, i.e. the red dashed-dotted solid line crosses zero at $\sigma_1 = 0$.

But what if the planner is intratemporal inequality averse, i.e. $\sigma_1 \gg 0$? For simplicity, assume a Rawlsian planner with $\sigma_1 = \infty$. Note that the red dashed-dotted line ($\hat{\kappa} = 1000$) runs now above the black solid line ($\hat{\kappa} = 2.6667$). So the planner allocates *relatively more* tax revenue to a worker if a saver supplies labor *less elastic*. The intuition is that the Rawslian planner's wants to compensate the saver for his losses from the capital income tax. A saver faces zero private costs from labor supply when $\hat{\kappa} = 1000$, and a planner needs minimal resources to compensate the saver for his losses. A large share of the tax revenues can then be allocated to a worker.

THE HOUSEHOLDS' LABOR PRODUCTIVITIES MATTER LESS.— Figure 4.5 also shows that (i) increasing the worker's labor income (i.e. the blue dotted line with equal household pre-tax labor income) or (ii) increasing the saver's labor supply elasticity (i.e. the red dashed line with equal labor supply elasticities) yields similar changes in the redistribution pattern. The logic for (i) is that a higher worker's labor value share increases the worker's wage. As a result, the worker's after-tax labor income changes less when the worker receives one dollar of tax revenue and the worker faces less private costs from adjusting his labor supply, and the planner allocates more tax revenue to the worker. The logic for (ii) is that the worker's private costs from labor supply are relatively smaller when savers supply labor more elastically. Allocating tax revenues to the worker is then less costly because the worker adjusts his labor supply relatively less, and the planner allocates more tax revenues to the worker. Numerically, the pattern changes stronger with the labor supply elasticities than with the labor value shares in production because the labor supply elasticities govern non-linearities in the private costs of labor supply. The labor value shares in production impact the private costs of labor supply, however, only linearly, and are thus smaller.

ROBUSTNESS IN THE MAGNITUDE'S SIZE.— Next, I come to (ii) that a saver receives at least $2x$ more tax revenues than a worker for a wide range of σ_1, κ, $\hat{\kappa}$, θ_{l_s}, θ_{l_w}. The black horizontal line at 40% denotes the threshold under which a planner re-allocates the capital tax revenue equally across all households.[12] Each households' labor income changes then identically with the capital tax. Yet, the worker's labor income subsidy increases by less than 40% across all scenarios, so a planner allocates less tax revenues to a single worker than to a single saver. In fact, a single saver receives at least $2x$ as much tax revenues than a single saver depending on the intratemporal inequality aversion, labor supply elasticities and labor shares in production. For the baseline calibration and given an inequality averse planner with $\sigma_1 \geq 1$, a single saver receives ca. $3x$ as much tax revenues.[13]

I also vary the capital income tax rates covering a range from 1% to above 50% (See the figures in Appendix 4.8). The redistribution pattern does not change. Overall, the result that a worker receives significantly less tax revenue than a saver is robust for a wide range of σ_1, σ_2, κ, $\hat{\kappa}$, θ_{l_s}, θ_{l_w}.

4.6 Conclusion

That a utilitarian SWF ignores equity considerations is well documented. It is therefore surprising that many economic models on optimal tax theory adopt a utilitarian perspective, thus leaving insufficient room for the analysis of social equity concerns. This paper recognizes that the perception of justice is subjective and depends on social norms, and that societies differ in these norms—for example, the Swedish society is more egalitarian than the libertarian society of the United States. The aim of this paper is hence to investigate how the optimal taxation of households' capital and labor income depends on this heterogeneous perception of inequality. The paper also investigates how capital income tax revenues are returned back to the households.

In this framework, I distinguish between two types of households. Consistent

[12] The household-size-weighted equal share amounts to 40% because $1/(1+\lambda) = 1/2.5 = 0.4$. All workers together receive $\lambda/(1+\lambda) = 60\%$ of the revenues, a single worker receives $1/(1+\lambda) = 40\%$ of the revenues.

[13] If a worker's labor income increases, for example, by 20%, the labor income to all workers increases by $1.5 * 20\% = 35\%$. A saver's labor income increases then by $100\% - 35\% = 65\%$. Consequently, a single saver receives $65\%/20\% = 3.25x$ more tax revenues than a single worker.

with empirical studies, savers are intertemporally more elastic than workers with respect to consumption. Savers demand then lower returns to capital and thus own all of it. Workers consume all of their disposable labor income. This captures the stylized fact of concentrated capital ownership. Moreover, the framework can account for various social preference structures reflecting libertarian, egalitarian, and utilitarian principles of justice.

The results shed new light on policy recommendations for optimal redistributive taxation to address social equity concerns. I show how the redistribution of income depends on the SWF. First, it is surprising that economists have so far overlooked the fact that the optimal taxation of capital income is related to how a planner substitutes consumption over time. I find that capital income is not taxed under a utilitarian SWF in which the planner (implicitly) adopts the savers' IES to value growth-related consumption gains. However, taxing capital income is optimal if, for example, the planner adopts the workers' IES instead of the savers' IES. The tax makes capital investment less attractive and capital accumulates as workers prefer (and no longer as savers prefer). Second, applying a utilitarian SWF greatly affects the degree of income redistribution among households. However, less (or more) redistribution is equally well justifiable under a more libertarian (or egalitarian) SWF. For instance, in an economy calibrated to the EU-28, a worker may receive a zero labor income subsidy or more than a 100% labor income subsidy depending on the planner's intratemporal inequality aversion. Third, I also show that capital income tax revenues are largely returned to savers, regardless of the planner's equity concerns. Empirically, a saver supplies labor less elastic than a worker, so a saver benefits more from a dollar of capital income tax revenue than a worker. Numerically, a saver receives then at least twice as much capital income tax revenues than a worker. The result is robust to a broad distribution of household labor income, households' labor supply elasticities and a wide range of social inequality aversion parameters.

It is, of course, an open question how social norms can be measured and translated into the inequality aversion parameters of the model. The focus of this paper is, however, to assess the policy implications of these parameters. To this end, I have considered broad ranges for the inequality aversion parameters to reflect a wide possible range of social norms.

4.7 Appendix A: Theoretical derivations and proofs

4.7.1 King-Plosser-Rebelo preferences

King-Plosser-Rebelo (KPR) preferences yield a constant labor supply when labor is the only income source. KPR preferences are given by

$$u(C_t, L_t) = \frac{\left(C_t(1-\psi L^{1+\kappa})\right)^{1-\sigma}}{1-\sigma},$$

where C_t denotes consumption and L_t denotes the supplied labor quantity. If labor is the only income source, the household maximizes $u(C_t, L_t)$ subject to the budget constraint $C_t = w_t(1-\tau_t)L_t$, where w_t is the wage and τ_t is the labor income tax rate

$$\max_{\{C_t, L_t\}} \quad u(C_t, L_t) + \lambda_t(-C_t + w_t(1-\tau_t)L_t).$$

The FOCs w.r.t. C_t and L_t yield

$$\begin{aligned} C_t : & \frac{1-\psi L_t^{1+\kappa}}{\left(C_t(1-\psi L_t^{1+\kappa})\right)^{\sigma}} = \lambda_t \\ L_t : & \frac{C_t(1+\kappa)\psi L_t^{\kappa}}{\left(C_t(1-\psi L_t^{1+\kappa})\right)^{\sigma}} = \lambda_t w_t(1-\tau_t). \end{aligned}$$

Combing the FOCs and multiplying with L_t yields

$$\frac{C_t(1+\kappa)\psi L_t^{1+\kappa}}{\left(C_t(1-\psi L_t^{1+\kappa})\right)^{\sigma}} = \frac{1-\psi L_t^{1+\kappa}}{\left(C_t(1-\psi L_t^{1+\kappa})\right)^{\sigma}} w_t(1-\tau_t)L_t,$$

which I re-arrange, using $C_t = w_t(1-\tau_t)L_t$, to

$$\psi(2+\kappa)L_t^{1+\kappa} = 1,$$

and L_t becomes a constant that is independent of the after-tax wage $w_t(1-\tau_t)$. In optimum, households supply thus a fixed labor quantity independent of the wage rate.

4.7.2 The solution to the households' problem

The savers' problem.—— Savers have preferences

$$u(C_t^s, \hat{L}_t) = \log\left(C_t^s - \hat{X}_t \hat{L}_t^{1+\hat{\kappa}}\right),$$

and solve

$$\begin{aligned}\max_{\{C_t^s, \hat{L}_t, K_{t+1}\}_{t=0}^{\infty}} \quad & \sum_{t=0}^{\infty} \left(\frac{1}{1+\zeta}\right)^t (u(C_t^s, \hat{L}_t) \\ & + \lambda_t(-C_t^s - K_{t+1} + (r_t(1-\tau_t^K) + 1 - \delta)K_t + w_t^s(1-\tau_t^{L_s})\hat{L}_t)) .\end{aligned}$$

The FOCs read

$$\begin{aligned} C_t^s :& \frac{1}{C_t^s - \hat{X}_t \hat{L}_t^{1+\hat{\kappa}}} = \lambda_t \\ K_{t+1} :& \lambda_t = \frac{1}{1+\zeta}\lambda_{t+1}\left(r_{t+1}(1-\tau_{t+1}^K) + 1 - \delta\right) \\ \hat{L}_t :& \frac{1}{C_t^s - \hat{X}_t \hat{L}_t^{1+\hat{\kappa}}}\hat{X}_t(1+\hat{\kappa})\hat{L}_t^{\hat{\kappa}} = \lambda_t w_t^s (1-\tau_t^{L_s}) .\end{aligned}$$

First, combining the FOCs for capital and consumption yields the intertemporal optimal Euler equation

$$\left(C_t^s - \hat{X}_t \hat{L}_t^{1+\hat{\kappa}}\right)^{-1} = \frac{1}{1+\zeta}\left(C_{t+1}^s - \hat{X}_{t+1}\hat{L}_{t+1}^{1+\hat{\kappa}}\right)^{-1}\left(r_{t+1}(1-\tau_{t+1}^K) + 1 - \delta\right) . \tag{4.20}$$

Second, combining the optimality condition for consumption and labor yields the intratemporal optimality condition. This equation denotes that the marginal utility from labor income equals the marginal dis-utility from labor supply.

$$\hat{X}_t(1+\hat{\kappa})\hat{L}_t^{\hat{\kappa}} = w_t^s(1-\tau_t^{L_s}) . \tag{4.21}$$

The equilibrium is then given by the firm's resource constraint and the firms' inter- and intratemporal optimality condition. Using the savers' inter- and intratemporal optimality conditions in the intertemporal budget constraint yields the implementability constraint

$$\sum_{t=0}^{\infty}\left(\frac{1}{1+\zeta}\right)^{t}\frac{C_t^s-(1+\hat{\kappa})\hat{X}_t\hat{L}_t^{1+\hat{\kappa}}}{C_t^s-\hat{X}_t\hat{L}_t^{1+\hat{\kappa}}}=\frac{1}{C_0^s-\hat{X}_0\hat{L}_0^{1+\hat{\kappa}}}(r_0(1-\tau_0^K)+1-\delta)K_0\,. \tag{4.22}$$

The savers' budget constraint and the implementability constraint are sufficient to describe the households behavior in the decentralized economy.

