University of Neuchatel

Institute of Economic Research

IRENE Working paper 16-15

A Nonparametric Analysis of the Healthy Immigrant Effect

Antonio Fidalgo *Alberto Holly*

Marco Pecoraro *Philippe Wanner*

UNIVERSITÉ DE NEUCHÂTEL

Institut de recherches économiques

A nonparametric analysis of the healthy immigrant effect

Antonio Fidalgo[a,b,*], Alberto Holly[c], Marco Pecoraro[a], Philippe Wanner[b]

[a] *Swiss Forum for Migration and Population Studies, CH-2000 Neuchâtel, Switzerland*
[b] *I-Démo, University of Geneva, CH-1211 Geneva, Switzerland*
[c] *University of Lausanne, CH-1015 Lausanne, Switzerland*

December 13, 2016

Abstract

This paper uses data from the Swiss Labour Force Survey to evaluate the existence of the *healthy immigrant effect* (HIE) which would translate in *i.* a health advantage of immigrants upon their arrival in Switzerland compared to individuals with similar characteristics but Swiss-born and *ii.* an erosion of that advantage over the time of residence until convergence in the levels of health between these two groups. Our original contribution is to address this issue by taking a nonparametric approach in order overcome any potential danger of misspecification that would preclude valid inference. We find little empirical support for the HIE: *i.* no initial advantage and *ii.* no convergence but the health status of immigrants is shown to deteriorate more than Swiss-born individuals with similar characteristics. Significant differences appear when disaggregating among immigrants' country of origin. Interestingly, we show that a standard parametric approach, in contrast to our findings, would fully confirm the existence of the HIE with the same data set.

1. Introduction

Concerns with the health of the migrant population have increased in the recent decades with the growing number of immigrants in advanced countries. The issue has attracted a great deal of attention from various fields of research including social sciences and epidemiology. But its direct policy implications have also hit the headlines as anxious voices question the burden on the public health system implied by high levels of immigration, especially from relatively poor countries.

Guidance from the current literature, however, ought to be reassuring.[1] The follow-

*We would like to thank Didier Ruedin for numerous invaluable suggestions and Sara Rellstab for her help with the data. We also gratefully acknowledge comments gathered at various conferences, including the 5th OECD-CEPII, Paris, EHH2016, Lugano and 8th ICPDU, Lausanne. The research leading to these results has received funding from the Swiss National Science Foundation under grant agreement number 157027 and in the context of the NCCR on the move. Possible remaining errors are of our own responsibility.

Email address: fidalgo.antonio@gmail.com (Antonio Fidalgo)

[1] From the point of view of the host countries, not necessarily the sending countries.

ing regularity, indeed, emerges across studies. Immigrants appear to be in better health upon their arrival than a native-born with similar socio-economic characteristics—a "fact" coined *healthy immigrant effect* (HIE, henceforth). Yet, this effect is also shown to be temporary as the initial health advantage of immigrants diminishes till disappearing over the duration of the immigrant's stay in the host country.

The present study offers new evidence by addressing the HIE hypothesis[2] in a way not explored so far. The main contribution of our analysis lies in the econometric approach adopted. To our knowledge, we are the first to use nonparametric techniques in order to analyze the functional relationship between locals *vs* immigrants measure of health and a set of explanatory variables. Parametric methods require—prior to estimation—that researchers impose a functional form on the underlying, unknown data generating process. Thus, the danger exists that the formulated models suffer from incorrect specification and therefore that reliable inference is precluded.[3] Our approach overcomes that danger by using data-driven techniques in order to select the most appropriate model in the given samples. Hence, any HIE evidence obtained after relaxing the specification assumptions would represent a strong support for the existing parametric models.

The use of Swiss data is also unprecedented in the analysis of the HIE. While the absence of any study for this country could stand as a reason *per se*, we argue that Switzerland is actually a particularly good candidate to evaluate the robustness of the HIE findings. This is because the inflows of population from all origins—though mainly from the European Union—are among the highest in Switzerland relative to its population. On the other hand, the Swiss health system undeniably provides premium universal health services. Hence, any HIE evidence found in this least favorable case would necessarily carry a greater weight in discussions concerned with less advanced health systems.

Notice that we must distinguish between the relative insights of these two lines of investigation. We do so by running *parametric* estimations under standard specifications of the model for health determinants. In other words, we replicate the usual procedure with our data. The impact of our alternative econometric approach is thus evaluated with respect to that benchmark.

Our results shed a new light on the HIE hypothesis. Using a common parametric estimation, in line with the usual findings, we find *prima facie* evidence of the HIE in our data from Switzerland. However, once we abandon any prior for the functional form estimated, results become dramatically different, both quantitatively and qualitatively. Our nonparametric estimations show that, as early as they arrive in Switzerland, foreign-born are *worse* off in terms of health than native-born with similar characteristics. We interpret the discrepancies between methods as a reminder of

[2]That is, an immigrant's initial but eroding health advantage.

[3]As Monfort (1996) puts it: "The possibility of misspecification is often forgotten, perhaps because it is always frustrating to realize that a brilliant result holds only if the model under consideration is well specified."

the key importance of model specification when evaluating the determinants of health.

The plan of the paper is as follows. Section 2 reviews the literature to provide both current evidence and explanations of the HIE hypothesis. Section 3 introduces the main data set of our analysis while Section 4 explains the methodology adopted. Results are presented in Section 5. Section 6 summarizes our findings and offers concluding remarks.

2. Current evidence and explanations of the HIE

Most studies of the HIE hypothesis concentrate on a limited set of countries—mainly USA (Jasso et al., 2004, Antecol and Bedard, 2006, Abraido-Lanza et al., 1999, Thomson et al., 2013), Canada (McDonald and Kennedy, 2004, Newbold, 2005, Ng et al., 2005, Kwak, 2016) and Australia (Biddle et al., 2007, Chiswick et al., 2008, Powles, 1990).[4] Results generally confirm the existence of such an effect despite the fact that immigrants come from countries with lower levels of income and health indicators. This apparent puzzle has triggered various alternative explanations. The following lists first those offered to account for the immigrant's initial health advantage.

Selection of immigrants by the local authorities has been suggested to play a role in explaining the health gap. Selectivity can arise via many criteria, including health, skills or education. Laroche (2000) shows that health screening is unlikely to serve as an explanation as the number of applicants to Canada rejected on that ground is very small. Also, Uitenbroek and Verhoeff (2002) argue that the lower mortality of immigrants in Amsterdam do not derive from selection by authorities based on health. On the other hand, Chiswick et al. (2008) find that the type of visa granted by the Australian authorities reflects in the immigrant's level of health as skilled individuals are less likely to be in poorer health. This last result must be read in light of the current policies of many destination countries that attempt to attract skilled immigrants—e.g., UK's point-based immigration system, USA and Germany's selective policies.

The second type of explanation refers to immigrants' behaviors prior to migration. To the extent that these latter are healthier than those of the average person in the host country, an advantage in health could appear in the data. For instance, Abraido-Lanza et al. (1999) mention lower cigarette and alcohol consumption by Latino immigrants as a cause of their lower mortality in the USA.

