UNIVERSITÄT
TRIER

Niklas Scheuer

Quantile Regression and Happiness Inequality: Evidence from Germany

Research Papers in Economics
No. 7/26

Quantile Regression and Happiness Inequality: Evidence from Germany

Niklas Scheuer[(a),1]
[(a)]Department of Economics, Trier University

June 15, 2026

Using data from the German Socio-Economic Panel, this paper examines how socioeconomic characteristics shape the distribution of happiness in Germany and how these effects translate into happiness inequality. I estimate quantile regressions for 2012 and show that key socioeconomic characteristics affect happiness differently across the happiness distribution. Building on this heterogeneity, I construct a counterfactual 2017 happiness distribution by evaluating the 2017 covariate distribution under the estimated 2012 quantile-regression coefficient structure. I then examine happiness inequality and find that, despite the stability of observed happiness inequality in the data, the counterfactual distribution predicts a more equal distribution of happiness, especially according to the Gini coefficient. To reconcile observed and counterfactual inequality, I distinguish between a mechanical effect arising from changes in observed characteristics under the estimated model and a residual effect capturing the remaining deviation. The results suggest that the persistence of happiness inequality reflects residual factors not captured by observed covariates or by systematic changes in estimated returns.

JEL codes: C14, C21, D39, D63, I31
Keywords: Life satisfaction, Happiness, Quantile regression, Income, Wealth, Distributional heterogeneity, Germany, Counterfactual

1 Introduction

Happiness[2] is one of the most central emotions in humans and a key component of economic behavior (see Frey (2018) for an overview). Consequently, economic papers analyzing happiness and its drivers have surged. Empirical studies typically identify the drivers of happiness using linear models, often treating it as cardinal, or ordered response models, when its ordinal nature is emphasized. While these approaches are informative, they focus on average effects or impose restrictive distributional assumptions, thereby potentially obscuring important heterogeneity across the happiness distribution (Binder, 2016; Binder and Coad, 2011, 2015; Koenker and Bassett, 1978; Koenker and Hallock, 2001).

This limitation is particularly relevant given that the distribution of happiness is typically left-skewed, meaning only a small fraction of individuals report low satisfaction, while most rate their lives positively. As a result, the determinants of happiness and their effects may

[1]Contact details: Niklas Scheuer, Trier University, Universitätsring 15, 54296 Trier, Germany, phone +49 651 201 2714, scheuer@uni-trier.de. I am grateful for comments from Lea Stulz and Ludwig von Auer. This paper, previously circulating under the title "The distribution of happiness in Germany - What do we know about happiness inequality?" uses data from the German Socio-Economic Panel. Instructions for accessing the data can be found here.

[2]Following much of the empirical happiness literature, I use the terms "life satisfaction" and "happiness" interchangeably throughout the paper (cf. Footnote 7 in Becchetti et al. (2014)). Strictly speaking, however, the empirical measure used in the analysis is reported life satisfaction from the SOEP, while "happiness" is used as shorthand for the broader underlying concept captured by this measure. For a detailed discussion of how these concepts differ from subjective well-being, which is beyond the scope of this analysis, see Diener (2000).

differ substantially across the distribution. For instance, relatively unhappy individuals might experience large gains in happiness from even modest increases in wealth, as these allow them to satisfy certain needs, thereby boosting happiness immediately. In contrast, highly satisfied individuals may derive little additional happiness from further wealth, having already met their essential aspirations. Similarly, commuting may be perceived as stressful and detrimental to happiness for individuals at the lower end of the happiness spectrum, while those who are generally content may be less affected by it.

This neglect of heterogeneity becomes especially important when analyzing potential causes of happiness inequality. Despite substantial changes in economic conditions, happiness inequality has remained remarkably stable in Germany (Clark et al., 2014) and many other countries[3]. To understand this apparent stability, a distributional perspective that goes beyond mean effects and examines how determinants shape the entire happiness distribution provides more detailed and informative insights. While the international literature has begun to address these questions, evidence for Germany remains limited, particularly with respect to counterfactual analyses.

Accordingly, this paper seeks to uncover the distributional determinants of happiness from German data and to quantify how much of the observed stability in happiness inequality between 2012 and 2017 can be explained by changes in observable characteristics and how much remains unexplained.

Therefore, I regress life satisfaction on a wide range of potential factors (logarithmized net equivalized household income, net equivalized household wealth transformed using the inverse hyperbolic sine (IHS), relative income and wealth, socializing on a scale from 1 to 4, ranging from "Every Week" to "Never", education in years, a *not* commuting dummy, subjective health on a scale from 1 to 5 meaning from "Very Good" to "Bad", employment status, age in years, squared age and a dummy for being married)[4] for different quantiles[5] $q \in \{0.1, 0.25, 0.5, 0.75, 0.9\}$ to show the heterogeneous effects along the underlying happiness distribution. The latter is based on survey data from the German Socio-Economic Panel (henceforth SOEP) from 2012. After familiarizing the reader with these distributional patterns, I build a bridge toward estimating happiness inequality and implement a tighter grid of 99 quantiles $q \in \{0.01, 0.02, \ldots, 0.99\}$ for the baseline model. Next, I apply the estimated 2012 coefficients to the realizations of the *same* exogenous variables for the *same* individuals (a balanced sample with $N = 9325$) observed in 2017 to project the counterfactual distribution of happiness and its associated inequality, measured primarily by the coefficient of variation (henceforth CV) and by the Gini coefficient, using a modified Machado-Mata approach (Machado and Mata, 2005). Finally, I compare these projections with the actual CV and Gini values in 2017 to evaluate the extent to which happiness inequality can be explained by observable factors (mechanical effect) versus changes unaccounted for by data or model (residual effect). I provide further robustness checks by varying the treatment of wealth in quantile regressions and juxtaposing the findings from the traditional CV with the absolute dispersion (variance and standard deviation) as inter alia suggested by Clark et al. (2016).

The paper's novelty and main contribution lie in linking quantile regression-based heterogeneity to inequality decomposition in the analysis of happiness. While related to the RIF-based inequality decomposition approach of Becchetti et al. (2014), which focuses on changes in inequality indices, my analysis shifts attention to the full happiness distribution by embed-

[3]See Pedersen and Schmidt (2011) for an overview of happiness in Europe.

[4]For more detailed information on the variables, please see Section 3.1 and Appendix A.1.

[5]Quantile heterogeneity identifies the distributional incidence of socioeconomic characteristics, i.e., where along the happiness distribution these characteristics exert their influence. This information is essential for understanding inequality dynamics, yet it is fundamentally unavailable in mean regressions, variance decompositions or Recentered Influence Function (RIF)-based approaches, which typically collapse distributional effects into a single scalar object (e.g., Firpo et al., 2009, Footnote 9).

ding quantile-specific effects into a counterfactual framework. My paper shows that observed changes in individuals' characteristics would predict a notable decline in happiness inequality, even though inequality in the data remains stable. This joint perspective reveals that the observed stability in happiness inequality over time reflects offsetting distributional forces and cannot be understood from mean regressions or standard inequality decomposition alone.

In a minor contribution, I examine life satisfaction and its determinants *jointly* and *across the entire distribution* of happiness scores for SOEP data of 2012. To the best of my knowledge, only D'Ambrosio et al. (2020) provide a similar cross-sectional analysis in that they also apply quantile regression to examine the influence of various wealth and income measures on life satisfaction in the SOEP data over time. I extend their work by incorporating a different wealth measure (IHS transformed net household equivalized wealth) as well as socializing, a variable which proves highly relevant.

The analysis yields a broad range of findings. Focusing on the cross-sectional estimation of happiness drivers for Germany, I find, most notably, that income has a significant positive effect on happiness, particularly in the lower segments of the happiness distribution. Wealth is also positively associated with happiness throughout most of the distribution, while social comparisons with peers concerning income and wealth have similar negative effects when one's income/wealth is lower relative to peers. However, across the entire distribution, the strongest happiness correlates are subjective health and socializing. Neglecting either factor has a highly significant negative impact on happiness for all quantiles of the happiness distribution, where a subjective health deterioration by one category has the severest consequences for the unhappy by potentially decreasing happiness by more than 1 point. Education oscillates around 0 with significant positive consequences among the unhappiest. Not commuting enhances happiness, although the estimate is statistically significant only in the upper tail. Interestingly, being employed is negatively associated with happiness for those individuals who are already very happy. Older respondents report higher happiness throughout, while men report significantly lower happiness in the upper half of the distribution. Lastly, those who are married report higher happiness throughout the happiness distribution with particularly strong effects among those who are unhappiest.
When focusing on the estimation and explanation of happiness inequality in a counterfactual analysis, I find that, firstly, happiness inequality in my balanced sample remained relatively stable from 2012 to 2017. Depending on the measure of inequality, however, slightly different changes in inequality emerge. The Gini decreases marginally while the CV increases slightly between 2012 and 2017. Secondly, when considering a balanced panel and the structural relationship from 2012, both measures are actually predicted (via a Machado-Mata counterfactual distribution) to decrease for 2017 (around -11% (Gini) and -0.6% (CV)). I do not find significant evidence that coefficients counteracted, which would offer an explanation for the stagnating trend in the data. This implies that unobserved forces (e.g., idiosyncratic shocks) constitute a residual effect overturning the positive influence of the change in covariate realizations.

The remainder of the paper is structured as follows. Section 2 reviews the relevant literature. Section 3 introduces the quantile regression framework and presents the distributional patterns in the determinants of happiness. Section 4 then turns to the analysis of happiness inequality and the counterfactual decomposition. Section 5 provides robustness checks, before Section 6 concludes.

2 Related Literature

I acknowledge that the determinants of happiness have been extensively studied in the broader literature. However, the following review focuses specifically on studies that examine the drivers of happiness in the German context first. I then discuss the literature on happiness inequality, with particular emphasis on Germany.

Frey and Stutzer (2000) use German data to identify key determinants of happiness. Regarding income, they employ distributional dummy variables and find a significant positive relationship between income and life satisfaction in Germany, even after controlling for a wide range of covariates. Frijters et al. (2004) examine the impact of rising real incomes on life satisfaction in East Germany following reunification applying ordered logit models with fixed effects. They further decompose their results by age and education concluding that income and employment status are major drivers of subjective well-being, indicating that higher income tends to increase happiness. D'Ambrosio et al. (2009) focus on the effects of permanent income and wealth on well-being using dummy variables to represent individuals' positions within the respective income and wealth quintiles. While previous studies often account for temporal dynamics of life satisfaction, my analysis contributes by explicitly including both income and wealth measures and by examining their heterogeneous effects across the entire happiness distribution. Since a large body of research emphasizes the role of relative income, i.e., income comparisons with peers in the spirit of Luttmer (2005) and Headey et al. (2010), I also account for that.

Following up on D'Ambrosio et al. (2009), research on the effects of wealth on happiness, in the sense of, e.g., Deaton and Paxson (1994), for the case of Germany is rather scarce (as also noted by Jantsch et al. (2024)), although it might be more important for happiness than income (Brulé and Suter, 2019). Kasinger et al. (2023) analyze the role of wealth in the happiness gap between East and West Germany. Building on D'Ambrosio et al. (2009), D'Ambrosio et al. (2020) show that both permanent income and wealth affect life satisfaction, primarily through social comparison (for which I also account). Most interestingly, they use a quantile regression framework and subsequently examine how these effects vary across the life satisfaction distribution (see their Figure 1) and find little difference between standard OLS and distributional estimates. I complement their analysis with cross-sectional evidence for 2012 using an alternative wealth measure. While past wealth levels cannot be included by design, my model further controls for a wider range of factors, including socializing and (not) commuting, which are key predictors of life satisfaction absent from their estimations. Jantsch et al. (2024) study relative and absolute wealth from the German Panel of Household Finances and its impact on life satisfaction. I add to their approach by considering the entire distribution in the estimation process.

When it comes to socializing, the evidence for a positive impact on happiness is overwhelming (Becchetti et al., 2012; Bruni and Stanca, 2008; Gunaydin et al., 2021; Kahneman et al., 2004; Powdthavee, 2008). Focusing on the case of Germany, Mader and Franzen (2025) analyze social capital, measured through the number of friends, over time. Although they control for a number of factors, they omit wealth as a control factor disregarding its status effect (Cole et al., 1992; Francis, 2009) on life satisfaction and overlook heterogeneity across the happiness distribution. I contribute on both fronts. However, potentially closest to me, Schmiedeberg and Schröder (2017) study socializing, i.e. meeting with friends, over time for the case of Germany. Their socializing variable contains similar realizations (see their Table 2), however, in contrast to my approach, they include every realization[6] as a dummy variable and find that weekly and daily meetings with friends have a significant effect on life satisfaction over time. While they are specifically targeting the effect of leisure activities, such as socializing, on life satisfaction over time, I provide a cross-sectional, more comprehensive picture including more factors and allowing for different distributional effects.