The balanced growth path version is obtained as follows. Re-arranging the intratemporal optimality condition and multiplying by $\hat{L}_t$ yields the net-value of labor income in units of consumption, namely the difference of labor income adjusted for the dis-utility from labor supply:

$$w_t^s(1-\tau_t^{L_s})\hat{L}_t-\hat{X}_t\hat{L}_t^{1+\hat{\kappa}}=\hat{\kappa}\hat{X}_t\hat{L}_t^{1+\hat{\kappa}}\,. \tag{4.23}$$

On the balanced growth path $C_t^s-\hat{X}_t\hat{L}_t^{1+\hat{\kappa}}$ grows with g, so rearranging the intertemporal optimality condition yields the balanced growth path version of the capital price including taxes:

$$r_t(1-\tau_t^K)=(1+g)(1+\zeta)+\delta-1\,. \tag{4.24}$$

Using (4.23) and (4.24) in $C_t^s-\hat{X}_t\hat{L}_t^{1+\hat{\kappa}}$ yields

$$C_t^s-\hat{X}_t\hat{L}_t^{1+\hat{\kappa}}=\underbrace{\hat{\kappa}\hat{X}_t\hat{L}_t^{1+\hat{\kappa}}}_{=w_t^s(1-\tau_t^{L_s})\hat{L}_t-\hat{X}_t\hat{L}_t^{1+\hat{\kappa}}}-K_{t+1}+\left((1+g)(1+\zeta)\right)K_t$$

The balanced growth path requires $K_{t+1}=(1+g)K_t$, so I obtain

$$C_t^s-\hat{X}_t\hat{L}_t^{1+\hat{\kappa}}=\hat{\kappa}\hat{X}_t\hat{L}_t^{1+\hat{\kappa}}+\zeta K_{t+1}\,.$$

On the balanced growth path, consumption is then given by

$$C_t^s = \zeta K_{t+1} + w_t^s(1-\tau_t^{L_s})\hat{L}_t\,,$$

capital accumulates according to

$$K_{t+1} = (1+g)K_t\,,$$

and labor supply is given by the intratemporal optimality condition. On the balanced growth path and using the governments budget constraint and the firms' optimality conditions, $C_t^s - \hat{X}_t\hat{L}_t^{1+\hat{\kappa}}$ is given by

$$C_t^s - \hat{X}_t\hat{L}_t^{1+\hat{\kappa}} = \zeta(1+g)\frac{\theta_k}{r_t}Y_t + \frac{\hat{\kappa}}{1+\hat{\kappa}}\left(\theta_{l_s} + \tau_t^K\theta_k + \tau_t^{L_w}\theta_{l_w}\right)Y_t\,. \tag{4.25}$$

The workers' problem.—— Workers have preferences

$$u(C_t^w, L_t) = \frac{\left(C_t^w - X_t L_t^{1+\kappa}\right)^{1-\sigma} - 1}{1-\sigma}\,,$$

where $1/\sigma$ denotes the workers' intertemporal elasticity of substitution. Workers are intertemporally less elastic than savers ($\sigma > 1$) and have no access to capital markets. They solve

$$\begin{aligned}\max_{\{C_t^w, L_t\}_{t=0}^{\infty}} \quad \sum_{t=0}^{\infty}\left(\frac{1}{1+\zeta}\right)^t &(u(C_t^w, L_t)\\ &+ \lambda_t(-C_t^w + w_t^w(1-\tau_t^{L_w})L_t))\,.\end{aligned}$$

Combining the FOCs w.r.t. C_t^w and L_t yield

$$(1+\kappa)X_t L_t^{\kappa} = (1-\tau_t^{L_w})w_t^w\,.$$

On the balanced growth path and using $\theta_{l_w} Y_t = \lambda w_t^w L_t$, the expression $C_t^w - X_t L_t^{1+\kappa}$ is then given by

$$C_t^w - X_t L_t^{1+\kappa} = \frac{\kappa}{1+\kappa}(1 - \tau_t^{L_w})\frac{\theta_{l_w}}{\lambda} Y_t \,.$$

4.7.3 The solution to the social planner's problem

The planner's problem is set up in primal form and given by

$$
\begin{aligned}
& max_{\{K_{t+1},L_t,\hat{L}_t,Y_t,C_t^s,C_t^w\}_{t=0}^{\infty}} W \qquad (4.26)\\
& + \lambda_t^1 \left(-C_t^s - K_{t+1} + (1-\delta)K_t + (1-\tau_t^{L_s})\theta_{l_s} Y_t + (1-\tau_t^K)\theta_k Y_t\right)\\
& + \lambda_t^2 \left(-C_t^w + (1-\tau_t^{L_w})\frac{\theta_{l_w}}{\lambda} Y_t\right)\\
& + \mu^1 \left(-\sum_t^{\infty}\left(\frac{1}{1+\zeta}\right)^t \frac{C_t^s - (1+\hat{\kappa})\hat{X}_t \hat{L}_t^{1+\hat{\kappa}}}{C_t^s - \hat{X}_t \hat{L}_t^{1+\hat{\kappa}}} + \frac{1}{C_0^s - \hat{X}_0 \hat{L}_0^{1+\hat{\kappa}}}(r_0(1-\tau_0^K) + 1 - \delta)K_0\right)\\
& + \mu_t^2 \left(-(1-\tau_t^{L_w})\frac{\theta_{l_w}}{\lambda} Y_t + (1-\kappa)X_t L_t^{1+\kappa}\right)\\
& + \mu_t^G \left(\tau_t^K \theta_k Y_t + \tau_t^{L_w}\theta_{l_w} Y_t + \tau_t^{L_s}\theta_{l_s} Y_t\right) + \mu_t^Y \left(-Y_t + Y(K_t, \lambda L_t, \hat{L}_t)\right),
\end{aligned}
$$

and given an initial capital stock K_0. W is defined as

$$
W = \sum_{t=0}^{\infty}\left(\frac{1}{1+\zeta}\right)^t \frac{\left[\left(C_t^s - \hat{X}_t \hat{L}_t^{1+\hat{\kappa}}\right)^{1-\sigma_1} + \lambda\left(C_t^w - X_t L_t^{1+\kappa}\right)^{1-\sigma_1}\right]^{\frac{1-\sigma_2}{1-\sigma_1}} - 1}{1-\sigma_2}.
$$

$Y(K_t, \lambda L_t, \hat{L}_t)$ is defined in (4.9). λ_t^1 is the Lagrangian multiplier of the savers' intertemporal budget constraint (4.6) and μ^1 is the Lagrangian multiplier of the savers' implementability constraint (4.8). λ_t^2 denotes the Lagrangian multiplier of the workers' budget constraint (4.2) and μ_t^2 denotes the Lagrangian multiplier on the workers' labor supply constraint (4.3). μ_t^G is the Lagrangian multiplier on governments budgets constraint (4.11) and μ_t^Y is the Lagrangian multiplier on the technology constraint (4.9).

I derive the optimal tax rates on labor and capital income by comparing the household's FOCs with the solution to the planers' problem. The FOCs with respect to C_t^s and C_t^w yield

$$C_t^s : \left[\left(C_t^s - \hat{X}_t \hat{L}_t^{1+\hat{\kappa}}\right)^{1-\sigma_1} + \lambda\left(C_t^w - X_t L_t^{1+\kappa}\right)^{1-\sigma_1}\right]^{\frac{1-\sigma_2}{1-\sigma_1}-1} \left(C_t^s - \hat{X}_t \hat{L}_t^{1+\hat{\kappa}}\right)^{-\sigma_1} \tag{4.27}$$

$$= \lambda_t^1 + \mu_1 \left(\frac{\hat{\kappa}\hat{X}_t \hat{L}_t^{1+\kappa}}{C_t^s - \hat{X}_t \hat{L}_t^{1+\hat{\kappa}}}\right)$$

$$C_t^w : \left[\left(C_t^s - \hat{X}_t \hat{L}_t^{1+\hat{\kappa}}\right)^{1-\sigma_1} + \lambda\left(C_t^w - X_t L_t^{1+\kappa}\right)^{1-\sigma_1}\right]^{\frac{1-\sigma_2}{1-\sigma_1}-1} \left(C_t^w - X_t L_t^{1+\kappa}\right)^{-\sigma_1} \lambda \tag{4.28}$$

$$= \lambda_t^2 \, .$$

The FOCs with respect to $\hat{L}_t$ and L_t, multiplied by $\hat{L}_t$ and L_t respectively, yields:

$$\hat{L}_t : \left[\left(C_t^s - \hat{X}_t \hat{L}_t^{1+\hat{\kappa}}\right)^{1-\sigma_1} + \lambda\left(C_t^w - X_t L_t^{1+\kappa}\right)^{1-\sigma_1}\right]^{\frac{1-\sigma_2}{1-\sigma_1}-1} \left(C_t^s - \hat{X}_t \hat{L}_t^{1+\hat{\kappa}}\right)^{-\sigma_1} \times \tag{4.29}$$

$$(1+\hat{\kappa})\hat{X}_t \hat{L}_t^{1+\hat{\kappa}} = \mu_t^Y \theta_{l_s} Y_t + \mu_1 \left(\frac{\hat{\kappa}(1+\hat{\kappa})\hat{X}_t \hat{L}_t^{1+\hat{\kappa}}}{\left(C_t^s - \hat{X}_t \hat{L}_t^{1+\hat{\kappa}}\right)^2} C_t^s\right)$$

$$L_t : \lambda_t^2 (1-\tau_t^{L_w})\theta_{l_w} Y_t = \mu_t^2 (1+\kappa)(1-\tau_t^{L_w})\theta_{l_w} Y_t + \lambda_t^Y \theta_{l_w} Y_t \, , \tag{4.30}$$

where the optimality condition for L_t was combined with the workers' labor supply constraint and with the optimality condition for C_t^w. Multiplying (4.27) with $(1+$

