The Effects of Income Transparency on Well-Being:
Evidence from a Natural Experiment
By Ricardo Perez-Truglia
Online Appendix

A Additional Results and Robustness Checks

A.1 Additional Event-Study Graphs

In this section, we present some additional event-study graphs.

Figure A.3 presents variations of the event study graphs from Figure 6. Figure A.3.a reproduces Figure 6.b, only that breaking down the coefficients at the yearly level instead of using pairs of years. Given that we are estimating twice as many coefficients with the same number of observations, each individual coefficient is less precisely estimated and thus one must be more careful with the interpretation.

As discussed in section 2.2, there were some factors during 2001–2013 that may have increased or decreased the degree of income transparency. The event-study analysis from Figure A.3.a allows for a closer look at this. We cannot reject the null hypothesis that, for individuals with higher Internet access, all the post-2001 coefficients from Figure A.3.a are equal (p-value=0.201). This evidence suggests that the effects of higher transparency were stable over time. One possible interpretation is that the events during 2001–2013 did not affect the degree of income transparency. An equally valid interpretation is that there were factors reducing and increasing the degree of transparency over time, but these positive and negative effects canceled each other out.

We can use Figure A.3.a to get a closer look at what happened in 2001. For individuals with lower Internet access, the coefficient on 2001 is close to zero and statistically insignificant. On the contrary, for individuals with higher Internet access, the coefficient on 2001 is positive, large and statistically significant. The point estimate for 2003 is higher than the point estimate for 2001, suggesting that perhaps the effects of the higher transparency took a bit more time to fully materialize.[53] However, the difference between these two coefficients is not statistically significant (p-value=0.1744). Moreover, this difference could be due to measurement error: due to the timing of the survey collection, a non-trivial share of the survey responses for 2001 may have been collected before the change in disclosure took place, which will bias the coefficient for 2001 downwards.

Figure A.3.a also allows for a closer look at what happened in 2004–2016, when the online tax lists were limited to the three-week period following the release of the data. Due to the

[53]It may take some time after the online tax lists are published for its effects on well-being to fully materialize. Consider for example the anecdote about kids being bullied at school because their classmates found out that their parents were poor through the website. Even though some kids may have been bullied the day after the publication of the online tax lists, most kids would probably not be bullied until months after the tax lists became available.

timing of our survey, this restriction is unlikely to affect our estimates. The survey was collected every two years, and thus only one of the survey years, 2005, falls in this sub-period. Moreover, in 2005 the income search tool was available during the last two weeks of October and the first week of November. Since the survey responses are collected from late September to early December, it is likely that most of the survey responses in 2005 were collected while the income search tool was available. And since the tool was not available during the rest of the year, the income search tool may have been especially salient in this period of 2005. Figure A.3.a shows a coefficient for higher Internet in 2005 that is positive and statistically significant, suggesting that the effects of higher transparency did not decline in 2005.[54]

We can use Figure A.3.a to get a closer look at what happened in 2011–2013, when the government introduced a hassle to use the tax lists: individuals had to log into the official website of the tax agency with a pin-code and a password. The coefficients for 2011 and 2013 suggest that the effects of the transparency policy persisted through 2011–2013. This finding is consistent with the evidence discussed in section 2.3, according to which the volume of searches was still substantial in 2011–2013. Moreover, even if the search volume was reduced, that should not necessarily undo the effects of transparency. Indeed, even if the online tax lists were removed altogether, it may take years for the effects on well-being to vanish. If an individual became unhappy because she found out that she is poorer than she thought, removing her access to the online tax lists will not make her suddenly forget how poor she is. Similarly, if the individual is being bullied by others, those others will not suddenly forget that the individual is poor after losing access to the website.

Figure A.3.b presents a robustness check for the analysis of the data from the German Socio-Economic Panel survey. The German survey was collected every year in 1985–2013, while the data from the Norwegian Monitor Survey was collected every odd year in 1985–2013. In the baseline specification, to maximize power, we use all the years available in the German data. As a robustness check, Figure A.3.b reproduces the German event study from Figure 6.c, only that restricting the German observations to the odd-numbered years, to mimic the frequency of the Norwegian Monitor Survey. The coefficients in Figure A.3.b are less precisely estimated because we are discarding roughly half of the data. However, the main result is still robust: Figure A.3.b shows that the life satisfaction-income gradient did not change in Germany post–2001.

Figure A.3.c reproduces Figure 6.a, only that breaking down the coefficients at the yearly level instead of using pairs of years. Again, given that we are estimating twice as many coefficients with the same number of observations, each individual coefficient is less precisely estimated and thus the results must be interpreted with that caveat in mind.

Figure A.4 shows the event-study graphs for the other two outcomes used in the analysis:

[54] On the other hand, if we measured the effect of higher transparency as the gap between the coefficient for lower and higher Internet, that estimate would suggest that the effects of transparency were smaller in 2005.

Perceived Rank and *Income Adequacy*. Note, however, that we have to be more careful when interpreting these results. First, these two questions were not included in the survey until 1993, and thus the results will be less precisely estimated than for *Happiness*. Second, the magnitude of the effects for these two outcomes (8.5% and 4.7%, respectively) are substantially smaller than the effects for Happiness (29%), and thus these effects would be more difficult to detect even if we could hold constant the number of observations.

Figure A.4 shows that, consistent with the results presented in regression form in section 4.2, the event-study analysis for *Perceived Rank* and *Income Adequacy* are suggestive but not nearly as sharp as the results for well-being. Figure A.4.a reproduces Figure 6.b, only that using *Perceived Rank* as dependent variable instead of *Happiness*. Figure A.4.a shows that the gradient between perceived and actual income rank evolved similarly for individuals with higher and lower Internet access, but then diverged after 2001. However, this divergence is less precisely estimated. Figure A.4.b reproduces Figure 6.b, only that using *Income Adequacy* as dependent variable instead of *Happiness*. Again, the event-study graph suggests a divergence in the gradient after 2001, but this finding is not nearly as precisely estimated as the corresponding finding for *Happiness*.

Last, for the sake of completeness, Figures A.4.c and A.4.d reproduce Figure 6.a, but instead of using Happiness as dependent variable they use *Perceived Rank* and *Income Adequacy*, respectively.

A.2 Effects on the Average Level of Well-Being

Section 4.1 reports the findings for the effect of transparency on the gradient between subjective well-being and income rank. In this section, we report the findings for the effects of transparency on the average level of subjective well-being.

To estimate these average effects, we follow a differences-in-differences strategy that is based on the exposure variable discussed in section 4:

$$SWB_{i,t} = \alpha_1 \cdot HigherInternet_{i,t} + \alpha_2 \cdot HigherInternet_{i,t} \cdot I_t^{01-13} + X_{i,t}\beta + \delta_t + \epsilon_{i,t} \tag{A.1}$$

$SWB_{i,t}$ denotes subjective well-being of individual i in year t. $HigherInternet_{i,t}$ takes the value 1 if individual i's observable characteristics in year t, such as the age and education, predict above-median Internet access at home. I_t^{01-13} is a dummy variable indicating the period of higher income transparency. $X_{i,t}$ is a vector with a set of control variables. δ_t denotes the year dummies. And $\epsilon_{i,t}$ denotes the error term.

The coefficient α_1 estimates the well-being gap between individuals with higher and lower Internet access during 1985–2000, while α_2 measures the change in that gap from 1985–2000 to 2001–2013. If we assume that individuals with $HigherInternet_{i,t} = 0$ were not affected by the change in transparency while individuals with $HigherInternet_{i,t} = 1$ were affected by it, then α_2 measures the effect of the 2001 change of disclosure on the average level of well-being.[55]

The results from section 4 suggest that the effects on the happiness-income gradient for individuals with higher Internet was 0.217 (column (3) of Table 3). If that effect was purely a redistribution of happiness from poorer to richer individuals, we would expect the effect on average happiness to be null (i.e., $\alpha_2 = 0$). However, it is possible that the effects were asymmetric. However, if poorer households lost more happiness than the happiness gained by richer households, then we would expect $\alpha_2 < 0$. On the contrary, if richer households gained more happiness than the happiness lost by poorer households, then we would expect $\alpha_2 > 0$.

Column (1) of Table A.1 reports the regression results using *Happiness* as the dependent variable. The estimated α_2 (0.014) is close to zero and statistically insignificant (p-value=0.460). This evidence is consistent with a pure redistribution of happiness from poorer to richer households ($\alpha_2 = 0$). As a robustness check, column (2) introduces the falsification test to assess the possibility that the parameter α_2 is biased because of differential pre-trends across individuals with high and low Internet access, as in the specification from equation (3) in section 3. The results from column (2) are consistent with column (1): the coefficient α_2 (0.019) is still small and statistically insignificant (p-value=0.362), and the coefficient on the interaction with *I{1997–2000}* (0.015) is also small and statistically insignificant (p-value=0.582). As an additional robustness check, column (3) uses life satisfaction as the dependent variable instead of happiness. Since this outcome was measured starting in 1999, the pre-treatment period consists of just one year of data and thus the findings must be taken with a grain of salt. The estimated α_2 (0.027) is still small and statistically insignificant (p-value=0.390). In sum, the estimates for happiness and life satisfaction suggest that the change in disclosure had a null effect on the average level of well-being.

