Online Resource 1

for

Differential Survival in Europe and the United States:

Estimates Based on Subjective Probabilities of Survival

Adeline Delavande
RAND Corporation, Nova School of Business and Economics, and University of Essex

Susann Rohwedder
RAND Corporation and NETSPAR

April 2010

Supplemental Material for VERIFYING THE VIABILITY OF OUR APPROACH: Comparison of Estimates of Differential Survival based on Actual Survival with Estimates based on Subjective Probabilities of Survival

In the first part of our study we investigate whether subjective probabilities of survival generate estimates of differentials in survival by socioeconomic status similar to those produced by actual survival data. In addition to wealth (which we present in our main paper), we also analyze two other measures of socioeconomic status: income and education.

Our measure of income sums all sources of income received during the last calendar year for the respondent and spouse if married. These sources include earnings and other income from investments, pensions, annuities, Social Security, transfers, and benefits.[1] The information on income is measured at the same time (i.e., in the same survey wave) as the subjective probability of survival to P75. We define income terciles over all respondents interviewed in the same wave, stratifying by marital status (singles vs. couples) and age category (60-64 and 65-69).

The following analyses are based on the analytical sample described in the main paper in the section entitled "HRS Data."

Non-parametric estimates of differential survival. Supplement Figure 1 presents the estimates of the kernel regressions showing actual and subjective survival conditional on income. Because couples have much higher levels of income than singles, we run the kernel regressions separately for each group. The figure focuses on couples, because the vast majority of our sample lives in a couple household.[2] In keeping with our findings for wealth, the relationship between actual survival and income follows a similar gradient as the same relationship using subjective survival, albeit a bit flatter at the highest income levels. Looking at average survival by categorical variables of income and education offers another non-parametric way of assessing the validity of our approach. Supplement Figure 2 presents the percentage alive at age 75 alongside the average of P75 by income terciles, as well as by education levels. For both measures, the slopes of actual survival and P75 are strikingly similar. This indicates that P75 captures the differentials in survival rather accurately.

[1] We use the RAND HRS data, version H. The RAND HRS Data file is an easy-to-use longitudinal data set based on the HRS data. It was developed at RAND with funding from the National Institute on Aging and the Social Security Administration. For further details of the definitions of income, see the RAND HRS documentation (http://www.rand.org/labor/aging/dataprod/randhrsh.pdf), variables hwatota and hwitot.

[2] Recall that older spouses of HRS-age-eligible respondents make up a large fraction of our analytical sample of persons who reach their target age by 2006.

Parametric estimates of differential survival. Supplement Tables 1 and 2 present the results of our parametric estimates for income and education respectively.[3] In addition to the variables of interest, we include categorical variables for age at the time that respondents were asked about their expectations of survival and for sex as independent variables. The estimated coefficients for income and education are quite close. When testing the hypothesis that the coefficients associated with the socioeconomic status variables in the model of actual survival are equal to those in the model of subjective survival, the resulting values of the test statistics are 2.31 for income and 1.70 for education, indicating that for each specification we cannot reject the null hypothesis at the 5 percent significance level. These results suggest once again that subjective probabilities of survival provide a suitable alternative for estimating differential survival by socioeconomic characteristics.

Robustness Checks. We verify that our validation results are robust to the econometric specification. Using non-linear least squares rather than the quasi-maximum likelihood method yields essentially the same results (see Supplement Table 3 for wealth). In addition, our results are robust to using a different distributional assumption for *G*. For example, using a normal rather than a logistic distribution in equations (3) and (4) presented in the main paper yields very much the same results again: the resulting coefficients associated with the second and third wealth tercile are 0.158 and 0.281 respectively in the specification using actual survival as dependent variable, and 0.130 and 0.265 respectively in the specification using the subjective probability of survival as dependent variable. We decided to present the logistics specification in this paper because the coefficients on wealth can be directly interpreted as log odds ratios.

Measurement error in subjective probabilities of survival deserves more detailed attention. Rounding to the nearest 5 percent and providing focal answers at 0, 50, and 100 percent are common patterns of answers to subjective probability questions. One concern about our methodology is whether the tendency to provide focal answers varies systematically by socioeconomic status. For example, if many respondents in the lowest wealth tercile do not know what their chance of survival is and just answer “50 percent,” the average subjective probability of survival may be biased for this group as opposed to the other wealth terciles. That would affect the estimates of differential survival.