Self-selection among potential immigrants has also been proposed as an explanation for the HIE. Thereby, individuals in the upper tail of the distribution in health and wealth—highly correlated with the former—are also the most likely to migrate from the poorer countries. Evidence for self-selection effects in other contexts such as labor market outcomes has already been identified (Borjas, 1987, Chiswick, 1999, Chiquiar and Hanson, 2005, McKenzie and Rapoport, 2010, to cite a few examples). Jasso et al. (2004), Martinez et al. (2015), Kennedy et al. (2015) or Farré (2016), among

[4]Kennedy et al. (2015) compare patterns among these three countries and the UK.

others, highlight the role of self-selection in the context of health differences between immigrants and locals.[5]

With respect to the second component of the HIE hypothesis, namely the erosion of the immigrant's health advantage over the years since migration, the following reasons have been put forward. As hypothesized by Jasso et al. (2004), the convergence in health might derive from a phenomenon of acculturation whereby immigrants adopt the local way of living. Alternatively, the mere fact of being exposed to common environmental conditions might also explain why the levels of health become similar for both groups (Stephen et al., 1994). Another aspect builds on possible barriers that immigrants face in the use of health services, at least upon their arrival. As these barriers lower with language acquisition or overtaking of cultural differences, more adverse conditions among immigrants may be detected, possibly in worse forms than they would have been should they be diagnosed earlier (see Leclere et al., 1994, for an argument along these lines). Finally, a so-called "salmon effect" would also have an impact on the apparent declining of immigrants' average level of health if less healthy immigrants—with consequent lower job outcome opportunities—were to return to their home country (Razum, 2006).

Notice that all of these potential explanations are interrelated and might simultaneously be at work. It is therefore beyond the scope of this paper to distinguish their relative contributions in the conundrum.[6] Instead, we restrict our study to establishing new reliable evidence.

Finally, it is worth mentioning that general agreement on the existence of a HIE does stand for unanimity. For instance, Moullan and Jusot (2014) find that immigrants report either poorer or better health depending on which European country is examined.[7] This suggests that the relevance of the HIE may depend on the status of immigrants and immigration policy in the host country. Focusing on a cohort of refugees and family reunited immigrants in Denmark (relative to local-born Danes), Norredam et al. (2014) generally do not support the HIE. In the case of Israel where immigration is based on ideological rather than economic reasons, Constant et al. (2015) show evidence of a "sick immigrant effect", according to which immigrants arrive with a lower health status than natives and this status persists for several decades *ceteris paribus*. On the other

[5]Rubalcava et al. (2008) explores the self-selection issue and finds weak evidence of the HIE, though.

[6]Further explanations for the HIE have also been mentioned in the literature. In a rather unique natural experiment, Gibson et al. (2013) compare the health of Tongan immigrants that were accepted in New Zealand through a random lottery with the health of the unsuccessful candidates. Divergence in various measures between the two groups are interpreted by the authors as evidence of a direct impact of migration *per se* on the immigrant's health. Yet another view attributes to differences in perceptions of health between groups—e.g., immigrants *vs* locals, one origin *vs* another—part of the observed HIE (McDonald and Kennedy, 2004).

[7]A possible explanation for this finding may lie in the choice of not controlling for the immigrant's duration of residence in the host country. As mentioned above, there are reasons and evidence to expect the assimilation pattern to affect the immigrant's health. Hence, pooling individuals over that criterion may result in compounding contradictory effects.

hand, Solé-Auró and Crimmins (2008) find little evidence of the HIE only at high ages in their sample of European countries. McDonald and Kennedy (2004), for immigrants in Canada, find that the HIE holds for some measures of health but not others.

3. Data

3.1. The sample

The data in this analysis consists of observations from the Swiss Labour Force Survey, spanning from 2010 to the latest available data in 2014. The data have a longitudinal structure but attrition appears to be high and the period length between each wave is too short. For that reason, along with computational considerations, we treat each year separately.

Data is further screened in the following two ways. First, only male individuals are included as there is evidence that different dynamics are at play with female immigrants (see, for instance, McDonald and Kennedy, 2004, Martinez et al., 2015). Second, individuals' age is restricted to the range 18–65 years. This is close to the usual choices[8] and reflects interest in the labour force as opposed to the peculiarities related to health problems of retired individuals. The sample sizes range from 22'720 to 25'216 observations per year.

3.2. Determinants of health

Following the general approach in the literature, we adopt the perspective outlined in Dunn and Dyck (2000) where the observed health is determined, along with age, by socio-economic and cultural factors instead of other measures related to medical inputs or service utilization.

We capture socio-economic factors by including the following three variables. First, education is measured by a categorical factor with three possible values according to the highest level of education attained by the individual: 'primary', 'secondary' and 'tertiary'.[9] Second, work status allows to distinguish between the employed, the unemployed and those out of the labor force. Finally, marital status is differentiated into four categories such as single, married, separated and widow individuals.

Cultural determinants are taken into account by including the nationality of the individuals, in appropriate groups.[10] Also, often-cited cultural differences within Switzer-

[8]For instance, it is 18–65 in Moullan and Jusot (2014); 18–64 in Martinez et al. (2015); 20–65 in McDonald and Kennedy (2004); 21–65 in Kennedy et al. (2015).

[9]'Primary' stands for complete or incomplete compulsory school. 'Secondary' includes mainly apprenticeship and high-school studies. 'Tertiary' covers all university degrees as well as vocational high schools.

[10]These include: Switzerland, Germany, Italy, Spain and Portugal, Eastern Europe (Albania, Bulgaria, Poland, Romania, Hungary, Slovakia, Czech Republic, Serbia, Montenegro, Croatia, Slovenia, Bosnia and Herzegovina, Macedonia, Kosovo, Estonia, Latvia, Lithuania, Moldova, Russia, Ukraine, Belarus), Western Europe (Andorra, Belgium, Denmark, Finland, France, Greece, United Kingdom, Ireland, Iceland, Liechtenstein, Luxembourg, Malta, Monaco, Netherlands, Norway, Austria, Sweden, San Marino, Cyprus) and Rest of the world.

land are proxied by the language used in the interview by the individual surveyed.[11]

Finally, the main variable of interest in order to evaluate the HIE hypothesis is the following. Three groups of individuals are identified. These are the Swiss-born individuals, "CH", our reference group, and two groups of foreign-born individuals. Among these latter, we call newcomers, "NC", those who have arrived in Switzerland within the previous five years and old-timers, "OT", those who had stayed in the country for longer than five years. Notice that the number of years since migration considered to reflect the "upon arrival" condition was borrowed from the literature where it ranges from 5 to up to 10 years (as in, for instance, Kennedy et al., 2015, McDonald and Kennedy, 2004).

3.3. Measure of health

Our chosen indicator of physical health is the individual's self-assessed health status—SAHS—on a five-level scale which we recode into a numeric variable: very good (100), good (80), so so (60), poor (40) or very poor (20). Two notes of caution are worth underlining. First, measures of SAHS are often used in the literature despite possible limitations related to variations in perceptions of own health over age, gender, origin or further characteristics. The validity of such a measure, however, relies on its documented strong correlation with more objective factors such as mortality and morbidity, as argued in Biddle et al. (2007) or in the review of twenty-seven international studies by Idler and Benyamini (1997). Second, we follow one of the two main options in the existing literature across fields for using a polychotomous survey variable by transforming it into a continuous variable (see, for instance, Ortega and Polavieja, 2012). The common alternative of using a 0/1 indicator (less than 'good' versus 'good' or higher SAHS) such as in Newbold (2005) or McDonald and Kennedy (2004) could reduce the inherent information and is less suitable for our nonparametric approach.

