University of Heidelberg

Department of Economics

Discussion Paper Series | No. 654

Diverging Regional Climate Preferences and the Assessment of Solar Geoengineering

Tobias Pfrommer

August 2018

Diverging Regional Climate Preferences and the Assessment of Solar Geoengineering*

Tobias Pfrommer†
Heidelberg University

August 9, 2018

Abstract

Solar Geoengineering (SG) is a set of potential technologies to counteract climate change. While SG can only imperfectly compensate for temperature changes at the regional level, studies assessing regional SG impacts indicated so far that regional temperature disparities from SG may not be as severe as previously thought. A shortcoming of that literature is its assumption that regions' temperature preferences correspond to some historic baseline climate. I extend the main framework for examining regional SG impacts by allowing for regions to have temperature preferences diverging from the baseline climate, showing that the impact of these diverging preferences can be split into two components. The first component changes the optimal SG level, but does not affect regional disagreement over SG. The second component leaves the optimal SG level unaffected, but changes regional disagreement over SG. I identify three aspects of SG performance in the presence of diverging preferences. A numerical implementation of the extended model shows that the presence of diverging preferences may change SG performance in either direction and that the direction generally depends on which of the three aspects of SG performance is considered.

Keywords: Regional Climate Preferences, Solar Geoengineering, Regional Climate Disparities

*I want to thank Ben Kravitz for providing me with the climate model data used in this study. Furthermore, I want to thank Florian Diekert and Daniel Heyen for helpful comments. I gratefully acknowledge funding from the Priority Programme 'Climate Engineering: Risks, Challenges, Opportunities?' (SPP 1689) of the German Research Foundation (DFG).

†Email: pfrommer@eco.uni-heidelberg.de. Postal address: Department of Economics, Bergheimer Str. 20, 69115 Heidelberg, Germany. Phone: +49 6221 548014, Fax: +49 6221 548020.

1 Introduction

Techniques to increase the earth's reflection of solar radiation, so-called solar geoengineering (SG), have received increasing attention as potential means to reduce climate change risks. While SG may well be able to compensate for increased temperatures on a global level, SG has regionally heterogeneous impacts (Lunt et al. 2008, Robock et al. 2008, Irvine et al. 2010, Ricke et al. 2010). These heterogeneous impacts are widely regarded as a source of substantial SG governance problems and as a potential source of conflict (Robock 2008, Shepherd 2009, Weitzman 2015, Heyen 2016, Pasztor 2017).

Regional differences in SG impacts have been the focus of a number of studies (Moreno-Cruz et al. 2012, Kravitz et al. 2014, Yu et al. 2015, Pfrommer 2018). Generally, the results indicate that regional temperature disparities from SG may not be as severe as previously thought. However, this literature employs the assumption that regional climate preferences derive from a common baseline climate, e.g. 1990 or preindustrial climate conditions. This 'change-is-bad-assumption' has been criticized (Heyen et al. 2015) and research has provided empirical evidence that certain temperature levels may generally be more conducive to economic activity than others, irrespective of historic regional climate conditions. Schlenker and Roberts (2009) provide evidence for a non-linear, inversely U-shaped, relationship between crop yield and daily temperature. Graff and Neidell (2014) find a similar relationship between labor supply and daily temperature. On an aggregate level, empirical evidence points towards a globally generalizable, inversely U-shaped relationship between overall economic productivity and regional mean temperature (Burke et al 2015, Newell et al. 2018). Additionally, there are region-specific reasons why a deviation from regional baseline climate conditions may be beneficial (Heyen et al. 2015), e.g. that more natural resources in high northern latitudes may become available as temperatures rise.

The main purpose of the paper is to gain a conceptual understanding of the impact that regionally diverging temperature preferences have on SG outcomes. To this aim, I extend the Residual Climate Response (RCR) model developed by Moreno-Cruz et al. (2012). The RCR model is the main framework for examining SG impacts on the regional level in the literature. The extension allows for regions to have temperature preferences diverging from the baseline climate and builds on an illustrative example by Heyen et al. (2015) which demonstrates that regional temperature disparities may be substantially higher when regions have diverging temperature preferences. The extended model calculates the welfare maximizing SG level depending on the diverging temperature preferences that regions have. Regional damages, and therefore regional climate welfare, are determined by regional residual temperature. Regional residual temperature is the difference between regional temperature given optimal SG and a region's desired temperature level.

The key theoretical insight of the extended model is that the impact of diverging preferences can be split into two components. The first component changes the optimal SG level, but does not affect the set of residual temperatures. This component therefore does not change regional disagreement over SG. The second component leaves the optimal SG level unaffected, but changes the set of residual temperatures. This decomposition helps in understanding how specific diverging preferences affect globally optimal SG and the regional disagreement over SG.

In order to quantify the changes in SG outcomes due to diverging preferences, I propose four metrics. These metrics cover three different aspects of SG performance. The first aspect is the relative effectiveness of SG in reducing damages. Relative effectiveness measures in percent the share of regional damages in the high CO_2 climate[1] that optimal SG can compensate for. There are two potential damage reference levels for measuring relative effectiveness. The first level is total damages (arising from the combination of CO_2 driven temperature changes and diverging preferences), giving rise to the *total damage reduction metric M1* – the metric considered by Heyen et al. (2015). The second level is purely CO_2 driven damages, giving rise to the *CO_2 damage reduction metric M2.* While the *total damage reduction metric* measures how well optimal SG compensates for damages originating from the difference between regional temperatures in a high CO_2 climate and regionally preferred temperatures, the *CO_2 damage reduction metric* measures how well optimal SG compensates for damages caused by CO_2 induced temperature changes only. While the two damage reference levels – and the two metrics – are identical in the absence of diverging preferences, they constitute different ways of assessing the relative effectiveness of SG in the presence of diverging preferences.

The second aspect is the change of the minimum climate damages, or equivalently, of the maximum climate welfare that SG can implement. It is captured by comparing residual damages (the sum of regional damages given optimal SG) for a given diverging preferences scenario to residual damages in the absence of diverging preferences (the 'baseline scenario'). This second aspect gives rise to the *minimum climate damage metric M3.* The third aspect is the change in the gross value of SG relative to the baseline scenario. It is captured by comparing the maximum reduction of damages that SG can achieve for a given diverging preferences scenario to the maximum reduction in the baseline scenario and gives rise to the *gross value metric M4.* While relative effectiveness measures SG performance with respect to one fixed diverging preferences scenario, the other two aspects measure SG performance relative to SG performance in the baseline scenario. Thereby, they make absolute SG outcomes – in case of *M3* the maximum climate welfare that SG can implement, in case of *M4* the gross value of SG – comparable across different diverging preferences scenarios.

A diverging preferences scenario can be expressed as a combination of a 'profile' and of a 'preference strength'. A profile describes regions' diverging temperatures preferences relative to each other, the preference strength determines the magnitude of the desired temperature deviations from the baseline climate. Profiles in which high-latitude regions generally prefer higher temperatures than in the baseline climate and low-latitude regions generally prefer lower temperatures are of specific interest, since they are plausible and, at least for aggregate economic performance, are supported by empirical evidence (Burke et al. 2015, Newell et al. 2018). In order to develop a basic understanding

[1] In the entire paper, CO_2 is intended to mean "CO_2 equivalent", i.e. CO_2 represents all greenhouse gases.

of how diverging preferences deriving from such profiles affect SG outcomes, I numerically implement two concrete such profiles with data from the Geoengineering Model Intercomparision Project (Kravitz et al. 2011). For each profile, I implement different strengths of diverging preferences in order to capture effects that are non-linear in the magnitude of the desired temperature deviations from the baseline climate.

