

ORIGINAL ARTICLE

Open Access



Employment chances of immigrants and their children in Germany: does sense of personal control matter?

Anna Thum-Thysen

Correspondence:

anna.thum@ec.europa.eu

Most of the research underlying this article was undertaken at the European University Institute under a grant from the DAAD and does not necessarily reflect the views of the European Commission.
Rue de la Loi 170, 1040 Brussels, Belgium

Abstract

Labour market integration is a social process suggesting that personality traits are relevant. This paper explores whether immigrants with a higher belief in their ability to control outcomes tend to be more likely to be employed. This trait is known in psychology as the locus of control (LOC). I employ a model framework that allows LOC to depend on a set of observable determining variables. Results indicate that sense of personal control is correlated with employment probability and that there is a risk that immigrants and their children tend to have a more externalized LOC. Mother's education is a key factor correlated with the sense of personal control.

JEL Classification: J15, J24, M59, Z13

Keywords: Personality, Labour market, Unemployment

1 Introduction

Does personality matter for the integration of immigrants on the labour market? Integration is also a social process, which indeed suggests that personality traits are relevant. This paper explores whether and to what extent immigrants and their children with stronger beliefs in their ability to control success tend to be more likely to be employed. In personality psychology, the strength of this belief is referred to as the locus of control (LOC).

I am particularly interested in this personal trait firstly because immigrants and their children tend to face obstacles when searching for a job such as language difficulties, cultural differences and potentially discrimination. The sense of personal control can be a very useful personal trait to overcome these obstacles, in particular in Germany, where discrimination in particular in terms of formal degrees was found to be comparatively pronounced (see Section 2). Secondly, it was proven to predict several crucial economic outcomes including the employment probability (see for instance Heckman et al. 2006). As Groves (2005) notes, LOC is potentially the most widely used personality variable in economic research.

LOC was developed by Rotter (1966) and is a measure of the degree to which an individual believes he or she has control over the happenings in life. It is represented by a scale spreading from "external" to "internal". A highly *external* LOC indicates that the

individual believes that life is controlled by forces outside of his or her own influence and he or she does not have a high feeling of controlling life. A highly *internal* LOC however indicates that the individual believes strongly in the ability to control life.

LOC is linked to the concept of motivation: A person with a more internal LOC believes that a given level of effort will lead to a higher reward and thus *ceteris paribus* will work harder. However, LOC is not equal to motivation as motivation will also be affected by the cost of effort, which is not necessarily correlated with LOC. The LOC of an individual develops over time, with foundations being laid in childhood and adolescence. Education, family background, religion and personal immigration history can play a role to determine an immigrant's LOC.

The econometric approach in this paper for testing the link between LOC and employment probability follows a methodology developed in Carneiro et al. (2003) and was developed in Heckman et al. (2006). I extend their model framework by allowing the measure of LOC to depend on a set of observable determining variables. This extension is based on work by Fahrmeir and Raach (2006). The model is estimated using a Bayesian Markov Chain Monte Carlo (MCMC) algorithm. To test the model for Germany, I use the German Socioeconomic Panel (SOEP) database.

I find that a more *internal* LOC has a positive effect on the probability of being employed. Further, I find that being an immigrant is significantly and negatively correlated with having a more *internal* LOC. This means that there is a risk that immigrants and their children face a double disadvantage on the labour market: they are disadvantaged because of their status as an immigrant and they tend to have a more external position on the LOC scale. Mother's education is a key factor determining the sense of personal control.

The paper is structured as follows. Section 2 discusses labour market integration of immigrants in Germany and the link between personality traits, in particular the LOC, and labour market outcomes and integration. In Section 3, I introduce the data used to test the model. Section 4 presents the econometric model. In Section 5, I analyse the results and Section 6 concludes.

2 Previous findings

Belief in the ability to control outcomes in life has been identified as a key personality trait determining labour market success, and it is an established indicator of personality traits in economic research (Groves 2005). Both in terms of wages (see for instance Duncan and Dunifon 1998; Groves 2005; Heineck and Anger 2010) and in terms of employment probability and job search (see for instance Andrisani 1977; Caliendo et al. 2015; Uhlendorff 2004), the literature has established the importance of the LOC. Indeed, the importance of personality traits for socioeconomic success has been established in the economic literature recent years (Almlund et al. 2011; Borghans et al. 2008; Bowles 2001a, b; Brunello and Schlotter 2011; Heckman 1995).

Based on the Panel Study of Income Dynamics (PSID), Duncan and Dunifon (1998) find that a sense of personal control affects earnings positively and significantly. Groves (2005) finds that personality traits such as the LOC significantly determine wages of white women. Cobb-Clark (2015) examines the role of LOC on the labour market in the light of three main channels: decisions to invest in human capital, decisions to hire new employees and finally the conception of incentive contracts. She highlights the important role

that the recognition of personality traits on the labour market has had on a re-formulation of theoretical and empirical economic models.

To embed these findings in economic theory, previous literature in economics has attempted to link psychometric questions to economic parameters (Borghans et al. 2008). They come to the conclusion that personality traits introduced both on an empirical and a theoretical level into economics can be fruitful for economic theory. According to the authors classical economic theory should incorporate the fact that economic preferences might be consequences of constraints imposed by cognitive skills and personality traits. They name the example that a high rate of time preference might be due to the fact that the agent cannot imagine the future. This ability to imagine would be interpreted as a personal trait.¹

In this line of research, an attempt to economically model the content of the LOC was made by Bowles et al. (2001a). They develop a theoretical model to explain the advantage for an employer to employ a motivated worker by setting up a signalling model. They interpret LOC as an employee's preference, which reduces the employer's cost to induce the employee's effort. This simply means that a worker, who believes more in his own success is more motivated to induce effort. This could also be interpreted that a worker with an intrinsic motivation (such as the belief in his success) will not need an external source to make him provide an effort. This is a desirable trait of a worker for an employer and the authors call such a trait an "incentive enhancing preference".

In personality psychology, the belief in one's own ability to achieve success together with neuroticism, self-efficacy and self-esteem is a trait that belongs to the four dimensions of the fundamental appraisal of oneself—or core self-evaluations, a key personal trait predicting job performance and job satisfaction (Dormann et al. 2006). It has been argued that the belief in one's ability to control one's success measures the same single factor as neuroticism, self-efficacy and self-esteem (Judge et al. 2002). Some authors argue that this trait can have a higher predictive power than the FFM personality traits (Judge et al. 2008). It can however also be argued that it is one of the FFM personality traits as it measures the same concept as the neuroticism component (Judge and Bono 2001) or it can be allocated to conscientiousness, which includes the facet "striving to achievement" (Borghans et al. 2008). Note that the conceptualization of neuroticism in the FFM personality traits does not include measures of self esteem and is therefore more narrow than the core self-evaluations concept.

Several authors study the effect of personality traits on the German labour market—also with respect to integration of immigrants. Flossmann et al. (2008) analyse the effect of non-cognitive skills—in particular the LOC and self-esteem—on wages in Germany using the German Socioeconomic Panel (SOEP) and an MCMC algorithm. They find that this type of skills indeed matters for labour market success. Also based on the German SOEP, Heineck and Anger (2010) study the joint relationship between cognitive skills, personality traits, and earnings in Germany. They find that, while other personality traits yield heterogenous results, an external LOC can have negative effects on wages for both men and women. Sociability is also studied as a key factor for integration by de Palo et al. (2006).

Several papers focus on personality and labour market integration in Germany. Fertig (2004) analyses the differences in leisure-time activities and attitudes of foreign immigrants, ethnic Germans and different generations and finds that both generations

have differences in attitudes compared to Germans. Second generation immigrants seem to be the most fatalist and pessimist. Bonev et al. (2010) study whether the difference in unemployment rates between German natives and migrants in Germany can be explained by differences in personality traits. Using the German SOEP database and an instrumental variable methodology to address endogeneity issues, they find that there seem to be significant differences between Turks and Germans in terms of the Big Five personality traits.

