TAL TECH

DOCTORAL THESIS

Essays on Language Skills and Labour Market Outcomes

Svetlana Ridala

TALLINNA TEHNIKAÜLIKOOL
TALLINN UNIVERSITY OF TECHNOLOGY
TALLINN 2021

TALLINN UNIVERSITY OF TECHNOLOGY
DOCTORAL THESIS
11/2021

Essays on Language Skills and Labour Market Outcomes

SVETLANA RIDALA

TAL TECH PRESS

TALLINN UNIVERSITY OF TECHNOLOGY
School of Business and Governance
Department of Economics and Finance
This dissertation was accepted for the defence of the degree of Doctor of Philosophy in Economics on 21/01/2021

Supervisors:

Professor Kadri Männasoo, PhD
Department of Economics and Finance
Tallinn University of Technology
Tallinn, Estonia

Associate Professor Ants Aasma, PhD
Department of Economics and Finance
Tallinn University of Technology
Tallinn, Estonia

Lecturer Ott-Siim Toomet, PhD
Information School
University of Washington
Seattle, USA

Opponents:

Prof David G. Bivin
Department of Economics
School of Liberal Arts
Indiana University – Purdue University Indianapolis
Indianapolis, USA

Associate Prof Michiru Nagatsu, PhD
Faculty of Social Sciences
University of Helsinki
Helsinki, Finland

Defence of the thesis: 04/03/2021, Tallinn

Declaration:
Hereby I declare that this doctoral thesis, my original investigation and achievement, submitted for the doctoral degree at Tallinn University of Technology has not been submitted elsewhere for a doctoral or equivalent academic degree.

Svetlana Ridala ----------
signature

European Union
European Regional
Development Fund

Investing
in your future

Copyright: Svetlana Ridala, 2021
ISSN 2585-6898 (publication)
ISBN 978-9949-83-664-2 (publication)
ISSN 2585-6901 (PDF)
ISBN 978-9949-83-665-9 (PDF)
Printed by Koopia Niini & Rauam

TALLINNA TEHNIKAÜLIKOOL
DOKTORITÖÖ
11/2021

Esseed keeleoskusest ja tööturu väljunditest

SVETLANA RIDALA

TAL TECH
KIRJASTUS

Contents

List of Publications

The list of the author's publications, on the basis of which the thesis has been prepared:

I Bormann, S.-K.; Ridala, S.; Toomet, O.-S., 2019. Language Skills in an Ethnically Segmented Labour market: Estonia 1989-2012. *International Journal of Manpower*, vol. 40, no. 2, pp. 304–327. DOI https://doi.org/10.1108/IJM-06-2017-0115 (ETIS 1.1)

II Ridala, S.; Toomet, O.-S., 2019. Wage gap in Ethnically Segmented Labour Market: The Role of Cognitive Skills. *Eastern European Economics*, vol. 57, no. 1, pp. 20–30. DOI https://doi.org/10.1080/00128775.2018.1524713 (ETIS 1.1)

III Ridala, S., 2020. Language skills and labour market returns: A meta-regression analysis. *Language Problems and Language Planning*, vol. 44, no. 2, pp. 200–241. DOI https://doi.org/10.1075/lplp.20012.rid (ETIS 1.1)

Author's contribution to the publications

Contribution to the papers in this thesis are:

I The author of the thesis was responsible for the part *Does the Effect Differ by Age?* and contributed with the co-authors to formulating the research question, describing the theoretical background, and choosing the research methodology.

II The author of the thesis was the corresponding author of the article, conducted the empirical analysis, and wrote the text with the co-author.

III The author of the thesis is the sole author of the article.

Introduction and motivation for the research

The globalisation of the world economy and the transformation of labour markets under the spur of rapid progress in information and communication technologies, automation and artificial intelligence has provoked growing interest in how language skills affect labour market outcomes. Growing global instabilities coupled with an ageing population and a shrinking workforce in Europe and in other advanced economies have triggered expanding migration flows, and these trends call for ethnic and linguistic minorities to be better included and integrated in labour markets.

The literature on economics of language, which has been comprehensively surveyed and analysed by Chiswick (2008), Gazzola and Wickström (2016), and Ginsburgh & Weber (2016 and 2020) among others, has helped to give a better understanding of the relationship between economic returns and language skills, and of the disadvantages and benefits of ethnicity and language. The research on economics of language focuses on how language is related to economic status, the dynamics of the social and economic impact of languages, and language policies and planning, and more recently on game theory and pragmatic aspects (Klamer, 2003; Rubinstein, 2000). Zhang and Grenier (2013) propose the term "language and economics" and define it as: "Language and economics is a branch of economics that uses economic theory, principles and methods to study language and speech acts, which are considered as widespread social and economic phenomena" (p. 219).

Language skills are an integral part of human capital. Knowing languages contributes to the productivity, social capital, and eventual labour market success of an individual. Studies of labour market integration analyse how immigrants are affected by their skills in the local majority language (Bleakley & Chin, 2004; Budría et al., 2017; Chiswick, 1998; Chiswick & Miller, 1995; Dustmann & Fabbri, 2003; Dustmann & van Soest, 2001). There is also an increasing body of literature that studies the economic returns to major global or professional languages, predominantly the returns to skills in English (Azam et al., 2013; Casale & Posel, 2011; Di Paolo & Tansel, 2019; Fabo et.al., 2017; Isphording, 2013; Stöhr, 2015; Wang et al., 2017; Williams, 2011). The literature on multiple local languages in a multilingual society and their roles, interactions and economic returns is quite limited and the existing studies vary widely in their linguistic and socio-economic contexts (Armstrong, 2015; Christofides & Swidinsky, 2010; Cappellari & Di Paolo, 2015; Di Paolo & Raymond, 2012; Williams, 2011; Bormann et al., 2019; Ridala, 2021; Männasoo & Ridala, 2021). The research in the current thesis contributes to this strand of literature by investigating the labour market returns to language skills in Estonia, which is one of the most linguistically diverse countries in Europe.

Estonia is one of the smallest member states of the European Union, and one of the most rapidly growing and transforming. Estonia witnessed an era of soviet domination from 1940–1991, then it regained independence in 1991, and joined the European Union in 2004 and the euro area in 2011. A legacy of the soviet era was extensive immigration from the east, and the Russian-speaking first, second and third-generation minority community is about a quarter of the Estonian population. The homogeneous Russian-speaking community has maintained strong ties to Russia, and it is quite compact and mainly geographically concentrated in the cities and towns on the Northern coast and in the North-East region of Estonia that borders Russia.

Preserving Europe's cultural, ethnic and linguistic diversity is a founding principle in the European Union[1]. Europe has a rich heritage of linguistic and ethnic minorities. The growing flows of immigration, particularly after the escalation of the Syrian civil war, have however posed a challenge for many European countries. Europe and its small countries, most with their own native language, have witnessed rapid growth in linguistically homogeneous and sizable Turkish, Spanish, Arab, Chinese and Russian-speaking communities. This can be seen in the Estonian context, which serves as a good example of a small country with its own native language hosting a sizable, homogeneous linguistic minority from a large neighbouring country, which for Estonia is Russia. This makes the Estonian case highly relevant for understanding the potential dynamics and challenges that European countries may face in the future as immigration grows.

The Linguistic Diversity Index (LDI) puts Estonia in fourth place in Europe after Belgium, Luxembourg and Latvia. Figure 1 shows that the probability that any two people in Estonia selected at random will have different mother tongues is 47%.

