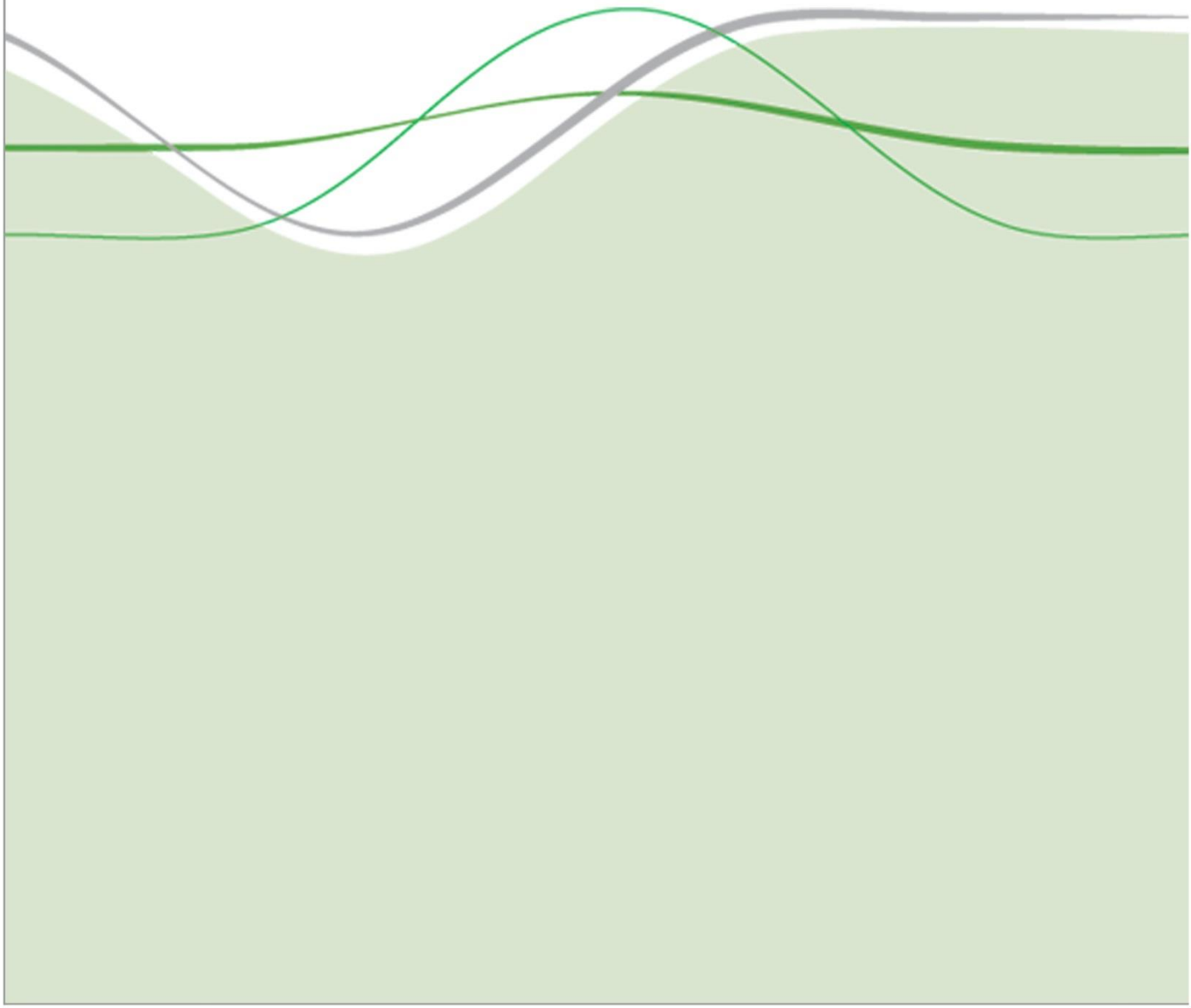
Office of Health Economics Research

Research Paper 17/05

Why do immigrants report lower life satisfaction?

July 2017

Firat Yaman, Patricia Cubi-Molla



Why do immigrants report lower life satisfaction?

Firat Yaman[a], Patricia Cubi-Molla[b]

[a]City, University of London [b]Office of Health Economics

For further information please contact:

Patricia Cubi-Molla

pcubi-molla@ohe.org

The Office of Health Economics (a company limited by guarantee of registered number 09848965)

Southside, 105 Victoria Street

London SW1E 6QT

United Kingdom

Tel: +44 207 747 8858

About OHE

The Office of Health Economics (OHE) has over 50 years' experience of conducting high quality research on the economics of innovation and the life sciences industry, the organisation and financing of health care, and the role for outcomes research and health technology assessment. OHE is a registered charity in England and Wales (registration number 1170829).

About OHE Research Papers

OHE Research Papers are intended to provide information on and encourage discussion about a topic in advance of formal publication.

Any views expressed are those of the authors and do not necessarily reflect the views or approval of OHE, its Editorial Committee or Research and Policy Committee, or its sponsors.

Once a version of the Research Paper's content is published in a peer reviewed journal, that supersedes the Research Paper and readers are invited to cite the published version in preference to the original version.

Acknowledgements

We are grateful to Hugh Gravelle for his comments on an earlier draft.

Table of Contents

ABSTRACT

We investigate the evolution of life satisfaction of immigrants over the course of their stay in the host country. Using data from 1984 to 2010 we find that recently arrived immigrants are more satisfied than comparable natives, but that their life satisfaction relative to comparable natives decreases over time. When analysing the data year by year we find that in the initial waves of our data (the late 1980s) immigrants were becoming relatively more satisfied, but in later waves they were becoming less satisfied the longer they stayed. This effect is not caused by differences in the type of immigrant that arrives at different periods, since it also arises when analysing the evolution of a particular immigrant cohort. Changes in reported self-assessed health can explain some, but not all, of the effect. It is also not explained by selective attrition of immigrants, nor by inclusion of variables proxying the immigrants' integration or the permanence of their stay. We also explore the possibility of explaining this effect by changes in reporting behaviour. For this purpose we propose and estimate an innovative model which allows years since migration to affect reporting behaviour over time. We demonstrate that the estimates of these effects depend on how accurately individuals remember their past life satisfaction.

1. INTRODUCTION

Questions of happiness and well-being[1] have increasingly been drawing the attention of economists, with the implicit or explicit understanding that its measures approximate utility, or at any rate is one of its major components. In this vein, happiness measures have been used to "price" non-market goods (e.g. air traffic noise in Van Praag and Baarsma (2005) and air quality in Levinson (2012)), put forward to improve on quality of life measures used in health-care provision (Brazier and Tsuchiya, 2015), and been advocated (or at least discussed) as a measure or proxy of utility or welfare (Frey and Stutzer, 2002; Di Tella et al., 2003; Layard, 2006; Fleurbaey, 2009; Anand et al., 2011).[2]

Immigration and the assimilation of immigrants has been a long-standing question of interest for economists, dating back at least to Chiswick (1978). Since then immigration and its consequences have become one of the most important phenomena in almost all developed countries. Germany received more than one million refugees over the course of 2015[3] against a background of wide scepticism about the potential of the country to absorb and assimilate such a huge inflow of immigrants. The happiness of immigrants is likely to be an indicator of the success or failure of this effort, either as a cause or consequence.

In this paper we build on a number of findings (discussed below) which suggest that immigrants are in general unhappier than natives in their host country, or that at any rate their happiness tends to decrease over time compared to the natives. Similar to Obucina (2013) and using the same data, the German Socio-Economic Panel (henceforth SOEP), we find that recently arrived immigrants are happier than similar natives, and that their happiness approaches and finally drops below the happiness of natives (we refer to this as years-since-migration, henceforth YSM, effect). Contrary to the existing literature, we explore the robustness and origin of this finding with regard to 1) econometric specifications that allow for correlation between years since migration and unobserved heterogeneity, 2) the stability of the effect over time, 3) the inclusion of different health related variables, and 4) selective attrition (for example only happy immigrants returning to their country). Finally, we introduce an innovative approach to distinguish between threshold effects and genuine satisfaction effects of YSM in an ordered logit model, based on the idea in Easterlin (2001) that individuals apply current thresholds to questions about both their current as well as past satisfaction.

We find that 1) regression models performed over the whole sample period (pooled OLS, fixed effects, ordered logit, fixed effects ordered logit), suggesting a negative YSM effect on the life satisfaction of the immigrant, 2) year-specific models show that the negative YSM effect appears only in the late 1980s when it switches from being positive to

[1] This paper considers life satisfaction. We will however use the words life satisfaction, well-being, and happiness interchangeably. The distinctions are important, but not sharp and a more careful distinction would, we think, be an unnecessary diversion from our main interest: namely the trends in the evolution of life satisfaction of immigrants.

[2] The advocates of using-happiness-as-utility also have their detractors and sceptics. In a recent paper Benjamin et al. (2014) find important deviations between anticipated well-being of available choices and the actual choices that were made, Loewenstein and Ubel (2008) draw attention to the adaptation of individuals with serious chronic health conditions, and Bertrand and Mullainathan (2001) point to the biases induced by random and non-random noise in self-reported happiness.