In this paper, I shed some more light on the role of one particular personality trait, namely the LOC, which can be of particular importance for immigrants and the second generation as they might need to provide extra effort to overcome barriers on the labour market as mentioned in the introduction.

3 Data

The German Socioeconomic Panel (SOEP) is a valuable data set, which is especially rich through questions going beyond purely observable characteristics. It is particularly of interest for this study since it includes personality questions as well as migrant-related and detailed data on labour market returns and educational history of individuals.

In order to capture the effect of LOC on employment, the LOC measure should be cleaned as much as possible from labour market experiences. Therefore, I construct a sample that consists of individuals aged 17–30, who are still in education or at an early stage of their career and employment outcomes are measured 8 years later. Given that we need to take this lagged structure into account, that LOC is measured in 1999, 2005 and 2010 and that the German SOEP is available until 2014, it is either possible to use the 1999 measure or the 2005 measure of LOC. The immigrant sample size is too small with 36 individuals when constructing a sample of individuals aged 17–30 in 2005 and measuring their employment outcomes in 2013. Therefore, I construct a sample of individuals aged 17–30 in 1999 and measure their employment outcomes in 2007.

Hence, employment D , age and education are measured in 2007. Psychometric measures and importance of religion are measured in 1999. The final sample size is 1812. There are 111 immigrants (6.1 % of the sample) and 243 children of immigrants (13.4 % of the sample). The German statistical office reports a percentage of 8.8 % of “foreign population” (inhabitants of Germany with foreign nationality) in Germany in 2008.²

An immigrant is defined as “foreign born with no German nationality at birth”. Note that the sample does not include foreign born with German nationality at birth. Immigrants’ children (“second generation”) are defined as “born in Germany with no German nationality at birth”. Note that this definition excludes second generation immigrants who were born from immigrants, who have obtained German citizenship.

Current employment status is measured by a dummy variable taking the value “1” if the individual is either full-time, halftime or marginally employed. The variable takes the value “0” if the individual is not in employment. Individuals for which this variable was inapplicable and those in vocational training in 2007 are dropped from the sample.

Educational attainment is measured by three dummy variables based on three categories according to the ISCED³ classification. ISCED 0–2 includes education up to the level of general elementary schooling and indicates a low education level, ISCED

3–4 includes “middle vocational schooling” and “vocational plus Abitur” and indicates a medium education level and ISCED 5–6 includes “higher vocational schooling” and “higher education” and indicates a high education level.

Three geopolitical nationality groups take into account the different nationalities present in the sample: “EU15”, “Central Europe and former Soviet Union” and “Turkey”. Turkish immigrants are a large group among German non-nationals. A foreign language indicator takes the value one, if the only language spoken at home is the foreign language.

The data set further includes dummy variables for whether individuals are married (as opposed to single or divorced) and whether individuals have one or more children under 16 at their charge.

The measurement of the LOC goes back to Rotter (1966), who developed an item battery of 23 items. Selected waves of the German Socioeconomic Panel measure LOC in 10 items (see for instance Caliendo et al. 2015; Flossmann et al. 2008; Gallo et al. 2003; Heineck and Anger 2010; Piatek and Pinger 2016), which are given in Appendix 1. The answers in the 2005 wave of the German SOEP are given on a four-point Likert scale and range from “totally disagree” over “slightly disagree” and “slightly agree” to “totally agree”. I merge the first two categories to improve identification of the model, since the first two categories are characterized by low frequencies. Agreement with questions 1, 4 and 6 is seen as an internal LOC whereas agreement with questions 2, 3, 5, 7, 8 and 10 is seen as an external LOC. The questions are chosen using the correlation matrix of the 10 items. Five items display bivariate correlations which are considered sufficiently large in this context: items 2, 3, 5, 7 and 10. Below I test for robustness of the results when relaxing this restriction to 5 out of 10 items (see Section 5 and Appendix 5).

4 Econometric strategy

The econometric strategy is based on a methodology developed in Carneiro et al. (2003) and Heckman et al. (2006). I extend their model framework by allowing the measure of LOC to depend on a set of observable determining variables. This extension is based on work by Fahrmeir and Raach (2006). The model is estimated using a Bayesian Markov Chain Monte Carlo algorithm. Borghans et al. (2008) indeed see latent factor theory as a crucial connecting tool between psychology and economics. Borghans et al. (2008) refer to the work of Carneiro et al. (2003) as a successful example for incorporating psychometric questions in an economic outcome model in a way that addresses the problem of endogeneity.

The econometric model is based on two elements: a personality model—based on a traditional factor model—and on an employment model including latent factors as explanatory variables. The two models are estimated simultaneously, allowing for recognition of the unobservable nature of the latent factor. Indeed, treating the latent factor as observable is a less efficient method than a method estimating all parameters simultaneously and taking into account that the latent variable is an estimated and not an observable entity. On the other hand, if the latent factor is estimated in a wrong way, any mistake is carried on to the estimation of the remaining parameters. The model allows the measure of LOC to depend on observable variables. Especially, we are interested in whether immigrants and their children have different positions on the LOC scale.

4.1 The LOC model

The first element of the model is a classic factor model. Factor models have been developed in psychology to measure intelligence (Spearman 1904). Latent factor models were also used to measure other personality traits, in political science for measuring concepts and in financial economics to measure latent concepts which influence financial markets.

The main idea of factor models is to use a set of measures for the concept “intelligence”, “discipline”, “peace” or “beliefs on the stock market” and to divide the joint variation among these measures into a common part θ and a random part ε and to estimate the common part θ and its effect on the measures, indicated by α . θ indicates in this paper the LOC, which is measured using a set of questions related to the LOC (Rotter 1966). The model is a simultaneous equation model of the five psychometric questions above. Each psychometric question is modelled as an ordered probit model. All five questions are assumed to depend on a latent factor θ , the LOC, and an independent random error term ε^M . The psychometric questions all depend differently on the latent factor—each question has a different factor loading α^M , which can be interpreted as a coefficient of the latent factor in the regression of M on θ . The model takes the form:

$$\begin{aligned} M_1 &= \{1, 2, 3\} \\ M_1^* &= \alpha^{M_1}\theta + \varepsilon^{M_1} \\ M_2 &= \{1, 2, 3\} \\ M_2^* &= \alpha^{M_2}\theta + \varepsilon^{M_2} \\ M_3 &= \{1, 2, 3\} \\ M_3^* &= \alpha^{M_3}\theta + \varepsilon^{M_3} \\ M_4 &= \{1, 2, 3\} \\ M_4^* &= \alpha^{M_4}\theta + \varepsilon^{M_4} \\ M_5 &= \{1, 2, 3\} \\ M_5^* &= \alpha^{M_5}\theta + \varepsilon^{M_5} \end{aligned}$$

Appendix 2 explains how this model can be identified on the basis of the identification of latent factor models.

4.2 The employment model

The second element of the model is an employment model. The latent factor θ , estimated through the model above, is treated as an additional explanatory variable in the employment equation. The model takes the form:

$$\begin{aligned} D &= \{0, 1\} \\ D^* &= \beta_0^D + \alpha^D\theta + \beta^DX + \varepsilon^D \end{aligned}$$

4.3 The simultaneous equation model

Both models described above are estimated simultaneously using a simultaneous equation model. The model is a linear parametric simultaneous equation model with an embedded factor model structure, as described above. The simultaneous equation model contains the equations for the economic outcome D and for the measures M . In this paper,

the latent concept LOC is endogenized and so I add another equation in the simultaneous equation model to determine θ .

The model then takes the following form:

$$\begin{aligned} D &= \{0, 1\} \\ D^* &= \beta_0^D + \alpha^D \theta + \beta^D X + \varepsilon^D \\ M &= \{1, 2, 3\} \\ M^* &= \alpha^M \theta + \varepsilon^M \\ \theta &= \gamma W + \varepsilon^\theta \end{aligned}$$

where D is an employment indicator and M signifies psychometric measures for LOC. Since M and D are categorical variables, we need to impose a probit structure on the variables, so D^* and M^* indicate the latent underlying variables for the probit models for M and D . X comprises the control variables in the employment equation (called *direct effects*) age, gender, immigrant status, education level, language spoken at home, marital status and whether there are any children under 16 at home. These variables have been chosen on the basis of a Mincer equation with typical control variables for immigrants. W comprises the control variables for the latent factor equation (called *indirect effects*) age, gender, whether religion is important, immigrant status, education level, language spoken at home and the time spent in Germany. The importance of religion was added as it is a factor commonly identified as correlated with the LOC (Kahoe 1974). Note that identification of the model X and W could be identical.