[6]Realizations for the variable I use are [1] "Every Week", [2] "Every Month", [3] "Less Frequently" and [4] "Never", whereas their realizations are "Never", "Less than 1×/month", "At least 1×/month", "At least 1×/week" and "Daily". This discrepancy stems from using the family panel of the SOEP and me using the standard private survey data. See also Section A.1.2 for more information.

It is well established that education positively influences a range of life outcomes (Heckman et al., 2018). Focusing on Germany and studies using SOEP data, Bertermann et al. (2023) examine the impact of education on life satisfaction, exploring heterogeneity across both the schooling distribution and employment status. They report no effect along the schooling distribution, which is consistent with my analysis. However, they identify an indirect effect of education on life satisfaction mediated through higher income.[7] Finally, Durst (2021) emphasizes the positional nature of education and reports a positive association between education and life satisfaction, even when controlling for income and other covariates. She also considers distributional aspects related to income. I extend her work by analyzing years of schooling, rather than degree dummies (vocational or university, with no vocational degree as the reference category), and by examining effects across the entire distribution of happiness.

In general, lower subjective, as well as objective health, is associated with lower life satisfaction (Easterlin, 2005, Section 2.1). Focusing on research on the impact of health on happiness in Germany, the evidence is rather scarce. Ruseski et al. (2014) establish a causal link between exercise and self-reported happiness using German survey data, hinting at a positive effect of having good health on happiness. A similar finding is attained by Headey et al. (2010, Table 1) using the SOEP and exercise frequency as a measure for healthy lifestyle. Pagán-Rodríguez (2010) analyzes the effect (ex ante, ex nunc and ex post) of a negative health shock (disability) on life satisfaction. Using objective health measures (time spent in hospital), Frijters et al. (2004) find negative effects of bad (objective) health on life satisfaction. Comparing to the present study, I confirm the findings by Frijters et al. (2004) and complement their results (and those of the previously mentioned authors) by including subjective health and estimating impacts across the happiness distribution.

Studies on commuting in Germany cover a broad range of topics, from determinants of flows in general (Chen et al., 2021) or between states (Piliuk et al., 2023) to the persistence of habits (Jost, 2024) and sustainability effects of working from home as opposed to commuting (Böhnen and Kuhnimhof, 2024). Focusing on life satisfaction, Stutzer and Frey (2008) comprehensively analyze commuting behavior in Germany focusing on the average relationship between commuting and life satisfaction (although they refer to it as subjective well-being) using pooled OLS and fixed-effects regressions. Lorenz (2018) provides a similar analysis for a more comprehensive subjective well-being measure and a slightly different time frame. Having multiple robustness checks, their results describe effects at the conditional mean of the happiness distribution, i.e., how commuting affects the "average individual". They do not explore whether their effects differ across the happiness distribution. I contribute by extending this to the distributional level.

Lastly, focusing on happiness inequality, there are notable declines in happiness inequality across the world. While there is plenty of evidence on the US (Brooks, 2008; Stevenson and Wolfers, 2008), German evidence remains relatively scarce. To the best of my knowledge, Becchetti et al. (2014), Clark et al. (2014) and Clark et al. (2016) are the only studies focusing on Germany in the context of happiness inequality (with partly focusing on the role of income). Becchetti et al. (2014) analyze the drivers of happiness inequality over time considering two time windows (mid 90s vs. mid 2000s) using an RIF approach combined with an Oaxaca-Blinder decomposition. They decompose changes in inequality indices into components attributable to shifts in observable characteristics and changes in coefficients, finding that the observed decline in happiness inequality is driven almost entirely by changes in covariates, with coefficient effects playing a negligible role. More generally, my approach is also related to regression-based methods for inequality analysis. Aaberge et al. (2005) provide a framework for studying how covariates affect (pseudo-)Lorenz- and Gini-type inequality objects directly. Clark et al. (2014) find that, as GDP rises, the share of people reporting the lowest happiness levels shrinks, but

[7] I conducted a robustness check for this interaction effect and found an insignificant effect around 0.

the highest happiness category is also becoming less common across the SOEP. Instead, more people cluster in the middle categories, meaning the decline in happiness inequality reflects a mean-preserving contraction toward the center. Clark et al. (2016) focus on six surveys, among which is the SOEP, and show descriptively the impact of GDP growth (see their Figure 1 and 2, excluding Germany) and income heterogeneity (see their Figure 3, upper right corner) on happiness inequality. Aligning with my research approach, they find that happiness varies less among individuals with higher incomes than among those with lower incomes, suggesting the top earners are better shielded from shocks. Higher income, in particular, helps people maintain their consumption even when their earnings fluctuate and potentially cushions the blow of different life shocks (e.g., divorce).

Firstly, I contribute to all of these previous studies. I explicitly incorporate wealth, which is an important yet largely neglected determinant of happiness inequality in the German context. Secondly, I complement Becchetti et al. (2014) by shifting the focus from inequality indices to the full happiness distribution, emphasizing how individual characteristics shape different parts of the distribution and thereby influence inequality. This allows me to identify where distributional changes originate, rather than only how much inequality changes (they do observe the distributional pattern of education, however.). I also include subjective health and socializing, which, based on the conditional quantile regression, are vital for an individual's happiness and Becchetti et al. (2014) neglect. Lastly, whereas Becchetti et al. (2014) find that the observed decline in happiness inequality is driven almost entirely by changes in covariates (Becchetti et al., 2014, p. 438), I find for 2012-2017 that observed happiness inequality remains broadly stable even though changes in covariates would have predicted a decline under the 2012 coefficient structure. Reconciling my findings with Becchetti et al. (2014), differences in time periods and methodological approaches (counterfactual exercises, considering the same individuals et cetera) likely account for these contrasting findings. Concerning Aaberge et al. (2005), my analysis is related in spirit, because it also links covariates to inequality through distributional regression methods. However, it follows a different route. I estimate conditional quantiles of happiness, construct a counterfactual marginal distribution and then evaluate inequality measures on that counterfactual distribution, rather than estimating pseudo-Gini regression coefficients directly. Compared to Clark et al. (2014, 2016), I use a more econometrically intensive approach and also include the Gini aside from the CV, giving a broader picture of happiness inequality.

3 Quantile Regression

This section has several purposes. The first is to familiarize the reader with the data I focus on (cf. Section 3.1). Secondly, it provides an overview of the principles, estimation and interpretation of the quantile regression (cf. Section 3.2) and, thirdly, it covers the distributional aspects (cf. Section 3.3).

3.1 Preliminaries

Life satisfaction is my dependent variable. It is measured on an 11-point Likert scale[8] from 0-10, where respondents rate their satisfaction from "Completely dissatisfied" (0) to "Completely

[8]Ferrer-i-Carbonell and Frijters (2004) and Pagán-Rodríguez (2010) show that distinguishing between ordinal and cardinal scales yields negligible differences in results. In line with Kaiser and Vendrik (2020) and their findings, I treat the life satisfaction variable as a cardinal measure to facilitate interpretation, which is also appealing when turning to happiness inequality as Bérenger and Silber (2022) pointed out.

satisfied" (10). Its distribution for the survey years I concentrate on, i.e., 2012 and 2017[9], is

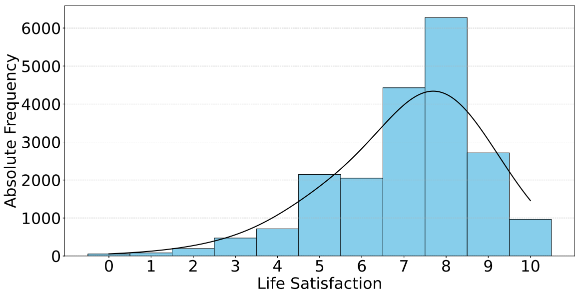

Figure 1: 2012

Figure 2: 2017

The results indicate that the distributions are strongly left-skewed, implying the majority of the sample is reporting relatively high levels of happiness. This skewness motivates the use of quantile regression as it allows us to examine the effects of independent variables across different points of the underlying distribution rather than relying solely on the mean. My primary independent variables include income, wealth, socializing with friends, education, health and commuting, along with several additional control variables. The following provides a brief description of the data and any required modifications. In general, variables were cleaned by removing observations where respondents either did not answer the question or where their questionnaire did not contain the question of interest. Regarding income, I use the natural logarithm of the equivalized net household income (Binder and Coad, 2011; Frijters et al., 2004). I transform net equivalized household wealth using the inverse hyperbolic sine (IHS) transformation (Bellemare and Wichman, 2020). Wealth is highly right-skewed, spans a wide support and includes zero as well as negative values (debt), making the logarithmic transformation inappropriate. The IHS behaves similarly to the natural logarithm for large positive values while remaining defined at zero and for negative observations, thereby compressing the upper tail. This transformation also ensures greater consistency in interpretation with log-transformed income, facilitating comparability of coefficients across economic resources. Aside from wealth itself, I, in the spirit of Luttmer (2005) and Van Praag (2011), also consider the influence that comparing with the reference group (peers) concerning income and wealth has on happiness. To proxy these reference-group positions in a parsimonious way, I classify individuals relative to the median income and wealth of their reference group (following, in general, Brown and Gray (2016)) defined by year, gender, age band, education and federal state. Specifically, and in line with Becchetti et al. (2014), individuals are coded as relatively low-income if equivalized household income is below 60% of the reference-group median and as relatively high-income if it exceeds 200% of that median. Analogously, indicators for wealth are constructed based on the reference-group wealth median. These relative-position measures capture social-comparison channels in income and wealth without imposing functional form assumptions. Socializing[10] is social time spent with friends and measured ordinally but included as a cardinal variable (cf. Binder and Coad, 2011) with realizations from "Weekly"(1) to "Never"(4). Education is measured in years of schooling. Subjective health is measured ordinally but included as a cardinal variable (cf. Binder and Coad, 2011) with a scale ranging from "Very Good"(1) to "Bad"

[9]The reader may wonder as to why I do not use more recent data. The reason is that data on equivalized household wealth are only available for 2002, 2007, 2012 and 2017 when concentrating on German survey data by the German Institute of Economic Research. Being of utmost importance for the analysis of individual happiness, the two most recent survey years I can focus on including wealth are 2012 and 2017.

[10]The 2012 SOEP wave does not measure frequency of socializing in a way comparable to 2017, instead it solely reports satisfaction with one's social life. To maintain comparability, I use the 2011 frequency-of-socializing measure for the *same* individuals of the entire analysis, assuming no change between 2011 and 2012.

(5). Commuting is included as dummy "Not Commuting" indicating whether the individual has no commute[11] to work (dummy "Not Commuting" = 1) or has to commute (dummy "Not commuting" = 0). Lastly, I also account for age, gender via a Male dummy, employment status via a dummy and marital status via a dummy "Married" as individual-specific controls[12].

3.2 Estimation

For the regression, let

$$\mathbf{X}_i = \begin{bmatrix} \text{LogEquivNetHHInc}_i, & \text{IHSEquivNetHHWealth}_i, & \text{Social}_i, & \text{Educ}_i, \\ \text{RelIncLow}_i, & \text{RelIncHigh}_i, & \text{RelWealthLow}_i, & \text{RelWealthHigh}_i, \\ \text{NotCommuting}_i, & \text{SubjHealth}_i, & \text{Employed}_i, & \text{Age}_i, \\ \text{Age}_i^2, & \text{Male}_i, & \text{Married}_i & \end{bmatrix}$$

denote a $k \times 1$ vector of covariates for individual i and let $\boldsymbol{\beta}$ be the corresponding vector of coefficients. Then the standard OLS specification with life satisfaction as the dependent variable can be written compactly as

$$\text{LifeSa}_i = \alpha + \mathbf{X}_i^\top \boldsymbol{\beta} + \epsilon_i, \quad \epsilon_i \sim \text{i.i.d. } (0, \sigma^2)\,. \tag{1}$$

To allow the effects of covariates to vary across the conditional distribution of life satisfaction, I estimate the q-th conditional quantile using quantile regression. Let $Q_q(\text{LifeSa}_i \mid \mathbf{X}_i)$ denote the regression for conditional q-th quantile. The quantile regression model then reads (following the notation of inter alia Koenker and Bassett (1978), Buchinsky (1995), Chernozhukov et al. (2013) or Binder and Coad (2011))

$$Q_q(\text{LifeSa}_i \mid \mathbf{X}_i) = \alpha_q + \mathbf{X}_i^\top \boldsymbol{\beta}_q. \tag{2}$$

Estimating Equation (2) yields the quantile-specific coefficients reported in Table 1 below[13]:

---

[11]First, the "Not Commuting" indicator aggregates all individuals for whom commuting is either not applicable or effectively zero. This group includes remote workers, individuals whose workplace is located at home and non-employed respondents. What these cases share is the absence of *daily travel to work*, which represents the only harmonizable dimension of the commuting measure between 2012 and 2017. While this aggregation masks underlying heterogeneity in employment arrangements, it provides the closest conceptually consistent treatment across waves concerning commuting.

