Submitted: 27th of August 2014
Published: 9th of October 2014

# Crime, income, educational attainment and employment among immigrant groups in Norway and Finland

Emil O. W. Kirkegaard*

Open Differential Psychology

**Abstract**

I present new predictive analyses for crime, income, educational attainment and employment among immigrant groups in Norway and crime in Finland. Furthermore I show that the Norwegian data contains a strong general socioeconomic factor (S) which is highly predictable from country-level variables (National IQ .59, Islam prevalence -.71, international general socioeconomic factor .72, GDP .55), and correlates highly (.78) with the analogous factor among immigrant groups in Denmark. Analyses of the prediction vectors show very high correlations (generally > ±.9) between predictors which means that the same variables are relatively well or weakly predicted no matter which predictor is used. Using the method of correlated vectors shows that it is the underlying S factor that drives the associations between predictors and socioeconomic traits, not the remaining variance (all correlations near unity).

**Keywords:** National IQ, intelligence, group differences, country of origin, Norway, Finland, Denmark, immigration, crime, spatial transferability hypothesis, income, employment, educational attainment, general socioeconomic factor, Islam, method of correlated vectors

## 1 Introduction

Recent studies show that criminality and other socioeconomic traits such as educational attainment among immigrant groups is predictable from their country of origin[1, 2, 3, 4]. This study attempts to replicate and generalize these findings.

The theoretical impetus for testing country-level predictors[1] is the spatial transferability hypothesis.[5] In brief, it proposes that:

1. A country's performance on a variety of metrics is due to some degree to the psychological makeup of its inhabitants;
2. People retain their psychological attributes to some degree when they migrate; and hence
3. The psychological attributes of groups determine to some degree their relative performance on a variety of socioeconomic variables, such as crime, educational attainments, income, and employment rate, in the countries that receive them.

For instance, when people from a poorer country move to a wealthier country, they will tend to be relatively poor in that country as well. This is because part of the reason the country is poor to begin with is that the people living there are low (or high) in the relevant behavioral traits. When they move to a new country, they will generally still be low in the relevant behavioral traits, and this will cause them to be relatively poor in that country as well. This is of course still allowing for other causes (e.g. culture or religion people tend to bring with them) as well as improvements on an absolute scale. Somalis living in Denmark are far richer than those who have stayed behind in Somalia, but they are nonetheless poorer than ethnic Danes, just as Somalia is poorer as a whole than Denmark. Note also that the hypothesis does not specify why these traits tend to be preserved. Both genetic and non-genetic models are possible.

The most obvious way of testing the spatial transferability hypothesis involves looking at immigrant performance on a variety of measures grouped by country of origin, and then checking how predictable this performance is from country-level variables such

* Department of Culture and Society, University of Aarhus, E-mail: emil@emilkirkegaard.dk

1 A note about terminology. "predictor" is used here to mean the same as "independent variable". No causality is implied, merely linguistic convenience.

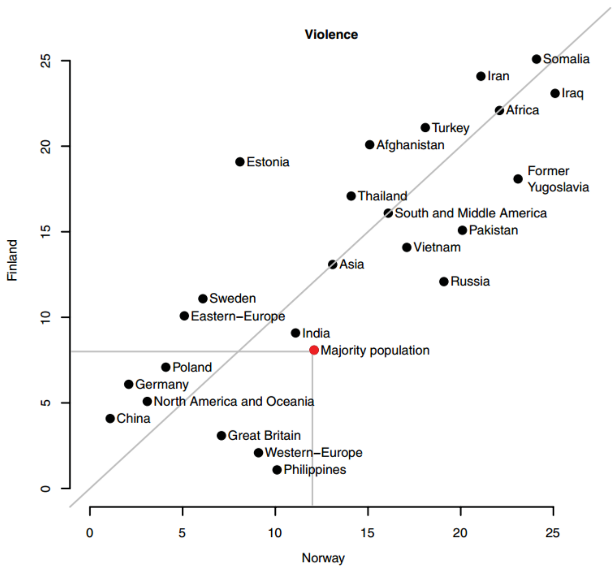


**Figure 1:** Violent crime in Norway and Finland by country of origin. From [6].

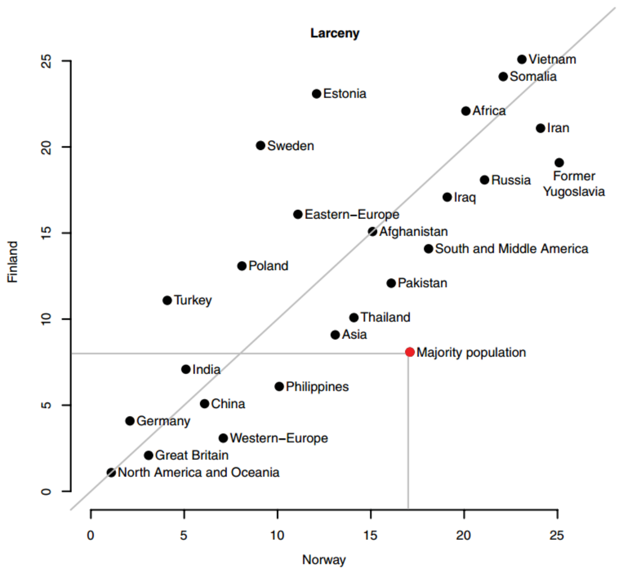


**Figure 2:** Property crime (larceny) in Norway and Finland by country of origin. From [6].

as national IQ and national prevalence of Islam. In this study I explore a number of datasets in this fashion.