Appendix 3 lists the assumptions needed to identify this model.

4.3.1 Estimation: the Gibbs sampler

The model is estimated by a Bayesian Markov Chain Monte Carlo routine. The likelihood function of the model under the assumption of independently and identically distributed observations is given by

$$\begin{aligned} & \prod_{i=1}^N f(M_i, D_i, M_i^*, D_i^*, \theta_i | X_i, W_i, \alpha, \beta, \gamma, c) \\ &= \prod_{i=1}^N f(M_i^*, D_i^*, \theta_i | X_i, W_i, \alpha, \beta, \gamma, c) \prod_{i=1}^N f(M_i, D_i | \theta_i, M_i^*, D_i^*, X_i, W_i, \alpha, \beta, \gamma, c) \\ &= \prod_{i=1}^N f(M_i^*, D_i^*, \theta_i | X_i, W_i, \alpha, \beta, \gamma, c) \prod_{i=1}^N f(M_i, D_i | c) \end{aligned}$$

where the factor loadings are written as $\alpha = (\alpha^M, \alpha^D)$ and the coefficients as $\beta = \beta^D$. The first simplification follows from exploitation of the product rule. The second step follows from the fact that ordinal responses are solely determined by the underlying variables D_i^* and M_i^* and by the cutpoints c . We can factor out the likelihood function $f(M_i^*, D_i^*, \theta_i | X_i, W_i, \alpha, \beta, \gamma, c)$ into $f(M_i^*, \theta_i | \cdot) f(D_i^*, \theta_i | \cdot)$ due to the conditional independence assumptions above. The likelihood functions of D_i^* and M_i^* written separately are

$$\prod_{i=1}^N \left[f(M_i^*, \theta_i | \alpha, \gamma, c, M_i, W_i) \left\{ \sum_{k_M=1}^{K_M} 1(M_i = k_M) 1(c_{k_M-1} < M_i^* < c_{k_M}) \right\} \right]$$

$$\prod_{i=1}^N \left[f(D_i^*, \theta_i | \alpha, \beta, \gamma, D_i, X_i, W_i) \left\{ \sum_{k_D=1}^{K_D} 1(D_i = k_D) 1(c_{k_D-1} < D_i^* < c_{k_D}) \right\} \right]$$

Each of the factors $f(M_i^*, \theta_i | \cdot)$ and $f(D_i^*, \theta_i | \cdot)$ needs to be multiplied by two indicators—an indicator which equals one if the observation M_i (D_i) falls in category k_M (k_D) and an operator indicating that M_i^* (D_i^*) must fall between the two cutpoints c_{k_M-1} (c_{k_D-1}) and c_{k_M} (c_{k_D}) according to its category.

θ is unobservable and will be estimated. To make the mechanism by which θ_i influences M_i^* and of D_i^* perspicuous, we integrate out θ_i and obtain the distributions of M_i^* and D_i^* conditional on the parameters of the model and on the data.

$$f(M_i^* | \alpha, c, \gamma, M_i, W_i) = \int_{\theta} f(M_i^* | \alpha, c, \theta_i, M_i) f(\theta_i | \gamma, W_i) d(\theta_i)$$

$$f(D_i^* | \alpha, \beta, \gamma, D_i, X_i, W_i) = \int_{\theta} f(D_i^* | \alpha, \beta, c, \theta_i, D_i, X_i) f(\theta_i | \gamma, W_i) d(\theta_i)$$

It becomes obvious that the likelihood function of the model is a high-dimensional integral, which cannot be solved analytically and needs to be solved by numerical methods. Markov Chain Monte Carlo methods provide a way to estimate the parameters of interest by sampling from the integral. The main advantage of the Gibbs sampler is its relative computational ease.

The Gibbs sampler is a Bayesian method. The Bayesian paradigm specifies statistical models as a posterior joint distribution, composed of the two elements prior distribution and likelihood function. The prior distribution contains the beliefs of the researcher about the parameters before taking into account the information in the data. The prior is combined with the likelihood function, which contains the information of the data. The posterior joint distribution is obtained by simply multiplying the priors with the likelihood and it can be written as

$$f(\beta, \alpha, \gamma, \theta_i, M^*, D^*, c | M, D, X, W)$$

$$\propto f(\beta) f(\alpha) f(\gamma) f(c) \prod_{i=1}^N f(M_i, D_i, M_i^*, D_i^*, \theta_i | X_i, W_i, \alpha, \beta, \gamma, c)$$

where $f(\beta) f(\alpha) f(\gamma) f(c)$ are the priors for the coefficients of X , the factor loadings, the coefficients of W and the cutpoints.

The Gibbs sampler is an algorithm which samples from this joint posterior distribution in a sequential way. The idea of the Gibbs sampler is to sample one of the elements among $M_i^*, D_i^*, \beta, \alpha, \gamma, c$ and θ at a time, conditioning on the last sampled values for the remaining elements and on the data. This procedure is equivalent to sampling from a set of conditional distributions sequentially. Each conditional distribution is a conditional posterior distribution of a parameter value given the last sampled values of the other parameters and the data. These conditionals—each of them constitutes one step of the Gibbs sampling algorithm—are called “full conditionals”. The closed form of the full conditionals follows from the properties of the model. After a sufficient amount of iterations, the algorithm converges under a set of regularity conditions and the sampled values are samples

from the true posterior.⁴ The algorithm for the model in this paper ran for 100,000 iterations, and convergence statistics do not indicate that the algorithm has not converged. In the following, I derive the full conditionals of the model.

First, a value is sampled from the posterior conditional distribution (or full conditional) of the latent underlying variables, then from the posterior conditional distribution of the factor loadings and so forth. For the second iteration, the same procedure is repeated, conditioning on the sampled values from the first iteration. The very first iteration starts with a set of specified initial values. The algorithm is not sensitive to the choice of the starting values.

Appendix 4 describes the conditional posterior distributions underlying the respective elements of the model; namely of the latent underlying variables, the factor loadings, the direct coefficients, the cutpoints, the latent factors and the indirect coefficients.

5 Results

Table 1 shows the results of a simple Mincer equation for employment including controls for country of origin and nationality.

Table 1 Estimates of the employment equation

β^D	(1)	(2)
Intercept	−1** (−1.99, −0.04)	−1, 34** (−2.31, −0.31)
Age	0, 03** (0.01, 0.05)	0, 04** (0.02, 0.06)
Gender	−0, 56** (−0.70, −0.421)	−0, 59** (−0.72, −0.44)
Immigrant	−0, 4** (−0.67, −0.13)	
Second generation	−0.16 (−0.34, 0.036)	
Turkish immigrant		−0, 65** (−1.06, −0.23)
Central European immigrant		−0.06 (−0.60, 0.50)
EU15 immigrant		0.5 (−0.18, 1.23)
Turkish second generation		−0, 45** (−0.87, −0.02)
Central European second generation		0.32 (−0.28, 0.95)
EU15 second generation		0.16 (−0.30, 0.64)
German second generation		−0.13 (−0.38, 0.12)
Immigrant foreign language at home		−0.53 (−1.19, 0.12)
Second generation foreign language at home		−0.25 (−0.89, 0.41)
Low education	−0, 4** (−0.58, −0.21)	−0, 37** (−0.56, −0.17)
High education	0, 31** (0.14, 0.47)	0, 28** (0.11, 0.45)
Marital status	0.08 (−0.07, 0.24)	0.1 (−0.06, 0.25)
Children under 16	0, 23** (0.08, 0.37)	0, 26** (0.10, 0.41)