[12]For more details regarding the data, please see Appendix A.1.

[13]For analytical details on the estimation of quantile regression, including the role of the so-called "check function" and the linearization of the underlying non-linear optimization problem, see Koenker and Bassett (1978) for the original contribution and Binder and Coad (2011) for a more recent applied exposition. To solve this, I follow the recommendation from Koenker (2005, ch. 6) for larger quantile regression problems and estimate the parameters numerically applying the interior point algorithm (based originally on Karmarkar (1984)) using R, specifically the command "rq.fit.fn". This solving algorithm is best-suited for larger estimations compared to the standard simplex (Koenker and Bassett, 1978) or iterative weighted least square approach (convergence issues).

	Dependent variable: Life Satisfaction				
	$q = 0.1$	$q = 0.25$	$q = 0.5$	$q = 0.75$	$q = 0.9$
	(1)	(2)	(3)	(4)	(5)
Intercept	4.235***	5.413***	7.976***	7.272***	7.975***
	(0.727)	(0.527)	(0.693)	(0.389)	(0.393)
Log Equiv Net HH Inc	0.292***	0.283***	0.125**	0.201***	0.197***
	(0.071)	(0.052)	(0.055)	(0.037)	(0.039)
IHS Equiv Net HH Wealth	0.014***	0.021***	0.029***	0.012***	0.003
	(0.005)	(0.004)	(0.008)	(0.003)	(0.004)
Relative Income (low)	-0.241***	-0.139**	-0.117	0.017	0.081
	(0.093)	(0.070)	(0.096)	(0.048)	(0.069)
Relative Income (high)	0.014	0.009	0.004	0.068	0.081
	(0.110)	(0.084)	(0.039)	(0.060)	(0.061)
Relative Wealth (low)	-0.189***	-0.172***	-0.022	-0.076**	-0.105**
	(0.068)	(0.049)	(0.032)	(0.036)	(0.048)
Relative Wealth (high)	-0.035	-0.039	-0.015	-0.014	0.007
	(0.069)	(0.051)	(0.019)	(0.034)	(0.042)
Socializing	-0.373***	-0.317***	-0.134***	-0.178***	-0.116***
	(0.030)	(0.028)	(0.043)	(0.019)	(0.021)
Education (years)	0.062***	0.035***	0.007	0.002	-0.009
	(0.011)	(0.008)	(0.005)	(0.005)	(0.006)
Not Commuting	0.059	0.036	0.026	0.038	0.071*
	(0.058)	(0.046)	(0.020)	(0.033)	(0.038)
Subjective Health	-1.173***	-1.063***	-0.918***	-0.638***	-0.562***
	(0.026)	(0.023)	(0.028)	(0.020)	(0.018)
Employed	-0.018	-0.107**	-0.054**	-0.123***	-0.123***
	(0.068)	(0.052)	(0.027)	(0.035)	(0.044)
Age	0.015***	0.014***	0.006**	0.012***	0.012***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Age Squared	-0.000***	-0.000***	-0.000**	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Male Dummy	0.027	0.006	-0.022	-0.111***	-0.146***
	(0.048)	(0.036)	(0.017)	(0.026)	(0.032)
Married	0.271***	0.152***	0.059**	0.088***	0.043
	(0.057)	(0.048)	(0.029)	(0.030)	(0.036)
Observations	14908	14908	14908	14908	14908
Pseudo R^2	0.210	0.213	0.169	0.072	0.058
Note:	*p<0.1; **p<0.05; ***p<0.01				

Table 1: Quantile Regressions of Life Satisfaction

The quantile regression is estimated on a sample of 14908 observations. Inference is based on bootstrap standard errors with 999 replications, which is standard in quantile regression due to the unknown conditional error density. Model fit, measured by the pseudo R^2, is highest at the lower quantiles (around 0.21 for $q = 0.1$ and $q = 0.25$) and declines monotonically towards the upper tail. This indicates that observed socioeconomic characteristics explain substantially more variation among individuals with lower happiness than among those who are already highly satisfied.

Log equivalized net household income is positive and highly significant across all quantiles. The coefficient decreases modestly from about 0.29 at $q = 0.1$ to 0.2 at $q = 0.9$ indicating that additional income is always positively associated with happiness (aligning with, e.g., Killingsworth et al. (2023)) and especially so for those who are unhappiest. Since income enters the specification in logarithmic form, the corresponding coefficient can be interpreted as a semi-elasticity. This means that, for instance, a 10% change in income, $\Delta \log(\text{Income}) = \log(1.10) \approx 0.095$ for someone in the first happiness decile with coefficient $\beta = 0.292$ translates to a gain in predicted happiness of roughly $0.095 \cdot 0.292 \approx 0.028$ points. The economic reason is likely that additional income enables individuals to satisfy immediate needs boosting happiness of those who are relatively unhappy more strongly than those who are relatively happier.

The coefficient on IHS equivalent net household wealth is positive and statistically significant across most quantiles, but economically modest in magnitude. For example, at the median ($q = 0.5$), the coefficient equals 0.029, implying that a one-unit increase in the inverse hyperbolic sine of household wealth raises median happiness by 0.029 points. Since the IHS transformation behaves approximately like a logarithm for positive wealth values, a doubling of wealth corresponds roughly to an increase of $\log(2) \approx 0.693$ in the transformed variable. This implies that doubling household wealth increases median happiness by approximately $0.029 \cdot 0.693 \approx 0.02$ points on the 0-10 scale. The effect is slightly smaller at the lower quantile (0.014 at $q = 0.1$) and becomes economically negligible at the upper tail (0.003 at $q = 0.9$ and statistically insignificant). Overall, absolute wealth exerts a positive but quantitatively small effect on happiness, especially when controlling for income and relative wealth position (cf. Section 5).

Concentrating on the relative income and wealth measures a clear pattern emerges, when individuals have relatively low earnings (60% of the median) compared to their peers, their happiness significantly drops when they belong to relatively unhappy individuals (quantiles 0.1 and 0.25) hinting at envy. However, when individuals are earning relatively high incomes (2 times the median) compared to their peers, it does not affect individuals across all quantiles (effect nearly 0 and insignificant). Interestingly, wealth follows the same pattern. If an individual is relatively poor (same definition as before only for wealth), happiness drops significantly across all but one quantile with the highest decrease among the unhappiest hinting, again, at a negative comparison effect (e.g., Luttmer (2005)), while those who are relatively rich are affected slightly, but always insignificantly, negative.

Reduced social interaction is associated with lower happiness throughout the distribution. The negative effect of socializing is strongest at the lower quantiles (-0.37 at $q = 0.1$) and weakens substantially towards the upper tail (-0.1 at $q = 0.9$), highlighting a potential protective or supportive role of social ties for particularly unhappy individuals.

Education has only small effects once income, employment and health are controlled for[14]. It is significantly positive at the lower tail and becomes negative at higher quantiles, suggesting limited (direct) happiness returns to additional schooling among already highly satisfied individuals, but a significant positive effect of more schooling for relatively unhappy individuals.

Not Commuting is positively associated with happiness across all quantiles, with the strongest effects in the upper tail. Thus, having the comfort of not sitting in traffic or overcrowded trains and/or even working from home is positively related to happiness of individuals.

Subjective health exhibits the largest coefficients in absolute value. A one-category deterioration in self-rated health is associated with a decline of more than one point in happiness at the lower quantiles (-1.15 at $q = 0.1$) and about half a point at the upper quantiles (-0.5 at $q = 0.9$). This pattern indicates that happiness is strongly health-constrained, particularly

[14] Several studies find that much of education's association with happiness operates through higher earnings and improved employment (income and employment channels), so once those channels are controlled for the remaining direct effect is usually small or statistically weak (cf. Bertermann et al. (2023) or Araki (2022)).

among less-satisfied individuals.

Being employed is negatively associated with happiness. While insignificant at lower quantiles, it becomes significantly negative at the upper tail, possibly reflecting stress or time constraints which highly satisfied individuals dislike.

Age follows a slightly concave pattern, with a positive term throughout and a negative squared term. Men report slightly lower happiness than women from the median upward, which is consistent with several empirical findings showing that, on average, women frequently report equal (Zweig, 2015, Table 8) or somewhat higher happiness than men (Montgomery, 2022, e.g., Table 1). Nonetheless, caution is warranted as the observed difference may stem from systematic gender variation in the interpretation or use of the response scale (Montgomery, 2022, p. 66). The coefficient on the married indicator is positive across the distribution and statistically significant at most quantiles, with particularly strong effects in the lower tail. This implies that among relatively unhappy people, married individuals report nearly 0.27 higher happiness on the 0-10 scale compared to otherwise similar non-married individuals.

3.3 Distributional Implications

Overall, the quantile regression results reveal pronounced heterogeneity for a variety of factors for different parts of the happiness distribution. These heterogeneous effects of factors examined so far allow us to consider happiness inequality in a more intuitive manner. By identifying which variables exert stronger or weaker influences across different parts of the distribution, we can better understand how they shape overall inequality in happiness. Figure 3 depicts the coefficients from the quantile regression (cf. Table 1) along the happiness distribution:

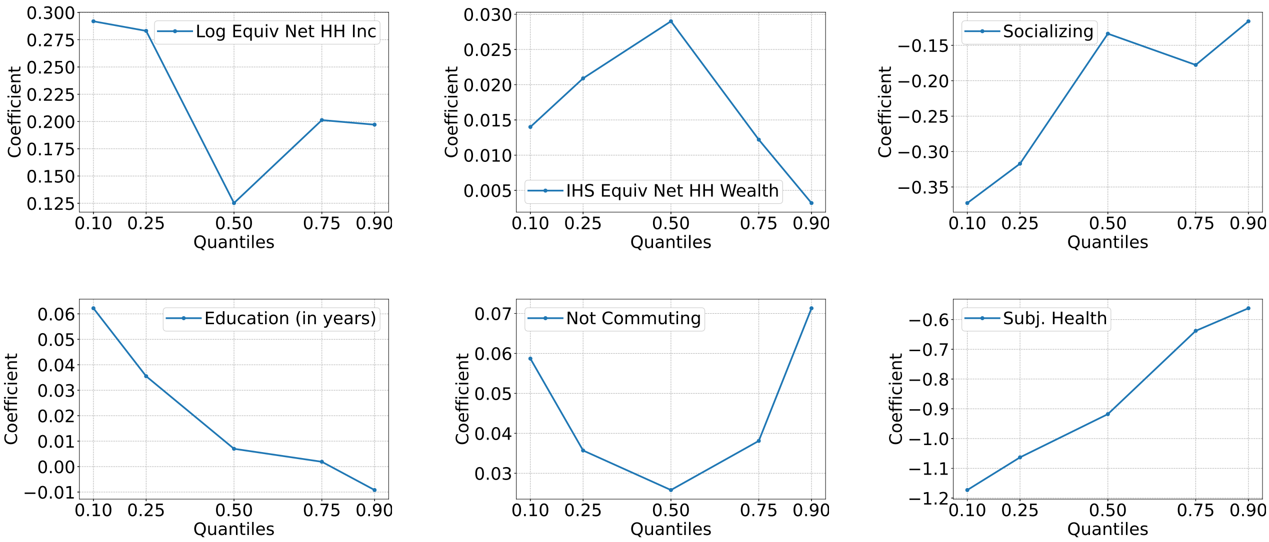

Figure 3: A selection of estimated coefficients across the happiness distribution

From a distributional perspective, income is positively associated with happiness. The impact is larger in the lower tail than at the median, and somewhat higher again in the upper tail. This suggests that an increase in income affects the happiness distribution more strongly from the lower tail "pushing" more individuals towards the center and thereby reducing happiness inequality. Wealth tends to benefit individuals who are relatively less happy more than those at the higher end of the happiness distribution. This suggests that increases in wealth have a stronger marginal effect on the unhappier segment of the population, likely reducing inequality from below. A similar but even more pronounced pattern emerges for education: Individuals at the lower end of the distribution gain significantly from an additional year of schooling, whereas those who are already relatively happy experience little benefit. These patterns suggest that changes in education may exert inequality-reducing pressure, especially if gains are concentrated among individuals in the lower part of the happiness distribution. This finding is also reported, and is particularly pronounced, in Becchetti et al. (2014, p. 429), although education is measured using ISCED categories[15].