$\hat{\kappa})C_t^s$, subtracting (4.29) and re-arranging yields

$$\left[\left(C_t^s - \hat{X}_t \hat{L}_t^{1+\hat{\kappa}}\right)^{1-\sigma_1} + \lambda \left(C_t^w - X_t L_t^{1+\kappa}\right)^{1-\sigma_1}\right]^{\frac{1-\sigma_2}{1-\sigma_1}-1} \left(C_t^s - \hat{X}_t \hat{L}_t^{1+\hat{\kappa}}\right)^{1-\sigma_1} \tag{4.31}$$
$$= \lambda_t^1 C_t^s - \mu_t^Y \frac{\theta_{l_s} Y_t}{1+\hat{\kappa}} \,.$$

The FOC w.r.t Y_t reads

$$\mu_t^Y = \lambda_t^1 (1-\tau_t^{L_s})\theta_{l_s} + \lambda_t^2 (1-\tau_t^{L_w}) \frac{\theta_{l_w}}{\lambda} - \mu_t^2 (1-\tau_t^{L_w}) \frac{\theta_{l_w}}{\lambda} + \lambda_t^1 (1-\tau_t^K)\theta_k \tag{4.32}$$

which I re-arrange using (4.28) and (4.30) to

$$\mu_t^Y = \lambda_t^1 (1-\tau_t^{L_s})\theta_{l_s} + \frac{\kappa}{1+\kappa} \lambda_t^2 (1-\tau_t^{L_w}) \frac{\theta_{l_w}}{\lambda} + \frac{1}{1+\kappa} \mu_t^Y \theta_{l_w} \,. \tag{4.33}$$

The FOC with respect to K_{t+1} reveals

$$\begin{aligned} \lambda_t^1 = & \frac{1}{1+\zeta} \lambda_{t+1}^1 \left(1 - \delta + \frac{\partial Y(K_{t+1}, \lambda L_{t+1}, \hat{L}_{t+1})}{\partial K_{t+1}}\right) \\ & + \frac{1}{1+\zeta} \mu_{t+1}^Y \frac{\partial Y(K_{t+1}, \lambda L_{t+1}, \hat{L}_{t+1})}{\partial K_{t+1}} \,. \end{aligned} \tag{4.34}$$

I am know equipped to derive the tax rates' optimality conditions. I first derive the optimality conditions for the savers' capital tax rate, I then proceed to the workers' labor tax rate and the savers' labor tax rate.

Optimality condition for the savers' capital tax rate.—— I obtain the optimal capital tax rate on the balanced growth path using (4.33) in (4.34) and comparing it to the savers' intertemporal optimality condition (4.20). This yields

$\lambda_t^1 = \frac{1}{1+\zeta}\lambda_{t+1}^1 \left(1 - \delta + \frac{\partial Y(K_{t+1}, \lambda L_t t+1 \hat{L}_{t+1})}{\partial K_{t+1}}\right)$ and I obtain

$$\frac{\partial Y(K, \lambda L, \hat{L})}{\partial K} = r = (1+g)^{\sigma_2}(1+\zeta) + \delta - 1\,.$$

The decentralized equilibrium, however, reads in (4.24)

$$r_t(1-\tau_t^K) = (1+g)(1+\zeta) + \delta - 1\,,$$

so the balanced-growth path capital income tax rate is given by

$$\tau^K = (1+\zeta)\frac{(1+g)^{\sigma_2} - (1+g)}{r}\,. \tag{4.35}$$

(4.33) holds if $\mu_t^Y = \lambda_t^1$ and (4.33) can be further simplified to

$$\lambda_t^1\left(\theta_{l_w}\frac{\kappa}{1+\kappa} - \tau_t^{l_w}\theta_{l_w}\right) = \frac{\kappa}{1+\kappa}\lambda_t^2(1-\tau_t^{L_w})\frac{\theta_{l_w}}{\lambda}\,, \tag{4.36}$$

where λ_t^1 is given by

$$\left[\left(C_t^s - \hat{X}_t\hat{L}_t^{1+\hat{\kappa}}\right)^{1-\sigma_1} + \lambda\left(C_t^w - X_tL_t^{1+\kappa}\right)^{1-\sigma_1}\right]^{\frac{1-\sigma_2}{1-\sigma_1}-1}\left(C_t^s - \hat{X}_t\hat{L}_t^{1+\hat{\kappa}}\right)^{-\sigma_1}\times$$
$$\frac{\left(C_t^s - \hat{X}_t\hat{L}_t^{1+\hat{\kappa}}\right)}{C_t^s - \frac{\theta_{l_s}Y_t}{1+\hat{\kappa}}} = \lambda_t^1\,, \tag{4.37}$$

and λ_t^2 is given by

$$\left[\left(C_t^s - \hat{X}_t\hat{L}_t^{1+\hat{\kappa}}\right)^{1-\sigma_1} + \lambda\left(C_t^w - X_tL_t^{1+\kappa}\right)^{1-\sigma_1}\right]^{\frac{1-\sigma_2}{1-\sigma_1}-1}\left(C_t^w - X_tL_t^{1+\kappa}\right)^{-\sigma_1}\lambda$$
$$= \lambda_t^2\,. \tag{4.38}$$

I can now obtain the optimality condition for the workers labor tax rate.

Optimality condition for the workers' labor tax rate.—— I use the balanced growth path versions of savers' consumption and re-arrange (4.36) using (4.37), (4.38) and the governments' budget constraint. The optimal workers' labor income tax on the balanced growth path (and thus dropping the time subscript) is then given by:

$$\begin{aligned}&\left(\zeta(1+g)\frac{\theta_k}{r}+\frac{\hat{\kappa}}{1+\hat{\kappa}}\left(\theta_{l_s}+\tau^K\theta_k+\tau^{L_w}\theta_{l_w}\right)\right)^{1-\sigma_1}\left(1-\frac{1+\kappa}{\kappa}\tau^{L_w}\right)\\ =&\left(\frac{\kappa}{1+\kappa}(1-\tau^{L_w})\frac{\theta_{l_w}}{\lambda}\right)^{-\sigma_1}\left(1-\tau^{L_w}\right)\left(\zeta(1+g)\frac{\theta_k}{r}+\frac{\hat{\kappa}}{1+\hat{\kappa}}\theta_{l_s}+\tau^K\theta_k+\tau^{L_w}\theta_{l_w}\right).\end{aligned} \tag{4.39}$$

Or alternatively

$$\begin{aligned}&\left(\frac{\zeta(1+g)\frac{\theta_k}{r}+\frac{\hat{\kappa}}{1+\hat{\kappa}}\left(\theta_{l_s}+\tau^K\theta_k+\tau^{L_w}\theta_{l_w}\right)}{\left(\frac{\kappa}{1+\kappa}(1-\tau^{L_w})\frac{\theta_{l_w}}{\lambda}\right)}\right)^{\sigma_1}\\ =&\frac{\left(\zeta(1+g)\frac{\theta_k}{r}+\frac{\hat{\kappa}}{1+\hat{\kappa}}\left(\theta_{l_s}+\tau^K\theta_k+\tau^{L_w}\theta_{l_w}\right)\right)\left(1-\frac{1+\kappa}{\kappa}\tau^{L_w}\right)}{\left(\zeta(1+g)\frac{\theta_k}{r}+\frac{\hat{\kappa}}{1+\hat{\kappa}}\theta_{l_s}+\tau^K\theta_k+\tau^{L_w}\theta_{l_w}\right)(1-\tau^{L_w})}.\end{aligned} \tag{4.40}$$

Optimality condition for the savers' labor tax rate.—— Given the savers' capital tax rate in (4.35) and the workers' labor tax rate from (4.39), the savers' optimal labor tax rate is given by the governments budget constraint:

$$\tau^K\theta_k+\tau^{L_w}\theta_{l_w}+\tau^{L_s}\theta_{l_s}=0\,. \tag{4.41}$$

4.7.4 Proof to Proposition 9

PROOF: The optimal workers' labor income tax rate is given by

$$\left(\frac{\zeta(1+g)\frac{\theta_k}{r}+\frac{\hat{\kappa}}{1+\hat{\kappa}}\left(\theta_{l_s}+\tau^K\theta_k+\tau^{L_w}\theta_{l_w}\right)}{\left(\frac{\kappa}{1+\kappa}(1-\tau^{L_w})\frac{\theta_{l_w}}{\lambda}\right)}\right)^{\sigma_1}$$
$$=\frac{\left(\zeta(1+g)\frac{\theta_k}{r}+\frac{\hat{\kappa}}{1+\hat{\kappa}}\left(\theta_{l_s}+\tau^K\theta_k+\tau^{L_w}\theta_{l_w}\right)\right)\left(1-\frac{1+\kappa}{\kappa}\tau^{L_w}\right)}{\left(\zeta(1+g)\frac{\theta_k}{r}+\frac{\hat{\kappa}}{1+\hat{\kappa}}\theta_{l_s}+\tau^K\theta_k+\tau^{L_w}\theta_{l_w}\right)(1-\tau^{L_w})},$$

which can be re-arranged to

$$\begin{aligned}F=\ &\sigma_1\log\left(\zeta(1+g)\frac{\theta_k}{r}+\frac{\hat{\kappa}}{1+\hat{\kappa}}\left(\theta_{l_s}+\tau^K\theta_k+\tau^{L_w}\theta_{l_w}\right)\right)\\&-\sigma_1\log\left(\left(\frac{\kappa}{1+\kappa}(1-\tau^{L_w})\frac{\theta_{l_w}}{\lambda}\right)\right)\\&-\log\left(\left(\zeta(1+g)\frac{\theta_k}{r}+\frac{\hat{\kappa}}{1+\hat{\kappa}}\left(\theta_{l_s}+\tau^K\theta_k+\tau^{L_w}\theta_{l_w}\right)\right)\left(1-\frac{1+\kappa}{\kappa}\tau^{L_w}\right)\right)\\&+\log\left(\left(\zeta(1+g)\frac{\theta_k}{r}+\frac{\hat{\kappa}}{1+\hat{\kappa}}\theta_{l_s}+\tau^K\theta_k+\tau^{L_w}\theta_{l_w}\right)\left(1-\tau^{L_w}\right)\right)=0\,.\end{aligned}\tag{4.42}$$

I apply the implicit function theorem to investigate how τ^{L_w} changes with σ_1:

$$\frac{\partial\tau^{L_w}}{\partial\sigma_1}=-\frac{\frac{\partial F}{\partial\sigma_1}}{\frac{\partial F}{\partial\tau^{L_w}}}\,.\tag{4.43}$$

