How to predict environmental consciousness?

A test of the most influential predictors and best performing methods.

Jan Smelik

Abstract

To help organisations model the extent to which consumers are environmental conscious, this study identifies the best practices in predicting individual-level environmental consciousness. A data set of 12.820 observations over 27 countries allows drawing generalised conclusions. Two predictor sets are tested, sociodemographic factors and psychological characteristics, and three modelling methods are applied. These methods are ordinary least squares regressions, random intercept models and random forest models. A cross-comparison is made to identify the strongest combination. Also the influence of individual predictors is evaluated.

It is found that psychological characteristics have better predictive capabilities than sociodemographics in this case. The recommendation is to steer away from the use of sociodemographcis if possible. Random forests have proven to be superior to simple techniques in predicting environmental consciousness, regardless of the set of predictors included. Personal values, especially universalism, achievement, and conformity are very strong predictors. Also the consumer-specific values ethnocentrism, quality consciousness and health consciousness are influential in the prediction. Regarding the sociodemographics, only income stands out. These concrete findings might help organisation improving their targeting and segmentation strategies.

ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics
Master Thesis Data Science and Marketing Analytics

Supervisor: M.G. de Jong
Second assessor: K. Gruber
Date final version: February 28, 2022
Student number: 448345

1 Introduction

Today's world is facing enormous problems. One pressing nonetheless unresolved issue is that of climate change. The projections for an increase in global temperature of over 1.5 ºC at 2100 range from frightening to devastating. Only if we pull out all the stops, we might still be able to limit global warming to 1.5 ºC (Masson-Delmotte et al., 2019, 2021). Therefore, researchers, policymakers, and businesses need to join forces to do everything possible to avoid this doomsday scenario. Meanwhile, the world of predictive modelling is evolving rapidly with the advent of powerful computers, big data and machine learning. The variety of applications is hardly imaginable, and indeed, in many fields of scientific research and regarding societal problems, opportunities remain unexploited (Zhou et al., 2017).

This study is one of many small steps in the right direction by utilising some possibilities offered by predictive modelling. More specifically, it examines the best approach to predict whether and to which extent consumers are environmentally conscious. To do so, two important predictor sets are tested, namely sociodemographics and psychological characteristics. These predictors will be modelled using different predictive methods, that is an ordinary least squares regression, mixed models and random forests. The aim is to gain insights into the best practices of predicting environmental consciousness, in order to propose concrete recommendations. The findings could offer valuable insights for policymakers and businesses when implementing environmental policies such as marketing strategies for green products or increasing sustainable energy use of households. Hopefully, this will help to avert "the greatest threat to global security" (Parry, 2020).

To specify the research problem, it is essential to define the concept of environmental consciousness. Schlegelmilch et al. (1996) defined it as: "an evaluation of, or an attitude towards facts, one's own behaviour or other's behaviour with consequences for the environment." Similarly, Kollmuss and Agyeman (2002) define it as a complex system of environmental attitudes, knowledge, and values. Sánchez and Lafuente (2010) did great work in conceptualising environmental consciousness. They defined four dimensions: the affective, dispositional, cognitive, and active dimensions. The first dimension includes the perceived importance of the environment, the second the personal norm and perceived costs, the third environmental knowledge and the fourth encompasses pro-environmental behaviour. In the current study, environmental consciousness includes the affective and dispositional dimensions of Sánchez and Lafuente (2010). These dimensions will be measured with three questions: one question on the personal costs of pro-environmental measures and two on feelings of anger regarding environmental misconduct.

From the above, the research question follows:

Which model predicts environmental consciousness most accurately?

This question is answered by comparing the predictive power of two sets of predictors: sociodemographics and psychological characteristics. Also, the influence of individual predictors is reviewed. Simultaneously, the performance of the predictive methods is compared. This is summarised by the three following subquestions:

RQ1: Which set of predictors - sociodemographics or psychological characteristics - predicts environmental consciousness the most accurately?

RQ2: Which method has the best capabilities to predict environmental consciousness given the set of predictors?

RQ3: Which individual predictors are most influential in predicting environmental consciousness?

It is found that psychological characteristics have better predictive capabilities in this case than sociodemographics. The evidence even suggests that there is no need to include sociodemographic variables when including psychological variables. It is also proven that random forests have far better capabilities in predicting environmental consciousness than the other methods, regardless of the predictors included. Personal values, especially universalism, achievement, and conformity are very strong predictors. Also the consumer-specific values ethnocentrism, quality consciousness and health consciousness are influential in the prediction. Regarding the sociodemographics, only income stands out.

A great deal of work has been done in identifying which factors influence environmental consciousness. The relationship between individual sociodemographics and environmental consciousness is explored in great detail (see for example the meta-analysis of Gifford and Nilsson (2014)). Also, outstanding work has been done on how much of the variance of environmental consciousness is explained by sociodemographics together (Diamantopoulos et al., 2003). In the last two decades, the effect of personality traits on environmental consciousness is addressed (see Gifford and Nilsson, 2014; Soutter et al., 2020). Some researchers also looked into the effect of the country-level variables on the environmental consciousness of individuals, but the evidence is rather scarce (see Wang, 2017; Gifford and Nilsson, 2014; Spaargaren, 2011; Peattie, 2010). Moreover, as Soutter et al. (2020) have recently concluded: "Only a few studies have attempted to combine elements from separate disciplines". This is exactly the void that the current study aims to fill.

Notwithstanding the outstanding work that has been done, there is another issue regarding the environmental literature. Diamantopoulos et al. (2003) note that about eighty per cent of the sustainability studies focus on US consumers, while climate change is of course a global problem (also see Vicente-Molina et al., 2013). Sharma and Bansal (2013) state the obvious: consumers in upcoming economies are different from consumers in more developed countries. Also, Wang (2017) stresses the use of cross-national data since this can mitigate possible biases between countries. Thus, it appears that environmental research requires a more cross-national approach. The current study, therefore, makes use of data from 27 countries across 4 continents.

Finally, part of the existing literature fails to address some methodological issues. Xiao and McCright (2007) note that the literature on the relationship between socio-demographics and environmental concern is very inconsistent. They wonder whether these inconsistencies might be due to the violation of two statistical assumptions that researchers in this area often make: the interval-level measurement and parallel regression assumption. After testing this empirically using data from the Earth Day Survey, they find that a violation of the parallel regression assumption can cause serious problems. Therefore, this study does not only apply parametric methods but also implement the non-parametric random forests method. This method does not make strong assumptions on the functional form and might increase prediction accuracy dramatically (James et al., 2013, p. 303).

As it seems, it is about time environmental consciousness is researched from a broad perspective, including various groups of determinants and taking a multinational approach. The current study adds to the extensive amount of existing literature, by looking at the effects of both sociodemograpics and psychological characteristics on environmental consciousness. Furthermore, due to the data used, it has the outstanding position of taking a multi-country approach. Finally, to the best of my knowledge, it is one of the rare studies that look into the predictive power of non-parametric methods in this area.

From a more practical point of view, this study might deliver value to businesses and policymakers. As sustainable consumerism becomes mainstream, it is essential for businesses to respond to this all-encompassing change of the business environment (Coppola et al., 2019). Evidence is piling up that environmental considerations play a large part in the consumer's decision-making process (Golob and Kronegger, 2019). These considerations play a role in assessing information sources or gathering information (Oates et al., 2008) and it becomes clear that consumers are likely to pay more for socially and environmentally friendly products (Didier and Lucie, 2008). Businesses missing out on this trend could be taken over by competitors.

Yet, there is still considerable heterogeneity between consumers regarding environmental consciousness (Golob and Kronegger, 2019). Even within the group of sustainability-conscious consumers, it is "well advised to recognize multiple types of sustainability-conscious consumers (. . .) and to implement targeting strategies that do not rest on the assumption of homogeneity." (Balderjahn et al., 2018). Companies should take into account that being environmental conscious is not binary. Some consumers do not care about the environment at all, others choose green products in case of equal prices, and part of the consumers will only consume sustainably. To address this targeting challenge, this study offers companies the tools to predict environmental consciousness by pointing out the most important determinants and the most effective methods.

This study aims to provide data analysts with the tools to make these predictions, so managers are able to make data-driven decisions. Also for this purpose, it is necessary to consider the multinational perspective. Already long ago it is recognised that standardised strategies across countries are a major opportunity (Buzzel, 1968; Yilmazsoy et al., 2015). If strategies can be standardised, this would lead to great cost reduction. However, in this context, it is all the more important since environmental consciousness varies across countries (Golob and Kronegger, 2019; de Maya et al., 2011; Yilmazsoy et al., 2015; García-Álvarez and Moreno, 2018). This makes it uncertain whether standardised strategies will be effective. Which countries show similar levels of environmental consciousness and which not? Despite the importance of the matter, most studies have been carried out among U.S. consumers and cross-country research is relatively sparse (Diamantopoulos et al., 2003; Vicente-Molina et al., 2013). Thus, the use of multi-national data in the current study could lead to essential insights for (multinational) companies.

Furthermore, it should be noted that also for policymakers, who often struggle satisfying citizens with opposing environmental views, this study could be of great value. Analogous to businesses, policymakers could make use of the insights gained by this study to get an idea of the environmental consciousness in their geographical area, which is essential in assessing the support for possible environmental policies. Additionally, policymakers can make use of the determinants of environmental consciousness by targeting more specific population subgroups. By these means, the insights gained in this study can help policymakers implement environmental policies more effectively.

Unfortunately, the ever-increasing consumerism plays a big role in climate change (Ivanova et al., 2016). To ensure the planet remains the comfortable place it currently is, it is essential to move the consumer to become a more sustainable version of itself. Hopefully, this study is a small step in that direction.

The remainder of this paper is structured as follows. In Chapter 2, the theoretical framework, the concept of environmental consciousness is be reviewed. Also, the literature is laid out as strong basis for hypothesising the effect of (sets of) predictors on environmental consciousness. The data and measurement model are elaborated on in Chapter 3. After that, in the methodological section, the predictive methods are discussed (Chapter 4). Subsequently, the model output is compared in Chapter 5. Conclusions are drawn and discussed in Chapter 6.

2 Theoretical framework

2.1 Conceptualising environmental consciousness

Environmental consciousness has seen a long history of conceptualisation. According to Schlegelmilch et al. (1996), the concept of environmental consciousness includes attitudinal and behavioural components. Also Stern (1999) identified the personal and the behavioural component of environmental consciousness. Dunlap et al. (2000) developed the New Ecological Paradigm Scale, which finds its ground in the social-psychological theories of attitudes structures. The scale reflects environmental attitudes, beliefs, and values. Kollmuss and Agyeman (2002) analyse the most influential frameworks for environmental consciousness, reviewing factors like environmental knowledge, locus of control, attitudes, values and external factors. They find that environmental attitudes, knowledge, values and emotional involvement make up a big part of the concept of environmental consciousness. Sánchez and Lafuente (2010) settled the debate: environmental consciousness is made up out of four dimensions: the affective, dispositional, cognitive and active dimensions. The affective dimension reflects concern for the environment and having a pro-environmental worldview. The dispositional dimension consists of self-efficacy, the perception of individual responsibility, and the willingness to accept costs. The cognitive dimension catches the knowledge of an individual. Finally, the active dimension can also be referred to as pro-environmental behaviour, consisting of environmental activism, low-cost behaviours, such as recycling, and high-cost behaviour (also known as green-consumerism), such as reducing aeroplane travelling. So, environmental consciousness is a multidimensional concept consisting of four dimensions.

The multidimensionality of the concept is highlighted by research from Diamantopoulos et al. (2003). They reviewed environmental literature over the past 25 years, and note that as a consequence of the research which has been done in a wide range of disciplines, many different concepts have been used. As a result, the effect Diamantopoulus and his co-authors are looking into, namely, the effect of sociodemographics on environmental consciousness is very ambiguous according to the literature. In the second part of their study, they aim to address this issue. To do so, the authors set up a survey with environmental consciousness conceptualised into five measures: environmental knowledge, environmental attitudes, recycling behaviour, political action behaviour, and purchasing behaviour. The first reflects the cognitive dimension, the second reflects the affective and dispositional dimensions, and the last three the active dimension. 9700 UK households filled out the survey and supplied sociodemographic information. The findings are telling: the effect of sociodemographics on environmental consciousness is varying over the dimensions. An important lesson is learnt: environmental consciousness is a broad concept and one should be careful when interpreting findings due to its multidimensional face.

In the next subsections, the literature on determinants of environmental consciousness is reviewed. First, the effect of sociodemographics, then the effect of psychological characteristics and finally national-level factors are looked into. The focus lies on the dimensions used in this study, namely the affective and dispositional dimensions, and the determinants that are available in the data set.

2.2 Sociodemographics

There is much evidence on the effect of sociodemographics on environmental consciousness, however, the findings are generally ambiguous (Diamantopoulos et al., 2003). Fortunately, Diamantopoulos et al. (2003) set up a very decent study and Gifford and Nilsson (2014) performed a meta-analysis on 18 personality and social factors influencing environmental consciousness. Diamantopoulos et al. (2003) confirm the hypothesis that younger people are generally more concerned about the environment. Also, Gifford and Nilsson (2014); Dietz et al. (1998); Straughan and Roberts (1999) find that almost all research indicate the same. On the other hand, a meta-analysis performed by Wiernik et al. (2013) concludes that most studies point in the direction of a positive relationship between being environmentally conscious and age. Secondly, it is found that females tend to be more environmentally concerned (Diamantopoulos et al., 2003; Gifford and Nilsson, 2014; Dietz et al., 1998; Straughan and Roberts, 1999). Furthermore, the number of children per household is only researched a few times. Diamantopoulos et al. (2003) report that three studies indicate that having more children is associated with higher concern in this context (Grunert, 1993; Brooker, 1976). However, their own empirical research showed an insignificant, negative relationship (Diamantopoulos et al., 2003). It has been proposed that environmental concerns could be concerns of high social classes (Buttel and Flinn, 1978, p. 436). Diamantopoulos et al. (2003) hypothesise based on research that the higher the social class, the more concerned the class members are. Gifford and Nilsson (2014) find that middle and middle-upper class members are generally more environmental concerned. However, the empirical findings of Diamantopoulos et al. (2003) show no significant relationship. The expectations are rather ambiguous, although the well set-up study of Diamantopoulos et al. (2003) indicate that social class is not a strong predictor. The effect of education on environmental consciousness is found to be homogeneous and strong: almost all studies reviewed by Diamantopoulos et al. (2003); Gifford and Nilsson (2014) indicate a strong, positive relationship (Straughan and Roberts, 1999). However, the empirical evidence of Diamantopoulos et al. (2003) shows no significant relationship.

All in all, some sociodemographics have been researched more extensively than others. Although most findings are ambiguous, one can generally say that younger, higher educated females from a higher social class tend to be more environmentally concerned. However, it is known that demographics do not explain environmental consciousness very well, only about 6 per cent of the variance (Diamantopoulos et al., 2003). Other predictors might be necessary to accurately predict environmental consciousness.

2.3 Psychological characteristics

As the prediction of environmental consciousness based on sociodemographics alone is not satisfactory, adding a set of more powerful predictors is necessary. Psychological characteristics have been known as important predictors of many attitudinal and behavioural outcomes (Parks-Leduc et al., 2015). It speaks for itself that many researchers are calling for the use of psychological characteristics to predict environmental consciousness (Straughan and Roberts, 1999; Diamantopoulos et al., 2003; Golob and Kronegger, 2019; Verain et al., 2012; Balderjahn et al., 2018; Wang, 2017; Dietz et al., 1998; Soutter et al., 2020; Yilmazsoy et al., 2015).

The concept of an individuals' personality is generally considered to include personality traits and personal values (Parks-Leduc et al., 2015). Personality traits are often defined as "descriptions of people in terms of relatively stable patterns of behavior, thoughts, and emotions" (Parks-Leduc et al., 2015; McCrae and Costa, 2003). Personal values are typically described as "rather stable broad life goals that are important to people in their lives and guide their perception, judgments, and behavior" (Parks-Leduc et al., 2015; Rokeach, 1973; Schwartz, 1992). Personality traits are thus rather descriptive, whereas values are motivational (Parks-Leduc et al., 2015). Furthermore, two levels of abstractions of values are used. The differentiation is made between consumer domain-specific values and general values (Vinson et al., 1977; Fred van Raaij and Verhallen, 1994; Homer and Kahle, 1988). Where general values are broad and promoted in various situations, consumer domain-specific values are about concrete goals specifically in the consumer domain. Therefore, domain-specific values add additional value because they narrow the broad concept of values down to the specific subject under investigation. These three groups - personality traits, general values and consumer domain-specific values - together would make up a sound theoretical framework for one's personality, of which an overview can be found in Table 2.1.

Table 2.1: Overview of the definitions, references and structure of the psychological characteristics used for modelling. The theoretical structure consists of personality traits, personal values, and consumer-specific values.