The numerical implementation shows that the performance of optimal SG in the presence of diverging preferences relative to the baseline scenario may change in either direction and that the direction generally depends on which of the three aspects of SG performance is considered. The latter implies that the aspects and metrics developed are in fact independent and convey complementary information about how SG performance changes in the presence of diverging preferences. Furthermore, optimal climate welfare in the presence of diverging preferences generally compares differently relative to different welfare reference levels. Compared to optimal climate welfare in the baseline scenario, optimal climate welfare in the presence of diverging preferences is higher when the magnitude of the desired temperature deviations is small, but smaller for moderate and large desired temperature deviations. In contrast, compared to climate welfare in the baseline climate (i.e. in the climate before CO_2 driven temperature changes set in), optimal climate welfare in the presence of diverging preferences is higher for all but very small magnitudes of desired temperature deviations.

I proceed as follows. In section 2, I introduce the RCR model. Section 3 extends the RCR model. The implementation of the scenarios follows in section 4. Section 5 concludes.

2 The Residual Climate Response Model

The Residual Climate Response (RCR) model (Moreno-Cruz et al. 2012) is a simple framework for evaluating regional effects of solar geoengineering (SG). The main purpose of the RCR model is to examine how well SG can compensate for CO_2 induced climate change on the regional level. Due to its simplicity, the model is intuitively accessible (compare Figure 1). Its strength lies in identifying first-order effects in the assessment of regional SG impacts and in providing a framework for conceptually thinking about regional disparities in SG outcomes.

The RCR model operates on a fixed number of regions. In the model, regional climate damages derive from the difference in mean temperature between actual climate conditions and climate conditions in some baseline climate. The CO_2 vector T_{CO2} consists of the regional temperature increases due to CO_2 emissions relative to the baseline climate.[2] One unit of SG is defined as the amount of SG restoring global mean temperature to the level of the baseline climate. The SG vector T_{SG} consists of the regional mean temperature changes due to one unit of SG relative to the high CO_2 climate. Since the CO_2 and SG vectors are not congruent (i.e. the relative effects of CO_2 and SG on tem-

[2]While other metrics, like regional precipitation, are generally relevant and employed as well, I focus on the metric of regional temperature. Baseline climates employed in the literature are preindustrial and 1990 climate conditions (Moreno-Cruz et al. 2012, Kravitz et al. 2014, Yu et al. 2015, Pfrommer 2018).

peratures differ across regions), SG can only imperfectly compensate for CO_2 induced temperature increases on a regional level. The model assumes linearity in the effect of SG on regional temperature.[3] The SG level x is defined as a fraction of one unit of SG. Due to the linearity assumption, the residual vector T_{RES} of regional temperatures, given the SG level x, is

$$T_{\mathrm{RES}}(x) = T_{\mathrm{CO2}} + x \cdot T_{\mathrm{SG}}.$$

The global welfare measure is residual damages, i.e. the sum of regional damages, and corresponds to the squared length of the residual vector T_{RES} for a given level of SG:

$$D(x) = \sum_i d^i(x) = |T_{\mathrm{RES}}(x)|^2 \quad \text{with} \quad d^i(x) = (T^i_{\mathrm{CO2}} + x \cdot T^i_{\mathrm{SG}})^2.$$

The optimal SG level, minimizing residual damages, is

$$x^* = -\frac{\sum_i T^i_{\mathrm{SG}} \cdot T^i_{\mathrm{CO2}}}{|T_{\mathrm{SG}}|^2} = -\frac{T_{\mathrm{SG}} \bullet T_{\mathrm{CO2}}}{|T_{\mathrm{SG}}|^2} = \frac{|T_{\mathrm{CO2}}| \cdot \cos(\varphi)}{|T_{\mathrm{SG}}|},$$

where φ is the angle between the CO_2 and the SG vector and $(\bullet)$ the dot product between vectors.[4] The smaller φ, the more similar the CO_2 and the SG vector and the better SG can compensate for the temperature changes on a regional level caused by CO_2: In case the vectors are parallel ($\varphi = 0°$), compensation is perfect. Consequently, we then have $x^* = 1$ and $D(x^*) = 0$. In case the vectors are perpendicular ($\varphi = 90°$), compensation is not possible at all. Consequently, we then have $x^* = 0$ and $D(x^*) = |T_{\mathrm{CO2}}|^2$. If not stated otherwise, the terms residual damages and residual vector from now on refer to the respective outcomes given the optimal SG level.

The metric M used for assessing SG is the relative effectiveness of optimal SG in compensating for CO_2 induced damages on the regional level.[5] Due to the linearity assumption, M does not depend on the length of the CO_2 vector, but only on φ. Since the dot product between the residual vector $T_{\mathrm{RES}}(x^*)$ and the SG vector T_{SG} is zero, the two are perpendicular (see Figure 1) and M evaluates to

$$\frac{D(0) - D(x^*)}{D(0)} = 1 - \frac{|T_{\mathrm{RES}}(x^*)|^2}{|T_{\mathrm{CO2}}|^2} = 1 - \sin^2(\varphi).$$

[3]Evidence for the reasonableness of this assumption is provided, among others, by Moreno-Cruz et al. (2012) and Kravitz et al. (2014).

[4]Technically, φ as depicted in Figure 1, is not the angle between the CO_2 and the SG vector, but between the CO_2 vector and the negative of the SG vector. φ is to be understood as defined by Figure 1. The reason for doing so is that I do not want to deviate from the definition in Moreno-Cruz et al. (2012). The only implication is that $\cos(\varphi)$ picks up a minus sign:

$$\cos(\varphi) = -\frac{T_{\mathrm{CO2}} \bullet T_{\mathrm{SG}}}{|T_{\mathrm{CO2}}| \cdot |T_{\mathrm{SG}}|}.$$

[5]The damages different regions experience may also be weighted. Such weights may, for example, reflect differences in population or economic output. However, the analysis does not fundamentally change in the presence of welfare weights and the possibility of welfare weights is therefore not further pursued in this paper.

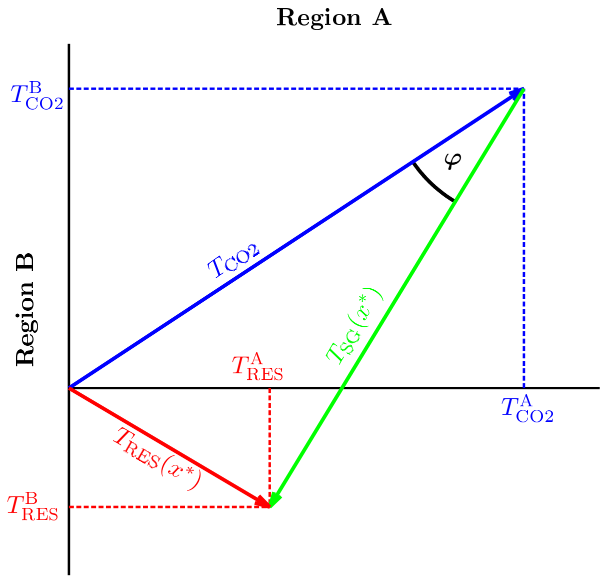

Figure 1: Two-region representation of the Residual Climate Response model. The horizontal axis shows changes in temperature for region A, the vertical axis shows changes in temperature for region B. The blue CO_2 vector represents regional temperature changes due to CO_2. The green SG vector represents regional temperature changes due to optimal SG. The red residual vector points to regional temperatures under optimal SG. Since regional damages are quadratic, the squared length of the residual vector is proportional to residual damages. The angel φ between CO_2 and SG vector represents the extent to which SG can compensate for regional temperature changes due to CO_2 and determines the socially optimal SG level as well as the metric M.

3 Extension of the Residual Climate Response Model

The following extension of the RCR model includes the possibility of diverging regional preferences from the baseline climate. The vector T_{DIV} of diverging temperature preferences consists of each region's preferred temperature relative to the baseline climate. From a societal perspective, the aim in deploying SG in the baseline model is to compensate for the CO_2 induced regional temperature deviation from the baseline climate as represented by the CO_2 vector. In the face of diverging preferences, the aim changes to compensating for the differences between regional temperatures in the high CO_2 climate and regions' preferred temperature levels. The residual vector in the extended model is therefore

$$T_{\mathrm{RES}}(x, T_{\mathrm{DIV}}) = (T_{\mathrm{CO2}} - T_{\mathrm{DIV}}) + x \cdot T_{\mathrm{SG}}$$

and the welfare goal is minimizing

$$D(x, T_{\mathrm{DIV}}) = |T_{\mathrm{RES}}(x, T_{\mathrm{DIV}})|^2.$$

When referring to the residual vector and the residual damages in the absence of diverging preferences, I will leave out the second argument.