From (4.42) follows $\frac{\partial F}{\partial \sigma_1} \geq 0$

$$
\begin{aligned}
&\frac{\partial F}{\partial \sigma_1} \geq 0 \\
&\Leftrightarrow \Bigg[\log \left(\zeta(1+g)\frac{\theta_k}{r} + \frac{\hat{\kappa}}{1+\hat{\kappa}} \left(\theta_{l_s} + \tau^K \theta_k + \tau^{L_w} \theta_{l_w} \right) \right) \\
&\quad - \log \left(\left(\frac{\kappa}{1+\kappa}(1-\tau^{L_w})\frac{\theta_{l_w}}{\lambda} \right) \right) \Bigg] \geq 0 \\
&\Leftrightarrow \frac{1}{\sigma_1} \log \left(\left(\zeta(1+g)\frac{\theta_k}{r} + \frac{\hat{\kappa}}{1+\hat{\kappa}} \left(\theta_{l_s} + \tau^K \theta_k + \tau^{L_w} \theta_{l_w} \right) \right) \left(1 - \frac{1+\kappa}{\kappa} \tau^{L_w} \right) \right) \\
&- \frac{1}{\sigma_1} \log \left(\left(\zeta(1+g)\frac{\theta_k}{r} + \frac{\hat{\kappa}}{1+\hat{\kappa}} \theta_{l_s} + \tau^K \theta_k + \tau^{L_w} \theta_{l_w} \right) \left(1 - \tau^{L_w} \right) \right) \geq 0
\end{aligned}
$$

$$
\begin{aligned}
&\Leftrightarrow \left(\zeta(1+g)\frac{\theta_k}{r} + \frac{\hat{\kappa}}{1+\hat{\kappa}} \left(\theta_{l_s} + \tau^K \theta_k + \tau^{L_w} \theta_{l_w} \right) \right) \left(1 - \frac{1+\kappa}{\kappa} \tau^{L_w} \right) \\
&\quad > \left(\zeta(1+g)\frac{\theta_k}{r} + \frac{\hat{\kappa}}{1+\hat{\kappa}} \theta_{l_s} + \tau^K \theta_k + \tau^{L_w} \theta_{l_w} \right) \left(1 - \tau^{L_w} \right) \\
&\Leftrightarrow \frac{\left(\zeta(1+g)\frac{\theta_k}{r} + \frac{\hat{\kappa}}{1+\hat{\kappa}} \left(\theta_{l_s} + \tau^K \theta_k + \tau^{L_w} \theta_{l_w} \right) \right) \left(1 - \frac{1+\kappa}{\kappa} \tau^{L_w} \right)}{\left(\zeta(1+g)\frac{\theta_k}{r} + \frac{\hat{\kappa}}{1+\hat{\kappa}} \theta_{l_s} + \tau^K \theta_k + \tau^{L_w} \theta_{l_w} \right) \left(1 - \tau^{L_w} \right)} \\
&\quad = \left(\frac{\zeta(1+g)\frac{\theta_k}{r} + \frac{\hat{\kappa}}{1+\hat{\kappa}} \left(\theta_{l_s} + \tau^K \theta_k + \tau^{L_w} \theta_{l_w} \right)}{\left(\frac{\kappa}{1+\kappa}(1-\tau^{L_w})\frac{\theta_{l_w}}{\lambda} \right)} \right)^{\sigma_1} \geq 1 \, .
\end{aligned}
$$

$\frac{\partial F}{\partial \sigma_1} \geq 0$, follows because $\zeta(1+g)\frac{\theta_k}{r} + \frac{\hat{\kappa}}{1+\hat{\kappa}} \theta_{l_s}(1-\tau^{L_s}) \geq \frac{\kappa}{1+\kappa}(1-\tau^{L_w})\frac{\theta_{l_w}}{\lambda}$ is true for all σ_1, given $\hat{\kappa} > \kappa$, $\theta_{l_s} > \theta_{l_w}/\lambda$ and that only savers have a capital income.

Further, from (4.42) follows $\frac{\partial F}{\partial \tau^{L_w}} > 0$

$$\frac{\partial F}{\partial \tau^{L_w}} = \underbrace{\sigma_1 \left[\frac{\frac{\hat{\kappa}}{1+\hat{\kappa}}\theta_{l_w}}{\zeta(1+g)\frac{\theta_k}{r} + \frac{\hat{\kappa}}{1+\hat{\kappa}}\left(\theta_{l_s} + \tau^K\theta_k + \tau^{L_w}\theta_{l_w}\right)} + \frac{\frac{\kappa}{1+\kappa}\frac{\theta_{l_w}}{\lambda}}{\left(\frac{\kappa}{1+\kappa}(1-\tau^{L_w})\frac{\theta_{l_w}}{\lambda}\right)} \right]}_{\geq 0} \tag{4.44}$$

$$+ \underbrace{\frac{\frac{1+\kappa}{\kappa}}{1 - \frac{1+\kappa}{\kappa}\tau^{L_w}} - \frac{1}{1-\tau^{L_w}}}_{>0 \text{ if } \tau^{L_w}<0}$$

$$+ \frac{\theta_{l_w}}{\zeta(1+g)\frac{\theta_k}{r} + \frac{\hat{\kappa}}{1+\hat{\kappa}}\theta_{l_s} + \tau^K\theta_k + \tau^{L_w}\theta_{l_w}}$$

$$- \frac{\frac{\hat{\kappa}}{1+\hat{\kappa}}\theta_{l_w}}{\zeta(1+g)\frac{\theta_k}{r} + \frac{\hat{\kappa}}{1+\hat{\kappa}}\left(\theta_{l_s} + \tau^K\theta_k + \tau^{L_w}\theta_{l_w}\right)} > 0\,.$$

From (4.43) and using $\frac{\partial F}{\partial \sigma_1} \geq 0$ and $\frac{\partial F}{\partial \tau^{L_w}} > 0$, it follows $\frac{\partial \tau^{L_w}}{\partial \sigma_1} < 0$ so that a planner subsidizes the workers labor income stronger if the planner is more intratemporal inequality averse. In fact, $\tau^{L_w} \leq 0$ always holds for all σ_1, σ_2. Taxing the workers' labor income is never optimal, because then the planner would transfer money to the saver—who are better off anyways implying a smaller marginal welfare from the saver's consumption (See Corollary 2).

4.7.5 Proof to Proposition 10

PROOF: I note that

$$\frac{\partial r}{\partial \sigma_2} = (1+\zeta)(1+g)^{\sigma_2}\log(1+g) > 0 \tag{4.45}$$

$$\frac{\partial \tau^K}{\partial \sigma_2} = \frac{1}{r}\frac{\partial r}{\partial \sigma_2}(1-\tau^K)\,. \tag{4.46}$$

Savers then receive more capital income tax revenues if:

$$\begin{aligned}&\frac{\partial\left(\theta_{l_s}+\tau^K\theta_k+\tau^{L_w}\theta_{l_w}\right)}{\partial\sigma_2} > \frac{\partial\left((1-\tau^{L_w})\frac{\theta_{l_w}}{\lambda}\right)}{\partial\sigma_2}\\ \Leftrightarrow &\frac{\lambda}{1+\lambda}\frac{\partial\tau^K}{\partial\sigma_2}\theta_k > \left(\frac{\partial-\tau^W}{\partial\sigma_2}\right)\theta_{l_w}\,.\end{aligned} \tag{4.47}$$

Further, I obtain the expression $\left(\frac{\partial-\tau^W}{\partial\sigma_2}\right)$ from 4.7.6 and I re-arrange the inequality (4.47) to:

$$\begin{aligned}&\theta_k\frac{1}{r}\frac{\partial r}{\partial\sigma_2}(1-\tau^K)\\ >&\frac{1+\lambda}{\lambda}\theta_{l_w}\left(\frac{1}{r}\frac{\partial r}{\partial\sigma_2}\right)\\ &\times\frac{(\sigma_1-1)\frac{-\zeta(1+g)\frac{\theta_k}{r}+\frac{\hat\kappa}{1+\hat\kappa}\theta_k(1-\tau^K)}{\zeta(1+g)\frac{\theta_k}{r}+\frac{\hat\kappa}{1+\hat\kappa}\theta_{l_s}(1-\tau^{L_s})}+\frac{-\zeta(1+g)\frac{\theta_k}{r}+\theta_k(1-\tau^K)}{\zeta(1+g)\frac{\theta_k}{r}+\frac{\hat\kappa}{1+\hat\kappa}\theta_{l_s}-\tau^{L_s}\theta_{l_s}}}{(\sigma_1-1)\left(\frac{\frac{\hat\kappa}{1+\hat\kappa}\theta_{l_w}}{\zeta(1+g)\frac{\theta_k}{r}+\frac{\hat\kappa}{1+\hat\kappa}\theta_{l_s}(1-\tau^{L_s})}+\frac{1}{1-\tau^{L_w}}\right)+\frac{1}{\frac{\kappa}{1+\kappa}-\tau^{L_w}}+\frac{\theta_{l_w}}{\zeta(1+g)\frac{\theta_k}{r}+\frac{\hat\kappa}{1+\hat\kappa}\theta_{l_s}-\tau^{L_s}\theta_{l_s}}}\,.\end{aligned}$$

From there I show that (4.47) holds with $\sigma_1 = 1$, using $\zeta > 0$. The LHS and RHS in (4.47) simplify to

$$\theta_k(1-\tau^K) > \frac{1+\lambda}{\lambda}\theta_{l_w}\frac{-\zeta(1+g)\frac{\theta_k}{r}+\theta_k(1-\tau^K)}{\zeta(1+g)\frac{\theta_k}{r}+\frac{\hat{\kappa}}{1+\hat{\kappa}}\theta_{l_s}-\tau^{L_s}\theta_{l_s}}$$
$$\Leftrightarrow\left(\frac{1}{\frac{\kappa}{1+\kappa}-\tau^{L_w}}+\frac{\theta_{l_w}}{\zeta(1+g)\frac{\theta_k}{r}+\frac{\hat{\kappa}}{1+\hat{\kappa}}\theta_{l_s}-\tau^{L_s}\theta_{l_s}}\right)\theta_k(1-\tau^K)$$
$$>\frac{1+\lambda}{\lambda}\theta_{l_w}\frac{-\zeta(1+g)\frac{\theta_k}{r}+\theta_k(1-\tau^K)}{\zeta(1+g)\frac{\theta_k}{r}+\frac{\hat{\kappa}}{1+\hat{\kappa}}\theta_{l_s}-\tau^{L_s}\theta_{l_s}}\,.$$

I further simplify using $\theta_k(1-\tau^K) > 0$

$$1 \geq \frac{\left(\frac{\kappa}{1+\kappa}-\tau^{L_w}\right)\frac{\theta_{l_w}}{\lambda}}{\left(\zeta(1+g)\frac{\theta_k}{r}+\frac{\hat{\kappa}}{1+\hat{\kappa}}\theta_{l_s}+\tau^K\theta_k+\tau^{L_w}\theta_{l_w}\right)}\,. \tag{4.48}$$

Because

$$\frac{\left(\frac{\kappa}{1+\kappa}-\tau^{L_w}\right)\frac{\theta_{l_w}}{\lambda}}{\left(\zeta(1+g)\frac{\theta_k}{r}+\frac{\hat{\kappa}}{1+\hat{\kappa}}\theta_{l_s}+\tau^K\theta_k+\tau^{L_w}\theta_{l_w}\right)}$$
$$=\left(\frac{\left(\frac{\kappa}{1+\kappa}\frac{\theta_{l_w}}{\lambda}(1-\tau^{L_w})\right)}{\zeta(1+g)\frac{\theta_k}{r}+\frac{\hat{\kappa}}{1+\hat{\kappa}}\left(\theta_{l_s}+\tau^K\theta_k+\tau^{L_w}\theta_{l_w}\right)}\right)^{1-\sigma_1}=1\,,$$

and if $\sigma_1 = 1$, I know that (4.48) hold with equality so (4.47) will hold with inequality, given $\zeta > 0$.