Psych. characteristic	Definition	Reference
Personality traits	"Descriptions (...) in terms of relatively stable patterns of behavior, thoughts, and emotions"	(Parks-Leduc et al., 2015)
Five factor model		(McCrae and John, 1992)
Openness to experience	"Curious, intellectual, imaginative, creative, innovative, and flexible (vs. closed-minded, shallow, and simple)"	(Parks-Leduc et al., 2015)
Agreeableness	"Helpful, good-natured, cooperative, sympathetic, trusting, and forgiving (vs. rude, selfish, hostile, uncooperative, and unkind)"	(Parks-Leduc et al., 2015)
Extraversion	"Sociable, talkative, optimistic, ambitious, assertive, reward-seeking, outgoing, and energetic (vs. introverted, shy, reserved, quiet, and unadventurous)"	(Parks-Leduc et al., 2015)
Conscientiousness	"Organized, responsible, dependable, neat, efficient, and achievement-oriented (vs. disorganized, lazy, irresponsible, careless, and sloppy)"	(Parks-Leduc et al., 2015)
Emotional Stability	"Calm, self-confident, stable, resilient, and well-adjusted (vs. neurotic, nervous, insecure, fearful, and anxious)"	(Parks-Leduc et al., 2015)
General values	"Rather stable broad life goals that are important to people in their lives and guide their perception, judgments, and behavior"	(Parks-Leduc et al., 2015)
Personal values	Personal values are broad motivational goals, and their relative importance vary.	(Schwartz, 1992)
Self-direction	"Independent thought and action–choosing, creating, exploring."	(Schwartz, 2012)
Stimulation	"Excitement, novelty, and challenge in life."	(Schwartz, 2012)
Hedonism	"Pleasure or sensuous gratification for oneself."	(Schwartz, 2012)
Achievement	"Personal success through demonstrating competence according to social standards."	(Schwartz, 2012)
Power	"Social status and prestige, control or dominance over people and resources."	(Schwartz, 2012)
Security	"Safety, harmony, and stability of society, of relationships, and of self."	(Schwartz, 2012)
Conformity	"Restraint of actions, inclinations, and impulses likely to upset or harm others and violate social expectations or norms."	(Schwartz, 2012)
Tradition	"Respect, commitment, and acceptance of the customs and ideas that one's culture or religion provides."	(Schwartz, 2012)
Benevolence	"Preserving and enhancing the welfare of those with whom one is in frequent personal contact (the 'in-group')."	(Schwartz, 2012)
Universalism	"Understanding, appreciation, tolerance, and protection for the welfare of all people and for nature."	(Schwartz, 2012)
Consumer values	"Values specific to economic transactions through economic exchange and consumption."	(Vinson et al., 1977)
Product		
Materialism	Having valuable possessions is important.	(Richins and Dawson, 1992)
Consumer innovativeness	"Predisposition to buy new products and brands, rather than to remain with previous choices and consumption patterns."	(Steenkamp et al., 1999)
Normative influence	Having a "moral obligation to perform or refrain from specific actions"	(Schwartz and Howard, 1984)
Quality consciousness	Preference for high-quality products.	(Sproles and Kendall, 1986)
Health consciousness	Preference for healthy products.	
Time		
Nostalgia	"A longing for the past, a yearning for yesterday, or a fondness for possessions and activities associated with days of yore"	(Holbrook, 1993)
Place		
Consumer ethnocentrism	"Beliefs held by (...) consumers about the appropriateness, indeed morality, of purchasing foreign made products."	(Shimp and Sharma, 1987)

2.3.1 Personality traits

Five-Factor Model - As mentioned, personality traits are a description of a person and are relatively stable over time. Research on which traits should be considered as 'personality' has been going on during the second part of the last century (Matthews et al., 2003). However, roughly three decades ago the Five-Factor Model (McCrae and John, 1992) was introduced and since then this is the most widely used model to capture personality traits (Matthews et al., 2003; Parks-Leduc et al., 2015; Allik, 2005). This model merges various traits into five domains (also known as the Big Five): openness, conscientiousness, extraversion, agreeableness and neuroticism (or opposite emotional stability). Openness describes how eager a person is to have new experiences. A conscientious person is well-organised and achievement-orientated. Extraversion includes outgoingness and sociable traits. An agreeable person is generally helpful and cooperative. Finally, neuroticism describes whether a person is calm or anxious.

How do these traits relate to environmental consciousness? According to Gifford and Nilsson (2014), who performed a meta-analysis, openness and agreeableness are positively related to environmental consciousness. Conscientiousness and neuroticism are to a lesser extent related to environmental consciousness, whereas the latter relationship is negative. Soutter et al. (2020) also performed a meta-analysis, totalling up to 44993 participants from 38 sources, representing at least 19 countries across 4 continents. The findings indicate a positive relationship between all domains of the Big Five and environmental attitudes (which consists of various environmental scales). Openness has the strongest association, followed by agreeableness, conscientiousness, and extraversion. Neuroticism is least associated with environmental attitudes. Overall, Soutter et al. (2020) conclude that the Big Five explain a rather high part of the variance in environmental attitudes.

2.3.2 General values

Personal values - Personal values are broad motivational goals, and their relative importance varies. They are relatively stable over time and are an antecedent of attitudes and behaviour (Parks-Leduc et al., 2015; Schwartz, 2012). The most used concept is that of Schwartz (1992, 2006). He introduced ten personal values and structured these in Figure 2.1. Values that have less distance to each other in the circle, are generally stronger associated with each other. On the other hand, values that have an opposite position, are negatively associated. The circle is surrounded by a higher-order structure, which represents two bipolar dimensions: openness versus conservation and self-transcendence versus self-enhancement. The ten values all have a defining goal: self-direction, with as goal independent thought and action; stimulation, with goals of excitement, novelty and challenge; hedonism, with pleasure and sensuous gratification being important; achievement, where personal success is important; power, with the defining goal being social status and prestige; security, concerning safety and harmony; conformity, which aims not to upset others and violate social norms and expectations; tradition, where following traditions and accepting customs and ideas is important; benevolence, to enhance the welfare of the in-group, and universalism, for which acceptance of all people and nature is central. Especially important in the international context this research is done, the personal values are proven to be universal over different cultures (Schwartz, 1994, 2011).

As values show what is important to a person, they can lead to having specific attitudes or behaviour (Schwartz, 2012), like environmental consciousness. Especially the relationship with the dimension of self-enhancement versus self-transcendence (enhancing your own welfare or enhancing the welfare of others) seems to be important. As one would expect, environmental consciousness has been associated with values that promote self-transcendence and contrary with values that promote self-enhancement (Gifford and Nilsson, 2014; Milfont and Gouveia, 2006; Nilsson et al., 2004; Nordlund and Garvill, 2002; Schultz and Zelezny, 1999; Stern, 2000; Stern et al., 1995; Stern and Dietz, 1994; Ahmad et al., 2020). Also, in China, the conser-

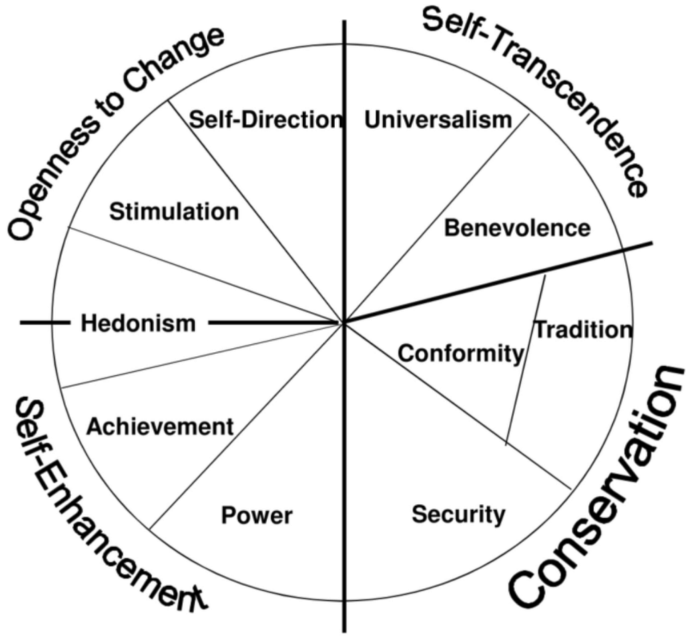


Figure 2.1: Ten personal values in relationship with each other. Values that are nearer to each other, exhibit stronger associations. Opposing values are negatively associated. Outside of the circle are the higher-order dimensions.

Adapted source: Schwartz (1992, 2006)

vation dimension is positively associated with environmental attitudes (Ahmad et al., 2020). When zooming in on specific values, a study across 174 Danish teachers show that the values self-direction, universalism and benevolence are positively related to pro-environmental purchasing behaviour, stimulation, hedonism, and achievement are unrelated to pro-environmental purchasing behaviour, and power, security, conformity and tradition are negatively related to pro-environmental purchasing behaviour (Grunert and Juhl, 1995). Although the proof is not unambiguously with regards to the conservation dimension, the values that fall within the self-transcendence dimension are found to be positively related to environmental consciousness.

2.3.3 Consumer domain-specific values

Consumer domain-specific values are unfortunately not exhaustive and as structured as Schwartz's general values. This is not due to a lack of effort put in, but to the fact that there are generally hundreds of domain-specific values (Vinson et al., 1977). Selecting and structuring these values will thus be somewhat more arbitrary. After a quick sweep through a sample of the values, a handful is somehow expected to be connected to environmental consciousness. To structure these values, the classification of Steenkamp and de Jong (2010) is used: the products-time-space framework. The following consumer domain-specific values are used. The products category consists of materialism, consumer innovativeness, normative influence, health consciousness, and quality consciousness. In the time dimension, there is nostalgia. In the space dimension, the value of consumer ethnocentrism is included. Their relationship to environmental consciousness is elaborated on in the next paragraphs.

Materialism - Materialistic people find having plenty of valuable possessions important (Richins and Dawson, 1992). On the other hand, there are postmaterialistic values. According to Inlgehart's postmaterialism thesis, postmaterialist give priority to broader issues, whereas materialists care about physical sustenance (Inglehart, 1977). Therefore, one would expect postmaterialists would be more environmentally conscious than materialists. Unfortunately, only two studies examine the difference between having materialistic and post-materialist values on being environmental consciousness. In Turkey, postmaterialists appear to be more concerned about global environmental issues, materialists about local environmental issues (Gökşen et al., 2002). However, in Israel materialistic values do not appear to be a strong determinant of environmental consciousness (Drori and Yuchtman-Yaar, 2002). More research is needed, but it is expected that having materialistic values, has a weak, negative correlation with environmental consciousness.

Consumer innovativeness - Innovative consumers are the so-called front-runners. They are keen to buy the newest and most innovative products. This value is known as consumer innovativeness (see Tellis et al., 2009; Im et al., 2003; Steenkamp et al., 1999). In theory, innovative consumers would be expected to adopt environmental attitudes and behaviour earlier than their not so innovative counterparts. This is due to the fact that environmentally conscious products are often innovations, such as the electric car, solar panels and plastic-free packaging. According to Thøgersen et al. (2010); Thogersen (2002); Tews et al. (2003), one should consider eco-labelling as innovation, for example. Based on the previous reasoning, one could thus say that innovative consumers are more likely to buy eco-labelled products. This effect is found in research on Danish shopping mall visitors (Thøgersen et al., 2010). Also, Englis and Phillips (2013), who performed a survey across 1443 US consumers, found that some dimensions of consumer innovativeness, strengthen the link between environmental attitudes and behaviour. Although the evidence is altogether not indisputable, it is expected that consumer innovativeness has a positive effect on environmental consciousness.

Normative influence - Normative influence is defined as having a "moral obligation to perform or refrain from specific actions" (Schwartz and Howard, 1984). The norm activation theory of Schwartz (1977) describes how normative influence leads to altruism. Over the past decades, evidence has been found that indicate that the same holds for environmental behaviour (Gifford and Nilsson, 2014; De Groot and Steg, 2009). This concerns for example the case of recycling (Matthies et al., 2012), protesting against nuclear energy (De Groot and Steg, 2010), and reduction of car use (Abrahamse et al., 2009). Since environmental behaviour is related to environmental consciousness (Schlegelmilch et al., 1996; Roberts and Bacon, 1997; Minton and Rose, 1997; Bamberg and Möser, 2007; Klöckner, 2013) (with a correlation of around 0.40 (Hines et al., 1987), it is assumed there is an association between normative influence and environmental consciousness.

Quality and health consciousness - Quality conscious consumers are consumers who search for products of good quality (Sproles and Kendall, 1986). Health-conscious consumers care about healthy products, which is especially the case for food (e.g. Magnusson et al., 2001). Unfortunately, not much is known about the relationship between quality and health-conscious consuming and environmental-conscious consuming. Yet, it is imaginable that there is a relationship. Buying organic food, for example, is motivated by health as well as environmental concerns (Kriwy and Mecking, 2012). The relationship needs more research, but it is expected that being quality and health consciousness, is positively related to being environmental consciousness.

Nostalgia - Holbrook (1993) almost poetically describes nostalgia as "a longing for the past, a yearning for yesterday, or a fondness for possessions and activities associated with days of yore". Back then the scientific literature still had that magic touch. In this description the dissonance with environmental conscious consuming behaviour is tangible. New developments and innovations that the environmental nonsense brings along, is probably rejected by nostalgic people. Research is absent, yet the expectation is there is a negative relationship between nostalgic values and environmental consciousness.

Consumer ethnocentrism - Consumer ethnocentrism is a value stemming from the sociological concept of ethnocentrism, which makes a distinction between the in and out-group. Consumer ethnocentric people prefer buying national products instead of foreign products, based on several reasons. Ethnocentric consumers, for example, find that buying foreign products hurts the national economy (Shimp and Sharma, 1987). One can imagine that there is a negative link between global problems like climate change and the local focus of ethnocentrism. However, this effect might be captured by personal values like benevolence and universalism. Yet, some evidence is found that ethnocentrism has a significant influence on environmental conscious consumption (Paladino, 2005).

2.4 National-level variables

From the previous sections, it becomes clear that sociodemographics and especially psychological characteristics determine environmental consciousness to a great extent. However, that is not the entire story. People are typically social beings, living in social groups and being affected by their surroundings. Not surprisingly, research suggests that variations over time in the Big Five personality traits and motivational constructs are determined by genetics for only 30 per cent. The rest can be attributed to contextual effects (Bleidorn et al., 2010). In the environmental literature, it is sometimes stressed that the focus on individual factors, can cause contextual factors to be overlooked (Peattie, 2010; Spaargaren, 2011). To account for contextual effects, this study takes the national-level variables into account.

So, which national-level variables should one consider? I would propose the following dimensions. First of all, environmental governance is important in mitigating environmental problems and influencing individual environmentalism (Wang, 2017). Secondly, the risk of being affected by climate change might have an effect on an individual's environmental perception. Thirdly, also the share of activities leading to climate change is expected to influence individual attitudes. Finally, there are some country-level confounding factors.

Environmental governance - Environmental governance is "the set of regulatory processes, mechanisms and organizations through which political actors influence environmental actions and outcomes," (Lemos and Agrawal, 2006). One can think of governmental actors, but also non-governmental parties, like businesses and non-governmental organisations (NGOs). It is an indirect effect that goes via the spread of culture, awareness and concern (Boli and Thomas, 1997; Frank et al., 2000; Wapner, 1996). A study over 37 countries found for example that actions from NGOs and environmental governmental groups increase individual environmentalism (Givens and Jorgenson, 2013). Wang (2017) looked into the effect of environmental governance on individual sustainable consumption, analysing almost 40 thousand respondents over 31 countries, and found that environmental governance had a positive effect in high-income countries and a negative effect in other countries.

Natural hazard risk - From 1970 until 2010, the estimated damage caused by natural disasters have been increasing a great deal (Mucke et al., 2011). The new report of the Intergovernmental Panel on Climate Change (IPCC) makes no bones about it: human activities have led to climate change, which in turn leads to an increase in extreme conditions, and eventually results in an increase in natural disasters (Masson-Delmotte et al., 2021). Via subjective risk perception, these objective events can lead to more environmental consciousness. The link between risk perception of natural hazards and environmental attitudes has been researched in Taiwan (Lee et al., 2019). The results support the existence of the link.

Impact on climate change - Since human activity is the main driver of climate change (Masson-Delmotte et al., 2021), the share a country is contributing to climate change, could be an important driver of individual attitudes towards the environment. The most important cause of climate change is greenhouse gas emission (Masson-Delmotte et al., 2021). On one hand, people living in a country that is emitting a lot of greenhouse gas per capita can have feelings of guilt. On the other hand, emitting many greenhouse gases can also mean that people living in that country are not concerned with the environment. Either way, the effect is expected to be evident.

Confounding factors - Finally, some factors are confounding. Research suggests GDP per

Sociodemographics

- Age (-)
- Female (+)
- Number of children per household (0)
- Social class (0)
- Educational level (0)

Psychological characteristics

Personality trais

Five Factor model
- Openness (+)
- Agreeableness (+)
- Conscientiousness (+)
- Extraversion (+)
- Neuroticism (+)

General values

Personal values
- Power (-)
- Achievement (0)
- Hedonism (0)
- Stimulation (0)
- Self-direction (+)
- Universalism (+)
- Benevolence (+)
- Tradition (-)
- Conformity (-)
- Security (-)

Consumer domain-specific values

Product
- Materialism (-)
- Consumer innovativeness (+)
- Normative influence (+)
- Health consciousness (+)
- Quality consciousness (+)

Time
- Nostalgia (-)

Space
- Consumer ethnocentrism (+)

National-level variables

National environmental situation

- Environmental governance (+)
- Natural hazard risk (+)
- Impact on climate change (+)

National confounding factors

- GDP per capita (+/-)
- Population density (-)

Figure 2.2: Overview of expected effects of predictors on environmental consciousness. '+' is positive, '-' is negative, '0' is not statistically significant, and '+/-' is ambiguous.

capita has an influence on environmental attitudes (Wang, 2017). It is believed that a higher GDP per capita increases awareness and it allows people to invest money in environmental positive behaviour (Ewers, 2006). On the other hand, some researchers believe that it has a negative influence on environmentalism (Givens and Jorgenson, 2011; Parikh and Shukla, 1995; York et al., 2003). Either way, the effect should not be ignored. Also, there is proof that a higher population density leads to a higher rate of environmental degradation (York et al., 2003; Vasi, 2007).

All in all, it seems that environmental governance, natural hazard risk and impact on climate change all have an impact on individual environmental attitudes. Figure 2.2 summarises the expected effects of sociodemographics, psychological characteristics and national-level variables on individual environmental consciousness. In the next chapters, these relationships are going to be tested empirically.

3 Data

3.1 Individual-level data

For the sociodemographics and psychological characteristics, primary data are used. These data comes originally from a study by Steenkamp and de Jong (2010). The data are collected in 2004 by the marketing agencies GfK and TNS. They performed an online survey among 13,321 respondents in 28 countries. However, in some countries internet was rarely used back then, so some questionnaires were filled out in malls on a laptop or paper. The following countries are included: Argentina, Austria, Belgium, Brazil, China, Czech, Denmark, France, Germany, Hungary, Ireland, Italy, Japan, the Netherlands, Norway, Poland, Portugal, Romania, Russia, Slovakia, Spain, Sweden, Switzerland, Taiwan, Thailand, the United Kingdom, the United States of America and Ukraine. Unfortunately, the national-level data did not contain Taiwan, so these rows are removed. Also, 149 rows contained missing data and are omitted, resulting in 12,820 rows.

3.1.1 Measurement of the sociodemographics

A total of twelve sociodemographics is included. This entails the ones discussed in the theoretical section, namely age, gender, number of children per household, the number of children under 18 living in the household, social class (6 classes), and educational level (7 levels). Also, additional sociodemographics are included, despite the fact, there is no literature on the effect of these variables on environmental consciousness. These demographics are the size of the household (concerning all people living in the household for at least 4 days a week), the number of inhabitants living in the village (6 classes), the highest level of education of the partner of the respondent (8 classes), the yearly income of all wage earners in the household (8 categories), and a categorical variable that reflects the income change over the past 3 years. This last variable has five levels, starting from "gone down a lot" to "gone up a lot" (Steenkamp and de Jong, 2010). The specific class levels can be found in the second column of Table 7.2 in the appendix.