Denote the angle between the SG vector and the diverging preferences vector by ϑ (see Figure 2). Furthermore, denote optimal SG in the baseline model as x^* and in the extended model as x^*_{DIV}. The optimal amount of SG in the presence of diverging preferences is then

$$x^*_{\mathrm{DIV}} = -\frac{T_{\mathrm{SG}} \bullet (T_{\mathrm{CO2}} - T_{\mathrm{DIV}})}{|T_{\mathrm{SG}}|^2} = x^* + \frac{|T_{\mathrm{DIV}}| \cdot \cos(\vartheta)}{|T_{\mathrm{SG}}|}.$$

The vector of diverging preferences can be decomposed into a component parallel and a component perpendicular to the SG vector. In case ϑ is larger than 90°, the parallel component points in the opposite direction of the SG vector. The diverging preferences vector then (partially) substitutes for the cooling the SG vector provides and less SG is optimal than in the absence of diverging preferences. The contrary is true in case ϑ is smaller than 90° and the parallel component points in the same direction as the SG vector. The decomposition of the diverging preferences vector into the parallel and perpendicular components is

$$T_{\mathrm{DIV}} = T^{\perp}_{\mathrm{DIV}} + T^{\parallel}_{\mathrm{DIV}},$$

$$\text{with} \quad T^{\parallel}_{\mathrm{DIV}} = \frac{T_{\mathrm{SG}} \bullet T_{\mathrm{DIV}}}{|T_{\mathrm{SG}}|^2} \cdot T_{\mathrm{SG}} \quad \text{and} \quad T^{\perp}_{\mathrm{DIV}} = T_{\mathrm{DIV}} - T^{\parallel}_{\mathrm{DIV}}.$$

The component parallel to the SG vector changes regional temperature preferences in the same proportions as SG changes regional temperatures. Increasing or decreasing SG relative to optimal SG in the baseline model can therefore perfectly compensate for the parallel component. This implies that the parallel component does not change the residual vector. Therefore, optimal SG changes, but regional damages and hence residual damages do not change relative to the absence of diverging preferences. In particular, when the diverging preferences vector is (anti)parallel to the SG vector, regions disagree about SG in exact the same way as in the baseline model.

In contrast, changing the SG level cannot compensate for the perpendicular component at all and the perpendicular component does therefore not affect optimal SG. It must then completely be taken up by the residual vector, thereby changing how regions disagree about SG relative to the baseline model. In particular, when the diverging preferences vector is perpendicular to the SG vector, optimal SG does not change compared to the baseline model.

The residual vector in the presence of diverging preferences may be longer or shorter than the residual vector in the baseline model – reflecting an increase or decrease in residual damages – depending on its perpendicular component's length and direction relative to the baseline residual vector. The angle between the perpendicular component and the baseline residual vector is from now on denoted by γ. The tighter γ and the shorter the perpendicular component relative to the baseline residual vector, the more likely it is that residual damages are smaller in the presence of diverging preferences. When γ is smaller than 90°, the vectors largely point into the same direction and partially cancel out. When γ is larger than 90°, both vectors point in opposing directions and the

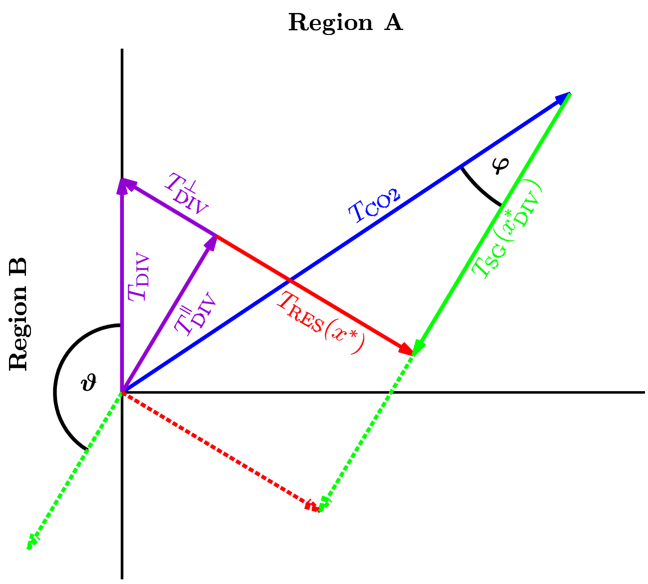

Figure 2: Extension of the Residual Climate Response model. The combination of the solid green vector and the dashed green vector represents regional temperature changes due to optimal SG in the baseline model, the dashed red vector represents the regional residual temperatures in the baseline model. The solid green vector represents regional temperature changes due to optimal SG in the presence of the purple diverging preferences vector. The latter can be decomposed into a component parallel and a component perpendicular to the SG vector. The parallel component changes the optimal SG level. The perpendicular component changes the residual vector. Since the angle ϑ between the diverging preferences vector and the SG vector is larger than 90°, optimal SG decreases relative to the baseline model. The residual vector in the presence of diverging preferences is the red baseline residual vector less the purple perpendicular component. Since the perpendicular component points into the opposite direction of the baseline residual vector, the residual vector in the presence of diverging preferences is longer than the baseline residual vector, i.e. residual damages increase.

perpendicular component necessarily increases residual damages. In Figure 2, we have $\gamma = 180°$, hence residual damages in the example depicted are larger in the presence of diverging preferences. When the perpendicular component is at least twice at large as the baseline residual vector, the perpendicular component necessarily overcompensates for the baseline residual vector and residual damages increase even if the vectors are parallel ($\gamma = 0°$).

Theoretical Result. *Let $T_{\text{DIV}} = T^{\perp}_{DIV} + T^{\parallel}_{DIV}$ be the vector of diverging preferences and its decomposition into the perpendicular and the parallel component.*

1. *The optimal SG level is*

$$x^*_{\text{DIV}} = x^* + \frac{|T_{DIV}| \cdot \cos(\vartheta)}{|T_{SG}|} = x^* + \frac{|T^{\parallel}_{\text{DIV}}|}{|T_{\text{SG}}|} \cdot \text{sign}(\cos(\vartheta)).$$

2. The residual vector is

$$T_{\mathrm{RES}}(x^*_{\mathrm{DIV}}, T_{\mathrm{DIV}}) = T_{\mathrm{RES}}(x^*) - T^{\perp}_{\mathrm{DIV}}$$

3. Residual damages in presence of diverging preferences are smaller than those in absence of diverging preferences if and only if

$$\cos(\gamma) > \frac{1}{2}\frac{|T^{\perp}_{\mathrm{DIV}}|}{|T_{\mathrm{RES}}(x^*)|},$$

where γ is the angle between the perpendicular component $T^{\perp}_{\mathrm{DIV}}$ and the residual vector $T_{\mathrm{RES}}(x^*)$ in the baseline model.