For the sake of completeness, columns (4) through (7) of Table A.1 reproduce the results from columns (1) and (2), but with the additional outcomes (*Perceived Rank* and *Income Adequacy*) as dependent variables. Column (4) suggests that the effect on the average perceived income rank is positive (0.061) and statistically significant (p-value=0.001). This finding is robust to the alternative specification, reported in column (5): the coefficient on the interaction with *I{1997–2000}* (-0.017) is close to zero, statistically insignificant (p-value=0.570), and statistically different from the interaction with *I{2001–2013}* (p-value=0.002). These results imply that individuals found out through the online tax lists that, on average, their income rank was higher than they thought. This evidence is consistent with the findings from Karadja, Mollerstrom, and Seim (2017) that, on average, households under-estimate their own income

[55] If individuals with lower Internet access were affected less than individuals with higher Internet access but still affected to some extent, then the coefficient α_2 would still measure the average effect of transparency but suffer from attenuation bias.

rank and thus will update their perceptions upwards when provided with accurate information.

Last, columns (6) and (7) show the results for *Income Adequacy* as the dependent variable. Column (6), which reports the most basic specification, indicates that the coefficient on the interaction with *I{2001–2013}* (0.037) is close to zero, although statistically significant (p-value=0.069). In the alternative specification, reported in column (7), this coefficient is even smaller (0.003) and becomes statistically insignificant (p-value=0.924). These results suggest that were no significant effects on the average level of *Income Adequacy*.

A.3 Probit-OLS versus OLS and Ordered Probit

In the baseline specification, we coded the subjective questions using the Probit-OLS method. In this section, we show that the results are robust under alternative econometric models.

When constructing the happiness outcome, instead of arbitrarily assigning values 1, 2, 3, and 4 to the four possible answers to the happiness question, we employ the Probit-OLS method to assign these values (van Praag and Ferrer-i-Carbonell, 2008). This method consists of assigning values to match the distribution of responses to a normal distribution. For example, if a fraction q reports the lowest category ("not at all satisfied"), the Probit-OLS method assigns the lowest category an score of $E[z|z < q]$, where z is distributed standard normal. The resulting values for the happiness scores are 1.36 ("very happy"), -0.17 ("quite happy"), -1.67 ("not particular happy") and -2.79 ("not at all happy").

Table A.2 explores the robustness of the results to different treatments of the subjective data. Columns (1) and (2) denote the baseline specifications for Happiness and Life Satisfaction, respectively (these results are identical to columns (3) and (5) from Table 3). Columns (3) and (4) from Table A.2 correspond to the specifications from columns (1) and (2), but with the responses to the happiness and life satisfaction questions coded from 1 to 4 and 1 to 5, respectively. The results from columns (3) and (4) are qualitatively consistent with the results from columns (1) and (2). The magnitudes of the coefficients from columns (3) and (4) are not directly comparable to the coefficients from columns (1) and (2), because of the differences in scales of the dependent variables. With that caveat in mind, the findings are quantitatively robust across the two specifications: column (3) suggests that the change in disclosure increased the happiness-income gradient by 27% ($= \frac{0.049}{0.179}$), which is close to (and statistically indistinguishable from) the 29% increase implied by the coefficients from column (1); and column (4) indicates that the change in disclosure increased the life satisfaction-income gradient by 20% ($= \frac{0.089}{0.452}$), which is close to (and statistically indistinguishable from) the 21% increase implied by the coefficients from column (2).

Columns (5) and (6) from Table A.2 estimate the same specifications from columns (1) and (2), except that they use an Ordered Probit model instead of the OLS model from the baseline specification. Columns (5) and (6) report the raw coefficients from the Ordered Probit model, which cannot be compared in magnitude directly to the OLS coefficients from columns (1) and (2). In terms of signs and statistical significance, the results from columns (5) and (6) are consistent with the results from columns (1) and (2). The results are also quantitatively robust: column (5) suggests that the change in disclosure increased the happiness-income gradient by 30% ($= \frac{0.113}{0.380}$), which is very close to (and statistically indistinguishable from) the 29% increase implied by the coefficients from column (1). Similarly, column (6) indicates that the change in disclosure increased the life satisfaction-income gradient by 25% ($= \frac{0.166}{0.667}$), which is close to (and statistically indistinguishable from) the 21% increase implied by the coefficients from column (2).

A.4 Binned Scatterplots

In the baseline specification, we assume a linear relationship between subjective well-being and *Income Rank*. In this section, we present results under a more flexible specification using binned scatterplots.

In addition to exploring the role of outliers and non-linearities, the binned scatterplots can also shed light on the distribution of the effects of transparency. These effects could be unevenly distributed along the income distribution due to an uneven exposure to the online tax lists. Indeed, the evidence from the triple-differences specification already indicates that the effects were concentrated in individuals with higher Internet access. Since the individuals with higher Internet access tend to be richer than the rest of the population, then we would expect effects that are stronger for richer individuals.[56]

We follow the baseline regression specification but, instead of letting *Income Rank* enter linearly in the right hand side of the regression equation, we include this variable as a set of dummies for nine equal-sized income groups: one set of dummies for the post-2001 period, and another one for the 2001–2013 period. The middle groups are set as omitted categories, and thus their coefficients are normalized to zero.

Figure A.5 presents the results for the binned scatterplot analysis. Figure A.5.a shows the results for *Happiness*, and Figures A.5.b through A.5.d show the results for the rest of the outcomes. Most important, these figures suggest that the linear specification used in the baseline specification for *Income Rank* provides a fair approximation. These figures also confirm that the results are not driven by outliers or non-linearities. Last, the results are consistent

[56]Additionally, richer individuals may be more salient in the online tax lists – for example, it was common for the online search tools to provide rankings with the richest individuals in each city. Also, the effects may be stronger for some income groups due to the nature of income comparisons. For example, if individuals are last-place-averse (Kuziemko et al., 2014), then the effects of transparency may be particularly strong at the bottom of the income distribution. Similarly, if well-being is a concave function of relative pay (Fehr and Schmidt, 1999), the effects of transparency may be weaker among richer individuals (Card et al., 2012).

with the expectation that, due to differences in Internet access, the effects were stronger in the upper part of the income distribution.

A.5 Local vs. National Income Rank

In the baseline specification, *Income Rank* corresponds to the position of the respondent in the national distribution of household income. To the extent that this does not fully capture the income comparisons that people care about, this can be a source of measurement error and thus introduce an attenuation bias. In this section we present results under an alternative definition, based on the local instead of the national income rank.

Using the local alternative for *Income Rank* may be a more appropriate specification for a number of reasons. For example, due to the Balassa-Samuelson effect, the local income rank may predict purchasing power better than the national income rank. And regarding income comparisons, individuals may care the most about the comparisons to their social contacts, which are probably drawn disproportionately from the same area of residence.

The most disaggregated geographic identifiers in the Norwegian survey data correspond to the county identifiers. There are 19 counties in Norway, with populations in 2001 ranging from 73,417 in Finnmark to 599,230 in Oslo, with a median of 233,705 in Vestfold (Source: Statistics Norway). We constructed a local *Income Rank*, based on the within-county rank instead of the national rank. The income distributions across these 19 regions are fairly similar, and as a result the national and county ranks are highly correlated to each other (correlation coefficient of 0.9681).

The results are presented in Table A.3. Columns (1) and (2) denote the baseline specifications for happiness and life satisfaction, respectively, which are identical to columns (3) and (5) from Table 3. Columns (5) and (6) from Table A.3 reproduce the specifications from columns (1) and (2) but using the local version of *Income Rank* instead. The results from columns (5) and (6) are similar to (and statistically indistinguishable from) the results from columns (1) and (2) – if anything, and consistent with the argument of attenuation bias, the estimated effects of transparency are slightly larger under the local definition of *Income Rank*.

A.6 Alternative Definitions of Income Rank

This section provides some robustness checks related to the construction of *Income Rank*.

One potential source for concern is that the income question added a bin in 1999, which could contaminate the comparison of the happiness-income gradient around 2001. Since only 1.55% of respondents fell in the ninth bin in 1999, this is probably a minor concern. To address any remaining concerns, Table A.4 presents a sharp robustness check. The results from columns (1) through (3) correspond to the baseline definition of *Income Rank* (identical to columns (1),

(3) and (4) of Table 3). The results from columns (4) through (6) are based on the same specifications but with an alternative version of *Income Rank*, which ignores the distinction between the eight and ninth bin. That is, we treat the data as if the ninth bin was never introduced: we start by pooling the responses to the ninth and eight bins in the raw data, and then we replicate the data construction and analysis starting from this revised dataset. The results are robust to this check: the coefficients from columns (1) through (3) are almost identical to the corresponding coefficients from columns (4) through (6). For example, column (1) indicates that transparency increased the happiness-income gradient by 0.090 (p-value=0.005), while the corresponding coefficient in column (4) is 0.090 (p-value=0.005).