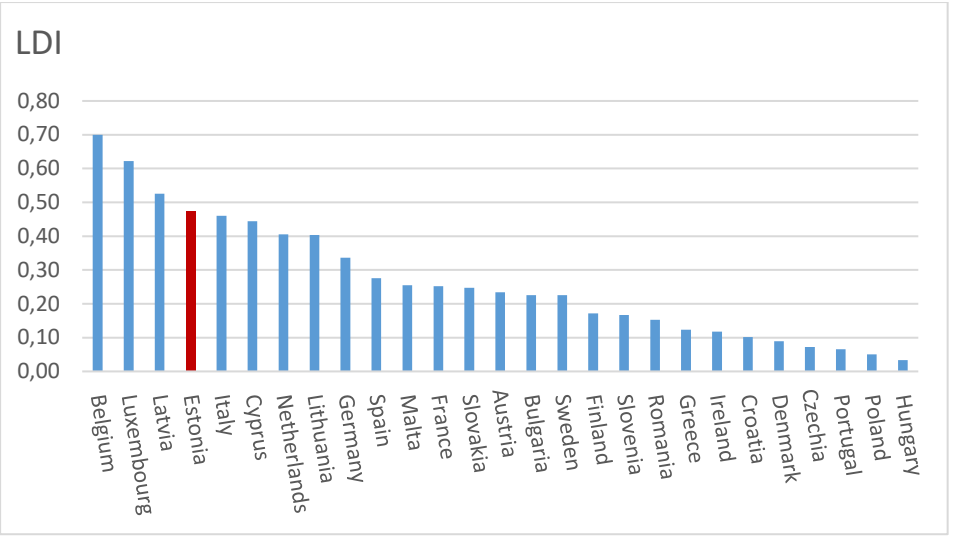


Figure 1: Author's calculations. Linguistic Diversity Index (LDI) for the EU countries. Source: Ethnologue, 2017

Chan (2016) proposed the Power Language Index (PLI) as a ranking of languages by their global influence that incorporates multiple dimensions or opportunities provided by language, including easier travel geographically, and enhanced economic participation, communication and diplomacy. The most influential language in the world is English, which has a PLI score of 0.889, or more than 3.5 times the global influence of Russian, which has a PLI score of 0.244, and 50 times the PLI score of 0.017 for Estonian, which is one of the smallest official languages in the European Union and is spoken only in Estonia (see Figure 2). The globalisation of the world economy and new all-embracing world-wide communication platforms have made English, as May (2016) claims, the 'lingua mundi'.

[1] https://europa.eu/european-union/about-eu/eu-languages_en

Using country and language-specific contexts provides a rich basis for the empirical research. Equally however, heterogeneity in the contexts and designs of the studies can challenge the generalisations found from the line of research.

In consequence the current thesis not only provides primary empirical evidence in Paper I and Paper II, but also incorporates a meta-regression research in Paper III on 61 primary empirical studies with 258 distinct estimates of returns to language. Taking the original empirical studies and the meta-regression analysis, the author proposes a conceptual framework for a holistic understanding of how skills in a second language affect labour market outcomes (see Figure 12 in Section 5).

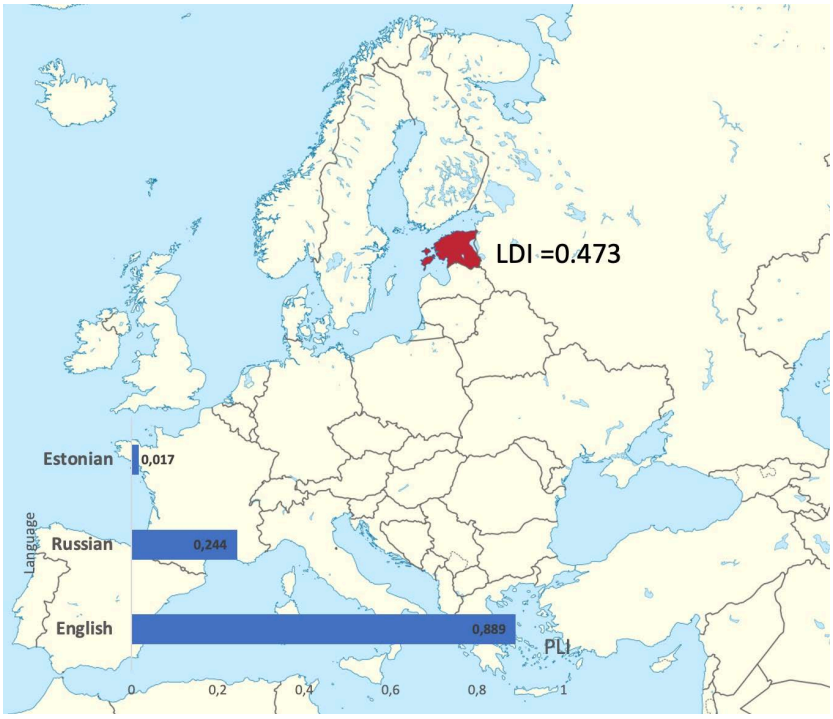


Figure 2: Estonia in Europe with the LDI and PLI indexes

The thesis proceeds with the following structure (see also Graphical outline of the study in the Appendixes). The next section surveys the literature on the theory and empirical evidence of economics of language. Section 2 describes the Estonian context for the study. Section 3 presents the main research hypotheses and contribution of the study. Section 4 explains the research data and methodology and outlines the key elements of the three published papers that underlie this research thesis. Section 5 synthesizes the institutional and socio-linguistic contextual factors that have an important effect on language returns and proposes a holistic concept. Section 6 concludes with final comments that refer to the limitations of the studies and provide ideas for further research. Appendices reprint the three publications of the thesis.

Abbreviations

AES	Adult Education Survey
ELFS	Estonian Labour Force Survey
EU	European Union
IMS	Integration Monitor Survey
LPM	Linear probability model
OECD	Organisation for Economic Co-operation and Development
OLS	Ordinary least squares
PIAAC	Programme for the International Assessment of Adult Competencies
PS-TRE	Problem solving in technology-rich environments
PLI	Power Language Index
LDI	Language Diversity Index
KOFGI	KOF (Center of Economic Research at ETH Zurich) Index of Globalisation

1. Economics of language research

Marschak (1965) was the pioneer of research into economics of language when he introduced economic concepts such as value, utility, costs and benefits into linguistic analysis. A series of articles followed that investigated the relationship between language and economics and highlighted its importance (Bretton, 1976; Carliner, 1981; Chiswick, 1977; Fishman, 1967; Grenier, 1987; Hocevar, 1975; Pool, 1972). Economics of language research intersects closely with the literature on human capital. The first elements of human capital theory were introduced by Walsh (1935), Mincer (1958), and Friedman and Kuznets (1945). Formal human capital theory then emerged in the early 1960s and built theories focusing on schooling, on-the-job training, individual awareness, and health issues, all of which can transform a person, and on migration, which transforms how a person interacts with their environment. The key contributions here were the supplementary volume *Investment in Human Beings* published by the Journal of Political Economy in 1962, the ground-breaking papers by Schultz (1962) and Mincer (1962), and the Gary Becker's study *Human Capital* issued in 1964 (Becker, 1962, 1964). Mincer followed this with his classic book *Schooling, Experience, and Earnings* (Mincer, 1974).

Before human capital theory took a central place, the stress in the literature on economic growth was on the accumulation of physical capital in the neoclassical growth model. Schultz (1962) opened the new discussion and helped to shift this stress towards investment in human skills and knowledge with a major line of new research that followed both the theoretical and empirical strands (Islam, 1995; Mankiw et al., 1992) and had strong implications for the modern literature on economic growth (Lucas, 1988; Razin, 1972; Uzawa, 1965).

Increasing recognition of the economic advantage from language skills (Grin, 1996a, 2003) gave rise to an empirical literature on labour market integration that mainly focused on how the language skills of immigrants as a form of human capital affect their economic outcomes (Carliner, 1981; Grenier, 1987; Grenier & Vaillancourt, 1983; McManus et al., 1983; McManus, 1990; Tainer, 1988). *The Economics of Language*, edited by Lamberton (2002), claims that language is the "oldest and most public of technologies". Ginsburgh and Weber (2020) see language as a social technology.

The research in economics of language has identified the economic benefits that arise from language skills as better employment prospects, higher earnings, more effective cost budgeting, stronger political involvement, and larger social networks, among others (Bleakley & Chin, 2004; Duncan & Mavisakalyan, 2015; Greiner & Vaillancourt, 1983; Grenier, 2015). However, the rapidly growing literature in economics of language is fragmented and heterogeneous (Ginsburgh & Weber, 2020). The work on theoretical frameworks and on empirical measurement of the role of language skills in determining labour income is ongoing. The most commonly used theoretical framework "is derived from the combination of the economics of language perspective with human capital theory developed in education economics" (Grin, 2003).

The literature on the economics of language status has grown fast in recent years and is highly heterogeneous. This thesis attempts to add new knowledge to the empirical strand of the literature on applied economics of language by focusing on labour market returns to language learning. The literature of interest on the nexus between language, earnings and the labour market, and of employment prospects may be broadly divided

into three strands: (1) labour market integration; (2) the labour market effects of globalisation; and (3) the economic value of the second most commonly spoken language.