[3] See our main paper for details on the estimation strategy.

Focal answers of 0 and 100 percent are different from answers of 50 percent in that they most likely convey that these individuals consider their chances of surviving to the target age to be either extremely low or very high. In our analytical sample, the fraction of those who answer 50 and 100 percent is similar across wealth terciles (between 20 and 23 percent),[4] but there is a difference in the prevalence of zeros (10 percent in the lowest wealth tercile and 2-3 percent in the higher terciles). This is consistent with the fact that, on average, respondents in the lowest wealth tercile do not survive as long as other respondents. One might expect similarly noticeable differences across wealth terciles for the 100-percent answers, but this is not the case.

Even though the overall fraction of 50-percent answers does not vary across wealth terciles, the fraction of those 50-percent answers that reflects "don't knows" could still vary by wealth tercile. This would potentially introduce bias into the estimates based on subjective probabilities of survival. The fact that these estimates match well with the results based on actual survival makes this unlikely. But we nevertheless investigate this possibility.

We use variables shown to correlate strongly with subjective probabilities of survival to impute P75 and replace the 50-percent answers with the imputations. The set of covariates for imputation includes basic demographics, a number of health-related variables, and parental mortality.[5,6] Once again, we find that the coefficients on the wealth terciles are very similar in both the logit regression on actual survival at 75, and the quasi-maximum likelihood estimator on P75 where we replaced any original 50-percent answers with the imputation (tables not shown).

In our analytical sample, item non-response to the subjective probabilities is very low, but it is not randomly distributed by survival status or socioeconomic characteristics. Respondents with lower socioeconomic status and respondents who died by age 75 were less likely to respond to P75. To deal with this issue, we impute the missing P75 using the same set

[4] Using a simple ANOVA test, we cannot reject the hypothesis that the average fraction of those answering 50 and 100 percent is equal across wealth terciles.

[5] The health variables include information on drinking alcohol, smoking, the number of chronic conditions, self-rated health, number of Activities of Daily Living limitations, and Body Mass Index.

[6] Different interpretations of the 50% would suggest slightly different implementations of the imputation. If one believes that respondents who answer 50% merely express epistemic uncertainty, one would want to exclude them from the regression to produce the imputations. However, if one believes that the 50s are informative, one would keep them. The truth is probably a mix of the two, but we have no way in our data to separate respondents who truly believe they have about a 50% chance of survival to their target age from respondents who are simply uncertain. We follow both strategies and find that the results are not affected by this issue. In HRS 2006, respondents who answer 50% are asked a follow-up question: whether they just do not know their survival probabilities or whether their belief is really that the chances are about 50 percent. It turns out that the fraction of 50s being simply uncertain is high – a little over 60 percent. Unfortunately, we do not have this information for the earlier waves of HRS, which provide the baseline observations in our analysis.

of covariates described above. Again, we find that the estimations based on subjective and actual survival produce very similar coefficients on the wealth terciles (tables not shown).

Supplemental Material for APPLYING OUR APPROACH: A COMPARISON OF DIFFERENTIAL SUBJECTIVE SURVIVAL BY WEALTH IN EUROPE AND THE UNITED STATES

Can Heterogeneity in the Distribution of Wealth between Countries Explain the Variation in Differential Survival by Wealth?

In our main paper, we find that there is substantial heterogeneity in subjective survival by wealth terciles across European countries. This heterogeneity could simply reflect heterogeneity in the wealth distribution. For example, Figure 4 in the main paper shows that the odds ratio for the second wealth tercile is higher in Germany than in Austria. Is this due to the fact that there is greater dissimilarity between the first and second wealth terciles in Germany than in Austria? Supplement Table 4 summarizes the distribution of wealth within each country. It shows the ratio of the median wealth of each wealth tercile to the median of the lowest wealth tercile. These ratios suggest that heterogeneity in the wealth distributions cannot solely explain the heterogeneity in differential survival. Germany and Austria have a similar relative difference of wealth between the first and the second wealth terciles. The same conclusions hold when considering additional percentiles of the wealth distributions within wealth terciles.