4.7.6 Derivation of $\partial\tau^{L_w}/\partial\sigma_2$

To obtain a closed form expression for $\partial\tau^{L_w}/\partial\sigma_2$, I first use (4.41) to obtain

$$
\begin{aligned}
F = {} & \sigma_1 \log\left(\zeta(1+g)\frac{\theta_k}{r} + \frac{\hat{\kappa}}{1+\hat{\kappa}}\left(\theta_{l_s} + \tau^K\theta_k + \tau^{L_w}\theta_{l_w}\right)\right) \\
& - \sigma_1 \log\left(\left(\frac{\kappa}{1+\kappa}(1-\tau^{L_w})\frac{\theta_{l_w}}{\lambda}\right)\right) \\
& - \log\left(\left(\zeta(1+g)\frac{\theta_k}{r} + \frac{\hat{\kappa}}{1+\hat{\kappa}}\left(\theta_{l_s} + \tau^K\theta_k + \tau^{L_w}\theta_{l_w}\right)\right)\left(1 - \frac{1+\kappa}{\kappa}\tau^{L_w}\right)\right) \\
& + \log\left(\left(\zeta(1+g)\frac{\theta_k}{r} + \frac{\hat{\kappa}}{1+\hat{\kappa}}\theta_{l_s} + \tau^K\theta_k + \tau^{L_w}\theta_{l_w}\right)\left(1-\tau^{L_w}\right)\right) = 0
\end{aligned}
\tag{4.49}
$$

I then apply the implicit function theorem to obtain:

$$
\frac{\partial\tau^{L_w}}{\partial\sigma_2} = -\frac{\frac{\partial F}{\partial\sigma_2}}{\frac{\partial F}{\partial\tau^{L_w}}}
\tag{4.50}
$$

I obtain two expressions:

$$
\frac{\partial F}{\partial\sigma_2} = (\sigma_1 - 1)\,\frac{\frac{1}{r}\frac{\partial r}{\partial\sigma_2}\left(-\zeta(1+g)\frac{\theta_k}{r} + \frac{\hat{\kappa}}{1+\hat{\kappa}}\theta_k(1-\tau^K)\right)}{\zeta(1+g)\frac{\theta_k}{r} + \frac{\hat{\kappa}}{1+\hat{\kappa}}\left(\theta_{l_s} + \tau^K\theta_k + \tau^{L_w}\theta_{l_w}\right)}
\tag{4.51}
$$

$$
+ \frac{\frac{1}{r}\frac{\partial r}{\partial\sigma_2}\left(-\zeta(1+g)\frac{\theta_k}{r} + \theta_k(1-\tau^K)\right)}{\zeta(1+g)\frac{\theta_k}{r} + \frac{\hat{\kappa}}{1+\hat{\kappa}}\theta_{l_s} + \tau^K\theta_k + \tau^{L_w}\theta_{l_w}}\,.
\tag{4.52}
$$

and

$$\frac{\partial F}{\partial \tau^{L_w}} = \sigma_1 \left[\frac{\frac{\hat{\kappa}}{1+\hat{\kappa}}\theta_{l_w}}{\zeta(1+g)\frac{\theta_k}{r} + \frac{\hat{\kappa}}{1+\hat{\kappa}}\left(\theta_{l_s} + \tau^K\theta_k + \tau^{L_w}\theta_{l_w}\right)} + \frac{\frac{\kappa}{1+\kappa}\frac{\theta_{l_w}}{\lambda}}{\left(\frac{\kappa}{1+\kappa}(1-\tau^{L_w})\frac{\theta_{l_w}}{\lambda}\right)} \right] \tag{4.53}$$

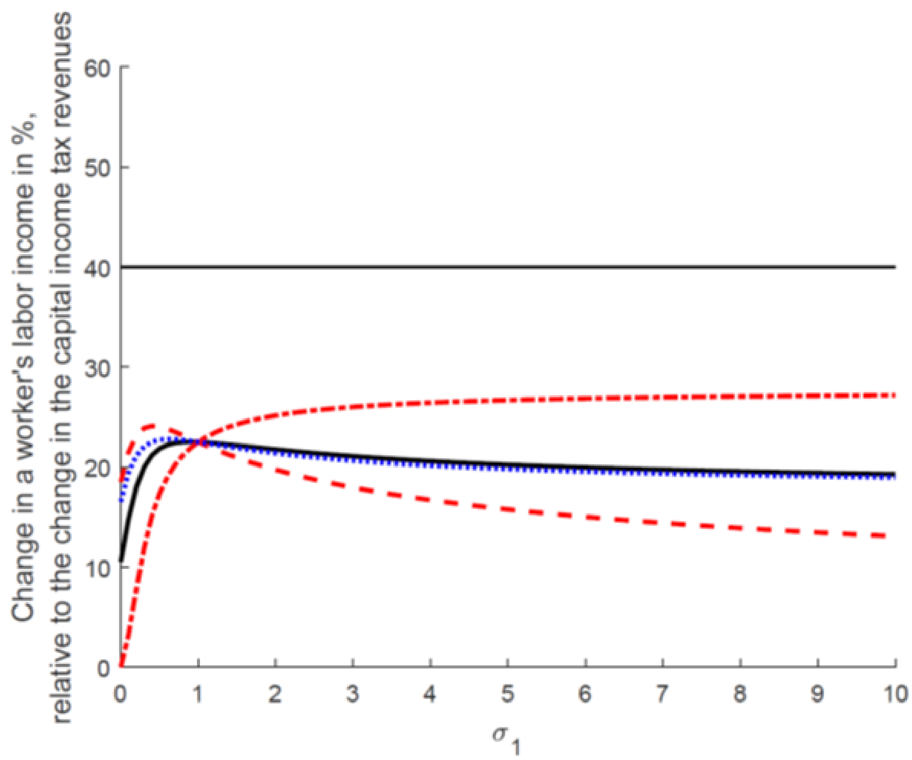
$$+ \frac{\frac{1+\kappa}{\kappa}}{1 - \frac{1+\kappa}{\kappa}\tau^{L_w}} - \frac{1}{1-\tau^{L_w}}$$

$$+ \frac{\theta_{l_w}}{\zeta(1+g)\frac{\theta_k}{r} + \frac{\hat{\kappa}}{1+\hat{\kappa}}\theta_{l_s} + \tau^K\theta_k + \tau^{L_w}\theta_{l_w}}$$

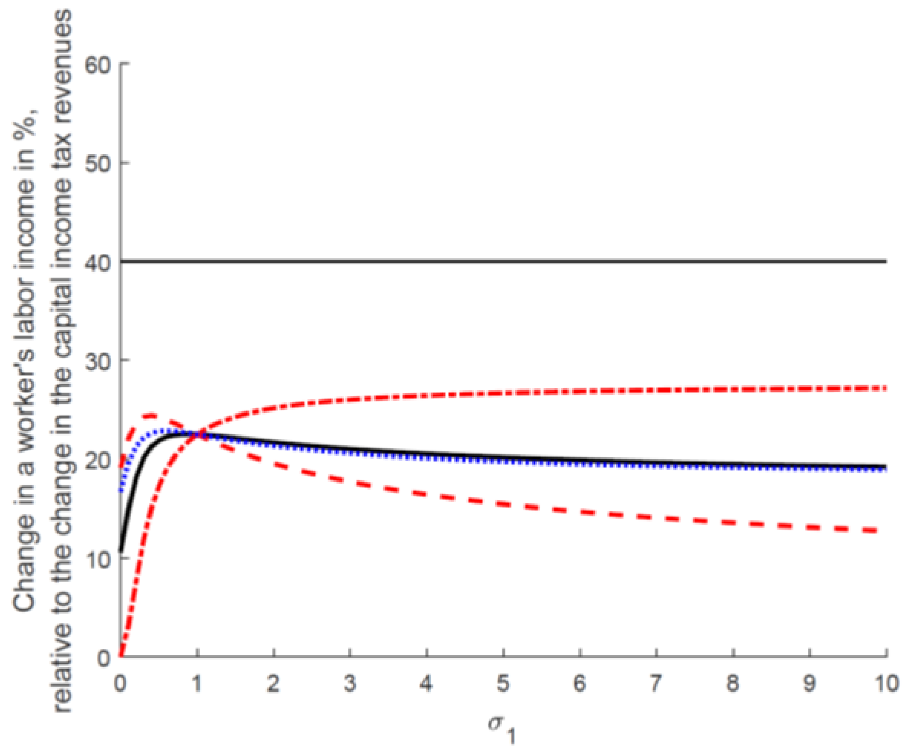
$$- \frac{\frac{\hat{\kappa}}{1+\hat{\kappa}}\theta_{l_w}}{\zeta(1+g)\frac{\theta_k}{r} + \frac{\hat{\kappa}}{1+\hat{\kappa}}\left(\theta_{l_s} + \tau^K\theta_k + \tau^{L_w}\theta_{l_w}\right)}\,.$$

