

Learn English, not the Local Language! Ethnic Russians in the Baltic States

By OTT TOOMET*

Previous immigration literature indicates that one of the most important skills on the labor market is the knowledge of the local language. Different studies suggest that fluency in the language of the destination country is typically associated with 10-30% wage premium (Chiswick and Miller, 1995; Chiswick, 1998; Leslie and Lindley, 2001; Chiswick and Miller, 2002; Shields and Price, 2002; Rooth and Saarela, 2007). The positive association may be related to easier access to better-paying jobs outside of the limited ethnic labor market, or making a minority worker genuinely more productive, for instance through improved communication with customers or colleagues.

In order for any of these mechanisms to be employed, one needs both willingness and opportunity for communication. Unless the majority is willing to communicate and work together with the immigrants, the language skills are of little use. The language premium may also depend on the size and sophistication of the minority labor market. A small, mainly low-skilled minority group is unable to create high-productivity jobs in the corresponding ethnic economy and hence the language premium for skilled minority members may be large. The opposite may be true for a large and high-skilled group.

The bulk of the studies analyzing the language premium are related to immigrants in developed economies. These studies typically indicate that despite of the substantial wage premium associated with fluency in the local language, it still does not close the immigrant-native income differential (Leslie and Lindley, 2001). The remaining gap is often loosely attributed to country-specific human capital and discrimination. Studies on anglophones in Quebec indicate that French skills are indeed related to income premium, while there is little residual income differential (Albouy, 2008). This outcome can be explained by the fact that both anglophones and francophones

mostly share similar human capital while ethnic discrimination is virtually absent.

In this study, we extend this literature by analyzing Russian-speaking minorities in the Baltic States (Estonia and Latvia). These countries house large Russian-speaking minority groups that are similar to the titular populations in the most relevant dimensions of human capital. At the same time, tense relations between the dominant and Russian-speaking populations, stemming from the World War II era, ensure that a lot of public attention is paid to the questions of language and ethnic background.

We find that ethnic Russian men earn very little premium on the proficiency of the local language, only the public administration sector shows a clear positive association. These results suggest that even fluency in the majority language does not guarantee access to more productive jobs in presence of ethnic tensions.

I. Background

This section briefly reviews the unique institutional background of Estonia and Latvia.

Before the Second World War, around 90% of the Estonian and about 75% of the Latvian population were ethnic Estonians and Latvians, respectively. After the war, the Baltic States, now incorporated into the USSR, became a destination of substantial migration. This resulted in the immigrant population reaching about 40% in Estonia and 50% in Latvia by 1989. The large inflow of mainly Russian-speaking labor led to having two *de facto* official languages by the 1970s. Certain areas in the economic and public sphere, such as the army, railways, and the merchant fleet were completely dominated by Russian-speaking workers. The widening use of Russian caused increasing concerns about the future of the local people and language among the ethnic Estonians and Latvians. One particular outcome of these concerns was an unwillingness to participate in mainstream Soviet society and in this way the countries remained fairly segregated by ethnicity in everyday life.

* Toomet: Department of Economics, Tartu University, Narva 4, Tartu EE51009, ottoomet@ut.ee. I am grateful for financial support from MD-RDE project in NORFACE framework.

The tide turned during the last years of *perestroika*. The titular population became organized relatively quickly and grasped the upper hand after the collapse of the Soviet Union in 1991. Both Estonian and Latvian were instated as the sole official languages of the respective countries. Due to the new role of language in the new nation-states, the younger generation of the titular population became less and less fluent in Russian while the opposite was true for ethnic Russians.

Both countries have retained segregated schools, divided according to the language of instruction. In this way one important institution to boost social ties in relatively early age has been deficient. In everyday life, both ethnic groups are largely living on their own with limited contacts. The separate worlds are also reflected in media which may present quite different and occasionally antagonistic viewpoints depending on the language (Korts and Kõuts, 2002).

The relationship between the two main ethnic groups has been somewhat tense. Most notably, tensions exploded into large-scale riots in Tallinn, Estonia, in spring 2007 due to relocation of a Soviet WWII monument. There is no direct evidence of ethnic discrimination. However, substantial ethnic income disparities arose (and still persist) around the time of the collapse of the Communist Bloc (Leping and Toomet, 2008).