Estimates of Differential Subjective Survival by Income across Countries

We define income terciles separately within each country, by marital status (single/couple) and age band (50-58; 59-65). To address issues potentially arising from focal and missing answers, we present results with imputations for missing subjective probabilities and answers of 50 percent. Supplement Figure 3 shows the exponential of the coefficients attached to the middle and highest income terciles. This can be interpreted as the odds ratio of survival

compared with the lowest income tercile. The largest gradients are found for France, Sweden, and England. The Netherlands and Spain have the smallest gradients. The coefficients for the low-gradient countries are statistically significantly different (at 10%) from those for the high-gradient countries when testing for the joint hypothesis of equality of the coefficients associated with the second and third income terciles for each pair of countries. Note that for Austria, Spain, Italy, and the Netherlands, the odds ratio of the middle income tercile in comparison with the lowest equals one, or even a little less than one. This suggests that these two income groups have no difference in survival.[7]

Again, heterogeneity in the distribution of income could explain the different results across countries. But as we concluded for wealth, heterogeneity in income inequality across countries cannot be the sole explanation for our results. For example, Supplement Table 4 shows that there is more inequality in income in Italy than in France (the ratio of median income of the highest income tercile to that of the lowest tercile is 20 in Italy, compared with 6 in France). But the gradient of subjective survival by income is steeper in France than in Italy.

Across Europe, there is less inequality in differential subjective survival by income than by wealth. The only country in our comparison for which this did not hold true was the United States. Several reasons may explain this, and why the differences between the income and wealth gradients may vary by country. The first thing to note is that wealth and income measure very different concepts: wealth is a stock measure that captures asset accumulation over a long time horizon while current income is a flow measure that can change from year to year and tends to drop when a person retires. The classification of households by wealth will therefore differ – in some cases substantially – from that of income, and the extent of the differences is influenced by variation in institutions across countries. In the age groups we consider, among respondents who work, current income comes primarily from wages; and among respondents who are retired, from retirement benefits. Because the proportion of retirees increases with age, income tends to decrease with age in our data. However, the probability of surviving until age 75, conditional on being currently alive, increases with age. Combined, these two effects lead mechanically to a relatively flat impact of income on survival to age 75. Controlling for age in our regression does

[7] The coefficients for the countries where the difference is slightly less than one are negative, but close to zero, and not statistically significantly different from zero at any conventional level.

not fully account for this, because income tends to decrease with age within the age categories as well.

This mechanical effect will impact the countries we study differently because the proportion of respondents retired varies greatly across countries. For example, only 4% of the 51-54 year olds and 37% of the 59-61 year olds are retired or disabled in Germany, compared with 14% and 62% of the same groups in Austria.

Individuals in a given wealth tercile are not necessarily in the equivalent income tercile.[8] Again, the retirement decision may be a factor here. The decision to retire (which in our data affects current income) depends on many factors, but in particular on health and wealth. People who retire at a young age may either be in poor health and no longer able to work, or they may be wealthy and healthy, with no need for wage income. In the former case, poor health may also have limited the total amount of wealth people have accumulated. This disparity is particularly marked in countries where differential survival by wealth and income differs greatly. For example, in the Netherlands, among respondents in the lowest income tercile, 35% are in the lowest wealth tercile, 30% in the second wealth tercile, and 35% in the highest wealth tercile. However, in England where differences in survival by income and wealth are more comparable, these numbers are 53%, 29%, and 17% respectively.

Finally, local factors such as taxation, the health care system, the pension replacement rate, and progressiveness of the applicable pension system may play a role in an individual's decision to retire at a certain age, the incentive to accumulate assets throughout one's working life, and the differential impact of wealth and income on survival. All of these factors vary greatly in the countries we consider.[9]

Estimates of Differential Subjective Survival by Education across Countries

To facilitate comparisons by education across countries with different educational systems, we organize respondents' reports of their highest degree into three categories of schooling: (1) less-than-secondary, (2) secondary, and (3) tertiary. Supplement Table 5 presents the resulting distribution of the education categories by country, also stratified by wealth tercile. Note the heterogeneity in the distribution of education across the various countries we consider.

[8] Differential survival by income and wealth would be identical if people in a given wealth tercile were in the equivalent income tercile.

[9] For example, the pension of a worker with average lifetime gross earnings replaces 45.8% of pre-retirement earnings in Germany, but 78.8% of pre-retirement earnings in Italy (OECD, 2005).

For example, in Germany 12% of the respondents have less than secondary schooling, compared with 75% in Spain.

We present estimation results with imputations for missing probabilities and 50-percent answers. Supplement Figure 4 shows the exponential of the coefficients attached to secondary and tertiary education. This can be interpreted as the odds ratio between subjective survival at those levels and subjective survival for those with less-than-secondary schooling.

