

I Survey data

To analyze the impact of tautological factors on impactful sustainable behavior, we re-analyzed data from a previous study by the authors [redacted]. In this study, we conducted a representative survey among German households between April to July 2021. Quotas were set for gender, age and federal state based on statistics by the Federal Statistical Office of Germany (DESTATIS, 2019). In total, $N = 10813$ (5491 female, $M_{age} = 46.05$, $SD = 15.39$) respondents were surveyed after removing respondents not conforming to the quotas.

In this survey, we asked respondents to report on different demographic, political and behavioral aspects. Relevant to the present study, we assessed the following situation related factors: Household size, dwelling type and state, living space, fuel heating type, living area, possibility of energetic renovation and added travel time using public transportation instead of car use. We further assessed person-related factors: Personal income, age, education, job status, car possession, number of cars in possession, driven distance by car per year and perceived control of using public transportation. An overview of all these factors as well as the specific questions and answer formats is presented in Table 6. Further information about data processing and item sources can be found in [redacted].

To calculate mobility and shelter (heating) related CFPs, we used the Carbon Footprint Calculator developed by the German Federal Environmental Office (Schunkert et al., 2022; Umweltbundesamt, 2022) in cooperation with non-profit organization ifeu gGmbH (Schächtele & Hertle, 2007) and KlimAktiv gGmbH (Schunkert et al., 2022). The following surveyed factors were needed to calculate heating related CFP: Household size, housing type, year of construction, living space, ownership status, heating system (fuel type) and annual fuel consumption (e.g., litres of heating oil). If respondents did not know their annual fuel consumption, details about their specific heating system and heating behavior was used (e.g., central heating unit, thermostat settings) to estimate CFPs. For the mobility domain, CFPs were calculated using the following information: Household size, car possession and number of cars, car specifics (drive type, average fuel consumption), total distance driven per car/year, use of other mobility options and total distance/year (e.g., bus, train, bike, air travel). If individuals were unable to report total distances, they were estimated based on number and distance of weekly trips to points of interest (e.g., work, shopping, leisure). An overview of all factors used to calculate the domain specific CFP, including wordings in the survey and answer options is presented in Table 7. A more detailed description of the survey methodology, data processing and calculation of CFP can be found in [redacted].

II Models and analysis procedure

Analogous to the majority of previous studies analyzing behavioral drivers of CFPs, we ran OLS Regression models although they might not be the best suited models for CFP analyses due to potential non-linear relationships between predictors and CFPs (Brandenstein et al., 2023). For our results reported in Figure 1 and Figure 2 in the main text, we entered the respective tautological predictors in a step-wise fashion and plotted the resulting model error (MSE) and explained variance (R^2) for each step. Since we were interested in the overall impact of all tautological factors on domain-specific CFPs used in previous studies, we manually entered the tautological factors and did not use a stepwise regression procedure with variable selection. In the present analysis, we were only interested in showing descriptive *changes* of overall model performance and explained sum of squared when including (non-) tautological predictors and did not conduct any formal statistical (frequentistic) analysis or hypothesis testing. Therefore, and since OLS regression slopes and predicted values are robust to most potential OLS assumption violations (Altman & Krzywinski, 2016b; James et al., 2021), we primarily checked for multi-collinearity between model predictors which none of our models showed. However, to get a more accurate estimate of the numerical extent that tautological factors may influence model error and sum of squares, we used Cook's distance to detect further outliers in the OLS context (Altman & Krzywinski, 2016a; Cook, 1977). To be as conservative as possible, we only labeled cases with $D_i > 1$ as outliers and removed them from all models (e.g., Stevens, 1984). Notably, this only led to the removal of one case in mobility-related CFP analyses. Tables 2 and 3 depict the full model results, including model comparisons, errors and explained variance.