Notes: (1) 95 % Bayesian confidence intervals in brackets. These intervals contain 95 % of the posterior probability. If this interval contains 0, the effect measured by the parameter estimate is insignificant. (2) Marginal effects at any $\{\theta, X\}$ value can be derived from the information given (i.e. the estimates of alpha and beta). It is in the nature of probit/logit/etc. that the marginal effects are high when the probability is around 0.5 and low when it is close to zero or one

**Significance at the 95 % confidence level

In the first column, two dummy variables are included for immigrants and for the second generation, respectively. The base category are native-born who have the German nationality at birth. All coefficients display the expected signs: Being an immigrant is significantly associated with a lower employment probability. This effect is attenuated for the second generation whereas the coefficient for the second generation is not significant. Being married does not significantly increase the probability of being employed whereas having children does.⁵

The second column displays the results for a model taking into account the ethnic background and the language spoken at home. The sample is split into four ethnic groups—four dummy variables control for Turkish, central European (this group includes the former Soviet Union), EU15 (including Switzerland and the USA) and German nationalities. The base category are again the German-born with German nationality at birth.⁶ The results for the control variables are similar to the results in the first column. The ethnic variables show that being part of the Turkish first or second generation is significantly associated with a lower employment probability compared to native Germans. It must be noted, however, that the sample sizes for the different ethnic groups are small and this might affect the significance of the coefficients. In order to compare with more traditional estimation methods, Table 8 in Appendix 5 shows that probit regressions yield comparable results.

5.1 Adding LOC

Table 2 shows the results for a simple model of employment—for natives, immigrants and their children, adding a measure of LOC. As outlined above, LOC is treated here as a variable that is correlated with socioeconomic conditions and possibly partly also with genetical heritage.

The results in Table 2 are quite similar to those of Table 1. The coefficient for immigrants is slightly less negative than without controlling for LOC and the coefficient for the second generation is much less negative but still insignificant. Again, being an immigrant is associated with a lower employment probability and this disadvantage is attenuated for the second generation. The last row in Table 2 shows the results for LOC. They show that having a more internal LOC is positively and significantly associated with the employment probability. We can see that an increase by 2.5σ LOC units can compensate for being an immigrant and an increase by 1σ LOC unit for being a second generation immigrants. About four σ LOC units can compensate for having a low educational attainment level as opposed to a medium one.

Table 3 shows the estimations of the factor loadings. They are positive for all items and using both ways of estimating the model (see columns 1 and 2).

In Table 4, I show the coefficients of the determinants of LOC. The results show that being an immigrant is negatively correlated with a more internal LOC. The same is true for the second generation but the effect is attenuated. Age and the duration of stay in Germany are slightly positively and significantly correlated with the LOC. Mother's education—interestingly, as opposed to father's education—is positively and significantly associated with the LOC. Education does seem to be significantly associated with the LOC, but the fact of still being in education does.

The second column shows a model in which I add nationalities and an indicator of whether German is spoken at home. The coefficients on the first set of control variables

Table 2 Estimates of the employment equation with LOC

β^D, α^M	(1)	(2)
Intercept	-0,75 (-1.71, 0.23)	-1,41** (-2.41, -0.39)
Age	0,03** (0.01, 0.05)	0,04** (0.02, 0.06)
Gender	-0,57** (-0.70, -0.42)	-0,58** (-0.72, -0.44)
Low education	-0,4** (-0.59, -0.21)	-0,36** (-0.55, -0.17)
High education	0,27** (0.10, 0.44)	0,26** (0.09, 0.42)
Immigrant	-0,34** (-0.62, -0.05)	
Second generation	-0.11 (-0.31, 0.08)	
Turkish immigrant		-0,6** (-1.02, -0.18)
Central European immigrant		-0,06 (-0.60, 0.51)
EU15 immigrant		0,56 (-0.15, 1.26)
Turkish second generation		-0.41 (-0.85, 0.01)
Central European second generation		0,32 (-0.29, 0.92)
EU15 second generation		0.18 (-0.28, 0.66)
German second generation		-0.11 (-0.36, 0.15)
Immigrant foreign language spoken at home		-0.47 (-1.13, 0.19)
Second generation foreign language spoken at home		-0.16 (-0.82, 0.49)
Marital status	0.02 (-0.13, 0.18)	0.09 (-0.08, 0.24)
Children under 16	0,15** (0.001, 0.29)	0,26** (0.11, 0.41)
LOC	0,13** (0.05, 0.21)	0,1** (0.02, 0.18)

Note: 95 % Bayesian confidence intervals in brackets. These intervals contain 95 % of the posterior probability. If this interval contains 0, the effect measured by the parameter estimate is insignificant

**Significance at the 95 % confidence level

Table 3 Estimates of the psychometric equations

α^M	(1)	(2)
Not achieved what I deserve	0,62** (0.56, 0.69)	0,62** (0.57, 0.69)
Achievements are question of luck	0,44** (0.39, 0.49)	0,44** (0.39, 0.49)
Other people influence my life	0,72** (0.65, 0.79)	0,72** (0.65, 0.79)
Doubt my abilities	0,58** (0.53, 0.65)	0,58** (0.53, 0.65)
Little control over my life	1,37** (1.19, 1.56)	1,33** (1.17, 1.51)

Note: 95 % Bayesian confidence intervals in brackets. These intervals contain 95 % of the posterior probability. If this interval contains 0 the effect measured by the parameter estimate is insignificant

** denotes significance at this confidence level

Table 4 Estimates of determinants of the locus of control

γ	(1)	(2)
Age	0,04** (0.03, 0.04)	0,04** (0.03, 0.04)
Gender	-0.03 (-0.15, 0.08)	-0.04 (-0.16, 0.07)
Religion important	-0.03 (-0.17, 0.1)	-0.04 (-0.17, 0.10)
Low education in 1999	-0.03 (-0.17, 0)	-0.02 (-0.16, 0.12)
High education in 1999	0.16 (-0.02, 0.35)	0.17 (-0.02, 0.35)
Father highly educated	0 (-0.19, 0.19)	-0.01 (-0.20, 0.18)
Mother highly educated	0,35** (0.11, 0.59)	0,34** (0.10, 0.58)
In education in 1999	0,22** (0.05, 0.37)	0,15** (0.00, 0.30)
Immigrant	-0,8** (-1.38, -0.18)	
Second generation	-0.23 (-0.44, 0)	
Turkish immigrant		-0,55** (-1.72, -0.39)
Central European immigrant		-0.19 (-0.76, 0.38)
EU 15 immigrant		-0.6 (-1.39, 0.16)
Turkish second generation		-0,55** (-0.86, -0.22)
Central European second generation		0.02 (-0.37, 0.42)
EU15 second generation		-0.22 (-0.61, 0.15)
German second generation		-0.12 (-0.39, 0.15)
Immigrant foreign language spoken at home		-0.4 (-0.97, 0.18)
Second generation foreign language spoken at home		-0.1 (-0.74, 0.53)
Time stayed in Germany	0,04** (0.00, 0.07)	0.03 (-0.00, 0.07)

Note: 95 % Bayesian confidence intervals in brackets. These intervals contain 95 % of the posterior probability. If this interval contains 0 the effect measured by the parameter estimate is insignificant
 **Significance at the 95 % confidence level

do not change much in size or sign compared to column 1. Age has a small and significantly positive effect on LOC, and mother's education seems to be an important determinant. The results show that there is a risk that in particular immigrants from some regions have a more external position on the LOC scale. It should be noted that, as mentioned above, the sample sizes for the separate ethnic groups are small. A Turkish immigrant can compensate his disadvantage on the labour market by six σ units of belief