3.1.2 Measurement of the psychological constructs

Measurement model

The psychological constructs involved, consist in most cases of multiple items measured on an ordinal scale. The constructs themselves are thus latent. This requires choosing how the items should be represented in the latent constructs. The relationships between the latent construct and the items are represented by a measurement model. A simple measurement model would be to sum the item scores and divide the total by the number of items. Another option is to use explanatory factor analysis (EFA) or confirmatory factor analysis (CFA) (Jöreskog, 1971, 1969). These methods are specifically meant to combine several observed items into a smaller number of common factors (Thurstone, 1947). The main difference is, however, that CFA makes assumptions on the number of common factors and covariance of the factors, whereas EFA uses exploratory techniques to determine this. CFA thus requires a sound theoretical basis and aims to confirm or deny the theory, EFA can be used as a data-driven starting point to develop a theory (Brown and Moore, 2012).

In this case, CFA is used to estimate the measurement model. The constructs in the data have, as can be read in Chapter 2.3, theoretical grounds and are often empirically tested. Moreover, CFA has the advantage over EFA in that it can model measurement errors and can deal with the interrelatedness of the constructs (Brown and Moore, 2012). Thus, based on the theory, the structure of the relationships of the constructs and their items can be used to estimate the relationships with CFA.

CFA works as follows (see Brown and Moore, 2012). The aim is to end up with a set of parameter estimates that result in a prediction of a variance(-covariance) matrix that best approaches the matrix inputted. The factor loadings (λ) and factor variances (ϕ) are estimated and they model the relationship between the items and the latent factors. The first is the slope and the latter is the sample variability. Next to these parameters, the unique variance reflects the measurement error. If there is more than one factor, the factor covariances (ψ) should be set (possibly to 0), similar if items load onto multiple factors. The CFA solution is the estimation of these parameters.

However, not all parameters are freely estimated. In fact, there are three kinds of parameters: fixed, free, and constrained parameters. Whereas fixed parameters are to be set, often either to 1 or 0, free parameters are freely estimated simultaneously with other parameters, so that the difference between the predicted observed matrices of variances and covariances are minimised. A constrained parameter is in essence a free parameter, yet often constrained so that it equals another parameter. The output is generally standardised.

An important concept in CFA is model identification. A model needs to have enough observed information (which is the variances and covariances of the input matrix) to find a solution for the parameters. The first step in identifying the model is to scale the latent factor. This is necessary since the latent factors have no known values. There are two methods. The first method is called the marker method and it sets one of the factor loadings to 1. The variance standardisation method sets one of the factor variances to 1. Both methods generally give similar outcomes. The next step is to statistically identify the model. Now that one of the parameters is fixed, a number of parameters remain to be freely estimated. The *degrees of freedom* are the number of known elements from the input matrix minus the number of free parameters. If the degrees of freedom are positive, the model is *over-identified*. If they are negative, the model is *under-identified*. If there are no degrees of freedom, the model is *just-identified*. An under-identified model cannot be estimated. For evaluative statistics, in order to check the validity of the structure, an over-identified model is required. Next to the statistical identification, a model can be empirically under-identified due to various reasons (see Wothke, 1993; Brown, 2015), but this theory is somewhat out of the scope of this research.

Since multiple parameters need to be estimated simultaneously, a *fitting function* is needed. Although there are multiple methods, the most used, also for this purpose, is *maximum likelihood* (ML). As the name states, the function tries to maximise the likelihood that the parameters fit the observed data. It starts at a certain value for the parameters and it iterates while trying to increase the likelihood until the outcome converges. Maximum likelihood assumes a large sample size, that the indicator distribution is normally multivariate, and that the indicators are at least measured at an approximation of an interval scale. Minor violations lead to biased standard errors. In that case, ML with robust standard errors should be estimated (Bentler, 1995). If the normality assumption is grossly violated, an alternative fitting function should be used. This is for example the case if most respondents gave the lowest answer on the indicator scale. In most cases, maximum likelihood, possibly with robust standard errors, will work properly (Brown and Moore, 2012).

Finally, for some purposes, the model fit is relevant. In this particular case, most constructs are already evaluated and used often, and the purpose is not to confirm any hypothesis but to use CFA as a measurement model. Despite this, model evaluation is used to validate the results, if possible. Note, however, that in order to produce evaluation measures, enough information is necessary. In other words, a model should be over-identified. There are a few fit measures that are generally reported together, as they all provide a different aspect of model fit. These are the standardised root mean square residuals (SRMR), the root mean square error of approximation (RSMEA), the Tucker-Lewis index (TLI), and the comparative fit index (CFI) (Brown and Moore, 2012). Following Hu and Bentler (1999), a reasonable fit would be SRMR lower or equal to 0.08, RSMEA lower or equal to 0.06, and the latter two indices higher or equal to 0.95.

Mathematically, CFA works as follows. The variance-covariance matrix is described by $\Sigma(\Theta)$,

Table 3.1: Output and statistics of the environmental consciousness construct estimated using CFA

Construct or item	**Mean**	**SD**	**Loading**	**Variance**
Environmental consciousness construct	0	0.551	NA	0.371**
Q1 I would be willing to stop buying products from companies guilty of polluting the environment, even though it might be inconvenient for me.	3.551	1.049	1	0.729**
Q2 I become incensed when I think about the harm being done to plant and animal life by pollution.	3.554	1.009	1.341**	0.341**
Q3 When I think of the ways industries are polluting the environment, I get frustrated and angry.	3.484	1.015	1.376**	0.328**

Note. * $p < 0.05$, ** $p < 0.01$.

the model-implied covariance matrix:

$$\Sigma(\theta) = \mathbf{\Lambda}\mathbf{\Psi}\mathbf{\Lambda}' + \Theta_\epsilon$$

The first parameter, Λ, is a matrix of factor loadings, the second parameter, Ψ, is a matrix with the variances and covariances of the factors, and Θ_ϵ is a matrix of the variances and covariances of the residuals. When taking environmental consciousness for example, which is a one-factor construct with three items, using the marker method, $\Sigma(\Theta)$ would be:

$$\Sigma(\theta) = \psi_{11}\begin{pmatrix}1\\ \lambda_2\\ \lambda_3\end{pmatrix}\begin{pmatrix}1 & \lambda_2 & \lambda_3\end{pmatrix} + \begin{pmatrix}\theta_{11} & 0 & 0\\ 0 & \theta_{22} & 0\\ 0 & 0 & \theta_{33}\end{pmatrix}$$

with λ_2 being how the second items loads on the construct, ψ_{11} is the first factor variance[1], and θ_{11} is the variance of the first residual. The zeros in the variance-covariance matrix mean that the residual covariances are fixed to 0. As explained, the marker method fixes the first factor leading (λ_1) to 1. If the variance standardisation method is used, ψ_{11} would be fixed to 1. From this example, it also becomes clear that there are 6 free parameters to be estimated, so there need to be at least 3 items to make sure there is enough information to come to a solution (Brown and Moore, 2012).

Estimating the constructs

For the implementation of CFA, `R` statistical software and the `Lavaan` package (Rosseel et al., 2021) are used. After inspecting the data, it appears that the distributions of all indicators follow roughly a normal multivariate distribution, so maximum likelihood with robust standard errors is used as a fitting function. To set the scale, the marker method is being used. As noted, both the marker method and the variance standardisation method often give the same results, and after inspecting, it also appears to be the case in this situation. The `predict` function of the `Lavaan` package is used to eventually estimate the construct values. In the following paragraphs, the outcomes of the individual constructs are discussed.

Environmental consciousness - The measurement of the constructs is predetermined. Table 3.1 shows the output of the CFA and a description of the items. The questions are measured on a five-point Likert scale and come from a study by Grunert and Juhl (1995). The first question reflects the dispositional dimension since it concerns the willingness to accept the cost associated with environmental consciousness. The second and third questions relate to the affective dimension since it concerns the feeling that the environment is threatened (Sánchez and Lafuente, 2010). The items are roughly normally distributed. The output is standardised, so the construct mean is 0. Furthermore, all loadings and variances are significant. Since the marker

[1]When considering multiple factors, the variance of the second factor would be ψ_{22} and the covariance of the first and second factor ψ_{21}

method is used, the factor loading of the first item is set to 1. Table 3.2 shows additional descriptive statistics and fit measures for all constructs. Since there are three items, this model is just identified, with zero degrees of freedom. Therefore, evaluative measures cannot be calculated.

Personality traits - Personality is represented with the Big Five (McCrae and John, 1992). Each of the five constructs is measured with six items. The items of the Five-Factor Model and other constructs can be found in Table 7.1 in the appendix. The questions are randomly ordered and answers should be given on a five-point Likert scale. The difficulty in performing CFA, in this case, is that we now have five factors and there might be inter-relatedness. The five factors together reflect the personality but might be slightly overlapping or some characteristics might be often seen together. The developers of the Five-Factor model speak of inter-relatedness of the factors (McCrae and John, 1992). Cooper et al. (2010) found that the cross-loading of items onto factors is negligible, but that there is a small to moderated inter-relatedness of the factors. Therefore, the factor covariances are estimated to account for the inter-relatedness. The model is over-identified with 395 degrees of freedom. This allows for a model evaluation. The values in Table 3.2 does not show a perfect fit, but the model is not drastically under-performing.

General values - The general values are surveyed using Schwartz (1992) items. The survey contains 45 values, which are supposed to cover all domains of values. Schwartz (1992) his guidelines are closely followed: the descriptive name and a short description are listed. The respondents could rate these values from -1 to 7, for which -1 means that the value opposes the respondents value, 0 means that the value is not important, 3 means that the value is important, and 7 means that the value is of "supreme importance" (Steenkamp and de Jong, 2010). The number of items per factor range from 2 to 8, with 45 items in total, which can be found in Table 7.1 in the appendix. As known from the theory of Schwartz (1992), values are inter-related. Factor covariances are thus estimated. The total parameters estimated is 134, which leaves 856 degrees of freedom. Also in this case the model fit is not perfect, but reasonable for this purpose, see Table 3.2.

Consumer values - All items were asked randomly on a five-point Likert scale. The items for the consumer values are taken from the following authors: materialism from Richins and Dawson (1992), consumer innovativeness from Steenkamp and Gielens (2003), health and quality consciousness (Steenkamp and de Jong, 2010), consumer ethnocentrism from Shimp and Sharma (1987), and nostalgia from Holbrook (1993), although for most constructs a short version has been used (as in Alden et al., 2006; Steenkamp et al., 1999).

Not all constructs have degrees of freedom to calculate evaluative measures. In fact, quality consciousness and nostalgia were measured by only 2 items and were thus under-identified. Therefore, the factor loadings of the items were set to be equal (so they are constrained parameters), which releases one degree of freedom. Concerning the other constructs, it can be found that normative influence and consumer ethnocentrism have a very good fit, whereas consumer innovativeness has a rather poor fit. Finally, it should be noted that a manual calculation of the constructs, by adding the item scores up and dividing it by the number of items, might also be sufficient. The last column of Table 3.2 shows the correlation between the constructs calculated manually and the CFA prediction, and in general, these correlations are rather high.

3.2 National-level data

Environmental governance - Wang (2017) already researched the effect of environmental governance on individual-level environmental attitudes, although in a different research context. His set-up is gratefully used in this study as well. Wang (2017) stresses that environmental governance is a multi-dimensional concept and that it thus is hard to capture in one variable. Fortunately, outstanding work has been done in order to come up with data that reflects environmental governance: the Environmental Sustainability Index (ESI) (WEF et al., 2002). This index consists of 20 indicators, each build up out of multiple variables, that represent 5 broad components: environmental systems, reducing environmental stresses, reducing human vulner-

Table 3.2: Descriptives and fit measures of the constructs after performing CFA.

Construct	**SD**	**Min**	**Max**	**SRMR**	**RSMEA**	**TLI**	**CFI**	**Corr**
Environmental consciousness	0.551	-1.582	0.927	NA	NA	NA	NA	0.978
Five Factor model				0.072	0.070	0.675	0.705	
Openness	0.817	-3.469	2.140					0.901
Conscientiousness	0.855	-3.894	2.263					0.960
Extraversion	0.864	-3.455	2.412					0.963
Agreeableness	0.847	-4.318	2.298					0.917
Neuroticism	0.880	-3.160	2.420					0.979
Personal values				0.059	0.063	0.785	0.805	
Power	0.882	-2.943	2.613					0.951
Achievement	0.931	-4.689	2.248					0.922
Hedonism	0.864	-4.108	2.127					0.877
Stimulation	0.904	-3.618	2.120					0.945
Self-direction	0.968	-5.295	2.390					0.824
Universalism	0.939	-5.034	1.924					0.973
Benevolence	0.944	-5.112	1.974					0.961
Tradition	0.925	-4.086	2.304					0.881
Conformity	0.964	-5.007	2.113					0.887
Security	0.929	-5.231	2.196					0.893
Materialism	0.844	-1.817	2.953	0.041	0.080	0.887	0.932	0.991
Consumer innovativeness	0.896	-2.813	1.689	0.106	0.134	0.566	0.690	0.736
Normative influence	0.939	-1.461	3.114	0.027	0.067	0.959	0.971	0.992
Health consciousness	0.914	-2.508	1.824	NA	NA	NA	NA	0.981
Quality consciousness	0.814	-3.078	1.245	NA	NA	NA	NA	0.997
Consumer ethnocentrism	0.956	-1.557	2.245	0.009	0.063	0.997	0.990	0.992
Nostalgia	0.722	-1.623	1.544	NA	NA	NA	NA	1.000

Note. SD is standard deviation, Min is minimum value, Max is maximum value, SRMR are the standardised root mean squared residuals, RSMEA is the root mean square error of approximation, TLI is the Tucker-Lewis index, and CFI is the comperative fit index. Corr is the correlation between the CFA construct and the constructs calculated manually by taking the average of the items.

ability, social and institutional capacity, and global stewardship (Esty et al., 2002). Due to the sound methodology used, it is possible to aggregate these components into one score: the ESI score. Important to this case, is that the ESI score is standardised in order to make cross-country comparisons. The data come from 2002, which is no problem, due to a potential lagged impact. Unfortunately, data on Taiwan is missing, so this country is dropped from the analysis.

What should be noted is that the impact of environmental governance can differ for high-income versus low-income countries, due to different crowding-out effects. Indeed, Wang (2017) finds that, although the effect of environmental governance on individual-level attitudes is always positive, the slope is less steep among non-high-income countries. A dummy variable indicating which countries are high-income is retrieved from the World Bank (2021).

Natural hazard risk - In order to capture the natural hazard risk, the World Risk Index is used (Mucke et al., 2011). This index is a calculation of the risk as a function of exposure on one hand and vulnerability on the other. Unfortunately, the oldest data available come from 2011. Without any doubt, situations have changed over these six years, but it is assumed that between-country differences have roughly stayed the same, so the data will still contain important information.

Impact on climate change - Since it is clear that human activity is the main driver of climate change, the country-level impact on climate change is added to the analysis. According to the most recent IPCC report, the very most important factor that is driving climate change is emitting greenhouse gases (Masson-Delmotte et al., 2021). Data from the World Bank on greenhouse gas emission (in kiloton of C02 equivalents) is used and divided by the population.

National-level confounding factors - GDP per capita (set to the value of a dollar in 2021) and population density (number of people per square kilometre of land) are known to be confounding factors (Wang, 2017). These variables are also taken from the World Bank data set.

3.3 Data exploration

This section aims to explore the data and the underlying relationships independent of model choice. First, consider the demographics. Table 7.2 in the appendix shows a comparison of means of different groups of demographics (in the form of an ANOVA or t-test). It is shown that for all demographics, there are significant differences between groups. Especially income and income change have a high F-value. The first exploration point in the direction that females of lower social classes, living in bigger cities with lower income, being housewife, retired, unemployed or working less than 8 hours, lower educated, having a lower educated partner, and having experienced negative income changes, are generally more environmentally conscious.

Figure 3.1 is a heat map showing the correlations between all the individual-level variables. The categorical variables are added as continuous variables since these are in most cases ordered. A yellow square at the intersection between a row and a column, indicate a high correlation between these variables. Blue colours indicate negative correlations. The less colour, the lower the correlation.

The bottom row shows the correlations between environmental consciousness and the other variables. Age has a small, but positive association with environmental consciousness, children per household and household size show both a very small, negative association. Unsurprising correlations are found between demographic factors. More children per household and household size are strongly and positively related. Social class is positively associated with education and income. The latter two are also positively correlated.