Assessment Metrics

In the baseline RCR model, the performance of optimal SG is captured by a single aspect, the relative effectiveness in damage compensation. In contrast, at least three different aspects of SG performance are of interest in relation to diverging preferences. Analogously to the baseline model, the first aspect is the relative effectiveness of optimal SG in reducing damages for a given diverging preferences scenario. Measuring relative effectiveness necessarily involves the definition of a damage reference level. In the baseline model, the obvious choice for that reference level is damages in absence of SG or, equivalently, damages caused by CO_2. However, damages in absence of SG and damages caused by CO_2 are different damage levels in the presence of diverging preferences. These two reference levels for measuring relative effectiveness give rise to two different metrics in the extended model. The first is the *total damage reduction metric M1*, using total damages (arising from the combination of CO_2 induced temperature changes and diverging preferences) as damage reference level. The *total damage reduction metric* is the one used by Heyen et al. (2015) for illustrating the potential impact of diverging preferences on SG performance. The second one is the *CO_2 damage reduction metric M2*, using damages purely caused by CO_2 induced temperature changes (i.e. total damages less damages from the mere presence of diverging preferences) as damage reference level. The *total damage reduction metric* measures how well optimal SG compensates for damages arising from the differences between regional temperatures in the high CO_2 climate and regionally preferred temperatures. Since the *total damage reduction metric* measures SG effectiveness relative to the regional optima, its maximum value is 100%. The *CO_2 damage reduction metric* measures how well optimal SG compensates for damages arising from differences between regional temperatures in the high CO_2 climate and regional temperatures in the baseline climate. The relative effectiveness in compensating for damages purely caused by CO_2 induced temperature changes can therefore be higher than 100%. In those cases, optimal SG compensates for more damages than CO_2 causes, meaning that residual damages are lower than damages in the baseline climate. Analytical definitions of the metrics can be found in Table 1.

At least two different aspects of SG performance concerning the change in performance across different diverging preferences scenarios are of interest. The first aspect is the minimum climate damages, or, equivalently, the maximum climate welfare that SG can implement. This aspect is captured by residual damages, i.e. damages given optimal SG. In order to compare SG performance across different diverging preferences scenarios, residual damages are normalized for a given diverging preferences scenario to baseline residual damages. This aspect gives rise to the *minimum climate damage metric M3*, measuring how maximum climate welfare changes with the presence of diverging preferences. The second aspect is the gross value of SG, which is captured by the maximum damage reduction SG can achieve. For comparing the gross value across different diverging preferences scenarios, they are normalized to the gross value of SG in the baseline model. This aspect gives rise to the *gross value metric M4*, which measures how the gross value of SG changes with the presence of diverging preferences. Note that in the baseline model, or for any other fixed preferences scenario, maximum climate welfare and the gross value of SG are redundant. However, they are independent across climate preferences, since total damages vary across different climate preferences.

Metrics in the Presence of Diverging Preferences

M1	*M2*	*M3*	*M4*
$\frac{\Delta D(T_{\mathrm{DIV}})}{D(0,T_{\mathrm{DIV}})}$	$\frac{\Delta D(T_{\mathrm{DIV}})}{D(0,T_{\mathrm{DIV}})-\lvert T_{\mathrm{DIV}}\rvert^2}$	$\frac{D(x^*_{\mathrm{DIV}},T_{\mathrm{DIV}})}{D(x^*)}$	$\frac{\Delta D(T_{\mathrm{DIV}})}{D(0)-D(x^*)}$

Table 1: $\Delta D(T_{\mathrm{DIV}})$ expresses the difference in damages between no SG and optimal SG for a given vector of diverging preferences: $\Delta D(T_{\mathrm{DIV}}) = D(0, T_{\mathrm{DIV}}) - D(x^*_{\mathrm{DIV}}, T_{\mathrm{DIV}})$. The metrics are the *total damage reduction metric M1*, the *CO_2 damage reduction metric M2*, the *minimum climate damage metric M3* and the *gross value metric M4*.

The metric in the baseline model is determined by φ, whereas the metrics in the extended model additionally depend on ϑ and the relative length of the diverging preferences vector and the CO_2 vector.

4 Exemplary Implementation of Preferences Scenarios

The exemplary implementation of the extended RCR model delivers a closer examination of how a specific class of diverging preferences scenarios affects SG outcomes. I follow the relevant literature (Moreno-Cruz et al. 2012, Kravitz et al. 2014, Yu et al. 2015, Pfrommer 2018) in defining the regions of the RCR model according to Giorgi and Francisco (2000). I use data from the thirteen climate models which participated in the G1 experiment as defined in the Geoengineering Model Intercomparision Project (Kravitz et al. 2011) for the implementation. Each model performed a run simulating preindustrial climate conditions, a run simulating a climate with four times elevated CO_2 concentration levels, and a run in which SG is used to restore the global mean temperature to the preindustrial level. The regional mean temperatures, averaged over the thirteen models, of the three runs are used to calculate the CO_2 vector and the SG vector.

The SG literature usually normalizes regional temperatures to each region's preindustrial interannual variability when assessing SG (Heyen et al. 2015). However, potential sources for diverging regional preferences derive from phenomena related to absolute temperatures (Burke et al. 2015, Heyen et al. 2015, Newell et al. 2018). Therefore, I use absolute temperatures for the implementation. For comparison, I provide the results when using normalized temperatures in the appendix. The results I obtain for the baseline RCR model when using normalized temperatures are in line with the literature (Moreno-Cruz et al. 2012, Yu et al. 2015). The main difference in results between using normalized and absolute temperatures in the presence of diverging preferences is that optimal SG levels are substantially higher when using normalized temperatures. The difference is explained by low-latitude regions having on average a much smaller interannual variability in mean temperature than high-latitude regions. The upshot is that absolute temperatures should be used for the extended RCR model, unless one has evidence that the impacts one is interested in are better captured by normalized temperatures.

The Scenarios

I focus on scenarios which are based on the premise that, as a rule, high-latitude regions prefer a warmer climate and low-latitude regions prefer a cooler climate relative to the baseline climate. This general scenario structure is both plausible and supported by empirical evidence (Burke et al. 2015, Newell et al. 2018). I group regions into high-latitude, high-mid-latitude, low-mid-latitude and low-latitude bins (see Table 2).

I consider two scenario profiles. In scenario profile A, regions pertaining to the high-latitude group desire a warmer climate than in the baseline climate and regions in the low-latitude group desire a colder one. All mid-latitude regions are content with the baseline climate. The difference in scenario profile B is that regions in the high-mid-latitude group also desire a warmer climate. Implementing two scenario profiles reduces the risk of picking up results which are idiosyncratic to a specific profile and essentially serves as a robustness check. Regions' diverging preferences in the scenario profiles are always equally strong or not present: Each region's desired temperature is one unit above temperatures in the regional baseline climate, one unit below temperatures in the regional baseline climate or equals temperatures in the regional baseline climate.

The last part of the theoretical result implies non-linear effects in the relative length of the diverging preference vector on the change in SG outcomes. In order to capture these effects, I consider different strengths of diverging preferences (i.e. different sizes of what constitutes "one unit") for each scenario profile. In other words, a scenario profile defines regions' diverging preferences relative to each other (the direction of the diverging preference vector) and the preference strength defines the magnitude of the desired temperature deviations from the baseline climate (the vector's relative length). Jointly, the two determine a diverging preferences scenario. I define preference strength in percent of the average regional temperature change caused by CO_2. Due to the RCR model's linearity, this relative definition leads to well-defined results. As an example, say CO_2 driven average regional warming is 4°C. A strength of 50% in diverging preferences

Grouping of Regions and Scenario Profiles

	Each Region's			Scenario Profile	
	Avg. Lat.	Opt. SG Level	Res. Temp.	A	B
Greenland	67.5N	1.09	0.44	1.0	1.0
Alaska	66.0N	1.08	0.50	1.0	1.0
North. Europe	61.5N	1.09	0.33	1.0	1.0
North Asia	60.0N	1.12	0.66	1.0	1.0
WN America	45.0N	1.04	0.06	0.0	1.0
CN America	40.0N	1.07	0.23	0.0	1.0
Central Asia	40.0N	1.02	-0.08	0.0	1.0
Tibet	40.0N	0.99	-0.29	0.0	1.0
Mediterranean	39.0N	1.04	0.05	0.0	0.0
SS America	38.0S	1.02	-0.04	0.0	0.0
EN America	37.5N	1.03	0.01	0.0	0.0
South. Australia	37.5S	1.01	-0.09	0.0	0.0
East Asia	35.0N	1.03	0.02	0.0	0.0
Sahara	24.0N	0.99	-0.21	-1.0	-1.0
Central America	20.0N	0.97	-0.29	-1.0	-1.0
South Asia	17.5N	0.96	-0.35	-1.0	-1.0
Southern Africa	11.5S	0.97	-0.27	-1.0	-1.0
North. Australia	9.5S	0.95	-0.38	-1.0	-1.0
Southeast Asia	4.5N	0.93	-0.35	-1.0	-1.0
Amazon Basin	4.0S	0.99	-0.20	-1.0	-1.0
Western Africa	3.0N	0.97	-0.31	-1.0	-1.0
Eastern Africa	3.0N	0.97	-0.30	-1.0	-1.0

Table 2: Regions are ordered according to their average latitude and grouped into 'high-latitude', 'high-mid-latitude', 'low-mid-latitude' and 'low-latitude' bins. The left columns states each region's average latitude, each region's preferred SG level and each region's residual temperature under optimal SG. The right columns states the scenario profiles A and B.

then means that regions desiring a temperature change relative to the regional baseline climate do so by 2°C.