## 2 Dataset 1: Norway and Finland

Skardhamar et al (2014)[6] presented new crime data for immigrant groups by country and macro-region of origin for Norway and Finland and compared the two countries. In doing so they adjusted for both gender ratio and age structure in the populations. Their main findings are shown in Figures 1 and 2.

These findings indicate that people from the same areas of origin are similarly disposed to criminal behaviour in Norway and Finland.

## 3 Dataset 2-4: Norway

I took a closer look at Statistisk Sentralbyrå's (SSB)[2] website for data that could be useful for testing the spatial transferability hypothesis. To be useful, the data must concern a variable of considerable social interest and contain information on immigrant group performance by country of origin with at least a small sample (my threshold was ≥ 10) of countries.

I searched for "landbakgrunn" (country background) on the website[3], limited the results to publicly available datasets ("statistikbanken") and looked through all 124 results. I identified three useful datasets:

1. Income after tax, measured as a percent of the national mean (income index).[4] No information about age is given. I included all available countries (N=23). Generally, SSB limits the available countries to the ones with samples large enough to give reliable results.

2. Registered unemployed persons by sex aged 15-74, as a percent of the working population.[5] As before I included all available countries (N=120) and both sexes separately. There is no fine-grained age information, so the data are not well-adjusted for age.

3. Tertiary educational attainment per capita for persons aged 16 and above in 2013.[6] This table was in absolute numbers, so I supplemented it with the population size by country of origin to calculate a pseudo per capita value.[7] The reason it is a pseudo per capita is that population sizes were not available by age groups, so I had to use the entire age group, even though the educational attainment data concerned only people aged 16 and above. This introduces error if the age structures are different between groups. The data are also not broken down by gender, so there is possibly gender ratio bias as well. To examine

2 The official statistics bureau of Norway. http://www.ssb.no/

3 I could not find out how the SSB determines the origin country. Which country does a person who has a Swedish and a Norwegian parent count as? Likewise there is no information about immigrant generation.

4 Tabell: 10489: Innvandreres inntekt etter skatt per forbruksenhet, etter landbakgrunn

5 Tabell: 07117: Registrerte arbeidsledige 15-74 år, etter landbakgrunn og kjønn. Absolutte tall og i prosent av arbeidsstyrken

6 Tabell: 09623: Innvandrere 16 år og over, etter utdanningsnivå og landbakgrunn. Absolutte tall

7 Tabell: 05184: Innvandrere, etter kjønn og landbakgrunn

effects of including small samples, I used two different versions of this variable. The first includes all groups with a population ≥ 200 (N=118). The second only includes groups with ≥ 1000 to reduce sampling error (N=67).

## 4 Predictive analyses

I did all analyses with R.[8] The primary question was whether crime was predictable from country-level variables as previously found. To test this, I used the following predictors in a correlation analysis:

- Prevalence of Islam in 2010 (as estimated by the Pew Research Center).[7]
- Lynn and Vanhanen's national IQs with changes based on the work of Jason Malloy. When a value is changed, it is noted in the datafile.[8, 9]
- Altinok's educational achievements.[10]
- The World Bank's GDP per capita (2013).[11]
- Kirkegaard's country-level general socioeconomic (S) factor scores.[12]

Table 1 shows the correlations of interest. Generally, all predictors do well when two conditions are satisfied:

1. The sample of countries is large enough to have significant inter-country variation.
2. The sample of countries is not so large as to introduce significant sampling error in estimates.

The reason this introduces error is that the more countries covered in a variable means that the value must be based on a smaller sample of persons from that country.

Findings of note include: Violent crime is easier to predict than property crime, just as in the Danish dataset.[1] The poor predictive ability of Altinok with the crime and income variables seems to be due to sampling error (N's 13-14). The educational attainment variable which includes only large samples ("Tert. Ed. Att. Big") has higher correlations than the one with smaller samples too. This is probably because the smaller ones introduce sampling error. Islam is a better predictor of female unemployment than of male, which may be related to the role of women in Islam.

[8] R is a free, powerful, easy to use programming language designed for data mining and statistics. See http://www.r-project.org/.

The intercorrelations between predictors is shown in Appendix C. IQ, Altinok, logGDP and international S have high intercorrelations, with a minimum of .72 and a mean of .84. Islam correlates weakly to moderately with the others (-.14 to -.43, mean -.29).