Since the income question is elicited in bins, our baseline specification uses the standard method from the happiness literature to impute the values of *Income Rank* within each bin (Stevenson and Wolfers, 2008; Kahneman and Deaton, 2010). Table A.4 also assesses the sensitivity of the results to this imputation method. Again, columns (1) through (3) correspond to the results with the baseline definition of *Income Rank*, which uses the imputation. Columns (7) through (9) show the results under the alternative definition of *Income Rank* using the raw data (i.e., without the within-bin imputation). Most important, the results are qualitatively and quantitatively robust across the imputed and non-imputed versions of *Income Rank*. For example, column (1) indicates that the happiness-income gradient increased by 29% after 2001, while column (7) indicates that it increased by 25% – moreover, we cannot reject the null hypothesis that these two estimates are equal. Additionally, the comparison of the coefficient on *Income Rank* between columns (1)–(3) and (7)–(9) suggests that the imputation method is indeed helping to ameliorate the measurement error: the happiness-income gradients are almost 10% higher when using the imputation.

A.7 Alternative Definitions of Higher Internet

This section shows the robustness of the results under alternative definitions of *I{Higher Internet}*.

The results are presented in Table A.5. Column (1) corresponds to the baseline definition of *I{Higher Internet}* (identical to column (4) of Table 3). Columns (2) through (5) are based on the same specification but with alternative versions of *I{Higher Internet}*. In column (2), *I{Higher Internet}* is identical to the baseline definition except that, instead of using responses to *Internet Access* for 2001, it is based on responses for 1999. In column (3), *I{Higher Internet}* is identical to the baseline definition except that it is based on responses to *Internet Access* for the entire period for which the *Internet Access* is available (1999–2013). In column (4), *I{Higher Internet}* is identical to the baseline definition except that it splits individuals by the median value of predicted Internet access over the entire sample, rather than splitting them by the median value for the respective year. In column (5), *I{Higher Internet}* is identical to the baseline definition except that we predict Internet access using a Probit regression instead of an OLS regression. The coefficients reported in Table A.5 indicate that the results are qualitatively and quantitatively robust between the baseline specification (column (1)) and the alternative specifications (columns (2)–(5)).

A.8 Evolution of Inequality and Sample Composition

Since the survey is a repeated cross-section, we need to check that the composition of the survey respondents has not changed abruptly around 2001. To address this concern, Table A.6 presents the yearly averages of some individual characteristics. Consistent with the gradual changes in the composition of the universe of Norwegians, over the 28 years there were gradual changes in the composition of of the survey respondents: a gradual increase in age and education, and a gradual decrease in marriage rates. Most important, the results confirm that the composition of the survey respondents has not changed abruptly around 2001.

Relatedly, Table A.3 explore the robustness of the results to the use of sampling weights. Columns (1) and (2) denote the baseline specifications for happiness and life satisfaction, respectively, which are identical to columns (3) and (5) of Table 3. Columns (3) and (4) of Table A.3 reproduce the same regressions from columns (1) and (2), but using individual-specific sampling weights computed by the team in charge of collecting the survey data. As expected, using sample weights does not change the results: the coefficients in columns (3) and (4) are very similar to (and statistically indistinguishable from) the coefficients from columns (1) and (2).

Another potential source for concern is that the change in the happiness-income gradient is mechanically driven by an increase in income inequality. In other words, richer individuals may become happier because they are able to afford more stuff, while poorer individuals become less happy because they can afford less stuff. This possibility seems highly unlikely, because it would require a large, sudden and persistent increase in inequality that would be unprecedented in a developed country. To address this concern more directly, Table A.7 presents data on the evolution of income inequality in Norway during the sample period.

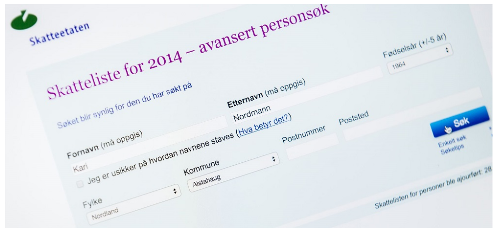
Table A.7 shows the evolution of income inequality according to the data from the Norwegian Monitor Survey. The coefficient of variation in incomes did not change abruptly in 2001 – furthermore, it was stable during the entire 1985–2001 period. For example, we can mimic the comparison from the event-study analysis: the coefficient of variation went from an average of 0.521 in 1997/99 to an average of 0.520 in 2001/03, amounting to a mere reduction of 0.21%. This change in inequality is tiny compared to the estimated 29% increase in the happiness-income gradient around 2001.

Table A.7 also presents results for two measures of inequality constructed with administrative data.[57] The *Gross Gini Index* is an index of inequality based on the gross household income, while the *Net Gini Index* is an index of inequality based on the net household income. Just like the survey data, the administrative data indicates that there was no abrupt change in income inequality around 2001, and that the levels of inequality were stable over the entire 1985–2013 period. Moreover, we can mimic the comparison from the event-study analysis with these alternative measures of inequality. The *Gross Gini Index* increased from an average of 44.24 in 1997/99 to an average of 44.79 in 2001/03. This amounts to a mere 1.3% increase in income inequality, which is insignificant relative to the 29% increase in the happiness-income gradient measured around 2001. Similarly, the *Net Gini Index* increased from an average of 24.55 in 1997/99 to an average of 25.15 in 2001/03, amounting to a mere 2.4% increase in income inequality. Again, these changes in income inequality are minuscule compared to the estimated 29% increase in the happiness-income gradient around 2001.

[57] These measure of inequality were obtained from the Chartbook of Economic Inequality and are publicly available in the following URL: https://www.chartbookofeconomicinequality.com/.

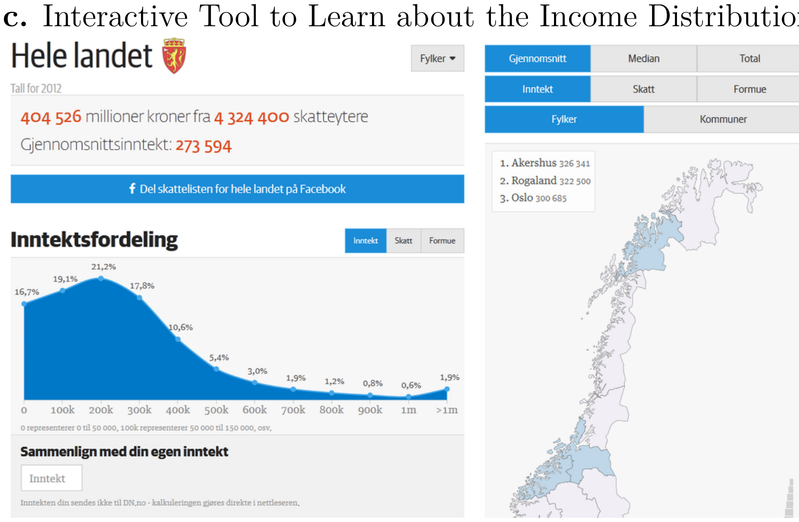
Figure A.1: More Screenshots of Websites and Apps Designed to Search the Tax Records