4.8 Appendix B: Graphs and figures

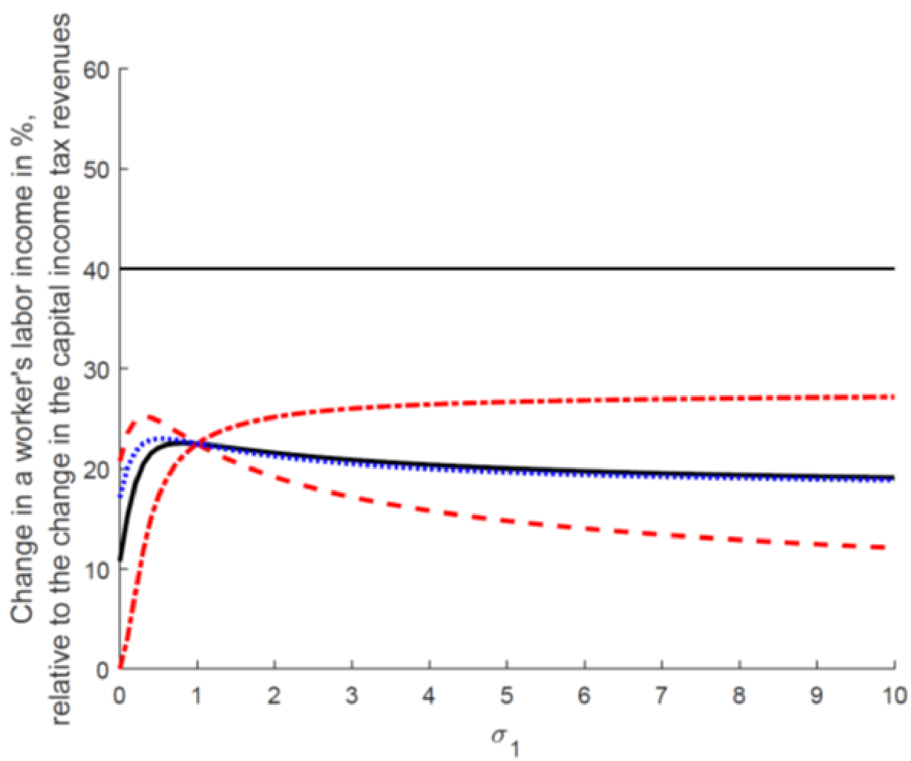
Figure 4.6 shows that the patterns to redistribute the capital income tax revenues back to the househlds do not change with the capital income tax rate. Workers receive the "same" share of capital tax revenues, also when employing a different capital income tax (due to a different σ_2). The figure visualizes the patterns for capital tax rates of ca. $1\%, 10\%, 33\%$ and 53%. A detailed explanation of the single lines is given in section 4.5.6.



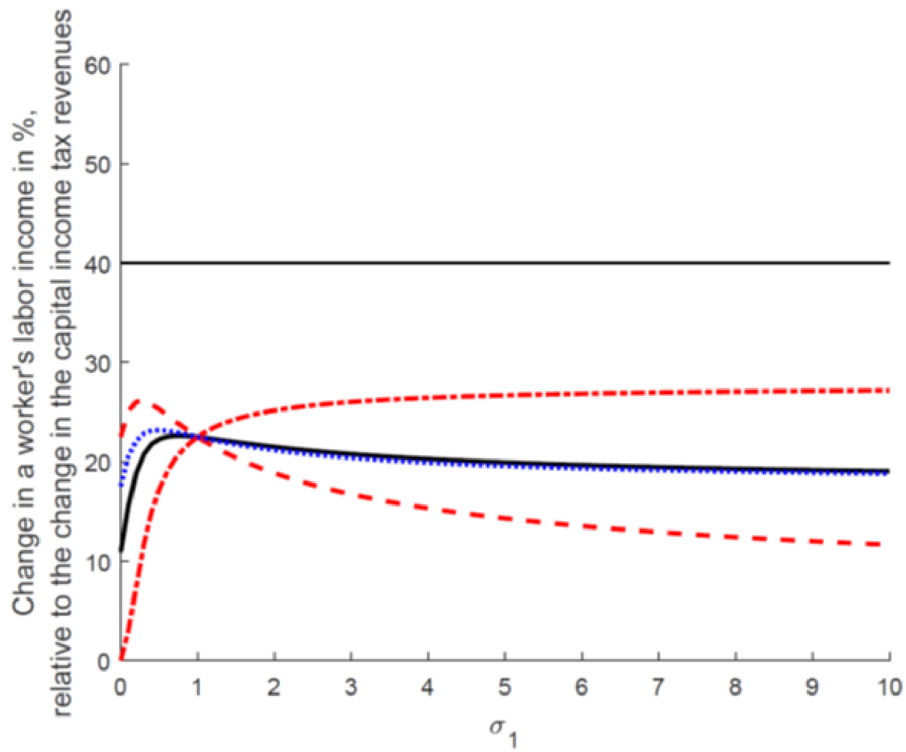
(a) Change in a worker's labor income due to a capital tax of ca. 1% (or $\sigma_2 = 1.1$).



(b) Change in a worker's labor income due to a capital tax of ca. 10% (or $\sigma_2 = 2$).



(c) Change in a worker's labor income due to a capital tax of ca. 33% (or $\sigma_2 = 5$).



(d) Change in a worker's labor income due to a capital tax of ca. 53% (or $\sigma_2 = 10$).

FIGURE 4.6. Change in a worker's labor income due to capital taxation.

Notes: **Black horizontal lines**: Threshold under which all households' receive the equal shares of capital income tax revenues. **Black solid lines**: Baseline calibration, savers respond at the intensive margin ($\hat{\kappa} = 2.6667$) and workers respond at the intensive and extensive margin ($\kappa = 1.1429$). **Blue dotted lines**: Same labor supply elasticities as in the baseline calibration ($\hat{\kappa} = 2.6667$, $\kappa = 1.1429$) plus a worker and a saver have the same pre-tax labor income ($\theta_l = \theta_{l_w}/\lambda = 0.2448$). **Red dashed lines**: Savers and workers respond both at the intensive and extensive margin ($\hat{\kappa} = \kappa = 1.1429$). **Red dashed-dotted lines**: Workers respond at the intensive and extensive margin ($\kappa = 1.1429$) and savers supply labor perfectly inelastically ($\hat{\kappa} = 1000$).

Bibliography

Abbott, J. and Fenichel, E.: 2014, Following the golden rule: Negishi welfare weights without apology. Unpublished manuscript.

Abrell, J. and Rausch, S.: 2017, Combining price and quantity controls under partitioned environmental regulation, Journal of Public Economics **145**, 226–242.

Abrell, J., Rausch, S. and Schwarz, G. A.: 2018, How robust is the uniform emissions pricing rule to social equity concerns?, Journal of Environmental Economics and Management **92**, 783–814.

Acemoglu, D., Golosov, M. and Tsyvinski, A.: 2011, Political economy of ramsey taxation, Journal of Public Economics **95**, 467–475.

Aguiar, A., Chepeliev, M., Corong, E., McDougall, R. and van der Mensbrugghe, D.: 2019, The gtap data base: Version 10., Journal of Global Economic Analysis **4**, 1–27.

Aiyagari, S. R.: 1994, Uninsured idiosyncratic risk and aggregate saving, The Quarterly Journal of Economics **109**, 659–684.

Atkinson, A.: 1970, On the measurement of inequality., Journal of economic theory **2**, 244–263.

Atkinson, A.: 1973, How progressive should income tax be?, M. Parkin and A.R. Nobay, eds. Essays in modern economics, Longman, London, pp. 90–109.

Attanasio, O., Banks, J. and Tanner, S.: 2002, Asset holding and consumption volatility, Journal of Political Economy **110**, 771–792.

Attanasio, O., Orazio, P. and Browning, M.: 1995, Consumption over the life cycle and over the business cycle, American Economic Review **85**, 1118–1137.

Baccianti, C.: 2013, Estimation of sectoral elasticities of substitution along the international technology frontier, ZEW Discussion Papers . No. 13-092.

Barrage, L.: 2018, Be careful what you calibrate for: Social discounting in general equilibrium, Journal of Public Economics **160**, 33–49.

Barrage, L.: 2019, Optimal Dynamic Carbon Taxes in a Climate–Economy Model with Distortionary Fiscal Policy, The Review of Economic Studies **87**(1), 1–39. **URL:** *https://doi.org/10.1093/restud/rdz055*

Barrage, L.: 2020, Optimal dynamic carbon taxes in a climate-economy model with distortionary fiscal policy, Review of Economic Studies **87**, 1–39.

Bastani, S. and Waldenström, D.: 2020, How should capital be taxed?, Journal of Economic Surveys **34**, 812–846.

Belfiori, M. E.: 2017, Carbon pricing, carbon sequestration and social discounting, European Economic Review **96**, 1–17.

Bentham, J.: 1776, A fragment on government.

Bernheim, B. D.: 1989, Intergenerational altruism, dynastic equilibria, and social welfare, Review of Economic Studies **56**, 119–128.

Bilbiie, F. O.: 2008, Limited asset markets participation, monetary policy and (inverted) aggregate demand logic, Journal of Economic Theory **140**, 162–196.

Bilbiie, F. O., Monacelli, T. and Perotti, R.: 2013, Public debt and redistribution with borrowing constraints, The Economic Journal **123**, 64–98.

Bilbiie, F. O. and Straub, R.: 2013, Asset market participation, monetary policy rules, and the great inflation, Review of Economics and Statistics **95**, 377–392.

Blanchard, O. and Tirole, J.: 2021, Major future economic challenges. Accessed in December 2021: https://www.strategie.gouv.fr/sites/strategie.gouv.fr/files/atoms/files/fs-2021-rapport-anglais-les_grands_defis_economiques-juin.pdf.

Blinder, A.: 1976, Intergenerational transfers and life cycle consumption, The American Economic Review, Papers and Proceedings of the Eighty-eighth Annual Meeting of the American Economic Association **66**, 87–93.

Blundell, R., Browning, M. and Meghir, C.: 1994, Consumer demand and the life-cycle allocation of household expenditures, The Review of Economic Studies **61**, 57–80.

Boeters, S.: 2014, Optimally differentiated carbon prices for unilateral climate policy, Energy Economics **45**, 304–312.

Böhringer, C., Carbone, J. C. and Rutherford, T. F.: 2016, The strategic value of carbon tariffs, American Economic Journal: Economic Policy **8**, 28–51.

Böhringer, C., Hoffmann, T. and de Lara-Peñate, C. M.: 2006, The efficiency costs of separating carbon markets under the EU emissions trading scheme: A quantitative assessment for Germany, Energy Economics **28**(1), 44–61.
URL: *https://doi.org/10.1016/j.eneco.2005.09.001*

Böhringer, C., Lange, A. and Rutherford, T. F.: 2014, Optimal emission pricing in the presence of international spillovers: Decomposing leakage and terms-of-trade motives, Journal of Public Economics **110**, 101–111.