Like differential survival by income, differential survival by education also differs from that by wealth, although education and wealth share similar magnitudes. By education, the United States, Austria, and Italy show the largest gradients in differential survival, while the Netherlands, France, and Belgium show the smallest. The coefficients for the low gradient countries are statistically significantly different (at 10%) from those for the high gradient countries when testing for the joint hypothesis of equality of the coefficients associated with secondary and tertiary education for each pair of countries. The difference between Europe and the United States is very large.

Again, there may be various explanations as to why differential survival by education differs from that of wealth. Factors such as access to education, returns to schooling, and each nation's system of taxation may play a role. Moreover, the relationship between wealth tercile and education is heterogeneous across countries. Supplement Table 5 presents the relationship between wealth terciles and education levels by country. It shows, for example, that in Spain and the Netherlands—where differential survival by education is flatter than that by wealth—a high proportion of people in the high wealth tercile (61 and 41% respectively) have less than a secondary education.

Estimates of Differential Subjective Survival by Sex and Socioeconomic Status across Countries

To investigate whether differential subjective survival varies by sex, we estimate regressions similar to those described above, but interact our measures of socioeconomic status with an indicator variable for sex. We find that overall, the coefficients associated with the measures of socioeconomic status for men are not systematically significantly different at the 10% level from those for women (tables not shown). Only in a few countries are there

exceptions for the interactions of sex with wealth and education, which we detail below. In the case of income, we never reject equality of the coefficients for men and women at the 10% significance level.

Wealth. In the case of wealth, we find differences by sex in the United States, where the gradient in subjective probabilities of survival is steeper for women than for men. In Sweden, we find similar differentials in subjective survival for men and women in the highest wealth tercile, compared with those in the lowest wealth tercile. But the differential is larger for women in the second wealth tercile than for men in that tercile.

Education. We find sex differences in subjective survival by education in the United States and Belgium, where the gradient is steeper for women than for men. For Italy, we find that having a tertiary as opposed to a primary education is associated with similar differentials in subjective survival for men and women. However, the differential for individuals with a secondary education is larger for women than for men.

REFERENCES

OECD. (2005). *Pensions at Glance 2005*. Paris, France: Organisation for Economic Co-operation and Development.

RAND HRS Data, Version H. 2008. Produced by the RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration. Santa Monica, CA

SUPPLEMENT TABLES

Supplement Table 1: Estimates of differential survival by income tercile

	Logit on actual survival to 75		Quasi maximum-likelihood on subjective survival to 75	
	Coefficients	P-value	Coefficients	P-value
Income terciles				
Lowest	(ref)		(ref)	
2nd	**0.381**	0.013	**0.305**	0.001
Highest	**0.823**	0.000	**0.576**	0.000
Age at baseline				
61	-1.043	0.000	-0.290	0.083
62	-0.731	0.000	0.088	0.338
63	(ref)		(ref)	
64	0.354	0.134	0.083	0.501
65	0.048	0.840	0.285	0.022
66	-0.150	0.590	0.289	0.081
Female	0.048	0.741	0.056	0.502
Constant	0.671	0.000	0.229	0.008
N	*1,219*		*1,219*	

Supplement Table 2: Estimates of differential survival by education

	Logit on actual survival to 75		Quasi maximum-likelihood on subjective survival to 75	
	Coefficients	P-value	Coefficients	P-value
Less than High School	(ref)		(ref)	
High School & GED	**0.162**	0.292	**0.146**	0.115
Some college	**0.421**	0.027	**0.380**	0.000
College graduate	**0.818**	0.000	**0.540**	0.000
Age at baseline				
61	-1.074	0.000	-0.310	0.061
62	-0.677	0.000	0.125	0.177
63	(ref)		(ref)	
64	0.354	0.133	0.088	0.481
65	0.088	0.709	0.316	0.010
66	-0.129	0.642	0.308	0.066
Female	0.026	0.857	0.040	0.636
Constant	0.792	0.000	0.307	0.001
N	*1,219*		*1,219*	

Supplement Table 3: Estimates of differential survival by wealth tercile using non-linear least squares