### 4.1 *Predictor vector intercorrelations*

Are some predictors just generally better at predicting than others, or is there specificity such that while predictor A may be better at predicting outcome X, predictor B is better at predicting outcome Y? An example of this would be that the prevalence of Islam predicts crime better than (national) IQ at predicting crime, while IQ is better at predicting educational attainment. To investigate this, I correlated the prediction vectors (rows in Table 1) for each predictor with the vectors of each other predictor. Correlations at ±1 indicate that predictor performance is general, while correlations near ±0 indicate specificity. Table 2 shows the results.

Surprisingly, even though there are problems with small sample sizes of the predictive correlations and the length of the vectors (N=9), the results strongly suggest that what is well-predicted by one predictor is also well-predicted by other predictors, no matter which two were compared. The mean abs. r=.92.

## 5 A general socioeconomic factor among immigrant groups in Norway

Similarly to my previous study of immigrant groups in Denmark[3], I wanted to investigate the possibility of a general socioeconomic factor at the group level (S factor).[12] To do this, I used all the variables concerning Norway except for the educational attainment with smaller groups to avoid duplicating variables.

### 5.1 *Handling missing values*

Factor analytic methods require that there are no missing values. The easiest and most common way to deal with this is to limit the data to the subset with complete cases. This, however, produces biased results if the data are not missing completely at random, which they rarely are. Furthermore, it heavily reduces sample sizes. Lastly, it wastes non-redundant information and potentially resources spent gathering it. If a case has values for 5 out of 6 chosen variables, removing the case wastes 5 pieces of useful information.[13, 14, 15, 16] Table 3 shows the distribution of missing values.

For the above reasons, I used four methods for handling missing cases:

**Table 1:** Correlation matrix for country-level predictors and socioeconomic variables.

| | ViolentCrimeNO | ViolentCrimeFI | LarcenyNO | LarcenyFI | Tert. Ed. Att. | Tert. Ed. Att. Big | Unemployment men | Unemployment women | Income |
|---|---|---|---|---|---|---|---|---|---|
| IQ | -0.63 | -0.64 | -0.38 | -0.29 | 0.35 | 0.47 | -0.5 | -0.43 | 0.61 |
| n | 20 | 19 | 20 | 19 | 117 | 67 | 119 | 119 | 23 |
| Altinok | -0.24 | -0.14 | -0.14 | 0.19 | 0.31 | 0.41 | -0.49 | -0.46 | 0.29 |
| n | 14 | 13 | 14 | 13 | 92 | 54 | 95 | 95 | 15 |
| Islam | 0.82 | 0.83 | 0.52 | 0.25 | -0.26 | -0.37 | 0.45 | 0.68 | -0.52 |
| n | 19 | 18 | 19 | 18 | 118 | 67 | 120 | 120 | 23 |
| GDP (log) | -0.33 | -0.07 | -0.26 | -0.06 | 0.33 | 0.46 | -0.44 | -0.36 | 0.64 |
| n | 17 | 16 | 17 | 16 | 103 | 58 | 105 | 105 | 20 |
| S.scores | -0.59 | -0.38 | -0.4 | 0.04 | 0.35 | 0.41 | -0.5 | -0.53 | 0.64 |
| n | 16 | 15 | 16 | 15 | 103 | 58 | 105 | 105 | 18 |

**Table 2:** Correlation matrix of predictor vectors. N=9

| Var | Altinok.cors | Islam.cors | GDP.cors | S.cors |
|---|---|---|---|---|
| IQ.cors | 0.83 | -0.99 | 0.93 | 0.96 |
| Altinok.cors | | -0.86 | 0.9 | 0.92 |
| Islam.cors | | | -0.92 | -0.97 |
| GDP.cors | | | | 0.97 |

**Table 3:** The distribution of missing values in the Norwegian dataset.

| Number of missing values | Number of cases |
|---|---|
| 0 | 15 |
| 1 | 3 |
| 2 | 8 |
| 3 | 41 |
| 4 | 61 |
| 6 | 141 |

1. Complete cases only (N=15)
2. Imputing[9] data in cases with 1 missing value (N=18)
3. Imputing data in cases with 2 or fewer missing values (N=26)
4. Imputing data in cases with 3 or fewer missing values (N=67)

Table 4 shows descriptive statistics for each dataset. The imputed datasets are similar to both the full datasets and the complete cases although there were changes in both the skew and kurtosis.

KMO tests show that all datasets are suitable to factor analysis, KMO's .68-.75. Note that the method of imputation used is probabilistic, i.e. does not result in the same imputation every time. Therefore, any researcher who replicates the analyses will find that the numbers deviate somewhat from the shown results. The KMO values were always around these values in my tests.

## 5.2 *Number of factors to extract*

To find out how many factors to extract, I ran *nScree()* from the **nFactors** package.[10] For each dataset, all four tests within that function suggested to extract only one factor.

## 5.3 *Strength of the general factor*

Previous studies show that principal component analysis tends to overestimate factor loadings when used on a small number of variables, but that other factor methods yield very similar results.[12, 18, 19, 20] I used minimum residuals (the default) to extract the first factor from each dataset.[11]

Revelle and Wilt[21] showed that one cannot solely rely on the size of the first factor in a normal analysis as a measure of the strength of the general factor. They advocate five other methods, of which I have used four here:

1. Hierarchical omega and its asymptotic value
2. The amount of variance accounted for by the first factor in a Schmid-Leiman transformation

[9] I used the **VIM** package 4.00. The *irmi()* function imputes values.[17] I used the default settings. http://cran.r-project.org/web/packages/VIM/index.html.