Another strand of the research into the relationship between language or ethnicity and earnings considers the theory of discrimination, in particular discrimination in labour markets (Lang, 1986). A substantial portion of the literature documents that the labour earnings of ethnic minorities are lower because their qualifications are lower and their command of the local language is deficient (Dustmann & Fabbri, 2003; Dustmann & Theodoropoulos, 2010), but also because communication between ethnically or linguistically dissimilar groups is costly and minorities bear the bulk of these costs (Lang, 1986). Other literature sources point to family background (Black et al., 2006), labour market discrimination (Becker, 2010; Elliott & Lindley, 2008; Leping & Toomet, 2008; Schafgans, 1998) and the asymmetric effects of macroeconomic shocks (Borjas et al., 1996; Bound & Freeman, 1992) as causes of labour market disadvantage for immigrant and minority ethnic populations.

Empirical literature has found that the ethnic wage gap remained present even after a rich set of explanatory factors was controlled for (see Kahanec, 2007; Leping and Toomet, 2008; Miranda and Zhu, 2013a). Examples have been found for black and middle eastern workers in Europe (Blackaby et al., 2005; Carlsson & Rooth, 2011; Clark & Drinkwater, 2009; Lehmer & Ludsteck, 2011; Longhi et al., 2013; Miranda & Zhu, 2013a), or black people in the southern US (Altonji & Blank, 1999; Black et al., 2006). Although lower salaries may, in principle, be compensated for by lower unemployment and higher fringe benefits, the bulk of the evidence suggests that this is not the case.

The empirical literature has also investigated the effects of language on aggregate economic performance, and shows that sharing a common language or wide knowledge of foreign languages promotes bilateral trade (Egger & Lassmann, 2012; Frankel & Rose, 2002; Head & Mayer, 2014; Melitz, 2008; Melitz & Toubal, 2014). Kim et al. (2015) argue that language is a dynamic instrument that reduces transaction costs and so fosters capital allocations between investors, generating a significant and robust relationship between FDI and the absence of language barriers.

Grin (2016) and Ginsburgh and Weber (2020) contribute excellent, comprehensive literature surveys on economics of language, and they identify and structure the different lines of investigation in this heterogeneous literature. The latter survey reviews the impact of language on economic outcomes and public policies and focuses on how language influences human thinking and behaviour, trade, migration, financial markets, language learning, and the returns from it. Gazzola et al. (2016) stress the economics of language literature on language, earnings and the labour market, dividing it further into literature on indigenous labour and literature on migrant labour (Gazzola et al., 2016, pp. 65–75). Zhang and Grenier (2013) suggest that the traditional strand of contemporary research in economics of language could be described as falling roughly into three major categories of "language and economic status (mainly earnings), economic analysis of the dynamic development of languages, and the economic approach to language policy and planning". Grin (2016) confirms this in his "mental map of economics of language" by outlining the category of "economics of language status" within the literature on economics of language and dividing it into a) influence of linguistic variables on economic variables ({L} → {E}), and b) influence of economic variables on linguistic variables ({E} → {L}). All three papers in the current thesis fall under the broader category of literature investigating the relationship between language and economic status with a particular focus on language-related wage premiums.

1.1. Theoretical framework

A theoretical framework proposed by McManus (1990) for the US labour market, which has enclaves of Spanish speakers, can be adapted for the Estonian labour market, which is multilingual and has a large Russian-speaking minority. The supply and demand model of McManus (1990) has two labour markets, one for proficient speakers of a language and the other for individuals with limited skills in the language or languages that are most in demand in the labour market. The labour market for individuals who lack language skills is limited and so there is an excess supply of workers whose language skills are deficient.

The Estonian labour market has demand for multilingual individuals with skills in Estonian, Russian and English. The labour market for monolinguals is limited and the firms offering jobs for monolinguals have lower capital endowment than firms offering jobs for multilingual employees. This generates incentives for monolinguals to learn languages so they can enter the labour market for the multilingual, a market that offers a higher number and wider variety of jobs, and has jobs that are better paid. The wages in the labour market for the multilingual are directly linked with language proficiency or a degree of multilingualism.

Let R_i be the substitution coefficient for the *i*th individual that measures the rate at which that individual can take up a job in the market for multilinguals, where $0 \leq R_i \leq 1$ is proportional to language proficiency and $R_i=0$ denotes monolingualism and $R_i=1$ denotes perfect multilingualism. The equilibrium condition for the wage in the multilingual labour market is:

$$W_i = R_C W_A,$$

where W_i is the *i*th individual's wage in the multilingual labour market, W_A is the wage for a perfectly multilingual worker and R_C is the substitution coefficient of the marginal individual *i*, who is indifferent between the two labour markets.

Figure 3 below shows the excess demand for perfectly multilingual workers, ED, and the excess supply of monolinguals, *ES*. *N* presents the number of individuals entering the multilingual labour market. The curve *ED'* takes into account the reduction in wages from multilingual skills being less than perfect. First, potential entrants into the multilingual labour market are ordered so that *R*, the premium for multilingual skills, is a decreasing function of *N*, the number choosing to enter the labour market for multilinguals. Then the *R(N)* function is multiplied by the excess demand, *ED*, for multilinguals measured in wage terms. Thus,

$$ED' = R(N)ED.$$

By equating ES and ED', it is possible to determine the equilibrium level of entry N_1 into the multilingual labour market and the equilibrium wages, W_A and W_i, for the perfect multilingual worker and worker *i* respectively. The difference between the wages of a perfect multilingual worker, W_A, and those of worker *i,* W_i, is:

$$P = W_A - W_i = W_A\big(1 - R(N_1)\big).$$

From this, the wage differential, P, decreases as the language proficiency of individual *i* increases, and it increases together with the number of individuals entering the labour market with the same level of skills.

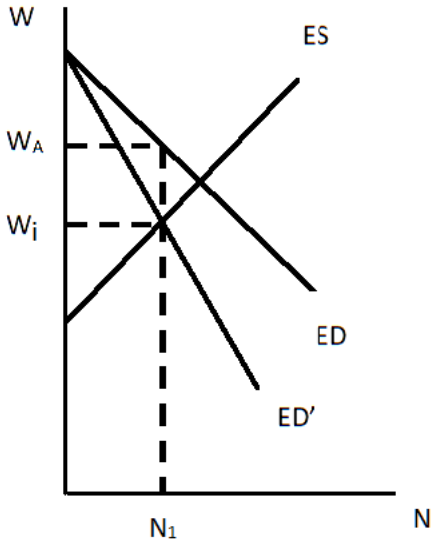


Figure 3: Labour market equilibrium based on the underlying theoretical framework by McManus (1990)

The straightforward supply-demand model of McManus (1990) does not account for labour market frictions such as job search or non-convexities in the labour supply, which are aspects that may be associated with language skills or ethnic identity.

1.2. Literature on labour market integration

Labour mobility has increased over recent decades and migration flows to Europe have exploded. This calls for a better understanding of how to integrate the newly-arrived individuals into the society of the host country[2]. Recent research shows that smoothly assimilating migrants and integrating them with the labour force of the host country produces economic benefits such as higher productivity and lower public spending (Aiyar et al., 2016; Jaumotte et al., 2016). Chiswick (1978) opened a strand in the literature on the labour market that investigates the benefits for immigrants of knowing the host country's language (Carliner, 1981; Reimers, 1983, 1985; Bleakley & Chin, 2004; Budría et al., 2017; Budría & Swedberg, 2012; Chiswick, 1998; Chiswick & Miller, 1995; Dustmann & Fabbri, 2003; Dustmann & van Soest, 2001). Most studies in this line of the literature find positive returns to proficiency in the host country language, though the study results vary because of differences in the sample data, the estimation techniques, and the specific research focuses.

[2] There are several concepts of migrant adaptation, which have varied across time. The concept of assimilation is more commonly used in the USA and the concept of integration is more common in Europe (Borooah & Mangan, 2009; Modood, 2005)

The bulk of the studies on immigrant labour are on the US labour market and show that knowing English is associated with a wage premium of up to 59% (Bleakley & Chin, 2004; Bratsberg & Ragan, 2002; Carnevale et al., 2001; Chiswick & Larsen, 2015; Chiswick & Miller, 2002a, 2010; Gonzalez, 2005; Hall & Farkas, 2008; Hellerstein & Neumark, 2003; Isphording & Sinning, 2012; Kim, 2003; Levanon, 2014; Lewis, 2011; Trejo, 2001; Zhen, 2015). The return to English for immigrants in the UK is a wage premium of up to 46% (Dustmann & Fabbri, 2003; Leslie & Lindley, 2001; Lindley, 2002; Miranda & Yu Zhu, 2013a; Miranda & Yu Zhu, 2013b). Guven and Islam (2015) find that English proficiency produces gains for the wages and career outcomes of immigrants in Australia, particularly for adults who moved to the country in their childhood.