a. Official Search Tool for the Tax Records of 2014



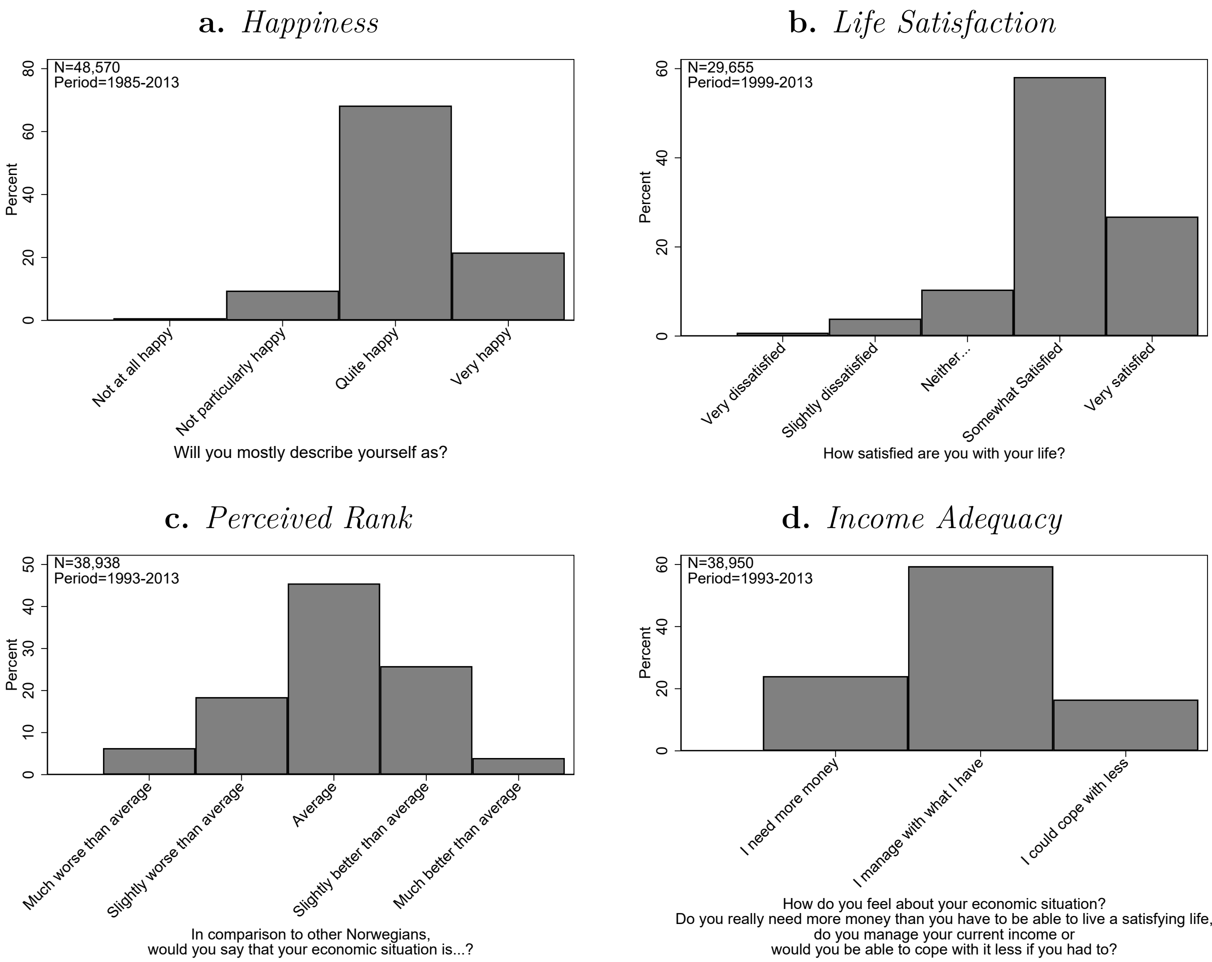
b. TV2s Skattelisten Iphone App

c. Interactive Tool to Learn about the Income Distribution



Source: (a) and (b): Origo (2010). (c) web.archive.org.

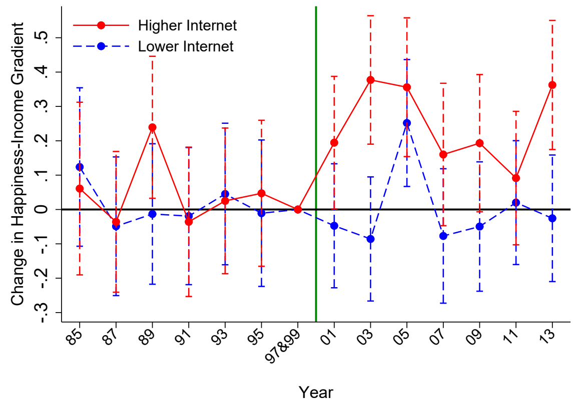
Figure A.2: Histograms for All the Outcome Variables



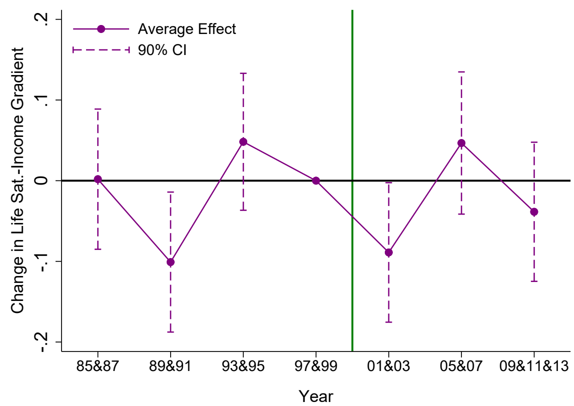
Notes: The histograms show the raw distribution of dependent variables used in the analysis. See Table 1 for further data definitions.

Figure A.3: Alternative Event-Study Graphs for Subjective Well-Being

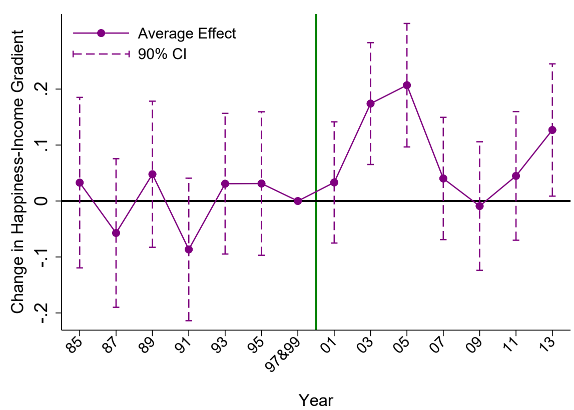
a. Norway (yearly, by Internet)



b. Germany (odd years)

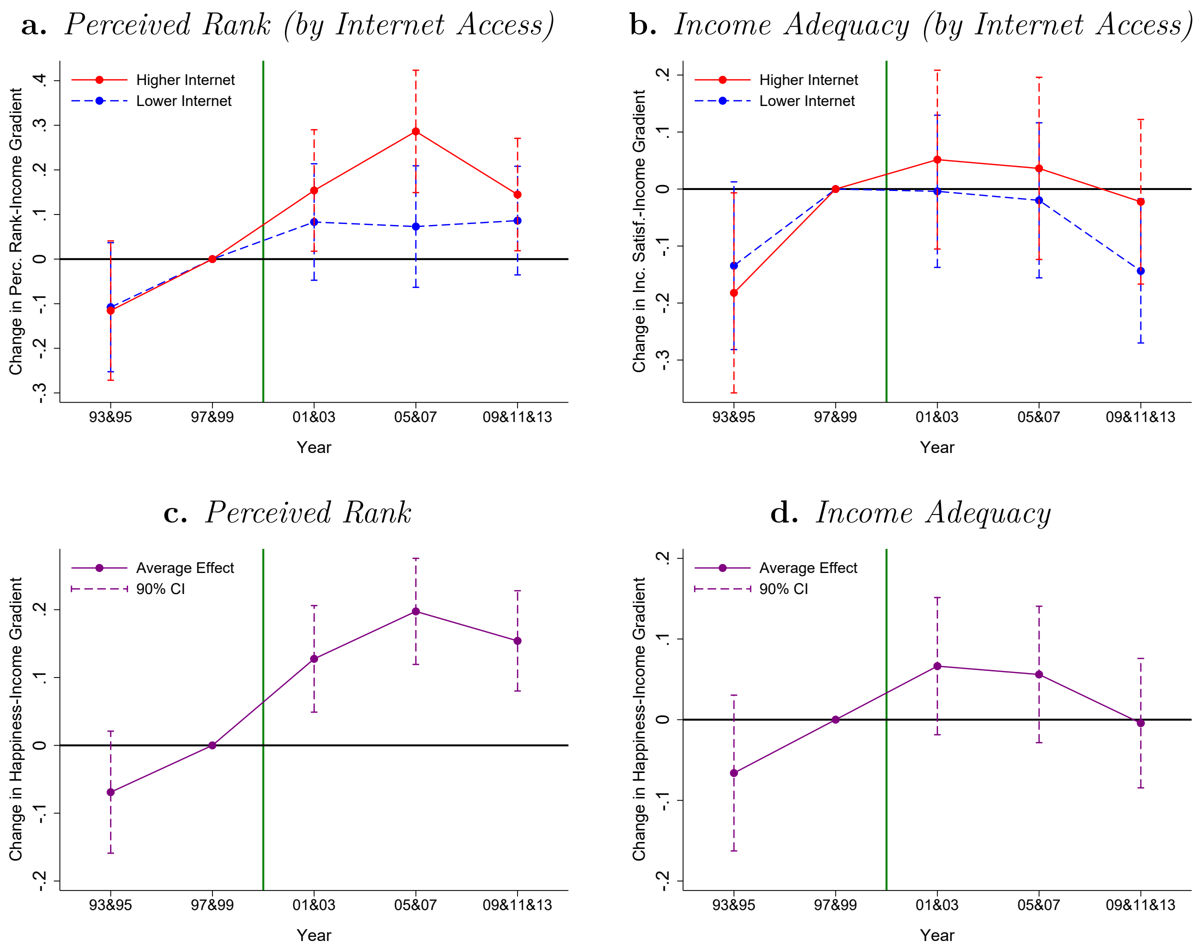


c. Norway (yearly)