[10] Version 2.3.3 http://cran.r-project.org/web/packages/nFactors/index.html.

[11] I used the *fa()* function from **Psych** package. http://cran.r-project.org/web/packages/psych/index.htmlVersion: 1.4.8.11

**Table 4:** Descriptive stats by dataset

| Var name | Dataset | n | mean | sd | min | max | skew | kurtosis |
|---|---|---|---|---|---|---|---|---|
| Violent crime | Full | 26 | 1.31 | 0.87 | 0.2 | 3.2 | 0.55 | −0.83 |
| | Complete cases | 15 | 1.41 | 0.99 | 0.2 | 3.2 | 0.39 | −1.25 |
| | Impute 1 | 18 | 1.33 | 0.93 | 0.2 | 3.2 | 0.57 | −0.94 |
| | Impute 2 | 26 | 1.28 | 0.81 | 0.2 | 3.2 | 0.7 | −0.27 |
| | Impute 3 | 67 | 1.23 | 0.63 | 0.2 | 3.2 | 1.05 | 1.41 |
| Larceny | Full | 26 | 0.77 | 0.56 | 0.1 | 2 | 0.56 | −1.09 |
| | Complete cases | 15 | 0.78 | 0.55 | 0.2 | 1.6 | 0.38 | −1.74 |
| | Impute 1 | 18 | 0.72 | 0.53 | 0.1 | 1.6 | 0.55 | −1.42 |
| | Impute 2 | 26 | 0.69 | 0.56 | −0.29 | 1.96 | 0.62 | −0.58 |
| | Impute 3 | 67 | 0.69 | 0.34 | 0.1 | 1.6 | 0.71 | 0.19 |
| Tert. ed. att. | Full | 67 | 0.12 | 0.08 | 0.01 | 0.31 | 0.42 | −0.91 |
| | Complete cases | 15 | 0.1 | 0.07 | 0.01 | 0.23 | 0.39 | −1.32 |
| | Impute 1 | 18 | 0.1 | 0.07 | 0.01 | 0.23 | 0.5 | −1.12 |
| | Impute 2 | 26 | 0.09 | 0.07 | 0.01 | 0.24 | 0.75 | −0.63 |
| | Impute 3 | 67 | 0.12 | 0.08 | 0.01 | 0.31 | 0.42 | −0.91 |
| Unemployment, men | Full | 120 | 7.05 | 4.18 | 1.38 | 22.08 | 1.26 | 1.69 |
| | Complete cases | 15 | 7.4 | 5.36 | 2.68 | 22.08 | 1.38 | 1.18 |
| | Impute 1 | 18 | 7.31 | 4.89 | 2.68 | 22.08 | 1.57 | 2.16 |
| | Impute 2 | 26 | 6.88 | 4.3 | 2.66 | 22.08 | 1.8 | 3.72 |
| | Impute 3 | 67 | 6.81 | 4.11 | 1.66 | 22.08 | 1.39 | 2.18 |
| Unemployment, women | Full | 120 | 7.5 | 4.97 | 1.32 | 31.82 | 1.92 | 5.11 |
| | Complete cases | 15 | 8.9 | 6.2 | 1.98 | 22.42 | 0.83 | −0.58 |
| | Impute 1 | 18 | 8.17 | 5.93 | 1.9 | 22.42 | 1.03 | −0.04 |
| | Impute 2 | 26 | 7.4 | 5.25 | 1.56 | 22.42 | 1.3 | 1.14 |
| | Impute 3 | 67 | 7.41 | 5.39 | 1.32 | 31.82 | 1.97 | 5.23 |
| Income | Full | 23 | 79.86 | 14.58 | 53.25 | 108.25 | −0.01 | −0.92 |
| | Complete cases | 15 | 78.78 | 14.78 | 53.25 | 108.25 | 0.19 | −0.78 |
| | Impute 1 | 18 | 82.4 | 16.3 | 53.25 | 112.29 | 0.11 | −0.95 |
| | Impute 2 | 26 | 79.09 | 13.9 | 53.25 | 108.25 | 0.13 | −0.75 |
| | Impute 3 | 67 | 81.71 | 13.86 | 31.96 | 108.25 | −0.76 | 0.95 |

3. The explained common variance

4. the squared multiple correlation of regressing the first factor on the original variables.[12]

Table 5 shows the comparison statistics including data from the reanalysis of the Danish data presented in the next section.

The data makes it clear that the S factors at the group-level among immigrants in Norway and Denmark are very strong, even compared to the international S factor and the general factor of cognitive ability (g) in 5 classic datasets. The imputation of data has little effect on the measures of general factor strength.

[12] I used the *omega()* function from **Psych** package to extract the information.