Chiswick and Repetto (2001) find a wage premium for immigrants with command of Hebrew in Israel. Berman et al. (2003) study how language acquisition affects the growth in income of Soviet immigrants in Israel and find that language complements high-skill occupations. They suggest that the positive estimated returns to language acquisition in cross-sections may suffer from ability bias in low-skilled occupations. Cohen-Goldner and Eckstein (2008) estimate that the increase in wages for Soviet immigrants in Israel was about 1.16% in 1992–1995 during the first five years after arrival because of the availability of vocational training programmes provided by the government.

Dustmann and van Soest (2001) use data from 1984–1993 and report German language premiums of 0.9–7.3% estimated with ordered probit for immigrant workers in Germany. OLS estimates on the same data gave a wage premium of 5% for men and 4% for women, whereas the IV estimates reached a wage premium of 14% for men and 12% for women (Dustmann & van Soest, 2002).

Yao and van Ours (2015) study the labour market performance of immigrants in the Netherlands in 2008-2014 and find that deficiencies in language skills lower the hourly wages of female workers, but have no significant effect on the wages of male workers. Isphording (2013) finds no compelling evidence for a host country language premium for immigrants in Spain in 2005. Chiswick and Miller (2002b) and Hwang et al. (2010) note that large language enclaves may relieve the issue of deficiency in the host country language to some degree by generating employment opportunities within the immigrant language enclaves. Bormann et al. (2019) and Toomet (2011) find no evidence of a host country language premium for Russian minority men in Estonia, though there is a positive host country language premium for Russian-speaking women. The dominant evidence on language enclaves suggests, however, that knowing the language of the host country is an important factor in the labour market of the host country (Clark & Drinkwater, 2000; Zhou & Logan, 1989). A number of studies (Aldashev et al., 2009; Dustmann & Fabbri, 2003; Shields & Price, 2002) also find that immigrants who are fluent in the host country language are less likely to become unemployed.

1.3. Literature on labour market globalisation

Globalisation has occurred at an unprecedented magnitude since the end of the twentieth century, and has had an enormous role in shaping human activity across the world. Globalisation is associated with increasing multi-ethnicity and multiculturalism as expressed in the process of people, companies, and governments interacting and integrating around the world. Globalisation has been strongly driven by new technologies emerging and entering into various forms of human activity to create network societies, global platforms and markets for goods and capital (Castells, 1996).

In a similar manner, labour markets have been very much affected by globalisation, and this has had far-reaching welfare consequences.

The globalisation of labour markets has increased the importance of international communication and a common language or 'lingua mundi' (Angouri, 2014; May, 2016). English has emerged as the 'lingua mundi' and its influence has been growing. English is the main working language of many international organisations, such as the United Nations, the European Union institutions, the OECD (Organisation for Economic Co-operation and Development), the BSPC (Baltic Sea Parliamentary Conference), ASEAN (Association of Southeast Asian Nations), SEAMEO (Southeast Asian Ministers of Education Organization) and several others (Duchêne, 2008). English also plays a major role in virtual communication as a language of computer programs and social media.

Mufwene (2010), Gil (2010), Gazzola and Grin (2013), and Ginsburgh and Moreno-Ternero (2018) stress that English does not rule the scene alone in Europe, and that the role of German and French is not to be underestimated and is growing in the European Union after Brexit.

The importance of the leading business languages has led to an empirical literature that estimates the return to knowing these languages, most often as the English skills premium in non-English speaking countries. Ginsburgh and Prieto-Rodriguez (2011) estimated the returns to language proficiency of male workers in 1994–2001 in nine EU countries: Portugal, Finland, Austria, France, Greece, Germany, Spain, Denmark and Italy. They found that English is the dominant 'lingua franca' in Northern Europe (Austria, Denmark, Finland and Germany), with little competition from other foreign languages. In Southern Europe (France, Greece, Italy, Portugal and Spain) however, there is also an earnings premium for foreign languages other than English. Ginsburgh and Prieto Rodriguez (2013) show for Denmark, Finland, France, Germany, Italy and Spain that the reward from using English at work is in general higher for men than for women. They investigate whether this result has a link with the gender wage discrimination in Southern European countries (Spain, France and Italy) and detect a glass ceiling for female workers in Spain and Italy, and to some extent in Denmark. Grin and Sfreddo (1998) use OLS to estimate a 10% English skill premium for male workers in Switzerland. Klein (2004) finds a 2% premium for male workers and a 0.3% premium for female workers in Luxembourg who possess English skills. Williams (2011) reports a 5–30% English skill premium across the non-English speaking EU-15 countries.

Stöhr (2015) looks at German data from 2000 and estimates the wage premium of occupations that require expert-level English, finding a 12% return to English for native Germans and a 26 % return for immigrant workers. Drydakis (2012) reports a low English language premium of 1% for Romanians working in Greece in 2009–2010. Isphording (2013) shows that the English language premium for immigrants in Spain varies between 9% and 47%, and this variation depends on the estimator selected, gender and occupational sub-sample. Cohen-Goldner and Eckstein (2008) also demonstrate that occupation matters and that the return to English for soviet immigrants in Israel in 1992–1995 was 14% in white collar occupations, whereas the return was negative at -4% in blue collar occupations.

There are only very few studies that investigate the English language wage premium in post-communist European countries that have experienced a rapid transition to a market economy and economic integration with the rest of the world (Adamchik et al., 2019a). Kroncke and Smith (1999) use OLS to estimate the returns to English in the Estonian labour market and they find a premium for native Estonian men and for

Russian-speaking minority men in the range of 4–5% in 1989 and 20% in 1994. Toomet (2011) shows that the Russian minority men in Estonia and in Latvia prefer to acquire skills in English rather than in the local state languages Estonian and Latvian. The English language premium for Russian-speaking men in Estonia was 14% according to the Estonian Labour Force Survey 2000–2010. In Latvia, the return to English was as high as 62% according to the Paths of a Generation (PoG) panel.

Fabo et al. (2017) use online job portal data and assess the high economic value of English skills in Czechia, Hungary, Poland and Slovakia. For Poland, Adamchik et al. (2019a) estimate the return to English for the native-born population to be 50% for women with 'good' or 'very good' English and 60% for men over the wages of those with 'no English' for the period 2013 to 2017 using the IV approach, and Adamchik et al. (2019b) estimate a return of 32% for women and 36% for men with 'advanced' English for 2014 employing an OLS regression. Results for Hungarian youth found using the IV approach showed that speaking English increased wages by 6% and speaking German did so by 4% in 1998–1999 (Galasi, 2003).

Research has also found positive returns to English skills in non-European, non-English speaking countries. Di Paolo and Tansel (2015, 2019) report positive returns to English of between 11% and 25% in Turkey. Levinsohn (2007), Cornwell and Inder (2008) find an English premium of 18–40% in South Africa, while Wang et al. (2017) estimate an earnings effect of 7–19% for Chinese women and one of 7–12% for Chinese men.

Chiswick and Miller (2016), Di Paolo and Tansel (2015), Isphording (2013), Saiz and Zoido (2005), and Williams (2011) have studied the role and economic impact of other major international business languages beyond English. Williams (2011) shows that the effect varies considerably across countries, ranging from non-existent to 25% for Italian, 32% for French, and 29% for German, and from 16% to 27% for Spanish. Di Paolo and Tansel (2015) show an earnings premium of 19% from knowing Russian in Turkey, a 6% wage premium for German, and no premium for French or Bulgarian.

1.4. Economics of language studies in bilingual and multilingual contexts

There is only a relatively small number of studies on the role and economic value of skills in the second most commonly spoken language in multilingual countries with large minority languages or a second official language. After the collapse of the Soviet Union, the communist states of Eastern Europe and the states of the former Soviet Union regained their independence and reinstated their native languages as titular national languages. See Bormann et al. (2019) for the Estonian case. Independence movements have also raised issues of language and ethnicity in other parts of Europe including Catalonia (Di Paolo & Raymond, 2012; Cappellari & Di Paolo, 2015), Belgium and Luxembourg (Williams, 2011).