The figure also shows the correlations between psychological characteristics and environmental consciousness. It appears that agreeableness, a few of the personal values, in particular universalism, and the consumer values health consciousness, quality consciousness and consumer ethnocentrism are positively associated with environmental consciousness. Neuroticism, power, materialism and normative influence are negatively associated with environmental consciousness. The high correlations between the personal values might lead to collinearity problems. The plot

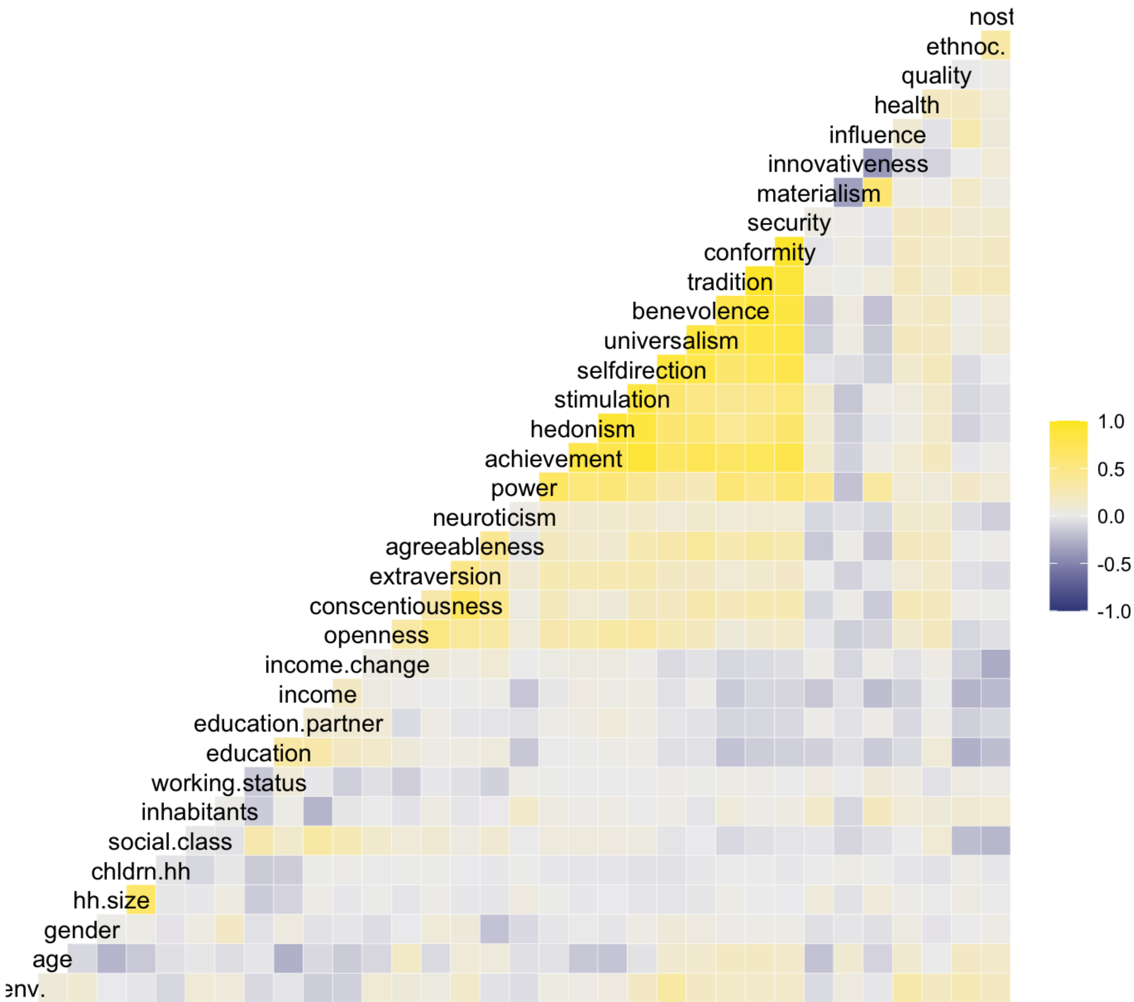


Figure 3.1: Heat map of correlations between all individual-level variables

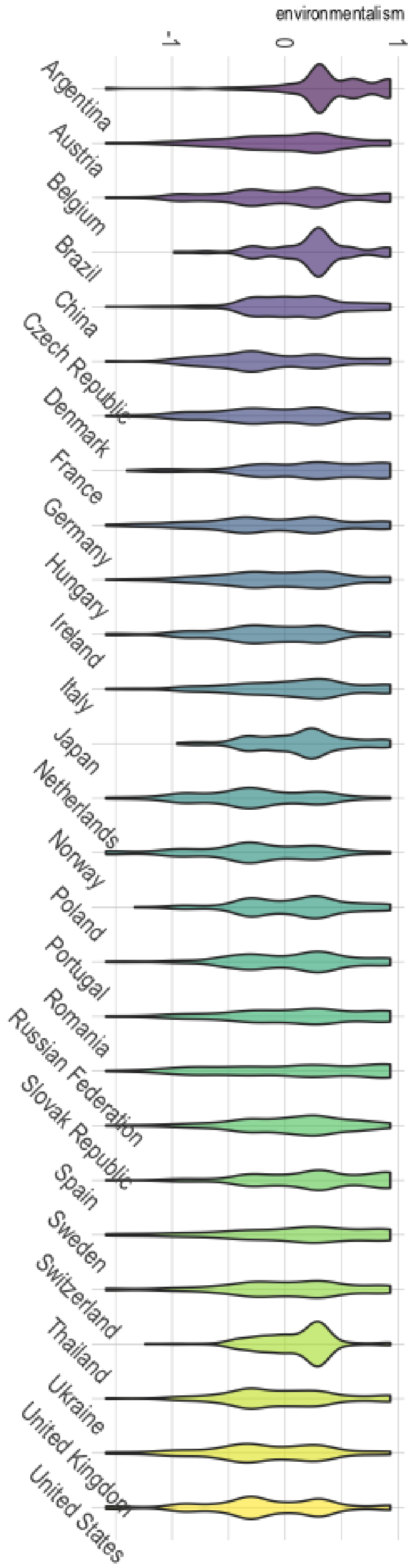


Figure 3.2: Distribution of environmental consciousness per country, standardised as a whole.

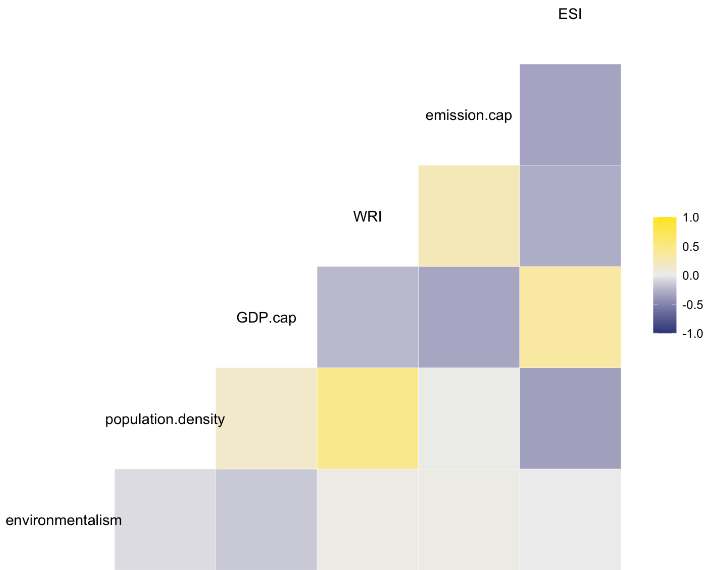


Figure 3.3: Correlations between the national-level variables and the average environmental consciousness per country.

also offers insight into cross-correlations. Personality traits are positively related to each other, the same holds for the personal values. This contradicts the Value Theory of Schwartz (1992), which poses that values opposite in the value structure of Figure 2.1, should be negatively associated. The consumer values show understandable relationships. Materialism and normative influence show a positive correlation. This makes sense, since both a triggered by a desire to fit in socially by consuming. On the contrary, both are negatively related to consumer innovativeness, which is driven by intrinsic motivation. Ethnocentrism and nostalgia are also positively related, which can be explained by having conservative feelings. Similarly, they are also positively correlated to conservative personal values, like tradition and conformity.

Figure 3.2 shows violin plots per country. The size of the distributions on the vertical axis shows the extent of environmental consciousness per country. Note that the values are standardised as a whole, meaning that the mean of all values is set to 0, and the standard deviation of all values together to 1. Some of the results are rather surprising. Western, high-income countries like Norway, the Netherlands, the UK and US are generally not very environmentally conscious, whereas less western countries like Argentina, Brazil, China and Russia perform quite well. An ANOVA test shows that at least one of the means is significantly different from the other at a 1% significance level. The plot also offers insights into the distribution of environmentalism per country. Although the distributions differ somewhat per country, some of them are more or less bimodal, others are more normally distributed, one can say there are not very big differences in distribution. Looking at country differences, one can also inspect the intra-class correlation (ICC). This metric shows which proportion of the total variance in environmentalism is attributed by the between-country variance (Hausknecht et al., 2008; Goldstein, 2010). Only 4.45% of the total variance can be allocated to country differences. All the above indicate that there are some group effects per country, but the overall effect is rather small.

Finally, Table 7.3 in the appendix offers insights into the values of the national-level variables. The variables GDP per capita, population density, emission per capita, and the world risk index are log-transformed in order to adjust for non-normality. Additionally, figure 3.3 gives an overview of the correlations between the average level of environmental consciousness per country, and some (untransformed) national-level variables. Environmental consciousness is most strongly, negatively related to GDP per capita. It is also negatively related to the environmental sustainability score, which is a proxy for environmental governance. At a first glance, it seems that a higher GDP per capita and more environmental governance, crowd out individual environmental attitudes. On the other hand, environmental consciousness is positively associated with the emission per capita and climate risk. All in all, the correlations are rather weak. This suggests that the predictive power of these variables are low.

4 Methodology

This section explains the methods used to answer the research question. Three methods are used to make a prediction with three sets of variables: a set of sociodemographics and national-level variables, a set of psychological characteristics and national-level variables, and a combination of these variables. The baseline method is a simple ordinary least squares (OLS) regression. However, an OLS regression ignores that individuals are clustered in groups. It might be the case that individuals living in a specific country, are in a different stage of becoming more environmentally conscious. In other words, the error terms might be correlated due to group effects. As a result, the OLS estimates might be biased. It is possible to estimate a model with country fixed effects, but this will not allow to include other national-level variables. Mixed-models might offer a solution. Still, mixed-models are, as well as OLS regressions, linear models. These models make strong assumptions (Goldstein, 2010). In practice, human behaviour or attitudes might not be determined linearly at all. Non-parametric models do not make these assumptions and thus might increase prediction accuracy. Therefore, the third method is random forest. To evaluate the models, the data set is split up into a train (80%) and test (20%) set and the outcomes are compared by the root mean square error (RSME) of the test set.

$$RSME = \frac{1}{n}\sqrt{\sum_{i=1}^{n}(y_i - \hat{f}(x_i))^2}, \tag{1}$$

with n observations and $\hat{f}(x_1)$ being the prediction of $\hat{f}$ for observation i.

4.1 OLS

An ordinary least squares (OLS) regression estimates a linear relationship between one or more independent variables and a dependent variable. This is done by choosing the coefficients that minimise the sum of squared residuals (RSS). The OLS regression formula looks as follows:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + ... + \hat{\beta}_p x_p, \tag{2}$$

with $\hat{y}$ being the estimation of the dependent variable, $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, ..., \hat{\beta}_p$ the estimated coefficients, and $\hat{x}_1, \hat{x}_2, ..., \hat{x}_p$ the predictors. The residual sum of squares is calculated as follows:

$$RSS = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2} - ... - \hat{\beta}_p x_{ip}). \tag{3}$$

In order to reduce the complexity of the model, redundant variables should be left out of the model. To do so, mixed stepwise selection is used. This method first starts with an only-intercept model, and then ads variables that have a good fit, meanwhile removing variables that become less relevant. As criterion for a good fit, often the Akaiki informatie criterion (AIC) or Bayesian information criterion (BIC) are used (James et al., 2013, p. 71-79).

OLS performs very well if it models a well-measured, linear relationship, for example when modelling the gravity formula. It performs drastically worse when the relationships get more complicated and noisy. More specifically, six assumptions have to hold. First of all, a linear relationship should exist between the predictors and response variables. Secondly, the error terms should be uncorrelated. Thirdly, the variance of error terms should be constant. In other words, this assumption is violated if the variance of the error terms increases if the value of the response increases. Fourth and fifth, an OLS is very sensitive to outliers and high-leverage points (outliers in the response variable). Finally, a solution might be hard to find if there is (multi)collinearity. If all these assumptions hold, the predictions will be accurate (James et al., 2013, p. 92-102).

4.2 Mixed-models

The second step is to estimate mixed-models. Mixed-models allow modelling cluster effects, by estimating a different intercept and/or slope per cluster. Random slopes models are not estimated in this case, so no explanation is necessary. Mixed-models have a few advantages over the simple linear model described before. First, the coefficient estimates will be more accurate if cluster effects are big. Second, the standard error of a previously described simple linear model might be overstated. Mixed models will produce more conservative, more precise standard errors. Finally, multi-level modelling offers insights into the variation between clusters (Goldstein, 2010).

Mathematically, mixed-models are different to an OLS. A mixed-model with random intercepts is explained as follows. In a random intercept model, with m clusters ($j = 1, ..., m$), the intercept β_{0j} becomes a random variable: $\beta_{0j} = \beta_0 + u_j$. u_j is a random variable and has the following parameters: $E(u_j) = 0$ and $var(u_j) = \sigma_u^2$. Then we can consider the linear relationship:

$$y_{ij} = \beta_{0j} + \sum_{h=1}^{p} \beta_h x_{hi} + e_{ij} = \beta_0 + \sum_{h=1}^{p} \beta_h x_{hi} + (u_j + e_{ij}) \tag{4}$$

with $var(e_{ij}) = \sigma_e^2$ and p predictors. In the rewritten part of the equation, the random part is expressed within the brackets, and the fixed part outside the brackets (Goldstein, 2010).

Also, parameter estimation is different. Considering Equation (4), the parameters that need to be estimated are the fixed ones, β_0, ..., β_p, and the random parameters: the variances σ_u^2 and σ_e^2. Although the model is an extension to the linear model explained before, an OLS estimation procedure would not suffice to estimate the fixed and random parameters together. Therefore, maximum likelihood is used to estimate the parameters. To reiterate, the log-likelihood of the outcome for every observation is determined based on some starting values and are summed up. This process is iterated until the log-likelihood of the sample converges. In a normal regression model, maximum likelihood works since the independence of observations are assumed. However, observations are inherently not independent in a random intercept model. Conceptually, with random parameters, an estimated error covariance matrix is used in order to allow correlation of observations within clusters. Since the observations are assumed to be independent between clusters, the likelihood is calculated for one cluster at a time. This procedure allows us to estimate the parameters, but maximum likelihood produces biased estimations for the random parameters because it ignores sampling variation of the fixed parameters. Restricted maximum likelihood (REML) is an often-used alternative that produces unbiased estimates (Goldstein, 2010). ML and REML are good methods for the estimations, but still, the assumptions need to be taken into account.

Essentially, a mixed-model is similar to a normal regression model. Thus, the same assumptions as with OLS need to hold. For example, the normality assumption needs to hold. If it does not hold, the estimated will be consistent, but the standard errors will not. But the very nature of a mixed model is that observations are allowed to be clustered within groups. Therefore, the independence of observations assumption can be relieved to the extent that observations can be correlated within clusters, but not between clusters (Goldstein, 2010).

4.3 Random forests

OLS and mixed models are simple, yet effective ways to model relationships. Unfortunately, they make strong assumptions about the functional form. In the case that these assumptions do not hold, non-parametric methods often outperform parametric models. A very powerful one is an extension of a simple decision tree. Decision trees divide the predictor space into regions. The predictor space is the combination of possible values for the predictors X_1, X_2, ..., X_p. The regions that divide the predictor space, $R_1, ..., R_J$, must be distinct, not-overlapping, and for reasons of simplicity high-dimensional rectangles, also called boxes. All observations that fall within one box, will have the same prediction. Decision trees are grown by recursive binary splitting. In other words, the predictor space is iteratively divided into the boxes $R_1, ..., R_J$, with the aim to minimise the RSS. The RSS is similar to Equation 3, although repeated for the boxes:

$$RSS = \sum_{j=1}^{J} \sum_{i \in R_j} (y_i - \hat{y}_{R_j})^2. \tag{5}$$

However, it is not feasible to consider all different divisions of the predictor space in J boxes. To solve this computational challenge, a top-down approach is used. Only one split at a time is being considered, and the best split at that point is made. It is also called the greedy approach since it does not consider other divisions of the predictors' space in the future. Unfortunately, decisions trees generally have high variance (James et al., 2013).

Bootstrap aggregation, or *bagging*, is a method to reduce the variance of a statistical learning method, like decision trees. The rationale is that averaging a set of observations will reduce the variance. Bootstrapping is taking samples of size n from a data set with replacement, which is often used to quantify the uncertainty of statistical methods. In this case, however, the bootstrap samples are used to make predictions. A method $\hat{f}^{*b}(x)$ is trained on B bootstrap samples. The average of the predictions, so the aggregated prediction of the bootstrap sample, will be the bagging prediction. Mathematically, this looks as follows (James et al., 2013):

$$\hat{f}_{bag}(x) = \frac{1}{B} \sum_{b=1}^{B} \hat{f}^{*b}(x) \tag{6}$$

Bagging trees, however, has a downside. Bagged trees often look similar, especially when there is one strong predictor. This predictor will be at the top of the decision tree in most cases and will make up the most influential division of the predictor space. The trees will thus be highly correlated, and averaging highly correlated trees will not reduce the variance that much. Fortunately, this can be solved with a simple trick. By considering only a subset of m predictors per bagged tree, choosing $m < p$, then the trees will be grown based on different sets of predictors. This will decorrelate the trees and reduce the variance of the aggregated prediction. This method is called random forests (James et al., 2013).

There are two hyperparameters to be tuned. First, the number of trees grown, B. Fortunately, random forests cannot be overfitted by setting B too high, thus it is only required to set B at a sufficiently high number. Secondly, the number of parameters considered for every split, m needs to be tuned. Recall that random forest uses a subset m of the total number of predictors p. Luckily, there is a natural validation set available: the out-of-bag observations. When drawing a bootstrap sample, a subset of all observations will not be used for the estimation. This is the subset of the out-of-bag observations. Generally, this is one-third of the observations. These can be used to calculate the out-of-bag error, which in turn allow for model validation or hyperparameter tuning. The out-of-bag error is generally almost equal to the leave-one-out cross-validation error (James et al., 2013). With only these two hyperparameters to tune, random forest is a relatively simple method to implement.

Random forests often lead to a great reduction of variance in comparison to decision trees or linear regressions and are great in dealing with noisy data. Whereas the former is the effect of not taking into account all predictors, the latter is due to not taking into account all observations. If there is an influential outlier in the data, it will not be considered in one-third of the trees, and thus will not influence the outcomes that much. The same holds for multicollinearity. Since only a subset of the predictors is considered for every tree grown, multicollinearity is reduced. A downside of random forests, however, is that it is difficult to interpret the model. There are tricks to get some insights, like calculating the variable importance. The variable importance is generally calculated by the amount a variable decreases the accuracy averaged out over all the trees. Also, the partial dependence can be calculated, which illustrates the marginal effects of certain predictors. The partial dependence of predictor x_s is calculated by the formula:

$$\hat{f}_S(x_S) = E_{X_C}\left[\hat{f}(x_S, X_C)\right] = \int \hat{f}(x_S, X_C) d\mathbb{P}(X_C) \tag{7}$$

with X_c being the other predictors and $\hat{f}$ the machine learning model used (Friedman, 2001). All in all, random forest is a powerful predictive method (James et al., 2013).

5 Results

In this Chapter, first the estimation procedure of the models is elaborated on, Section 5.1. Second, after the models are estimated, they are evaluated in Section 5.2. In this Section RQ1 and RQ2 are answered. Third, RQ3 is answered when the influence of the individual predictors is assessed in Section 5.3. Finally, the stability, reliability, and robustness of the results is discussed in Section 5.4.