II. Data and Empirical Model

This study is based on two data sources: Estonian Labor Force Survey (LFS) and Paths of a Generation (PoG). Besides common socioeconomic information, the LFS includes two variables, crucial for this study. First, it reports the “ethnic nationality”, the ethnic identity of the respondents. As most of those who do not consider themselves ethnic Estonians use Russian as their first language, we call them “Russians” below. Second, the data contains information on language skills. The fluency level is coded as “understanding”, “speaking”, “writing”, and “spoken at home”. Below, we use the linear scale: 0 (no knowledge), 0.33 (understanding), 0.66 (speaking), and 1.00 (writing or home language).¹ We use the monthly wage at the main job as our income variable. We select LFS observation from years 2000-2010 Q1, and focus on minority men of 20-60 years of age. This

¹Similar scale was used by Bleakley and Chin (2004) and Chiswick and Miller (2010). Using dummies for different fluency levels gives qualitatively similar results.

gives us about 8,000 observations across 3,500 distinct individuals.

As the Latvian Labor Force Survey does not include information on language skills, we cannot conduct the same analysis for Latvia in similar detail. However, we rely on another dataset, the 1997 wave of Paths of a Generation (PoG), a panel study of high school graduates of 1982 (Titma and Tuma, 1995). That dataset contains similar background information as LFS, including socioeconomic variables, ethnic nationality and language skills. We use gross income for previous months as our income variable. The information on language command is coded in a broadly similar way, namely, the survey distinguishes between fluency (assigned numeric code 1), speaking but not writing (0.75), limited communication (0.5), basic phrases (0.25), and no knowledge at all (0).

We follow a simple empirical strategy, estimating the labor market outcomes for the Russian minority as a function of the language skills and other workplace and human capital variables of individual i :

$$(1) \quad \log w_{it} = \alpha' X_{it} + \beta L_{it} + \eta_i + v_t + \varepsilon_{it}.$$

Here L is the variable capturing language skills, X are the other individual and job-specific characteristics, η_i is the individual-specific effect, v_t is the time period specific effect (we use both year dummies and a flexible polynomial), and ε_{it} is the idiosyncratic error term. We include the individual random effects η only for LFS data as the PoG wave only contains a single observation for each individual. We employ quantile regressions (Koenker and Bassett, 1978), using the same specification as above, to obtain additional information about the distribution of income, depending on variables of interest. The panel version of the quantile regression is estimated using Koenker (2004) type methodology.

III. Results

Table 1 provides the estimates for the language skills variables.² The results based on LFS data

²Here, and below, the control variables include (5-th degree polynomial) time trend, age, education, immigrant status, family status, children, and capital city dummies. Two variables, residence in the capital city and education, are highly correlated with both language skills and income. Including or excluding the other variables does not qualitatively influence the results.

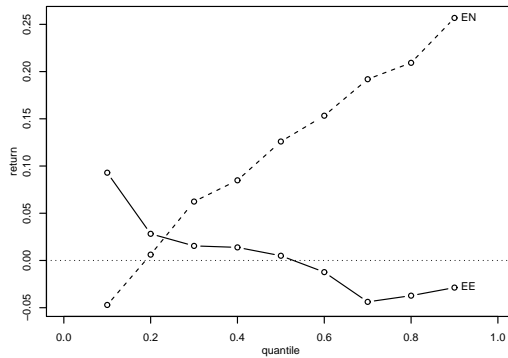


FIGURE 1. ESTIMATED RETURNS TO LANGUAGE SKILLS BY QUANTILES. EE – ESTONIAN, EN – ENGLISH.

Notes: Random effect models, controls variables are the same as in the Table 1.

(column 1) suggest that command of the local language generates returns, that are consistently lower than those for English fluency. Even more, the command of the local language seems to be statistically completely unrelated to income. The point estimate (0.005) is even smaller than those, Chiswick and Miller (2007) find for English in the US for individuals who do not fully use their language skills at work. The estimates, based on PoG data (columns 2 and 3), suggest somewhat higher returns, in the range of 10-20%. Unfortunately, the standard errors are large, rendering these figures statistically not different from zero. In contrast, the corresponding figures for English are all statistically significant at least at the 10% level. To re-conciliate the quantitatively different estimates, based on LFS and PoG data, we also estimate a LFS based model, selecting a subgroup, similar to PoG sample (column 4). These results are quite similar to those of the Estonian PoG sample (column 2).

Next, we look at the association at different quantiles. Due to the sample size, we only use the LFS data below. We provide the results in a graphic form (Figure 1). We use the same specification as in the Table 1 above.