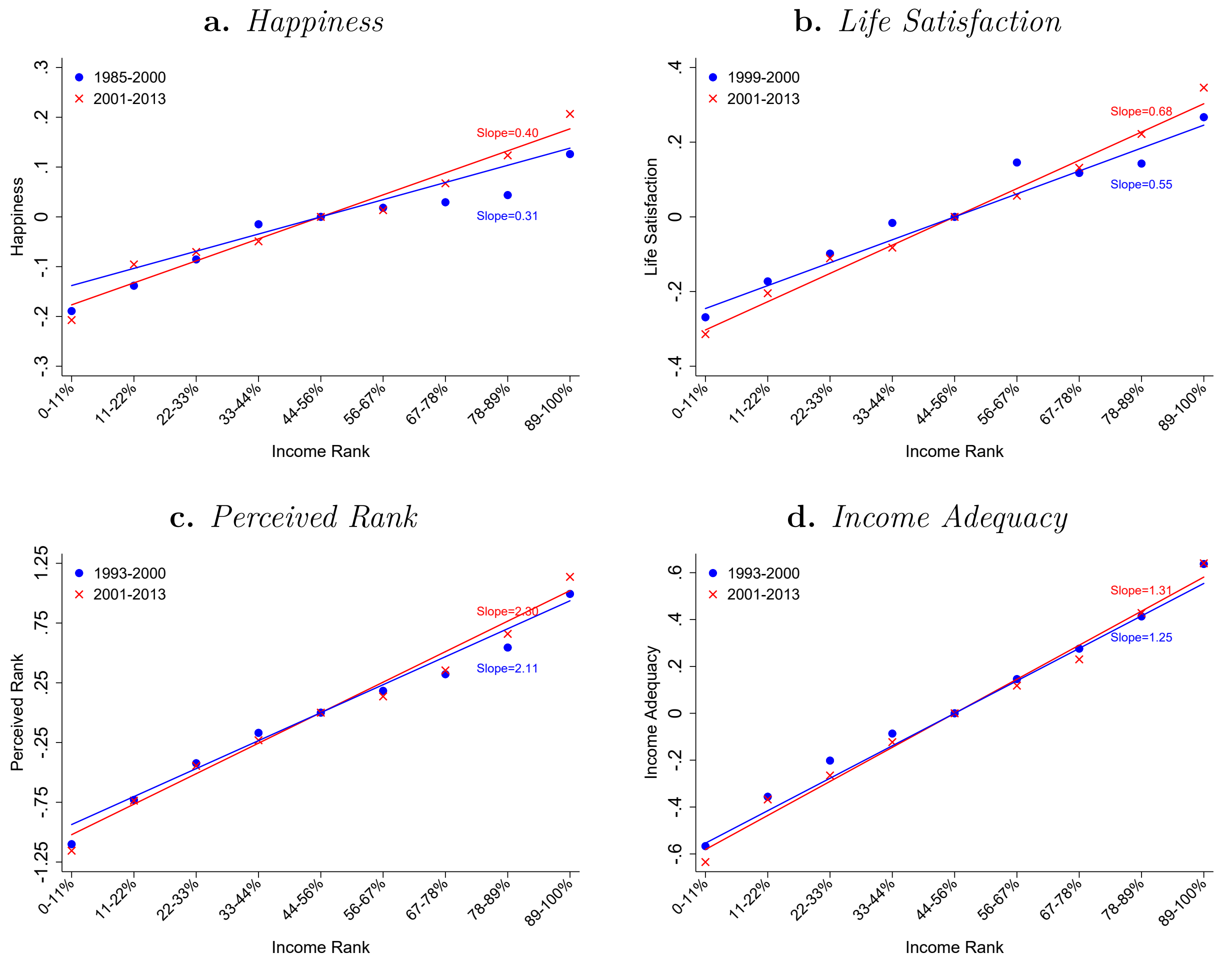
Notes: Variations of the event study graphs presented in Figure 6. Panel (a) reproduces panel (b) from Figure 6, except that it breaks down the coefficients at the year level. Panel (b) reproduces panel (c) from Figure 6, only that it restricts the data to odd-numbered years in 1985–2013. Panel (c) reproduces panel (a) from Figure 6, except that it breaks down the coefficients at the year level. See notes to Figure 6 for more details about the specification and the data.

Figure A.4: Event-Study Graphs for Perceived Income Rank and Income Adequacy



Notes: Panel (a) reproduces panel (b) from Figure 6, only that it uses *Perceived Rank* as the dependent variable instead of *Happiness*. Panel (b) reproduces panel (b) from Figure 6, only that it uses *Income Adequacy* as dependent variable instead of *Happiness*. Panel (c) reproduces panel (a) from Figure 6, only that it uses *Perceived Rank* as the dependent variable instead of *Happiness*. Panel (d) reproduces panel (a) from Figure 6, only that it uses *Income Adequacy* as dependent variable instead of *Happiness*. All the dependent variables have been normalized to have mean 0 and standard deviation of 1. See notes to Figure 6 for more details about the specification and the data.

Figure A.5: Binned Scatterplot Showing the Change in Gradient Between Happiness and Income Rank



Notes: Panel (a) corresponds to the same regression from column (1) of Table 3, only that *Income Rank* is introduced as two sets of nine equally-sized dummies (one set for the post-2001 period and another one for the pre-2001 period), with the coefficients on the middle categories normalized to zero. Panel (b), (c) and (d) are identical to panel (a), only that instead of *Happiness* they use the dependent variables *Life Satisfaction*, *Perceived Rank* and *Income Adequacy*, respectively. All these dependent variables have been normalized to have mean 0 and standard deviation of 1, with higher values denoting higher happiness/satisfaction/rank/adequacy. See notes to Table 3 for more details about the regression specification and the data.

Table A.1: Effects on the Average Level of Well-Being

	(1) Happiness	(2) Happiness	(3) Life Satisf.	(4) Perc. Rank	(5) Perc. Rank	(6) Income Adequacy	(7) Income Adequacy
I{Higher Internet}	-0.010	-0.016	-0.037	-0.078***	-0.067**	-0.048**	-0.013
	(0.017)	(0.020)	(0.033)	(0.019)	(0.027)	(0.021)	(0.029)
I{Higher Internet} * I{2001-2013}$^{(i)}$	0.014	0.019	0.027	0.061***	0.050*	0.037*	0.003
	(0.018)	(0.021)	(0.032)	(0.018)	(0.026)	(0.020)	(0.029)
I{Higher Internet} * I{1997-2000}$^{(ii)}$		0.015			-0.017		-0.055*
		(0.027)			(0.030)		(0.034)
P-value (i)=(ii)		0.861			0.002		0.016
Period	85-13	85-13	99-13	93-13	93-13	93-13	93-13
Observations	48,570	48,570	29,655	38,938	38,938	38,950	38,950

Notes: *** p<0.01, ** p<0.05, * p<0.1. Heteroskedasticity-robust standard errors in parenthesis. Each column corresponds to a separate OLS regression. *Happiness, Life Satisfaction*, *Perceived Rank* and *Income Adequacy* were normalized to have mean 0 and standard deviation of 1, with higher values denoting higher happiness/satisfaction/rank/adequacy. *I{Higher Internet}* is a dummy variable that takes the value 1 if the respondent's predicted *Internet Access* is above the median value for a given year. *I{2001–2013}* takes the value 1 for 2001–2013. *I{1997–2000}* takes the value 1 for 1997–2000. All regressions control for year dummies, age, age squared, a gender dummy, three education dummies, four dummies for marital status, four dummies for household size, three dummies for number of working household members and *Income Rank* (i.e., the respondent's position in the national distribution of household income for that year). Data from the Norwegian Monitor Survey, which has been collected every other year in 1985–2013. See Table 1 for a summary of data definitions and Table 2 for descriptive statistics.

Table A.2: Robustness to Alternative Econometric Models

	(1) Happiness	(2) Life Satisf.	(3) Happiness	(4) Life Satisf.	(5) Happiness	(6) Life Satisf.
Income Rank	0.310***	0.585***	0.179***	0.452***	0.380***	0.667***
	(0.032)	(0.056)	(0.018)	(0.043)	(0.039)	(0.064)
Income Rank * I{2001-2013}$^{(i)}$	0.090**	0.122**	0.049**	0.089**	0.113**	0.166***
	(0.037)	(0.055)	(0.021)	(0.042)	(0.045)	(0.062)
Income Rank * I{1997-2000}$^{(ii)}$	0.001		0.001		-0.000	
	(0.048)		(0.028)		(0.059)	
P-value (i)=(ii)	0.043		0.053		0.034	
Model	OLS	OLS	OLS	OLS	O-Probit	O-Probit
POLS Transformation	Yes	Yes	No	No	No	No
Period	85-13	99-13	85-13	99-13	85-13	99-13
Observations	48,570	29,655	48,570	29,655	48,570	29,655

Notes: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Heteroskedasticity-robust standard errors in parenthesis. Column (1) through (4) report coefficients from OLS regressions, while columns (5) and (6) report raw coefficients from an Ordered Probit model. *Happiness* and *Life Satisfaction* are responses to subjective questions where higher value denotes higher happiness/satisfaction. In columns (1) and (2), responses to these questions were coded using the Probit-OLS method, and then normalized to have mean 0 and standard deviation of 1. In columns (3) and (5), responses to the happiness question are assigned values from 1 (not at all happy) to 4 (very happy). In columns (4) and (6), responses to the life satisfaction question are assigned values from 1 (very dissatisfied) to 4 (very satisfied). *Income Rank* denotes the position of the respondent's household relative to all the other respondents for that year, from 0 to 1. *I{2001–2013}* takes the value 1 for 2001–2013. *I{1997–2000}* takes the value 1 for 1997–2000. All regressions control for year dummies, age, age squared, a gender dummy, three education dummies, four dummies for marital status, four dummies for household size and three dummies for number of working household members. Data from the Norwegian Monitor Survey, which has been collected every other year in 1985–2013. See Table 1 for a summary of data definitions and Table 2 for descriptive statistics.