[3] See http://www.welt.de/politik/deutschland/article150678614/1-1-Millionen-Fluechtlinge-kamen-2015-nach-Deutschland.html.

significantly negative, 3) changes in objective health-related measures do not fully explain the effect, whereas subjective self-assessed health measures seem to be to a large extent driven by the same underlying variables as changes in life satisfaction, 4) the effect is not driven by selective attrition and 5) the effect may be explained by changes in the reporting behaviour. For the latter part, we propose a model that has the potential to decompose the YSM effect into an effect on the latent variable (true change in life satisfaction) and into an effect on the threshold (change in reporting behaviour).

There is a nascent literature that links the issues of happiness and immigration. Many of these recent studies have sought to establish a link between the migration experience and the happiness of immigrants, and most of them have found a negative association between being an immigrant and being happy. Bartram (2011) finds a negative coefficient on being an immigrant in the United States in a cross-section of the World Values Survey. Safi (2010) provides the most comprehensive international evidence on this negative association: using the European Social Survey immigrants report lower levels of life satisfaction in all of the 13 European countries in the study, with life satisfaction being lowest among immigrants who arrived 5 to 10 years ago in the host country. The paper does not discuss whether the latter finding is a cohort effect, or a genuine change in life satisfaction of immigrants over time. This finding is replicated by Olgiati et al. (2013) using the Gallup World Poll in a pooled sample of European and Anglo-Saxon countries. Whether these findings relate to the actual immigration experience rather than, for example, immigrants being selected from particularly unhappy people or coming from "unhappy" countries, could not be established with the used data, since the immigrants' (or an appropriately chosen control group's) happiness before migration was not available. Stillman et al. (2015) exploit the random assignment of immigration status to immigration applicants from Tonga through a lottery. Tongans who won the lottery and therefore gained the right to immigrate and subsequently immigrated report the same level of happiness as Tongans who did not win the lottery and stayed in Tonga, despite having improved economically and on other measures of well-being.

In contrast to the above papers, Obúcina (2013) uses panel data from Germany from 1994 to 2009 to analyse the dynamic aspects of immigration and happiness. In a Mundlak-type random effects model (that is including person-specific mean values of time-variant variables as independent variables) he finds that recently arrived immigrants report higher levels of happiness than German natives, but their level of happiness drops linearly with time and ends up well below the level of their German peers.[4] He offers a plausible (but untested) explanation that an immigrant who arrived in Germany many years ago will feel unsatisfied if his income is the same as an otherwise equal but recently arrived immigrant. While the data and methodology is more suited to answer dynamic aspects of happiness, many possible explanations of the years-since-migration effect have remained unexplored. This paper fills this gap.

2. DATA AND DESCRIPTIVE EVIDENCE

We use data from the German Socio-Economic Panel (SOEP) from the initial wave 1984 to 2010. The SOEP is a well-known dataset among social scientists. We refer the reader to Wagner et al. (2007) and Wagner et al. (2008) for a comprehensive overview. The SOEP asks the following question in every wave: "How satisfied are you at present with

[4] A negative years-since-migration effect is also found by Taengnoi (2014) for the US.

your life as a whole?" The respondents can then answer this question with an integer number between 0 and 10, with 0 being the lowest and 10 the highest level of life satisfaction. We include in our analysis all observations which have non-missing values in the life satisfaction question and all independent variables that we use in our model. The independent variables we consider are immigration status (equal to 1 if the person was born outside of Germany), years since migration (defined as survey year minus year of immigration for immigrants, and set to 0 for natives[5]), age in years, age squared divided by 100, sex, living with a partner, years of education, the log of equivalised household income, having children in the household, being employed, being unemployed, and being retired (not being retired and not being in the labour force being the omitted category). The main health variable that we include is the number of nights that the person spent in hospital in the previous year, for the simple reason that it is available in almost all survey years, but we also repeat the main regressions with a number of alternative health variables. We include only observations from respondents aged 18 to 65 and exclude respondents resident in East Germany. Table 1 gives an overview of the descriptive statistics of the pooled sample separately for natives and immigrants.

Table 1. Descriptive statistics - averages for natives and immigrants

	Immigrant	**Native**	**Difference***
Life satisfaction (0-10)	6.99	7.16	-0.17
Years since migration	19.09		
Age	41.37	41.22	0.15
Male (%)	0.51	0.49	0.02
Has partner (%)	0.81	0.72	0.10
Education (years)	10.13	11.98	-1.85
Ln(income)	6.70	7.02	-0.31
Nights in hospital	1.76	1.43	0.33
Has children in household (%)	0.56	0.39	0.17
Employed (%)	0.67	0.73	-0.06
Unemployed (%)	0.09	0.04	0.05
Retired (%)	0.00	0.01	-0.002

* All differences in the third column are significant at the 5% level.

We see that important differences between natives and immigrants exist in all dimensions: Immigrants are significantly less satisfied, older (recall that the sample is restricted to 18 to 65 year old individuals), more likely to be male, more likely to be living with a partner (spouse), less educated, have less income, less healthy, more likely to have children in the household, less likely to be employed, and more likely to be unemployed.

We begin the analysis by presenting basic findings on the evolution of life satisfaction in the SOEP covering the survey years 1984 to 2010. Figure 1 depicts the age – life satisfaction profile of natives and immigrants in the pooled data. We see that young immigrants are more or less as satisfied as natives. However life satisfaction of immigrants older than 25 is consistently below that of natives, and this distance suddenly increases for observations in their 50s. We also see the U - pattern in age that has been documented in several studies (e.g. Blanchflower and Oswald (2008) or Frijters et al. (2004)).

[5] Note that this variable is also 0 for recently arrived immigrants –less than a year before the survey date.

Figure 1. Age - life satisfaction profile for natives and immigrants

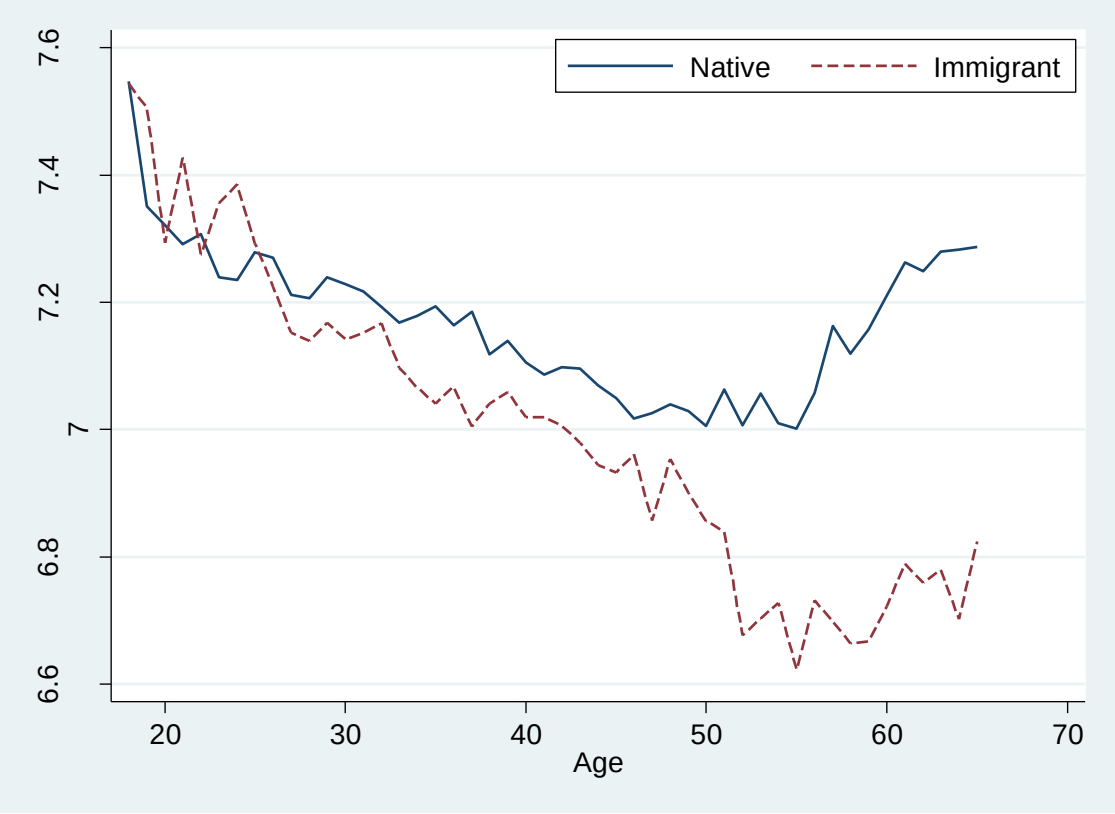


Figure 1 does not necessarily imply that immigrants who have lived longer in Germany are less satisfied. Immigrants who arrived late in their life in Germany might have been less satisfied already at arrival. In the extreme we could imagine all immigrants arriving in the same year, but the older ones being already much less satisfied. Figure 2 reveals that this pattern is indeed related to years since migration. The graph depicts the life satisfaction of immigrants compared to a same-aged native in the same year. In other words, it depicts the YSM coefficients (plus the immigrant coefficient) in a regression of life satisfaction on an immigrant dummy, a full set of YSM dummies, a full set of age dummies, and a full set of year dummies. The graph mimics the pattern of figure 1. We conclude that immigrants who have lived longer in Germany are genuinely less satisfied than their native peers, and this gap is increasing with YSM. Figure 3 shows that this pattern is not uniform across all immigrant groups. The YSM effect is most pronounced among Turkish and Yugoslavian immigrants, and weakly present among immigrants from the main Mediterranean countries (Spain, Greece, and Italy).[6]

The graphical analysis still cannot rule out the presence of confounding factors or that the YSM effect is just a cohort effect, that is immigrants who arrived a long time ago in Germany were already less satisfied than their native peers at their arrival, whereas those who arrived recently have satisfaction levels comparable to natives. The ensuing regression analyses demonstrate that the YSM effect truly reflects a within-individual rather than between-individual effect and that it is robust to inclusion of control variables.