Di Paolo and Raymond (2012) find that the return to proficiency in Catalan is virtually zero for public sector workers, but can add a premium of up to 5-13% of the monthly salary in the private sector. Cappellari and Di Paolo (2015) estimate a sizable Catalan language premium of 22% for non-native Catalan speakers in Spain in 2006–2011. Using the 1991 and 1996 censuses in Catalonia, Rendon (2007) estimates that individuals who read and speak Catalan have a 3–5% higher probability of being in employment, and that the probability is 2–6% higher for those who can also write in Catalan.

Williams (2011) finds that workers in Belgium and Luxembourg who speak a second language at work earn more. The language premium in Belgium from knowing German was 20% and that from Dutch was 16%. In Luxembourg, knowing French brought a language premium of 28%, and knowing German added 29%.

Drinkwater and O'Leary (1997) report that bilingual Welsh and English speakers have lower unemployment in Wales than English monolinguals. Saarela and Finnäs (2003) show that the unemployment rate of the Swedish-speaking minority in Finland is 30% lower than that of the Finnish-speaking majority.

Grin and Sfreddo (1998) find no evidence of an income differential for German, French or Italian-speaking women in Switzerland. They do, however, find a negative association between Italian as a mother tongue and wages for men residing in non-Italian speaking regions in Switzerland.

Grin (1994, 1996b) refers to the "empirical Canadian tradition", which is a literature that studies the relationship between the two official languages of Canada and investigates the wage differentials of Anglophones and Francophones (Armstrong, 2015; Breton, 1964, 1978; Christofides & Swidinsky, 2010; Raynauld & Marion, 1972; Vaillancourt, 1980). Christofides and Swidinsky (2010) report substantial, statistically significant rewards to English skills of about 14% for female workers and 13% for male workers in Quebec. In 2001, no language rewards for skills in French were found in the rest of Canada. Armstrong (2015) uses a model where wages depend both on the worker's innate productive ability and on their language skills, and demonstrates a significant bilingualism premium for native French speakers both inside and outside the province of Quebec.

2. The Estonian case: context and trends

Background information on population trends, linguistic trends, labour market structure and education policy is crucial for understanding the context of empirical results for how language skills affect labour market outcomes and for making good judgements on them. Estonia has been in transition for nearly 30 years and major political, demographic and socio-economic transformations have shaped its population, economy and labour market over that period. The most evident linguistic trends in Estonia over the three decades of independence may be summarised as follows:

- The Russian-speaking population in Estonia, though decreasing, remains significant both in size and in economic impact. The recent westward immigration from Ukraine and Belarus may counteract the declining trend and increase the Russian-speaking population as a share of the whole;
- An increasing share of Russian speakers are becoming local citizens and citizens of the EU, which opens up better opportunities for them in the local and EU labour markets;
- The ageing of the population and other de-population trends, stemming from a natural birth rate below replacement level and from emigration, are increasing demand for labour and so improving the labour market outlook for both Estonian-speaking and Russian-speaking jobseekers;
- The local-language skills of Russian speakers are constantly improving and this is particularly evident for the younger generations;
- The birth cohort remains a very significant determinant in explaining disparities in labour market outcomes driven by language and ethnicity, as the disadvantages are substantially more prevalent among the older generations and much less so among the younger generations.

2.1 Population trends

Before the Second World War, Estonia had a highly homogeneous ethnic composition, as about 88% of the total population were native Estonians. After the war, the Russian-speaking population grew rapidly through migration from Russia and other parts of the Soviet Union. The share of Russian speakers in the population in Estonia rose to 38.5% by 1989 and the share of ethnic Estonians in the total population declined to 61.5% (Tammaru, 1997). Russian gained linguistic dominance in inter-ethnic communication and as an official language by the 1970s, though Estonian preserved its status in culture, education, media and science, and the country had two de facto official languages. The soviet immigrants largely settled in the capital city Tallinn and in a mineral-rich industrial area in the north-east of the country and this led to a high concentration of Russian-speaking minority communities, creating a spatial ethnic segregation that has been very persistent since then. See also Tammaru and Kulu (2003).

After the collapse of the Soviet Union in 1991, Estonia experienced substantial depopulation, partly following the withdrawal of soviet army officials and their families, but predominantly because natural population growth was below the replacement level during the early years of the transition (see the evolution of the workforce in Estonia during the pre-EU accession period by Hazans, 2007). This decline in the population slowed down in 2001–2002 in advance of EU accession, but it re-accelerated in 2004–2005

after the EU enlargement. Accession to the European Union in 2004 led to rapid changes in the Estonian economy, including changes in the labour market (Veebel & Loik, 2012). The early years of EU membership brought major structural changes and rapid economic growth that increased demand for labour and wages. Outward migration of the working-age population to high-income member states of the EU aggravated the labour shortage and helped bring unemployment down. Figure 4 below shows that the Estonian population was decreasing throughout the transition period, but the decline slowed down after the recovery from the Global Financial Crisis of 2008-2009, and the trend has since reversed, with the country's population increasing over the past five years because of positive net migration (Figure 4).

Nevertheless, the low birth rate and emigration pulled the ageing Estonian population down to 1.32 million people by the beginning of 2018, which was 82,117 fewer than in 2000. The share of the population aged 65 or older rose from 13.6% in 2000 to 19.8% in 2019, and the percentage of working-age individuals aged 15–64 dropped from 67.3% in 2000 to 63.8%. These demographic trends have left a strong imprint on labour markets over the recent decades.

According to Statistics Estonia, the share of ethnic Estonians increased sharply in the first decades of transition from 61.5% in 1989 to 67.9% in 2000, though this trend has since slowed and the share had risen only to 68.4% by 2019. The opposite trend has been seen for Russian speakers, whose proportion dropped from 30.3% before the transition to 25.6% afterwards in 2000, with the most recent statistics from 2019 reporting the share of Russian speakers in Estonia at 24.7%.

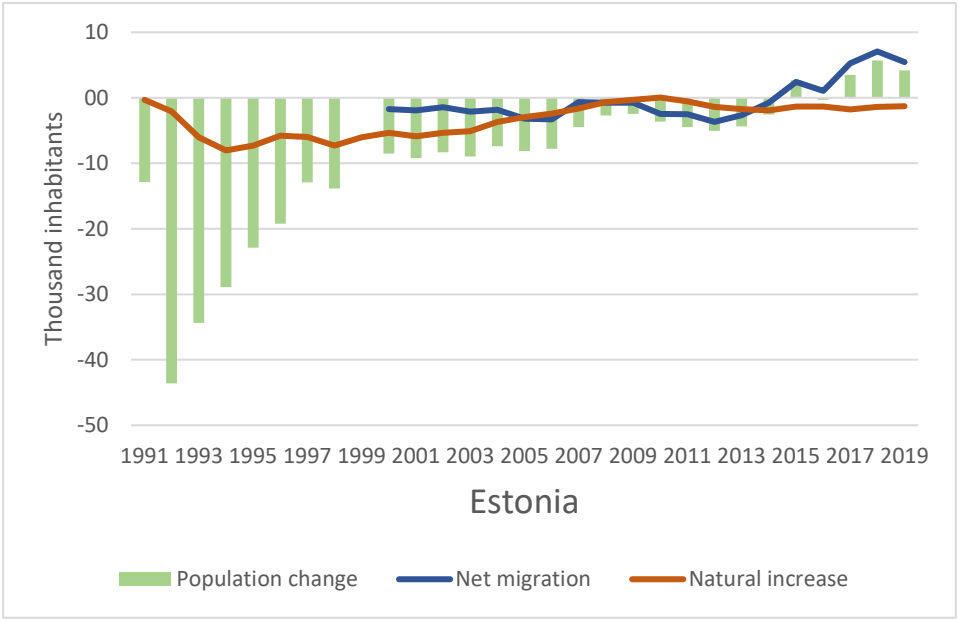


Figure 4: Population change in Estonia (in thousand inhabitants).
Source: Statistics Estonia

The political dispute about citizenship policy has also had significant implications for language and labour market integration. The conditions and requirements for obtaining Estonian citizenship have been debated heatedly for more than two decades. For more

detailed information about the legal and policy positions, and the conditions and requirements, and procedural aspects of obtaining citizenship see Pajumets (2019). The share of non-citizens and foreign citizens, mainly citizens of Russia, who are predominantly Russian-speaking and reside in Estonia has dropped constantly. There were 256 non-citizens or foreign citizens in 2000, but this had dropped to 199.2 thousand in the 2019 statistics, of whom 72.4 thousand had non-defined citizenship[3]. This trend has mostly occurred as Russian-speaking residents have gradually been integrated into Estonian society and are increasingly motivated to pursue not only Estonian but also EU citizenship. See also Vihalemm and Hogan-Brun (2013) for their survey of the challenges and policies of integrating language communities in Estonia.