Table A.3: Robustness to the Local Definition of Income Rank and the Sample Weights

	(1) Happiness	(2) Life Satisf.	(3) Happiness	(4) Life Satisf.	(5) Happiness	(6) Life Satisf.
Income Rank	0.310***	0.585***	0.306***	0.563***	0.268***	0.554***
	(0.032)	(0.056)	(0.036)	(0.063)	(0.032)	(0.056)
Income Rank * I{2001-2013}$^{(i)}$	0.090**	0.122**	0.087**	0.112*	0.098***	0.115**
	(0.037)	(0.055)	(0.043)	(0.061)	(0.037)	(0.055)
Income Rank * I{1997-2000}$^{(ii)}$	0.001		-0.022		0.006	
	(0.048)		(0.056)		(0.048)	
P-value (i)=(ii)	0.043		0.034		0.035	
Weights	No	No	Yes	Yes	No	No
Income Rank	Nation	Nation	Nation	Nation	County	County
Period	85-13	99-13	85-13	99-13	85-13	99-13
Observations	48,570	29,655	48,570	29,655	48,570	29,655

Notes: *** p<0.01, ** p<0.05, * p<0.1. Heteroskedasticity-robust standard errors in parenthesis. Each column corresponds to a separate OLS regression. *Happiness* and *Life Satisfaction* are responses to subjective questions where higher value denotes higher happiness/satisfaction, normalized to have mean 0 and standard deviation of 1. *Income Rank* denotes the position of the respondent's household, from 0 to 1, relative to all the other respondents for that year in the nation (columns (1) through (4)) or county (columns (5) and (6)). *I{2001–2013}* takes the value 1 for 2001—2013. *I{1997–2000}* takes the value 1 for 1997—2000. In columns (3) and (4), the regressions use population weights computed by the group in charge of conducting the survey. All regressions control for year dummies, age, age squared, a gender dummy, three education dummies, four dummies for marital status, four dummies for household size and three dummies for number of working household members. Data from the Norwegian Monitor Survey, which has been collected every other year in 1985–2013. See Table 1 for a summary of data definitions and Table 2 for descriptive statistics.

Table A.4: Robustness to Alternative Definitions of Income Rank

	(1) Happiness	(2) Happiness	(3) Happiness	(4) Happiness	(5) Happiness	(6) Happiness	(7) Happiness	(8) Happiness	(9) Happiness
Income Rank	0.311***	0.310***	0.331***	0.313***	0.313***	0.333***	0.291***	0.283***	0.305***
	(0.028)	(0.032)	(0.040)	(0.028)	(0.032)	(0.040)	(0.026)	(0.030)	(0.037)
Income Rank * I{2001-2013}$^{(i)}$	0.090***	0.090**	-0.004	0.090***	0.091**	-0.006	0.072**	0.081**	-0.003
	(0.032)	(0.037)	(0.051)	(0.032)	(0.037)	(0.052)	(0.031)	(0.036)	(0.048)
Income Rank * I{2001-2013} * I{Higher Internet}			0.217***			0.223***			0.168**
			(0.073)			(0.073)			(0.069)
Income Rank * I{1997-2000}$^{(ii)}$		0.001			0.001			0.024	
		(0.048)			(0.048)			(0.048)	
P-value (i)=(ii)		0.043			0.042			0.196	
Income Rank Definition	I	I	I	II	II	II	III	III	III
Observations	48,570	48,570	48,570	48,570	48,570	48,570	48,570	48,570	48,570

Notes: *** p<0.01, ** p<0.05, * p<0.1. Heteroskedasticity-robust standard errors in parenthesis. Each column corresponds to a separate OLS regression. *Happiness* is normalized to have mean 0 and standard deviation of 1, with higher values denoting higher happiness. *Income Rank* denotes the position of the respondent's household in the national income distribution, from 0 to 1. In columns (1) through (3), *Income Rank* is defined as in the baseline specification described in Table 1. In columns (4) through (6), the ninth bin (introduced in 1999) is merged with the eight bin before constructing *Income Rank*. In columns (7) through (9), *Income Rank* is constructed without the imputation of the within-bin ranks. *I{2001–2013}* takes the value 1 for 2001—2013. *I{1997–2000}* takes the value 1 for 1997—2000. *I{Higher Internet}* is a dummy variable that takes the value 1 if the respondent's predicted *Internet Access* is above the median value for a given year. All regressions control for year dummies, age, age squared, a gender dummy, three education dummies, four dummies for marital status, four dummies for household size and three dummies for number of working household members. Data from the Norwegian Monitor Survey, which has been collected every other year in 1985–2013. See Table 1 for a summary of data definitions and Table 2 for descriptive statistics.

Table A.5: Robustness to Alternative Definitions of Higher Internet

	(1) Happiness	(2) Happiness	(3) Happiness	(4) Happiness	(5) Happiness
Income Rank	0.331***	0.336***	0.382***	0.351***	0.334***
	(0.040)	(0.040)	(0.041)	(0.040)	(0.040)
Income Rank * I{2001-2013}$^{(i)}$	-0.004	0.016	-0.023	-0.010	0.007
	(0.051)	(0.050)	(0.053)	(0.053)	(0.052)
Income Rank * I{2001-2013} * I{Higher Internet}	0.217***	0.152**	0.209***	0.211***	0.187**
	(0.073)	(0.071)	(0.075)	(0.074)	(0.074)
I{Higher Internet} Definition	I	II	III	IV	V
Observations	48,570	48,570	48,570	48,570	48,570

Notes: *** p<0.01, ** p<0.05, * p<0.1. Heteroskedasticity-robust standard errors in parenthesis. Each column corresponds to a separate OLS regression. *Happiness* is normalized to have mean 0 and standard deviation of 1, with higher value denoting higher happiness. *Income Rank* denotes the position of the respondent's household in the national income distribution, from 0 to 1. *I{2001–2013}* takes the value 1 for 2001—2013. *I{Higher Internet}* is a dummy variable that takes the value 1 if the respondent's predicted *Internet Access* is above the median value for a given year. In column (1), *I{Higher Internet}* is defined as in the baseline specification described in Table 1. In columns (2) through (5), *I{Higher Internet}* is based on alternative definitions. In column (2), *I{Higher Internet}* is identical to the baseline definition except that it is based on responses to *Internet Access* for 1999 instead of 2001. In column (3), *I{Higher Internet}* is identical to the baseline definition except that it is based on responses to *Internet Access* for 1999–2013 instead of 2001. In column (4), *I{Higher Internet}* is identical to the baseline definition except that instead of splitting the sample within each year, we split the sample using the median value over the entire 1985–2013 period. In column (5), *I{Higher Internet}* is identical to the baseline definition except that it is constructed using a Probit model instead of an OLS model. All regressions control for year dummies, age, age squared, a gender dummy, three education dummies, four dummies for marital status, four dummies for household size and three dummies for number of working household members. Data from the Norwegian Monitor Survey, which has been collected every other year in 1985–2013. See Table 1 for a summary of data definitions and Table 2 for descriptive statistics.