Boskin, M. J. and Sheshinski, E.: 1978, Optimal redistributive taxation when individual welfare depends upon relative income, The Quarterly Journal of Economics **92**, 589–601.

Bovenberg, A. L. and Goulder, L. H.: 1996, Optimal environmental taxation in the presence of other taxes: General- equilibrium analyses, The American Economic Review **86**(4), 985–1000.
URL: *http://www.jstor.org/stable/2118315*

Bovenberg, L. A. and de Mooij, R. A.: 1994, Environmental levies and distortionary taxation, American Economic Review **94**, 1085–1089.

Bovenberg, L. A. and de Mooij, R. A.: 1997, Environmental tax reform and endogenous growth, Journal ofPublic Economics **63**, 207–237.

Bovenberg, L. A. and Smulders, S.: 1996, Transitional impacts of environmental policy in an endogenous growth model, International Economic Review **37**, 861–893.

Bovenberg, L. A. and van der Ploeg, F.: 1994, Environmental policy, public finance and the labour market in a second-best world, Journal of Public Economics **55**(3), 349–390.

Broome, J.: 1994, Discounting the future, Philosophy and Public Affairs **23**, 128–156.

Campbell, J. Y. and Mankiw, N. G.: 1991, The response of consumption to income: A cross-country investigation, European Economic Review **35**, 723–767.

Caplin, A. and Leahy, J.: 2004, The social discount rate, Journal of Political Economy **112**, 1257–1268.

Capros, P., van Regemorter, D., Paroussos, L., Karkatsoulis, P., Fragkiadakis, C., Tsani, S., Charalampidis, I. and Revesz, T.: 2013, Gem-e3 model documentation. Publications Office of the European Union. European Commission Joint Research Centre Institute for Prospective Technological Studies Technical Reports. https://publications.jrc.ec.europa.eu/repository/handle/JRC83177. Editors: Perry M, Abrell J, Ciscar Martinez J, Pycroft J, Saveyn B.

Challe, E., Matheron, J., Ragot, X. and Rubio-Ramirez, J.: 2017, Precautionary saving and aggregate demand, Quantitative Economics **8**, 435–478.

Chamley, C.: 1986, Optimal taxation of capital income in general equilibrium with infinite lives, Econometrica **54**(3), 607–622.

Chari, V. V. and Kehoe, P. J.: 1999, Handbook of Macroeconomics, Elsevier Science, chapter Optimal fiscal and monetary policy, pp. 1671–1745.

Chari, V. V., Nicolini, J. P. and Teles, P.: 2020, Optimal capital taxation revisited., Journal of Monetary Economics **116**, 147–165.

Chen, Y., Paltsev, S., Reilly, J., Morris, J. and Babiker, M.: 2015, MIT EPPA6 Model: Economic Growth, Energy Use, and Food Consumption, The.

Chetty, R., Guren, A., Manoli, D. and Weber, A.: 2002, Are micro and macro labor supply elasticities consistent? a review of evidence on the intensive and extensive margins., American Economic Review **101**, 471–475.

Chichilnisky, G. and Heal, G.: 1994, Who should abate carbon emissions? an international perspective, Economic Letters **44**, 443–449.

Conesa, J. C., Kitao, S. and Krueger, D.: 2009, Taxing capital? not a bad idea after all!, American Economic Review **99**, 25–48.

Costantini, V., Crespi, F. and Paglialunga, E.: 2019, Capital-energy substitutability in manufactoring sectors: methodological and policy implications, Eurasian Business Review **9**, 157–182.

Crippa, M., Oreggioni, G., Guizzardi, D., Muntean, M., Schaaf, E., Lo Vullo, E., Solazzo, E., Monforti-Ferrario, F., Olivier, J. and Vignati, E.: 2019, Fossil co_2 and ghg emissions of all world countries - 2019 report. data file “CO2 excl short-cycle org C” retrieved from https://edgar.jrc.ec.europa.eu/overview.php?v=50_GHG on 02.02.2019.

Cropper, M. and Oates, W.: 1992, Environmental economics: A survey, Journal of Economic Literature **30**(2), 675–740.

Dasgupta, P.: 2008, Discounting climate change, Journal of Risk and Uncertainty **37**, 141–169.

Dasgupta, P., Sen, A. and Starrett, D.: 1973, Notes on the measurement of inequality, Journal of Economic Theory **6**, 180–187.

Denning, F. and Emmerling, J.: 2017, A note on optima with negishi weights. Unpublished manuscript.

Diamond, P.: 1965, National debt in a neoclassical growth model, American Economic Review **60**, 1126–1150.

Diamond, P. and Mirrlees, J. A.: 1971, Optimal taxation and public production i: Production efficiency, American Economic Review **61**, 8–27.

Diamond, P. and Spinnewijn, J.: 2011, Capital income taxes with heterogeneous discount rates., American Economic Journal: Economic Policy **3**, 52–76.

Dissou, Y., Karnizova, L. and Sun, Q.: 2015, Industry-level econometric estimates of energy-capital-labor substitution with a nested ces production function, Atlantic Economic Journal **43**, 107–121.

Duffy, J., Papageorgiou, C. and Perez-Sebastian, F.: 2004, Capital-skill complementarity? evidence from a panel of countries, The Review of Economics and Statistics **86**, 327–344.

European Central Bank: 2020. The Household Finance and Consumption Survey: Results from the 2017 wave, accessed in November 2020, https://www.ecb.europa.eu/pub/pdf/scpsps/ecb.sps36 0245ed80c7.en.pdf.

European Commission: 2020, Regulation of the european parliament and of the council establishing the framework for achieving climate neutrality and amending regulation (eu) 2018/1999 (european climate law). https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1588581905912&uri=CELEX:52020PC0080.

European Commission: 2021a. Taxes on Capital-Total, accessed in August 2021: from https://ec.europa.eu/taxation_customs/taxation-1/economic-analysis-taxation/data-taxation_en.

European Commission: 2021b. Environment and energy tax revenues-Total, accessed in August 2021: https://ec.europa.eu/taxation_customs/taxation-1/economic-analysis-taxation/data-taxation_en.

European Commission: 2021c, Proposal for amending directive 2003/87/ec, com(2021) 551 final. Accessed in July 2021: https://carbonpricingdashboard.worldbank.org/map_data.

Fahri, E. and Werning, I.: 2007, Inequality and social discounting, Journal of Political Economy **115**, 365–402.

Fair, R. C.: 1971, The optimal distribution of income, The Quarterly Journal of Economics **85**, 551–579.

Fawcett, A. A., Clarke, L. C., Rausch, S. and Weyant, J. P.: 2014, Overview of emf 24 policy scenarios, The Energy Journal **35**, 33–60.

Federal Constitutional Court: 2021, Constitutional complaints against the federal climate change act partially successful. Press Release No. 31/2021 of 29 April 2021. https://www.bundesverfassungsgericht.de/SharedDocs/Pressemitteilungen/EN/2021/bvg21-031.html.

Feldstein, M.: 1973, On the optimal progressivity of the income tax, Journal of Public Economics **2**, 357–376.

Gahvari, F. and Micheletto, L.: 2016, Capital income taxation and the atkinson - stiglitz theorem, Economics Letters **147**, 86–89.

Gali, J., Lopez-Salido, J. and Valles, J.: 2007, Understanding the effects of government spending on consumption, Journal of the European Economic Association **5**, 227–270.

Gollier, C. and Hammitt, J. K.: 2014, The long-run discount rate controversy, Annual Review of Resource Economics **6**, 273–295.

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

Golosov, M., Troshkin, M., Tsyvinski, A. and Weinzierl, M.: 2013, Preference heterogeneity and optimal capital income taxation., Journal of Public Economics **97**, 160–175.

Goulder, L. H., Hafstead, M. A. C. and Williams III, R. C.: 2016, General equilibrium impacts of a federal clean energy standard, American Economic Journal: Economic Policy **8**, 186–218.

Goulder, L. and Parry, I. W.: 2008, Instrument choice in environmental policy, Review of Environmental Economics and Policy **2**(2), 152–174.

Goulder, L. and Williams, R. I.: 2012, The choice of discount rate for climate change policy evaluation, Clim. Chang. Econ **3**, 125002.

Greenwood, B., Hercowitz, Z. and Huffmann, G.: 2011, Investment, capacity utilization, and the real business, American Economic Review **78**, 402–417.

Hahn, R. W.: 1984, Market power and transferable property rights, Quarterly Journal of Economics **99**, 753–765.

Halsnaes, K., Shukla, P., Ahuja, D., Akumu, G., Beale, R., Edmonds, J., Gollier, C., Gruebler, A., Duong, M. H., Markandya, A., McFarland, M., Nikitina, E., Sugiyama, T., Villavicencio, A. and Zou, J.: 2007, Framing issues, Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth

Assessment Report of the Intergovernmental Panel on Climate Change, B. Metz, O. R. Davidson, P. R. Bosch, R. Dave, L. A. Meyer (eds)], Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 117–167.

Hassler, J. and Krusell, P.: 2012, Economics and climate change: Integrated assessment in a multi-region world, Journal of the European Economic Association **10**, 970–1000.

Henningsen, A., Henningsen, G. and van der Werf, E.: 2019, Capital-labour-energy substitution in a nested ces framework: A replication and update of kemfert (1998), Energy Economics **82**, 16–25.

Hobijn, B. and Nechio, F.: 2019, Sticker shocks: Using vat changes to estimate upper-level elasticities of substitution, Journal of the European Economic Association **18**, 799–833.

Hoel, M.: 1996, Should a carbon tax be differentiated across sectors?, Journal of Public Economics **59**, 17–32.

Hubbard, R., K. Judd, R., Hall, E. and Summers, L.: 1986, Liquidity constraints, fiscal policy, and consumption, Brookings Papers on Economic Activity **1986**, 1–59.

Jacobs, B. and Bovenberg, A.: 2010, Human capital and optimal positive taxation of capital income., International Tax and Public Finance **17**, 451–478.

Jacobs, B., Gerritsen, A. and Spiritus, K.: 2020, Optimal taxation of capital income with heterogenous rates of return. CESifo Working Paper Servies 8395, CESifo.

Jaimovich, N. and Rebelo, S.: 2009, Can news about the future drive the business cycle?, American Economic Review **99**, 1097–1118.

Judd, K. L.: 1985, Redistributive taxation in a simple perfect foresight model, Journal of Public Economics **28**(1), 59–83.

Kalsbach, O. and Rausch, S.: 2021, Pricing carbon in a multi-sector economy with social discounting, ZEW Discussion Paper, No. 21-060 .

Kaplan, G., Moll, B. and Violante, G. L.: 2018, Monetary policy according to hank, American Economic Review **108**, 697–743.