5.1 Estimation procedure

A total of 11 models is estimated. All estimations are done in `R` statistical software. Models 1 to 4 are OLS models, respectively an only-intercept model, a model with sociodemographic factors, a model with psychological characteristics and a model with all predictors. Models 5 to 8 are are mixed-models with random intercepts, estimated using the `lme4` package (Bates et al., 2015) and the significance levels are calculated using the `lmerTest` package (Kuznetsova et al., 2017). Model 5 is a model with only random intercepts and no predictors, model 6 with random intercepts and sociodemographic factors, model 7 with random intercepts and psychological characteristics, and model 8 with all predictors. The sociodemographic factors are regarded as numerical data in the foregoing models. This makes the interpretation more straightforward and is made possible by the fact that almost all categorical variables are ordered. Only gender and job status are included as dummies since there is no natural ordering in these variables. To reduce possible multicollinearity in the parametric models, the variables are selected with a stepwise selection procedure, with the AIC as selection criterium. Model 9 up to and including 11 are random forests models. Due to the nature of the random forests method, there is no intercept model. All models contain national-level variables.

The random forest models are estimated using the `randomForest` package (Liaw and Wiener, 2002). Two hyperparameters need to be set, B and m. Since the number of trees grown, B, does not lead to overfitting, it is set to a thousand for all models. The number of predictors used per tree, m, is tuned by defining a grid of all possible values (1 to p). For all values, a model is trained and evaluated using the out-of-bag RMSE. The model with the lowest out-of-bag RMSE is chosen. In order to perform this grid search, the `caret` package is used (Kuhn, 2021). The optimal values for m are 17, 11 and 28, for the model with only sociodemographics, only psychological characteristics, and all variables respectively.

5.2 Model comparison

Eventually, the best way to assess the predictive power of a model is to put it to a test. All 11 models are tested on 2455 unused test observations and the root mean square error of every model is calculated. These test RSME scores can be found in Table 5.1. First, two intercept only models are estimated and the RSME scores are given in the first column. The upper cell only has one intercept for all countries, the middle cell a different intercept for every country. There is no intercept only model for the random forest. The RMSE of the OLS model is 0.555, with random intercepts 0.533. The consecutive columns tell us what the effect is of the different sets of predictors. Comparing the other columns, a major difference in performance between the

Table 5.1: Test RMSE scores of the OLS, mixed model, and random forests over the three sets of predictors.

Method	**Only intercept**	**Sociodemographics**	**Psychological characteristics**	**All predictors**
OLS	.555	.540	.441	.440
Random intercept	.533	.524	.435	.435
Random forest	NA	.389	.315	.319

sociodemographic set and the psychological characteristics stands out, regardless of the method. The decrease in RMSE ranges from 0.99 for the OLS model, to 0.89 for the random intercept model, and 0.74 for the random forests model. There is, however, no significant difference between the prediction with only psychological characteristics and the full set of predictors, with a difference of at most 0.004. Moreover, even a model with different intercepts per country outperforms a model based on sociodemographic factors. These findings are evidence that psychological characteristics are explanatory predictors, sociodemographic factor rather superfluous in most cases.

The difference between the rows shows how the methods perform. It appears that a mixed model with random intercepts outperforms a simple OLS, and the random forest outperforms both. In fact, the difference between the performance of the random forest and the performance of the other methods is so big, that the difference between the OLS and the mixed model pales into insignificance. The differences between the OLS and random intercept models range from 0.022 to 0.015. The differences between the OLS and the random forests model range from 0.150 to 0.121. The gain of using random forests is greatest in the event that only sociodemographic factors are used for the prediction, with a difference of 0.150. Yet in all cases, random forests prove to generate the strongest predictions by far.

5.3 Predictor influence

Table 5.2 shows the estimations of the OLS models (2 to 4) and the mixed-models (6 to 8). In general, the estimates are fairly stable over the different models, concerning magnitude, sign and significance. Still, a few things stand out. First, almost all sociodemographic predictors decrease with at least half their size and/or become insignificant when adding the psychological predictors to the model, regardless of adding random intercepts. Only the factors number of inhabitants, job status and income stand their ground. In contrast, the psychological characteristics remain fairly stable when adding the sociodemographic factors. This suggests that most of the sociodemographic factors only reflect the effect of psychological characteristics. Secondly, where the national-level variables are most significant and stable over the OLS models, they are removed from the random intercept models. Apparently, the country intercepts cannibalise the effect of the country-level predictors.

Table 5.2: Model output of the OLS and random intercept models

	Model					
Predictor	(2)	(3)	(4)	(6)	(7)	(8)
Intercept	.608**	.753**	.313**	-.012	.580*	.137
Sociodemographics						
Age	.003**		-.001*	.004**		
Female	.113**		.028**	.118**		.035**
Household size	.013*					
Children per household	-.021**			-.013*		
Social class				-.020**		
# of inhabitants	.018**		.013**	.010**		.009**
Job status (base: full-time)						
Part-time (8<hrs<29)	-.003		-.025	-.005		-.031
Part-time (<8 hrs)	.137**		.099**	.104**		.086**
Unemployed	.032		.016	.002		-.002
Sick/disabled	.019		.027	.086*		.050
Retired	.056**		.046*	.032*		.020
Housewife	.094**		.044*	.034		.007

Student	.049**		.054**	.054**		.060**
Education						
Education partner						
Income	-.013**		-.010**	-.014**		-.023**
Income change	-.035**		-.010*	-.022**		
Personality traits						
Openness		.020**	.022**		.028**	.030**
Conscientiousness						
Extraversion		.001				
Agreeableness		.046**	.042**		.044**	.040**
Neuroticism		-.048**	-.040**		-.041**	-.034**
Personal values						
Power		.137**	.140**		.134**	.134**
Achievement		-.343**	-.349**		-.363**	-.362**
Hedonism		-.032	-.041*		-.065**	-.069*
Stimulation		-.184**	-.180**		-.155**	-.150**
Selfdirection		.440**	.447**		.468**	.463**
Universalism		.284**	.273**		.244**	.238**
Benevolence						
Tradition						
Conformity		-.126**	-.124**		-.110**	-.107**
Security		.129**	-.119**		-.109**	-.104**
Consumer values						
Materialism						
Innovativeness		.016**	.018**		.020**	.021**
Normative influence						
Health consciousness		.100**	.099**		.094**	.095**
Quality consciousness		.076**	.078**		.075**	.076**
Consumer ethnocentrism		.115**	.118**		.118**	.119**
Nostalgia		.059**	.052**		.058**	.055**
National-level variables						
ESI	.005**	.002*	.002**			
Log WRI	-.222**	-.129*	-.158**			
Log population density	.090**	.068**	.080**			
Log emission per capita			.013			
Log GDP per capita	-.223**	-.102**	-.108**		-.131*	
Not high income country	.025	.-032*	.030*			
High income * ESI						

Note. * $p < 0.05$, ** $p < 0.01$.

(2) OLS with sociodemographic predictors, (3) OLS with psychological predictors,

(4) OLS with both predictor sets, (6) random intercept model with sociodemographic predictors,

(7) random intercept model with psychological predictors,

(8) random intercept model with both predictor sets.

Which factors are most important? Although model 10, random forest with psychological characteristics performs best, it is not easy to interpret this model. With a thousand decision trees aggregated, it is hard to assess how the model came to its predictions. Therefore, to identify which factors are most important in the predictions, the best performing parametric model is used: the random intercept model with all predictors.

Only a few sociodemographics remain significant. Generally, women living in bigger cities, working part-time or being students and with a lower income are more environmental consciousness. No evidence has been found that age, household size, children per household, social class, educational level and income change have any effect on environmental consciousness. If not only the magnitude of the predictors are taken into account, but also the extent the variables vary, income seems to be the most important sociodemographic predictor.

The psychological factors are more influential in the prediction. Looking at the magnitude and significance of the variables, it is remarkable that values, especially personal values, have a major effect on environmental consciousness. Self-direction, achievement en universalism stand out. Benevolence and tradition, however, are not important. Out of the consumer values, ethnocentrism and health consciousness have the greatest impact. Materialism and normative influence have no impact.

Finally, the effect of the national-level variables is suggested to reflect country effects, since the variables are removed when adding random intercept. When not including the country effects, the following stands out. The ESI score appears to be not very influential, yet positive and significant. This is as expected. However, the interaction effect seems to be absent. The WRI index is surprisingly negative. Apparently, countries that have a higher risk of being damaged by climate disasters, are less environmental consciousness. Emission per capita does not seem to have any effect on the environmental consciousness score. The income variables, GDP per capita and being a high-income country both point in the direction that more income leads to less environmental concern.

Although random forest is a black-box method, some tools can be used to uncover some of

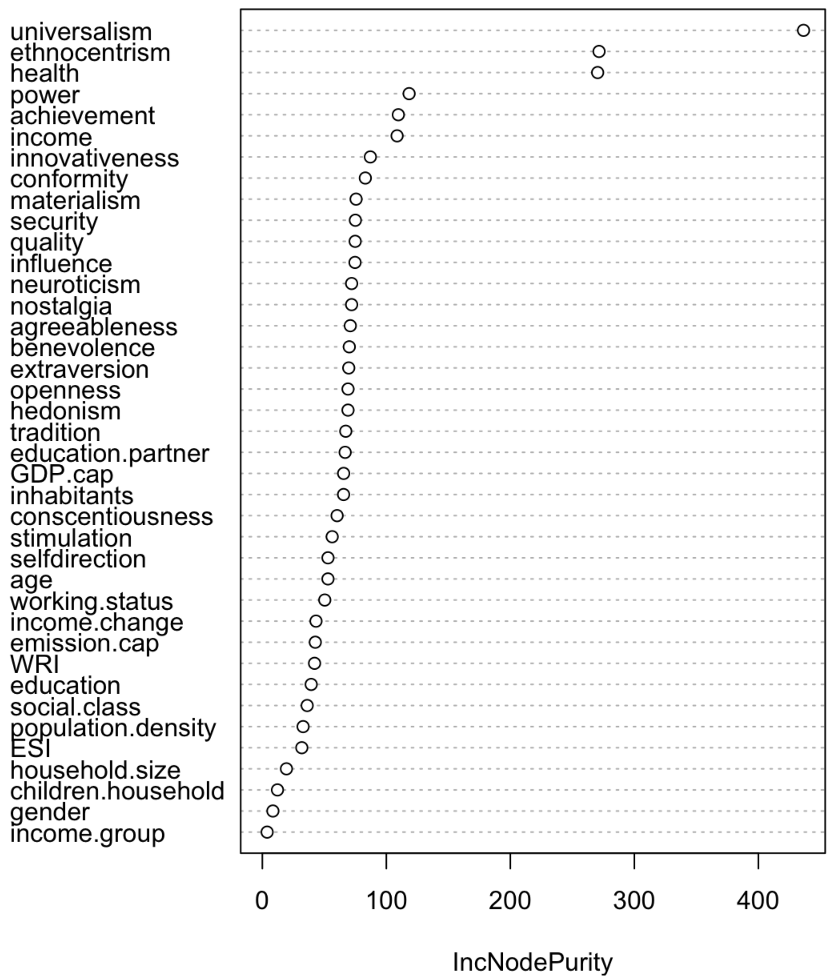


Figure 5.1: Variable importance of random forests model with all variables

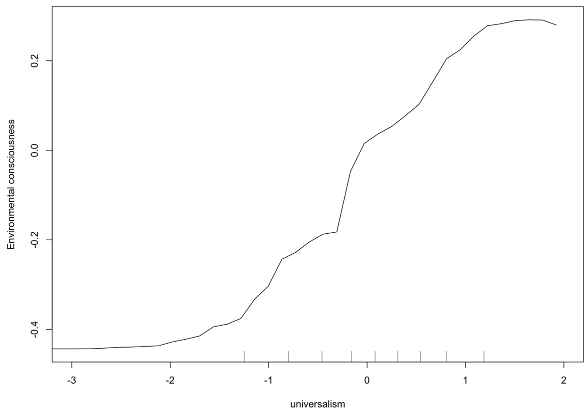
the internal calculations. Figure 5.1 shows the variable importance calculations of the random forest model with all variables. Similar to the parametric models, it suggests that the most important are the personal values, followed by the consumer values. The national-level variables and sociodemographic factors are less influential. Surprisingly though, not the value of self-direction, but universalism is far most important. Similar to the findings of the parametric models, however, are the importance of the consumer values health consciousness and ethnocentrism. Also in accordance with the findings of the parametric models, the most important sociodemographic predictor is income level. Figure 5.2 shows that the direction of the effects of the most important predictors is similar to the findings of the parametric models. The association between universalism and environmental consciousness is positive, just as ethnocentrism, and health consciousness. Power, achievement, and income are negatively associated.

5.4 Assumptions and robustness

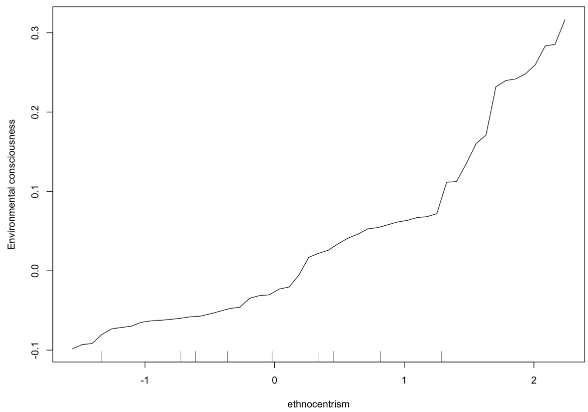
Regarding the parametric models, only one assumption seems to be grossly violated. The residuals plots of the parametric models show few potential problems. The relationship between the predictors and response seems to be roughly linear, the error terms are uncorrelated and heteroskedasticity is absent. However, there seems to be multicollinearity. Multicollinearity occurs when more than two variables correlate with each other. The issue is not that of prediction accuracy, but it leads to problems when identifying the effect of individual predictors (James et al., 2013). The extent of multicollinearity can be measured with the variance inflation factors (VIF). The VIF is the ratio of the variance of the estimated coefficient of a specified predictor of the full model and the variance of the estimated coefficient of that parameter when estimating the model with only that predictor. It will be 1 if there is no multicollinearity. A rule of thumb is that the VIF should not exceed 5 or 10 (James et al., 2013). Figure 5.3a shows the VIF values of the variables of model 8, the random intercept model with all variables, calculated using the `car` package (Fox and Weisberg, 2019). All and only the personal values exceed the maximum VIF values. Especially the VIF value of self-direction is high. The same appears in the other models in which the personal values are included, although these plots are omitted due to brevity. The cause of the multicollinearity is most likely the high correlations between the personal values, which can be found in Figure 3.1. When removing the most problematic variable, self-direction, from the model, the VIF values shrink to an acceptable level, as can be seen in Figure 5.3b.

To what extent is this multicollinearity problematic? In theory, the prediction accuracy should not be compromised by the influence of multicollinearity. Indeed, when evaluating the performance of the parametric models, not including the variable self-direction, the RSME scores remain roughly unchanged (see Figure 7.1, panel a and b in the appendix). The problem at hand, however, is that OLS and random intercept models will have problems identifying the actual effect of the variables that are multicollinear. It is not possible to separate the effect of these variables. Thus, coefficient estimates will not be accurate. Unfortunately, when the extent of multicollinearity is reduced by excluding the variable self-direction, the multicollinear coefficients, those of the personal values, change a great deal, which appears when comparing Figure 5.3c and 5.3d. When looking at this in more detail, something remarkable catches the eye. In the original case, the influence of the personal values is inconsistent between the parametric and the random forests models. When reducing the extent of multicollinearity, however, the influence of the personal values are more in line with each other. This appears for example for the variable universalism, which is most important according to Figure 5.1, and also according Figure 5.3d. To understand this, recall that random forests are not affected much by multicollinearity. Since every decision tree that is estimated only uses a subset of the predictors, the amount of multicollinearity per tree is reduced. The checks offer some insight into the actual coefficients, but unfortunately, multicollinearity keeps from drawing hard conclusions.

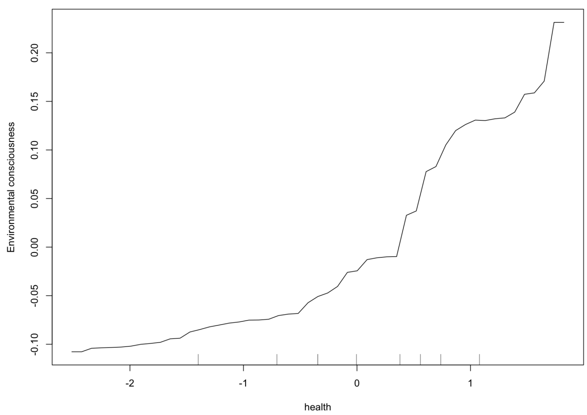
There are three options to deal with the foregoing. First, one could remove the variable or variables that cause multicollinearity. Unfortunately, this is not desirable, since this would mean