2.2 Linguistic trends

The soviet era in Estonia resulted in one-sided bilingualism, in that the native Estonian population became fluent in Russian, while the immigrants remained monolingual and expected the local population to use Russian (Vihalemm & Hogan-Brun, 2013). Russian dominated as an official language and as the main language at work in sectors and jobs that were dominated by Russian-speaking workers.

Conscription to the soviet army was a universal military obligation and was mandatory for all male citizens with only a few exceptions. The 24-month service took Estonian men to other parts of the Soviet Union, where there was mostly no alternative to communication in Russian. Consequently Estonian men had better command of the Russian language than Estonian women.

There was language segregation in education and in the media, with Russian speakers enrolling their children in schools that taught in Russian and retaining their preference for media in Russian and originating from Russian providers (Mägi, 2018; Tammaru, 1999). Unlike in many other states of the Soviet Union, Estonians remained loyal to schools with teaching in Estonian and kept following media in Estonian, and they retained the good reputation of Estonian ethnic, cultural and linguistic identity (Kaljund, 2018).

The collapse of the Soviet Union dramatically changed the situation. Even before the restoration of the independence of Estonia in 1989, the first Language Acts defined the official status of the Estonian language. The Language Law of 1995 removed the official status of Russian as the second language of communication in Estonia. The constitution of Estonia set preserving the Estonian language and culture as a national priority. The status of the Estonian language was also strengthened by other reforms in education (OECD, 2018), in public administration (Reinholde, 2004) and in labour market policy[4]. With accession to the EU in 2004, the Estonian language gained new status as an official EU language.

Estonians quickly took back the leading position in society, while the Russian-speaking population found it increasingly challenging to define its identity and status in independent Estonia (Brubaker, 1996; Pettai, 2006). Leppik (2020) refers to “the segmented integration and mediated transnationalism” that comes from the division of political power and ethnic disparities in income and in opportunities for education, with these discrepancies being particularly severe in the early 1990s (Leping & Toomet, 2008; Lindemann & Saar, 2011; Ridala & Toomet, 2019). The language segregation has had a strong impact on

[3] https://www.siseministeerium.ee/et/siseturvalisuse-valdkond/kodakondsus-ja-ranne

[4] New Trade Union Act, amendments to the Labour Law, Posted Workers Enforcement Directive, and Employment Contracts Act, “Incapacity to work” reform, Aliens Act in Estonia (ETUI, 2017).

ethnic occupational segregation (Leppik & Vihalemm, 2015), with Russian speakers over-represented in industrial blue-collar jobs and particularly in the mining industry in the north-east of Estonia (Helemäe & Saar, 2015).

Vihalemm and Hogan-Brun (2013) note that the rapid strengthening of the status of the Estonian language caused a reversal in asymmetrical bilingualism, whereby the native Estonian youth became less and less fluent in Russian while the Russian-speaking youth started rapidly to improve their skills in Estonian. The first Estonian State Integration Programme in 2000–2007 stressed learning Estonian as the main measure for ethnic integration, but the next State Integration Programme in 2008–2013 augmented the set of integration measures with objectives for civic integration and participation in policy. Vihalemm and Hogan-Brun (2013) also point out that proficiency in Estonian is largely related to the social status of the Russian-speaking population and their self-confidence about their own economic prospects.

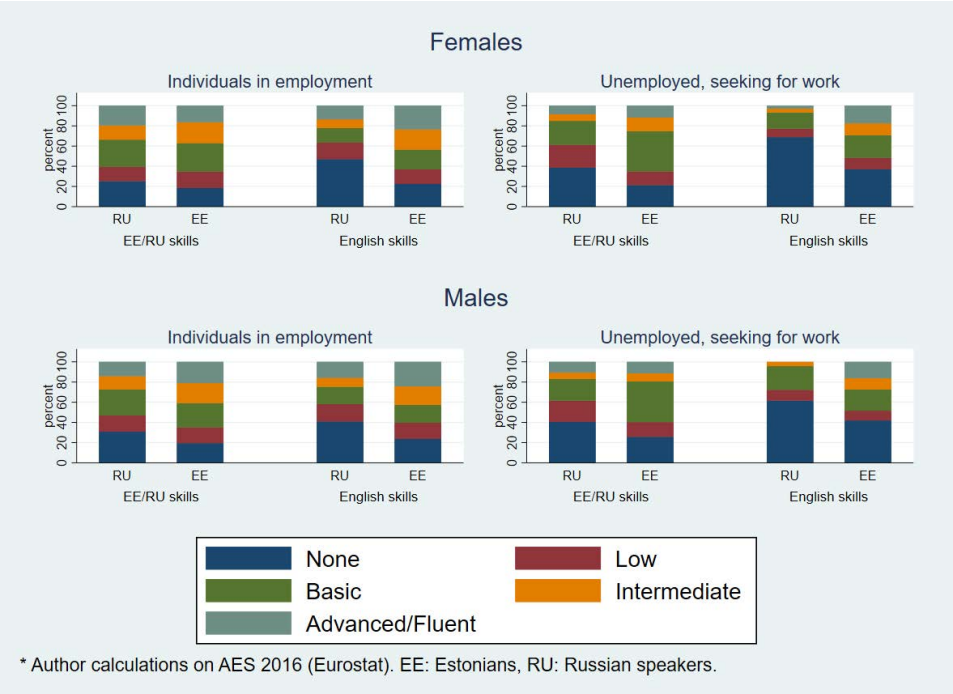


Figure 5: Language skills in Estonian, Russian and in English. Breakdown by males and females and by employment status.
Source: Adult Education Survey 2016, Eurostat

The transition from the soviet command economy to a western market economy was successful, and the rapid reorientation towards competitive European and global markets demanded new skills. Fluency in English became a highly valued skill, and one that was scarce at the beginning of the transition. Russian speakers faced the dilemma of whether to learn Estonian or English, as there were good incentives to prefer to learn English for both the local labour market (Toomet, 2011; Vihalemm & Hogan-Brun, 2013) and the labour markets in Western countries (Aptekar, 2009).

Figure 5 shows the distribution of the language skills of men and women who are native speakers of Estonian and Russian in the other local language and in English,

grouped by their labour market status in 2016. Even 25 years of independence had not reversed the bilingualism asymmetry, as Estonians on average had better skills in Russian than Russian speakers had in Estonian, and this was particularly evident for men and for individuals who were unemployed. Estonians also report consistently higher skills in English, across both gender and labour market status. Evidently the higher degree of multilingualism among Estonians than among Russian speakers not only comes from the country's historical background, but is also determined by the identity of Estonians as a small nation, an identity that holds knowledge of foreign languages in high regard, together with other skills that improve an individual's competitiveness internationally.

2.3 Labour market trends

Despite the challenging demographic trends, the labour force in Estonia has remained relatively stable since 1997, with the drop in population compensated for by a significant and consistent increase in the share of the population that is active. Figure 6 below presents the trends in labour market participation and in unemployment by gender and citizenship in 1997-2019. From the beginning of the 21st century, the share of females who are inactive in the labour market has decreased by a third. The growth in labour immigration over recent years has also contributed to a modest and steady increase in the size of the workforce.

By the beginning of the new millennium, Estonia had overcome the largest structural economic problems of the early transition period and unemployment was consistently dropping except in the years of the Global Financial Crisis and its aftermath. The repercussions of the unfolding economic crisis triggered by Covid-19 will be seen in the near future. A fertility rate below replacement level and sustained ageing of the population imply, however, that the labour supply within Estonia will shrink further if it is not counterbalanced by an inflow of immigrant workers.

The average unemployment rate of non-Estonian natives has been consistently higher than that of Estonian natives and it remains so, though the gap shrinks slightly at times when unemployment is low. The labour market disparities between the ethnic groups in Estonia are still noticeable and warrant further consideration, research and commensurate policy action.