3.4. Descriptive features

Figure 1 provides descriptive features of the individuals average of self-assessed health status plotted in three relevant dimensions. The top panel represents, for each year, the unadjusted average health of the three identified groups: Swiss-born, newcomers and old-timers individuals. It reveals that, before accounting for further individual characteristics, newcomers in Switzerland are, on average, in better health than the locals while the old-timers seem worse of in terms of health. Focusing on the assimilation patterns of the immigrants as measured by their health status, this first plot tends to indicate that the health of immigrants deteriorates over their stay in Switzerland.

It is not clear, however, that a causal mechanism is specifically at work for the immigrant population. Indeed, holding age of arrival constant, a longer duration of residence implies an older age. Hence, the declining health could be an effect of age alone. As illustrated by the bottom panel of Figure 1 for all individuals in the sample, health status seem unsurprisingly negatively related to age.

[11] These are one of the following: German, French, Italian and English.

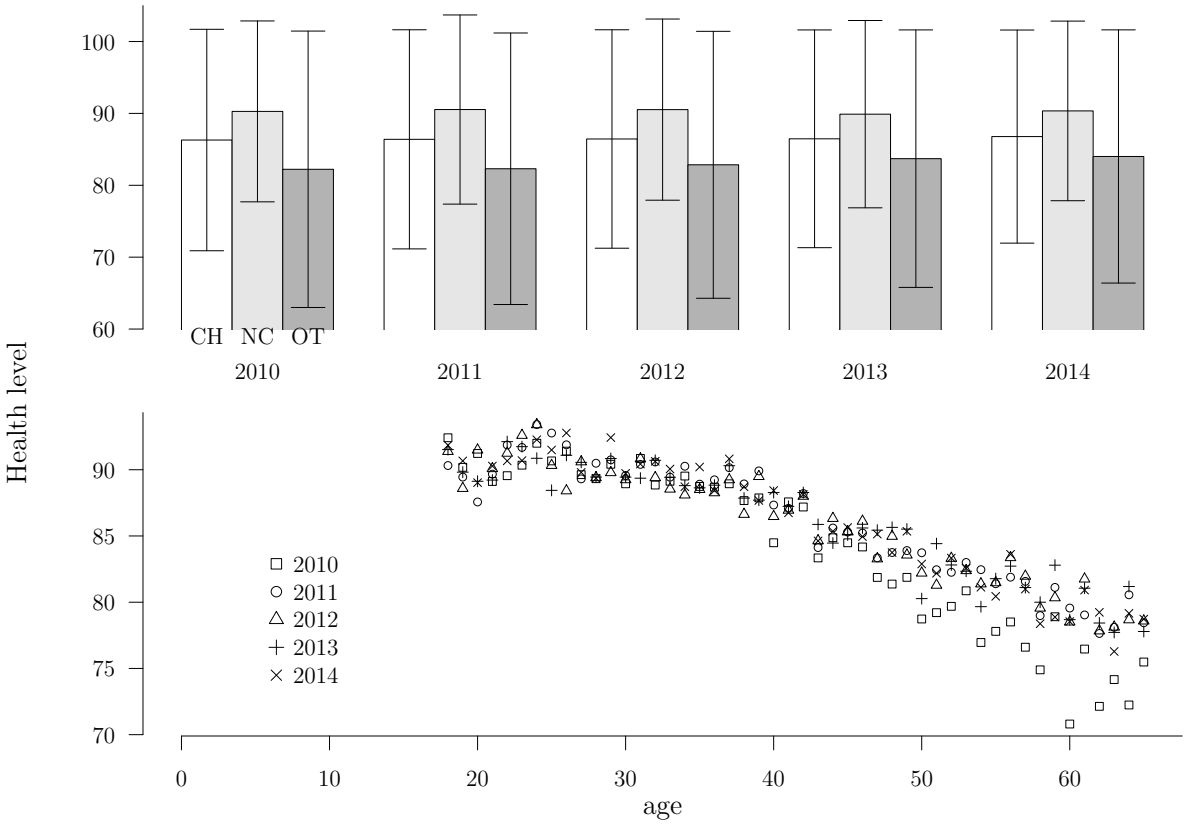

Figure 1: Simple averages of the health level for each identified group of individuals—Swiss-born 'CH', newcomers 'NC' and old-timers 'OT'; bounds of one standard deviation from the mean are drawn (top panel) and for each ceiling of age of all individuals (bottom panel).

4. Methodology

A major drawback in previous studies is that their conclusions hinge on the validity of the particular forms of their estimated models. The danger therefore exists of falling in either or both of two traps associated with the presence of misspecification. On one hand, using an incorrect functional form can lead to a quantitatively unknown bias in the estimates. The latter could amount up to a change in sign of a coefficient, a case that could prove pivotal in the context of the estimation of the HIE effect. On the other hand, as is well established, the usual tests—e.g., t statistics—do not exhibit correct size and power, leading to a potentially erroneous evaluation of the statistical significance of a variable in the model. Our empirical investigation avoids that twofold danger by adopting a nonparametric approach whereby no assumption is made on the functional relationship between the explained variable and its determinants.

Let y_i be the self-assessed health status of individual i, hg_i the health group of the individual and X_i the set of explanatory variables mentioned above. Then, the function

of health may be written as:

$$y_i = f(hg_i, X_i) + u_i, \quad i = 1, \ldots, n, \tag{1}$$

where u_i is a disturbance term and where the sample realizations (y, hg, X) are i.i.d. The function f is unknown but presumed twice continuously differentiable.

The object of interest is $f(hg, X)$, the conditional mean of the variable y given a realization (hg, X). We use a local constant kernel estimator described in Li and Racine (2007, pp. 60–63) as a consistent estimate of the conditional mean, $\hat{f}(hg, X)$.

The core of this method is the local averaging/smoothing of the values of y which are "close" in terms of the values taken by the regressors. Its consistency relies on the amount of local information—the number of observations—available in each neighborhood whose length is controlled by a bandwidth parameter. In high-dimensional spaces,[12] however, the number of total "well-behaved" observations required to ensure enough local information in small neighborhoods increases dramatically. This problem with nonparametric smoothing estimation is known as the *curse of dimensionality*.

A few comments to justify our method are the following. First, the sizes of our samples are relatively large as they include between 22'720 and 25'216 observations. Second, let q be the number of continuous variables of our estimation. It can be shown (Li and Racine, 2003) that the rate of convergence of the estimation bias is, for the mixed data case with the least squares cross-validation of bandwidths adopted here,

$$\hat{f}(\cdot) - f(\cdot) = \mathcal{O}_q(n^{-2/(q+4)}). \tag{2}$$

While slower than the usual parametric methods, this rate only applies to the *continuous* variables involved and not the categorical variables. In the present problem, only one—age—of the regressors is continuous. The convergence of our bias towards zero is therefore not desperately slow and likely to be achieved with our samples sizes. More fundamentally, it is worth emphasizing that the alternative is not between slow and fast procedures but between misspecified and therefore inconsistent parametric models and relatively inefficient but consistent nonparametric models. The present studies is the first to propose one of the latter that can then be compared with all the existing examples of the former.