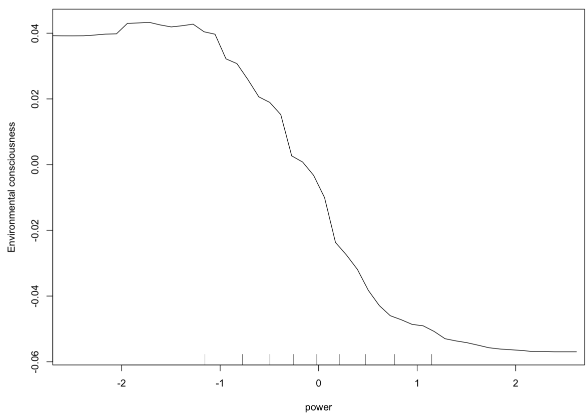
(a) Universalism



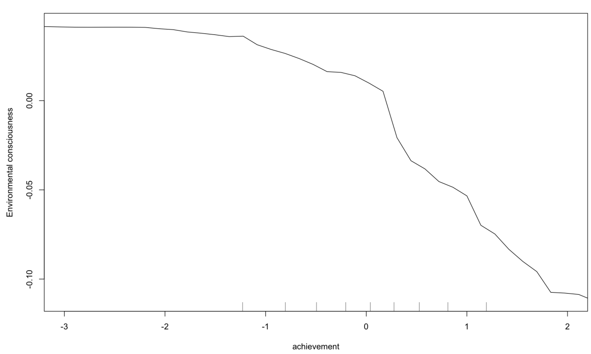
(b) ethnocentrism



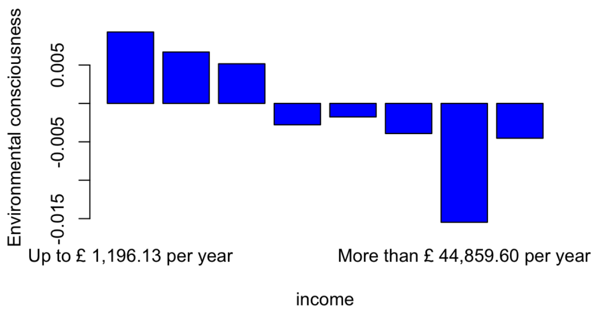
(c) Health consciousness



(d) Power

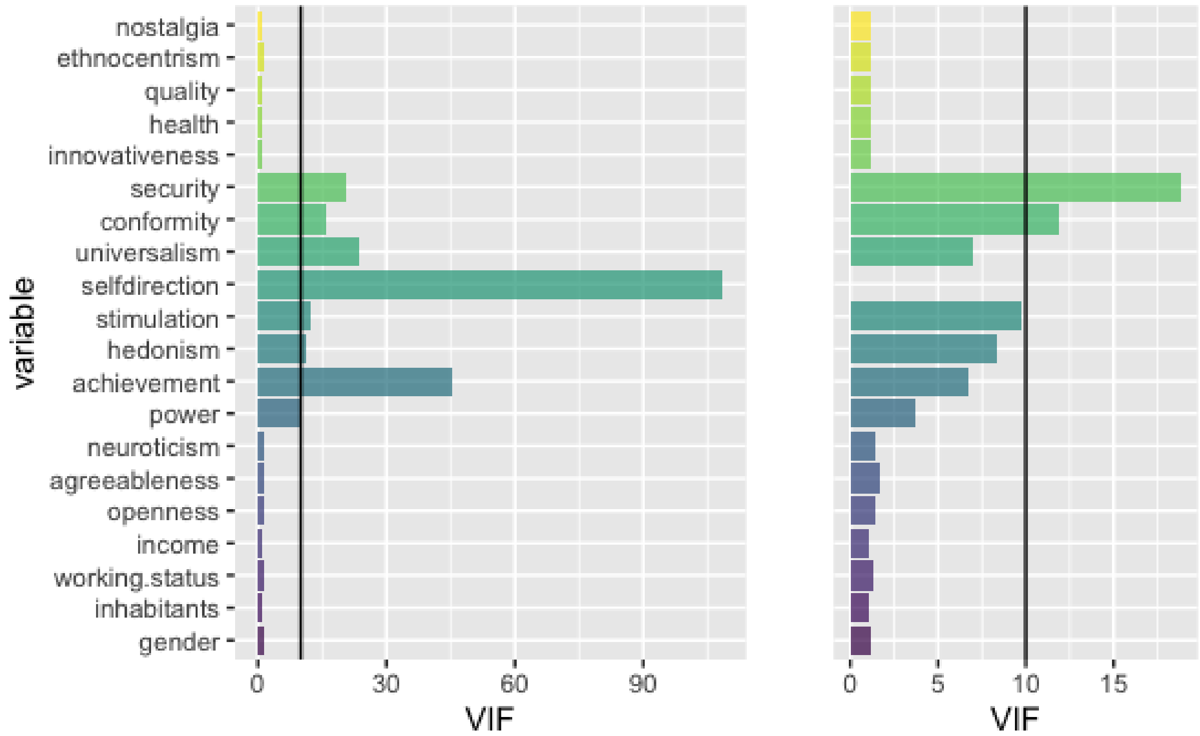


(e) Achievement

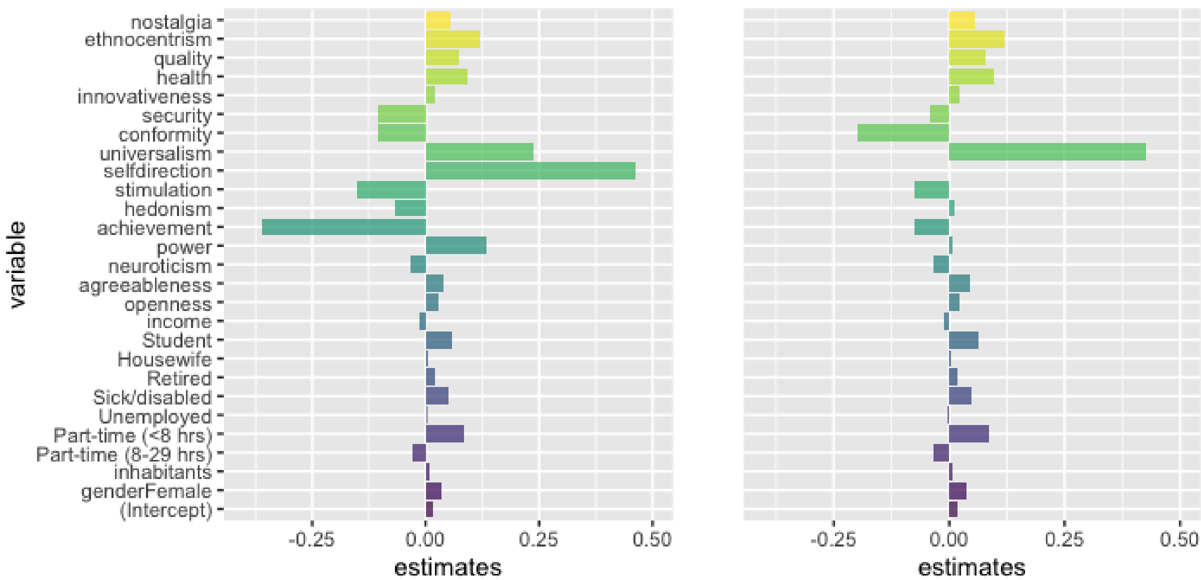


(f) Income

Figure 5.2: Partial dependence plots of four most influential predictors of the full random forests models



(a) VIF values including self-direction (b) VIF values excluding self-direction



(c) Coefficients including self-direction (d) Coefficients excluding self-direction

Figure 5.3: The top part shows the VIF values, which measures the extent of multicollinearity. The left part is the full model including self-direction, the right part is the full model excluding self-direction. The bottom part shows the coefficients with and without self-direction. All figures are based on model 8.

removing possible important predictors from the model. Second, the multicollinear predictors can be combined into less high-order factors. Also this solution is not liberating, although could be recommended for future research of some kind. In this case, however, it would prevent interpreting the influence of the individual personal values. Finally, one could accept a certain level of multicollinearity, knowing that the coefficients might be inaccurate. In this case, a combination between the latter two is possible, also keeping in might the variable importance of the random forest can give some extra confidence.

In order to establish the robustness of the results, the models have been tested on their stability with regards to changes in approach. First, another measurement method is used for the psychological constructs and dependent variable. Recall that in Table 5.1 confirmatory factor analysis has been used to extract the constructs from the survey items. For this test, the construct score is the average score of the items that the construct is represented by. So, the item scores are added up and divided by the number of items. The blue bars in Figure 7.1 in the appendix shows the results. The relative differences to the original RMSE scores are similar over the models. This gives confidence that regardless of the measurement model, the results hold. Second, the green bars of the same Figure shows the RMSE scores when the categorical variables are added as dummies, instead of continuous variables. The scores are similar to that of the original models. The third set of estimations is done without the national-level variables. The reason for this test is to establish the influence of the national-level variables on the performance of the other predictor sets. As the yellow bars in Figure 7.1 in the appendix shows, the results are very similar, indicating that the choice to include national-level variables is not important for the outcome. The tests show the results are robust to changes in the methodology.

6 General discussion

This study attempts to provide concrete tools to accurately predict environmental conscious consumerism. It uses data from 27 countries and models different predictor sets and modelling methods for a cross-comparison of the test RMSE-scores. It appears that psychological variables are stronger predictors than sociodemographic factors. The evidence suggests that there is no need to include sociodemographic factors when psychological characteristics are included. Values seem to be the most influential predictors of the set of psychological characteristics. Both consumer-specific values and personal values add a lot of predictive power, although the latter more than the former. Concerning the methods, it is crystal clear that random forests perform best in every situation. Without any doubt, the best predictions will be made by this method. For reasons of interpretability and simplicity, the relationship might be modelled with a parametric method. In that case, it would be suggested to use a random intercept model with psychological characteristics. In that case, it is not necessary to consider sociodemographic factors or national-level variables. Bringing this all together, I would recommend predicting environmental consciousness using psychological predictors and random forest as modelling method.

So which individual predictors are most important? Frankly, difficulties arise when answering this question. Multicollinearity makes it difficult to separate the effect of certain variables, although this is only limited to the personal values. Based on robust findings in both the parametric and random forests model, it is found that universalism is a very strong predictor. Also achievement, and conformity appear to be strong predictors. Concerning the other personal values, the findings are less consistent, and the findings must be interpreted with care. Also three of the consumer values are consistently strong predictors, namely ethnocentrism, quality and health consciousness. Out of the sociodemographics, the only important predictors are being a student or working part-time under eight hours, and having a lower income. Most other sociodemographic factors are diminished into insignificance when adding the psychological characteristics. In general, we can say that a person is more likely to be environmental consciousness, if he or she is universalistic, not achievement orientated and does not value conformity. Other characteristics associated are being ethnocentric, quality and health consciousness, and having a low income.

Can these findings be explained? The finding that psychological characteristics weigh out sociodemographic factors is in line with the expectations. Recall the finding of Diamantopoulos et al. (2003) that sociodemographic factors only explain 6% of the variance in environmental consciousness. The fact that the effect of sociodemographic factors is cannibalised by the effect of psychological characteristics, is an indication that environmental consciousness is not driven by the sociodemographic background of an individual, but by the psychological characteristics of one. The significant effect of sociodemographic factors that many studies found (e.g. Diamantopoulos et al., 2003; Gifford and Nilsson, 2014; Dietz et al., 1998; Straughan and Roberts, 1999; Wiernik et al., 2013; Buttel and Flinn, 1978), might be understood as a reflection of the underlying relationship between one's values and personality and being environmental consciousness or not.

The sociodemographic factors that remain significant even when psychological characteristics are included, are, next to income, being a student or working part-time (less than 8 hours), being a woman and living in a bigger city. The negative effect of income on environmental consciousness is remarkable. One could say less wealthy people have other concerns than the environment, and being environmentally conscious is a privilege of the wealthy. The opposite is true, which would be an interesting topic for further research. The finding that students are more environmentally conscious, conceptually easier to explain since students are generally more progressive. The same holds for living in bigger cities. The finding that women are more likely to be environmentally conscious is in line with most literature (Diamantopoulos et al., 2003; Gifford and Nilsson, 2014; Dietz et al., 1998; Straughan and Roberts, 1999). It is remarkable that age and education have little to no effect on being environmental consciousness when adding the psychological characteristics since plenty of research find significant effects (Diamantopoulos et al., 2003; Gifford and Nilsson, 2014; Dietz et al., 1998; Straughan and Roberts, 1999). All in all, the findings shed new light on the effect of sociodemographic factors on environmental consciousness.

How are the effects of the psychological characteristics explained? As Soutter et al. (2020) found, the personality traits are all positively related to environmental consciousness and have a reasonably high impact. The reason that personality has less impact on environmental consciousness than values might be explained by the definitions of the two. Recall that personality traits are a description of people, whereas values are motivational (Parks-Leduc et al., 2015; Schwartz, 1992). Values are guiding principles in perceptions, judgments, and behaviour (Parks-Leduc et al., 2015), and they are thus expected to be more explanatory to factors like environmental conscious consumption than personality traits. Interestingly, environmental consciousness is explained a little better by personal values than by consumer values. This could be explained as follows. Schwartz (1992) his Value Theory is exhaustive. If consumer values are eventually covering the same motivational factors as personal values, on a different level, then the consumer value environmental consciousness must be, in theory, fully overlapping with the personal values. Consumer values, in contrast, are exhaustive nor exclusive (Vinson et al., 1977). Associations between consumer values can occur simply because they are overlapping coincidentally. So, this explanation would mean that personal values must be overlapping with consumer values like environmental consciousness. Consumer values can be overlapping with each other.

Zooming in to the personal values, the question is whether the values follow the sinusoid pattern of Figure 2.1 and if so, which side is negatively and which site is positively associated with environmental consciousness. According to most research, especially the self-transcendence dimension should be positively related to environmental consciousness (Gifford and Nilsson, 2014; Milfont and Gouveia, 2006; Nilsson et al., 2004; Schultz and Zelezny, 1999; Stern, 2000; Stern et al., 1995; Stern and Dietz, 1994; Ahmad et al., 2020). Unfortunately, also in this case, one should be very careful interpreting the results due to multicollinearity. The results show no clear sinusoid shape, with the values on the dimension openness to change versus conservation both being negative. What stands out, however, is that the self-transcendence value universalism is strongly positive and the self-enhancement value achievement is strongly negative. This also makes sense. It is understandable that people that are focused on individual achievement, are not focused on the public environment. On the other hand, people that are universalistic, are expected to be focused on their surroundings.

Concerning the consumer values, ethnocentrism and health consciousness are important predictors. The finding that ethnocentrism is positively associated with environmental consciousness is surprising at first glance. However, the perception of the environment can be heterogeneous. Ethnocentric people can be conscious about the local environment, whereas universalistic people can be conscious about the global environment. It was expected that people being health-conscious, are environmentally conscious too. The foregoing findings are not completely in line with what one would expect. What is more or less expected, is the lack of effect of materialism, consumer innovativeness and normative influence (see 2.3.3) It also does not show any clear relationship with the product-time-space framework (Steenkamp and de Jong, 2010). It could be an interesting topic for future research on how these specific values relate to environmental consciousness.

The performance of the methods is mainly as expected. In theory, since there is some heterogeneity in the level of environmental consciousness per country (see 3.2), it is expected that the random intercept models outperform the OLS models. This is what has been found. Thus, the empirical results are in accordance with the theoretical expectations. Expectations have also been met regarding the performance of the random forests. Random forests are capable of dealing with complicated and noisy relationships. Although non of the assumptions of the parametric models seem to be grossly violated (except for multicollinearity, which has no effect on the model performance), apparently random forests are better equipped to deal with the complexity of the predictor-response relation. This gap might be partly explained by the following limitations.

6.1 Limitations

As with all research, this paper has its limitations. Although the rich data is a major benefit to the research, it also comes with drawbacks. First of all, the data is gathered in 2004, and it is the question that the found relationships can be extrapolated to here and now. Back then, being environmentally conscious was something for the front-runners. At the current day, it has become more conventional, with high approval ratings for environmental policies and green consumerism becoming mainstream. It cannot be ruled out that this change also has an effect on the relationship under investigation. On the other hand, psychological characteristics are known to be stable over time (Parks-Leduc et al., 2015). At least the finding that psychological characteristics are important in predicting environmental consciousness, is expected to stand the test of time. Also, it should be noted that the data comes from a sample of countries. These countries are mostly developed. The question arises whether the same mechanisms are present in undeveloped countries. Still, it should be noted that the Value Theory is proven to be universal, which might suggest similar patterns can be found in less developed countries, with respect to the most important predictor set. Altogether, there is no reason to fully question the external validity, however, some caution is required when interpreting the findings.

As becomes clear in Section 2.1, it is not obvious how to define environmental consciousness. The concept encompasses several dimensions (Sánchez and Lafuente, 2010), which can be measured with multiple items. It should be noted that in the current research, specific dimensions are measured, and other dimensions are not. If one is interested in pro-environmental behaviour, for example, this study does not offer relevant insights. As Diamantopoulos et al. (2003) state in their research on the effect of sociodemographics on environmental consciousness: "While, on the face of it, one might anticipate that the relationships hold regardless of the component of the environmental domain at issue, this is by no means the case." The use of the conclusions is thus restricted to the measurement of environmental consciousness.

Not only environmental consciousness is a latent construct; the other psychological characteristics as well. CFA is performed in order to extract these constructs. However, evaluation of these measurement models shows that some measurements models do not have a good fit (recall Table 3.2). This could lead to the information in the predictive models being biased or noisy.

Traditionally, structural equation modelling is used when dealing with latent variables. With structural equation modelling, the measurement model and the structural model are estimated simultaneously. This allows to correct in the structural modelling for measurement errors. In the current case, however, measurement errors are ignored in order to compare multiple modelling techniques. Nevertheless, it is naive to assume there are no measurement errors. These measurement errors evidently cause the models to be biased. Although random forests might be able to deal with these measurement errors better than the parametric models, there is some evidence that also this method might suffer from a loss of prediction accuracy in that case (Jiang et al., 2021). In any case, the presence of biases cannot be excluded.

The variable selection procedure used for the parametric models is known to have major flaws. These follow from the greediness of the approach. Stepwise selection procedures only consider a set of variables at a time and permanently delete or add a variable based on that set. Of course, this approach is very naive. Variables that are deleted at the start of the procedure, might actually add value when the final model is defined. But in the stepwise selection procedure, these variables are not considered later on anymore. This drawback is proven to lead to the standard errors and p-values being biased toward zero, overcomplexity of models, and inflation of coefficients (Harrell, 2001). The application of this procedure might have led to the gap between the performance of the parametric models and the random forests models.

Finally, perhaps needless to repeat, the multicollinearity is causing a serious problem for interpreting the coefficient estimates of the personal values. It is clear that the multicollinearity is causing problems, but not to what extent.

6.2 Managerial recommendations

Although the aforementioned limitations should be taken into account, the conclusions still offer practical recommendations. Previous research already indicated that sociodemographic factors only explain a small fraction of the variance in environmental consciousness (Diamantopoulos et al., 2003). The current study provides new evidence that sociodemographic factors are not satisfactory in predicting environmental consciousness. This is relevant when segmenting the consumer base on environmental consciousness. The recommendation is to steer away from using sociodemographic factors. Psychological factors, especially values, are more explanatory.

Companies often try to diversify by launching green products or services. Big fashion brands nowadays have a sustainable line. Also companies with green products or services as the sole focus are ubiquitous. The strategy is to distinguish oneself in order to raise the price above competition level, as it is established that environmental considerations are an important factor in the decision-making process of consumers (Golob and Kronegger, 2019). It has indeed been found that some consumers are willing to pay more for green products (Didier and Lucie, 2008). Essential to that strategy is targeting environmentally conscious consumers. Logically, only those who value the environment, might be willing to pay extra. Figure 3.2 shows that most countries have a bimodal distribution of environmental consciousness. There is apparently a significant part in every country that is above average environmental conscious. It is recommended to target this group for both companies launching green products and entirely green companies. In order to do so, advanced modelling techniques might be an outcome. Random forests have proven to be superior to simple techniques regardless of the set of predictors used. They outperform parametric models, even only including less explanatory predictor sets. In conclusion, targeting consumers by using advanced modelling techniques is a cost-effective strategy.

The global roll-out of new green products or services is a strategic challenge. Since companies as a matter of course face budget constraints, they are required to make a choice on where to start. The choice depends on multiple factors, but vital when introducing green products or services is obviously targeting green consumers (Cooper, 1979). Unfortunately, individual-level data on environmental conscious consumerism is often not available. Country-level data and cultural components, on the other hand, are almost always available. Based on the findings of the current study, green consumers are often found in cultures that are universalistic, not focused on achievement, and do not value conformity. Also, countries of interest have a lower natural hazard risk, a higher population density, and a lower GDP per capita. For a successful roll-out strategy for new green products or services, countries that fit this image are a good starting point.