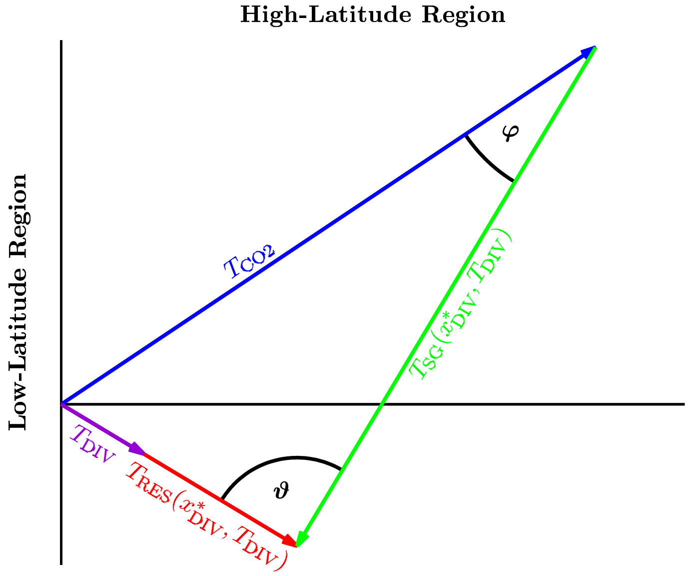

Figure 3: Two-region representation of the basic structure underlying the scenarios. Given baseline climate preferences, optimal SG undercompensates for temperature in the high-latitude region, but overcompensates for temperature in the low-latitude region. The high-latitude region desires a warmer climate and the low-latitude region a colder climate compared to the baseline climate. In this specific example, the SG vector and the diverging preferences vector are perpendicular ($\vartheta = 90°$), hence the optimal SG level does not change relative to the baseline model. Due to the structure of the scenario, the perpendicular component of the diverging preferences vector (note that there is no parallel component in this specific example) and the baseline residual vector (which here is the combination of the violet diverging preferences vector and the red residual vector in the presence of diverging preferences) point into the same direction. The angle γ between the two vectors is 0°, since they are parallel, and the former compensates in part for the latter. In a two-region example, γ can only attain 0°and 180°, which is not the case when a higher number of regions is involved.

The theoretical result mirrors the result from the two-region example (Figure 3). It states that, for residual damages to decrease due to diverging preferences, the angle γ between the perpendicular component of the diverging preferences vector and the baseline residual vector has to be between $-90°$ and 90°, while at the same time the former vector has to be short enough relative to the latter. Taking into account that in the two-region example γ is 0°, the theoretical result exactly predicts the outcomes in two-region example. The theoretical result tells us that the pattern will qualitatively hold beyond two-region examples if only γ is between $-90°$ and 90°. It is self-evident that for two-region examples γ is always in that range (even exactly 0°). Whether this is the case in a specific multi-region scenario ultimately depends on the details of the scenario profile, such as which regions are considered high-latitude, which are considered low-latitude, how strong individual regions' preferences are and the differential impacts of SG along other dimensions than latitude. However, for the type of scenario profile under consideration, it seems likely that γ is in that range – at least for many scenario profiles. The implementation delivers the actual γ for two specific scenario profiles and delivers the range of preference strengths for which residual damages are smaller than in the baseline scenario, i.e. for which the *minimum climate damage metric* evaluates to below 100%. Lastly, the implementation may reveal whether there are similar patterns concerning other aspects of SG performance.

Results of the Implementation

I report the angles φ, ϑ and γ and the optimal SG level in absence and presence of diverging preferences. I report the results for the four metrics for several preference strengths between 0% and 100%. The angle φ and optimal SG in absence of diverging preferences x^* are independent of the scenario profile. Given x^*, ϑ (which is determined by the scenario profile) and the strength of diverging preferences jointly determine the SG level in presence of diverging preferences. Optimal SG is linear in the preference strength and it suffices to only explicitly report optimal SG for one particular strength. The optimal SG level x^*_{DIV} reported refers to a preference strength of 100%.[6]

Optimal SG Levels and Angles between Vectors

Baseline Model			Scenario Profile A			Scenario Profile B		
x^*	φ	M	x^*_{DIV}	ϑ	γ	x^*_{DIV}	ϑ	γ
1.03	2.9°	99.7	1.16	80.1°	22.6°	0.96	94.5°	38.2°

Table 3: The left columns state the optimal SG level x^* in the baseline model, the angle φ between the CO_2 and the SG vector and the metric M (in percent) from the baseline model. The middle columns and the right columns state, for the scenario profiles A and B, respectively, the optimal SG level x^*_{DIV} when preference strength is 100%, the angle ϑ between the SG vector and the diverging preferences vector and the angle γ between the perpendicular component of the diverging preferences vector and the baseline residual vector.

The angle φ between the CO_2 vector and the SG vector is small (2.9°). In the absence of diverging preferences, optimal SG is therefore close to one and the relative effectiveness of SG in reducing CO_2 induced damages (metric M) is close to 100%. For both scenario profiles, the angle ϑ between the SG vector and the diverging preferences vector is close to 90° (scenario profile A: $\vartheta = 80.1°$, scenario profile B: $\vartheta = 94.5°$). Consequently, optimal SG differs only moderately between the absence and presence of diverging preferences. Even for very strong diverging preferences (100%), optimal SG is only 0.13 higher than in the baseline model (+12.6%) for scenario profile A and only 0.07 lower than in the baseline model (−6.8%) in scenario profile B. Optimal SG is larger

[6] Optimal SG for other preference strengths can be obtained by linear interpolation. For example, optimal SG for a strength of 50% is $x^* + 0.5 \cdot (x^*_{\text{DIV}} - x^*)$.

in scenario profile A than in scenario profile B, corresponding to the on net stronger preferences for higher temperatures in scenario profile B.

The *total damage reduction metric M1* and the *minimum climate damage metric M3* exhibit a similar pattern for both scenario profiles. According to both metrics, the performance of optimal SG increases for weak diverging preferences relative to the baseline scenario and then decreases again for strong diverging preferences. Minimum residual damages are attained for a preference strength of 6.4% in scenario profile A and for a preference strength of 4.7% in scenario profile B. Residual damages start to exceed baseline residual damages for a preference strength of 12.8% in scenario profile A and for 9.4% in scenario profile B. The corresponding preferences strengths for the *total damage reduction metric* are very similar (differences $< 0.2\%$). Residual damages become very large relative to baseline residual damages for very strong diverging preferences, up to a factor of 184 and of 255 higher in scenario profile A and B, respectively. Relative effectiveness in compensating for total damages is substantially lower than in the baseline scenario for very strong diverging preferences, down to 72.9% and 56.9% in scenario profile A and B, respectively. However, for both scenario profiles, SG can compensate for at least 86.1% of total damages when the preference strength is no higher than 50% and for at least 96.9% when the preference strength is no higher than 25%.