Also, when evaluating the precision of our estimates, we do not rely on the asymptotics of well-behaved data but use a bootstrap method instead.

The parametric benchmark for comparison, in the this context, is borrowed from McDonald and Kennedy (2004) or Biddle et al. (2007). Using the notation above, the parametric model for health may be written as:

$$y_i = \alpha + \beta hg_i + \tilde{X}_i'\gamma + u_i, \quad i = 1, \ldots, n, \tag{3}$$

where $\tilde{X}_i$ is X_i with an extra variable age^2.

[12] The dimensions being the regressors considered in the model.

5. Results

Our nonparametric estimates of the HIE are given in Figure 2. For each yearly model, we provide estimated health levels for each health group. We represent these latter in the form of boxplots with notches as they allow a direct comparison of the medians between the three groups. Indeed, as argued for instance in Chambers et al. (1983), if two notches do not overlap, then there is strong evidence that their medians differ.

5.1. Health differential upon arrival

The first component of the HIE hypothesis states that, upon arrival, immigrants enjoy a health advantage with respect to comparable locals. In our plots, such advantage would translate into a different estimate between Swiss-born and Newcomers. Close inspection of the yearly graphs fails to find any significant evidence of that effect. If, in some occurrences, the median health level of Newcomers is higher than Swiss-born individuals with similar characteristics, the reverse is also found in other yearly models.

A common feature of these opposite estimates, however, is their lack of significance. Indeed, the differences in the estimated levels are small and the medians remain in the range of the notch of the other group. Our results therefore suggest that immigrants in our data do not have a health advantage upon their arrival in Switzerland.

5.2. Immigrant's pattern of the health level over the duration of their stay

The second part of the HIE hypothesis lends itself to a few interpretations. On one hand, if the acculturation explanation is to hold, the HIE hypothesis implies that the health of immigrants *converges* towards the health of the similar local. Presumably, these two levels ought ultimately to be very close. On the other hand, in the less strict interpretation of the hypothesis given by the barriers to health view or the "salmon effect", the level of health of the immigrants is only expected to *decrease* over the duration of their stay.

We address this hypothesis—in either weak or strong form—by comparing the estimated health levels of Swiss-born individuals and immigrants with similar characteristics who lived in the country for more than five years. Again, Figure 2 provides that comparison.

Our results clearly indicate significant differences between these two groups. After accounting for other variables, foreign-born individuals who have stayed a long time in Switzerland appear to be in worse health than Swiss-born individuals. Recall that, according to our results above, there was no difference in health upon their arrival. Therefore, our evidence suggests that the health of foreign-born individuals deteriorates faster than that of Swiss-born individuals. This finding, in turn, supports the second part of the HIE hypothesis whereas the level of health of the immigrants decreases over the duration of their stay for reasons beyond those controlled for in the model such as age.

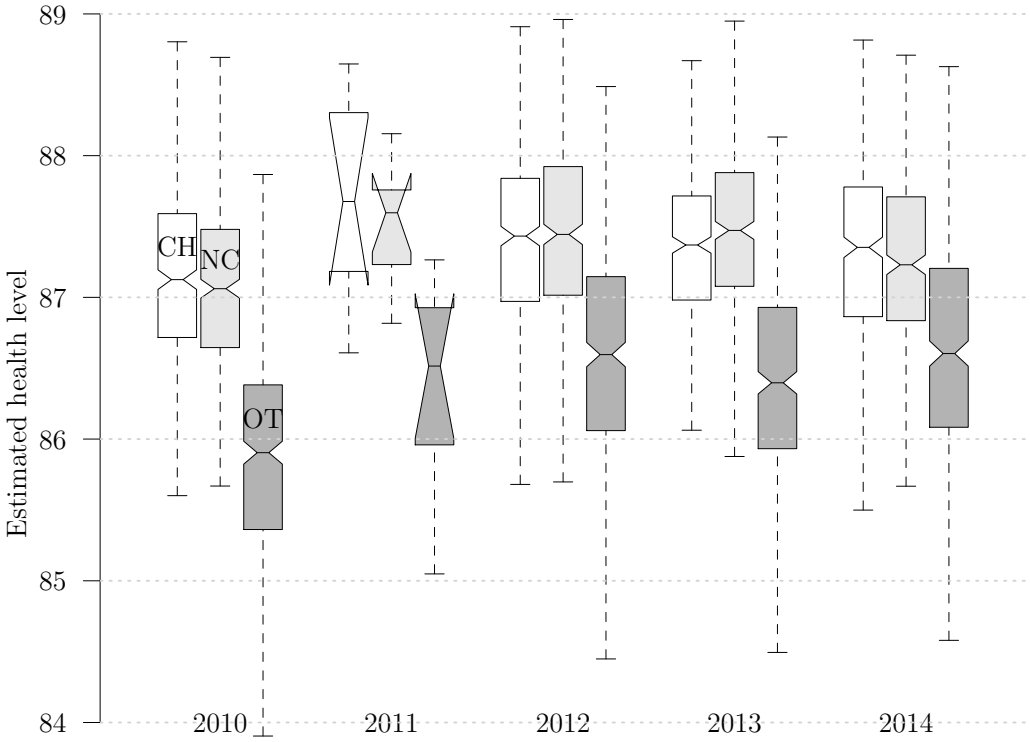

Figure 2: Nonparametric estimates of the health for different health groups—Swiss-born 'CH', newcomers 'NC' and old-timers 'OT'

5.3. Parametric benchmark

Using a nonparametric approach, we show mixed support for the HIE hypothesis in our data set: no health advantage of the immigrants upon arrival but a declining health with the number of years since migration. The reader may wonder whether this result can be attributed to our technical approach or whether standard approaches would have been able to match this result.

We answer that concern by conducting a conventional OLS estimation of the parametric model of McDonald and Kennedy (2004) and Biddle et al. (2007) as outlined in Equation 3. The main results of this exercise are given in Table A.3, focusing on the main variable of interest, the health group. The interpretation of the coefficients is unambiguous. Other things equal, newcomers are significantly in better health than the reference group of Swiss-born individuals. For the old-timers, the negative and significant coefficient indicates that individuals in this group are worse off in terms of health.

Overall, our parametric results would provide strong evidence for the HIE hypothesis. In turn, this finding highlights the pivotal of the assumptions on the functional form of the conditional mean of health in the population.

Table 1: OLS estimation of the parametric model in Equation 3.

Variable	Yearly parametric model				
	2010	2011	2012	2013	2014
NEWCOMER	1.37***	1.57***	1.74***	1.08***	1.34***
	(0.339)	(0.342)	(0.334)	(0.357)	(0.367)
OLDTIMER	−1.92***	−1.94***	−1.56***	−1.17***	−1.1***
	(0.252)	(0.238)	(0.229)	(0.228)	(0.227)
Adj. R^2	0.17	0.17	0.17	0.15	0.14
Num. obs.	22720	24510	25216	24397	24068

Robust standard errors in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.
Control variables include: age, age squared, work status, civil status and language of interview.

5.4. Health differences and origin of the immigrants

In this section, we evaluate whether the country of origin of the immigrants affects the health differentials with respect to Swiss-born individuals.