## 6 Reanalysis of immigrant performance in Denmark

To better investigate the question of the strength of the S factor within another country, I repeated the analyses discussed above on the dataset from Kirkegaard and Fuerst (2014)[3]. Since I used the same methods on this dataset as I did on the Norwegian ones discussed above, I will keep the description short.

I analyzed the data with the *fa()* and *omega()* functions just as before. Results are shown in Table 5 above. KMO is .73 in the complete cases dataset and .83 in the imputed dataset.

**Table 5:** Measures of general factor strength. The cognitive and personality data is from Revelle and Wilt (2013)[21], the international S factor data is from Kirkegaard (2014)[12], and the Danish comparison data is from a reanalysis of the datasets from Kirkegaard and Fuerst (2014)[3] presented in the next section.

| Dataset | Var% MR | Var% MR SL | Omega h. | Omega h. a. | ECV | R2 |
|---|---|---|---|---|---|---|
| NO Complete cases | 0.68 | 0.65 | 0.87 | 0.91 | 0.78 | 0.98 |
| NO Impute 1 | 0.66 | 0.62 | 0.86 | 0.9 | 0.74 | 0.96 |
| NO Impute 2 | 0.64 | 0.60 | 0.85 | 0.89 | 0.75 | 0.95 |
| NO Impute 3 | 0.63 | 0.59 | 0.82 | 0.87 | 0.73 | 0.99 |
| DK complete cases | 0.57 | 0.51 | 0.83 | 0.85 | 0.68 | 0.99 |
| DK impute 4 | 0.55 | 0.51 | 0.86 | 0.88 | 0.73 | 0.99 |
| International S factor | 0.43 | 0.35 | 0.76 | 0.77 | 0.51 | 0.81 |
| Cognitive data | | 0.33 | 0.74 | 0.79 | 0.57 | 0.78 |
| Personality data | | 0.16 | 0.37 | 0.48 | 0.34 | 0.41 |

### 6.1 *Handling missing values*

There are a few missing values in the dataset. I used two methods to deal with this:

1. Complete cases (N=31)
2. Imputation via the **VIM** package for cases with 4 or fewer missing values (N=70)

Table 6 shows the distribution of missing values.

**Table 6:** Distribution of missing values in the Danish dataset.

| Number of missing values | Number of cases |
|---|---|
| 0 | 31 |
| 1 | 9 |
| 2 | 23 |
| 3 | 6 |
| 4 | 1 |
| 23 | 1 |

### 6.2 *Predictor vector intercorrelations*

I repeated the analysis from Section 4.1 on the Danish data. The Danish data allows for a better test of the general vs. specificity models. This is because the Danish data has more variables (25 instead of 9; Appendix B contains a list of the variables), and they include more countries (N's close to 70), and they are age controlled. Table 7 shows the correlations between prediction vectors.

The correlations are even closer to unity than they are in the Norwegian data. The mean abs. r=.97. This is probably because the error sources are smaller in these data.

**Table 7:** Correlation matrix of predictor vectors in the Danish data. N=25

| Predictors | Altinok | Islam | logGDP | S.score |
|---|---|---|---|---|
| **IQ** | 0.99 | -0.96 | 0.98 | 0.99 |
| **Altinok** | | -0.94 | 0.98 | 0.98 |
| **Islam** | | | -0.94 | -0.96 |
| **logGDP** | | | | 0.99 |

## 7 Predictive analyses of S scores

I wanted to know how well S factor scores in Norway are predictable from the country-level predictor variables. Table 8 shows the correlation matrix with the Danish scores as a comparison.

The results indicate that S factor scores are about equally predictable by predictor values in the full Danish and Norwegian datasets. The size of the correlation decreases with the amount of imputation and increasing sample size. This may be because the imputation introduces error or that the correlations are artificially high due to sampling error. The only discrepancy is the predictive power of the National S (.54 vs. .72). I don't have any good guess for why this is.

## 8 Method of correlated vectors

Arthur Jensen invented the method of correlated vectors (MCV) in 1983 to find out if g is responsible for mean differences in measures of intelligence.[22, 23, 24] Today, the method is mostly used with g (e.g. [25, 26, 27, 28]), and in context to mean differences, but it can be used for any latent variable and

**Table 8:** Correlation matrix of predictor variables and S factor scores in Denmark (with imputed values) and the four Norwegian datasets with varying amounts of imputation.

| Variable | DA S imp. | NO S complete | NO S imp. 1 | NO S imp. 2 | NO S imp. 3 |
|---|---|---|---|---|---|
| National IQ | 0.54 | 0.72 | 0.73 | 0.66 | 0.59 |
| Altinok | 0.55 | 0.26 | 0.31 | 0.2 | 0.6 |
| Islam | −0.71 | −0.79 | −0.79 | −0.72 | −0.71 |
| log(GDP) | 0.51 | 0.35 | 0.4 | 0.44 | 0.55 |
| National S | 0.54 | 0.7 | 0.71 | 0.63 | 0.72 |
| DA S imp. | | 0.91 | 0.91 | 0.79 | 0.78 |
| NOS complete | | | 1 | 0.99 | 0.99 |
| NO S imp. 1 | | | | 0.99 | 0.99 |
| S imp. 2 | | | | | 0.99 |

in other contexts such as prediction of grade point averages.[29] I have previously used it on international rankings data and found that the international S factor was driving the correlations with predictors such as national IQs.[12]