[6] For the remaining immigrants there appears to be no YSM effect.

Figure 2. Years since migration - life satisfaction profile for immigrants (natives' life satisfaction as baseline)

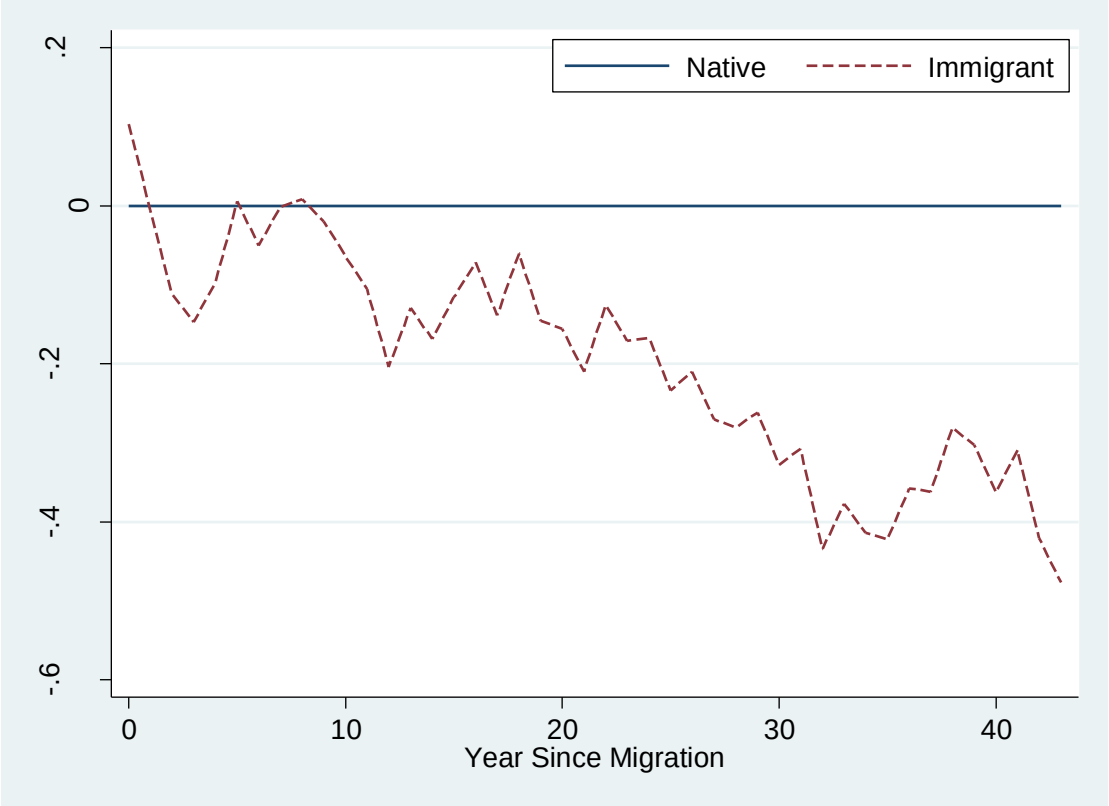


Figure 3. Years since migration - life satisfaction profile, by origin

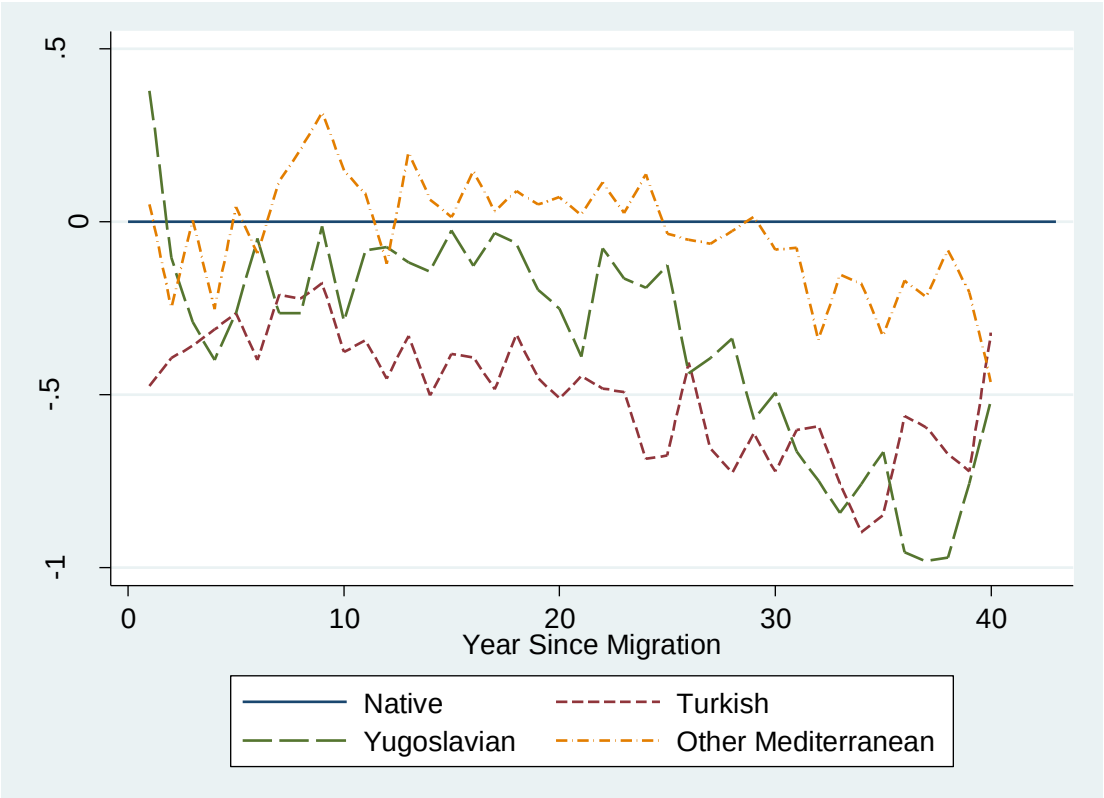


Table 2. Determinants of life satisfaction: 1984 to 2010

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Pooled OLS		Fixed Effects (FE)		Ordered Logit		FE Ordered Logit	
Immigrant	0.024	0.197***			0.023	0.212***		
	(0.018)	(0.018)			(0.019)	(0.019)		
Years since migration	-0.010***	-0.008***	-0.014***	-0.010***	-0.011***	-0.009***	-0.015***	-0.011***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.004)	(0.004)
Age	-0.040***	-0.115***	0.115***	0.073***	-0.044***	-0.122***	-0.055***	-0.110***
	(0.002)	(0.002)	(0.019)	(0.019)	(0.002)	(0.002)	(0.006)	(0.007)
$Age^2/100$	0.042***	0.125***	0.002	0.047***	0.047***	0.133***	0.004	0.063***
	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.007)	(0.008)
Male		-0.072***				-0.084***		
		(0.007)				(0.007)		
Partner		0.451***		0.346***		0.448***		0.460***
		(0.009)		(0.013)		(0.010)		(0.028)
Ln(income)		0.625***		0.380***		0.663***		0.526***
		(0.009)		(0.011)		(0.009)		(0.021)
Nights in hospital		-0.020***		-0.010***		-0.019***		-0.012***
		(0.000)		(0.000)		(0.000)		(0.001)
Children		0.202***		0.113***		0.214***		0.145***
		(0.008)		(0.010)		(0.009)		(0.021)
Employed		0.111***		0.104***		0.076***		0.137***
		(0.009)		(0.011)		(0.010)		(0.020)
Unemployed		-0.885***		-0.564***		-0.855***		-0.667***
		(0.017)		(0.017)		(0.018)		(0.031)
Retired		-0.071		-0.028		-0.087		-0.030
		(0.049)		(0.045)		(0.053)		(0.069)
Education		0.021***				0.024***		
		(0.001)				(0.002)		
Observations	250,363	250,363	250,363	250,363	250,363	250,363	250,363	250,363
R-squared	0.009	0.084	0.023	0.046				
Persons			33,240	33,240			33,240	33,240

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The R-squared for the fixed effects models are squared correlations between the de-meaned life satisfaction and predicted de-meaned life satisfaction from the fixed effects estimation. Stata reports this measure as R-squared within.