The diversity of these countries of origin, however, calls for a first choice about some appropriate grouping. We choose the following six groups according to the nationality of the immigrants: *i.* Germany, *ii.* Italy, *iii.* Spain or Portugal, *iv.* other Western Europe, *v.* Eastern Europe and *vi.* rest of the world.

Figure 3 provides the estimates for the 2010 yearly model as a typical result of this exercise. Several conclusions can be drawn: we mention two. First, these estimates confirm our finding above of a stronger decline of immigrants' health over the duration of their stay in the host country. Old-timers in all but one group have a significantly lower median health level compared to both the Swiss-born and the newcomers with otherwise similar characteristics.

The new and main insight from Figure 3, however, is that immigrants' origin matters a great deal when evaluating their health status relative to the individuals born in the country. This is particularly true when it comes to evaluate the immigrants' health advantage upon arrival. Consider, for instance, the group of German individuals. Their health level does not seem to differ from the health level of their Swiss-born counterpart—an absence of any immigrants' advantage that we have already illustrated above. Notice, however, the striking contrast with that respect with the other national groups. Individuals from other countries than Germany are already in significantly poorer health than similar locals when they arrive in Switzerland—a health disadvantage that moreover grows over time.

This finding, in turn, is strongly at odds with the largest part of the literature whereas support for the HIE hypothesis is provided by parametric estimations. Furthermore, it also draws attention to the importance of the composition of the overall immigrant population—i.e., the origin of the immigrants matters,[13] an aspect that is

[13]Removing the German immigrants from the analysis, for instance, unsurprisingly reveals a signif-

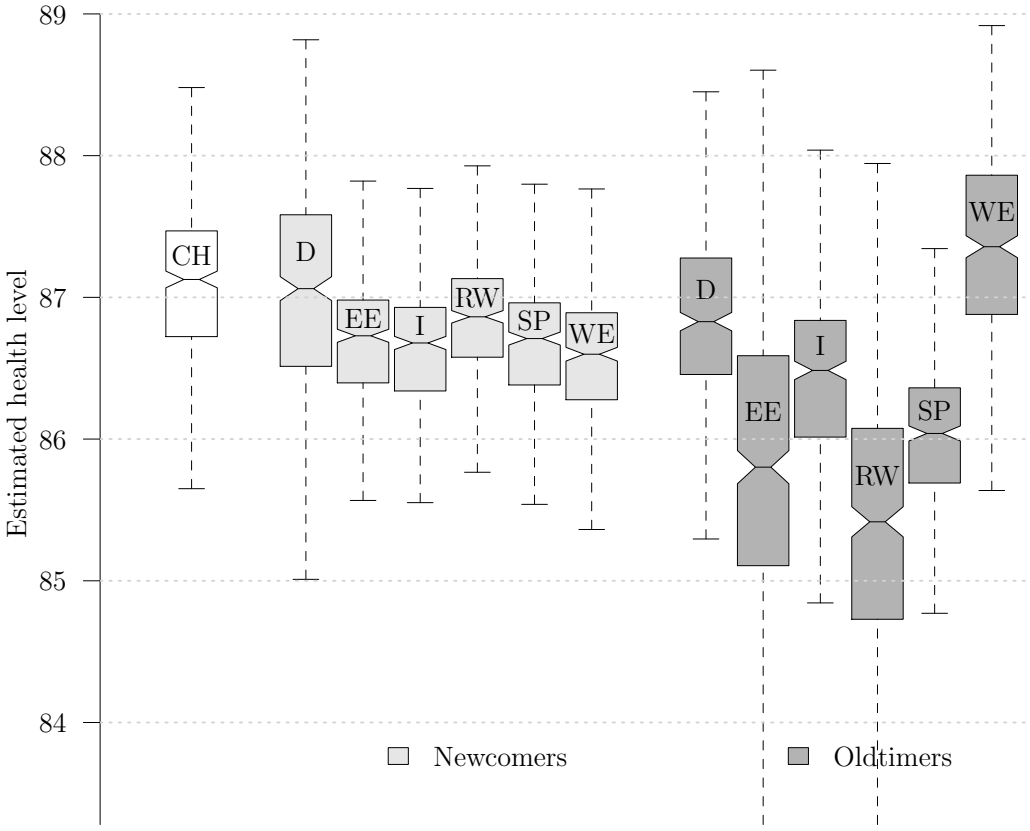

Figure 3: Nonparametric estimates of the health for different origins of immigrants. 2010 data. Results for other years are displayed in the appendix. Key: CH, Swiss-born individuals; D, Germany; EE, Eastern Europe; I, Italy; RW, Rest of the world; SP, Spain and Portugal; WE, Western Europe.

sometimes neglected (see Moullan and Jusot, 2014, however, for an example that reaches a similar conclusion).

6. Conclusion

This paper presents new empirical evidence on the validity of the HIE hypothesis. Using the data from the Swiss Labour Force Survey and taking a nonparametric approach in order overcome potential misspecification that would preclude valid inference, we find a number of robust results. First, we fail to find evidence of a health advantage for immigrants upon their arrival in Switzerland. Subsequently, their level of health is shown to significantly decrease during their stay in the host country. Therefore, instead of converging towards the level of health of a Swiss-born individual with similar characteristics, as current support of the HIE hypothesis has it, we find clear evidence

icant health disadvantage of the immigrants in several samples.

of divergence between Swiss-born and foreign-born individuals.

The findings in this study are at odds with most of the results established so far in the literature. However, we show that a traditional econometric strategy with our data would also have reached an opposite result and confirmed the HIE. This apparent puzzle highlights the key importance model specification when evaluating the determinants of health.

Notice that the confirmation of the HIE is often transposed into comforting policy implications such as "immigration is not likely to represent a financial burden on the public health system" (Farré, 2016). In the light of our results, we would argue that the health of immigrants in Switzerland deserves closer scrutiny than previously thought.

References

Abraido-Lanza, A. F., Dohrenwend, B. P., Ng-Mak, D. S., and Turner, J. B. (1999). The Latino mortality paradox: A test of the "salmon bias" and healthy migrant hypotheses. *American Journal of Public Health*, 89(10):1543–1548.

Antecol, H. and Bedard, K. (2006). Unhealthy assimilation: Why do immigrants converge to american health status levels? *Demography*, 43(2):337–360.

Biddle, N., Kennedy, S., and McDonald, J. T. (2007). Health assimilation patterns amongst Australian immigrants. *Economic Record*, 83(260):16–30.

Borjas, G. J. (1987). Self-selection and the earnings of immigrants. *The American Economic Review*, 77(4):531–553.

Chambers, J., Cleveland, W., Kleiner, B., and Tukey, J. (1983). Comparing data distributions. *Graphical Methods for Data Analysis (Bell Laboratories)*, pages 47–73.

Chiquiar, D. and Hanson, G. H. (2005). International migration, self-selection, and the distribution of wages: Evidence from mexico and the united states. *Journal of Political Economy*, 113(2):239–281.

Chiswick, B. R. (1999). Are immigrants favorably self-selected? *The American Economic Review*, 89(2):181–185.

Chiswick, B. R., Lee, Y. L., and Miller, P. W. (2008). Immigrant selection systems and immigrant health. *Contemporary Economic Policy*, 26(4):555–578.