Results for the Four Metrics

Preference	Scenario Profile A				Scenario Profile B			
Strength	*M1*	*M2*	*M3*	*M4*	*M1*	*M2*	*M3*	*M4*
0%	99.7	99.7	100.0	100.0	99.7	99.7	100.0	100.0
5%	99.9	100.0	18.7	101.2	99.9	100.0	38.5	99.3
10%	99.8	100.4	42.3	102.5	99.7	100.4	117.3	98.6
15%	99.5	100.7	170.7	103.8	99.1	100.8	336.5	98.0
25%	98.2	101.4	742.1	106.5	96.9	101.5	1196	96.6
50%	91.6	103.2	4005	113.1	86.1	103.6	5801	93.3
75%	82.4	105.1	9890	120.0	71.5	106.0	13915	90.1
100%	72.8	106.9	18396	127.1	56.9	108.7	25539	86.9

Table 4: The metrics are the *total damage reduction metric M1*, the *CO_2 damage reduction metric M2*, the *minimum climate damage metric M3* and the *gross value metric M4*. *M1* and *M2* are in percent of their respective reference damage level. *M3* and *M4* are in percent relative to the respective outcome in the baseline scenario. Results are given in incremental steps in the strength of diverging preferences. For example, a strength of 50% means that one unit of diverging preferences corresponds to half the average regional temperature change from climate change. If regional mean temperature rises by 4°C in absence of SG, a strength of 50% corresponds to low-latitude regions having a diverging temperature preference of -2°C and high-latitude regions having a diverging temperature preference of $+2$°C in scenario profile A.

The changes in the *CO_2 damage reduction metric M2* and the *gross value metric M4* are monotone in the strength of diverging preferences. The relationship between the two metrics and preference strength is almost linear. Relative effectiveness in compensating for damages caused by CO_2 increases in both scenarios. For very strong diverging preferences, it is 106.9% in scenario A and 108.7% in scenario B. Relative effectiveness in compensating for damages caused by CO_2 is at least 100% in both scenarios when the strength of diverging preferences is at least 5%. The gross value of SG increases in scenario A, but decreases in scenario B. For very strong diverging preferences, the gross value of SG is 27.1% higher than the gross value of SG in absence of diverging preferences in the former and 13.1% lower than the gross value of SG in absence of diverging preferences in the latter.

Explanation and Interpretation of the Results

The *minimum climate damage metric M3* measures the maximum climate welfare SG can implement in a scenario relative to maximum climate welfare in the baseline scenario and is determined by the ratio of the respective residual damages. Since $\gamma = 22.6°$ in scenario profile A and $\gamma = 38.2°$ in scenario profile B, the *minimum climate damage metric* follows the pattern presumed in advance: The metric is smaller than 100% for weak diverging preferences and larger than 100% for moderate to strong diverging preferences in both scenario profiles. In both scenario profiles, γ is well within the range of $(-90°, 90°)$. One can therefore expect γ to be generally within this range for scenario profiles following the basic premise, implying that one can expect the pattern to generally hold for such scenario profiles. The values $\cos(\gamma)$ attains in the scenario profiles are close to one (0.92 in scenario profile A, 0.78 in scenario profile B). The preference strengths for which the *minimum climate damage metric* is smaller than 100% in scenario profile A are therefore close to the maximum of such strengths possible for any scenario profile. For strong diverging preferences, the perpendicular component of the diverging preferences vector is, irrespective of its direction, much larger than the baseline residual vector. The residual vector is then almost identical to the perpendicular component. Hence, the *minimum climate damage metric* becomes very large for strong diverging preferences, irrespective of the scenario profile.

The *total damage reduction metric M1* measures the relative effectiveness of optimal SG in compensating for total damages. It can be expressed as

$$1 - \frac{|T_{\text{RES}}(x^*) - T_{\text{DIV}}^{\perp}|^2}{|T_{\text{CO2}} - T_{\text{DIV}}|^2}.$$

In both scenario profiles, the relative effectiveness in compensating for total damages is higher than in the baseline scenario for weak diverging preferences and lower for moderate to strong ones. Since the CO_2 vector is comparatively large, the relative change in total damages (the denominator in the expression) from absence to presence of diverging preferences is comparatively small, while the relative change in residual damages (the nominator in the expression) is comparatively high. Therefore, residual damages govern the qualitative behavior of the *total damage reduction metric* at least for weak and moderate preferences, resulting in the observed pattern in both scenarios. This intuition is valid except for the rather unrealistic case of the perpendicular component being trivial.

The CO_2 *damage reduction metric M2* measures the relative effectiveness of optimal SG in compensating for damages caused by CO_2. It can be expressed as

$$1 + \frac{|T_{\text{DIV}}|^2 - |T_{\text{RES}}(x^*_{\text{DIV}}, T_{\text{DIV}})|^2}{|T_{\text{CO2}} - T_{\text{DIV}}|^2 - |T_{\text{DIV}}|^2}.$$

From this characterization, it is evident that the CO_2 *damage reduction metric* evaluates to 100% when the diverging preferences vector and the residual vector are of the same length, i.e. when SG restores the damage level of the baseline climate, thereby exactly compensating for pure CO_2 damages.[7] In both scenario profiles, the CO_2 *damage reduction metric* exceeds 100% for even weak preferences. When $\gamma \in (-90°, 90°)$, the residual vector becomes shorter for weak preferences. The closer γ is to 0°, the faster the diverging preference vector becomes longer than the residual vector with increasing preference strength.[8] Since the baseline residual vector is very short and γ is in both scenarios of small to moderate size, the CO_2 *damage reduction metric* becomes larger than 100% for even weak preferences in both scenario profiles.

The *gross value metric M4* measures the gross value of SG in a scenario relative to its gross value in the baseline scenario. It can be expressed as

$$\frac{|T_{\text{CO2}} - T_{\text{DIV}}|^2 - |T_{\text{RES}}(x^*) - T^{\perp}_{\text{DIV}}|^2}{|T_{\text{CO2}}|^2 - |T_{\text{RES}}(x^*)|^2}.$$

The *gross value metric* exceeds 100% for all preference strengths in scenario A and is below 100% for all preference strengths in scenario B, coinciding with optimal SG increasing in scenario profile A and decreasing in scenario profile B relative to optimal SG in the baseline scenario. As a heuristic, the parallel component of the diverging preferences vector and the CO_2 vector can be considered as (anti)parallel, since the angle φ between the SG and the CO_2 vector is very small. Intuitively, the parallel component then increases total damages when it causes optimal SG to rise and reduces total damages when it causes optimal SG to fall, while not affecting the residual vector – in other words, the parallel component increases the gross value of SG when it causes optimal SG to rise and decreases the gross value of SG when it causes optimal SG to fall. However, this reasoning leaves out the influence of the perpendicular component on the change in gross value of SG. This means that, while a positive change in optimal SG increases the gross value of SG *ceteris paribus*, a positive (negative) change in optimal SG relative to optimal SG in the baseline scenario does not necessarily correspond to a gross value of SG higher (lower) than in the baseline scenario.

The numerical implementation and its discussion yield two main results. Firstly, the performance of optimal SG in presence of diverging preferences may increase or

[7] Note that the distribution of damages among regions will then be different than in the baseline climate.

[8] It can be easily be shown that the diverging preference vector always becomes longer than the residual vector for strong enough preferences when $\gamma \in (-90°, 90°)$: It holds that

$$|T_{\text{DIV}}|^2 - |T_{\text{RES}}(x^*_{\text{DIV}}, T_{\text{DIV}})|^2 = |T^{\parallel}_{\text{DIV}}|^2 - |T_{\text{RES}}(x^*)|^2 + T_{\text{RES}}(x^*) \cdot T^{\perp}_{\text{DIV}}.$$

The last term on the right-hand side is positive if $\gamma \in (-90°, 90°)$, from which the statement follows.

decrease relative to its performance in absence of diverging preferences. Which one is the case depends on the diverging preferences scenario, the aspect of SG performance one is interested in and, for the aspect of relative effectiveness of optimal SG, on the damage reference level against which relative effectiveness is measured. This underpins the importance of specifying the aspect of SG performance and, if applicable, the damage reference level, when assessing SG in presence of diverging preferences. Specifically, the *absolute damage reduction metric* and the *minimum climate damage metric* are different ways of measuring how well SG can minimize residual damages – in the former case for a specific diverging preferences scenario relative to total damages, in the later case relative to residual damages in the absence of diverging preferences. Both indicate a reduction in SG performance in both scenario profiles for diverging preferences which are at least moderately strong. In contrast, the *CO_2 damage reduction metric* measures how well SG can bring back damages to pre-climate change levels. This metric indicates an increase in SG performance in the presence of diverging preferences for both scenario profiles. The *gross value metric* measures by how much SG can reduce damages, indicating that the change in its ability to do so in the presence of diverging preferences is linked to the change in optimal SG.