Table 2: Model results tautological predictors - CFP Shelter (heating)

Model	Factors	Known consumption				Unknown consumption			
		<i>RSS</i>	<i>SSQ</i>	<i>MSE</i>	<i>R²</i>	<i>RSS</i>	<i>SSQ</i>	<i>MSE</i>	<i>R²</i>
0	Null model (Intercept only)	3094		1	0	6933		1	0
1	Fuel type	2199	895	0.71	0.29	5157	1776	0.74	0.25
2	Fuel type + household size	2056	142	0.66	0.34	4542	615	0.66	0.34
3	Fuel type + household size + living space	1834	223	0.59	0.41	2698	1844	0.39	0.61
4	Fuel type + household size + living space + dwelling type	1809	24	0.58	0.42	2586	112	0.37	0.63
5	Fuel type + household size + living space + dwelling type + dwelling state	1793	16	0.58	0.42	1759	827	0.25	0.77

Note. Stepwise inclusion of tautological predictors in OLS Regression models. Dependent variable: Heating CFP (standardized), separate for respondents with either known and unknown consumption. Known consumption = Annual consumption of heating fuel (e.g., oil, gas) known, Unknown consumption = Annual consumption of heating fuel (e.g., oil, gas) unknown. $N_{known} = 3013$, $N_{unknown} = 6849$. *RSS* = Residual sum of squares, *SSQ* = Explained sum of squares per predictor, *MSE* = Mean Squared error (model error), R^2 = Explained variance. *MSE* and R^2 values rounded to two decimals, all other rounded to nearest one.

Table 3: Model results tautological predictors - CFP Mobility

Model	Factors	<i>RSS</i>	<i>SSQ</i>	<i>MSE</i>	<i>R²</i>
0	Null model (Intercept only)	8801		1	0
1	Car possession	7339	1462	0.83	0.17
2	Car possession + No. of cars	5982	1358	0.68	0.32
3	Car possession + No. of cars + Driven distance: Car	1528	4454	0.17	0.83

Note. Stepwise inclusion of tautological predictors in OLS Regression models. Dependent variable: Mobility CFP (standardized). $N = 8803$. *RSS* = Residual sum of squares, *SSQ* = Explained sum of squares per predictor, *MSE* = Mean Squared error (model error), R^2 = Explained variance. *MSE* and R^2 values rounded to two decimals, all other rounded to nearest one.

III Additional results

To further analyse how the presence of tautological predictors impacts the conclusions drawn from CFP analyses, we ran different OLS regression models. First, we included some distal (personal and situational) predictors previously defined as relevant to individual behavior and emissions into the full tautological predictor models described in [section II](#). For shelter related CFPs, we included income, living area, possibility of renovation, age, job status and education. For mobility-related CFPs, we included income, perceived control over car usage and added travel time when using public transportation. [Table 7](#) provides an overview of the distal factors used in our models, including a detailed description of the scales/items used to assess them. Note that we use these distal factors for illustration purposes only, they might not be the only or most important ones driving CFPs (see also [Table 1](#) in the main text). Analogous to the tautological predictor models

reported in [section II](#) and the main text, we were interested in showing descriptive changes in model performance and particularly importance of distal predictors during presence of tautological predictors. In doing so, we calculated overall model error (MSE) and explained variance (R^2) and calculated changes in both metrics when adding distal predictors to the tautological models as well as when no tautological predictors are included. As shown in [Figures Figure 3](#) and [Figure 4](#) in the main text, the inclusion of distal predictors neither seems to influence model error nor explained variance very much under the presence of tautological predictors. However, when removing the tautological predictors from the models, many of the distal predictors do seem to substantially reduce model error and explained variance in both shelter- and mobility-related CFPs.

To further illustrate the extent to which tautological predictors can mask important distal predictors, we calculated the explained sum of squares (SSQ) for each distal predictors in the presence and absence of the tautological predictors ([Table 4](#) and [5](#)). In line with the findings depicted in [Figures Figure 3](#) and [Figure 4](#) in the main text, the results show that the explained sum of squares of many distal predictors decrease substantially in the presence of tautological predictors.

As a result, including (controlling for) tautological predictors can lead to a reduced impact of distal (personal, situational) predictors in statistical models aimed to explain CFP. Researchers employing tautological predictors in their models may thus assume these distal predictors to be less relevant, although they exhibit substantial explanatory power in the absence of tautological predictors. This result illustrates how the latter predictors may distort the predictive validity of the former which can impose biases to subsequent conclusions and recommended mitigation strategies (see main text).