To apply the method, one correlates the indicator variables' (IVs) loading on the latent variable of interest with their correlations with the criteria variable. If the general factor is 'driving' the association and is positively correlated with the criteria variable, then the correlation between factor loadings on it and the effect sizes of the predictor-criterion associations will be positive. However, if the association is driven by the variance not attributable to the general factor, the correlation will generally be negative. And it will generally be somewhere in between if the association is driven by a mix of general and non-general factors.

Since the method relies on the IVs of the latent variable, it is susceptible to IV sampling error. There are three main sources of error:

1. When the number of IVs is small, the correlation will be unstable from sample to sample.
2. If the IVs are unrepresentative of the total population of IVs, then the correlation can be biased either way.
3. If the variance in IVs' loadings on the latent variable is restricted, the correlations will be smaller.[30]

The $N_{IV}$ is quite small for the Norwegian data (6), but reasonable for the Danish data (25). The standard deviation of loadings in the Norwegian and Danish datasets are .83 and .75, respectively, so range restriction does not appear to be a problem. The IVs are reasonably representative of things considered socioeconomically important, especially for the Danish dataset. Table 9 shows the MCV correlations.

**Table 9:** Results from method of correlated vectors applied to the S factor in Norway and Denmark. Note that Islam prevalence has been reversed because it is negatively related to the S factor.

| Predictor | Norway (N=6) | Denmark (N=25) |
|---|---|---|
| IQ | 0.99 | 0.97 |
| Altinok | 0.95 | 0.95 |
| Islam | 0.99 | 0.99 |
| logGDP | 1 | 0.95 |
| Int. S factor | 0.99 | 0.96 |

In every case, the result is close to unity. Strong correlations can result from nonlinearity in the data, so I examined the scatterplots. However, they were all very linear.

## 9 Discussion

The simple predictive analyses give results similar to those found earlier. They serve as a successful replication and generalization to two new countries.

The analyses of general factor strength show that the local S factors are generally very strong, surpassing even the g factor and the international S factor. This is due in part to the grouped nature of the data as group correlations tend to go towards ±1 when there aren't sampling errors or a non-linear relationship.[31]

Surprisingly, the analysis of the predictor vectors show very high near unity intercorrelations. The results were even stronger in the reanalyzed Danish data, which makes it probable that the somewhat lower correlations in the Norwegian data are due to statistical artifacts. I interpret this as showing that predictors are very general in their predictive ability. In accordance with Occam's Razor, causal theories of these correlations should be similarly general, not specific.

Applying the method of correlated vectors showed that the relationships between the socioeconomic variables and the predictor variables were driven by the latent trait (S factor), not the remaining variance (correlations near unity).

Generally the results strongly confirm the spatial transferability hypothesis.[5]

Limitations include the small sample sizes and the lack of adjustment for age and sex in some of the variables. This probably introduces some bias in an unknown direction. Note however that the Danish data is age-controlled, and yet the results are very similar to the Norwegian ones, showing that bias due to age is unlikely to be a large source of error.

## Supplementary material

All datasets and source code are available in the submission thread at http://openpsych.net/forum/showthread.php?tid=136 and on OSF https://osf.io/emfag/. Most of the data used in the study can be found in version 1.5 extra of the Worldwide Megadataset.

The appendix contains a list of S scores by group in Norway and Denmark.

## References

[1] Kirkegaard EOW. Criminality and fertility among Danish immigrant populations. Open Differential Psychology. 2014;.

[2] Kirkegaard EOW. Criminality among Norwegian immigrant populations. Open Differential Psychology. 2014;.

[3] Kirkegaard EOW, Fuerst J. Educational attainment, income and use of social benefits among 71 immigrant groups in Denmark. Open Differential Psychology. 2014;.

[4] Proto E, Oswald AJ. National Happiness and Genetic Distance: A Cautious Exploration. IZA Discussion Papers. 2014;July 2014.

[5] Fuerst J, Kirkegaard EOW. Do National IQs Predict U.S. Immigrant Cognitive Ability and Outcomes? An Analysis of the National Longitudinal Survey of Freshman. Open Differential Psychology. 2014;.

[6] Skardhamar T, Aaltonen M, Lehti M. Immigrant crime in Norway and Finland. Journal of Scandinavian Studies in Criminology and Crime Prevention. 2014;(ahead-of-print):1–21.

[7] Research P. The Future of the Global Muslim Population; 2011. Available from: http://features.pewforum.org/muslim-population/.

[8] Lynn R, Vanhanen T. Intelligence: A unifying construct for the social sciences. Ulster Institute for Social Research; 2012.

[9] Malloy J. Human Varieties Global IQ;. Available from: http://humanvarieties.org/category/hv-global-iq/.