3. ESTIMATION AND RESULTS

Most estimations are straightforward. We estimate models of life satisfaction applying pooled OLS, fixed effects, and ordered logit estimators. The first two assume cardinality of life satisfaction. The fixed effects estimator allows for time-invariant unobserved heterogeneity which might correlate with the independent variables. We also estimate a fixed effects ordered logit model, which combines both the non-cardinality and the unobserved heterogeneity of the ordered logit and the fixed effects model. The method has been labelled "blow-up and cluster" by Baetschmann et al. (2015) and is based on a dichotomization of life satisfaction at every possible cut-off-point and applying the conditional logit of Chamberlain (1980) to the dichotomized dependent variable. We briefly describe the method in Appendix A and refer the reader to Mukherjee et al. (2008) and Baetschmann et al. (2015) for a detailed discussion of this estimator. All regressions include a full set of year-fixed effects.

The results from these regressions are reported in Table 2. We see that being an immigrant with 0 YSM carries a positive coefficient (columns 2 and 6), implying that recently arrived immigrants report higher levels of life satisfaction than similar natives. However, every additional year in Germany reduces life satisfaction. The results from the fixed effects regressions (columns 3, 4, and 7, 8) demonstrate that this is a genuine YSM rather than a composition effect. As has been found in other studies, for example Ferrer-i-Carbonell and Frijters (2004), the choice of cardinal or ordinal methods is of no relevance. The OLS / fixed effects results are close to the ordered logit results. The remaining variables also carry the expected results. Being female, having a partner, more income, better health, and (not) being (un)employed are all associated with higher life satisfaction. Little more than 8% of the variation in life satisfaction is explained by the model, a typical number for life satisfaction models (Diener et al., 1999). The effect of YSM on life satisfaction is significant, but modest in absolute value. This is apparent when we standardize life satisfaction: including all controls, an additional 10 years in Germany implies a decrease in life satisfaction of only 6% of a standard deviation in life satisfaction (fixed effects), and a recently arrived immigrant is 11% of a standard deviation more satisfied than an otherwise comparable native (pooled OLS). We have also estimated a model including YSM squared. This did not change the monotonic negative relationship between YSM and life satisfaction and did not significantly improve the model fit, so we proceed with models linear in YSM.

3.1. Health

We next take a closer look at the health of immigrants. Since life satisfaction is higher among recently arrived immigrants but decreases faster than for natives, a natural explanation is that life satisfaction follows the same path as the health of immigrants. Health has also been reported to be higher among recent immigrants (the "healthy immigrant" effect, see for example Farré (2016) and Giuntella and Mazzonna (2015)). Since immigrants in Germany have been working in physically more demanding jobs, this health advantage might have shrunk and disappeared over time. We have used nights in hospital in our benchmark model because of the availability of this variable in most years. The use of alternative health measures comes with the disadvantage of treating different subsamples for each variable included in the analysis.

Table 3 reports results from fixed effects regressions which use alternative measures. Here, "physical health" and "mental health" are composed of answers to a battery of questions and range between 0 and 100. Their construction is discussed in Andersen et al. (2007). We briefly describe the construction of these variables in Appendix B. First we observe that the inclusion of alternative health-related variables does not affect the significance of any coefficient but the one for YSM. The significance of the YSM coefficient is only preserved when we use a dummy indicating whether the respondent is hindered in his daily activities by his health (column 4), and when we use a dummy indicating whether the respondent has not been to a doctor in the last year (column 6). In column 5 the YSM effect does not exist. A reason for this could be that the "chronic illness" variable is only available in the 1980s. We will see that the YSM effect appears only beginning in the late 1980s/early 1990s.

Table 3. Different health determinants of life satisfaction

	Fixed effects					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Years since migration	-0.001	-0.008	-0.005	-0.009***	0.003	-0.009***
	(0.002)	(0.006)	(0.006)	(0.002)	(0.007)	(0.001)
Age	0.092***	-0.034***	-0.004	-0.035**	0.047***	0.070***
	(0.018)	(0.013)	(0.012)	(0.018)	(0.016)	(0.019)
Age2/100	0.042***	0.052***	0.024**	0.057***	0.061***	0.049***
	(0.004)	(0.010)	(0.010)	(0.005)	(0.014)	(0.003)
Partner	0.339***	0.389***	0.303***	0.387***	0.483***	0.348***
	(0.014)	(0.030)	(0.028)	(0.021)	(0.039)	(0.013)
Ln(income)	0.331***	0.288***	0.250***	0.374***	0.334***	0.384***
	(0.012)	(0.024)	(0.023)	(0.018)	(0.028)	(0.011)
Children	0.103***	0.096***	0.089***	0.087***	0.006	0.112***
	(0.011)	(0.025)	(0.023)	(0.016)	(0.027)	(0.010)
Employed	0.103***	0.100***	0.092***	0.096***	0.091***	0.111***
	(0.012)	(0.025)	(0.024)	(0.017)	(0.028)	(0.011)
Unemployed	-0.466***	-0.505***	-0.470***	-0.597***	-0.633***	-0.563***
	(0.018)	(0.037)	(0.036)	(0.027)	(0.044)	(0.017)
Retired	-0.086	0.034	0.043	0.027	0.052	-0.028
	(0.052)	(0.125)	(0.119)	(0.073)	(0.095)	(0.046)
Self-Assessed Health	0.489***					
	(0.005)					
Physical health		0.019***				
		(0.001)				
Mental health			0.052***			
			(0.001)			
Hindered by health				-0.520***		
				(0.011)		
Chronic illness					-0.326***	
					(0.022)	
No doctor visits						0.121***
						(0.007)
Observations	190,019	57,313	57,312	112,182	58,234	249,502
R-squared	0.098	0.035	0.124	0.063	0.031	0.044
Number of persons	27,838	19,098	19,098	25,062	13,053	33,220

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.The R-squared for the fixed effects models are squared correlations between the de-meaned life satisfaction and predicted de-meaned life satisfaction from the fixed effects estimation. Stata reports this measure as R-squared within.

Inclusion of the health variables in the survey: Self-Assessed Health (1992-2010), Physical health and Mental health (2002-2010 every other year), Hindered by health (1984-1987; 1992; 1995-2001); Chronic illness (1984-1991), No doctor visits (1984-2010).

We also observe that health variables more closely related to self-assessed perceptions of health (e.g. self-reported health, SF-12-based measures of physical and mental health) explain more of the variation of life satisfaction, and reduce the YSM coefficient as well as remove its significance. Self-assessed health and life satisfaction are psychological constructs.[7] How they are related to each other has generated an intense debate in recent years with a focus on specific patient cohorts (see Brazier and Tsuchiya, 2015). Less is known on the potential relation of health and life satisfaction for other socio-demographic groups such as immigrants. Our observation is that controlling for constructed health measures reduces the YSM effect to insignificance. One explanation of this is causal: immigrants' health declines, and therefore so does their life satisfaction. Alternatively, immigrants' perception of their health changes (without an underlying change in their health status), and this correlates with a change in the immigrants' perception of their life satisfaction. Thus constructed health measures are more likely to be endogenous. We therefore prefer the models that use objective health measures. We will return to the analysis of changes in the perception of life satisfaction in section 3.6.

3.2. Attrition

Another candidate explanation for the YSM effect is non-random attrition, especially in light of the return-migration option for immigrants. For example, if immigrants arrive in Germany with certain objectives and intend to return once these objectives are met, we might see successful immigrants leaving earlier and more satisfied than immigrants who fall short of their targets, become less satisfied, and opt to stay in Germany. The immigrants who become less-satisfied become over-represented in an unbalanced panel and pull down the YSM coefficient.

In a first step we analyse the factors predicting the attrition hazard. To this end we estimate a Weibull attrition hazard model. The results are presented in Table 4. The first column reveals that immigrants are more likely to drop out of the survey, undoubtedly also due to non-negligible return migration. Immigrants who have been in Germany longer are less likely to attrite. Adding life satisfaction as a variable also shows that the less satisfied are more likely to drop out. However, including an interaction term (column three) shows that there is no systematic difference between satisfied immigrants and satisfied natives to drop out.

Next we test for non-random attrition in a fixed effects model of life satisfaction by including a dummy in our benchmark regression which is equal to one if the respondent is not in the sample in the following year (as suggested in Wooldridge (2002), pp. 586–587). The result in Table 5, column 1 shows that respondents which are going to attrite in the next survey wave are less satisfied, echoing the result from the hazard model. Column 2 also shows that this effect is not significantly different between immigrants and natives: an immigrant who will attrite is only marginally more satisfied than a native who is going to attrite. To correct for potential attrition bias we again follow the steps in Wooldridge (2002): we run separately for each year a probit regression of whether the respondent will attrite in the next survey wave on current values of all independent variables. We then construct predicted inverse Mills ratios λ_{it} (based on variable values in $t-1$) which we include along with interactions with all year dummies in a fixed effects model. A loose but intuitive interpretation is that the λ proxy the probability that the

[7] I.e. not directly observable, non-tangible, and subjective.

respondent will attrite in the next period. Column 3 of Table 5 shows results from this attrition-corrected model. The YSM effect is slightly above the benchmark result. The YSM effect is thus not driven by any systematic attrition difference between natives and immigrants or within the immigrant sample.