Health and social interactions follow a comparable pattern. Poor health or reduced socializing disproportionately affects unhappier individuals, potentially increasing the gap between unhappy and happy individuals and, thus, happiness inequality. Conversely, improvements in health or increased social engagement yield the largest gains for the lower end of the happiness spectrum, thereby decreasing inequality. Avoiding a commute has a positive effect on happiness across all shown quantiles. The effect increases with the level of happiness and is strongest at the upper quantiles, indicating that avoiding a commute may tend to push happiness towards its upper bound and possibly bunch happiness scores there.

4 Happiness Inequality

This section turns from quantile-specific associations to the paper's main object of interest: Happiness inequality. I first document observed inequality in the balanced sample (cf. Section 4.1), then construct a counterfactual 2017 happiness distribution (cf. Section 4.2), decompose the resulting inequality changes (cf. Section 4.3) and finally test whether quantile-specific coefficients changed between 2012 and 2017 (cf. Section 4.4).

4.1 Observed Inequality

While these distributional patterns are instructive and often suggest clear welfare policy targets (e.g., based on my estimated pattern improving the health of the least happy sharply reduces happiness inequality), a more formal quantification of inequality is required to assess their overall impact when realizations change. To this end, I examine summary measures of dispersion, specifically the Gini coefficient and the CV (cf. Becchetti et al. (2014)). These measures provide a formal characterization of inequality allowing us to connect these heterogeneous effects identified in quantile regressions to broader distributional outcomes. Concentrating on a balanced sample[16] concerning the reference year 2017 and the base year 2012, Table 2 presents the Gini coefficient[17] and CV with the corresponding distributions in Figure 4 next to it.

[15]The ISCED categories come from the UNESCO framework called the International Standard Classification of Education, which is used worldwide to classify education levels in a consistent way.

[16]This means that I restrict myself to individuals with complete covariate information observed in both 2012 and 2017. This avoids composition bias and ensures that the counterfactual and actual distributions refer to the same set of agents ($n = 9325$).

[17]The Gini coefficient G can be defined in the classical form as $G = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}|x_i - x_j|}{2n^2\bar{x}}$ or equivalently using the rank-based formula $G = \frac{2\sum_{i=1}^{n} i\, x_{(i)}}{n\sum_{i=1}^{n} x_i} - \frac{n+1}{n}$, where $x_{(i)}$ denotes the i-th smallest observation (Dorfman, 1979; Lerman and Yitzhaki, 1984; Neumann and Scheuer, 2024). The values of the Gini coefficients for 2012 and 2017 appear to be quite accurate when comparing with previous studies, for instance, Becchetti et al. (2014, p. 8) find 0.146 for 2007.

Year	Gini	CV
2012	0.1280	0.2389
2017	0.1267	0.2406

Table 2: Gini Coefficient and CV of balanced sample

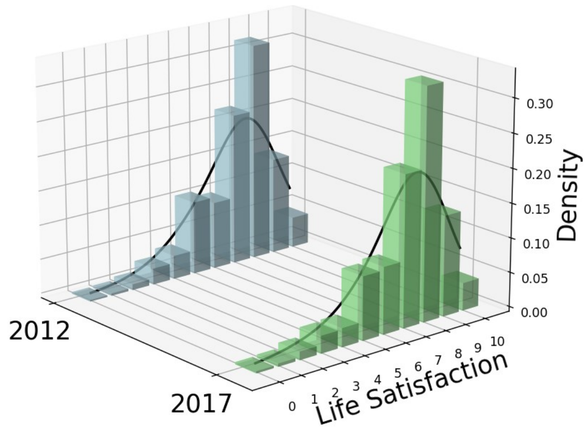

Figure 4: Happiness distributions for 2012 and 2017 of balanced sample

These measures reflect the left-skewed, right-steep distributions observable in Figure 4, highlighting a high degree of happiness concentration at the top, thereby implying a relatively low degree of dispersion and, thus, happiness inequality. This raises two important questions: To what extent can comprehensive models, such as Equation (2), explain observed inequality for 2012 and 2017 and how does that differ across alternative inequality measures?

4.2 Counterfactual Happiness

To address these questions, I derive a counterfactual happiness distribution using a variation of the method from Machado and Mata (2005), which was formalized by Melly (2005) and summarized in Fortin et al. (2011, p. 61). Machado and Mata (2005) determine counterfactual wage distributions and try to disentangle the effect of a change in coefficients (wage returns from schooling, tenure, age, gender, et cetera) and a change in covariates (years of schooling, years of tenure et cetera) on wage inequality. Their method is based on the estimation of the marginal density function of wages in a given year implied by counterfactual distributions of some or all the observed attributes. I focus on their algorithm (Machado and Mata, 2005, p. 449) and slightly adapt it to my purpose. Specifically, I alter Step 1, 3 and 4. Firstly, I do not draw a random number m of quantiles (Step 1) but, similar to Albrecht et al. (2003), focus on the first percentile, the second percentile, and so on up to the 99$^{\text{th}}$ percentile. Then I do not bootstrap the covariates once for every quantile creating a number m of covariate matrices (Step 3) which they then apply the coefficients from the quantile regression of the reference year to, thereby creating a (marginal) wage distribution (Step 4), but rather apply my entire covariate sample (not a bootstrapped subsample) from my reference year to every quantile's coefficient vector from my base year to create the counterfactual happiness distribution. Further applications of this approach include Gosling et al. (2000) but more importantly Albrecht et al. (2003) and Heinze (2010). Concerning the latter two, I share Step 1 with them but differ at Step 3 in that Albrecht et al. (2003) draw 100 times from their covariate matrix at random for each quantile and Heinze (2010) does so 10000 times. When focusing on econometric theory, my approach is comparable to Type 2 of estimating counterfactual effects from (Chernozhukov et al., 2013) with the change in the covariate distribution being the change from 2012 to 2017.

Eventually, I estimate Equation (2) for 2012 on a tighter grid of 99 quantiles[18] for a balanced panel, collect the estimated parameters and use these structural effects from 2012 to simulate

[18]Using the interior point optimization algorithm, predicted quantiles are stable and, concerning quantile crossings, a standard rearrangement step ensures valid conditional distributions (cf. Chernozhukov et al. (2010) and Appendix A.2.3 for details). Also, given the number of observations and a more precise description of the distribution, i.e., switching from 5 quantiles for illustration to 99, appears useful and logical.

a counterfactual happiness distribution for 2017. I simulate[19]

$$\widehat{Q}_q(\text{LifeSa}_i \mid \tilde{\mathbf{X}}_i^{2017}) = \tilde{\mathbf{X}}_i^{2017\top}\widehat{\boldsymbol{\beta}}^{2012}(q), \quad \text{for } q = 0.01, \ldots, 0.99 \tag{3}$$

where I maintain my balanced panel assumption[20] for the simulation analysis. This implies that each entry of the counterfactual distribution represents a "synthetic" happiness value at every given quantile using its 2012-based coefficients for 2017 covariates with 99 quantiles per person[21], leading to a total of 923175 pseudo-observations[22] ($9325 \cdot 99$). Intuitively, this procedure shows what the predicted 2017 happiness of the individuals would be at every point of the conditional distribution, had the structural relationship from 2012 remained unchanged. I then calculate the Gini and CV for this counterfactual distribution and compare them to the actual 2017 values (given my balanced sample).

The resulting inequality measures are reported in Table 3.

Year	Gini	CV
2012	0.1280	0.2389
2017	0.1267	0.2406
2017 (cf)	0.1113	0.2362

Table 3: Gini Coefficient and CV of balanced sample and counterfactual (cf)

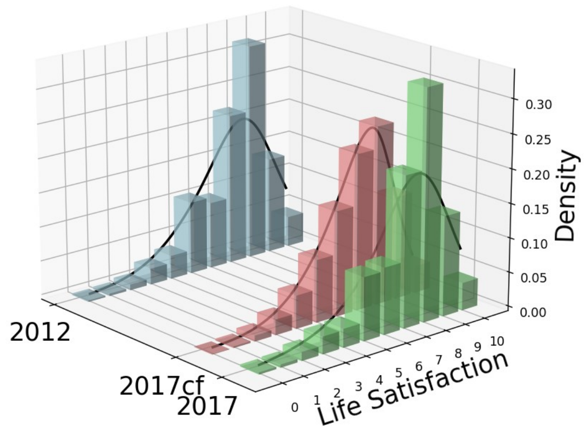

Figure 5: Happiness distributions for 2012 and 2017 – Balanced sample and counterfactual (cf)

If Germans in the balanced panel retained their 2017 characteristics but the 2012 happiness-formation process, then happiness inequality in 2017 would be 0.1113 (Gini) and 0.2362 (CV). Thus, the Machado-Mata counterfactual yields a substantially lower Gini coefficient[23] (also implied by Figure 5 with a steeper curve for 2017cf) and a slightly lower CV than the actual distribution. This, to some extent, is actually expected as the Gini is sensitive to changes in the shape of the distribution, especially in the middle (Davies et al., 2017; Neumann and Scheuer,

[19]Note that $\tilde{\mathbf{X}}_i = (1, \mathbf{X}_i)$. The remaining notation is based on Fortin et al. (2011). For more details regarding Equation (3) please see Appendix A.2.

[20]A slight caveat as individuals, who remain in panel surveys, tend to have more stable socioeconomic characteristics (Fitzgerald et al., 1998), which would imply that the sample does not cover the entire bandwidth of socioeconomic characteristics. However, this limitation is inherent in the available data.

[21]Put differently, each individual is characterized by one observed covariate vector in 2017, $\mathbf{X}_i^{2017}$, but the quantile-regression model maps this vector into different predicted outcomes depending on the quantile-specific coefficient vector $\hat{\boldsymbol{\beta}}^{2012}(q)$. Hence, for each individual, I obtain 99 predicted life-satisfaction values, one for each quantile $q \in 0.01, \ldots, 0.99$. These values should not be interpreted as repeated observations for the same individual, but as points of the estimated conditional distribution implied by the 2012 coefficient structure and the individual's 2017 covariates. Hence, I get one possible happiness distribution per person.

[22]These pseudo-observations are not interpreted as additional independent survey observations. They are model-based predicted values used to approximate the counterfactual marginal distribution implied by the estimated 2012 quantile-regression coefficients and the observed 2017 covariates.

[23]Please note that, although the counterfactual Gini is lower, the Lorenz curves cross. Hence, the counterfactual distribution is not Lorenz-dominant.

2024), while the CV reflects overall dispersion. Hence, the counterfactual distribution features a similar variance but redistributes mass much more evenly across the middle quantiles, reducing the Gini.

4.3 Decomposing Happiness Inequality

This result allows us to differentiate between a mechanical as opposed to a residual (idiosyncratic) effect over time. The *mechanical* effect represents the change in inequality predicted solely from changes in observable characteristics (that is, from changes in the realizations of the exogenous variables from 2012 to 2017) while the *residual* effect captures idiosyncratic, unobserved or non-systematic variation in happiness that is not accounted for by the model[24]. The former can be mathematically defined for inequality measure I as

$$I^{\text{mech}} \equiv I\left(\text{LifeSa}^{2017,\text{cf}}\right) \; - \; I(\text{LifeSa}^{2012}), \tag{4}$$

where this component reflects how the distribution of happiness would evolve if individuals' happiness responded exactly as specified by the estimated model, i.e., based on Equation (3).