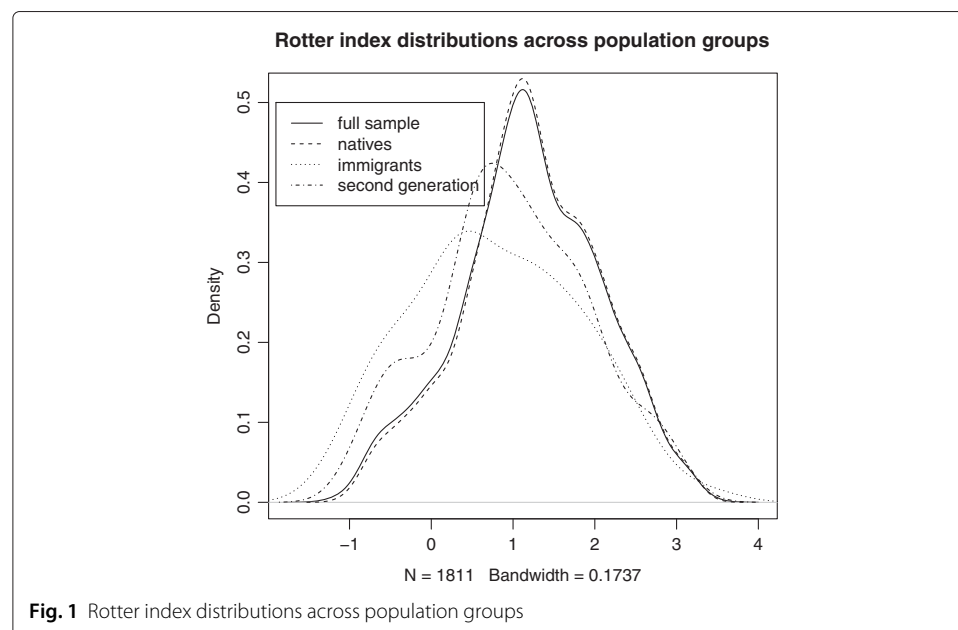
in being able to determine his success. As in Table 2, speaking only the foreign language at home is not significantly negatively associated with the employment probability. LOC still has a positive and significant effect.

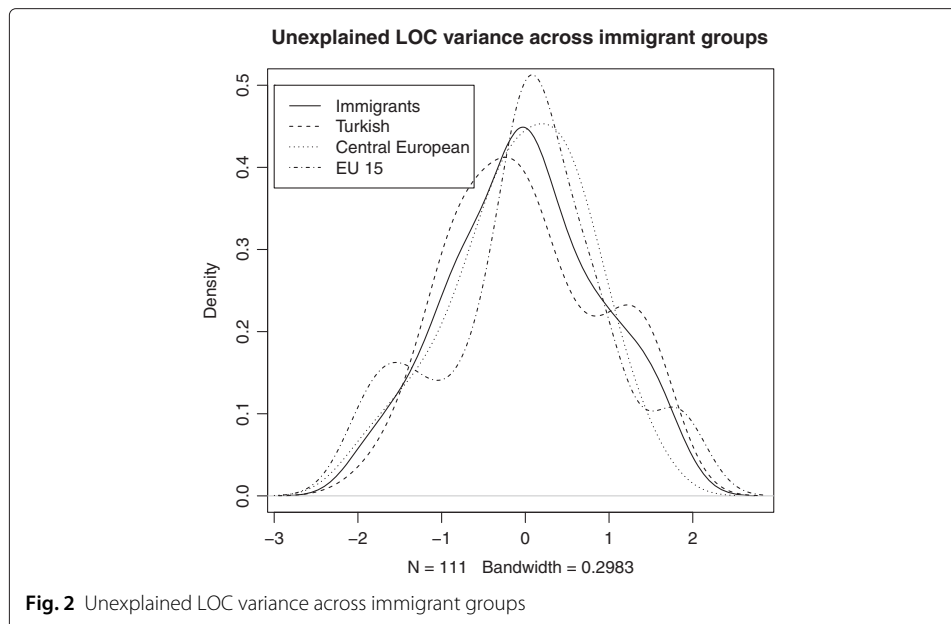
Figure 1 depicts the LOC distributions across population groups—the full sample, natives, immigrants and the second generation. The figure confirms that the mean LOC level is lower for both immigrants and the second generation compared to natives.

To shed some light on why the LOC level is lower among immigrants compared to the native population—and in particular on how important the role of factors other than individual-level factors W are, Fig. 2 shows the variation of the error term ε^θ among each immigrant group. This unaccounted variance could stem from destination country factors such as discrimination. Figure 2 shows that the probability mass of the unaccounted variance in LOC is more centred around zero for those immigrant groups that are culturally closer to Germany such as the EU15 group. This may point to the fact that discrimination—which is not explicitly taken into account in W —may play a larger role among Turkish and Central European immigrants. This finding should be taken with caution however, as W does not account exhaustively for all factors on the individual level.

The outlined results above suggest that the sense of personal control is positively and significantly associated with employment and that immigrants can overcome their disadvantage by believing in success. Immigrants—especially Turkish—indeed have a *double disadvantage* on the labour market: they are disadvantaged in terms of employment and additionally they tend to have a more external position on the LOC scale.

In Appendix 5 (see Tables 8, 9 and 10), I provide several comparisons with the more traditional probit model as well as robustness checks. I test whether using ten items instead of five and whether using the full age range instead of including only those individuals at the early stage of their career has a considerable impact on the main findings. The probit model was estimated with an estimate of θ obtained from the model in this section.





6 Conclusions

This paper set out to examine whether the sense of personal control is positively associated with the labour market performance of immigrants to Germany, using a concept well-known in psychology as the LOC. I find that a strong sense of personal control is positively and significantly associated with the probability of being employed. There is a risk that immigrants may tend to have a more external LOC than natives. The second generation tends to also have a more external LOC than natives, but it seems to be already more internal than that of the first generation. This is taken as evidence for a generational convergence of migrants' LOC towards that of natives.

The results of this article imply that there is a risk that immigrants may have a double disadvantage on the labour market: they are often disadvantaged by lower employment chances because of their status and they tend to have a more external position on the LOC. In addition, mother's education was found to potentially play an important role in cultivating sense of control.

Based on the association between a stronger sense of personal control and a higher employment probability identified in this paper, it can be concluded that policy interventions or programs that enhance the sense of personal control—such as simple questionnaires increasing awareness of the personal sense of control—or including training sessions on developing a more internal locus of control in integration classes can potentially be an important and simple public policy tool for the integration of immigrants.

Endnotes

¹As addressed in our econometric strategy below, this trait could in turn be determined by the family background and previous experiences.

²http://www.statistik-portal.de/Statistik-Portal/de_jb01_jahrta2.asp.

³UNESCO (2006): ISCED 1997 - International Standard Classification of Education, www.uis.unesco.org.

⁴For the theory MCMC algorithms and on the Gibbs sampler, see Robert and Casella (2004).

⁵It would be of interest to split the sample into men and women since it could certainly be the case that the coefficient for marital status differs largely between men and women and therefore renders the coefficient insignificant once taking the whole sample. But here the interest lies in the immigrant population and I consider the immigrant and the second generation samples as too small to be able to split the sample.

⁶There is no group for foreign born with German nationality since there are none in the sample.

⁷As above β and θ denote the last sampled values.

Appendix 1: LOC questions in the German SOEP

1. How my life goes depends on me.
2. Compared to other people, I have not achieved what I deserve.
3. What a person achieves in life is above all a question of fate or luck.
4. If a person is socially or politically active, he/she can have an effect on social conditions.
5. I frequently have the experience that other people have a controlling influence over my life.
6. One has to work hard in order to succeed.
7. If I run up against difficulties in life, I often doubt my own abilities.
8. The opportunities that I have in life are determined by the social conditions.
9. Inborn abilities are more important than any efforts one can make.
10. I have little control over the things that happen in my life

Appendix 2: Parametric identification of factor models

Here I give a brief outline of the identification of factor models. Factor models take the form of the measurement equation above:

$$M^* = \alpha^M \theta + \varepsilon^M$$

Consider M^* to be computable.