Constant, A. F., García-Muñoz, T., Neuman, S., and Neuman, T. (2015). A "healthy immigrant effect"or a "sick immigrant effect"? Selection and policies matter. Discussion paper no. 9338, Institute for the Study of Labor (IZA).

Dunn, J. R. and Dyck, I. (2000). Social determinants of health in Canada's immigrant population: Results from the National Population Health Survey. *Social Science & Medicine*, 51(11):1573–1593.

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

Gibson, J., Stillman, S., McKenzie, D., and Rohorua, H. (2013). Natural experiment evidence on the effect of migration on blood pressure and hypertension. *Health Economics*, 22(6):655–672.

Idler, E. L. and Benyamini, Y. (1997). Self-rated health and mortality: A review of twenty-seven community studies. *Journal of Health and Social Behavior*, 38(1):21–37.

Jasso, G., Massey, D. S., Rosenzweig, M. R., and Smith, J. P. (2004). Immigrant health: selectivity and acculturation. *Critical perspectives on racial and ethnic differences in health in late life*, pages 227–266.

Kennedy, S., Kidd, M. P., McDonald, J. T., and Biddle, N. (2015). The healthy immigrant effect: Patterns and evidence from four countries. *Journal of International Migration and Integration*, 16(2):317–332.

Kwak, K. (2016). An evaluation of the healthy immigrant effect with adolescents in canada: Examinations of gender and length of residence. *Social Science & Medicine*, 157:87–95.

Laroche, M. (2000). Health status and health services utilization of Canada's immigrant and non-immigrant populations. *Canadian Public Policy/Analyse de Politiques*, 26(1):51–75.

Leclere, F. B., Jensen, L., and Biddlecom, A. E. (1994). Health care utilization, family context, and adaptation among immigrants to the United States. *Journal of Health and Social Behavior*, 35(4):370–384.

Li, Q. and Racine, J. S. (2003). Nonparametric estimation of distributions with categorical and continuous data. *Journal of Multivariate Analysis*, 6(2):266–292.

Li, Q. and Racine, J. S. (2007). *Nonparametric Econometrics: Theory and Practice*. Princeton University Press.

Martinez, J. N., Aguayo-Tellez, E., and Rangel-Gonzalez, E. (2015). Explaining the Mexican-American health paradox using selectivity effects. *International Migration Review*, 49(4):878–906.

McDonald, J. T. and Kennedy, S. (2004). Insights into the 'healthy immigrant effect': Health status and health service use of immigrants to Canada. *Social Science & Medicine*, 59(8):1613–1627.

McKenzie, D. and Rapoport, H. (2010). Self-selection patterns in mexico-us migration: The role of migration networks. *The Review of Economics and Statistics*, 92(4):811–821.

Monfort, A. (1996). A reappraisal of misspecified econometric models. *Econometric Theory*, 12(04):597–619.

Moullan, Y. and Jusot, F. (2014). Why is the 'healthy immigrant effect' different between European countries? *The European Journal of Public Health*, 24(suppl 1):80–86.

Newbold, B. (2005). Health status and health care of immigrants in Canada: A longitudinal analysis. *Journal of Health Services Research & Policy*, 10(2):77–83A.

Ng, E., Wilkins, R., Gendron, F., and Berthelot, J.-M. (2005). Dynamics of immigrants' health: Evidence from Canada, 1994-95 to 2002-03. Summary Report. *Canadian Studies in Population*, 32(1):131–133.

Norredam, M., Agyemang, C., Hansen, H., Oluf, K., Petersen, J. H., Byberg, S., Krasnik, A., and Kunst, A. E. (2014). Duration of residence and disease occurrence among refugees and family reunited immigrants: Test of the 'healthy migrant effect' hypothesis. *Tropical Medicine & International Health*, 19(8):958–967.

Ortega, F. and Polavieja, J. G. (2012). Labor-market exposure as a determinant of attitudes toward immigration. *Labour Economics*, 19(3):298–311.

Powles, J. (1990). The best of both worlds: Attempting to explain the persisting low mortality of Greek migrants to Australia. *What we know about Health Transition: The Cultural, Social and Behavioural Determinants of Health*, pages 584–594.

Razum, O. (2006). Commentary: Of salmon and time travellers–musing on the mystery of migrant mortality. *International Journal of Epidemiology*, 35(4):919–921.

Rubalcava, L. N., Teruel, G. M., Thomas, D., and Goldman, N. (2008). The healthy migrant effect: New findings from the Mexican family life survey. *American Journal of Public Health*, 98(1):78–84.

Solé-Auró, A. and Crimmins, E. M. (2008). Health of immigrants in European countries. *International Migration Review*, 42(4):861–876.

Stephen, E. H., Foote, K., Hendershot, G. E., and Schoenborn, C. A. (1994). Health of the foreign-born population: United States, 1989-90. *Advance data*, (241):1–12.

Thomson, E. F., Nuru-Jeter, A., Richardson, D., Raza, F., and Minkler, M. (2013). The hispanic paradox and older adults disabilities: is there a healthy migrant effect? *International Journal of Environmental Research and Public Health*, 10(5):1786–1814.

Uitenbroek, D. G. and Verhoeff, A. P. (2002). Life expectancy and mortality differences between migrant groups living in Amsterdam, the Netherlands. *Social Science & Medicine*, 54(9):1379–1388.

Appendix

Table A.1: OLS estimation of the parametric model in Equation 3—full table.

Variable	Yearly parametric model				
	2010	2011	2012	2013	2014
Newcomer	1.37***	1.57***	1.74***	1.08***	1.34***
	(0.339)	(0.342)	(0.334)	(0.357)	(0.367)
Oldtimer	−1.92***	−1.94***	−1.56***	−1.17***	−1.1***
	(0.252)	(0.238)	(0.229)	(0.228)	(0.227)
Age	−1.08***	−1.11***	−1.14***	−1.1***	−1.1***
	(0.063)	(0.06)	(0.059)	(0.06)	(0.058)
Age^2	0.01***	0.01***	0.01***	0.01***	0.01***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Education					
Secondary	4.63***	4.23***	4.77***	4.63***	4.34***
	(0.365)	(0.339)	(0.34)	(0.354)	(0.347)
Tertiary	8.28***	8.06***	8.5***	8.52***	8.09***
	(0.376)	(0.348)	(0.347)	(0.358)	(0.353)
Work status					
Not active	−13.92***	−14.07***	−13.54***	−12.16***	−11.69***
	(0.519)	(0.504)	(0.497)	(0.492)	(0.491)
Unemployed	−5.14***	−4.75***	−4.33***	−2.36***	−4***
	(0.631)	(0.634)	(0.63)	(0.558)	(0.577)
Civil status					
Married	1.68***	1.94***	2.13***	1.58***	1.67***
	(0.257)	(0.246)	(0.24)	(0.24)	(0.244)
Separated	0.25	0.83	0.81	0.05	0.4
	(0.427)	(0.404)	(0.394)	(0.411)	(0.406)
Widow	1.4	−0.49	2.44	1.74	1.82
	(1.344)	(1.357)	(1.373)	(1.23)	(1.369)
Lang. of interview					
French	−1.1***	−1.18***	−0.6***	−0.47***	−0.35***
	(0.259)	(0.25)	(0.243)	(0.244)	(0.239)
Italian	−3.1***	−2.88***	−2.62***	−2.51***	−2.03***
	(0.411)	(0.403)	(0.379)	(0.386)	(0.368)
Other	1.98***	0.7***	0.93***	1.44***	0.8***
	(0.608)	(0.622)	(0.603)	(0.614)	(0.671)
Intercept	109.68***	110.47***	110.96***	109.87***	110.34***
	(1.176)	(1.129)	(1.113)	(1.143)	(1.088)
Adj. R^2	0.17	0.17	0.17	0.15	0.14
Num. obs.	22720	24510	25216	24397	24068

Robust standard errors in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.
Reference values for factor variables are Swiss-born, primary (education), active (work status), single (civil status) and German (language of interview).