[10] Altinok N, Diebolt C, Demeulemeester JL. A new international database on education quality: 1965–2010. Applied Economics. 2014;46(11):1212–1247.

[11] Bank TW. GDP per capita; 2014. Available from: http://data.worldbank.org/indicator/NY.GDP.MKTP.CD.

[12] Kirkegaard EOW. The international general socioeconomic factor: Factor analyzing international rankings. Open Differential Psychology. 2014;.

[13] Su YS, Yajima M, Gelman AE, Hill J. Multiple imputation with diagnostics (mi) in R: Opening windows into the black box. Journal of Statistical Software. 2011;45(2):1–31.

[14] van der Heijden GJ, T Donders AR, Stijnen T, Moons KG. Imputation of missing values is superior to complete case analysis and the missing-indicator method in multivariable diagnostic research: a clinical example. Journal of clinical epidemiology. 2006;59(10):1102–1109.

[15] Donders ART, van der Heijden GJ, Stijnen T, Moons KG. Review: a gentle introduction to imputation of missing values. Journal of clinical epidemiology. 2006;59(10):1087–1091.

[16] Janssen KJ, Donders ART, Harrell Jr FE, Vergouwe Y, Chen Q, Grobbee DE, et al. Missing covariate data in medical research: to impute is better than to ignore. Journal of clinical epidemiology. 2010;63(7):721–727.

[17] Templ M, Filzmoser P. Visualization of missing values using the R-package VIM. Reserach report cs-2008-1, Department of Statistics and Probability Therory, Vienna University of Technology. 2008;.

[18] Floyd RG, Shands EI, Rafael FA, Bergeron R, McGrew KS. The dependability of general-factor loadings: The effects of factor-extraction methods, test battery composition, test battery size, and their interactions. Intelligence. 2009;37(5):453–465.

[19] Major J. The dependability of the general factor of intelligence: Why g is not a first principal component. The University of Edinburgh; 2010.

[20] Major JT, Johnson W, Bouchard Jr TJ. The dependability of the general factor of intelligence: Why small, single-factor models do not adequately represent g. Intelligence. 2011;39(5):418–433.

[21] Revelle W, Wilt J. The general factor of personality: A general critique. Journal of research in personality. 2013;47(5):493–504.

[22] Jensen AR. Effects of inbreeding on mental-ability factors. Personality and Individual Differences. 1983;4(1):71–87.

[23] Jensen AR, Weng LJ. What is a good g? Intelligence. 1994;18(3):231–258.

[24] Jensen AR. The g factor: The science of mental ability. Praeger Westport, CT; 1998.

[25] Armstrong EL, Woodley MA, Lynn R. Cognitive abilities amongst the Sámi population. Intelligence. 2014;46:35–39.

[26] Peach H, Lyerly JE, Reeve CL. Replication of the Jensen effect on dysgenic fertility: An analysis using a large sample of American youth. Personality and Individual Differences. 2014;71:56–59.

[27] te Nijenhuis J, Jongeneel-Grimen B, Kirkegaard EO. Are Headstart gains on the< i> g</i> factor? A meta-analysis. Intelligence. 2014;46:209–215.

[28] Woodley MA, te Nijenhuis J, Must O, Must A. Controlling for increased guessing enhances the independence of the Flynn effect from< i> g</i>: The return of the Brand effect. Intelligence. 2014;43:27–34.

[29] Kirkegaard EOW. The personal Jensen coefficient does not predict grades beyond its association with g. Open Differential Psychology. 2014;.

[30] Hunter JE, Schmidt FL. Methods of meta-analysis: Correcting error and bias in research findings. 3rd ed. Sage; 2014.

[31] Lubinski D, Humphreys LG. Seeing the forest from the trees: When predicting the behavior or status of groups, correlate means. Psychology, Public Policy, and Law. 1996;2(2):363.