	Logit on actual survival to 75		Non-linear least squares on subjective survival to 75	
	Coefficients	P-value	Coefficients	P-value
Wealth terciles				
Lowest	(ref)		(ref)	
2nd	**0.256**	0.094	**0.207**	0.014
Highest	**0.466**	0.003	**0.421**	0.000
Age at baseline				
61	-1.041	0.000	-0.294	0.045
62	-0.705	0.000	0.097	0.299
63	(ref)		(ref)	
64	0.333	0.156	0.062	0.605
65	0.082	0.724	0.302	0.025
66	-0.147	0.594	0.094	0.067
Female	-0.004	0.978	0.015	0.855
Constant	0.830	0.000	0.324	0.000
N	*1,219*		*1,219*	

Supplement Table 4: Median wealth of each wealth tercile relative to the median of the lowest wealth tercile (first panel); and median income of each income tercile relative to the median of the lowest income tercile (second panel) by countries for respondents age 51 to 65, weighted

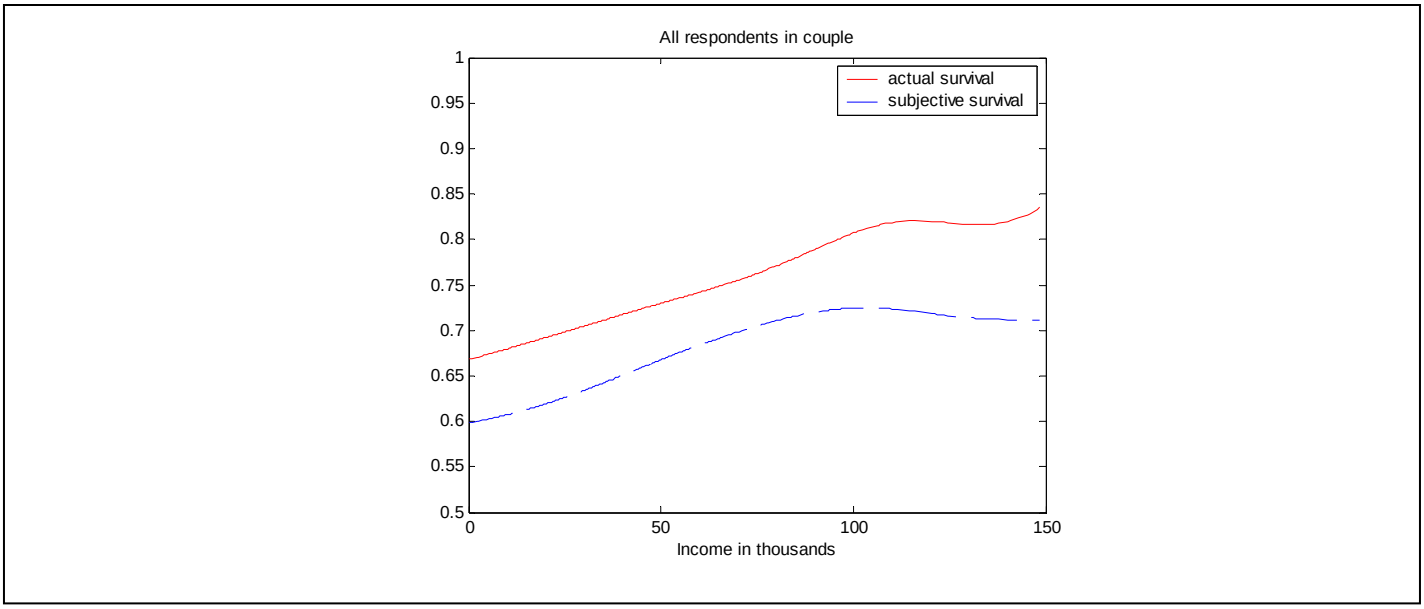
Wealth terciles	Austria	Germany	Sweden	Nether-lands	Spain	Italy	France	Denmark	Belgium	England	Europe all	U.S.
Lowest	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Second	26.6	27.3	11.1	16.7	5.5	11.9	10.0	4.7	2.9	4.3	12.7	7.6
Highest	72.4	82.6	34.0	43.4	17.4	32.0	29.6	12.1	8.6	9.1	33.8	27.8
Income terciles												
Lowest	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Second	4.7	4.0	2.1	3.8	6.3	8.7	2.6	2.4	3.0	2.5	3.6	3.0
Highest	10.2	9.0	3.4	7.9	16.6	20.2	6.0	3.9	8.3	4.6	8.2	6.1