Secondly, optimal climate welfare in the presence of diverging preferences generally compares differently relative to different welfare reference levels. Optimal climate welfare in the presence of diverging preferences is for both scenario profiles lower compared to optimal climate welfare in the baseline scenario, when the magnitude of the desired temperature deviations is at least moderate, but higher for small desired temperature deviations. In contrast, optimal climate welfare in the presence of diverging preferences is for both scenario profiles higher compared to climate welfare in the baseline climate (i.e. in the climate before CO_2 driven temperature changes set in) for all but very small magnitudes of desired temperature deviations. I have argued that these welfare comparisons are very likely (in case of the *minimum climate damage metric* even sure) to persist when the angle γ is small. Since γ can be expected to be small in scenario profiles in which high-latitude regions prefer a colder climate and low-latitude regions prefer a warmer one, these welfare comparisons can be expected to hold in scenarios based on such profiles more generally. These results demonstrate that the assessment of how diverging climate preferences affect optimal climate welfare very likely depends on which welfare reference levels one chooses to employ.

5 Conclusion

Solar geoengineering (SG) has the potential to compensate for increased temperatures from climate change on a global level. However, SG has heterogeneous impacts at the regional level. Until now, studies examining these regional differences have employed the assumption that regions' temperature preferences correspond to a common baseline climate, e.g. to preindustrial or 1990 climate conditions. This assumption has been criticized (Heyen et al. 2015) and conflicts with empirical evidence supporting globally generalizable relationships between economic productivity and absolute temperature levels (Burke et al. 2015, Newell et al. 2018).

In this paper, I extended the Residual Climate Response (RCR) model (Moreno-Cruz et al. 2012) for assessing regional SG differences by formally introducing the possibility of regional temperature preferences diverging from the baseline climate, building on an illustrative example by Heyen et al. (2015). In the key theoretical result, I showed that the impact of these diverging preferences can be split into two components. The first component changes the optimal SG level, but does neither change optimal climate welfare nor affect regional disagreement over SG. The second component leaves the optimal SG level unaffected, but changes regional disagreement over SG and optimal climate welfare. This decomposition helps in understanding how specific diverging preferences affect globally optimal SG and the disagreement over SG by different regions. I introduced metrics for measuring three independent aspects of SG performance. The first aspect is the relative effectiveness of SG in reducing damages, the second aspect is the change of optimal climate welfare relative to the absence of diverging preferences and the third aspect is the change in the gross value of SG relative to the absence of diverging preferences.

I numerically implemented the extended RCR model, focusing on scenarios in which high-latitude regions prefer a warmer climate and low-latitude regions prefer a cooler climate relative to the baseline climate – a scenario structure which is both plausible and supported by empirical evidence on aggregate economic output (Burke et al. 2015, Newell et al. 2018). The numerical implementation yields that the performance of optimal SG relative to the baseline scenario can change in either direction in the presence of diverging preferences and the change generally depends on the aspect of SG performance one is interested in. The results demonstrate that diverging climate preferences are not necessarily detrimental to central aspects of SG performance, at least when optimal SG can be implemented. However, one can expect optimal climate welfare to be lower in the presence of diverging temperature preferences of substantial magnitudes than in the absence of diverging preferences.

One should keep in mind that both the baseline and the extended RCR model are deliberatively simple in nature. They derive their usefulness from conceptual understanding and from identifying first-order effects. Additionally, the results I obtained for the type of scenario I considered are temperature specific and do not necessarily hold for other climate variables, for which diverging preferences may be relevant as well, like precipitation. Several lines of research should be pursued in the future for further increasing the understanding of the relationship between diverging climate preferences and the assessment of SG. Firstly, I concentrated on the outcomes for globally optimal SG. Examining the potential impact on Pareto optimal SG levels and on SG levels in the free-driver outcome (Weitzman 2015), as well as examining the respective welfare implications, will lead to a more complete picture regarding the relationship between diverging climate preferences and the assessment of SG. Secondly, investigating the potential impact of diverging preferences on coalition formation (Ricke et al. 2013), in particular coalitions based on similar latitudes, may lead to further insights into the strategic dimensions of SG. Lastly, a further conceptual extension seems desirable: Regional damages may be conceptualized as being a combination of damages deriving from the variability-adjusted temperature difference to regional baseline climate temperatures and damages deriving from the absolute temperature difference to some absolute temperature preference. A generalization incorporating both, however, necessarily opens up the question of the relative weight of both types of damages.

Appendix

A Implementation of the Extended RCR Model Using Normalized Temperatures

In this appendix, I provide the results for the extended RCR model when using temperatures normalized to regional interannual variability. The main difference in results between using normalized and absolute temperatures in the presence of diverging preferences is that optimal SG levels are substantially higher when using normalized temperatures (compare Table 3 and Table 5). The differences in optimal SG between using normalized and absolute temperatures is explained by low-latitude regions having on average a much smaller interannual variability in mean temperature than high-latitude regions. Using normalized temperatures, diverging temperature preferences of the same absolute magnitude then translate into a much larger magnitudes in units of interannual variability for low-latitude regions than for high-latitude regions, while this effect is not present when using absolute temperatures.

Optimal SG Levels and Angles between Vectors

Baseline Model			Scenario A			Scenario B		
x^*	φ	M	x^*_{DIV}	ϑ	γ	x^*_{DIV}	ϑ	γ
0.99	2.5°	99.8	1.68	59.1°	23.2°	1.60	72.3°	32.7°

Table 5: Results for the implementation of the extended RCR Model using normalized temperatures. Corresponds to Table 3. The left columns state the optimal SG level x^* in the baseline model, the angle φ between the CO_2 and the SG vector and the metric M (in percent) from the baseline model. The middle columns and the right columns state for the scenarios A and B, respectively, the optimal SG level x^*_{DIV} when preference strength is 100%, the angle ϑ between the SG vector and the diverging preferences vector and the angle γ between the perpendicular component of the diverging preferences vector and the baseline residual vector.

The results for the four metrics when using normalized temperatures (compare Table 6) are qualitatively the same as when using absolute temperatures. The same general patterns can be observed when using normalized temperatures as when using absolute temperatures, regarding the inversely U-shaped relationship between SG performance as measured by metric *M1*, as well as metric *M3* and preference strength, regarding *M2* exceeding 100% for all but very weak diverging preferences and regarding the relationship between the change in optimal SG and metric *M4*.

Results for the Four Metrics

Preference	Scenario A				Scenario B			
Strength	*M1*	*M2*	*M3*	*M4*	*M1*	*M2*	*M3*	*M4*
0%	99.8	99.8	100.0	100.0	99.8	99.8	100.0	100.0
5%	99.9	100.1	17.3	107.1	99.9	100.1	30.0	106.2
10%	99.9	100.7	58.4	114.5	99.7	100.6	134.7	112.7
15%	99.6	101.4	223.4	122.1	99.3	101.3	414.3	119.3
25%	98.7	103.1	924.9	138.0	97.9	102.7	1497	133.1
50%	95.3	108.7	4845	182.0	92.7	107.6	7265	170.8
75%	91.4	115.5	11863	232.2	86.9	113.4	17403	213.3
100%	87.7	122.9	21976	288.5	81.6	119.9	31911	260.5

Table 6: Results for the implementation of the extended RCR Model using normalized temperatures. Corresponds to Table 4. The metrics are the *total damage reduction metric M1*, the *CO_2 damage reduction metric M2*, the *minimum climate damage metric M3* and the *gross value metric M4*. *M1* and *M2* are in percent of their respective damage baseline. *M3* and *M4* are in percent relative to the respective outcome in the baseline scenario. Results are given in incremental steps in the strength of diverging preferences. For example, a strength of 50% means that one unit of diverging preferences corresponds to half the average regional temperature change from climate change. If regional mean temperature rises by 4°C in absence of SG, a strength of 50% corresponds to low-latitude regions having a diverging temperature preference of −2°C and high-latitude regions having a diverging temperature preference of +2°C in scenario A.