Table 4. Weibull attrition hazard models

	Weibull Duration Model		
	(1)	**(2)**	**(3)**
Immigrant	0.370***	0.375***	0.402***
	(0.033)	(0.033)	(0.068)
Years since migration	-0.011***	-0.011***	-0.011***
	(0.001)	(0.001)	(0.001)
Age	-0.162***	-0.165***	-0.165***
	(0.004)	(0.004)	(0.004)
$Age^2/100$	0.200***	0.204***	0.204***
	(0.005)	(0.005)	(0.005)
Male	0.080***	0.078***	0.078***
	(0.013)	(0.013)	(0.013)
Partner	-0.056***	-0.043**	-0.043**
	(0.017)	(0.017)	(0.017)
Ln(income)	-0.235***	-0.217***	-0.217***
	(0.016)	(0.016)	(0.016)
Nights in hospital	0.003***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)
Children	-0.067***	-0.062***	-0.062***
	(0.017)	(0.017)	(0.017)
Employed	-0.083***	-0.079***	-0.079***
	(0.017)	(0.017)	(0.017)
Unemployed	-0.083***	-0.107***	-0.107***
	(0.032)	(0.032)	(0.032)
Retired	1.530***	1.527***	1.527***
	(0.033)	(0.033)	(0.033)
Education	0.004	0.004	0.004
	(0.003)	(0.003)	(0.003)
Life satisfaction		-0.028***	-0.027***
		(0.004)	(0.004)
Life satisfaction x Immigrant			-0.004
			(0.008)
Observations	250,363	250,363	250,363

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 5. Non-random attrition models

	(1) Fixed effects	(2) Fixed effects	(3) Fixed effects with inverse Mills-ratios
Years since migration	-0.010***	-0.010***	-0.012***
	(0.001)	(0.001)	(0.001)
Age	0.061***	0.061***	-0.065***
	(0.019)	(0.019)	(0.003)
Age2/100	0.048***	0.048***	0.040***
	(0.003)	(0.003)	(0.003)
Partner	0.346***	0.346***	0.322***
	(0.013)	(0.013)	(0.013)
Ln(income)	0.378***	0.378***	0.373***
	(0.011)	(0.011)	(0.011)
Nights in hospital	-0.010***	-0.010***	-0.010***
	(0.000)	(0.000)	(0.000)
Children	0.112***	0.112***	0.117***
	(0.010)	(0.010)	(0.011)
Employed	0.103***	0.103***	0.109***
	(0.011)	(0.011)	(0.011)
Unemployed	-0.566***	-0.566***	-0.556***
	(0.017)	(0.017)	(0.017)
Retired	0.064	0.067	-0.068
	(0.047)	(0.047)	(0.047)
Attrition	-0.106***	-0.113***	
	(0.012)	(0.013)	
Attrition x Immigrant		0.029	
		(0.027)	
Observations	250,363	250,363	197,278
R-squared	0.046	0.046	0.039
Number of persons	33,240	33,240	27,941

Dependent variable: life satisfaction. Standard errors in parentheses; standard errors in column three are bootstrapped from 500 replications. *** p<0.01, ** p<0.05, * p<0.1

3.3. Time Period

Next we investigate whether the YSM - life satisfaction association is stable over time. In a year-by-year OLS estimation we find that the negative effect is only stable after the year 1992. Indeed, the explanatory power of the model seems to be much stronger in the latter third of the panel, and weaker in the first few waves. Running fixed effects regressions on different time periods reveals that there is no negative relationship in the first four survey years (see Table 6). Indeed, the YSM coefficient's p-value is just above 0.10, suggesting a positive relationship of years since migration and life satisfaction in the early years of the panel. The negative effect is most pronounced in the transformative years of 1988 to 1992 which saw the fall of the Berlin wall and German re-unification but also a rise in racist violence and rhetoric against immigrants and their descendants. We emphasize that the reversal of the YSM effect is not a result of new immigrants arriving in Germany, as the sample of immigrants over those years in the SOEP is fairly stable.

The variation of the YSM effect over time also explains why the inclusion of the physical health variable reduces the years since migration effect: the variable is only available from 2002 onwards and thus selects a time-period in which the YSM effect is smaller.

Table 6. Life satisfaction and immigration by year

Year	(1) Immigrant	(2) Years since migration	(3) R2
1984	-0.123	0.012*	0.069
1985	-0.126	0.012*	0.069
1986	0.118	-0.007	0.058
1987	0.417***	-0.016**	0.061
1988	0.379***	-0.009	0.049
1989	-0.053	0.007	0.052
1991	-0.186	0.005	0.062
1992	-0.434***	0.015***	0.071
1994	0.318***	-0.014***	0.065
1995	0.245***	-0.006	0.075
1996	0.230***	-0.013***	0.081
1997	0.247***	-0.010**	0.075
1998	0.264***	-0.010***	0.091
1999	0.178**	-0.007*	0.083
2000	0.253***	-0.015***	0.079
2001	0.254***	-0.011***	0.075
2002	0.205***	-0.009***	0.098
2003	0.310***	-0.014***	0.109
2004	0.231***	-0.009***	0.102
2005	0.059	-0.005	0.122
2006	0.181*	-0.008**	0.116
2007	0.285***	-0.011***	0.119
2008	0.256**	-0.010**	0.101
2009	0.306***	-0.011***	0.090
2010	0.196*	-0.007*	0.105
1984-1987 (FE)		0.028	0.031
1988-1992 (FE)		-0.083***	0.024
1994-2000 (FE)		-0.015**	0.026
2001-2010 (FE)		-0.008	0.033

FE = Fixed Effects; Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The R-squared for the fixed effects models are squared correlations between the de-meaned life satisfaction and predicted de-meaned life satisfaction from the fixed effects estimation. Stata reports this measure as R-squared within.

3.4. Ethnic Background

We next check whether there are important differences across ethnicities. To that end we run fixed effects regressions on separate immigrant samples and excluding the survey years 1984 to 1987. To be able to identify separate age and YSM effects we include the natives in each regression, too. Table 7 presents results from these sub-samples. The negative YSM effect is most pronounced for immigrants from the former Yugoslavia. This could be the effect of war as an external shock that occurred in the early 1990s and affected immigrants from the former Yugoslavia: the violence in their home countries might have reduced the life satisfaction of Yugoslavians over and above any age and proper YSM effect. Turks and immigrants from Spain, Greece, and Italy have the same YSM coefficient, but there is no clear effect for immigrants from other countries.

Table 7. Fixed effects regressions, by immigrant group

	(1) Turkish	(2) Yugoslavian	(3) Other Medi-terranean	(4) Other Immi-grant
Years since migration	-0.006**	-0.017***	-0.007***	-0.000
	(0.003)	(0.002)	(0.001)	(0.001)
Age	0.080***	-0.074***	-0.073***	-0.076***
	(0.019)	(0.003)	(0.003)	(0.003)
Age2/100	0.048***	0.076***	0.074***	0.077***
	(0.004)	(0.003)	(0.003)	(0.003)
Partner	0.324***	0.373***	0.368***	0.373***
	(0.014)	(0.012)	(0.012)	(0.012)
Ln(income)	0.363***	0.475***	0.468***	0.472***
	(0.012)	(0.010)	(0.010)	(0.010)
Nights in hospital	-0.011***	-0.012***	-0.013***	-0.013***
	(0.000)	(0.000)	(0.000)	(0.000)
Children	0.114***	0.140***	0.144***	0.138***
	(0.011)	(0.010)	(0.010)	(0.010)
Employed	0.081***	0.068***	0.071***	0.069***
	(0.012)	(0.011)	(0.011)	(0.011)
Unemployed	-0.537***	-0.620***	-0.621***	-0.612***
	(0.019)	(0.019)	(0.019)	(0.018)
Retired	-0.073	-0.049	-0.049	-0.054
	(0.054)	(0.049)	(0.048)	(0.048)
Observations	184,627	179,645	182,467	189,408
R-squared	0.041	0.039	0.039	0.039
Number of persons	25,790	25,269	25,670	26,826

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The R-squared for the fixed effects models are squared correlations between the de-meaned life satisfaction and predicted de-meaned life satisfaction from the fixed effects estimation. Stata reports this measure as R-squared within

3.5. Integration

The last set of variables we consider are related to how well integrated the immigrants are and/or where they see their future. Koczan (2013) finds that immigrants' life satisfaction is no different from natives' once a variable indicating whether the respondent works in a job he was trained in is controlled for. The first column of Table 8 shows that the YSM coefficient is not affected by inclusion of this variable. The second column shows that inclusion of self-assessed speaking fluency of German and one's native language also do not change the YSM coefficient.[8] Including a dummy for the intention to stay in Germany forever does reduce the YSM coefficient by 25% compared to the 0.008 benchmark value. Still, YSM remains significant.

3.6. Changes in the Perception of Life Satisfaction

Life satisfaction in the survey is reported as a rating scale. The common approach to manage rating scales is based on the assumption of a latent, unobserved true variable, which is converted into a particular category by the respondent, following a threshold partition. For instance, if LS is between two particular values (thresholds), the individual would choose one of the categories, and so on. The literature has been aware of a potential response heterogeneity (i.e. different individuals using different thresholds). Recent contributions have developed consistent and/or efficient estimators for the

[8] The language variables are only available for the immigrant sample.