The total observed change in inequality measure I is defined as

$$I^{\text{tot}} \equiv I(\text{LifeSa}^{2017}) - I(\text{LifeSa}^{2012}), \tag{5}$$

which implies the portion not explained by the model-generated projection, i.e. the residual component, to be the difference between total and mechanical change, i.e.,

$$I^{\text{res}} \equiv I^{\text{tot}} \; - \; I^{\text{mech}} \; = \; I(\text{LifeSa}^{2017}) - I\left(\text{LifeSa}^{2017,\text{cf}}\right). \tag{6}$$

Computing Equations (4)-(6), I get

Year	Gini	CV
I^{tot}	-0.0013	0.0015
I^{mech}	-0.0177	-0.0014
I^{res}	0.0164	0.0029

Table 4: Mechanical vs. Residual Effect

The decomposition reveals a striking pattern. Observable characteristics evolved between 2012 and 2017 in a way that, according to the 2012 structure, should have led to a substantial reduction, approximately 14%, in happiness inequality (mechanical effect: -0.018 for the Gini). However, the actual change in inequality is almost zero (total effect: –0.0013). This implies that unobserved factors or changes in the happiness-production process (shifts in the returns to characteristics) almost completely offset the predicted equalizing effect and results in a relatively large positive residual component (residual effect: +0.0164). In contrast, the CV decomposition is small throughout, which is expected given the bounded nature of happiness scores and the CV's insensitivity to changes in the middle of the distribution (Davies et al., 2017). What is reassuring, however, is that the mechanical effect for the CV points into the same direction and would have also predicted a decline in inequality by about 1%. In conclusion, if happiness were produced in 2017 the same way as in 2012, inequality measured by the Gini (CV) would be much (slightly) lower. The fact that it is not indicates that the residual effect, which apparently counteracts these equalizing tendencies, may stem from a change in the coefficients (the returns

[24]This includes unobserved heterogeneity, psychological adaptation, unexpected life events, shifts in aspirations, measurement noise and all individual-specific variation not captured by the observable covariates.

to happiness from the covariates) or other unobserved factors such as omitted variables, model misspecification or idiosyncratic shocks unaccounted for by the model. Concentrating on the former, the analysis next examines whether the quantile-specific coefficients changed between 2012 and 2017.

4.4 Changes in Quantile-specific Returns

Hence, in a final step, I examine whether the happiness-production process, i.e., the association between covariates and happiness (i.e., the regression coefficients), changed between 2012 and 2017 by estimating a pooled quantile regression with interactions between the 2017 indicator and the main covariates of interest. Specifically, I estimate

$$Q_q(\text{LifeSa}_i^s \mid \mathbf{X}_i^s, D_i^s) = \alpha_q + \mathbf{X}_i^{s\top}\boldsymbol{\beta}_q + D_i^s\gamma_q + (\mathbf{X}_i^s \cdot D_i^s)^\top\boldsymbol{\delta}_q, \tag{7}$$

where D_i^s is a dummy equal to 1 in $s = 2017$ and 0 in 2012. In this specification, $\boldsymbol{\beta}_q$ captures the vector of 2012 quantile-specific coefficients, while the vector $\boldsymbol{\delta}_q$ measures the change in these coefficients between 2012 and 2017. Consequently, the 2017 coefficient vector equals $\boldsymbol{\beta}_q+\boldsymbol{\delta}_q$. To assess whether returns to observed characteristics changed over time, I test the null hypothesis $H_0 : \boldsymbol{\delta}_q = \mathbf{0}$ for each quantile $q \in [0.05, 0.95]$[25]. Statistical inference is conducted using a cluster bootstrap at the individual level. Since the sample consists of a balanced panel in which the same individuals are observed in both waves, clustering accounts for within-individual dependence across time.

In principle, Equation (7) allows all covariates to interact with the time indicator. However, including a large number of interaction terms, particularly for binary indicators, induces substantial collinearity in the quantile regression design matrix. In finite samples, especially in the tails of the conditional distribution where effective sample sizes are smaller, this leads to numerical instability of the optimization routine and unreliable bootstrap inference. To ensure stable estimation of the quantile process while maintaining interpretability, I restrict the interaction terms to the most important predictors of happiness (based on coefficient level and significance) according to Table 1: Income, wealth, subjective health, socializing and education. All remaining variables enter the specification without interactions with the year indicator.

Figure 6 reports the estimated coefficient differences $\widehat{\boldsymbol{\delta}}_q$ from Equation (7) across the conditional distribution of happiness.

[25]Computationally cheaper and more feasible for the interior point solver, I use 95 quantiles.

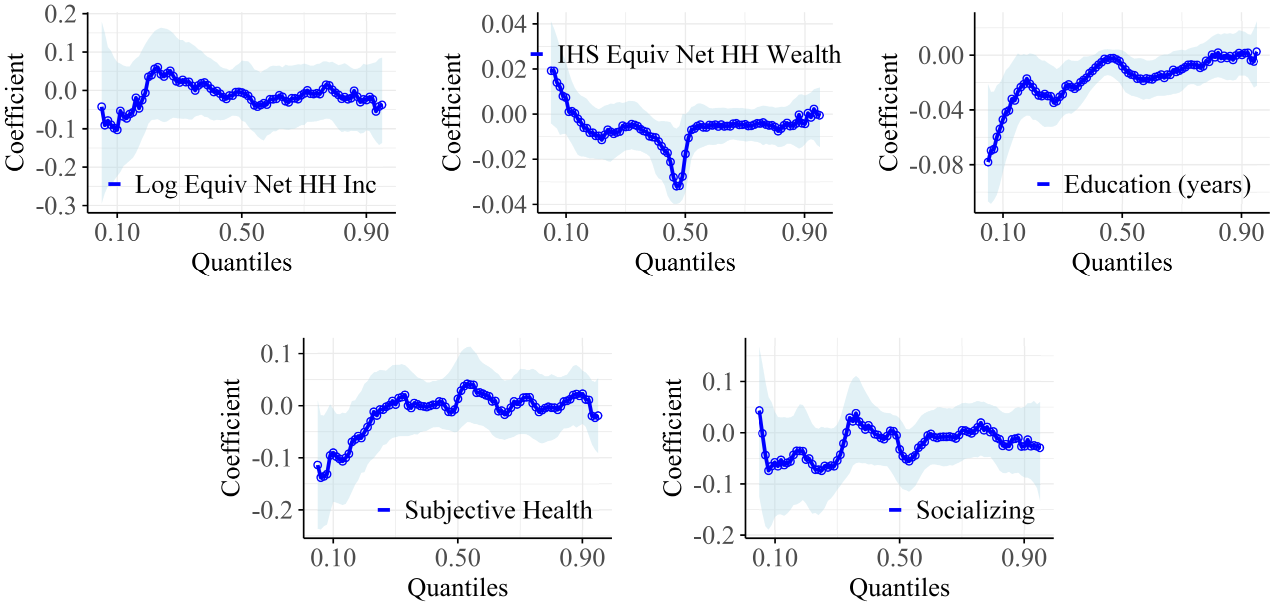

Figure 6: Estimated coefficient changes. Shaded areas depict 95% confidence intervals obtained from a cluster bootstrap at the individual level.

For each quantile $q \in [0.05, 0.95]$, $\widehat{\delta}_k(q)$ represents the estimated change in the marginal association between covariate X_k and happiness from 2012 to 2017, such that $\widehat{\beta}_{2017,k}(q) = \widehat{\beta}_{2012,k}(q) + \widehat{\delta}_k(q)$. The solid line traces $\widehat{\delta}_k(q)$ over 95 quantiles, while the shaded region depicts 95% confidence intervals (obtained from a cluster bootstrap at the individual level). If the confidence band excludes zero at a given quantile, the null hypothesis $H_0 : \delta_k(q) = 0$ is rejected, indicating a statistically significant change in the conditional quantile effect at that point of the distribution.

I find statistically significant coefficient changes only in limited parts of the distribution, most notably for subjective health at lower quantiles and, to a lesser extent, for education and wealth in selected regions of the distribution. For most covariates and quantiles, the estimated changes are statistically indistinguishable from zero. Inspecting the point estimates, income, subjective health (outside the lower tail), social interaction and wealth oscillate around zero, while the education gradient tends to decline in 2017 up to approximately the 80^{th} percentile.

Overall, these findings suggest that although some quantile-specific coefficients changed between 2012 and 2017, the magnitude and distributional pattern of these changes are limited. In particular, the observed coefficient shifts appear not sufficiently strong or pervasive to offset the inequality-reducing effect stemming from changes in observable characteristics. Consequently, the relative stability of happiness inequality over time appears to be driven primarily by unobserved forces, such as idiosyncratic shocks, or model misspecifications.

5 Robustness Checks

I conduct robustness checks which mainly focus on the treatment of wealth in this study and alternatives to the CV.

5.1 Alternative Wealth Measures

As a robustness check, I examine whether the baseline results are sensitive to alternative treatments of wealth. First, I replace the IHS-transformed wealth measure with a set of wealth-category indicators. Second, I use a dummy variable for positive net wealth. Third, I compare specifications with absolute wealth only to specifications that also include relative wealth positions.

5.1.1 Wealth in Categories

Using wealth in categories, I discretize equivalized net household wealth into seven mutually exclusive categories and construct a corresponding set of indicator variables $\mathbf{1}_{\{\text{Wealth category } k\}_i}$, where each indicator equals one if individual i belongs to category k and zero otherwise. The categories are defined as follows:

Category	Substantial debt	Small debt	Zero	Low	Middle	Upper middle	High
Wealth range in €	< -10000	$[-10000, 0)$	0	$(0, 50000]$	$(50000, 200000]$	$(200000, 500000]$	> 500000

Table 5: Wealth categories

In regression analyses, the coefficients on these dummy variables represent the difference in the specified happiness quantile relative to households in a reference category, i.e., "Substantial debt". For example, a positive coefficient on the "Low" wealth dummy indicates that households with wealth between € 0 and € 50000 have higher expected happiness at this specific quantile compared to households with substantial debt, holding other factors constant. The results are reported in Table 8. Reassuringly, the main predictors of happiness remain broadly stable. The wealth categories, however, reveal an unstable, non-monotonic coefficient pattern, also sometimes switching signs, when comparing before and after median quantiles. Thus, wealth in categories allows a more nuanced interpretation of its impact across the happiness distribution, but, due to the non-monotonic pattern, my baseline approach appears to be a better fit.

5.1.2 Wealth via Dummy

Using a simple yes/no dummy concerning whether a household's equivalized net wealth is positive, I obtain the results in Table 9. The results again show a similar pattern across the most significant predictors of happiness. Similar to when wealth is included in categories, this specification displays instability around the median. Overall, this suggests that the baseline with IHS-transformed wealth as well as comparison-with-peers parameters appears to be the most appropriate model choice.

5.1.3 Absolute vs. Positional Wealth

Concentrating on the impact of (net equivalized household) wealth on happiness across quantiles, I examine whether the impact of wealth stems from including it in absolute terms via an IHS transformation or whether its effect is driven through the comparison aspect, i.e., the relative wealth dummy variables. Running the quantile regression solely with IHS (see Table 10) and then as in Table 1 shows that, once the positional arguments are included, the coefficient of IHS transformed equivalized net household wealth drops sharply implying that the variation is absorbed by the positional arguments once they are included. This points towards a more important role of wealth through comparison with peers.

5.2 Absolute Dispersion

As an additional robustness check for the inequality analysis in Section 4, I consider variance and standard deviation as alternatives to the CV. In the spirit of Clark et al. (2016), I focus on the variance/standard deviation as the measure of dispersion instead of the CV as examining the raw variance (or standard deviation) of life satisfaction isolates "pure" dispersion, without normalization by a central moment. Since the CV can remain stable when both the mean and the standard deviation move proportionally, reporting the latter separately serves as an additional robustness check that captures changes in the overall spread of the distribution. This helps determine whether the counterfactual distribution differs from the actual one in terms of unadjusted variation rather than relative or mean-adjusted dispersion. Hence, I consider

Year	Variance	Standard deviation
2012	2.9151	1.7073
2017	2.9788	1.7259
2017 (estim.)	2.8274	1.6815

Table 6: Variance and standard deviation

First, one sees that variance and standard deviation are going up in reality but are decreasing when focusing on my projection, i.e., the counterfactual distribution. Thus, the result in Table 6 strongly suggests that the projection of a lower, mean-based CV for 2017 may be driven by a strong projected decrease in the variance, where it increases in the observed data. Second, Table 6 shows, similar to the CV, that the mechanical effect would have predicted lower absolute dispersion whereas in reality dispersion increased. Parallel to previous explanations, the mechanical effect is complemented by a residual which counteracts and drives variance and standard deviation upwards.