The estimates for both languages reveal a clear but completely different pattern. For Estonian, the estimates are positive but small for lowermost quantiles. Around the median, the parameter is close to zero, and turns negative for larger quantiles. This results in a clear falling trend, where the low wage earners seem to get some positive returns to their command of the local language. However, at median and above, the

TABLE 2—RETURN TO ESTONIAN SKILLS BY INDUSTRY AND OCCUPATION

dependent variable: log wage			
variable	estimate	std dev	pct
primary	-0.177**	0.073	7.5
trade	0.036	0.059	9.4
finance	0.067	0.066	6.5
public adm	0.294***	0.085	3.2
social	-0.004	0.078	3.3
other	-0.055	0.048	18.4
industry	-0.033	0.044	51.7
occ H	0.034	0.050	14.3
occ L	0.039	0.041	74.1
# obs	8639		
R^2	0.609		

Notes: We group the industries as follows: *primary* (ISIC 3 categories A, B, C), *industry* (D, E, F), *trade* (G, H), *finance* (J, K), *public adm* (L), *social* (M, N) and *other* (I, O, P, Q). The occupations were divided as *H* (codes 1,2), *M* (3,4,5,0, the base group) and *L* (6,7,8,9).

Controls variables are the same as in the Table 1.

return is non-existent. The picture for English is strikingly different. We see a steep increasing trend where the command of English is associated with no income premium for the lowest end, and with a sizable income boost of 25% for the upper decile.

Third, we analyze the impact by occupation and industry. We estimate the effect by random effects model including similar controls as in Table 1. We estimate the impact by industries and occupation by introducing two sets of controls: first industry and occupation dummies, and second the corresponding cross effects with the Estonian skills variable (note that we exclude middle occupations as the control group, see explanations in the Table). The estimated cross-effects are reported in the Table 2. The last column (*pct*) in the table also reports what percentage of minority men are employment in those sectors.

The table indicates that there is one clear exception among industries: In public administration, the command of Estonian language is clearly related to substantially larger (about 30%) wage. In the primary sector, the effect is negative, possibly due to selection of more able Estonian-speakers out of that sector. The estimates for the other industries are rather small, and not statistically significant. In particular, customer-oriented sectors, such as trade and finance

TABLE 1—RESULTS FOR LANGUAGE PARAMETERS IN MEAN REGRESSION MODELS.

dependent variable: log wage				
language:	LFS, EE	PoG, EE	PoG, LV	LFS1998, EE
	(1)	(2)	(3)	(4)
Local	0.005 <i>0.019</i>	0.128 <i>0.175</i>	0.209 <i>0.190</i>	0.093 <i>0.122</i>
English	0.144*** <i>0.027</i>	0.415* <i>0.233</i>	0.624*** <i>0.206</i>	0.455** <i>0.189</i>
# obs	8625	328	410	1259
R ²	0.507	0.204	0.276	0.406

Note: Standard errors in italics. Random effects/OLS models with controls for time, age, education, immigrant status, family status, children, and capital city. Statistical significance: * – 10%, *** – 1%. EE denotes results for Estonia, LV for Latvia

indicate now language premium either. However, as public administration employs only little over 3% of non-Estonian men, the number of those who can enjoy the benefits from the language fluency is limited.

Finally, we analyze the ethnic concentration effect. We know that the need and opportunity for language use depends on the presence of speakers of that language in one's social sphere. In ethnic concentration areas these opportunities may be limited. Previous research is ambiguous about the role of the majority language in the minority concentration areas. While Chiswick and Miller (2002) and Zhou and Logan (1989) find English equally important in minority neighborhoods, Hwang, Xi and Cao (2010) find the return to English diminishing in the concentration of immigrants across the US metropolitan areas. We look at the three different regions of Estonia. First, the capital city, the most active economic area where roughly 45% of the population are Russian-speakers. Second, the economically depressed Northeast, where ethnic Russians form the majority (about 85%). And last, the rest of the country, where Russians form about 9% of the population. We conduct analogous analysis by industry and occupation in all three regions separately. The results are given in Table 3.