The identification of factor models is based on the covariance matrix of the items:

$$\text{cov}(M^*) = \Lambda \Sigma_f \Lambda' + \Omega_e$$

where

$$\begin{aligned} \theta &\perp \varepsilon^M \\ \varepsilon^M &\sim N(0, 1) \end{aligned}$$

Λ - matrix of factor loadings α^M

Σ_f - variance-covariance matrix of the factors

Ω_e - diagonal matrix of “uniqueness”-variances of ε^M

K - number of factors θ

L - number of items Mx

The goal is to identify $K \times L$ factor loadings Λ and K variances of factors Σ_f . The elements of $\text{cov}(M)$ are observable and the elements of Ω_e are determined by our

distributional assumption on ε^M . So we can identify the unobservable elements Λ and Σ with the $(L(L - 1)/2)$ observable off-diagonal elements of $\text{cov}(M)$. So, we need that

$$L(L - 1)/2 \geq (L \times K) + K.$$

The number of unique terms in $\text{cov}(M)$ needs to be equal to or larger than the number of factor variances and factors. In our case $K = 1$ and $L = 5$. So we have

$$\begin{aligned} 5 * 4/2 &\geq 5 + 1 \\ 10 &\geq 6 \end{aligned}$$

Appendix 3: Identification assumptions of the simultaneous equation model

The identification strategy is parametric and we need to make assumptions on the distributions of the error terms and of the latent concept.

$$\begin{aligned} \varepsilon_D &\sim N(0, 1) \\ \varepsilon_M &\sim N(0, 1) \end{aligned}$$

We need to impose normalization conditions on D_i^* and M_i^* : $V(D_i^*)$ is normalized to 1, $V(M_i^*)$ is normalized to 1 and we impose normality on θ conditional on W .

Finally we need to impose conditional independence conditions:

$$\begin{aligned} \theta &\perp \varepsilon_M | W \\ \theta &\perp \varepsilon_D | W, X \\ X &\perp \theta | W \\ D &\perp M | \theta, X \\ M_j &\perp M_{j-1} | \theta \quad \forall j \end{aligned}$$

For all tri-categorical items the cut point between the first and the second category is $c_1 = 1$.

Appendix 4: The posterior conditional distributions of the model

The posterior conditional distributions of the latent underlying variables

Albert and Chib (1993) propose a data augmentation procedure to sample latent underlying variables in a threshold model. It follows from their work that the full conditional for the latent underlying variable of the binary response is

$$f(D^* | \alpha^D, \beta^D, \theta, D, X) \propto \prod_{i=1}^N f(D_i^* | \beta^D X_i^D + \alpha^D \theta_i, 1) \left\{ \sum_{k_D=1}^{K_D} 1(D_i = k_D) 1(c_{k_D-1} < D_i^* < c_{k_D}) \right\}$$

where $\alpha^D, \beta^D, \theta$ signify the last sampled values from the previous iteration of the algorithm. It follows from the normality assumptions on θ and ε that $f(D_i^* | \theta_i, \alpha, \beta, D_i, X_i)$ is normally distributed - with mean $\beta^D X_i^D + \alpha^D \theta_i$ and $V(D_i^*)$ normalized to one as indicated above.

The latent underlying variable is distributed as the following truncated normal distributions:

$$D_i^* | \alpha, \beta, \theta, D, X \sim TN_{(-\infty, 0)} (\beta^D X_i^D + \alpha^D \theta_i, 1) \text{ if } D_i = 0$$

$$D_i^* | \alpha, \beta, \theta, D, X \sim TN_{(0, \infty)} (\beta^D X_i^D + \alpha^D \theta_i, 1) \text{ if } D_i = 1$$

Similarly, the full conditionals for the polytomous variables are

$$f(M^* | \alpha, \theta, c, M, X) \propto \prod_{i=1}^N f(M_i^* | \alpha^M \theta_i, 1) \left\{ \sum_{k_M=1}^{K_M} 1(M_i = k_M) 1(c_{k_M-1} < M_i^* < c_{k_M}) \right\}$$

The latent underlying variables of the polytomous items is distributed as the following truncated normal distribution:

$$M_i^* | \alpha, \theta, c, M, X \sim TN_{(c_{k_M-1}, c_{k_M})} (\alpha^M \theta_i, 1)$$

The posterior conditional distribution of the factor loadings

The full conditional for the factor loadings for D and M can be written as⁷

$$f(\alpha^D | \beta, \theta, D, X, D^*) \propto f(\alpha^D) \prod_{i=1}^N f(D_i^* | \beta^D X_i^D + \alpha^D \theta_i, 1)$$

$$f(\alpha^M | \theta, M, X, M^*) \propto f(\alpha^M) \prod_{i=1}^N f(M_i^* | \alpha^M \theta_i, 1)$$

where we choose normal priors $f(\alpha^D) = N(0, 1)$ and $f(\alpha^M) = N(0, 1)$. If we rewrite the equation for D^* and M^* as

$$D_i^* - \beta^D X_i^D = \alpha^D \theta_i + \varepsilon_i^D$$

$$M_i^* = \alpha^M \theta_i + \varepsilon_i^M$$

we can treat it as a normal regression model and derive for M and D

$$\alpha^M | \theta_i, M_i, M_i^* \sim N \left[(\theta_i' \theta_i + 1)^{-1} \theta_i' (M_i^*), (\theta_i' \theta_i + 1)^{-1} \right]$$

$$\alpha^D | \beta, \theta_i, D_i, X_i, D_i^* \sim N \left[(\theta_i' \theta_i + 1)^{-1} \theta_i' (D_i^* - \beta^D X_i^D), (\theta_i' \theta_i + 1)^{-1} \right]$$

The posterior conditional distribution of the direct coefficients

Similarly to the procedure for the factor loadings, we can write the model as

$$D_i^* - \alpha^D \theta_i = \beta^D X_i^D + \varepsilon_i^D$$

For the coefficients, we choose to set diffuse priors as well. The full conditionals for the intercepts are, according to Albert and Chib (1993), p.671

$$\beta^D | \alpha, \theta_i, D_i, X_i, D_i^* \sim N \left[(X_i' X_i)^{-1} X_i' (D_i^* - \alpha^D \theta_i), (X_i' X_i)^{-1} \right]$$

The posterior conditional distribution of the cutpoints

We assume a uniform prior for the cutpoints and can write for the full conditionals for the polytomous responses

$$c^M | \alpha, \theta, M, M^* \sim unif \left[\begin{array}{l} \max \{ \max \{ M_i^* : M_i = k_M \}, c_{M-1} \}, \\ \min \{ \min \{ M_i^* : M_i = k_{M+1} \}, c_{M+1} \} \end{array} \right]$$

The posterior conditional distribution of the latent factors

Similarly as for the procedure for coefficients and factor loadings, we can rewrite the model as

$$\begin{aligned} D_i^* - \beta^D X_i^D &= \alpha^D \theta_i + \varepsilon_i^D \\ M_i^* &= \alpha^M \theta_i + \varepsilon_i^M \end{aligned}$$

and treat it as a normal regression model, where θ_i is the parameter to be estimated. Carneiro et al. (2003) specify a mixture of normals as the prior for the latent factors. We treat the latent factors as endogenous depending on γW_i . We treat θ_i in the same way as M_i^* and D_i^* for which the priors are implicitly determined by the prior distributions of the other parameters and by the assumptions on the distribution of ε_i^M and ε_i^D . The prior of θ_i is therefore implicitly determined by the priors of the other parameters of the model and by the assumptions on the distributions of ε_i^M , ε_i^D and ε_i^θ .

We can derive the full conditional for the latent factor as:

$$\begin{aligned} &f(\theta | \beta, \alpha, c, \gamma, X, W, D^*, M^*) \\ &\propto \prod_{i=1}^N f(M_i^* | \alpha^M \theta_i, 1) f(D_i^* | \beta^D X_i^D + \alpha^D \theta_i, 1) \end{aligned}$$

We do not need to condition on D and M since they are implicitly known through D^* and M^* and c . Our dependent variables are ordinal and for identification reasons their variances and error variances have been set to one.