Supplement Table 5: Distribution of education – all respondents age 51 to 65 by country and by wealth terciles, weighted

	Wealth tercile					Wealth tercile			
Education level	Lowest	2nd	Highest	All	Education level	Lowest	2nd	Highest	All
Austria					**France**				
Less than secondary	31.9	22.9	13.4	22.8	Less than secondary	57.5	39.6	26.4	41.3
Secondary	51.3	58.6	49.3	53.1	Secondary	29.2	41.5	33.8	34.8
Tertiary	16.8	18.6	37.3	24.1	Tertiary	13.3	18.9	39.8	23.9
Germany					**Denmark**				
Less than secondary	18.3	8.6	8.1	11.7	Less than secondary	19.0	16.3	12.3	15.9
Secondary	58.2	64.7	52.3	58.4	Secondary	53.6	48.0	37.5	46.4
Tertiary	23.5	26.7	39.7	30.0	Tertiary	27.5	35.7	50.2	37.7
Sweden					**Belgium**				
Less than secondary	52.2	43.6	29.7	41.9	Less than secondary	59.0	40.0	30.4	43.2
Secondary	30.6	29.1	34.8	31.5	Secondary	24.5	33.8	28.6	29.0
Tertiary	17.2	27.3	35.5	26.6	Tertiary	16.4	26.2	41.0	27.8
Netherlands					**United States**				
Less than secondary	64.5	47.8	41.2	51.3	Less than secondary	26.1	10.5	3.7	13.5
Secondary	24.3	25.5	22.5	24.1	Secondary	39.0	36.3	22.4	32.6
Spain					Tertiary	34.9	53.2	73.9	54.0
Less than secondary	82.1	83.3	61.1	75.4	**England**				
Secondary	12.0	9.4	18.9	13.5	Less than secondary	61.5	45.4	22.1	43.0
Tertiary	5.9	7.4	20.0	11.1	Secondary	33.0	43.2	49.7	42.0
Italy					Tertiary	5.5	11.4	28.1	15.0
Less than secondary	80.4	69.2	53.6	67.7					
Secondary	16.9	22.5	32.7	24.0					
Tertiary	2.7	8.3	13.7	8.2					