Table 4: Predictor impact - CFP Shelter (heating)

Factors	Known			Unknown		
	SSQ Distal	SSQ Distal + Tautological	SSQ reduction (%)	SSQ Distal	SSQ Distal + Tautological	SSQ reduction (%)
Income	7	1	85.71	15	12	20.00
Living area	194	4	97.94	138	5	96.38
Possibility of renovation	45	6	86.67	132	1	99.24
Age	120	15	87.50	270	4	98.52
Job status	2	5	00.00	7	1	85.71
Education	7	3	57.14	35	2	94.28
Model error (MSE)	0.88	0.57		0.91	0.25	
Explained Variance (R2)	0.12	0.43		0.09	0.75	

Note. Explained sum of squared for full models per predictor and model type (only distal/distal + tautological). For SSQ distal: All listed factors included as predictors in OLS model. For SSQ distal + tautological: All distal and all tautological predictors (see [Table 1](#)) included. Dependent variable: Shelter (heating) CFP (standardized). SSQ = Explained sum of squares per predictor, SSQ reduction = Reduction of explained sum of squares (percent) per predictor when tautological predictor included in model. MSE, R2 and SSQ reduction values rounded to two decimals, all other rounded to nearest one. $N_{known} = 3013$, $N_{unknown} = 6849$.

Table 5: Predictor impact - CFP Mobility

Factors	SSQ Distal	SSQ Distal + Tautological	SSQ reduction (%)
Income	1043	11	98.95
Perceived control: Car	984	6	99.39
Added travel time: PT	158	5	96.84
Model error (MSE)	0.75	0.17	
Explained Variance (R2)	0.25	0.83	

Note. Explained sum of squared for full models per predictor and model type (only distal/distal + tautological). For SSQ distal: All listed predictors are included in the OLS model. For SSQ distal + tautological: All distal and all tautological predictors (see Table 2) are included. Dependent variable: Mobility CFP (standardized). SSQ = Explained sum of squares per predictor, SSQ reduction = Reduction of explained sum of squares (percent) per predictor when tautological predictor included in model. MSE, R2 and SSQ reduction values rounded to two decimals, all other rounded to nearest one. $N = 8803$.

Table 6: Carbon Footprint calculation items for shelter (heating) and mobility domain

Life domain	Item/Question	Answer options
0. General factors	1. How many people live in your household (including you)?	Free text entry (number)
	2. In which type of house do you live?	Categorical: 1: Single-family house, 2: Apartment building, 3: Terraced house
	3. What is the size of your living space (in m ²)?	Free text entry (number)
	4. Are you a tenant or owner of the apartment/house?	Categorical: 1: Tenant, 2: Owner, 3: Other (rent-free with relatives, e.g. with parents, etc.)
	5. What was the state of construction of the house in which you lived in 2019?	Categorical:, 1: Until 1978 (not renovated), 2: 1979-1983 (not renovated), 3: 1984-1994 (not renovated), 4: 1995-2001 (not renovated), 5: 2002-2011, 6: From 2012, 7: Low energy house, 8: Passive house, 9: Fully energetically renovated (year of construction until 1919), 10: Fully energetically renovated (year of construction 1920-1948), 11: Fully renovated from an energetic point of view (year of construction 1949-2001), 12: Partially renovated old building (until 1969) under monument protection, 98: Unknown

Life domain	Item/Question	Answer options
<i>1. Heating</i>	1. What do you use for heating (type of heating)?	Categorical:, 1: Natural gas, 2: Heating oil, 3: District heating, 4: Wood pellets, 5: LPG, 6: Heat pump, 7: Heat pump operated with green electricity, 8: Solar thermal system, 9: Logs, 10: Wood chips, 11: Electricity (German electricity mix), 12: Green electricity, 13: Hard coal, 14: Brown coal briquettes, 98: Unknown
	2 I know my annual heating consumption.	Categorical: 1: Yes, 0: No
	2.1 How high is your annual consumption (given in the unit in the billing)?	Free text entry / unknown
	3.1 Which heating system do you have?	Categorical: 1: Central heating system, 2: Floor heating, 3: Individual heating, 98: Unknown
	3.2. For answer alternative 1, please specify (except: single heater):	Categorical:, 1: Before 1980, 2: From 1980, 3: From 1995, 4: From 2005, 5: Condensing boiler, 6: District heating, 7: Electric heat pump, 8: Direct electricity (e.g. night storage heating)
	4.1. At what temperature do you set the thermostat / room temperature on average (in ° C)?	Free text entry (number)
	4.2 How do you ventilate?	Categorical:, 1: Permanent ventilation (all windows), 2: Permanent ventilation (some windows), 3: Intermittent ventilation
	4.3 Which hot water use best describes your everyday behavior?	Categorical:, 4: Washing, 5: Showering regularly, but not every day, 6: Daily shower, 7: Frequent bathing
<i>2. Mobility</i>		