The table reveals a picture which is not qualitatively different from that of the national level. Only public administration shows clearly positive relationship between command of the official language and income. In the Russian-dominated Northeast we also identify a positive relationship for workers in education and health care sectors while fluent Estonian speakers in the primary sector seem to earn a little less. However, the results fail to indicate a clear positive association between skills of the official language

TABLE 3—RETURN TO ESTONIAN SKILLS BY INDUSTRY AND OCCUPATION, DIFFERENT REGIONS

dependent variable: log wage			
variable	Capital	NE	Rest
primary	0.307 <i>0.256</i>	-0.174* <i>0.104</i>	-0.057 <i>0.163</i>
trade	0.023 <i>0.090</i>	-0.010 <i>0.109</i>	0.094 <i>0.153</i>
finance	0.123 <i>0.094</i>	-0.014 <i>0.129</i>	-0.248 <i>0.217</i>
public adm	0.472*** <i>0.170</i>	0.280*** <i>0.105</i>	0.424* <i>0.237</i>
social	-0.163 <i>0.137</i>	0.293** <i>0.143</i>	-0.149 <i>0.185</i>
other	-0.113 <i>0.071</i>	0.067 <i>0.103</i>	0.063 <i>0.127</i>
industry	-0.053 <i>0.069</i>	0.046 <i>0.075</i>	-0.058 <i>0.122</i>
occ H	0.099 <i>0.083</i>	-0.026 <i>0.086</i>	0.054 <i>0.144</i>
occ L	0.061 <i>0.065</i>	-0.003 <i>0.073</i>	0.050 <i>0.117</i>
# obs	3770	3202	1650
R ²	0.644	0.646	0.366

Notes: Controls include individual characteristics, geographic location, time trend, industry and occupation. Random effects regression. Middle occupations is the control group.

and income in any of the three regions. Neither in case of being a minority, majority, or roughly equally distributed, the ethnic Russian men seem to gain sub-

stantially from the command of the official language.

IV. Discussion

We have shown that the skills of Estonian and Latvian, the sole official languages in the corresponding two Baltic Countries, are not closely related to income premium for the male minority workers. The language command is positively associated with wage only in the lower end of the income scale, and in public administration sector. This is in striking contrast to what we observe for English, and also with some other studies (Berman, Lang and Siniver, 2003).

Since the mid 1990s, Estonia experienced a substantial unexplained income differential, about 10-15%, in favor of the ethnic Estonians (Leping and Toomet, 2008). The current results strongly suggest that deficient language skills are not the main reason behind the minority income lagging behind that of the titular population. We also know that ethnic Russians share broadly similar human capital with the native population. Hence, explaining the lack of language premium as caused by foreign human capital, not valued in the host labor market (Friedberg, 2000), is not convincing. This conclusion is further strengthened by the fact that immigrant status has no statistically significant relationship to the wage. Rather, current results indicate that even when speaking the local language, the members of the minority community find themselves excluded from the better-paying jobs. This conclusion is analogous to Charles and Guryan (2008) and Waisman and Larsen (2009), who show that negative attitudes explain a large part of the income differential.

Although we cannot directly analyze the role of the institutional settings, it seems plausible that tense relations between the main ethnic groups play a role. In a context where the ethnic background is an issue, the implicit social hierarchies in the everyday discourse may create a "tilted" playing field, disfavoring the minority group (Gotsbachner, 2001), and in this way re-enforcing segregation. It is possible that problems of other minorities who are fluent in the language of their host country, such as blacks in the US or Latin-American immigrants in Spain (Bosch, Carnero and Farré, 2010), are related to similar mechanisms.

V. Conclusions

We analyze the Russian-speaking minority in the Baltic States and their wage premium, related to the

command of the official language. We focus on male workers for the period 2000-2010 based on the Estonian Labor Force Survey, and a Latvian and Estonian panel of 1982 high school graduates. We find the premium to be small and mostly not statistically significant. It is positive only in the lowest end of the income scale, and in the public sector. The results are similar for both minority and majority dominated regions. This is in striking contrast with the effect of English, which is related to substantial income premium of at least 15%.

Based on these findings, and previous studies indicating the presence of a sizable wage penalty for the members of the minority group, we argue that the most plausible explanations are related to discrimination and segregation. The members of the minority group find it hard to move toward the upper end of the income discrimination because of discrimination, and rely on less well paid jobs in the largely segregated Russian-speaking workplaces. Despite of labor market participation rate at least as high as that of the titular population for decades, the Russian minority still suffers from the glass-ceiling effect.