Several of the studies on the Estonian labour market have focused on wage determinants (Eamets & Tiwari, 2019; Ferraro et al., 2018; Masso & Krillo, 2010; Rõõm, 2003), gender differences in wealth (Meriküll et al., 2020), the wage gap between genders (Meriküll & Mõtsmees, 2017; Masso et al., 2020), and ethnic groups (Leping & Toomet, 2008; Ridala & Toomet, 2019). Since regaining independence, Estonia has witnessed a sizeable gender pay gap and ranks first in the EU as having the highest gross gender pay gap of 16%[5].

[5] https://ec.europa.eu/info/policies/justice-and-fundamental-rights/gender-equality/equal-pay/gender-pay-gap-situation-eu_en

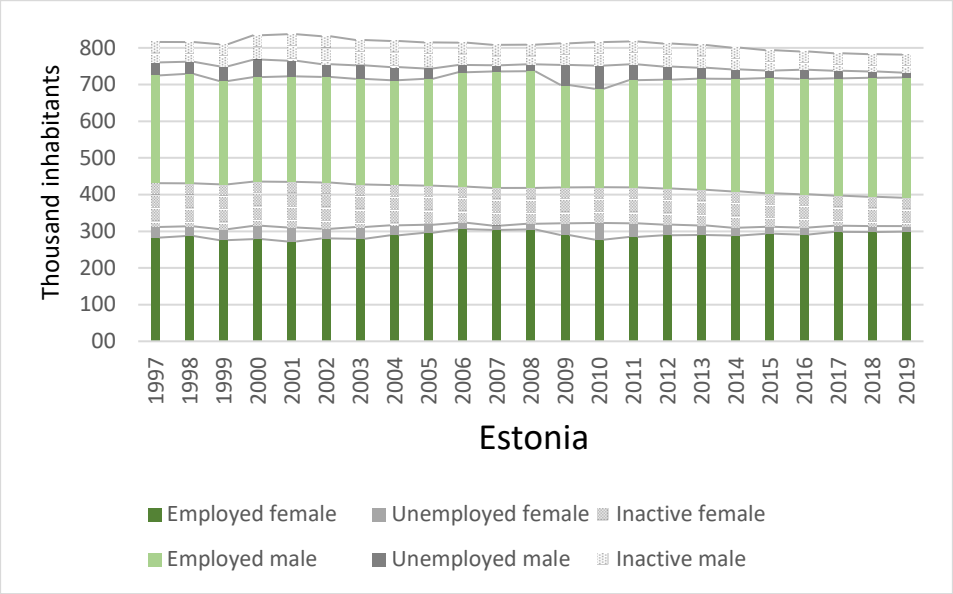


Figure 6: Working-age (20–64) males and females in Estonia and Latvia by employment status (thousands) for the years 1997–2019.
Source: Eurostat

The ethnic wage gap has been studied by Leping and Toomet (2008), who found this gap widened in Estonia among working-age males aged 20–60 in favour of the Estonian-speaking population in the period 1989–2005. Their analysis on wage gap differentials controlled for language skills, education, industry, and occupation. They claim that the main factors causing the unexplained wage gap include different returns to education for different ethnicities and better wages in the capital city. Additionally, Ridala and Toomet (2019) show that the three PIAAC cognitive skills in combination reduced the estimated unexplained wage differential between Estonian-speaking and Russian-speaking workers by approximately 20%, but did not manage to close the gap fully, so it still remains unexplained and considerable in size.

The wage inequalities may to a considerable extent stem from ethnic occupational segregation. The most common measure of occupational segregation is the Index of Dissimilarity (ID) by Duncan and Duncan (1955), which allows the concentration across occupational categories and occupational levels to be analysed in terms of ethnic groups. The Duncan segregation index has the form:

$$ID = \frac{1}{2}\sum_{i=1}^{n}\left|\frac{N_i}{N} - \frac{R_i}{R}\right|, \qquad i = 1,2,\dots,n$$

where n is the number of occupations, N_i and R_i are the number of native workers and Russian speakers respectively in occupation i, and N and R are the total numbers of native Estonians and Russian speakers in employment. ID ranges from 0 reflecting perfect ethnic integration across occupations to 1 indicating complete ethnic segregation within the workforce.

The Duncan index for Estonia (see Figure 7) used to assess the extent of horizontal segregation by occupation is calculated for eight occupational levels: (1) managers, (2) professionals, (3) technicians and associate professionals, (4) clerical support workers, (5) service and sales workers, (6) craft and related trades workers, (7) plant and machine operators, and assemblers, and (8) elementary occupations. Missing data and gaps mean that the calculations exclude armed forces occupations and skilled agricultural, forestry and fishery occupations. Overall the Duncan segregation index for 2000–2019 shows no clear trend, but rather a mean-reverting pattern (see Figure 7).

According to Duncan segregation index from 2019, about 17% of Russian speakers would have to change their occupations to achieve zero segregation or perfect ethnic integration. Occupational segregation is lowest among clerical support workers, and service and sales workers, and it is also relatively modest in technical and associate professionals. The greatest segregation is found among professionals, then among craft and related trade workers and elementary occupations, for whom it has held relatively stable. Segregation is considerable for managers, though a recent trend of de-segregation is becoming notable. Ethnic segregation of plant and machine operators decreased sharply in 2001–2004, but this trend has reversed.

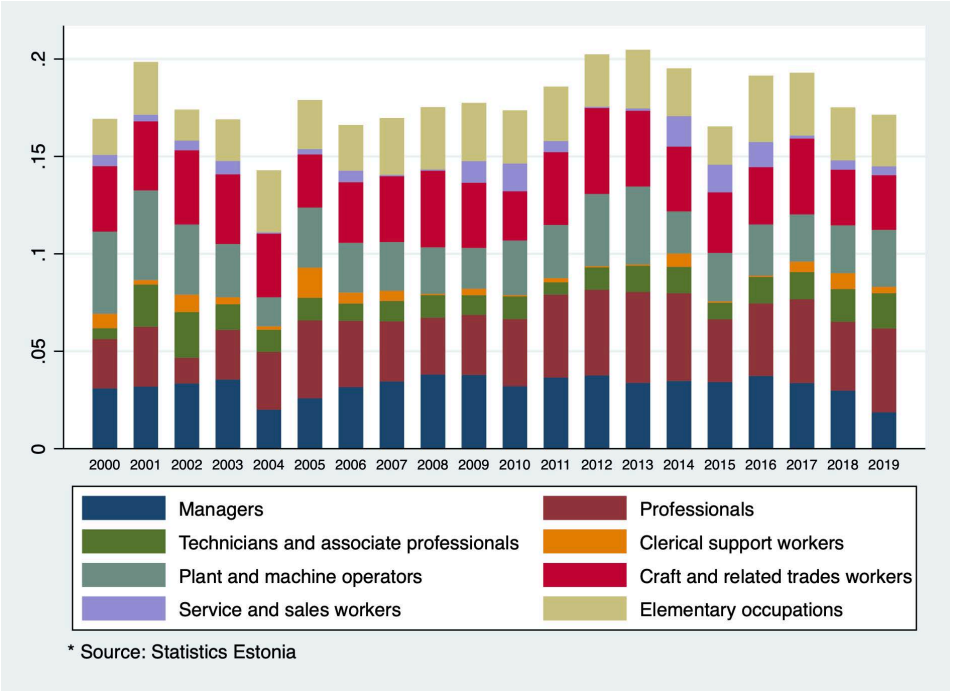


Figure 7: Structure and dynamics of the Duncan Occupational Segregation Index by nationality. Groups: Estonian native speakers and Russian speakers for the years 2000–2019

The occupational structure for 2019, presented in Figure 8, shows that the occupations dominated by Estonians remain the white-collar ones of managers, professionals and technicians, and associate professionals. Russian speakers meanwhile are overrepresented in the blue-collar occupations of elementary occupations, plant and machine operators, and assemblers, and among craft and related trades workers. This situation has been discussed by Bormann et al. (2019) and Ridala and Toomet (2019),

as segregation is one of the possible disadvantages for the Russian speakers and it was found that highly educated workers having negative attitudes toward Russians (Korts, 2009) may not be the issue.