Table 8. Integration variables on life satisfaction

	(1) Fixed effects (full sample)	(2) Pooled OLS (immigrant sample)	(3) Pooled OLS (immigrant sample)
Years since migration	-0.009***	-0.011***	-0.006***
	(0.002)	(0.002)	(0.001)
Age	0.090***	-0.068***	-0.095***
	(0.021)	(0.008)	(0.006)
Age2/100	0.032***	0.074***	0.099***
	(0.004)	(0.010)	(0.007)
Male		-0.052*	-0.040**
		(0.028)	(0.019)
Partner	0.290***	0.428***	0.468***
	(0.014)	(0.039)	(0.026)
Ln(income)	0.352***	0.586***	0.662***
	(0.014)	(0.036)	(0.024)
Nights in hospital	-0.010***	-0.018***	-0.020***
	(0.001)	(0.001)	(0.001)
Children	0.093***	0.207***	0.212***
	(0.012)	(0.033)	(0.022)
Employed	0.018	0.214***	0.282***
	(0.057)	(0.036)	(0.024)
Unemployed		-0.821***	-0.669***
		(0.053)	(0.035)
Retired		0.049	-0.254
		(0.280)	(0.166)
Education		-0.012*	0.012***
		(0.007)	(0.004)
German skill		0.197***	
		(0.015)	
Native language skill		0.089***	
		(0.019)	
Works in occupation	-0.018***		
	(0.006)		
Wants to stay in Germany			0.314***
			(0.019)
Observations	170,396	21,914	42,481
R-squared	0.033	0.086	0.096
Number of persons	26,274		

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The R-squared for the fixed effects models are squared correlations between the de-meaned life satisfaction and predicted de-meaned life satisfaction from the fixed effects estimation. Stata reports this measure as R-squared within.

ordered logit, which are all based on a dichotomization of the dependent variable and the application of Chamberlain's (1980) conditional logit model (as in Ferrer-i-Carbonell and Frijters (2004), Mukherjee et al. (2008) or Baetschmann et al. (2015)).

It is essential to note that all of these applications have assumed that the thresholds that divide one category from the other are fixed over time (but not necessarily across individuals). That is, for a particular latent value of life satisfaction, the individual will report the same category (e.g. "8") regardless the year of the survey. This assumption is necessary in ordered choice models, since the effect of an independent variable on the level of the latent variable cannot be separately identified from the independent variable's effect on the level of the thresholds. If thresholds do change systematically with observed

variables, then the coefficient estimates in the aforementioned papers are not effects on the level of the latent variable, but rather on the level change of the latent variable relative to threshold locations (which might have changed themselves), and thus contain very little information, since even the sign cannot be interpreted in its effect on the latent variable.

Nonetheless, the existence of a response-shift has been frequently found in the reporting behaviour linked to subjective psychological constructs (see for instance Schwartz et al. (2013)). In the context of life satisfaction for immigrants, we can find several papers pointing at that direction. For instance, Gelatt (2013) finds that immigrants change their reference group after some years living in the host country: they move from comparing themselves to those staying in their country, to other immigrants and/or natives. Also, Angelini et al. (2015) explore how cultural assimilation can have an impact on the immigrant's life satisfaction. In both scenarios, it is plausible that the immigrant may not be changing her level of life satisfaction but changing the thresholds which determine her response.

We take a closer look at the potential response shift and introduce an innovative approach to decompose the YSM coefficient of a fixed effects ordered logit regression into a part which shifts the latent variable (the "true" life satisfaction) and another part which shifts the thresholds to qualify a given level of life satisfaction into one of the answer categories (an integer between 0 and 10), as in Cubi-Molla and Yaman (2015). To do this, we first estimate the life satisfaction model with the blow-up and cluster estimator (as in Table 2, columns 7 and 8), but restrict the sample years to 1984 to 1987. This narrower time bracket is necessary, since we want to exploit the question about life satisfaction in the previous year, which was asked uniquely in those four years. Note that this is not the lagged life satisfaction variable, but rather how the respondent would today rate his or her life satisfaction in the previous year. Again, respondents have to rate their past life satisfaction on a scale from 0 to 10. We thus apply a latent variable formulation to this variable. Finally, we follow the model by Easterlin (2001) in assuming that respondents apply their current thresholds to assess both their current and their past life satisfaction: a choice that is supported by the fact that the question about past life satisfaction is asked right after the question about current life satisfaction. Table 9 demonstrates that the data on recalled and lagged life satisfaction is consistent with such a modelling choice (it might be consistent with other models, too). The table contrasts the recalled life satisfaction (rows) with the life satisfaction that was reported in the previous year (columns). Under perfect recall and stable reporting behaviour one would observe all responses only in the diagonal cell entries (100 percent). However, we see that most cases do not fall into those cells. Only 29% of the observations (not reported in table) report a recalled level of life satisfaction that coincides with what they had reported in the previous year, while 30% report recalled levels above, and 41% report recalled levels below the life satisfaction level they had reported in the previous year.

Table 9. Reported and recalled life satisfaction

		Life satisfaction for year t-1 recalled in year t											
		0	1	2	3	4	5	6	7	8	9	10	Total
Life satisfaction as reported in year t-1	0	25	10	14	28	27	41	15	17	16	4	3	200
	1	9	8	12	12	12	20	8	5	15	5	1	107
	2	7	15	15	32	32	75	26	24	21	6	7	260
	3	18	11	24	59	69	115	77	60	53	19	9	514
	4	13	13	33	71	72	165	108	98	60	24	11	668
	5	42	17	82	149	234	752	437	496	421	121	104	2,855
	6	14	6	26	73	134	372	361	504	396	119	79	2,084
	7	13	13	30	61	140	522	619	1,166	1,099	302	172	4,137
	8	18	8	34	68	110	533	566	1,436	2,357	829	428	6,387
	9	4	6	7	29	37	167	191	490	1,068	837	348	3,184
	10	12	6	13	23	47	184	142	335	872	672	1,206	3,512
	Total	175	113	290	605	914	2,946	2,550	4,631	6,378	2,938	2,368	23,908

We have thus the following building blocks of our regression model. The current latent life satisfaction is given by:

$$y_{it}^{*} = \beta_0' x_{it} + \beta_1 YSM_{it} + a_i + \varepsilon_{it} \quad (1)$$

where YSM_{it} is years since migration, x_{it} is a vector of other control variables, a_i is an individual fixed effect, and ε_{it} is an independent (of x) and transitory error term (say, the mood of the person on the survey day) following a logistic distribution with scale parameter 1. The previous life satisfaction, remembered from the present, is given by:

$$y_{it}^{t-1,*} = \delta_0\left(y_{i,t-1}^{*} - \varepsilon_{i,t-1}\right) + \delta_1 y_{it}^{*} + \tilde{u}_{it} \quad (2)$$

The first part of this reconstructed / recalled life satisfaction is based on the actual non-random life satisfaction that the respondent felt in the previous year (and which would be given by Equation 1 less the random term). The parameter δ_0 for an attenuated (or possibly amplified) effect of this on how the respondent remembers her past life satisfaction. The second term allows for the remembrance to be affected by current life satisfaction which is a function of current circumstances x_{it}. The psychological literature argues that the present bears heavily on people's memories of events and of their feelings (cf. Gilbert (2006), in particular Chapter 6, or O'Brien et al. (2012)). Finally, we also include an independent and transitory recall error $\tilde{u}_{it}$. The last building block is given by the threshold specification:

$$\lambda_{it}^{k} = \tau_i^{k} + \alpha_1 YSM_{it} \quad (3)$$

The thresholds are thus specific to individuals, but they are also allowed to vary over time with years since migration. Allowing for individual specific thresholds will also capture any scale differences that might exist due to cultural differences between immigrant groups. The threshold is applied to both Equations 1 and 2 by respondents who evaluate their current as well as past life satisfaction. The answers to the two questions can then be characterized by:

$$y_{it} = k \text{ iff } \tau_i^{k} - a_i < \beta_0' x_{it} + (\beta_1 - \alpha_1) YSM_{it} + \varepsilon_{it} < \tau_i^{k+1} - a_i \quad (4)$$

$$y_{it}^{t-1} = k' \text{iff } \tau_i^{k'} < \delta_0\left(\beta_0' x_{i,t-1} + \beta_1 YSM_{i,t-1} + a_i\right) + \delta_1(\beta_0' x_{it} + \beta_1 YSM_{it} + a_i + \varepsilon_{it}) - \alpha_1 YSM_{it} + \tilde{u}_{it} < \tau_i^{k'+1} \quad (5)$$