Life domain	Item/Question	Answer options
Car	1. How many vehicles does your household own (e.g. cars and motorcycles)?	Multiple Choice: 0 -10
	1.1. Filter if not 0 is specified for 1: Please specify the car class.	Categorical:, 1: Small car (e.g. VW Polo) with a combustion engine, 2: Mid-range car (e.g. VW Golf) with a combustion engine, 3: Upper class car (e.g. VW Passat) with a combustion engine, 4: Plug-in hybrid motor vehicle, 5: Electric vehicle, 6: Motorcycle, 7: Scooter / moped with combustion engine, 8: Electric scooter
	2. What system of drive do they have?	Categorical: 1: Gasoline, 2: Diesel, 3: Natural gas, 4: LPG
	How do you charge your vehicle?	Categorical: 1: Electricity mix in Germany, 2: Green electricity
	What is the average fuel consumption (in l per 100 km or kg per 100 km)? If you are not sure, you can look up the manufacturer's information on the Internet.	Free text entry (number)
	What is the average power consumption (in kWh per 100 km)? If you are not sure, you can look up the manufacturer's information on the Internet.	Free text entry (number)
	The power consumption of the vehicle is already covered by the indication of my annual energy consumption.	Categorical: 1: Yes, 0: No
	I know roughly how many kilometers were driven per vehicle.	Categorical:, 1: Yes, 0: No
	3. How many kilometers were driven per vehicle?	Free text entry (number)
	3.1 To estimate your mileage, we need some information about your behavior:	
	3.1.1 How many days of the week do you use the car to get to work / university / etc?	Single Choice: 0 -7 days, and free text entry

Life domain	Item/Question	Answer options
	3.1.2 How often do you drive your car for leisure activities (journeys less than 100 km)? You have various options for specifying your average driving frequency (per week, per month or per year).	Free text entry:, 1: Per week, 2: Per month, 3: Per year,
	3.1.3 How often do you use your car for shopping? You have various options for specifying your average driving frequency (per week, per month or per year).	Free text entry:, 1: Per week, 2: Per month, 3: Per year
	3.1.4 How often do you use your car to drive more than 100 km (e.g. to visit friends)? You have various options for specifying your average driving frequency (per week, per month or per year).	Free text entry:, 1: Per week, 2: Per month, 3: Per year,
<i>Public transport</i>	I know roughly how many kilometers I cover in public transport.	Categorical: 1: Yes, 0: No, 97: Does not apply, as no public transport was used
	4. How many kilometers do you cover using public transport?	Free text entry (number)
	What type of public transport did you mainly use?	Categorical: 1: Bus, 2: Train (regional and long-distance traffic, tram, underground), 3: Other (free text entry)
	4.1 For an estimate we need some information about your behavior:	
	4.1.1 How many days a week do you use public transport to get to work / university / etc.?	Single Choice: 0 - 7 days
	What type of public transport did you mainly use?	Categorical: 1: Bus, 2: Train (regional and long-distance traffic, tram, underground), 3: Other (free text entry)