Table A.6: Descriptive Statistics by Year

	Female	Age	Married	College	Oslo	Density
1985	0.499	40.861	0.721	0.231	0.118	185.719
	(0.011)	(0.361)	(0.010)	(0.010)	(0.007)	(9.388)
1987	0.505	40.733	0.731	0.249	0.119	188.673
	(0.010)	(0.325)	(0.009)	(0.009)	(0.006)	(8.338)
1989	0.508	39.759	0.684	0.280	0.108	176.713
	(0.010)	(0.318)	(0.009)	(0.009)	(0.006)	(7.789)
1991	0.501	39.352	0.622	0.270	0.114	182.496
	(0.010)	(0.333)	(0.010)	(0.009)	(0.006)	(8.063)
1993	0.494	39.989	0.527	0.288	0.122	192.982
	(0.010)	(0.304)	(0.010)	(0.009)	(0.006)	(8.182)
1995	0.490	40.906	0.483	0.311	0.119	188.013
	(0.010)	(0.334)	(0.010)	(0.009)	(0.006)	(8.371)
1997	0.517	41.212	0.488	0.360	0.113	181.017
	(0.008)	(0.245)	(0.008)	(0.007)	(0.005)	(6.358)
1999	0.537	42.492	0.486	0.394	0.117	186.894
	(0.008)	(0.253)	(0.008)	(0.008)	(0.005)	(6.597)
2001	0.554	44.571	0.496	0.412	0.114	183.217
	(0.008)	(0.252)	(0.008)	(0.008)	(0.005)	(6.542)
2003	0.542	45.205	0.510	0.421	0.114	185.332
	(0.008)	(0.257)	(0.008)	(0.008)	(0.005)	(6.626)
2005	0.559	47.563	0.539	0.461	0.112	179.919
	(0.008)	(0.261)	(0.008)	(0.008)	(0.005)	(6.674)
2007	0.531	51.840	0.573	0.489	0.117	188.027
	(0.008)	(0.253)	(0.008)	(0.008)	(0.005)	(6.727)
2009	0.561	48.434	0.546	0.521	0.116	187.730
	(0.008)	(0.293)	(0.008)	(0.008)	(0.005)	(7.004)
2011	0.531	50.763	0.579	0.598	0.124	198.645
	(0.008)	(0.249)	(0.008)	(0.008)	(0.005)	(6.850)
2013	0.535	47.821	0.503	0.577	0.137	215.520
	(0.008)	(0.304)	(0.008)	(0.008)	(0.006)	(7.310)

Notes: Data from the Norwegian Monitor Survey (48,570). *Female* takes the value 1 if the respondent is female. *Age* is the age in years. *Married* takes the value 1 if married. *College* takes the value 1 for College graduates. *Oslo* takes the value 1 for Oslo residents. *Density* corresponds to the population per square kilometer in the county of residence (Statistics Norway, 2011). Heteroskedasticity-robust standard errors in parenthesis.

Table A.7: Income Inequality by Year

	Administrative Data		Norwegian Monitor Survey						
	Gini Index		*Absolute Income*			*Income Rank*		*I{Higher Internet}*	
Year	*Gross*	*Net*	Mean	SD	CV	Mean	SD	Mean	SD
1985	40.37	21.00	5.10	2.96	0.581	0.50	0.29	0.50	0.50
1986	40.47	21.00							
1987	40.84	21.00	5.76	3.24	0.563	0.50	0.29	0.50	0.50
1988	40.72	21.10							
1989	40.07	22.80	5.69	3.09	0.543	0.50	0.29	0.50	0.50
1990	40.58	21.70							
1991	41.07	21.90	5.53	3.32	0.600	0.50	0.29	0.50	0.50
1992	42.59	22.30							
1993	43.98	22.90	4.44	2.40	0.540	0.50	0.29	0.50	0.50
1994	44.35	24.10							
1995	44.18	23.60	4.22	2.47	0.586	0.50	0.29	0.50	0.50
1996	44.36	24.50							
1997	44.72	24.90	4.79	2.48	0.518	0.50	0.29	0.50	0.50
1998	43.59	23.80							
1999	43.76	24.20	5.18	2.71	0.524	0.50	0.29	0.50	0.50
2000	45.57	26.20							
2001	43.51	22.90	5.32	2.79	0.524	0.50	0.29	0.50	0.50
2002	45.30	26.40							
2003	46.08	27.40	5.86	3.02	0.516	0.50	0.29	0.50	0.50
2004	46.90	28.30							
2005	50.62	32.70	6.06	3.11	0.513	0.50	0.29	0.50	0.50
2006	44.67	24.30							
2007	45.68	25.20	6.40	3.27	0.511	0.50	0.29	0.50	0.50
2008	44.90	24.80							
2009	44.68	24.10	6.70	3.44	0.513	0.50	0.29	0.50	0.50
2010	45.13	24.50							
2011	45.14	24.70	7.45	3.80	0.510	0.50	0.29	0.50	0.50
2012	45.22	24.90							
2013	45.44	25.00	7.72	4.52	0.585	0.50	0.29	0.50	0.50

Notes: The *Gini Index* are inequality indeces. The *Gross Gini Index* is based on gross household income, while the *Net Gini Index* is based on the net household income. Both of these measures were obtained from the Chartbook of Economic Inequality and are based on administrative data. The rest of the outcomes are based on data from the Norwegian Monitor Survey (N=48,570). *Absolute Income* is the gross household income measured in hundreds of thousands of Kroner, and converted to 2013 prices with the consumer price index from Statistics Norway. *SD* stands for standard deviation, and *CV* stands for coefficient of variation (i.e., the ratio between the standard deviation and the mean). *Income Rank* is respondent's position in the distribution of *Absolute Income* in a given year, and *I{Higher Internet}* is a dummy variable that takes the value 1 if the respondent's predicted *Internet Access* is above–median – see Table 1 for more detailed data definitions.

Table A.8: Auxiliary Regression Results: Predictors of Internet Access

Dependent Variable: *Internet Access * 100*			
Female	-9.294***	Number of HH Members (omitted: 1)	
	(1.391)	2	8.449***
Age	0.479*		(3.029)
	(0.266)	3	13.561***
Age Squared	-0.013***		(3.334)
	(0.003)	4	12.808***
Education (omitted: Primary School)			(3.656)
Middle School	11.688***	5+	16.999***
	(2.920)		(3.945)
High School	18.881***	Number of HH Workers (omitted: 0)	
	(2.732)	1	9.448***
College	30.729***		(2.609)
	(2.731)	2	18.859***
Marital Status (omitted: Married)			(2.794)
Cohabitant	-10.327***	3+	24.909***
	(2.218)		(4.055)
Single	-7.443**	Constant	33.856***
	(3.327)		(8.607)
Separated/Divorced	-4.310		
	(3.363)		
Widowed	2.287		
	(4.063)		

Notes: $N = 3,931$. *** p<0.01, ** p<0.05, * p<0.1. Heteroskedasticity-robust standard errors in parenthesis. Coefficients of an OLS regression. *Internet Access*100* takes the value 100 if the respondent has Internet access at home and 0 otherwise. Data from the Norwegian Monitor Survey for the year 2001. The average of the dependent variable is 60.75 percentage points.

B A Simple Model of the Effects of Income Transparency Through Income Comparisons

This section provides a simple model to illustrate how an increase in income transparency can increase the gradient between utility and income rank. This model is not intended to capture all the different forms that income comparisons can take. Instead, it is based on two specific channels (social-esteem and self-esteem) that have received some attention in the economics literature.

B.1 The Model

There is a continuum of individuals with a non-degenerate income distribution, where r_i^{true} denotes the true relative position in the income distribution of individual i.

Intrinsic Utility. We assume that intrinsic utility from income is a linear function of the true income rank of the individual:

$$U_i^{intrinsic} = \eta_0 \cdot r_i^{true} \tag{B.1}$$

Using this particular functional form for the intrinsic utility function is made just to simplify the notation – the intuitions would still apply under more standard functional forms, such as making intrinsic utility equal to the logarithm of absolute income.

We model income comparisons through two distinct channels: self-image utility and social-image utility. These two sources of utility depend on the social interactions with other individuals from the same population. The incomes of the individuals involved in each interaction are observable with some exogenous probability $\nu \in [0, 1]$. This parameter ν is a reduced-form representation of the degree of income visibility.[58]

Social-Image Utility. Each individual is paired with a random individual from the same population. A third party, the allocator, allocates some social-image utility worth $\eta_1 > 0$. The allocator wants to give the social-image utility to the individual in the pair with the higher income. With probability ν, the allocator can observe the incomes of the two individuals in the pair, in which case the allocator gives the social-esteem to the individual with the higher income. With probability $1-\nu$ the allocator cannot observe incomes, in which case the allocator simply randomizes who gets the social-esteem. As a result, the ex-ante utility from social-image is the following:

[58] We use a unique ν to simplify the notation. In reality, there should be at least two ν's: one relevant for the formation of self-image and another for the formation of social-image. For example, it is probably easier to observe information to infer one's position in the income distribution, for which it suffices to have access to aggregate income statistics, than to observe the income of a particular individual.

$$U_i^{social} = \nu \cdot r_i^{true} \cdot \eta_1 + (1-\nu) \cdot \frac{1}{2} \cdot \eta_1 \tag{B.2}$$

By taking the derivative of (B.2) with respect to r_i^{true}, we obtain the gradient between social-image utility and income. By taking an additional derivative with respect to ν, we show that this gradient increases with income visibility:

$$\frac{\partial^2 U_i^{social}}{\partial r_i^{true} \partial \nu} = \eta_1 > 0 \tag{B.3}$$

The intuition behind this result is simple. An increase in visibility would make an individual with a below-median income worse off, because with a higher probability her peers would observe her income and learn that she is poorer than they would have thought otherwise. On the other hand, a visibility increase would make an individual with above-median income better off, because with a higher probability the peers would observe her income and learn that she is richer than they would have thought otherwise.