6 Conclusion

This paper studies how socioeconomic characteristics shape the distribution of happiness in Germany and how these distributional effects translate into happiness inequality. Using SOEP data and quantile regressions, I show that several determinants of happiness, especially income, subjective health and socializing, differ substantially across the happiness distribution. These differences are most pronounced at the lower end of the distribution, suggesting that mean regressions miss important heterogeneity relevant for understanding happiness inequality.

The counterfactual analysis then shows that observed happiness inequality remained broadly stable between 2012 and 2017, although changes in observable characteristics would have predicted a decline under the 2012 happiness-production structure. This suggests that the equalizing effect of changing observables was offset by an unexplained component. Tests of quantile-specific coefficient changes provide little evidence that systematic changes in the returns to observed characteristics account for this offset. The remaining gap is therefore consistent with unobserved shocks, omitted variables, measurement variation or other forms of individual-specific heterogeneity not captured by the model.

These findings imply that stable aggregate happiness inequality can mask relevant distributional dynamics. Improving observable socioeconomic conditions may exert equalizing pressure, but such changes need not translate into lower observed happiness inequality when unobserved factors move in the opposite direction. The results should therefore be interpreted as a descriptive counterfactual rather than a causal policy simulation. Future research should investigate more directly which unobserved factors sustain happiness inequality over time.

References

Aaberge, R., S. Bjerve, and K. Doksum (2005): "Decomposition of Rank-Dependent Measures of Inequality by Subgroups," Tech. Rep. 437, Statistics Norway, Research Department, Oslo.

Albrecht, J., A. Björklund, and S. Vroman (2003): "Is There a Glass Ceiling in Sweden?" *Journal of Labor Economics*, 21, 145–177.

Araki, S. (2022): "Does Education Make People Happy? Spotlighting the Overlooked Societal Condition," *Journal of Happiness Studies*, 23, 587–629.

Becchetti, L., E. Giachin Ricca, and A. Pelloni (2012): "The Relationship Between Social Leisure and Life Satisfaction: Causality and Policy Implications," *Social Indicators Research*, 108, 453–490.

Becchetti, L., R. Massari, and P. Naticchioni (2014): "The drivers of happiness inequality: suggestions for promoting social cohesion," *Oxford Economic Papers*, 66, 419–442.

Bellemare, M. F. and C. J. Wichman (2020): "Elasticities and the Inverse Hyperbolic Sine Transformation," *Oxford Bulletin of Economics and Statistics*, 82, 50–61.

Bérenger, V. and J. Silber (2022): "On the Measurement of Happiness and of its Inequality," *Journal of Happiness Studies*, 23, 861–902.

Bertermann, A., D. A. Kamhöfer, and H. Schildberg-Hörisch (2023): "More Education Does Make You Happier – Unless You Are Unemployed," SOEPpapers on Multidisciplinary Panel Data Research 1192, Deutsches Institut für Wirtschaftsforschung (DIW Berlin), Berlin.

Binder, M. (2016): "Revisiting Cheerful Jane and Miserable John: the impact of income, good health, social contacts and education declines with increasing subjective well-being," *Applied Economics Letters*, 23, 544–553.

Binder, M. and A. Coad (2011): "From Average Joe's happiness to Miserable Jane and Cheerful John: using quantile regressions to analyze the full subjective well-being distribution," *Journal of Economic Behavior & Organization*, 79, 275–290.

——— (2015): "Heterogeneity in the Relationship Between Unemployment and Subjective Wellbeing: A Quantile Approach," *Economica*, 82, 865–891.

Blanchflower, D. G. and A. J. Oswald (2008): "Is well-being U-shaped over the life cycle?" *Social Science & Medicine*, 66, 1733–1749.

Böhnen, C. and T. Kuhnimhof (2024): "Working from home and commuter travel in germany – panel data analysis of long-term effects," *Transportation Research Part A: Policy and Practice*, 190, 104257.

Brooks, A. C. (2008): *Gross National Happiness: Why Happiness Matters for America—and How We Can Get More of It*, New York: Basic Books.

Brown, S. and D. Gray (2016): "Household finances and well-being in Australia: An empirical analysis of comparison effects," *Journal of Economic Psychology*, 53, 17–36.

Brulé, G. and C. Suter (2019): "Why wealth matters more than income for subjective well-being?" in *Wealth (s) and subjective well-being*, Springer, 1–13.

Bruni, L. and L. Stanca (2008): "Watching alone: Relational goods, television and happiness," *Journal of Economic Behavior & Organization*, 65, 506–528.

Buchinsky, M. (1995): "Estimating the asymptotic covariance matrix for quantile regression models a Monte Carlo study," *Journal of Econometrics*, 68, 303–338.

Buhmann, B., L. Rainwater, G. Schmaus, and T. M. Smeeding (1988): "EQUIVALENCE SCALES, WELL-BEING, INEQUALITY, AND POVERTY: SENSITIVITY ESTIMATES ACROSS TEN COUNTRIES USING THE LUXEMBOURG INCOME STUDY (LIS) DATABASE," *Review of Income and Wealth*, 34, 115–142.

Burbidge, J. B., L. Magee, and A. L. Robb (1988): "Alternative Transformations to Handle Extreme Values of the Dependent Variable," *Journal of the American Statistical Association*, 83, 123–127.

Chen, H., S. Voigt, and X. Fu (2021): "Data-Driven Analysis on Inter-City Commuting Decisions in Germany," *Sustainability*, 13.

Chernozhukov, V., I. Fernández-Val, and A. Galichon (2010): "Quantile and Probability Curves Without Crossing," *Econometrica*, 78, 1093–1125.

Chernozhukov, V., I. Fernández-Val, and B. Melly (2013): "Inference on Counterfactual Distributions," *Econometrica*, 81, 2205–2268.

Clark, A. E., S. Flèche, and C. Senik (2014): "The Great Happiness Moderation," in *Happiness and Economic Growth: Lessons from Developing Countries*, ed. by A. E. Clark and C. Senik, Oxford: Oxford University Press, 32–139.

——— (2016): "Economic Growth Evens Out Happiness: Evidence from Six Surveys," *Review of Income and Wealth*, 62, 405–419.

Clark, A. E., P. Frijters, and M. A. Shields (2008): "Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles," *Journal of Economic Literature*, 46, 95–144.

Clark, A. E. and A. J. Oswald (1994): "Unhappiness and Unemployment," *The Economic Journal*, 104, 648–659.

Cole, H. L., G. J. Mailath, and A. Postlewaite (1992): "Social Norms, Savings Behavior, and Growth," *Journal of Political Economy*, 100, 1092–1125.

D'Ambrosio, C., J. R. Frick, and M. Jäntti (2009): "Satisfaction with Life and Economic Well-Being: Evidence from Germany," *Journal of Contextual Economics – Schmollers Jahrbuch*, 129, 283–295, first published online: October 3, 2017.

D'Ambrosio, C., A. Lepinteur, and M. Jäntti (2020): "Money and Happiness: Income, Wealth and Subjective Well-Being," *Social Indicators Research*, 148, 47–66.

Davies, J. B., N. M. Fortin, and T. Lemieux (2017): "Wealth inequality: Theory, measurement and decomposition," *Canadian Journal of Economics/Revue canadienne d'économique*, 50, 1224–1261.

Deaton, A. and C. Paxson (1994): "Intertemporal Choice and Inequality," *Journal of Political Economy*, 102, 437–467.

Di Tella, R., R. J. MacCulloch, and A. J. Oswald (2001): "Preferences over Inflation and Unemployment: Evidence from Surveys of Happiness," *American Economic Review*, 91, 335–341.

Diener, E. (2000): "Subjective well-being: The science of happiness and a proposal for a national index," *American Psychologist*, 55, 34–43.

Dorfman, R. (1979): "A Formula for the Gini Coefficient," *The Review of Economics and Statistics*, 61, 146–149.

Durst, A. K. (2021): "Education as a Positional Good? Evidence from the German Socio-Economic Panel," *Social Indicators Research*, 155, 745–767.

Easterlin, R. A. (2005): "Building a Better Theory of Well-Being," in *Economics and Happiness: Framing the Analysis*, ed. by L. Bruni and P. L. Porta, Oxford: Oxford University Press, 29–64.

Ferrer-i-Carbonell, A. and P. Frijters (2004): "How Important is Methodology for the Estimates of the Determinants of Happiness?" *The Economic Journal*, 114, 641–659.

Firpo, S., N. M. Fortin, and T. Lemieux (2009): "Unconditional Quantile Regressions," *Econometrica*, 77, 953–973.

Fitzgerald, J., P. Gottschalk, and R. Moffitt (1998): "An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics," *The Journal of Human Resources*, 33, 251–299.

Fortin, N. M., T. Lemieux, and S. Firpo (2011): "Decomposition methods in economics," in *Handbook of Labor Economics*, Elsevier, vol. 4, 1–102.

Francis, J. L. (2009): "Wealth and the capitalist spirit," *Journal of Macroeconomics*, 31, 394–408.

Frey, B. S. (2018): *Economics of Happiness*, Springer.

Frey, B. S. and A. Stutzer (2000): "Maximising Happiness?" *German Economic Review*, 1, 145–167.

Frijters, P., J. P. Haisken-DeNew, and M. A. Shields (2004): "Money Does Matter! Evidence from Increasing Real Income and Life Satisfaction in East Germany Following Reunification," *American Economic Review*, 94, 730–740.

Goebel, J., M. M. Grabka, S. Liebig, M. Kroh, D. Richter, C. Schröder, and J. Schupp (2019): "The German Socio-Economic Panel (SOEP)," *Jahrbücher für Nationalökonomie und Statistik*, 239, 345–360.

Gosling, A., S. Machin, and C. Meghir (2000): "The Changing Distribution of Male Wages in the U.K." *The Review of Economic Studies*, 67, 635–666.

Gunaydin, G., H. Oztekin, D. H. Karabulut, and S. Salman-Engin (2021): "Minimal Social Interactions with Strangers Predict Greater Subjective Well-Being," *Journal of Happiness Studies*, 22, 1839–1853.

Headey, B., R. Muffels, and G. G. Wagner (2010): "Long-running German panel survey shows that personal and economic choices, not just genes, matter for happiness," *Proceedings of the National Academy of Sciences*, 107, 17922–17926.

Heckman, J. J., J. E. Humphries, and G. Veramendi (2018): "Returns to Education: The Causal Effects of Education on Earnings, Health, and Smoking," *Journal of Political Economy*, 126, S197–S246.

Heinze, A. (2010): "Beyond the Mean Gender Wage Gap: Decomposition of Differences in Wage Distributions Using Quantile Regression," Discussion Paper 10–043, ZEW – Centre for European Economic Research.

Huppert, F. A., N. Baylis, B. Keverne, J. F. Helliwell, and R. D. Putnam (2004): "The social context of well–being," *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 359, 1435–1446.

Jantsch, A., J. L. Blanc, and T. Schmidt (2024): "Beyond Income: Exploring the Role of Household Wealth for Subjective Well-Being in Germany," *Journal of Happiness Studies*, 25, 101.

Jost, R. (2024): "Persistence of commuting habits: context effects in Germany," *Annals of Regional Science*, 72, 837–862.

Kahneman, D., A. B. Krueger, D. A. Schkade, N. Schwarz, and A. A. Stone (2004): "A Survey Method for Characterizing Daily Life Experience: The Day Reconstruction Method," *Science*, 306, 1776–1780.

Kaiser, C. and M. Vendrik (2020): "How Threatening are Transformations of Happiness Scales to Subjective Wellbeing Research?" Discussion Paper 13905, IZA.

Karmarkar, N. (1984): "A new polynomial-time algorithm for linear programming," in *Proceedings of the Sixteenth Annual ACM Symposium on Theory of Computing*, New York, NY, USA: Association for Computing Machinery, STOC '84, 302–311.

Kasinger, C., L. Braunheim, M. Beutel, and E. Brähler (2023): "Closing the "happiness gap" by closing the wealth gap: The role of wealth on life satisfaction between east and west-Germans," *Journal of Public Health*, 31, 1433–1442.

Killingsworth, M. A., D. Kahneman, and B. Mellers (2023): "Income and emotional well-being: A conflict resolved," *Proceedings of the National Academy of Sciences*, 120, e2208661120.

Koenker, R. (2005): *Quantile regression*, vol. 38, Cambridge University Press.

Koenker, R. and G. Bassett (1978): "Regression Quantiles," *Econometrica*, 46, 33–50.