The posterior conditional distribution of θ_i is given by:

$$\theta_i | \beta, \alpha, \gamma, c, X_i, W_i, D_i^*, M_i^* \sim N \left[\begin{array}{l} \gamma W_i + (\alpha^{D'} \alpha^D + \alpha^{M'} \alpha^M + 1)^{-1} \\ (\alpha^{M'} (M_i^* - \alpha^{M'} \gamma W_i) + \alpha^D (D_i^* - \beta^D X_i^D - \alpha^D \gamma W_i)), \\ I - \alpha^{D'} (\alpha^{D'} \alpha^D + \alpha^{M'} \alpha^M + 1)^{-1} \alpha^D \\ - \alpha^{M'} (\alpha^{D'} \alpha^D + \alpha^{M'} \alpha^M + 1)^{-1} \alpha^M \end{array} \right]$$

The posterior conditional distribution of the indirect coefficients

The posterior we sample from can be written as

$$\begin{aligned} &f(\gamma | \theta, W) \\ &\propto f(\gamma) f(\theta | \gamma, W) \end{aligned}$$

The model for the latent variable is

$$\theta = \gamma W + \varepsilon^\theta$$

We assume a diffuse prior for the coefficient γ . Similar to the procedures above we get:

$$f(\gamma|\theta, W) \sim N((W'W)^{-1}W'\theta), (W'W)^{-1})$$

Appendix 5: Robustness checks

Table 5 Estimates of the employment equation, reduced-form estimation with probit

β^D	(1)	(2)
Intercept	-1.26* (0.51)	-1.33** (0.51)
Age	0.04*** (0.01)	0.04*** (0.01)
Gender	-0.57*** (0.07)	-0.58*** (0.07)
Immigrant	-0.41** (0.14)	
Second generation	-0.12 (0.10)	
Turkish immigrant		-0.65** (0.21)
Central European immigrant		-0.07 (0.27)
EU15 immigrant		0.46 (0.36)
Turkish second generation		-0.45* (0.22)
Central European second generation		0.30 (0.31)
EU15 second generation		0.15 (0.24)
German second generation		-0.13 (0.13)
Immigrant foreign language at home		-0.52 (0.33)
Second generation foreign language at home		-0.25 (0.34)
Low education	-0.41*** (0.10)	-0.36*** (0.10)
High education	0.29*** (0.09)	0.28** (0.09)
Marital status	0.08 (0.08)	0.09 (0.08)
Children under 16	0.26*** (0.08)	0.26*** (0.08)

Notes: (1) standard errors in brackets with codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 (2) marginal effects at any θ , X value can be derived from the information given (i.e. the estimates of alpha and beta). It is in the nature of probit/logit/etc. that the marginal effects are high when the probability is around 0.5 and low when it is close to zero or one

Table 6 Estimates of the employment equation with LOC, reduced-form estimation with probit

β_D	(1)	(2)
Intercept	-1.39** (0.51)	-1.43** (0.52)
Age	0.03*** (0.01)	0.04*** (0.01)
Gender	-0.57*** (0.07)	-0.58*** (0.07)
Immigrant	-0.33* (0.14)	
Second generation	-0.07 (0.10)	
Turkish immigrant		-0.58** (0.21)
Central European immigrant		-0.07 (0.28)
EU15 immigrant		0.54 (0.37)
Turkish second generation		-0.40 (0.22)
Central European second generation		0.29 (0.31)
EU15 second generation		0.18 (0.24)
German second generation		-0.10 (0.13)
Immigrant foreign language at home		-0.42 (0.33)
Second generation foreign language at home		-0.11 (0.34)
Low education	-0.39*** (0.10)	-0.35*** (0.10)
High education	0.26** (0.09)	0.25** (0.09)
Marital status	0.07 (0.09)	0.08 (0.08)
Children under 16	0.26*** (0.08)	0.26*** (0.08)
LOC	0.17*** (0.04)	0.15*** (0.04)

Note: standard errors in brackets with the following codes: 0 **** 0.001 *** 0.01 ** 0.05 * 0.1 ' ' 1

Table 7 Estimates of the employment equation with LOC, reduced-form estimation with probit

β^D	(1)	(2)
Intercept	-1.36** (0.51)	-1.41** (0.52)
Age	0.03*** (0.01)	0.04*** (0.01)
Gender	-0.56*** (0.07)	-0.57*** (0.07)
Immigrant	-0.52** (0.18)	
Second generation	-0.12 (0.14)	
Turkish immigrant		-0.73** (0.24)
Central European immigrant		-0.31 (0.34)
EU15 immigrant		0.40 (0.38)
Turkish second generation		-0.41 (0.24)
Central European second generation		0.28 (0.33)
EU15 second generation		0.17 (0.27)
German second generation		-0.11 (0.17)
Immigrant foreign language at home		-0.31 (0.35)
Second generation foreign language at home		-0.12 (0.35)
Low education	-0.39*** (0.10)	-0.35*** (0.10)
High education	0.26** (0.09)	0.25** (0.09)
Marital status	0.06 (0.08)	0.08 (0.08)
Children under 16	0.26*** (0.08)	0.26*** (0.08)
LOC	0.14** (0.05)	0.13** (0.05)
Immigrant* LOC	0.28 (0.16)	0.21 (0.17)
Second* LOC	0.04 (0.11)	0.01 (0.11)

Note: standard errors in brackets with the following codes: 0 **** 0.001 *** 0.01 ** 0.05 * 0.1 ' ' 1

Table 8 Estimates of the employment equation, robustness checks

β_D	Full age range (17–64)	10 items
Intercept	1.75** (1.16,2.36)	-1.55** (-2.58, -0.52)
Age	-0.03** (-0.03, -0.03)	0.03** (0.01,0.05)
Gender	-0.46** (-0.53, -0.39)	-0.57** (-0.71, -0.43)
Immigrant	-0.24** (-0.37, -0.10)	-0.30** (-0.58, -0.00)
Second generation	-0.24** (-0.37, -0.10)	-0.07 (-0.27,0.13)
Low education	-0.25** (-0.34, -0.15)	-0.40** (-0.59, -0.21)
High education	0.43** (0.34,0.51)	0.29** (0.11,0.45)
Marital status	0.15** (0.07,0.23)	0.07 (-0.09,0.23)
Children under 16	0.05 (-0.03,0.13)	0.27** (0.12,0.42)
LOC	0.18** (0.05,0.14)	0.09** (0.07,0.28)

Note: 95 % Bayesian confidence intervals in brackets. These intervals contain 95 % of the posterior probability. If this interval contains 0, the effect measured by the parameter estimate is insignificant
 **Significance at the 95 % confidence level

Table 9 Estimates of the psychometric equations, robustness checks

α^M	Full age range (17–64)	10 items
Not achieved what I deserve	0.60** (0.55,0.65)	0.01 (-0.00,0.04)
Achievements are question of luck	0.48** (0.45,0.52)	0.29** (0.23,0.35)
Other people influence my life	0.81** (0.76,0.87)	0.20** (0.15,0.24)
Doubt my abilities	0.75** (0.72,0.79)	0.61** (0.54,0.69)
Little control over my life	1.42** (1.30,1.53)	0.33** (0.26,0.40)
Lifecourse depends on me		-0.03** (-0.06, -0.01)
Social/political activity influences social conditions		0.23** (0.18,0.28)
Work hard to succeed		0.11** (0.08,0.13)
Opportunities determined by social conditions		0.21** (0.17,0.25)
Abilities more important than effort		0.48** (0.39,0.57)

Note: 95 % Bayesian confidence intervals in brackets. These intervals contain 95 % of the posterior probability. If this interval contains 0, the effect measured by the parameter estimate is insignificant
 **Significance at this confidence level