B Proofs

Proof of the Theoretical Result.

1. Since the residual damages are convex in the SG level, the first equation from the first part follows from computing the first order conditions of the residual damages with respect to the SG level. The second equation follows because $T^{\parallel}_{\text{DIV}}$ is the projection of T_{DIV} onto T_{SG}.

2. Plugging in the optimal SG level into

$$T_{\text{RES}}(x, T_{\text{DIV}}) = (T_{\text{CO2}} - T_{\text{DIV}}) + x \cdot T_{\text{SG}}$$

yields

$$\begin{aligned} T_{\text{RES}}(x^*_{\text{DIV}}, T_{\text{DIV}}) &= (T_{\text{CO2}} - (T^{\perp}_{\text{DIV}} + T^{\parallel}_{\text{DIV}})) + (x^* + \frac{|T_{\text{DIV}}| \cdot \cos(\vartheta)}{|T_{\text{SG}}|}) \cdot T_{\text{SG}} \\ &= (T_{\text{CO2}} + x^* \cdot T_{\text{SG}}) - T^{\perp}_{\text{DIV}} - T^{\parallel}_{\text{DIV}} + \frac{|T_{\text{DIV}}| \cdot \cos(\vartheta)}{|T_{\text{SG}}|} \cdot T_{\text{SG}} \end{aligned}$$

Because of

$$\cos(\vartheta) = \frac{T_{\mathrm{SG}} \bullet T_{\mathrm{DIV}}}{|T_{\mathrm{SG}}| \cdot |T_{\mathrm{DIV}}|},$$

it holds that

$$T_{\mathrm{DIV}}^{\parallel} = \frac{|T_{\mathrm{DIV}}| \cdot \cos(\vartheta)}{|T_{\mathrm{SG}}|} \cdot T_{\mathrm{SG}}$$

and the second part follows.

3. The third part follows from

$$\begin{aligned}
&|T_{\mathrm{RES}}(x^*) - T_{\mathrm{DIV}}^{\perp}|^2 < |T_{\mathrm{RES}}(x^*)|^2 \Leftrightarrow \\
&|T_{\mathrm{RES}}(x^*)|^2 - 2 \cdot T_{\mathrm{RES}}(x^*) \bullet T_{\mathrm{DIV}}^{\perp} + |T_{\mathrm{RES}}(x^*)|^2 < |T_{\mathrm{RES}}(x^*)|^2 \Leftrightarrow \\
&\tfrac{1}{2} \cdot |T_{\mathrm{DIV}}^{\perp}|^2 < T_{\mathrm{RES}}(x^*) \bullet T_{\mathrm{DIV}}^{\perp} = |T_{\mathrm{RES}}(x^*)| \cdot |T_{\mathrm{DIV}}^{\perp}| \cdot \cos(\gamma).
\end{aligned}$$

□

References

Ban-Weiss, G. A. and Caldeira, K. (2010). "Geoengineering as an optimization problem". In: *Environmental Research Letters* 5.3, p. 034009.

Burke, M., Hsiang, S. M., and Miguel, E. (2015). "Global non-linear effect of temperature on economic production". In: *Nature* 527.7577, pp. 235–239.

Giorgi, F. and Francisco, R. (2000). "Uncertainties in regional climate change prediction: a regional analysis of ensemble simulations with the HADCM2 coupled AOGCM". In: *Climate Dynamics* 16.2-3, pp. 169–182.

Graff Zivin, J. and Neidell, M. (2014). "Temperature and the allocation of time: Implications for climate change". In: *Journal of Labor Economics* 32.1, pp. 1–26.

Heyen, D. (2016). "Strategic conflicts on the horizon: R&D incentives for environmental technologies". In: *Climate Change Economics* 7.04, p. 1650013.

Heyen, D., Wiertz, T., and Irvine, P. J. (2015). "Regional disparities in SRM impacts: the challenge of diverging preferences". In: *Climatic Change* 133.4, pp. 557–563.

Irvine, P. J., Ridgwell, A., and Lunt, D. J. (2010). "Assessing the regional disparities in geoengineering impacts". In: *Geophysical Research Letters* 37.18.

Kravitz, B., Robock, A., Boucher, O., Schmidt, H., Taylor, K. E., Stenchikov, G., and Schulz, M. (2011). "The geoengineering model intercomparison project (GeoMIP)". In: *Atmospheric Science Letters* 12.2, pp. 162–167.

Kravitz, B., MacMartin, D. G., Robock, A., Rasch, P. J., Ricke, K. L., Cole, J. N., Curry, C. L., Irvine, P. J., Ji, D., Keith, D. W., et al. (2014). "A multi-model assessment of regional climate disparities caused by solar geoengineering". In: *Environmental Research Letters* 9.7: 074013, p. 074013.

Lunt, D. J., Ridgwell, A., Valdes, P. J., and Seale, A. (2008). ""Sunshade World": A fully coupled GCM evaluation of the climatic impacts of geoengineering". In: *Geophysical Research Letters* 35.12.

Moreno-Cruz, J. B., Ricke, K. L., and Keith, D. W. (2012). "A simple model to account for regional inequalities in the effectiveness of solar radiation management". In: *Climatic change* 110.3-4, pp. 649–668.

Newell, R., Prest, B., and Sexton, S. (2018). "The GDP-Temperature Relationship: Implications for Climate Change Damages". In: *RFF Working Paper.* Available at: http://www.rff.org/research/publications/gdp-temperature-relationship-implications-climate-change-damages.

Pasztor, J. (2017). "The Need for Governance of Climate Geoengineering". In: *Ethics & International Affairs* 31.4, pp. 419–430.

Pfrommer, T. (2018). "A model of solar radiation management liability". In: *AWI Discussion Paper Series* 644. Available at: https://www.uni-heidelberg.de/md/awi/institut/awlecture/dp644.pdf.

Ricke, K. L., Moreno-Cruz, J. B., and Caldeira, K. (2013). "Strategic incentives for climate geoengineering coalitions to exclude broad participation". In: *Environmental Research Letters* 8.1, p. 014021.

Ricke, K. L., Morgan, M. G., and Allen, M. R. (2010). "Regional climate response to solar-radiation management". In: *Nature Geoscience* 3.8, pp. 537–541.

Robock, A. (2008). "20 reasons why geoengineering may be a bad idea". In: *Bulletin of the Atomic Scientists* 64.2, pp. 14–18.

Robock, A., Oman, L., and Stenchikov, G. L. (2008). "Regional climate responses to geoengineering with tropical and Arctic SO2 injections". In: *Journal of Geophysical Research: Atmospheres (1984–2012)* 113.D16.

Schlenker, W. and Roberts, M. J. (2009). "Nonlinear temperature effects indicate severe damages to US crop yields under climate change". In: *Proceedings of the National Academy of sciences* 106.37, pp. 15594–15598.

Shepherd, J. G. (2009). *Geoengineering the climate: science, governance and uncertainty.* Royal Society.

Weitzman, M. L. (2015). "A Voting Architecture for the Governance of Free-Driver Externalities, with Application to Geoengineering". In: *The Scandinavian Journal of Economics* 117.4, pp. 1049–1068.

Yu, X., Moore, J. C., Cui, X., Rinke, A., Ji, D., Kravitz, B., and Yoon, J.-H. (2015). "Impacts, effectiveness and regional inequalities of the GeoMIP G1 to G4 solar radiation management scenarios". In: *Global and Planetary Change* 129, pp. 10–22.