# A S factor scores by country

**Table 10:** S factor scores by country.

| Name | ID | S Factor in NO | S factor in DK |
|---|---|---|---|
| Afghanistan | AFG | −1.09 | −1.38 |
| Argentina | ARG | | 0.75 |
| Australia | AUS | 1.03 | 1.13 |
| Austria | AUT | 1.02 | 0.95 |
| Burundi | BDI | −0.54 | |
| Belgium | BEL | 1.16 | 1.09 |
| Bulgaria | BGR | 0.17 | 0.81 |
| Bosnia and Herzegovina | BIH | 0.49 | −0.91 |
| Brazil | BRA | −0.34 | 0.46 |
| Canada | CAN | 1.03 | 1.14 |
| Switzerland | CHE | 1.13 | 1.12 |
| Chile | CHL | 0.25 | 0.28 |
| China | CHN | 0.61 | 0.63 |
| Congo Rep. | COG | −1.07 | |
| Colombia | COL | 0.26 | |
| Czech Republic | CZE | 0.43 | 0.25 |
| Germany | DEU | 1.04 | 0.85 |
| Denmark | DNK | 1 | |
| Algeria | DZA | −1.52 | −0.78 |
| Egypt Arab Rep. | EGY | | −0.24 |
| Eritrea | ERI | −0.43 | |
| Spain | ESP | 0.52 | 0.79 |
| Estonia | EST | 0.19 | 0.72 |
| Ethiopia | ETH | −0.16 | −0.59 |
| Finland | FIN | 0.78 | 0.89 |
| France | FRA | 0.97 | 1.1 |
| United Kingdom | GBR | 1.14 | 0.85 |
| Ghana | GHA | 0.03 | 0.16 |
| Gambia The | GMB | −0.84 | |
| Greece | GRC | 0.61 | 0.61 |
| Croatia | HRV | 0.54 | −0.12 |
| Hungary | HUN | 0.45 | 0.84 |
| Indonesia | IDN | 0.33 | 0.13 |
| India | IND | 0.63 | 0.53 |
| Ireland | IRL | | 0.88 |
| Iran Islamic Rep. | IRN | −0.35 | −0.69 |
| Iraq | IRQ | −2.26 | −1.65 |
| Iceland | ISL | 0.76 | 0.55 |
| Israel | ISR | | −0.06 |
| Italy | ITA | 0.86 | 0.77 |
| Jordan | JOR | | −1.19 |
| Japan | JPN | | 1.02 |
| Kenya | KEN | −0.24 | 0.09 |
| Kosovo | KSV | −0.43 | |
| Kuwait | KWT | | −2.62 |
| Lebanon | LBN | −1.03 | −2.03 |

Continued on next page

**Table 10 – continued from previous page**

| Name | ID | S Factor in NO | S factor in DK |
|---|---|---|---|
| Sri Lanka | LKA | −0.14 | −0.75 |
| Lithuania | LTU | −0.08 | 0.9 |
| Latvia | LVA | 0.06 | 0.68 |
| Morocco | MAR | −0.63 | −1.03 |
| Macedonia FYR | MKD | −0.19 | −0.44 |
| Myanmar | MMR | −0.27 | −1.81 |
| Nigeria | NGA | −0.53 | 0.34 |
| Netherlands | NLD | 1.11 | 1.12 |
| Norway | NOR | | 0.84 |
| Nepal | NPL | 0.75 | |
| Pakistan | PAK | −0.87 | −0.68 |
| Peru | PER | 0.1 | |
| Philippines | PHL | 0.58 | 0.36 |
| Poland | POL | −0.02 | 0.46 |
| Portugal | PRT | 0.54 | 0.63 |
| West Bank and Gaza | PSE | −3.8 | |
| Romania | ROU | 0.31 | 0.7 |
| Russian Federation | RUS | −0.44 | 0.45 |
| Sudan | SDN | −1.52 | |
| Somalia | SOM | −3.06 | −2.05 |
| Serbia | SRB | 0.46 | −1.93 |
| USSR | SUN | | 0.17 |
| Slovak Republic | SVK | 0.42 | |
| Sweden | SWE | 1.03 | 0.77 |
| Syrian Arab Republic | SYR | −1.62 | −2 |
| Thailand | THA | −0.03 | −0.23 |
| Tunisia | TUN | | −0.82 |
| Turkey | TUR | −0.52 | −1.42 |
| Tanzania | TZA | | −0.25 |
| Uganda | UGA | | −0.34 |
| Ukraine | UKR | 0.34 | 0.69 |
| United States | USA | 0.97 | 1.26 |
| Vietnam | VNM | −0.11 | −0.58 |
| Former Yugoslavia2 | YU2 | | −1.61 |
| Former Yugoslavia | YUG | | −1.25 |
| South Africa | ZAF | | 0.73 |

## B List of variables in the Danish dataset

**Table 11:** List of variables in the Danish dataset.

| Variable name |
|---|
| All.crime.age.15.19 |
| All.crime.age.20.29 |
| Income.15.19 |
| Income.20.29 |
| Income.30.39 |
| Income.40.49 |
| Income.50.59 |
| Income.60 |
| Basic.school.15.19 |
| Basic.school.20.29 |
| Basic.school.30.39 |
| Basic.school.40.49 |
| Basic.school.50.59 |
| Basic.school.60plus |
| Long.tert.edu.20.29 |
| Long.tert.edu.30.39 |
| Long.tert.edu.40.49 |
| Long.tert.edu.50.59 |
| Long.tert.edu.60plus |
| Social.benefits.16.19 |
| Social.benefits.20.29 |
| Social.benefits.30.39 |
| Social.benefits.40.49 |
| Social.benefits.50.59 |
| Social.benefits.60plus |

## C Intercorrelations between predictors

**Table 12:** The low correlations between Islam and the others is not due to sampling fluctuation. N's from 116 to 198. The full correlation matrix can be found in the supplementary material "correla- tions Norway2014.xlsx".

| Vars | Altinok | Islam | logGDP | International S |
|---|---|---|---|---|
| IQ | 0.91 | −0.27 | 0.72 | 0.86 |
| Altinok | | −0.43 | 0.76 | 0.87 |
| Islam | | | −0.14 | −0.33 |
| logGDP | | | | 0.9 |