The above also holds for environmental policies of governmental institutions. The first roll-out is often not cost-effective and the general success is dependent on the take-up. Take for example decreasing energy use of household by isolating houses. Isolating houses requires a current investment that only pays off in the future. Not all households are keen, or able, to make this investment. A good strategy is to reach environmental conscious households first, so that the policy can be optimised. After the first household adapted the policy, others might be more easily convinced to follow. Essential to this strategy, is targeting environmental conscious households. Governments can use district-level information on psychological characteristics to model the extent to which household are expected to be environmental conscious. By these means, environmental policies can be rolled out more effectively.

6.3 Future research

The significant effect of the values is the main starting point for further research. First of all, since it is known that personal values are the most important factors, but this research failed to identify the effect of individual values on environmental consciousness, a logical next step is to try to overcome the difficulties of the current study. Secondly, as becomes clear, the effect of sociodemographic factors disappears when adding the psychological factors, it would be interesting to find out which psychological factors are causing this. Are the sociodemographic factors merely reflecting the effect of personal values? Or that of personality? Answers to these questions would give insight into what sociodemographic factors actually mean.

Finally, there is cause to uncover the mechanisms of the effects of the values on environmental consciousness. First, the relationship between personal values and consumer values requires attention. Are the personal values indeed a broader concept of the consumer values, and do they cover similar parts of the whole value spectrum, or are the relationships more complicated? Secondly, why do certain personal values increase the chance of being environmentally conscious? Are universalistic people more environmental conscious, because their motivational goal is broader than, for example, achievement orientated people? And does this relate to the higher-order dimensions? Lastly, similar questions arise when looking at the association between the consumer values and environmental consciousness. Why are health and quality conscious consumers generally environmental conscious? Many questions emerge when trying to interpret these findings.

A straightforward, but interesting research would be to redo the study with updated data. Although values and personality traits are known to be rather stable over time (Parks-Leduc et al., 2015), it is still possible things have changed. Environmentalism has become rather mainstream over the years, and this might have had an effect on the relationships researched. Environmentalism could be triggered by other psychological characteristics for these front-runners in 2004 than it would be triggered for people who have a wait-and-see attitude.

To finish, an extension would be to include other sets of predictors or methods. Missing is, for example, the effect of social context on environmental consciousness. People are social beings, that are often affected by their social surroundings. Indeed, research finds that measures of the social context have a similar effect on environmental concern and behaviour, as do sociodemographic factors combined with political attitudes and environmental knowledge (Olli et al., 2001; Dietz et al., 1998; Stern, 2000). More generally, it has been found that social norms are an influential predictor (Klöckner, 2013). On the methodology side, there are many examples of generally very well-performing methods. Neural networks are known for their ability to predict well in very complicated situations. Also, other ensemble methods might be used, like boosting. It would be interesting to see whether these extensions could lead to an increase in prediction accuracy.

7 Appendix

Table 7.1: The psychological constructs and their respective items

Construct	Item
Environmental Consciousness	*I would be willing to stop buying products from companies guilty of polluting the environment, even though it might be inconvenient for me.*
	I become incensed when I think about the harm being done to plant and animal life by pollution.
	When I think of the ways industries are polluting the environment, I get frustrated and angry.
Openness	*Is original, comes up with new ideas.*
	Has an active imagination.
	Is inventive.
	Values artistic, aesthetic experiences.
	Has few artistic interests.
	Prefers work that is routine.
Conscientiousness	*Does a thorough job.*
	Can be somewhat careless.
	Tends to be disorganised.
	Tends to be lazy.
	Does things efficiently.
	Makes plans and follows through with them.
Extraversion	*Is talkative.*
	Generates a lot of enthusiasm.
	Is reserved.
	Tends to be quiet.
	Is sometimes shy, inhibited.
	Is outgoing, sociable.
Agreeableness	*Is helpful and unselfish with others.*
	Starts quarrels with others.
	Can be cold and aloof.
	Is considerate and kind to almost everyone.
	Likes to cooperate with others.
	Is sometimes rude to others.
Neuroticism	*Is relaxed, handles stress well.*
	Can be tense.
	Worries a lot.
	Is emotionally stable, not easily upset.
	Remains calm in tense situations.
	Gets nervous easily.
Power	*SOCIAL POWER (control over others, dominance)*
	AUTHORITY (the right to lead or command)
	WEALTH (material possessions, money)

Table 7.1 continued from previous page

	PRESERVING MY PUBLIC IMAGE (protecting my face)
Achievement	*SUCCESSFUL (achieving goals)*
	CAPABLE (competent, effective, efficient)
	AMBITIOUS (hard-working, aspiring)
	FLUENTIAL (having an impact on people and events)
Hedonism	*PLEASURE (gratification of desires)*
	ENJOYING LIFE (enjoying food, sex, leisure, etc.)
Stimulation	*AN EXCITING LIFE (stimulating experiences)*
	DARING (seeking adventure, risk)
	A VARIED LIFE (filled with challenge, novelty and change)
Selfdirection	*CREATIVITY (uniqueness, imagination)*
	FREEDOM (freedom of action and thought)
	INDEPENDENT
	CURIOUS (interested in everything, exploring)
	CHOOSING OWN GOALS (selecting own purposes)
Universalism	*BROADMINDED (tolerant of different ideas and beliefs)*
	WISDOM (a mature understanding of life)
	SOCIAL JUSTICE (correcting injustice, care for the weak)
	EQUALITY (equal opportunity for all)
	A WORLD AT PEACE (free of war and conflict)
	A WORLD OF BEAUTY (beauty of nature and the arts)
	UNITY WITH NATURE (fitting into nature)
	PROTECTING THE ENVIRONMENT (preserving nature)
Benevolence	*HELPFUL (working for the welfare of others)*
	HONEST (genuine, sincere)
	FORGIVING (willing to pardon others)
	LOYAL (faithful to my friends, group)
	RESPONSIBLE (dependable, reliable)
Tradition	*HUMBLE (modest, self-effacing)*
	ACCEPTING MY PORTION IN LIFE (submitting to life's circumstances)
	DEVOUT (holding to religious faith and belief)
	RESPECT FOR TRADITION (preservation of time-honoured customs)
	MODERATE (avoiding extremes of feeling and action)
Conformity	*POLITENESS (courtesy, good manners)*
	OBEDIENT (dutiful, meeting obligations)
	SELF-DISCIPLINE (self-restraint, resistance to temptation)
	HONOURING OF PARENTS AND ELDERS (showing respect)
Security	*FAMILY SECURITY (safety for loved ones)*
	NATIONAL SECURITY (protection of my nation from enemies)

Table 7.1 continued from previous page

	SOCIAL ORDER (stability of society) *CLEAN (neat, tidy)* *RECIPROCATION OF FAVOURS (avoidance of indebtedness)*
Materialism	*I admire people who own expensive homes, cars, and clothes.* *Some of the most important achievements in life* *include acquiring material possessions.* *I don't place much emphasis on the amount of* *material objects people own as a sign of success.* *The things I own say a lot about how well I am doing in life.* *I like to own things that impress people.* *I don't pay much attention to the material objects other people own.*
Consumer innovativeness	*When I see a new product on the shelf, I am reluctant to give it a try.* *In general, I am among the first to buy new products* *when they appear on the market.* *f I like a brand, I rarely switch from it just to try something new.* *I am very cautious in trying new and different products.* *I am usually among the first to try new brands.* *I rarely buy brands about which* *I am uncertain how they will perform.* *I enjoy taking chances in buying new products.* *I do not like to buy a new product before other people do.*
Normative influence	*If I want to be like someone,* *I often try to buy the same brands that they buy.* *It is important that others like the products and brands I buy.* *I rarely purchase the latest fashion styles until I am sure* *my friends approve of them.* *I often identify with other people by purchasing* *the same products and brands they purchase.* *When buying products, I generally purchase those* *brands that I think others will approve of.* *I like to know what brands and products* *make good impressions on others.* *If other people can see me using a product,* *I often purchase the brand they expect me to buy.* *I achieve a sense of belonging by purchasing* *the same products and brands that others purchase.*
Health Consciousness	*I consider myself as very health conscious.* *I think I do very much for my health.* *I value my health so much that I sacrifice many things for it.*
Quality consciousness	*Quality is decisive for me in purchasing products.* *I always aim at the best quality when purchasing products.*
Nostalgia	*We are experiencing a decline in the quality of life.*

	Table 7.1 continued from previous page
	Things used to be better in the good old days.
Consumer ethnocentrism	*A real Briton should always buy British-made products.*
	It is not right to purchase foreign products, because it puts British people out of work.
	British people should not buy foreign products, because this hurts British business and causes unemployment.
	We should purchase products manufactured in Britain instead of letting other countries get rich from us.

Table 7.2: Comparing means of environmental consciousness between sociodemographic groups

Variable	Mean environmental consciousness	F/T-value
Gender	Female: 0.068 Male: -0.057	-13.02**
Social class	Lower class: 0.157 Working class: 0.065 Lower middle class: 0.009 Middle class: -0.004 Upper middle class: -0.082 Upper class: -0.135	21.97**
Inhabitants	Below 10 000: -0.055 10 0000, to 50 000: -0.037 50 000 to 100 000: -0.038 100 000 to 500 000: -0.016 500 000 to 1 mln: -0.011 Over 1 mln: 0.100	31.12**
Working status	Full-time: -0.044 Part-time (9-29 hrs): -0.012 Part-time (under 8 hrs): 0.183 Unemployed: 0.074 Sick/disabled: 0.008 Retired: 0.106 Housewife: 0.156 Student: -0.023	27.95**
Education	No formal education: 0.151 Up to 12 yrs: 0.166 Up to 14 yrs: 0.077 Up to 16 yrs: 0.062 Up to 18 yrs: -0.009 Higher education: -0.031 University: -0.020	18.04**

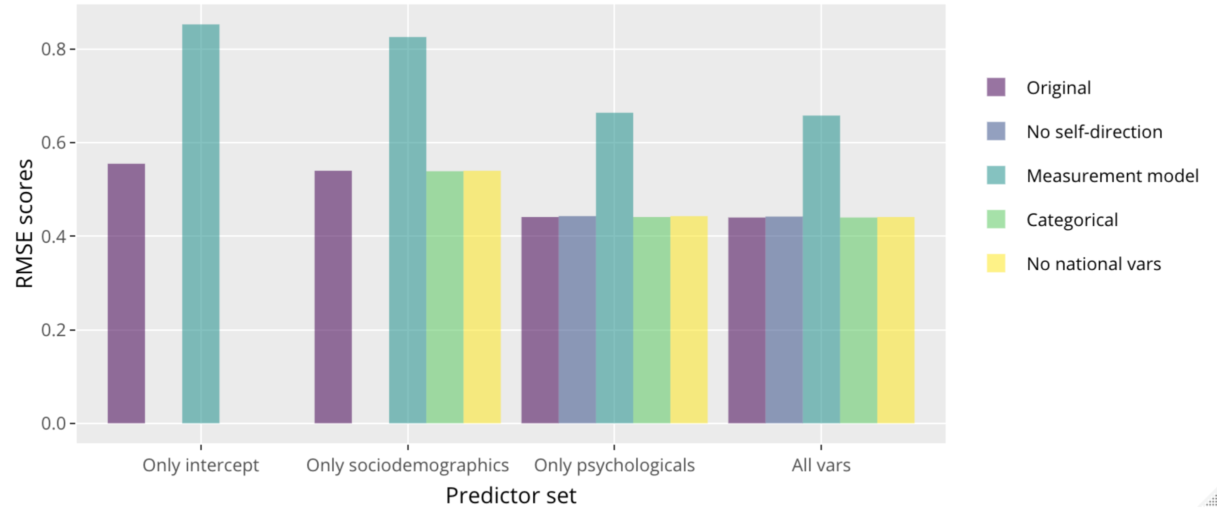
Education partner	Not formal education: 0.105 Up to 12: 0.195 Up to 14 yrs: 0.105 Up to 16 yrs: -0.010 Up to 18 yrs: -0.012 Higher education: -0.035 University: -0.024 Not applicable: 0.014	11.13**
Income	Below 1,196.13 pound per year: 0.167 1,196.13 to 2,990.46: 0.128 2,990.46 to 5,981.12: 0.053 5,981.12 to 11,962.23: 0.024 11,962.23 to 23,925.12: 0.008 23,925.12 to 44,859.60: -0.063 Over 44,859.60: -0.165 No answer: -0.018	50.82**
Income change	Gone down a lot: 0.123 Gone down: 0.052 No change: -0.016 Gone up: -0.060 Gone up a lot: -0.116	46.47***

Note. * $F/T < 0.05$, ** $F/T < 0.01$.

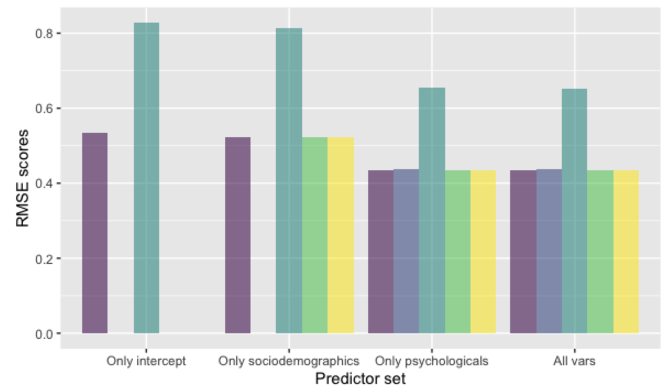
All values are F-values, except for the variable gender, for which a T-test is performed.

Table 7.3: Descriptive statistics of the national-level variables

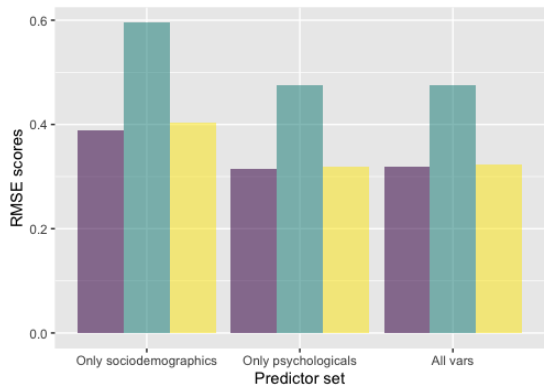
National variable	**Min**	**Max**	**Mean**
Population density (people per square km)	8.795	482.280	130.272
GDP per capita (in 2021 USD)	4426	42 515	25 979
World risk index	2.000	13.570	4.222
Emission per capita (kt in C02 equivalents)	0.000	1.197	0.050
ESI	35.00	73.00	54.24



(a) OLS



(b) Random intercepts



(c) Random forests

Figure 7.1: Effect of robustness tests on the RMSE scores. Purple is based on the original models. Dark blue is excluding self-direction to check the effect of multicollinearity. Light blue are the outcomes when the psychological variables are calculated as the average over their items, instead of using CFA. Green shows the outcomes when the categorical variables are imputed as dummies. Yellow are the models without national-level variables.

8 Bibliography

Abrahamse, W., Steg, L., Gifford, R., and Vlek, C. (2009). Factors influencing car use for commuting and the intention to reduce it: A question of self-interest or morality? *Transportation Research Part F: Traffic Psychology and Behaviour*, 12(4):317–324.

Ahmad, W., Kim, W. G., Anwer, Z., and Zhuang, W. (2020). Schwartz personal values, theory of planned behavior and environmental consciousness: How tourists' visiting intentions towards eco-friendly destinations are shaped? *Journal of Business Research*, 110:228–236.

Alden, D. L., Steenkamp, J.-B. E., and Batra, R. (2006). Consumer attitudes toward marketplace globalization: Structure, antecedents and consequences. *International Journal of Research in Marketing*, 23(3):227–239.

Allik, J. (2005). Personality Dimensions Across Cultures. *Journal of Personality Disorders*, 19(3):212–232.

Balderjahn, I., Peyer, M., Seegebarth, B., Wiedmann, K.-P., and Weber, A. (2018). The many faces of sustainability-conscious consumers: A category-independent typology. *Journal of Business Research*, 91:83–93.

Bamberg, S. and Möser, G. (2007). Twenty years after Hines, Hungerford, and Tomera: A new meta-analysis of psycho-social determinants of pro-environmental behaviour. *Journal of Environmental Psychology*, 27(1):14–25.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48.

Bentler, P. M. (1995). *EQS Structural Equations Program Manual*, volume 6. Multivariate software Encino, CA.

Bleidorn, W., Kandler, C., Hülsheger, U. R., Riemann, R., Angleitner, A., and Spinath, F. M. (2010). Nature and nurture of the interplay between personality traits and major life goals. *Journal of Personality and Social Psychology*, 99(2):366–379.

Boli, J. and Thomas, G. M. (1997). World culture in the world polity: A century of international non-governmental organization. *American sociological review*, pages 171–190.

Brooker, G. (1976). The Self-Actualizing Socially Conscious Consumer. *Journal of Consumer Research*, 3(2):107–112.

Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research*. Guilford publications.

Brown, T. A. and Moore, M. T. (2012). Confirmatory factor analysis. In *Handbook of Structural Equation Modeling*, pages 361–379.

Buttel, F. M. and Flinn, W. L. (1978). The Politics of Environmental Concern: The Impacts of Party Identification and Political Ideology on Environmental Attitudes. *Environment and Behavior*, 10(1):17–36.

Buzzel, R. D. (1968). Can You Standardize Multinational Marketing? *Harvard Business Review*, 46:103–113.

Cooper, A. J., Smillie, L. D., and Corr, P. J. (2010). A confirmatory factor analysis of the Mini-IPIP five-factor model personality scale. *Personality and Individual Differences*, 48(5):688–691.

Cooper, R. (1979). The Dimensions of Industrial New Product Success and Failure. *Journal of Marketing*, 43(3):93–103.

Coppola, M., Krick, T., and Blohmke, J. (2019). Feeling the heat? https://www2.deloitte.com/us/en/insights/topics/strategy/impact-and-opportunities-of-climate-change-on-business.html.

De Groot, J. I. M. and Steg, L. (2009). Morality and prosocial behavior: The role of awareness, responsibility, and norms in the norm activation model. *The Journal of Social Psychology*, 149(4):425–449.

De Groot, J. I. M. and Steg, L. (2010). Morality and nuclear energy: Perceptions of risks and benefits, personal norms, and willingness to take action related to nuclear energy. *Risk Analysis: An Official Publication of the Society for Risk Analysis*, 30(9):1363–1373.

de Maya, S. R., López-López, I., and Munuera, J. L. (2011). Organic food consumption in Europe: International segmentation based on value system differences. *Ecological Economics*, 70(10):1767–1775.