REFERENCES

- Albouy, David.** 2008. "The wage gap between Francophones and Anglophones: a Canadian perspective 1970-2000." *Canadian Journal of Economics*, 41(4): 1211-1238.
- Berman, Eli, Kevin Lang, and Erez Siniver.** 2003. "Language-skill complementarity: returns to immigrant language acquisition." *Labour Economics*, 10: 265-290.
- Bleakley, Hoyt, and Aimee Chin.** 2004. "Language skills and earnings: Evidence from childhood immigrants." *Review of Economics and Statistics*, 86(2): 481-496.
- Bosch, Mariano, M. Angeles Carnero, and Lúdia Farré.** 2010. "Information and discrimination in the rental housing market: Evidence from a field experiment." *Regional Science and Urban Economics*, 40(1): 11 - 19.
- Charles, Kerwin Kofi, and Jonathan Guryan.** 2008. "Prejudice and Wages: An Empirical Assessment of Becker's 'The Economics of Discrimination'." *Journal of Political Economy*, 116(5): 773-809.

- Chiswick, Barry R.** 1998. "Hebrew language usage: Determinants and effects on earnings among immigrants in Israel." *Journal of Population Economics*, 11: 253-271.
- Chiswick, Barry R., and Paul W. Miller.** 1995. "The Endogeneity between Language and Earnings: International Analyses." *Journal of Labour Economics*, 13(2): 246-288.
- Chiswick, Barry R., and Paul W. Miller.** 2002. "Immigrant earnings: Language skills, linguistic concentrations and the business cycle." *Journal of Population Economics*, 15: 31-57.
- Chiswick, Barry R., and Paul W. Miller.** 2007. "Matching Language Proficiency to Occupation: The Effect on Immigrants' Earnings." IZA Discussion Paper 2587, P.O.Box 7240, 53072 Bonn, Germany.
- Chiswick, Barry R., and Paul W. Miller.** 2010. "Occupational language requirements and the value of English in the US labor market." *Journal of Population Economics*, 23(1): 353-372.
- Friedberg, Rachel M.** 2000. "You Can't Take It with You? Immigrant Assimilation and the Portability of Human Capital." *Journal of Labor Economics*, 18(2): 221-251.
- Gotsbachner, Emo.** 2001. "Xenophobic normality: The discriminatory impact of habitualized discourse dynamics." *Discourse & Society*, 12(6): 729-759.
- Hwang, Sean-Shong, Juan Xi, and Yue Cao.** 2010. "The conditional relationship between English language proficiency and earnings among US immigrants." *Ethnic and Racial Studies*, 1-28.
- Koenker, Roger.** 2004. "Quantile regression for longitudinal data." *Journal of Multivariate Analysis*, 91: 74-89.
- Koenker, Roger, and Gilbert Bassett, Jr.** 1978. "Regression quantiles." *Econometrica*, 46(1): 33-50.
- Korts, Külliki, and Ragne Kõuts.** 2002. "Inter-ethnic communication in Estonian media (in Estonian)." In *Estonia and Estonians in comparative perspective (in Estonian)*. , ed. Aune Valk, 234-245. Tiigi 78, Tartu 50410, Estonia:Tartu University Press.
- Leping, Kristjan-Olari, and Ott Toomet.** 2008. "Emerging ethnic wage gap: Estonia during political and economic transition." *Journal of Comparative Economics*, 36(4): 599-619.
- Leslie, Derek, and Jannie Lindley.** 2001. "The Impact of Language Ability on Employment and Earnings of Britain's Ethnic Communities." *Economica*, 68(272): 587-606.
- Rooth, Dan-Olof, and Jan Saarela.** 2007. "Native language and Immigrant Labour Market Outcomes: An Alternative Approach to Measuring the Returns for Language Skills." *International Migration & Integration*, 8: 207-221.
- Shields, Michael A., and Stephen Wheatley Price.** 2002. "The English language fluency and occupational success of ethnic minority immigrant men living in English metropolitan areas." *Journal of Population Economics*, 15: 137-160.
- Titma, Mikko, and N. B. Tuma.** 1995. *Path of a Generation: A Comparative Longitudinal Study of Young Adults in the Former Soviet Union*. Stanford, CA:Stanford University Press.
- Waisman, Gisela, and Birthe Larsen.** 2009. "Do attitudes toward immigrants matter?" Stockholm University, Copenhagen School of Business.
- Zhou, Min, and John R. Logan.** 1989. "Returns on Human Capital in Ethnic Enclaves: New York City's Chinatown." *American Sociological Review*, 54(5): 809-820.