Life domain	Item/Question	Answer options
	4.1.2 How often do you use public transport for leisure activities (journeys less than 100 km)? You have various options for specifying your average driving frequency (per week, per month or per year). How big is this distance (one way in km)? What type of public transport did you mainly use?	Free text entry:, 1: Per week, 2: Per month, 3: Per year Categorical: 1: Bus, 2: Train (regional and long-distance traffic, tram, underground), 3: Other (free text entry)
	4.1.3 How often did you take public transport to go shopping in 2019? You have various options for specifying your average driving frequency (per week, per month or per year). How big is this distance (one way in km)? What type of public transport did you mainly use?	Free text entry:, 1: Per week, 2: Per month, 3: Per year, Categorical: 1: Bus, 2: Train (regional and long-distance traffic, tram, underground), 3: Other (free text entry)
	4.1.4 How often did you use public transport in 2019 to cover journeys of more than 100 km (e.g. to visit friends)? You have various options for specifying your average driving frequency (per week, per month or per year). How big is this distance (one way in km)? What type of public transport did you mainly use?	Free text entry:, 1: Per week, 2: Per month, 3: Per year Categorical: 1: Bus, 2: Train (regional and long-distance traffic, tram, underground), 3: Other (free text entry)
<i>Bicycle</i>	I know roughly how many kilometers I cycle.	Categorical: 1: Yes, 0: No, 97: Does not apply, as no bike was used
	5. How many kilometers do you cover by bike?	Free text entry (number)

Life domain	Item/Question	Answer options
	5.1 For an estimate we need some information about your behavior:	
	5.1.1 How many days a week do you cycle to work / university / etc.? How big is this distance (one way in km)?	Single Choice: 0 – 7 days, and free text entry (number)
	How often do you cycle to leisure activities (journeys less than 100 km)? You have various options for specifying your average driving frequency (per week, per month or per year). How big is this distance (one way in km)?	Free text entry:, 1: Per week, 2: Per month, 3: Per year
	5.1.3 How often do you use your bike to cover journeys of more than 100 km (e.g. to visit friends)? You have various options for specifying your average driving frequency (per week, per month or per year). How big is this distance (one way in km)?	Free text entry: 1: Per week, 2: Per month, 3: Per year
	5.1.4 How often did you use your bike in 2019 to cover journeys of more than 100 km (e.g., to visit friends)? You have various options for specifying your average driving frequency (per week, per month or per year). How big is this distance (one way in km)?	Free text entry:, 1: Per week, 2: Per month, 3: Per year
<i>Carsharing</i>	I know approximately how many miles I managed in 2019 using car sharing services.	Categorical: 1: Yes, 0: No, 97: Does not apply, as no carsharing was used
	6. How many kilometers do you cover with car sharing offers?	Free text entry (number)
	6.1 For an estimate we need some information about your behavior:	
	6.1.1 How many days of the week do you use car sharing offers to get to work / university / etc.? How big is this distance (one way in km)?	Single Choice: 0 - 7 days, and free text entry (number)

Life domain	Item/Question	Answer options
	6.1.2 How often do you use car sharing offers to get to leisure activities (journeys less than 100 km)? You have various options for specifying your average driving frequency (per week, per month or per year). How big is this distance (one way in km)?	Free text entry:, 1: Per week, 2: Per month, 3: Per year
	6.1.3 How often do you use car sharing offers for shopping? You have various options for specifying your average driving frequency (per week, per month or per year). How big is this distance (one way in km)?	Free text entry:, 1: Per week, 2: Per month, 3: Per year
	6.1.4 How often do you use car sharing offers to cover journeys of more than 100 km (e.g. to visit friends)? You have various options for specifying your average driving frequency (per week, per month or per year). How big is this distance (one way in km)?	Free text entry: 1: Per week, 2: Per month, 3: Per year
<i>Air travel</i>	7.1 How often do you fly privately within Europe?	Free text entry (number)
	7.1.1 How long do your European flights take on average (in minutes)?	Free text entry (number)
	7.2 How often do you fly privately with intercontinental flights?	Free text entry (number)
	7.2.1 How long do your European flights take on average (in minutes)?	Free text entry (number)

Note. All items and question refer to the year 2019. Further information can be found in [redacted]