Diamantopoulos, A., Schlegelmilch, B. B., Sinkovics, R. R., and Bohlen, G. M. (2003). Can socio-demographics still play a role in profiling green consumers? A review of the evidence and an empirical investigation. *Journal of Business Research*, 56(6):465–480.

Didier, T. and Lucie, S. (2008). Measuring consumer's willingness to pay for organic and Fair Trade products. *International Journal of Consumer Studies*, 32(5):479–490.

Dietz, T., Stern, P. C., and Guagnano, G. A. (1998). Social Structural and Social Psychological Bases of Environmental Concern. *Environment and Behavior*, 30(4):450–471.

Drori, I. and Yuchtman-Yaar, E. (2002). Environmental Vulnerability in Public Perceptions and Attitudes: The Case of Israel's Urban Centers. *Social Science Quarterly*, 83(1):53–63.

Dunlap, R. E., Van Lieren, K. D., Mertig, A. G., and Jones, R. E. (2000). Measuring Endorsement of the New Ecological Paradigm: A Revised NEP Scale. *Journal of Social Issues*, 56(3):425–442.

Englis, B. G. and Phillips, D. M. (2013). Does Innovativeness Drive Environmentally Conscious Consumer Behavior? *Psychology & Marketing*, 30(2):160–172.

Esty, D. C., Granoff, I. M. E., and Ruth, B. (2002). Yale Center for Environmental Law and Policy (YCELP). page 86.

Ewers, R. M. (2006). Interaction effects between economic development and forest cover determine deforestation rates. *Global Environmental Change*, 16(2):161–169.

Fox, J. and Weisberg, S. (2019). *An R Companion to Applied Regression*. Sage, Thousand Oaks CA, third edition.

Frank, D. J., Hironaka, A., and Schofer, E. (2000). The nation-state and the natural environment over the twentieth century. *American sociological review*, pages 96–116.

Fred van Raaij, W. and Verhallen, T. M. (1994). Domain-specific Market Segmentation. *European Journal of Marketing*, 28(10):49–66.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of statistics*, pages 1189–1232.

García-Álvarez, M. T. and Moreno, B. (2018). Environmental performance assessment in the EU: A challenge for the sustainability. *Journal of Cleaner Production*, 205:266–280.

Gifford, R. and Nilsson, A. (2014). Personal and social factors that influence pro-environmental concern and behaviour: A review. *International Journal of Psychology*, 49(3):141–157.

Givens, J. E. and Jorgenson, A. K. (2011). The effects of affluence, economic development, and environmental degradation on environmental concern: A multilevel analysis. *Organization & Environment*, 24(1):74–91.

Givens, J. E. and Jorgenson, A. K. (2013). Individual environmental concern in the world polity: A multilevel analysis. *Social Science Research*, 42(2):418–431.

Gökşen, F., Adaman, F., and ünal Zenginobuz, E. (2002). On Environmental Concern, Willingness to Pay, and Postmaterialist Values: Evidence from Istanbul. *Environment and Behavior*, 34(5):616–633.

Goldstein, H. (2010). *Multilevel Statistical Models*. John Wiley & Sons, Ltd, fourth edition.

Golob, U. and Kronegger, L. (2019). Environmental consciousness of European consumers: A segmentation-based study. *Journal of Cleaner Production*, 221:1–9.

Grunert, S. C. (1993). Everybody Seems Concerned About the Environment: But Is This Concern Reflected in (Danish) Consumers' Food Choice? *ACR European Advances*, E-01.

Grunert, S. C. and Juhl, H. J. (1995). Values, environmental attitudes, and buying of organic foods. *Journal of Economic Psychology*, 16(1):39–62.

Harrell, F. E. (2001). *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. Springer Science & Business Media.

Hausknecht, J. P., Hiller, N. J., and Vance, R. J. (2008). Work-Unit Absenteeism: Effects of Satisfaction, Commitment, Labor Market Conditions, and Time. *Academy of Management Journal*, 51(6):1223–1245.

Hines, J. M., Hungerford, H. R., and Tomera, A. N. (1987). Analysis and Synthesis of Research on Responsible Environmental Behavior: A Meta-Analysis. *The Journal of Environmental Education*, 18(2):1–8.

Holbrook, M. B. (1993). Nostalgia and Consumption Preferences: Some Emerging Patterns of Consumer Tastes. *Journal of Consumer Research*, 20(2):245–256.

Homer, P. and Kahle, L. (1988). A Structural Equation Test of the Value-Attitude-Behavior Hierarchy. *Journal of Personality and Social Psychology*, 54:638–646.

Hu, L.-t. and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, 6(1):1–55.

Im, S., Bayus, B. L., and Mason, C. H. (2003). An empirical study of innate consumer innovativeness, personal characteristics, and new-product adoption behavior. *Journal of the Academy of Marketing Science*, 31(1):61.

Inglehart, R. (1977). Values, Objective Needs, and Subjective Satisfaction Among Western Publics. *Comparative Political Studies*, 9(4):429–458.

Ivanova, D., Stadler, K., Steen-Olsen, K., Wood, R., Vita, G., Tukker, A., and Hertwich, E. G. (2016). Environmental Impact Assessment of Household Consumption. *Journal of Industrial Ecology*, 20(3):526–536.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An Introduction to Statistical Learning*, volume 112. Springer.

Jiang, T., Gradus, J. L., Lash, T. L., and Fox, M. P. (2021). Addressing Measurement Error in Random Forests Using Quantitative Bias Analysis. *American Journal of Epidemiology*, 190(9):1830–1840.

Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, 34(2):183–202.

Jöreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika*, 36(2):109–133.

Klöckner, C. A. (2013). A comprehensive model of the psychology of environmental behaviour—A meta-analysis. *Global Environmental Change*, 23(5):1028–1038.

Kollmuss, A. and Agyeman, J. (2002). Mind the Gap: Why do people act environmentally and what are the barriers to pro-environmental behavior? *Environmental Education Research*, 8(3):239–260.

Kriwy, P. and Mecking, R.-A. (2012). Health and environmental consciousness, costs of behaviour and the purchase of organic food. *International Journal of Consumer Studies*, 36(1):30–37.

Kuhn, M. (2021). *Caret: Classification and Regression Training.*

Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13):1–26.

Lee, Y.-J., Tung, C.-M., and Lin, S.-C. (2019). Attitudes to climate change, perceptions of disaster risk, and mitigation and adaptation behavior in Yunlin County, Taiwan. *Environmental Science and Pollution Research*, 26(30):30603–30613.

Lemos, M. C. and Agrawal, A. (2006). Environmental governance. *Annu. Rev. Environ. Resour.*, 31:297–325.

Liaw, A. and Wiener, M. (2002). Classification and regression by randomForest. *R News*, 2(3):18–22.

Magnusson, M. K., Arvola, A., Koivisto Hursti, U.-K., Åberg, L., and Sjödén, P.-O. (2001). Attitudes towards organic foods among Swedish consumers. *British Food Journal*, 103(3):209–227.

Masson-Delmotte, V., Intergovernmental Panel on Climate Change, WMO, and United Nations Environment Programme (2019). Climate change and land: An IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems : Summary for policymakers. Technical report, Intergovernmental Panel on Climate Change, Geneva.

Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, Ö., Yu, R., and Zhou, B. (2021). Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Technical report, Intergovernmental Panel on Climate Change.

Matthews, G., Deary, I. J., and Whiteman, M. C. (2003). *Personality Traits.* Cambridge University Press.

Matthies, E., Selge, S., and Klöckner, C. A. (2012). The role of parental behaviour for the development of behaviour specific environmental norms – The example of recycling and re-use behaviour. *Journal of Environmental Psychology*, 32(3):277–284.

McCrae, R. R. and Costa, P. T. (2003). *Personality in Adulthood: A Five-factor Theory Perspective.* Guilford Press.

McCrae, R. R. and John, O. P. (1992). An Introduction to the Five-Factor Model and Its Applications. *Journal of Personality*, 60(2):175–215.

Milfont, T. L. and Gouveia, V. V. (2006). Time perspective and values: An exploratory study of their relations to environmental attitudes. *Journal of Environmental Psychology*, 26(1):72–82.

Minton, A. P. and Rose, R. L. (1997). The Effects of Environmental Concern on Environmentally Friendly Consumer Behavior: An Exploratory Study. *Journal of Business Research*, 40(1):37–48.

Mucke, P., Birkmann, J., Welle, T., Krause, D., Wolfertz, J., Suarez, D.-C., Setiadi, N. J., Radtke, K., and Dickerhof, R. (2011). WorldRiskReport 2011. Technical report, Bündnis Entwicklung Hilft.

Nilsson, A., von Borgstede, C., and Biel, A. (2004). Willingness to accept climate change strategies: The effect of values and norms. *Journal of Environmental Psychology*, 24(3):267–277.

Nordlund, A. M. and Garvill, J. (2002). Value Structures behind Proenvironmental Behavior. *Environment and Behavior*, 34(6):740–756.

Oates, C., McDonald, S., Alevizou, P., Hwang, K., Young, W., and McMorland, L.-A. (2008). Marketing sustainability: Use of information sources and degrees of voluntary simplicity. *Journal of Marketing Communications*, 14(5):351–365.

Olli, E., Grendstad, G., and Wollebaek, D. (2001). Correlates of Environmental Behaviors: Bringing Back Social Context. *Environment and Behavior*, 33(2):181–208.

Paladino, A. (2005). Understanding the Green Consumer: An Empirical Analysis. *Journal of Customer Behaviour*, 4(1):69–102.

Parikh, J. and Shukla, V. (1995). Urbanization, energy use and greenhouse effects in economic development: Results from a cross-national study of developing countries. *Global environmental change*, 5(2):87–103.

Parks-Leduc, L., Feldman, G., and Bardi, A. (2015). Personality Traits and Personal Values: A Meta-Analysis. *Personality and Social Psychology Review*, 19(1):3–29.

Parry, E. J. (2020). The Greatest Threat To Global Security: Climate Change Is Not Merely An Environmental Problem. https://www.un.org/en/chronicle/article/greatest-threat-global-security-climate-change-not-merely-environmental-problem.

Peattie, K. (2010). Green consumption: Behavior and norms. *Annual review of environment and resources*, 35:195–228.

Richins, M. L. and Dawson, S. (1992). A Consumer Values Orientation for Materialism and Its Measurement: Scale Development and Validation. *Journal of Consumer Research*, 19(3):303–316.

Roberts, J. A. and Bacon, D. R. (1997). Exploring the Subtle Relationships between Environmental Concern and Ecologically Conscious Consumer Behavior. *Journal of Business Research*, 40(1):79–89.

Rokeach, M. (1973). *The Nature of Human Values*. The Nature of Human Values. Free Press, New York, NY, US.

Rosseel, Y., Jorgensen, T. D., Rockwood, N., Oberski, D., Byrnes, J., Vanbrabant, L., Savalei, V., Merkle, E., Hallquist, M., Rhemtulla, M., Katsikatsou, M., Barendse, M., Scharf, F., and Du, H. (2021). Lavaan: Latent Variable Analysis.

Sánchez, M. J. and Lafuente, R. (2010). Defining and measuring environmental consciousness. *Revista Internacional de Sociologia*, 68:731–755.

Schlegelmilch, B. B., Bohlen, G. M., and Diamantopoulos, A. (1996). The link between green purchasing decisions and measures of environmental consciousness. *European Journal of Marketing*, 30(5):35–55.

Schultz, P. and Zelezny, L. (1999). Values as predictors of environmental attitudes: Evidence for consistency across 14 countries. *Journal of Environmental Psychology*, 19(3):255–265.

Schwartz, S. (2006). Les valeurs de base de la personne : Théorie, mesures et applications. *Revue française de sociologie*, 47:929.

Schwartz, S. H. (1977). Normative Influences on Altruism. In Berkowitz, L., editor, *Advances in Experimental Social Psychology*, volume 10, pages 221–279. Academic Press.

Schwartz, S. H. (1992). Universals in the Content and Structure of Values: Theoretical Advances and Empirical Tests in 20 Countries. In Zanna, M. P., editor, *Advances in Experimental Social Psychology*, volume 25, pages 1–65. Academic Press.

Schwartz, S. H. (1994). Are There Universal Aspects in the Structure and Contents of Human Values? *Journal of Social Issues*, 50(4):19–45.

Schwartz, S. H. (2011). Values: Cultural and individual. In *Fundamental Questions in Cross-Cultural Psychology*, pages 463–493. Cambridge University Press, New York, NY, US.

Schwartz, S. H. (2012). An Overview of the Schwartz Theory of Basic Values. *Online Readings in Psychology and Culture*, 2(1).

Schwartz, S. H. and Howard, J. A. (1984). Internalized values as motivators of altruism. In *Development and Maintenance of Prosocial Behavior*, pages 229–255. Springer.

Sharma, K. and Bansal, M. (2013). Environmental consciousness, its antecedents and behavioural outcomes. *Journal of Indian Business Research*, 5(3):198–214.

Shimp, T. A. and Sharma, S. (1987). Consumer Ethnocentrism: Construction and Validation of the CETSCALE. *JMR, Journal of Marketing Research*, 24(3):280.

Soutter, A. R. B., Bates, T. C., and Mõttus, R. (2020). Big Five and HEXACO Personality Traits, Proenvironmental Attitudes, and Behaviors: A Meta-Analysis. *Perspectives on Psychological Science*, 15(4):913–941.

Spaargaren, G. (2011). Theories of practices: Agency, technology, and culture: Exploring the relevance of practice theories for the governance of sustainable consumption practices in the new world-order. *Global Environmental Change*, 21(3):813–822.

Sproles, G. B. and Kendall, E. L. (1986). A Methodology for Profiling Consumers' Decision-Marking Styles. *Journal of Consumer Affairs*, 20(2):267.

Steenkamp, J.-B. E. and de Jong, M. G. (2010). A Global Investigation into the Constellation of Consumer Attitudes toward Global and Local Products. *Journal of Marketing*, 74(6):18–40.

Steenkamp, J.-B. E. and Gielens, K. (2003). Consumer and market drivers of the trial probability of new consumer packaged goods. *Journal of Consumer Research*, 30(3):368–384.

Steenkamp, J.-B. E. M., ter Hofstede, F., and Wedel, M. (1999). A Cross-National Investigation into the Individual and National Cultural Antecedents of Consumer Innovativeness. *Journal of Marketing*, 63(2):55–69.

Stern, P. (1999). Information, incentives, and proenvironmental consumer behavior. *Journal of Consumer Policy*, 22(4):461–478.

Stern, P. C. (2000). Toward a Coherent Theory of Environmentally Significant Behavior. *Journal of Social Issues*, 56(3):407 424.

Stern, P. C. and Dietz, T. (1994). The Value Basis of Environmental Concern. *Journal of Social Issues*, 50(3):65–84.

Stern, P. C., Dietz, T., Kalof, L., and Guagnano, G. A. (1995). Values, beliefs, and proenvironmental action: Attitude formation toward emergent attitude objects. *Journal of Applied Social Psychology*, 26(18):1611–1636.

Straughan, R. D. and Roberts, J. A. (1999). Environmental segmentation alternatives: A look at green consumer behavior in the new millennium. *Journal of Consumer Marketing*, 16(6):558–575.

Tellis, G., Yin, E., and Bell, S. (2009). Global Consumer Innovativeness: Cross-Country Differences and Demographic Commonalities. *Journal of International Marketing - J INT MARKETING*, 17:1–22.

Tews, K., Busch, P.-O., and Jörgens, H. (2003). The diffusion of new environmental policy instruments1. *European Journal of Political Research*, 42(4):569–600.

Thogersen, J. (2002). Promoting green consumer behavior with eco-labels. *New Tools for Environmental Protection*, pages 83–104.

Thøgersen, J., Haugaard, P., and Olesen, A. (2010). Consumer responses to ecolabels. *European Journal of Marketing*, 44(11/12):1787–1810.

Thurstone, L. L. (1947). Multiple-factor analysis; a development and expansion of The Vectors of Mind.

Vasi, I. B. (2007). Thinking globally, planning nationally and acting locally: Nested organizational fields and the adoption of environmental practices. *Social Forces*, 86(1):113–136.

Verain, M. C., Bartels, J., Dagevos, H., Sijtsema, S. J., Onwezen, M. C., and Antonides, G. (2012). Segments of sustainable food consumers: A literature review. *International Journal of Consumer Studies*, 36(2):123–132.

Vicente-Molina, M. A., Fernández-Sáinz, A., and Izagirre-Olaizola, J. (2013). Environmental knowledge and other variables affecting pro-environmental behaviour: Comparison of university students from emerging and advanced countries. *Journal of Cleaner Production*, 61:130–138.

Vinson, D. E., Scott, J. E., and Lamont, L. M. (1977). The Role of Personal Values in Marketing and Consumer Behavior. *Journal of Marketing*, 41(2):44–50.

Wang, Y. (2017). Promoting Sustainable Consumption Behaviors: The Impacts of Environmental Attitudes and Governance in a Cross-National Context. *Environment and Behavior*, 49(10):1128–1155.

Wapner, P. K. (1996). *Environmental Activism and World Civic Politics*. Suny Press.

WEF, YCELP, and CIESIN (2002). 2002 Environmental Sustainability Index (ESI). Technical report, Yale Center for Environmental Law and Policy (YCELP)/Yale University, Palisades, NY.

Wiernik, M. B., Ones, S. D., and Dilchert, S. (2013). Age and environmental sustainability: A meta-analysis. *Journal of Managerial Psychology*, 28(7/8):826–856.

World Bank, . (2021). World Bank Country and Lending Groups. https://datahelpdesk-worldbank-org.eur.idm.oclc.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups.

Wothke, W. (1993). Nonpositive definite matrices in structural modeling. *Sage Focus Editions*, 154:256–256.

Xiao, C. and McCright, A. M. (2007). Environmental Concern and Sociodemographic Variables: A Study of Statistical Models. *The Journal of Environmental Education*, 38(2):3–14.

Yilmazsoy, B., Schmidbauer, H., and Rösch, A. (2015). Green segmentation: A cross-national study. *Marketing Intelligence & Planning*, 33(7):981–1003.

York, R., Rosa, E. A., and Dietz, T. (2003). Footprints on the earth: The environmental consequences of modernity. *American sociological review*, pages 279–300.

Zhou, L., Pan, S., Wang, J., and Vasilakos, A. V. (2017). Machine learning on big data: Opportunities and challenges. *Neurocomputing*, 237:350–361.