The composite error $\delta_1\varepsilon_{it} + \tilde{u}_{it}$ will not be logistically distributed even if ε_{it} and $\tilde{u}_{it}$ are, but it will approximately be logistically distributed if ε_{it} and $\tilde{u}_{it}$ are logistically distributed. Let the scale parameter of the composite error be σ. Denoting $\tilde{a}_i \equiv \frac{(\delta_0+\delta_1)a_i - \delta_0\beta_1}{\sigma}$, and substituting $YSM_{i,t-1} = YSM_{it} - 1$ for immigrants[9] we can rewrite Equation 5 as:

$$y_{it}^{t-1} = k' \text{ iff } \frac{\tau_i^{k'}}{\sigma} - \tilde{a}_i < \frac{\delta_0}{\sigma}\beta_0' x_{i,t-1} + \frac{\delta_1}{\sigma}\beta_0' x_{it} + \frac{\beta_1(\delta_0+\delta_1) - \alpha_1}{\sigma} YSM_{it} + u_{it} < \frac{\tau_i^{k'+1}}{\sigma} - \tilde{a}_i \quad (6)$$

where u_{it} now is logistically distributed with scale parameter 1. We start with the blow-up and cluster estimation of Equation 4 obtaining the parameters β_0 and the parameter $\beta_1 - \alpha_1$. With β_0 we can construct the variable $z_{it} = \beta_0' x_{it}$ and use z_{it} and $z_{i,t-1}$ in equation (6), which again we estimate with the blow-up and cluster estimator. This yields the two parameters $\frac{\delta_0}{\sigma}$ and $\frac{\delta_1}{\sigma}$, which can be interpreted as the influence of past life satisfaction and current life satisfaction in the respondent's memory of her past life satisfaction. The third parameter from this stage is $\frac{\beta_1(\delta_0+\delta_1)-\alpha_1}{\sigma}$. While we cannot point-identify β_1 and α_1, we can characterize these parameters as functions of σ.[10] That is, we would have point-estimates if we assume a certain value for σ. The only exception is if $\delta_0 + \delta_1 = 1$, in which case σ and $\beta_1 - \alpha_1$ are identified, but not β_1 and α_1 separately.

We think it helpful to classify any variable which has an effect on both the latent variable and on the threshold (in our case this is only YSM) into one of four mutually exclusive categories. Any variable which has a positive effect on the latent variable ($\beta_1 > 0$) is a "good", and otherwise a "bad". A "good" or a "bad" is "reinforcing" if its effect on the threshold is in the opposite direction (e.g. $\alpha_1 < 0$ for a good), and "adaptive" if it is in the same direction. If a variable is adaptive, an individual might not change her reported life satisfaction category even if her life satisfaction changed in response to a change of the concerned variable. If a variable is reinforcing an individual would change her life satisfaction more dramatically in response to a change of the concerned variable than she would have anticipated. To give an example, wealth might be a candidate for an adaptive good: More wealth increases satisfaction, but pushes out the thresholds for classifying oneself as wealthy. Table 10 summarizes the classification of a variable according to the sign of β_1 and α_1.

[9] For natives YSM is always zero.

[10] Point-identification would be possible if the variable were not linearly dependent over time. YSM increments by 1 from one year to the next. The function is an artefact of our ignorance of σ. It gives us information about what α_1 and β_1 must be if σ is of a certain value. The model specifies one true parameter value and singleton values for α_1 and β_1. In the model there is thus only a point rather than a function. The function does not tell us how α_1 and β_1 change if σ changes.

Table 10. Classification of variables

	$\beta_1 < 0$	$\beta_1 > 0$
$\alpha_1 < 0$	adaptive bad	reinforcing good
$\alpha_1 > 0$	reinforcing bad	adaptive good

We present the results from the threshold model in Table 11. The $\beta_1 - \alpha_1$ coefficient is 0.058, implying increasing reported life satisfaction with years since migration. All other coefficients have the usual and expected signs, maybe with the exception of age squared. The estimates for $\frac{\delta_0}{\sigma}$ and $\frac{\delta_1}{\sigma}$ also are in a range we would expect. The past is remembered quite accurately albeit with attenuated intensity, and the present bears on the memory as well: currently more satisfied people report higher past satisfaction and vice versa. Also, the ratio $\frac{\delta_0}{\delta_1}$ equals 1.8, thus past satisfaction is close to twice as relevant as current satisfaction in answering the question about past life satisfaction.

Table 11. Time-variant thresholds

	Coefficient
Years since migration	0.058**
	(0.024)
Age	-0.208***
	(0.038)
Age^2	-0.012
	(0.046)
Partner	0.590***
	(0.099)
Ln(income)	0.407***
	(0.057)
Nights in hospital	-0.004***
	(0.002)
children	0.114*
	(0.065)
employed	0.228***
	(0.064)
unemployed	-0.667***
	(0.088)
retired	0.211
	(0.170)
δ_0/σ	0.579***
	(0.070)
δ_1/σ	0.324***
	(0.078)
$\frac{\beta_1(\delta_0 + \delta_1) - \alpha_1}{\sigma}$	0.091**
	(0.038)
Observations	23,908
Persons	9,497

Standard errors in parentheses. Standard errors for the second stage estimates (lower panel) are bootstrapped from 500 replications. *** p<0.01, ** p<0.05, * p<0.1

Table 12 shows which values for β_1 and α_1 would be implied by different values for σ in a range between 0.4 to 2. The estimates imply that $\delta_0 + \delta_1 = 1$ if $\sigma \approx 1.1$. For this value of σ the parameters β_1 and α_1 would not be identified. Both parameters approach infinity when σ approaches 1.1 from the right, and negative infinity when σ approaches 1.1 from the left, which account for their high absolute values for σ = 1.0 and σ = 1.2. It is easy to show that for our estimates there is no value for σ that makes $\beta_1 < 0$ and $\alpha_1 > 0$ at the same time, so we can rule out the possibility that YSM is a reinforcing bad. YSM is a reinforcing good only if σ is less than ≈ 0.6, an adaptive good if σ is above 1.1, and an adaptive bad if σ is between 0.6 and 1.1.

Table 12. Threshold effects

σ	Implied β_1	Implied α_1
0.4	0.034	-0.024
0.6	0.007	-0.051
0.8	-0.053	-0.111
1.0	-0.340	-0.398
1.2	0.612	0.554
1.4	0.263	0.205
1.6	0.200	0.140
1.8	0.169	0.111
2.0	0.154	0.096

While we have eliminated a wide range of (α_1, β_1)[11], and the possibility of YSM being a reinforcing bad, the interval estimates for the pair (α_1, β_1) are still very wide. However, with more knowledge about σ, or equivalently about the variance of the recall error $\tilde{u}$, one could considerably narrow the range of values for α_1 and β_1. We think that learning more about the recall error will be an interesting and fruitful research area.

4. DISCUSSION AND CONCLUSION

We have analysed the pattern of life satisfaction of immigrants with attention to how it changes over the course of the stay of the immigrant in the host country. We found that on arrival immigrants report higher levels of life satisfaction than natives but that their life satisfaction decreases relative to comparable natives with time. Thus, there is a negative YSM effect. The magnitude of the effect is small, but it is statistically robust. Furthermore, it does not depend on the use of cardinal or ordinal models and increases when we use fixed effects models.

We have considered several explanations for the YSM effect. It does vanish when we control for self-assessed health measures, however given that these measures and life satisfaction are probably to a large extent driven by the same underlying variables we are sceptical about the appropriateness of including them in our model. Attrition also seemed a natural candidate to explain the YSM effect, but we saw that while systematic differences in attrition behaviour exist between natives and immigrants, and between satisfied and unsatisfied people, (un)satisfied immigrants are not more or less likely to attrite than (un)satisfied natives. Correcting for attrition bias leaves the YSM coefficient unaffected. Inclusion of other variables to proxy for the immigrants' integration or the permanence of their stay in Germany also didn't alter our YSM finding.

We also explore the existence of a response-shift on immigrants which may also explain

[11] Apart from sampling error.

the changes in life satisfaction related to the number of years since migration. We applied an innovative model to the time period 1984-1987, keeping in mind that the YSM effect is found to be positive in those years. In our model individuals reconstruct and evaluate their past life satisfaction. We showed that conditional on knowing the variance of the recall accuracy we can find point-estimates for the YSM effect on the "true" life satisfaction and on the thresholds to categorize their life satisfaction (their reporting behaviour). We ruled out the possibility that YSM is a reinforcing bad. We believe that more research on how and how accurately people can recall past events and states will enrich the field of life satisfaction research and might hold the key to identify changes in the latent variable from changes in category thresholds.

REFERENCES

Anand, P., Krishnakumar, J. and Tran, N.B., 2011. Measuring welfare: Latent variable models for happiness and capabilities in the presence of unobservable heterogeneity. *Journal of Public Economics*, 95(3), pp.205–215.

Andersen, H.H., Mühlbacher, A., Nübling, M., Schupp, J. and Wagner, G.G., 2007. Computation of standard values for physical and mental health scale scores using the SOEP version of SF-12v2. *Schmollers Jahrbuch*, 127(1), pp.171-182.

Angelini, V., Casi, L. and Corazzini, L., 2015. Life satisfaction of immigrants: does cultural assimilation matter? *Journal of Population Economics*, 28(3), pp.817-844.

Baetschmann, G., Staub, K.E. and Winkelmann, R., 2015. Consistent estimation of the fixed effects ordered logit model. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 178(3), pp.685-703.

Bartram, D., 2011. Economic migration and happiness: Comparing immigrants' and natives' happiness gains from income. *Social Indicators Research*, 103(1), pp.57-76.