Table 7: Overview of distal factors used in models

Factor (type)	Description (number and type of items)	Items and answer option
Income (person-related)	Monthly income in € (1 item, numerical)	"Monthly net income (in 2019)" 1: below 800 €, 2: 800 - 1000 €, 3: 1000 - 1500 €, 4: 1500 - 2000 €, 5: 2000 - 2500 €, 6: 2500 - 3000 €, 7: 3000 - 4000 €, 8: 4000 - 5000 €, 9: 5000 - 6000 €, 10: above 6000 €
Age (person-related)	Age of respondent (1 item, numerical)	Free entry
Job (person-related)	Professional status (end of 2019) (1 item, categorical)	1: Looking for a job, 2: Apprenticeship, 3: Student college / university, 4: Employee, 5: Civil servant, 6: Self-employed, 7: Retirement / Pensioner
Education (person-related)	Highest educational qualification (end of 2019) (1 item, categorical)	1: Left school without certificate, 2: Elementary school, 3: Second modern school (Hauptschule), 4: Second modern school (Realschule), 5: Technical or general university entrance qualification, 6: Bachelor, 7: State examination, 8: Master / Diploma, 9: PhD
Perceived behavioral control: Car (person-related)	Perceived behavioral control over using public transportation instead of car usage (2 items, numerical)	"For me, going without my car and using alternative means of transport (public transport, bicycle, etc.) is" 1. difficult - easy (5 point scale), 2. complicated - simple (5 point scale)

Factor (type)	Description (number and type of items)	Items and answer option
Added travel time PT (situation-related)	Change in time spent commuting to work when using alternative means of transportation (3 items, recoded to one item, categorical)	"How much less time did you have to spend on average if you had used alternative means of transport (e.g. public transport, bicycle, etc.) instead of a car to get to work?" 1: Less than half as much, 2: Half as much, 3: More than half as much, 4: Equal amount, 5: Less than twice as much, 6: Twice as much, 7: More than twice as much
Living area (situation-related)	Area of residency (urban/rural) (1 item, categorical)	"Which of the following categories describes your living situation at the end of 2019 best?" 1: Big city, 2: Suburban or outside of a large city, 3: (Small) town, 4: Village, 5: (Farm) farm in the country, 99: No answer
Possibility of renovation (situation-related)	Possibility to renovate apartment energetically (1 item, categorical)	"If I wanted, I could ensure that my (rented) apartment / house would be energetically renovated and / or a more efficient heating source would be installed" Yes/No

Note. Further information about the specific items, their source and wordings as well as data processing and recoding can be found in [redacted]

References

- Altman, N., & Krzywinski, M. (2016a). Analyzing outliers: Influential or nuisance? *Nature Methods*, 13(4), 281–282. <https://doi.org/10.1038/nmeth.3812>
- Altman, N., & Krzywinski, M. (2016b). Regression diagnostics. *Nature Methods*, 13(5), 385–386. <https://doi.org/10.1038/nmeth.3854>
- Brandenstein, N., Ackermann, K., Aeschbach, N., & Rummel, J. (2023). The key determinants of individual greenhouse gas emissions in Germany are mostly domain-specific. *Communications Earth & Environment*, 4(1), 422. <https://doi.org/10.1038/s43247-023-01092-x>
- Cook, R. D. (1977). Detection of influential observation in linear regression [Publisher: Taylor & Francis]. *Technometrics*, 19(1), 15–18.
- DESTATIS. (2019). Coordinated Population Projection for Germany.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *Statistical learning* (Casella, Fienberg, & Olkin, Eds.; Vol. 1). Springer.
- Schächtele, K., & Hertle, H. (2007). *Die CO2 Bilanz des Bürgers Recherche für ein internetbasiertes Tool zur Erstellung persönlicher CO2 Bilanzen* (tech. rep.).
- Schunkert, S., Siewert, J., Pitz, P., Paar, A., Hertle, H., Berg, F., Dittrich, M., & Dingeldey, M. (2022). *Der UBA-CO2-Rechner für Privatpersonen: Hintergrundinformationen* (Umweltbundesamt, Ed.; tech. rep.). Dessau-Roßlau.
- Stevens, J. P. (1984). Outliers and influential data points in regression analysis. [Publisher: American Psychological Association]. *Psychological bulletin*, 95(2), 334.
- Umweltbundesamt. (2022). UBA Carbon Calculator. Retrieved November 29, 2024, from https://uba.co2-rechner.de/en_GB/start#panel-calc