Table 10 Estimates of determinants of the locus of control, robustness checks

γ	Full age range (17–64)	10 items
Age	0.02** (0.02,0.02)	0.08** (0.067,0.09)
Gender	0.02 (–0.04,0.08)	–0.02 (–0.17,0.13)
Religion important	–0.07 (–0.13,0.00)	–0.19** (–0.37, –0.02)
Low education in 1999	–0.02 (–0.10,0.07)	0.05 (–0.12,0.23)
High education in 1999	0.23** (0.15,0.30)	0.01 (–0.23,0.25)
Father highly educated	0.10 (–0.01,0.21)	–0.13 (–0.38,0.11)
Mother highly educated	0.34** (0.18,0.50)	0.30 (–0.01,0.60)
In education in 1999	0.34** (0.20,0.47)	0.35** (0.15,0.56)
Turkish immigrant	0.14 (–0.18,0.45)	–0.00 (–1.99,0.30)
Central European immigrant	0.18 (–0.15,0.50)	0.01 (–0.71,0.75)
EU 15 immigrant	0.57** (0.19,0.91)	–0.36 (–1.35,0.61)
Turkish second generation	–0.26 (–0.51,0.01)	–0.57** (–0.98, –0.16)
Central European second generation	0.29 (–0.08,0.64)	0.24 (–0.27,0.76)
EU15 second generation	0.16 (–0.12,0.45)	–0.11 (–0.58,0.38)
German second generation	–0.04 (–0.19,0.11)	–0.05 (–0.39,0.31)
Immigrant foreign language spoken at home	–0.77** (–1.02, –0.50)	–0.53 (–1.25,0.20)
Second generation foreign language spoken at home	–0.17 (–0.66,0.31)	–0.06 (–0.89,0.75)
Time stayed in Germany	–0.02** (–0.03, –0.01)	0.03 (–0.02,0.07)

Note: 95 % Bayesian confidence intervals in brackets. These intervals contain 95 % of the posterior probability. If this interval contains 0, the effect measured by the parameter estimate is insignificant

**Significance at this confidence level

Competing interests

The IZA Journal of Migration is committed to the IZA Guiding Principles of Research Integrity. The author declares that she has observed these principles.

Authors' information

Anna Thum-Thyssen is an economic analyst at the European Commission (DG ECFIN). She has previously held positions at the Centre for European Policy Studies (CEPS), the Organization for Economic Cooperation and Development (OECD) and the New Zealand Institute for Economic Research (NZIER). She is an economist specialized in labour economics, economics of human capital and innovation.

Acknowledgements

This research was conducted at the European University Institute with a grant from the German Exchange Service (DAAD). I am very grateful to the valuable comments of an anonymous referee. I would like to thank Professor Richard Spady, Professor Jerome Adda, Professor Pascal Courty and Professor James Heckman for their support and insightful comments. I would also like to thank Macro Lombardi, Alicia Perez-Alonso as well as participants of the IZA summer school in 2007, the Royal Economic Society Meeting in 2010 and seminar participants at the University of Maastricht and at the University of Leuven for very helpful comments. Finally, I would like to thank Thomas Liebig for my time at the OECD International Migration Division.
Responsible editor: Denis Fougère.

Received: 23 December 2015 Accepted: 20 May 2016

Published online: 26 August 2016

References

- Albert JH, Chib S (1993) Bayesian analysis of binary and polychotomous response data. *J Am Stat Ass* 88(422):669–79
- Almlund M, Duckworth A, Heckman J, Kautz T (2011) Personality psychology and economics. In: Hanushek E, Machin S, Wößmann L (eds). *Handbook of the economics of education*. Elsevier, Amsterdam
- Andrisani PJ (1977) Internal-External Attitudes, Personal Initiative, and the Labor Market Experience of Black and White Men. *J Hum Res* 12:308–28
- Bonev P, Landmann A, Olapade M (2010) Can personality traits help to explain differences in employment dynamics between natives and migrants in Germany? Paper presented at the SOEP 2010 9th International German Socio-Economic Panel User Conference. Social Science Research Center (WZB), Berlin, 30 June - 1 July 2010
- Borghans L, Duckworth A, Heckman J, terWeel B (2008) The economics of psychology and personality traits. *J Hum Res* 43:972–1059
- Bowles S, Gintis H, Osborne M (2001) Incentive-enhancing preferences: personality, behavior and earnings. *Am Econ Rev* 91(2):155–8
- Bowles S, Gintis H, Osborne M (2001) The determinants of earnings: a behavioral approach. *J Econ Lit* 39(4):1137–76
- Brunello G, Schlotter M (2011) Non cognitive skills and personality traits: labour market relevance and their development in education and training systems. IZA Discussion Paper No. 5743. <http://ftp.iza.org/dp5743.pdf>
- Caliendo M, Cobb-Clark D, Uhlenhorff A (2015) Locus of control and job search strategies. *Rev Econ Stat* 97(1):88–103
- Carneiro P, Hansen K, Heckman J (2003) Estimating distributions of treatment effects with an application to the returns of schooling and measurement of the effects of uncertainty of college choice. *Int Econ Rev* 44(2):361–442
- Cobb-Clark D (2015) Locus of control and the labor market. *IZA J Lab Econ* 4:3:1–19
- de Palo D, Faini R, Venturini A (2006) The social assimilation of immigrants. IZA Discussion Paper No. 2439. <http://ftp.iza.org/dp2439.pdf>
- Dormann C, Fay D, Zapf D, Frese M (2006) A state-trait analysis of job satisfaction: on the effect of core self-evaluations. *App Psy: Int Rev* 55(1):27–51
- Duncan GJ, Dunifon R (1998) Long-run effects of motivation on labour market success. *Soc Psy Qu* 61(1):33–48
- Fahrmeir L, Raach A (2006) A Bayesian semiparametric latent variable model for mixed responses. *Psychometrika* 72(3):327–46
- Fertig M (2004) The societal integration of immigrants in Germany. In: Epstein GS, Gang IN (eds). *Migration and Culture, Frontiers of Economics and Globalization*. Emerald Publishing, Bingley Vol. 8
- Flossmann A, Piatek R, Wichert L (2008) Going beyond returns to education: the role of noncognitive skills on wages in Germany. DIW Discussion Papers No. 836. http://www.diw.de/documents/publikationen/73/diw_01.c.90384.de/dp836.pdf
- Gallo WT, Endrass J, Bradley EH, Hell D, Kasl SV (2003) The influence of internal control on the employment status of German workers. *Schmollers Jahrbuch* 123:71–82
- Heckman J (1995) Lessons from the bell curve. *J Pol Econ* 103(5):1019–120
- Heckman J, Stixrud J, Urzua S (2006) The effects of cognitive and non-cognitive abilities on labour market outcomes and social behavior. *J Lab Econ* 24(3):411–82
- Heineck G, Anger S (2010) The returns to cognitive abilities and personality traits in Germany. *Lab Econ*:535–46
- Judge T, Erez A, Bono J, Thoresen C (2002) Are measures of self-esteem, neuroticism, locus of control, and generalized self-efficacy indicators of a common core construct? *J Pers Soc Psy* 83(3):693–710
- Judge T, Heller D, Klinger D (2008) The dispositional sources of job satisfaction: a comparative test. *App Psy: Int Rev* 57(3):361–72
- Judge T, Bono J (2001) A rose by any other name: are self-esteem, generalized self-efficacy, neuroticism, and locus of control indicators of a common construct? In: Roberts BW, Hogan R (eds). *Personality Psychology in the Workplace*. American Psychological Association, Washington, DC
- Kahoe RD (1974) Personality and achievement correlates of intrinsic and extrinsic religious orientations. *J Pers Soc Psy* 29:812–18

- Groves MO (2005) How important is your personality? Labour market returns to personality for women in the US and UK. *J Econ Psy* 26:827–41
- Piatek R, Pinger P (2016) Maintaining (Locus of) Control? Data Combination for the Identification and Inference of Factor Structure Models. *J Appl Econ* 31:734–755
- Robert CP, Casella G (2004) Monte Carlo statistical methods. 2nd edition. Springer, New York
- Rotter JB (1966) Generalized expectancies for internal versus external control of reinforcement. *Psy Mon: Gen and App* 80(1):1–28
- Spearman C (1904) General intelligence, objectively determined and measured. *Am J Psy* 15:201–93
- Uhlendorff A (2004) Der Einfluss von Persönlichkeitseigenschaften und sozialen Ressourcen auf die Arbeitslosigkeitsdauer. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie* 56(2):279–303

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com