Self-Image Utility. Self-image utility is similar to social-image utility, only that the individual is comparing herself to others rather than being compared by the allocator. The individual is paired with another individual randomly chosen from the population. The individual must decide whether she deserves some self-image utility worth $\eta_2 > 0$. Whether the individual feels deserving of or not depends on whether she thinks she is richer than the individual she is paired with. With probability ν, she can observe the actual income of the peer, in which case she gets the self-image utility if and only if her income is higher than the income of the peer. From an ex-ante perspective, this happens with probability r_i^{true} (i.e., equal to the probability of being paired with someone poorer). With probability $1-\nu$, the income of the paired peer is not observable, in which case she will get a fraction of the self-image utility, equal to the perceived probability of being richer than the other individual. Let r_i^{self} be this prior perceived probability. We let the prior beliefs be heterogeneous and (possibly) correlated to the actual income ranks: $r_i^{self,prior} = \theta_0 + \theta_1 \cdot r_i^{true}$.

Let $r_i^{self,post}$ be the posterior belief about the own income rank. If the individual learns rationally, the expectation of the posterior beliefs should be as follows:

$$r_i^{self,post} = \nu \cdot r_i^{true} \cdot \eta_2 + (1-\nu) \cdot \theta \cdot r_i^{true} \tag{B.4}$$

Following the evidence on the middle class bias (Cruces, Perez-Truglia and Tetaz, 2013), we assume $\theta_1 < 1$.[59] As a result, the ex-ante expected utility from self-image is:

[59] To figure out whether θ_1 is equal, above or below 1, we can test an intermediate prediction of the model: $\frac{\partial^2 r_i^{self,post}}{\partial r_i^{true} \partial \nu} = (1-\theta_1)$. Thus, if $\theta < 1$, we would predict that higher income visibility increases the gradient between self-perceived income rank and actual income rank.

$$U_i^{self} = \nu \cdot r_i^{true} \cdot \eta_2 + (1-\nu) \cdot \left(\theta_0 + \theta_1 \cdot r_i^{true}\right) \cdot \eta_2 \tag{B.5}$$

By taking the derivative of (B.5) with respect to r_i^{true}, we obtain the gradient between self-image utility and income. By taking an additional derivative with respect to ν, we can show that this gradient is increasing in income visibility:

$$\frac{\partial^2 U_i^{self}}{\partial r_i^{true} \partial \nu} = \eta_2 \cdot (1-\theta_1) > 0 \tag{B.6}$$

The intuition for this result is also straightforward. When incomes are more easily observable, poor individuals learn that they are actually poorer than they thought, thus losing self-image utility; and rich individuals learn that they are actually richer than they thought, thus gaining self-image utility.

Finally, we can also explore the predictions of this model for the effect of income transparency on average well-being. Regarding social-image, it is straightforward to check that the average effect is zero. Intuitively, increasing visibility transfers social-image from poor to rich individuals, but no social-image utility gets created or destroyed in the process. Regarding self-image, it is straightforward to check that the average effect of higher visibility depends on whether θ_0 is above or below $\frac{1}{2}$. Intuitively, if $\theta_0 > \frac{1}{2}$, it means that on average individuals were over-estimating their own position in the income distribution. Since the higher transparency corrects this systematic bias, there is a net loss in utility from self-image. Similarly, if if $\theta_0 < \frac{1}{2}$, higher visibility would lead to an increase in average happiness; and if the average bias in self-perceived income rank was zero ($\theta_0 = \frac{1}{2}$), then higher income transparency would have no effect on the average utility from self-image.

B.2 Back-of-the-Envelope Calculations

To simplify the notation, we focus on the most obvious case of $\theta_1 = 0$: i.e., if all incomes were completely unobservable ($\nu = 0$), then self-perceptions about income rank would be orthogonal to actual income ranks. To obtain the overall utility, we must add up the three sources of utility: $U_i = U_i^{intrinsic} + U_i^{social} + U_i^{self}$. We add ((B.1), (B.2) and (B.5)) up and then re-arrange as follows:

$$U_i = (\beta_1 + \beta_2 \cdot \nu) \cdot r_i^{true} + \epsilon_i, \tag{B.7}$$

where $\beta_1 = \eta_0$, $\beta_2 = \eta_1 + \eta_2$ and $\epsilon_i = (1-\nu) \cdot \left(\frac{1}{2} \cdot \eta_1 + \theta_0 \cdot \eta_2\right)$. Note that β_1 measures the intrinsic utility from income, while $\beta_2 \cdot \nu$ measures the utility from income through income comparisons (in this model, self-image and social-image). As a result, $\frac{\beta_2 \cdot \nu}{\beta_1 + \beta_2 \cdot \nu}$ measures the value of income comparisons relative to intrinsic consumption. Intuitively, in a world where all incomes are unobservable ($\nu = 0$), an increase in one's income cannot increase the utility from self-image or social-image. On the other extreme, increasing one's income has the highest possible effect on self-image and social-image when incomes are perfectly visible ($\nu = 1$).

Let $\bar{\nu}_{t<2001}$ and $\bar{\nu}_{t\geq 2001}$ denote the visibility before 2001 and after 2001, respectively. Note that $\bar{\nu}_{t<2001}$ must be greater than zero, because even when the tax records were private, individuals could use other means to learn about the incomes of social contacts and about the income distribution. For instance, individuals could learn about the income distribution from school, from the media, or by talking with others about wages and consumption. Also, individuals reveal their own income to their social contacts, or signal it through conspicuous consumption. Similarly, $\bar{\nu}_{t\geq 2001}$ must be lower than 1, because even when the tax records were easily accessible online, there was still a small cost in attention, memory, and time to search those records. Thus, individuals did not search for the incomes of everyone with whom they interacted. Also, the online tax lists probably could not make incomes visible among total strangers – to find out someone's income in the tax records, you need to know the name of that person.

We can measure the importance of income comparisons as $s_t = \frac{\beta_2 \cdot \nu_t}{\beta_1 + \beta_2 \cdot \nu_t}$, which is the share of the happiness-income gradient that can be explained by income comparisons. Combining the regression model (1) with (B.7) and re-arranging:

$$s_{t<2001} = \frac{\frac{\alpha_2}{\alpha_1}}{\frac{\bar{\nu}_{t\geq 2001} - \bar{\nu}_{t<2001}}{\bar{\nu}_{t<2001}}}, \quad s_{t\geq 2001} = \frac{1 + \frac{1}{\frac{\bar{\nu}_{t\geq 2001} - \bar{\nu}_{t<2001}}{\bar{\nu}_{t<2001}}}}{1 + \frac{1}{\frac{\alpha_2}{\alpha_1}}} \tag{B.8}$$

The value of income comparisons depends on two parameters: $\frac{\alpha_2}{\alpha_1}$ and $\frac{\bar{\nu}_{t\geq 2001} - \bar{\nu}_{t<2001}}{\bar{\nu}_{t<2001}}$. The first parameter is the proportional growth in the happiness-income gradient as a result of the change in disclosure in 2001. This is the main parameter estimated in this study. The second parameter is the effect of the change in disclosure on income visibility.[60] Because we do not have an estimate of this second parameter, we present results assuming different values for it.

First, we estimate a lower bound for the value of income comparisons. Note that $s_{t\geq 2001}$ is strictly increasing in $\bar{\nu}_{t<2001}$. Thus, by assuming $\bar{\nu}_{t<2001} = 0$, we can estimate a lower bound on $s_{t\geq 2001}$. This is a conservative lower bound because it is highly unlikely that income information was completely private before 2001. Assuming that $\bar{\nu}_{t<2001} = 0$ implies that, after 2001, income comparisons explain at least 22% of the happiness-income gradient (i.e., $s_{t\geq 2001} = \frac{0.090}{0.311+0.090}$) and 17% of the life satisfaction-income gradient (i.e., $s_{t\geq 2001} = \frac{0.122}{0.585+0.122}$). These results suggest that the value of income comparisons is bound to be large.

Second, we estimate an upper bound for the value of income comparisons. As $\frac{\bar{\nu}_{t\geq 2001} - \bar{\nu}_{t<2001}}{\bar{\nu}_{t<2001}}$ approaches $\frac{\alpha_2}{\alpha_1}$ from above, both $s_{t<2001}$ and $s_{t\geq 2001}$ converge to 1. That is, a change in visibility

[60]The formula for $s_{t<2001}$ from (B.8) can be interpreted as a Wald estimate: i.e., the ratio between the effect on the happiness-income gradient (i.e., the reduced form effect) and the effect on visibility (i.e., the first stage effect).

of $\frac{\alpha_2}{\alpha_1}$ would imply that income comparisons explain the entire relationship between income and well-being. In the case of happiness, we would need to assume that visibility increased by 29% as a result of the change in disclosure. In the case of life satisfaction, we would have to assume that visibility increased by 21%. Given all the evidence about the widespread use of the search tool, it seems somewhat unlikely that the publication of tax records increased income visibility by just 21%. In this sense, our results suggest that income comparisons may not be the only factor mediating the effect of income on well-being.