Kaplow, L.: 2012, Optimal control of externalities in the presence of income taxation, International Economic Review **53**, 487–509.

Kaplow, L., Moyer, E. and Wiesbach, D.: 2010, The social evaluation of intergenerational policies and its application to integrated assessment models of climate change, The B.E. Journal of Economic Analysis and Policy **10**, Article 7.

Keane, M. P.: 2011, Labor supply and taxes: A survey, Journal of Economic Literature **49**, 961–1075.

Kehoe, P., Chari, V. and Atkeson, A.: 1999, Taxing capital income: A bad idea, Federal Reserve Bank of Minneapolis Quarterly Review, 23(**23**, 3–17.

Kemfert, C. and Welsch, H.: 2000, Energy-capital-labor substitution and the economic effects of co2 abatement: Evidence for germany, Journal of Policy Modeling **22**, 641–660.

Koesler, S. and Schymura, M.: 2015, Substitution elasticities in a constant elasticity of substitution framework- empirical estimates using nonlinear least squares, Economic Systems Research **27**, 101–12.

Kristjansson, A.: 2016, Optimal taxation with endogenous return to capital. Working Paper 06/2016, Oslo University.

Krutilla, K.: 1991, Environmental regulation in an open economy, Journal of Environmental Economics and Management **20**, 127–142.

Landis, F., Rausch, S. and Kosch, M.: 2018, Differentiated carbon prices and the economic cost of decarbonization, Environmental Resource Economics **70**, 483–516.

Lansing, K. J.: 1999, Optimal redistributive capital taxation in a neoclassical growth model, Journal of Public Economics **73**, 423–453.

Lansing, K. J.: 2015, Asset pricing with concentrated ownership of capital and distribution shocks, American Economic Journal: Macroeconomics **7**, 67–103.

Lu, Y. and Stern, D. I.: 2016, Substitutability and the cost of climate mitigation policy, Environmental and Resource Economics **64**, 81–107.

Mankiw, N. G.: 2000, The savers-spenders theory of fiscal policy., American Economic Review **90**, 120–125.

Markusen, J. R.: 1975, International externalities and optimal tax structures, Journal of International Economics **5**, 15–295.

Metcalf, G. E.: 2009, Market-based policy options to control u.s. greenhouse gas emissions, Journal of Economic Perspectives **23**, 5–27.

Mirrlees, J.: 1971, An exploration in the theory of optimum income taxation., Review of Economic Studies **38**, 175–208.

Moll, B.: 2014, Productivity losses from financial frictions: Can self-financing undo capital misallocation?, American Economic Review **104**, 3186–3221.

Nordhaus, W.: 2015, Climate clubs: Overcoming free-riding in international climate policy, American Economic Review **105**(4), 1339–70.
URL: *https://www.aeaweb.org/articles?id=10.1257/aer.15000001*

Nordhaus, W. D.: 2000, Warming the world: economic models of global warming, MIT Press, London.

Nordhaus, W. D.: 2007, A review of the stern review on the economics of climate change, Journal of Economic Literature **XLV**, 686–702.

Nordhaus, W. D.: 2011, Estimates of the social cost of carbon: Background and results from the rice-2011 model, Working Paper 17540, National Bureau of Economic Research.

Nordhaus, W. D.: 2017, Revisiting the social cost of carbon, Proceedings of the National Academy of Sciences of the United States of America **114**, 1518–1523.

Nordhaus, W. D. and Yang, Z.: 2006, Magnitude and direction of technological transfers for mitigating ghg emissions, Energy Economics **28**, 730–741.

Nozick, R.: 1974, Anarchy, state, and utopia. New York: Basic Books.

OECD: 2015, How much redistribution do governments achieve? the role of cash transfers and household taxes, Growing Unequal?: Income Distribution and Poverty in OECD Countries, pp. 97–121.

OECD: 2020. Exchange rates (indicator) doi: 10.1787/037ed317-en, Accessed in November 2019.

Okagawa, A. and Ban, K.: 2008, Estimation of substitution elasticities for cge models. OSIPP Discussion Papers in Economics and Business 08-16.

Paltsev, S., Reilly, J. M., Jacoby, H. D., Eckaus, R. S., McFarland, J., Sarofim, M., Asadoorian, M. and Babiker, M.: 2005, The mit emissions prediction and policy analysis (eppa) model: Version 4.

Papageorgiou, C., Saam, M. and Schulte, P.: 2017, Elasticity of substitution between clean and dirty energy inputs: A macroeconomic perspective, Review of Economics and Statistics **99**, 281–290.

Parry, I. W. H., Williams III, R. C. and Goulder, L. H.: 1999, When can carbon abatement policies increase welfare? the fundamental role of distorted factor markets, Journal of Environmental Economics and Management **37**, 52–84.

Phelps, E. S.: 1961, The golden rule of accumulation: A fable for growthmen, American Economic Review **51**(4), 638–642.

Phelps, E. S.: 1973, Taxation of wage income for economic justice, The Quarterly Journal of Economics **87**, 331–354.

Pigou, A.: 1932, Wealth and welfare. MacMillan and Co., 4th edition, London, UK.

Piketty, T. and Saez, E.: 2013, Optimal labor income taxation, Handbook of Public Economics, Vol. 5, pp. 391–474.

Ramsey, F.: 1928, A mathematical theory of saving, The Economic Journal **38**, 543–559.

Rauscher, M.: 1994, On ecological dumping, Oxford Economic Papers **46**, 822–840.

Rawls, J.: 1971, A theory of justice. Cambridge, MA: Belknap Press of Harvard University Press.

Reinhorn, L. J.: 2019, On optimal redistributive capital taxation, Journal of Public Economic Theory **21**, 460–487.

Saez, E.: 2001, Using elasticities to derive optimal income tax rates, The Review of Economic Studies **68**, 205–229.

Saez, E.: 2002, Optimal income transfer programs: intensive versus extensive labor supply responses, The Quarterly Journal of Economics **117**, 1039–1073.

Saez, E. and Stantcheva, S.: 2016, Generalized social marginal welfare weights for optimal tax theory., American Economic Review **106**, 24–45.

Saez, E. and Stantcheva, S.: 2018, A simpler theory of optimal capital taxation, Journal of Public Economics **162**, 120–142.

Sandmo, A.: 1975, Optimal taxation in the presence of externalities, Swedish Journal of Economics **77**, 86–98.

Sandmo, A.: 1999, Asymmetric information and public economics: The mirrlees-vickrey nobel prize, Journal of Economic Perspectives **13**, 165–180.

Schmookler, J.: 1966, Inventions and economics growth, Harvard University Press, Cambridge.

Schwartz, J. and Repetto, R.: 2000, Nonseparable utility and the double dividend debate: Reconsidering the tax interaction effect, Environmental and Resource Economics **15**, 149–157.

Sen, A. K.: 1973, On economic inequality. clarendon press, oxford.

Sheeran, K. A.: 2006, Who should abate carbon emission? a note, Environmental & Resource Economics **35**, 89–98.

Sidgwick, H.: 1874, The methods of ethics. Oxford University Press, Oxford, UK.

Solow, R.: 1974, The economics of resources or the resources of economics, The American Economic Review: Papers and Proceedings **64**, 1–14.

Stantcheva, S.: 2017, Optimal taxation and human capital policies over the life cycle., Journal of Political Economy **125**, 1931–1990.

Stanton, E.: 2011, Negishi welfare weights in integrated assessment models: the mathematics of global inequality, Climate Change **107**, 417–432.

Stern, N.: 1976, On the specification of models of optimum income taxation, Journal of Public Economics **6**, 123–162.

Stern, N.: 2006, The economics of climate change: The stern review.

Stiglitz, J.: 2018, Pareto efficient taxation and expenditures: Pre- and re-distribution., Journal of Public Economics **162**, 101–119.

Straub, L. and Werning, I.: 2020, Positive long run capital taxation: Chamley - judd revisited., American Economic Review **110**, 86–119.

The World Bank: 2020, Gdp growth - european union. Accessed in February 2021: https://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG?locations=EU.

Thompson, T., Rausch, S., Saari, R. and Selin, N.: 2014, A systems approach to evaluating the air quality co-benefits of us carbon policies, Nature Climate Change **4**, 917–923.

Timmer, M., Dietzenbacher, E., Los, B., Stehrer, R. and de Vries, G. J.: 2015, An illustrated user guide to the world input- output database: the case of global automotive production, Review of International Economics **23**, 575–605.

Tol, R. S. J.: 2009, The economic effects of climate change, Journal of Economic Perspectives **23**(2), 29–51.
URL: *https://www.aeaweb.org/articles?id=10.1257/jep.23.2.29*

van der Ploeg, F. and Rezai, A.: 2019, Simple rules for climate policy and integrated assessment, Environmental and Resource Economics **72**, 77–108.

van der Ploeg, F. and Withagen, C.: 1991, Pollution control and the ramsey problem, Environmental and Resource Economics **1**, 215–236.

van der Ploeg, F. and Withagen, C.: 2014, Growth, renewables and the optimal carbon tax, International Economic Review **55**, 283–311.

van der Werf, E.: 2008, Production functions for climate policy modeling: An empirical analysis, Energy Economics **30**, 2964–2979.

Vissing-Jorgensen, A.: 2002, Limited asset market participation and the elasticitiy of intertemporal substitution, Journal of Political Economy **110**, 825–853.

von Below, D.: 2012, Optimal carbon taxes with social and private discounting. Working Paper, SURED Conference, Ascona, Switzerland.

Weitzman, M. L.: 1974, Prices vs. quantities, Review of Economic Studies **41**, 477–491.

Weitzman, M. L.: 2007, A review of the stern review on the economics of climate change, Journal of Economic Literature **XLV**, 703–724.

Weyant, J., Davidson, O., Dowlatabadi, H., Edmonds, J., Grubb, M., Parson, E., Richels, R., Rotmans, J., Shukla, P., Tol, R., Cline, W. and Fankhauser, S.: 1996, Integrated Assessment of Climate Change: An Overview and Comparison of Approaches and Results, pp. 367–396.

World Bank: 2020, Gross savings (% of gdp)–european union. Accessed in February 2020: https://data.worldbank.org/indicator/NY.GNS.ICTR.ZS?locations=EU.

World Bank: 2021, Carbon pricing dashboard. Accessed in July 2021: https://carbonpricingdashboard.worldbank.org/map_data.

Curriculum Vitae

Born on May 26, 1992 in Motala, Sweden

2017 – 2022	Ph.D. in Economics - ETH Zurich
2011 – 2016	M.Sc. Industrial Engineering - Karlsruhe Institute of Technology