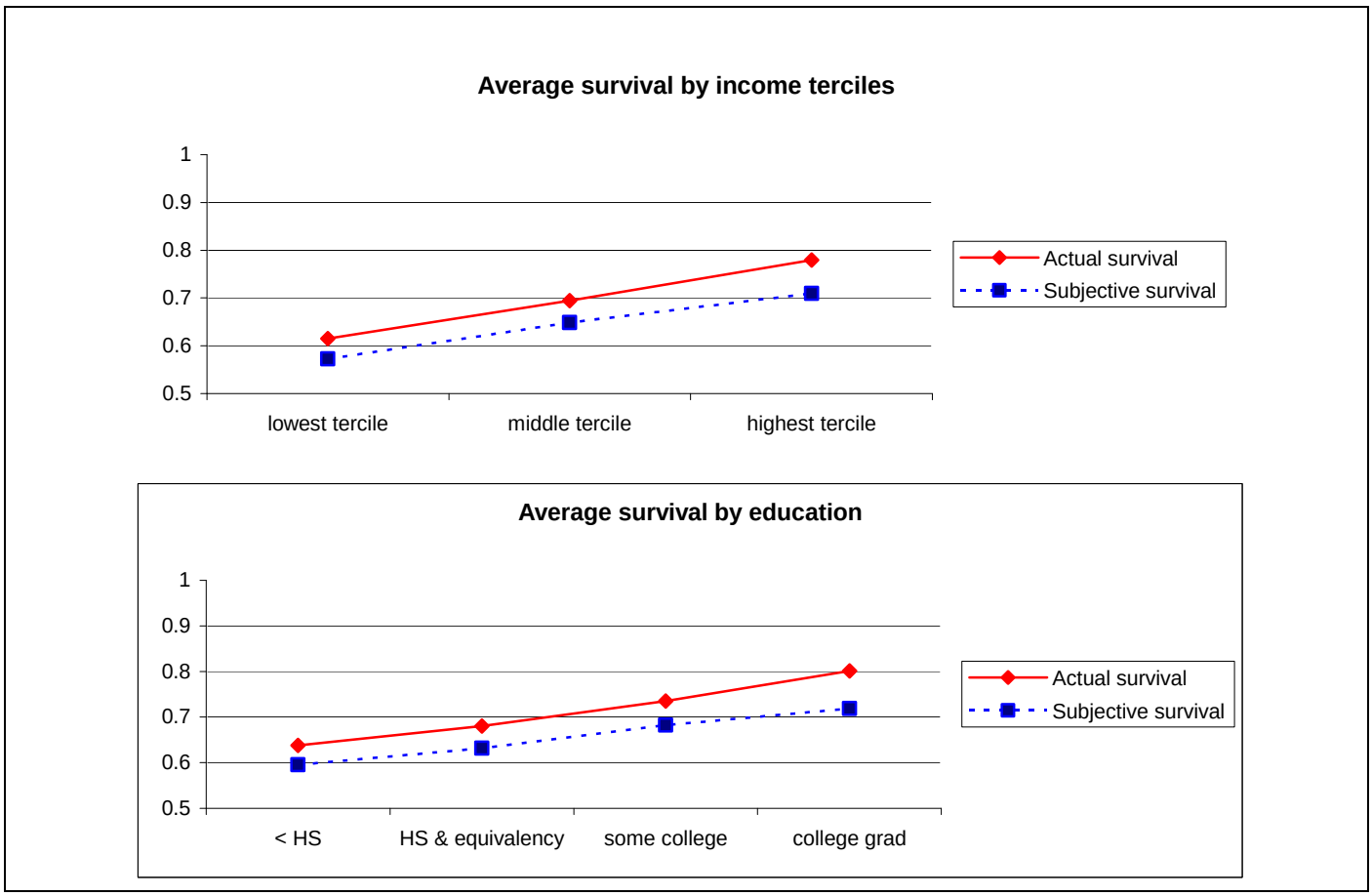
SUPPLEMENT FIGURES

Supplement Figure 1: Kernel regressions for income

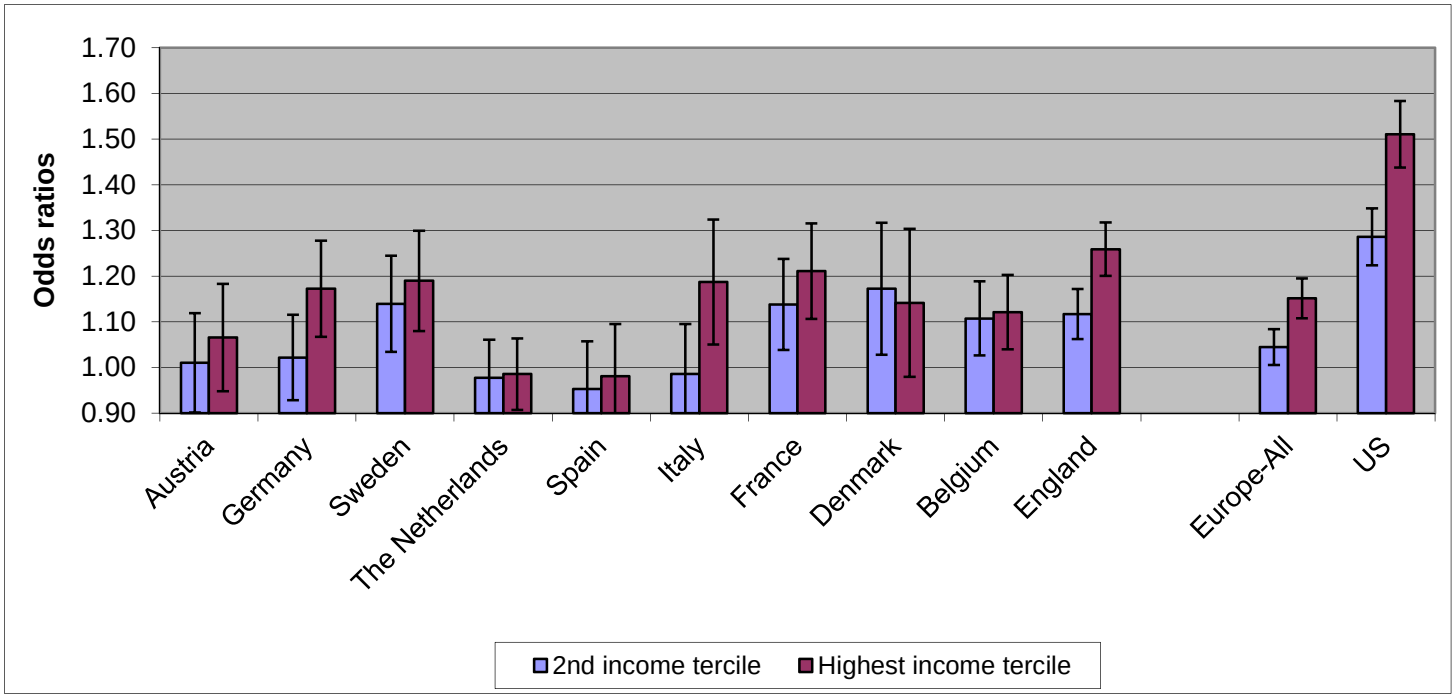


Regressions are based on Gaussian kernel with bandwidth equal to 10. Due to small sample size in the extreme, the kernel estimations are based on a truncated distribution of wealth with truncation at the bottom and top 2%.

Supplement Figure 2: Average survival to age 75 by income and education

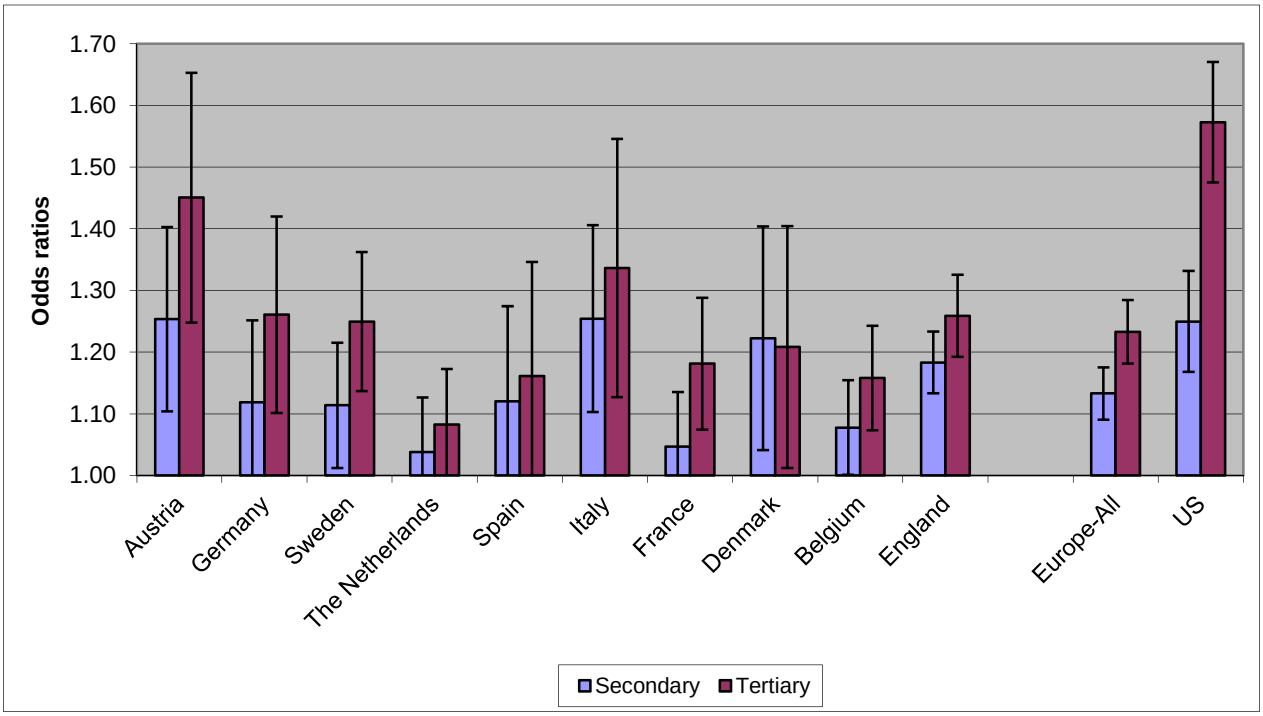


Supplement Figure 3: Differential subjective survival by income across countries – estimates and their 95% confidence intervals



Note: Missing probabilities of survival and 50% answers were replaced with imputed values for this estimation.

Supplement Figure 4: Differential subjective survival by education across countries – estimates and their 95% confidence intervals



Note: Missing probabilities of survival and 50% answers were replaced with imputed values for this estimation.