Table A.2: Summary of nonparametric estimations in Figure 2

Variable	Bandwidth in yearly nonparametric model				
	2010	2011	2012	2013	2014
Group	0.2595	0.2104	0.2282	0.2234	0.2802
Age	1.0326	1.1496	0.9959	1.2872	0.8006
Education	0.052	0.0378	0.049	0.0364	0.0465
Work status	0.0027	0.0046	0.0029	0.0025	0.0026
Civil status	0.2939	0.1899	0.2074	0.2291	0.3324
Lang. of int.	0.4229	0.3738	0.4069	0.4238	0.4496
R^2	0.29	0.29	0.29	0.28	0.29
Num. obs.	22676	24465	25169	24344	24022

Continuous kernel type: second-order Gaussian. Unordered categorical kernel type: Aitchison and Aitken. Fixed bandwidth type.
Group refers to the factor variable including three values: Swiss-born, Newcomers and Oldtimers.

Table A.3: Summary of nonparametric estimations in Figure 3

Variable	Bandwidth in yearly nonparametric model				
	2010	2011	2012	2013	2014
Country group	0.5131	0.4923	0.5513	0.4925	0.6241
Age	1.2138	1.8421	1.2521	1.777	0.8829
Education	0.0529	0.0378	0.0566	0.0317	0.0484
Work status	0.0026	0.0022	0.0026	0.0024	0.0037
Civil status	0.3453	0.2035	0.2382	0.2913	0.3447
Lang. of int.	0.5167	0.4261	0.424	0.4509	0.5209
R^2	0.31	0.29	0.31	0.27	0.28
Num. obs.	21670	23338	24007	23195	22733

Continuous kernel type: second-order Gaussian. Unordered categorical kernel type: Aitchison and Aitken. Fixed bandwidth type.
Country group refers to the factor variable including the country groups in Figure 3 and Figure A.1–A.4.

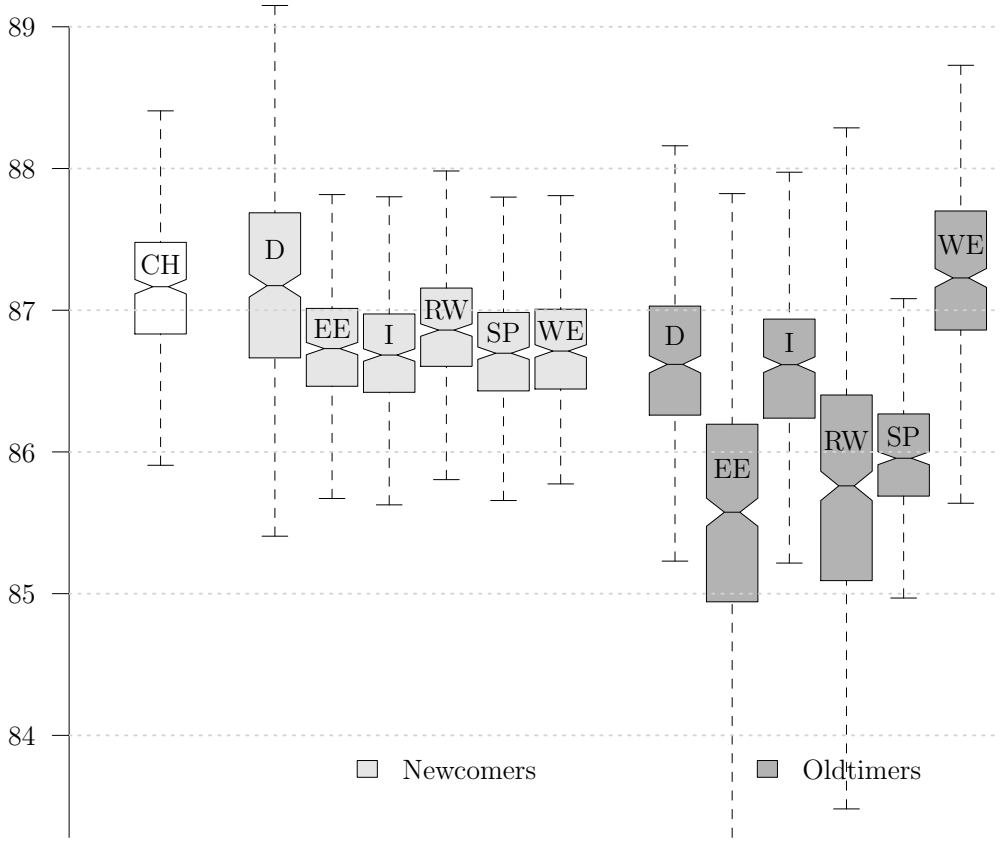

Figure A.1: Nonparametric estimates of the health for different origins of immigrants. 2011 data. Key: CH, Swiss-born individuals; D, Germany; EE, Eastern Europe; I, Italy; RW, Rest of the world; SP, Spain and Portugal; WE, Western Europe.