Benjamin, D.J., Heffetz, O., Kimball, M.S. and Rees-Jones, A., 2014. Can marginal rates of substitution be inferred from happiness data? Evidence from residency choices. *American Economic Review*, 104(11), pp.3498-3528.

Bertrand, M. and Mullainathan, S., 2001. Do people mean what they say? Implications for subjective survey data. *American Economic Review*, 91(2), 67-72.

Blanchflower, D.G. and Oswald, A.J., 2008. Is well-being U-shaped over the life cycle? *Social science & medicine*, 66(8), pp.1733-1749.

Brazier, J. and Tsuchiya, A., 2015. Improving cross-sector comparisons: going beyond the health-related QALY. *Applied health economics and health policy*, 13(6), pp.557-565.

Chamberlain, G., 1980. Analysis of Covariance with Qualitative Data. *The Review of Economic Studies*, 47(1), pp.225-238.

Chiswick, B.R., 1978. The effect of Americanization on the earnings of foreign-born men. *Journal of political Economy*, 86(5), pp.897-921.

Cubi-Molla, P. and Yaman, F., 2015. A fixed effects ordered choice model with flexible thresholds with an application to life-satisfaction. *City University London: Discussion Paper*, 14/10.

Di Tella, R., MacCulloch, R.J. and Oswald, A.J., 2003. The macroeconomics of happiness. *Review of Economics and Statistics*, 85(4), pp.809-827.

Ed, D., Suh, E.M., Lucas, R.E. and Smith, H.L., 1999. Subjective well-being: Three decades of progress. *Psychological Bulletin*, 125(2), pp.276-302.

Easterlin, R.A., 2001. Income and happiness: Towards a unified theory. *The economic journal*, 111(473), pp.465-484.

Farré, L., 2016. New evidence on the healthy immigrant effect. *Journal of Population Economics*, 29(2), pp.365-394.

Ferrer-i-Carbonell, A. and Frijters, P., 2004. How important is methodology for the estimates of the determinants of happiness? *The Economic Journal*, 114(497), pp.641-659.

Fleurbaey, M., 2009. Beyond GDP: The quest for a measure of social welfare. *Journal of Economic literature*, 47(4), pp.1029-1075.

Frey, B.S. and Stutzer, A., 2002. What can economists learn from happiness research? *Journal of Economic Literature*, 40(2), pp.402-435.

Frijters, P., Haisken-DeNew, J.P. and Shields, M.A., 2004. Investigating the patterns and determinants of life satisfaction in Germany following reunification. *Journal of Human Resources*, 39(3), pp.649-674.

Gelatt, J., 2013. Looking down or looking up: status and subjective well-being among Asian and Latino immigrants in the United States. *International Migration Review*, 47(1), pp.39-75.

Gilbert, D., 2006. *Stumbling on happiness.* Harper Perennial.

Giuntella, O. and Mazzonna, F., 2015. Do immigrants improve the health of natives? *Journal of Health Economics*, 43, pp.140-153.

Koczan, Z., 2013. Does integration increase life satisfaction? *Cambridge Working Papers in Economics*, 1313.

Layard, R., 2006. *Happiness: Lessons from a new science*. Allen Lane.

Levinson, A., 2012. Valuing public goods using happiness data: The case of air quality. *Journal of Public Economics*, 96(9), pp.869-880.

Loewenstein, G. and Ubel, P.A., 2008. Hedonic adaptation and the role of decision and experience utility in public policy. *Journal of Public Economics*, 92(8), pp.1795-1810.

Mukherjee, B., Ahn, J., Liu, I., Rathouz, P.J. and Sánchez, B.N., 2008. Fitting stratified proportional odds models by amalgamating conditional likelihoods. *Statistics in medicine*, 27(24), pp.4950-4971.

O'Brien, E., Ellsworth, P.C. and Schwarz, N., 2012. Today's misery and yesterday's happiness: Differential effects of current life-events on perceptions of past wellbeing. *Journal of Experimental Social Psychology*, 48(4), pp.968-972.

Obućina, O., 2013. The patterns of satisfaction among immigrants in Germany. *Social indicators research*, 113(3), pp.1105-1127.

Olgiati, A., Calvo, R. and Berkman, L., 2013. Are migrants going up a blind alley? Economic migration and life satisfaction around the world: Cross-national evidence from Europe, North America and Australia. *Social Indicators Research*, 114(2), pp.383-404.

Safi, M., 2010. Immigrants' life satisfaction in Europe: Between assimilation and discrimination. *European Sociological Review*, 26(2), pp.159-176.

Stillman, S., Gibson, J., McKenzie, D. and Rohorua, H., 2015. Miserable migrants? Natural experiment evidence on international migration and objective and subjective well-being. *World Development*, 65, pp.79-93.

Schwartz, C.E., Ahmed, S., Sawatzky, R., Sajobi, T., Mayo, N., Finkelstein, J., Lix, L., Verdam, M.G., Oort, F.J. and Sprangers, M.A., 2013. Guidelines for secondary analysis in search of response shift. *Quality of Life Research*, 22(10), pp.2663-2673.

Taengnoi, S., 2014. Economic man and social woman: Determinants of immigrants' life satisfaction. *Economics Bulletin*, 34(3), pp.1530-1549.

Van Praag, B. and Baarsma, B.E., 2005. Using happiness surveys to value intangibles: The case of airport noise. *The Economic Journal*, 115(500), pp.224-246.

Wagner, G.G., Frick, J.R. and Schupp, J., 2007. The German Socio-Economic Panel Study (SOEP)–Scope, Evolution and Enhancements. *Schmollers Jahrbuch: Journal of*

Applied Social Science Studies/Zeitschrift für Wirtschafts-und Sozialwissenschaften, 127(1), pp.139-169.

Wagner, G.G., Göbel, J., Krause, P., Pischner, R. and Sieber, I., 2008. Das Sozio-oekonomische Panel (SOEP): Multidisziplinäres Haushaltspanel und Kohortenstudie für Deutschland–Eine Einführung (für neue Datennutzer) mit einem Ausblick (für erfahrene Anwender). *AStA Wirtschafts-und Sozialstatistisches Archiv*, 2(4), pp.301-328.

Wooldridge, J.M., 2002. *Econometric analysis of cross section and panel data.* MIT press.

APPENDIX A - THE FIXED EFFECTS ORDERED LOGIT MODEL

We sketch here the fixed effects ordered logit model as proposed in Mukherjee et al. (2008) and Baetschmann et al. (2015). The latent variable y_{it}^{*} is modelled as:

$$y_{it}^{*} = X_{it}\beta + a_i + \varepsilon_{it} \tag{7}$$

and the respondent's answer to an ordered choice of options is characterized by:

$$y_{it}^{t} = k \text{ iff } \tau_i^{k-1} < X_{it}\beta + a_i + \varepsilon_{it} < \tau_i^{k} \tag{8}$$

The idea of the estimator is the following: for a given cut-off value *k*, dichotomize the ordinal variable, e.g. $\tilde{y}_{it} = \mathbb{I}(y_{it} > k)$. Chamberlain's conditional logit model is derived from the likelihood of the sequence of $\tilde{y}_{it} = (\tilde{y}_{i1}, \dots \tilde{y}_{iT})$ conditional on the number of ones in the sequence of answers being equal to $\sum_{t=1}^{T_i} \tilde{y}_{it}$. Denoting this by P_i^k, the estimator combines all the possible dichotomizations (with K categories, there are K-1 possible dichotomizations) in one log likelihood and maximizes it over *b*:

$$\max_b LL = \sum_{i=1}^{N} \sum_{k=1}^{K-1} \ln P_i^k(X, b) \tag{9}$$

If ε_{it} is a logistically distributed random variable with location 0 and scale 1, an individual's likelihood $P_i^k(X, b)$ is given by:

$$P_i^k(X, b) = \frac{\prod_{t=1}^{T_i} \exp(\tilde{y}_{it} * X_{it}\beta)}{\sum_{d_i \in D_i} \prod_{t=1}^{T_i} \exp(d_{it} * X_{it}\beta)} \tag{10}$$

where $d_{it} \in \{0,1\}$, $d_i = (d_{i1} \dots d_{iT_i})$ and D_i is the set of all distinct d_i such that $\sum_t d_{it} = \sum_t \tilde{y}_{it}$.

APPENDIX B - THE HEALTH VARIABLES IN THE SOEP

In 2002, and every other year after, that the SOEP contains a battery of questions which are designed after and differ only slightly from the SF-12 health questionnaire. The questionnaire contains 12 questions and is easily found online. The 12 questions are grouped into the following eight categories:

1. Physical functioning
2. Bodily Pain
3. General Health
4. Vitality
5. Mental health
6. Social functioning
7. Role functioning (physical)
8. Role functioning (emotional)

Each of these categories are normalized to a mean of 50 and standard deviation of 10. The two variables "physical health" (PH) and "mental health" (MH) were then constructed by explorative factor analysis. The detailed description of these variables can be found in Andersen et al. (2007). The health variables we use in our analysis are number of hospital visits in the last year, PH, MH, a dummy for whether the respondent is hindered by his health, a dummy for whether the respondent suffers from a chronic illness and a dummy for whether the respondent did not see a doctor in the last year.