Koenker, R. and K. F. Hallock (2001): "Quantile Regression," *Journal of Economic Perspectives*, 15, 143–156.

Lerman, R. I. and S. Yitzhaki (1984): "A note on the calculation and interpretation of the Gini index," *Economics Letters*, 15, 363–368.

Lorenz, O. (2018): "Does commuting matter to subjective well-being?" *Journal of Transport Geography*, 66, 180–199.

Luttmer, E. F. P. (2005): "Neighbors as Negatives: Relative Earnings and Well-Being*," *The Quarterly Journal of Economics*, 120, 963–1002.

Machado, J. A. F. and J. Mata (2005): "Counterfactual decomposition of changes in wage distributions using quantile regression," *Journal of Applied Econometrics*, 20, 445–465.

Mader, S. and A. Franzen (2025): "Friends make us happy: evidence from three European panel studies," *Journal of Happiness Studies*, 26, 55.

Melly, B. (2005): "Decomposition of differences in distribution using quantile regression," *Labour Economics*, 12, 577–590, european Association of Labour Economists 16th Annual Conference, Universidade Nova de Lisboa, Lisbon, 9th – 11th September, 2004.

Montgomery, M. (2022): "Reversing the gender gap in happiness," *Journal of Economic Behavior & Organization*, 196, 65–78.

Neumann, B. A. and N. Scheuer (2024): "The Impact of Bequest Taxation on Wealth Inequality: Theory and Evidence," Trier University Research Papers in Economics 5/24, Trier University.

Pagán-Rodríguez, R. (2010): "Onset of disability and life satisfaction: evidence from the German Socio-Economic Panel," *European Journal of Health Economics*, 11, 471–485.

Pedersen, P. J. and T. D. Schmidt (2011): "Happiness in Europe: Cross-country differences in the determinants of satisfaction with main activity," *The Journal of Socio-Economics*, 40, 480–489.

Pence, K. M. (2006): "The Role of Wealth Transformations: An Application to Estimating the Effect of Tax Incentives on Saving," *Contributions in Economic Analysis & Policy*, 5.

Piliuk, A., E. Semerikova, and A. Nastansky (2023): "Determinants of commuting flows in Germany," *Applied Econometrics*, 71, 99–127.

Powdthavee, N. (2008): "Putting a price tag on friends, relatives, and neighbours: Using surveys of life satisfaction to value social relationships," *The Journal of Socio-Economics*, 37, 1459–1480.

Ruseski, J. E., B. R. Humphreys, K. Hallman, P. Wicker, and C. Breuer (2014): "Sport Participation and Subjective Well-Being: Instrumental Variable Results From German Survey Data," *Journal of Physical Activity and Health*, 11, 396 – 403.

Schmiedeberg, C. and J. Schröder (2017): "Leisure Activities and Life Satisfaction: an Analysis with German Panel Data," *Applied Research in Quality of Life*, 12, 137–151.

Stevenson, B. and J. Wolfers (2008): "Happiness Inequality in the United States," *The Journal of Legal Studies*, 37, S33–S79.

——— (2009): "The Paradox of Declining Female Happiness," *American Economic Journal: Economic Policy*, 1, 190–225.

Stutzer, A. and B. S. Frey (2008): "Stress that Doesn't Pay: The Commuting Paradox," *The Scandinavian Journal of Economics*, 110, 339–366.

Van Praag, B. (2011): "Well-being inequality and reference groups: An agenda for new research," *The Journal of Economic Inequality*, 9, 111–127.

Zhang, M., J. C. Fortney, J. M. Tilford, and K. M. Rost (2000): "An Application of the Inverse Hyperbolic Sine Transformation—A Note," *Health Services and Outcomes Research Methodology*, 1, 165–171.

Zweig, J. S. (2015): "Are Women Happier than Men? Evidence from the Gallup World Poll," *Journal of Happiness Studies*, 16, 917–938.

A Online Appendix

A.1 Details on Section 3.1 – Data

I draw on data from the SOEP (Goebel et al., 2019), specifically the SOEP Core Sample, which provides the necessary variables for my analysis. Moreover, I focus on wave data from 2012 (indicated by the abbreviation *bc* in front of relevant data files) and 2017 (indicated by the abbreviation *bh* in front of the relevant data files).

A.1.1 Income and Wealth

In my baseline regression I use logarithmic equivalized net household income. This variable represents the combined income after taxes and government transfers in the previous year of all individuals in the household and is standard in the literature (Binder and Coad, 2011; Clark et al., 2008; Ferrer-i-Carbonell and Frijters, 2004; Mader and Franzen, 2025). For the income variable, negative and zero values indicating an omission of the income question or a missing response, are first set to NaN, as the natural logarithm is undefined for non-positive values. The remaining positive values are then log-transformed. This common transformation allows coefficients from a quantile regression to be interpreted approximately as changes in happiness points with respect to proportional changes in income (Binder and Coad, 2011). Excluding non-positive values avoids distortion from special survey codes such as "no answer" or "does not apply".

For wealth data, I access net household wealth of individuals. Using the number of household members, I, similar to income, equivalize net household wealth among them[26]. Since wealth data typically include zero and negative values due to indebtedness, the natural logarithm cannot be applied without discarding a substantial share of observations. To retain the full wealth distribution while preserving the interpretability of a log-like transformation, I use the Inverse Hyperbolic Sine (IHS) transformation[27] (Bellemare and Wichman, 2020; Burbidge et al., 1988; Pence, 2006)

$$asinh(y) = \ln\left(y + \sqrt{y^2 + 1}\right).$$

Since $asinh(y) \approx \ln(2y)$ for large values of y, the IHS behaves similarly to the natural logarithm at higher levels of wealth while remaining well-defined at zero and for negative values. However, the deviation at lower levels warrants caution in interpretation. Bellemare and Wichman (2020) provide a detailed discussion on how to reconcile estimates obtained using the IHS transformation with the variable's original scale, particularly for smaller values of y.

To proxy reference-group position in a parsimonious way, I classify individuals relative to the median income and wealth of a reference group defined by year, gender, age band, education and federal state. Then, the "class size" is the variable "cell_n". It denotes the number of individuals belonging to the same reference group and therefore measures the size of the peer group to which an individual is compared. When a reference group contains only very few observations, relative economic position is not well defined, as the group median would be driven by the individual's own outcome. To avoid mechanically assigning peer positions in such cases, peer indicators are set to missing whenever the reference-group size falls below a minimum threshold, ensuring that relative income and wealth comparisons are based only on sufficiently populated peer groups.

[26] Household wealth is equivalized following Buhmann et al. (1988), who propose the general functional form $E = (\text{household size})^{\theta}$ with $\theta \in [0.5, 0.7]$. In line with standard practice and OECD conventions, I set $\theta = 0.5$, corresponding to division by $\sqrt{\text{household size}}$. Moreover, children below the age of 16 are weighted as 0.5 of an adult when computing household size.

[27] For an empirical application, see Zhang et al. (2000).

Following Becchetti et al. (2014), individuals are coded as relatively low-income if equivalized household income falls below 60% of the reference-group median and as relatively high-income if it exceeds 200% of that median. This is done analogously for wealth. These relative-position measures capture social-comparison channels in income and wealth without imposing functional-form assumptions and enter the analysis as additional (dummy) covariates in the quantile-based counterfactual framework.

A.1.2 Socializing

The variable measuring social time with friends is cleaned by converting all negative survey codes (such as “No Answer” or “Does Not Apply”) to missing values. Retaining only valid non-negative responses ensures that subsequent regressions use meaningful information about the frequency of social contact, which is an established determinant of life satisfaction (Huppert et al., 2004). The realizations take integer values from 1 to 4, with 1 being “Every Week” to 4 “Never”. I include this variable in the baseline as being cardinal.

Additionally, please note that the 2012 SOEP wave does not report the frequency of socializing in the same manner as the 2017 questionnaire, which will be essential in the happiness inequality analysis when matching individual data points of the same variable. Instead, the 2012 wave contains a conceptually different item: satisfaction with one’s social life. To maintain comparability in the counterfactual exercise, I use the frequency-of-socializing measure from the 2011 wave for the **same** individuals, implicitly assuming no change between 2011 and 2012. This is admittedly a strong assumption, but it is necessary to retain a rich specification that includes variables known to be central determinants of life satisfaction for the baseline wave 2012 and reference wave 2017.

A.1.3 Individual-Specific Control Variables

Individual age is calculated as the survey year minus the reported year of birth. This approach ensures that age corresponds precisely to the respondent’s age at the time of each interview wave. The squared term (Age^2) is added to allow for a non-linear (typically concave) relationship between age and life satisfaction, capturing the well-known U-shaped pattern in life satisfaction across the life cycle (see Blanchflower and Oswald (2008)). The original year-of-birth variable is dropped thereafter, as age and its square suffice for modeling life-cycle effects.

Employment status is simplified into a binary indicator equal to one if the respondent is classified as employed (values 1–8) and zero otherwise. This transformation distinguishes between those participating in the labor market and those not employed, providing a parsimonious representation of employment while maintaining interpretability. The full categorical employment variable is discarded afterwards to avoid redundancy. Controlling for employment is standard in happiness regressions, as job status is one of the strongest correlates of life satisfaction (Clark and Oswald, 1994; Di Tella et al., 2001).

A dummy variable is created equal to one for male respondents (= 1) and zero for female respondents. Gender differences in happiness are well documented and including a gender dummy allows for controlling systematic mean differences (e.g., Stevenson and Wolfers (2009)).

The marital status question has seven possible answers ([1] Married, living together with my spouse, [2] Married, living permanently separated, [3] Unmarried, I was never married, [4] Divorced / registered same-sex partnership annulled, [5] Widowed / life partner from registered same-sex partnership deceased, [6] Registered same-sex partnership, living together, [7] Registered same-sex partnership, living separately), which I map into one dummy “Married” being equal to 1 if answers are [1] or [6] and 0 otherwise. In theory, one could create more categories, however, the more categories the more difficult it gets for the solving algorithm to obtain a numerically feasible solution.

A.2 Details on Section 4.2 – The Machado-Mata Counterfactual Construction

This appendix describes the procedure used to construct the counterfactual happiness distribution following the method of Machado and Mata (2005). The goal is to obtain the distribution of happiness in 2017 that would have prevailed if the conditional quantile relationship estimated in 2012 had remained unchanged, while individual characteristics evolved as observed.

A.2.1 Step 1: Estimation of Conditional Quantile Functions for 2012

Let y_i denote life satisfaction for this derivation only and X_i the vector of covariates for individual i. For each quantile level $q \in (0, 1)$, the conditional quantile function is

$$Q_q(y_i \mid \mathbf{X}_i^{2012}) = \mathbf{X}_i^{2012\top} \boldsymbol{\beta}(q).$$

The quantile coefficients are obtained by solving[28]

$$\widehat{\boldsymbol{\beta}}(q) = \arg\min_{b \in \mathbb{R}^k} \sum_{i=1}^{n} \rho_q \left(y_i - \mathbf{X}_i^{2012\top} b \right), \qquad \rho_q(u) = u \left(q - \mathbb{I}\{u < 0\} \right).$$

To approximate the entire conditional distribution, I estimate quantile regressions on a fine grid q_j with $j = 1, \ldots, J$ for $J = 99$ in the baseline specification.

A.2.2 Step 2: Predicting Conditional Quantiles for 2017

Let X_i^{2017} denote the covariates of individual i in 2017 who is also present in the 2012 sample. Using the coefficients[29] estimated from the 2012 model, I predict the conditional quantiles for 2017 covariates:

$$\widehat{Q}_i(q_j) = \tilde{\mathbf{X}}_i^{2017\top} \widehat{\boldsymbol{\beta}}^{2012}(q_j), \qquad j = 1, \ldots, J.$$

A.2.3 Step 3: Rearrangement to impose monotonicity

A quantile function must be weakly increasing in the quantile index. In the present setting, however, the conditional quantile functions are estimated separately for each q_j. This independent estimation does not mechanically impose monotonicity across quantiles. As a result, the predicted value at a higher quantile may occasionally lie below the predicted value at a lower quantile for the same individual. Such quantile crossing is not unusual in finite samples, but it has to be corrected before interpreting the predicted sequence as a valid conditional distribution.