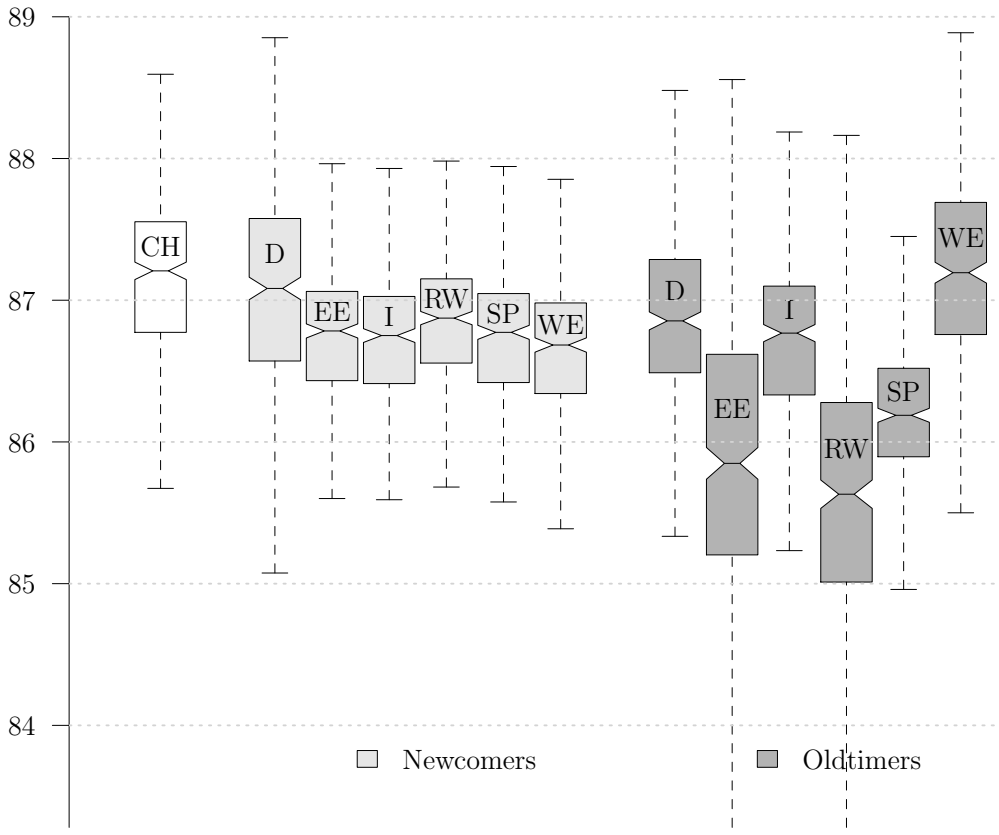

Figure A.2: Nonparametric estimates of the health for different origins of immigrants. 2012 data. Key: CH, Swiss-born individuals; D, Germany; EE, Eastern Europe; I, Italy; RW, Rest of the world; SP, Spain and Portugal; WE, Western Europe.

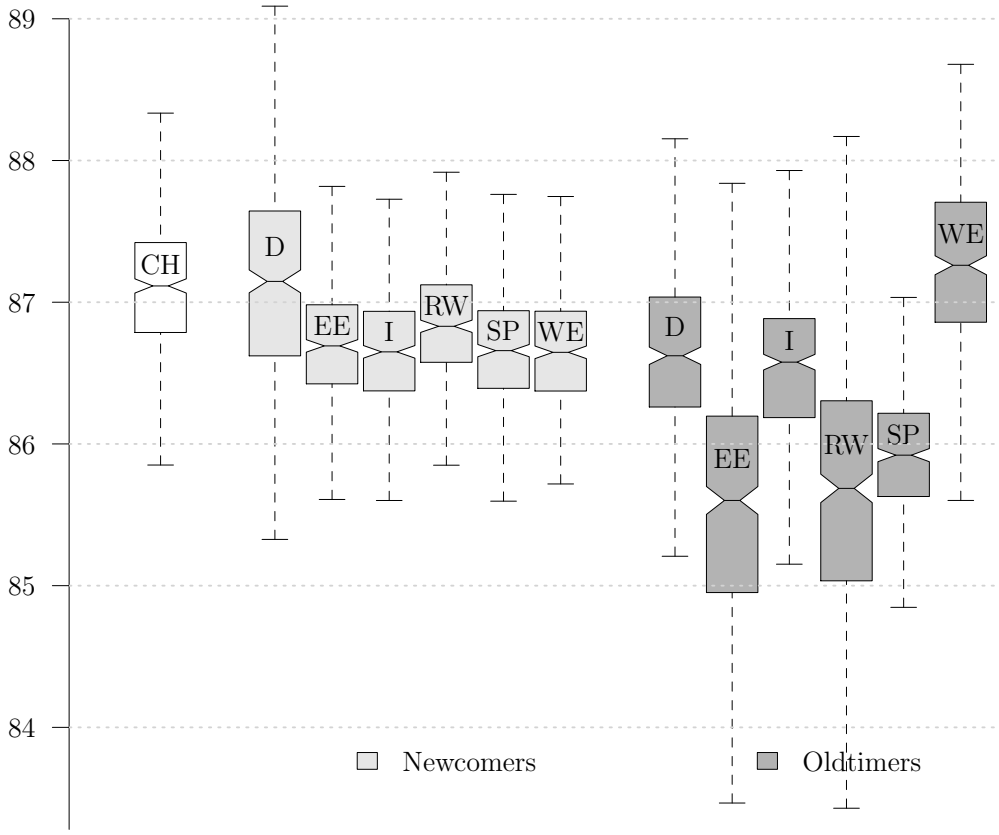

Figure A.3: Nonparametric estimates of the health for different origins of immigrants. 2013 data. Key: CH, Swiss-born individuals; D, Germany; EE, Eastern Europe; I, Italy; RW, Rest of the world; SP, Spain and Portugal; WE, Western Europe.

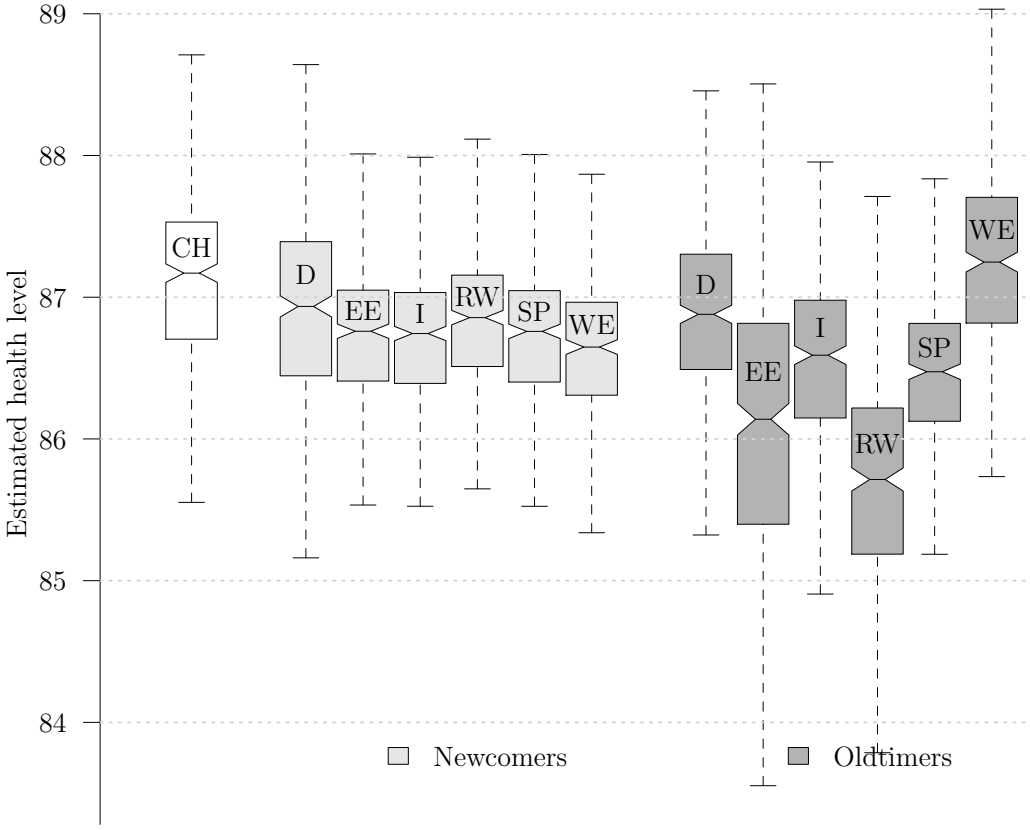

Figure A.4: Nonparametric estimates of the health for different origins of immigrants. 2014 data. Key: CH, Swiss-born individuals; D, Germany; EE, Eastern Europe; I, Italy; RW, Rest of the world; SP, Spain and Portugal; WE, Western Europe.