Following Chernozhukov et al. (2010), I therefore apply a rearrangement step. For each individual i, the predicted sequence $\left(\widehat{Q}_i(q_1), \ldots, \widehat{Q}_i(q_J) \right)$ is sorted in increasing order. This ensures that $\widetilde{Q}_i(q_j)$ is weakly increasing in q_j, so that the rearranged values can be interpreted as a valid individual-specific conditional quantile sequence.

In the present application, 6173 of 9325 individuals exhibit at least one crossing, corresponding to (66.2%) of the prediction sample. This incidence is not surprising given that the counterfactual distribution is constructed from 99 independently estimated conditional quantile functions, implying 98 adjacent quantile comparisons per individual. The incidence of crossings alone is therefore not very informative: even a very small local violation is sufficient for

[28]Notation, especially of loss function ρ, stems from Koenker (2005, p. 5).

[29]Note that $\tilde{\mathbf{X}}$ is defined as in Footnote 19.

an individual to be counted as having a crossing. I therefore also assess the magnitude of the crossings. For each individual, I compute

$$C_i = \max_j \left\{ \widehat{Q}_i(q_j) - \widehat{Q}_i(q_{j+1}), 0 \right\}$$

which records the largest adjacent downward violation in that individual's predicted quantile sequence. I get

Panel A: Crossing incidence and severity	
Statistic	Value
Individuals with any crossing	6,173
Share with any crossing	66.2%
Mean C_i	0.041
Median C_i	0.009
90th percentile of C_i	0.117
95th percentile of C_i	0.229
99th percentile of C_i	0.406
Maximum C_i	0.892

Panel B: Threshold exceedance	
Statistic	Value
Share with $C_i > 0.01$	48.5%
Share with $C_i > 0.05$	19.0%
Share with $C_i > 0.10$	11.2%
Share with $C_i > 0.25$	4.5%
Share with $C_i > 0.50$	0.5%

Table 7: Severity of quantile crossing

The diagnostics show that crossings are frequent in incidence but mostly limited in magnitude. While 66.2% of individuals exhibit at least one crossing, the median value of C_i is only 0.009, and the mean is 0.041 happiness points. Moreover, only 11.2% of individuals have a largest adjacent violation exceeding 0.10, and only 0.5% exceed 0.50. Hence, most crossings are small relative to the 0-10 happiness scale. Since rearrangement is applied within individuals, it changes only the ordering of predicted quantiles and not the pooled set of simulated values. Accordingly, the pooled Gini coefficient and CV are unchanged up to numerical precision, implying that the counterfactual inequality results are not driven by the rearrangement correction.

A.2.4 Step 4: Constructing the Counterfactual Distribution

The Machado-Mata method constructs a pseudo-sample by evaluating each individual's monotone conditional quantile function at all grid points:

$$y_{ij}^{CF} = \widetilde{Q}_i(q_j), \qquad j = 1, \ldots, J.$$

Aggregating across all individuals and quantiles yields the empirical counterfactual distribution[30]:

$$F_{CF}(y) = \frac{1}{nJ} \sum_{i=1}^{n} \sum_{j=1}^{J} \mathbf{1}\left\{ \widetilde{Q}_i(q_j) \le y \right\}.$$

From this counterfactual distribution I compute the CV and the Gini coefficient. These represent the inequality in 2017 that is mechanically implied by the structural relationship estimated in 2012.

[30]Put differently, the unconditional distribution can be approximated by integrating the conditional quantile function over both the covariate distribution and the quantile index.

A.3 Details on Section 5 – Robustness Checks

As structured in the main text, this appendix consists of detailed information for the robustness checks concerning the cross-sectional estimation and happiness inequality.

A.3.1 Alternative Wealth Measures

This appendix provides the detailed results for the arguments presented in Section 5.1. Firstly, I have the results for the wealth in categories check:

	Dependent variable: Life Satisfaction				
	q=0.1	q=0.25	q=0.5	q=0.75	q=0.9
	(1)	(2)	(3)	(4)	(5)
Intercept	4.443***	5.547***	9.000***	7.320***	8.242***
	(0.781)	(0.522)	(1.038)	(0.420)	(0.427)
Log Equiv Net HH Inc	0.265***	0.247***	0.000	0.200***	0.184***
	(0.074)	(0.049)	(0.092)	(0.039)	(0.039)
Wealth: Small Debt	0.087	0.165	0.000	-0.094	-0.076
	(0.143)	(0.155)	(0.323)	(0.136)	(0.164)
Wealth: Zero	0.280*	0.472***	1.000***	0.120	-0.084
	(0.144)	(0.148)	(0.351)	(0.134)	(0.161)
Wealth: Low	0.138	0.418***	1.000***	0.058	-0.092
	(0.142)	(0.148)	(0.363)	(0.134)	(0.154)
Wealth: Middle	0.346**	0.556***	1.000***	0.143	-0.023
	(0.145)	(0.149)	(0.350)	(0.133)	(0.154)
Wealth: Upper Middle	0.417***	0.715***	1.000***	0.220	0.070
	(0.161)	(0.160)	(0.338)	(0.138)	(0.165)
Wealth: High	0.189	0.678***	1.000***	0.197	0.057
	(0.216)	(0.182)	(0.330)	(0.152)	(0.173)
Relative Income (low)	-0.268***	-0.154**	-0.000	0.002	0.047
	(0.098)	(0.071)	(0.084)	(0.047)	(0.071)
Relative Income (high)	-0.003	0.037	-0.000	0.050	0.056
	(0.108)	(0.078)	(0.029)	(0.057)	(0.062)
Relative Wealth (low)	-0.165**	-0.118**	-0.000	-0.052	-0.084
	(0.069)	(0.054)	(0.034)	(0.041)	(0.052)
Relative Wealth (high)	-0.018	-0.062	-0.000	-0.029	-0.021
	(0.077)	(0.051)	(0.015)	(0.037)	(0.045)
Socializing	-0.365***	-0.321***	-0.000	-0.171***	-0.121***
	(0.030)	(0.028)	(0.086)	(0.018)	(0.021)
Education (years)	0.061***	0.035***	0.000	0.000	-0.011*
	(0.011)	(0.008)	(0.006)	(0.005)	(0.007)
Not Commuting	0.061	0.022	0.000	0.044	0.075*
	(0.059)	(0.046)	(0.020)	(0.034)	(0.039)
Subjective Health	-1.180***	-1.053***	-1.000***	-0.638***	-0.555***
	(0.026)	(0.023)	(0.052)	(0.020)	(0.019)
Employed	-0.013	-0.102**	-0.000	-0.125***	-0.107**
	(0.067)	(0.049)	(0.037)	(0.034)	(0.045)
Age	0.014***	0.012***	0.000	0.012***	0.011***
	(0.002)	(0.002)	(0.004)	(0.001)	(0.001)
Age Squared	-0.000***	-0.000***	-0.000	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Male Dummy	0.031	-0.013	-0.000	-0.116***	-0.148***
	(0.050)	(0.036)	(0.017)	(0.026)	(0.032)
Married	0.272***	0.156***	0.000	0.077**	0.029
	(0.060)	(0.045)	(0.044)	(0.031)	(0.037)
Observations	14908	14908	14908	14908	14908
Pseudo R^2	0.211	0.214	0.170	0.073	0.059
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 8: Quantile Regressions of Life Satisfaction (Wealth Categories)

Concerning the check using a wealth dummy, I get the following result table:

	Dependent variable: Life Satisfaction				
	q=0.1	q=0.25	q=0.5	q=0.75	q=0.9
	(1)	(2)	(3)	(4)	(5)
Intercept	4.074***	5.182***	9.000***	7.045***	7.908***
	(0.737)	(0.532)	(0.959)	(0.388)	(0.401)
Log Equiv Net HH Inc	0.300***	0.295***	0.000	0.217***	0.206***
	(0.072)	(0.053)	(0.094)	(0.037)	(0.039)
Has Wealth	0.215***	0.324***	1.000***	0.190***	-0.004
	(0.078)	(0.071)	(0.319)	(0.043)	(0.062)
Relative Income (low)	-0.203**	-0.122*	-0.000	0.028	0.091
	(0.096)	(0.070)	(0.078)	(0.047)	(0.070)
Relative Income (high)	0.013	0.007	-0.000	0.064	0.068
	(0.109)	(0.086)	(0.026)	(0.060)	(0.062)
Relative Wealth (low)	-0.214***	-0.200***	-0.000	-0.091***	-0.123***
	(0.066)	(0.047)	(0.040)	(0.035)	(0.046)
Relative Wealth (high)	-0.033	-0.024	0.000	-0.005	0.006
	(0.070)	(0.052)	(0.012)	(0.035)	(0.042)
Socializing	-0.372***	-0.318***	-0.000	-0.173***	-0.117***
	(0.031)	(0.028)	(0.082)	(0.018)	(0.021)
Education (years)	0.063***	0.038***	0.000	0.002	-0.009
	(0.011)	(0.008)	(0.006)	(0.005)	(0.006)
Not Commuting	0.051	0.029	0.000	0.040	0.073*
	(0.058)	(0.046)	(0.019)	(0.034)	(0.039)
Subjective Health	-1.172***	-1.065***	-1.000***	-0.643***	-0.563***
	(0.027)	(0.023)	(0.049)	(0.020)	(0.018)
Employed	-0.023	-0.119**	-0.000	-0.124***	-0.116***
	(0.068)	(0.051)	(0.034)	(0.035)	(0.044)
Age	0.015***	0.014***	0.000	0.013***	0.012***
	(0.002)	(0.002)	(0.004)	(0.001)	(0.001)
Age Squared	-0.000***	-0.000***	-0.000	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Male Dummy	0.024	0.002	-0.000	-0.118***	-0.144***
	(0.048)	(0.036)	(0.016)	(0.026)	(0.032)
Married	0.267***	0.154***	0.000	0.093***	0.037
	(0.058)	(0.047)	(0.041)	(0.029)	(0.035)
Observations	14908	14908	14908	14908	14908
Pseudo R^2	0.210	0.213	0.170	0.072	0.058
Note:			*p<0.1; **p<0.05; ***p<0.01		

Table 9: Quantile Regressions of Life Satisfaction

Lastly, concerning the importance of absolute vs. relative wealth, I get the following result for using IHS-transformed wealth only (so without the two relative wealth dummys):

	Dependent variable: Life Satisfaction				
	q=0.1	q=0.25	q=0.5	q=0.75	q=0.9
	(1)	(2)	(3)	(4)	(5)
Intercept	4.036***	5.200***	8.015***	7.083***	7.638***
	(0.706)	(0.495)	(0.657)	(0.378)	(0.410)
Log Equiv Net HH Inc	0.306***	0.288***	0.120**	0.214***	0.229***
	(0.069)	(0.049)	(0.051)	(0.036)	(0.040)
IHS Equiv Net HH Wealth	0.020***	0.028***	0.030***	0.015***	0.007**
	(0.004)	(0.004)	(0.007)	(0.003)	(0.003)
Relative Income (low)	-0.241***	-0.155**	-0.111	0.017	0.090
	(0.090)	(0.069)	(0.100)	(0.048)	(0.070)
Relative Income (high)	0.051	0.029	-0.000	0.064	0.066
	(0.115)	(0.087)	(0.038)	(0.058)	(0.064)
Socializing	-0.380***	-0.310***	-0.127***	-0.174***	-0.124***
	(0.032)	(0.029)	(0.038)	(0.018)	(0.021)
Education (years)	0.057***	0.037***	0.007*	0.000	-0.014**
	(0.010)	(0.008)	(0.004)	(0.005)	(0.006)
Not Commuting	0.064	0.038	0.023	0.047	0.064*
	(0.059)	(0.044)	(0.019)	(0.033)	(0.037)
Subjective Health	-1.172***	-1.060***	-0.921***	-0.644***	-0.570***
	(0.027)	(0.023)	(0.026)	(0.021)	(0.017)
Employed	-0.041	-0.121**	-0.057**	-0.119***	-0.142***
	(0.067)	(0.050)	(0.028)	(0.035)	(0.045)
Age	0.014***	0.013***	0.005***	0.013***	0.012***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Age Squared	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Male Dummy	0.009	-0.001	-0.021	-0.117***	-0.140***
	(0.046)	(0.035)	(0.017)	(0.025)	(0.031)
Married	0.270***	0.169***	0.063**	0.089***	0.027
	(0.056)	(0.044)	(0.027)	(0.030)	(0.036)
Observations	14908	14908	14908	14908	14908
Pseudo R^2	0.209	0.213	0.169	0.072	0.057
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 10: Quantile Regressions of Life Satisfaction