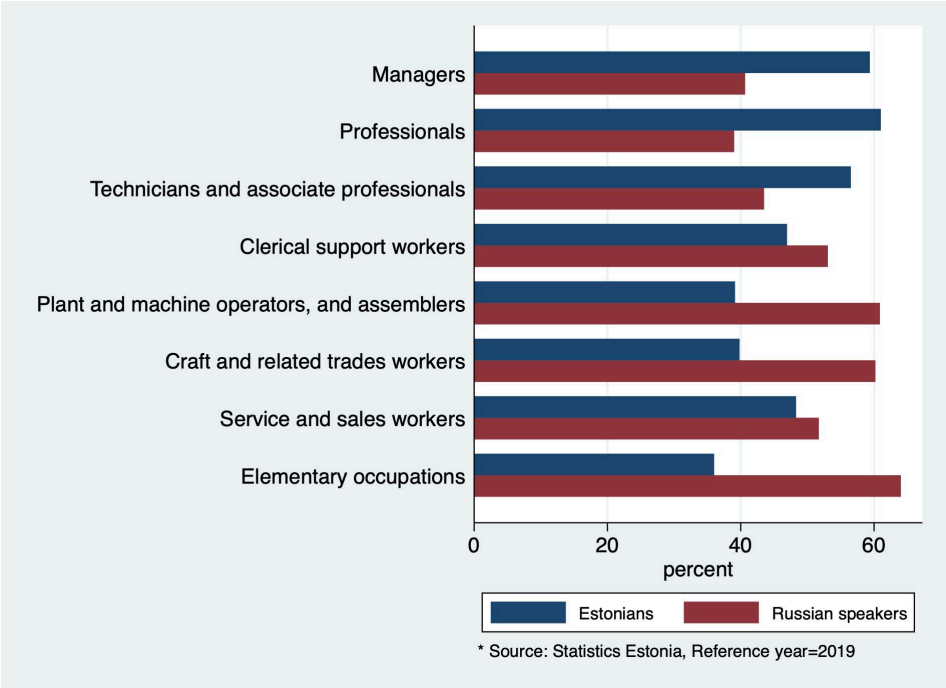


Figure 8: Occupational structure by nationality. Groups: Estonian native speakers and Russian speakers

2.4 Trends in education

The structural changes before and after accession to the EU were coupled with a strong increase in demand for qualified labour. In response, the share of employees with a tertiary degree has grown constantly, with the trend particularly noticeable after accession and more strongly so among women than among men (see the wage and employment trends in the context of education in Estonia by Rõõm, 2007). Figure 9 illustrates the trends in the share of individuals with tertiary education in Estonia by citizenship. Native Estonian women show the steepest growth in attaining tertiary education in 2004-2019 after accession to the EU.

The trend of attaining tertiary education is, however, also positive for men and for women who are not local citizens. The share of individuals with tertiary education fell dramatically among males who were not EU citizens after EU accession and has still not recovered, most probably because of emigration. One tentative explanation for this might be that there are no publicly-funded Russian-language university study programmes in Estonia. Attainment of tertiary education has rebounded in recent years for women in Estonia with non-EU citizenship. At the same time, the share of the population whose educational level is less than primary, primary, or lower secondary has declined steadily (see Figure 10, which presents the share of employees by level of educational attainment in Estonia).

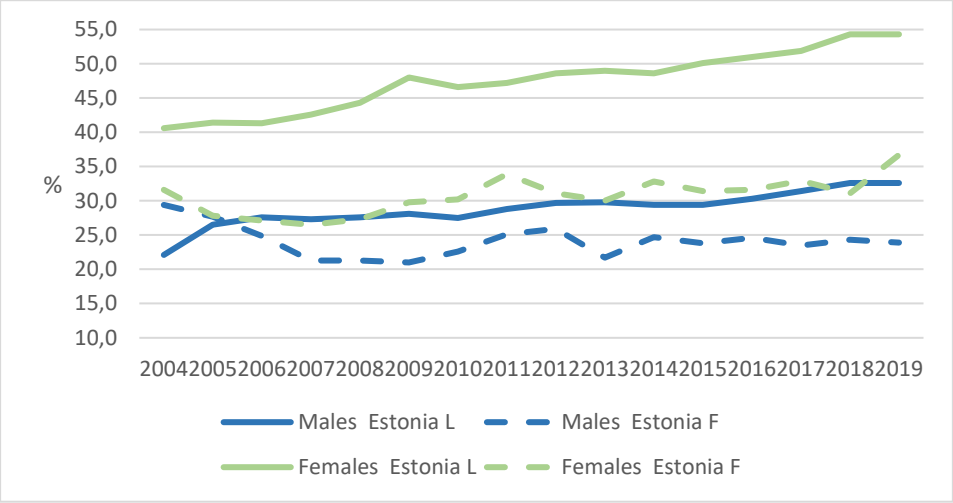


Figure 9: Tertiary educational attainment level of men and women by citizenship, aged 25–64 (L – local country, F – foreign country) (%).
Source: Eurostat

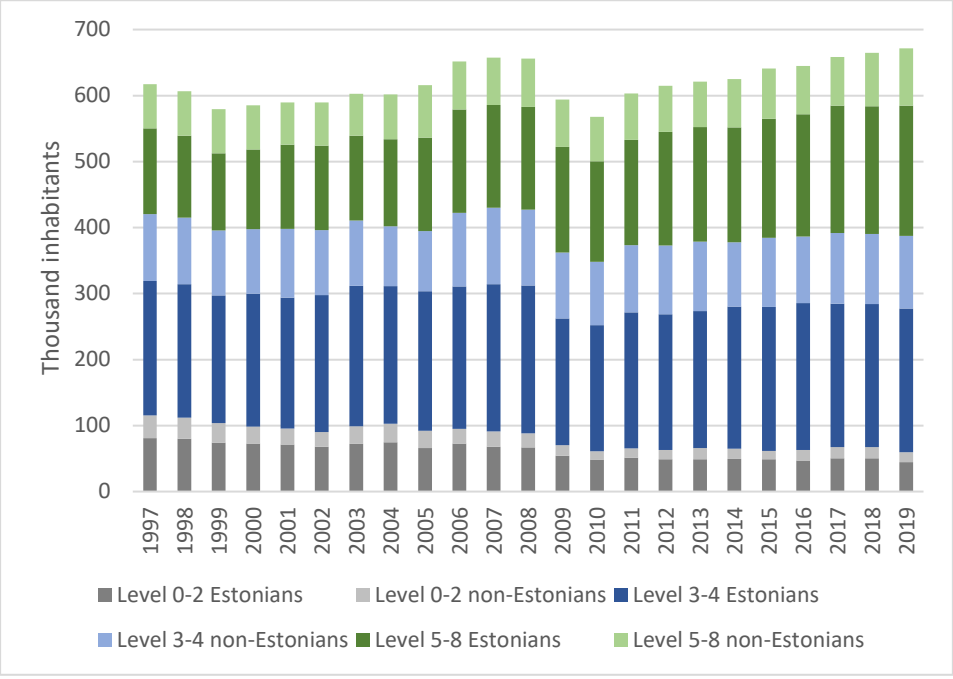


Figure 10: Share of employees by level of educational attainment and ethnicity in Estonia (Level 0-2 = less than primary, primary and lower secondary education; Level 3-4 = Upper secondary and post-secondary non-tertiary education; Level 5-8 = Tertiary education under ISCED 2011).
Source: Statistics Estonia

Figure 11 reflects the Duncan Employment Segregation Index for levels of education. Segregation for employment by educational attainment level during the period 2000–2019 is quite low, with the widest segregation at tertiary level. This rather mixed pattern reflects that Russian speakers need to increase their higher education attainment.

Estonia has maintained the school system inherited from the Soviet Union, in which education is segregated by language with separate schools providing education in the native language and in Russian. This setup reinforces the ethnic gaps in employment, wages and career opportunities. Political reforms have, however, introduced significant changes in recent years to education through language legislation.

The legislation passed in 1993 for basic schools (years 1–9) and high schools (years 10–12) declared that the language of instruction for high schools is Estonian; the language of instruction in basic schools may also be another language at the decision of the local government council. In practice the implementation stage of this reform took a long time, and the Government only established a gradual transition of the language of instruction in high schools from 2007, under which students whose language of instruction is Russian should study at least 60% of the minimum compulsory study volume in Estonian, while the remaining 40% of classes can be taught in the languages of national minorities or bilingually.

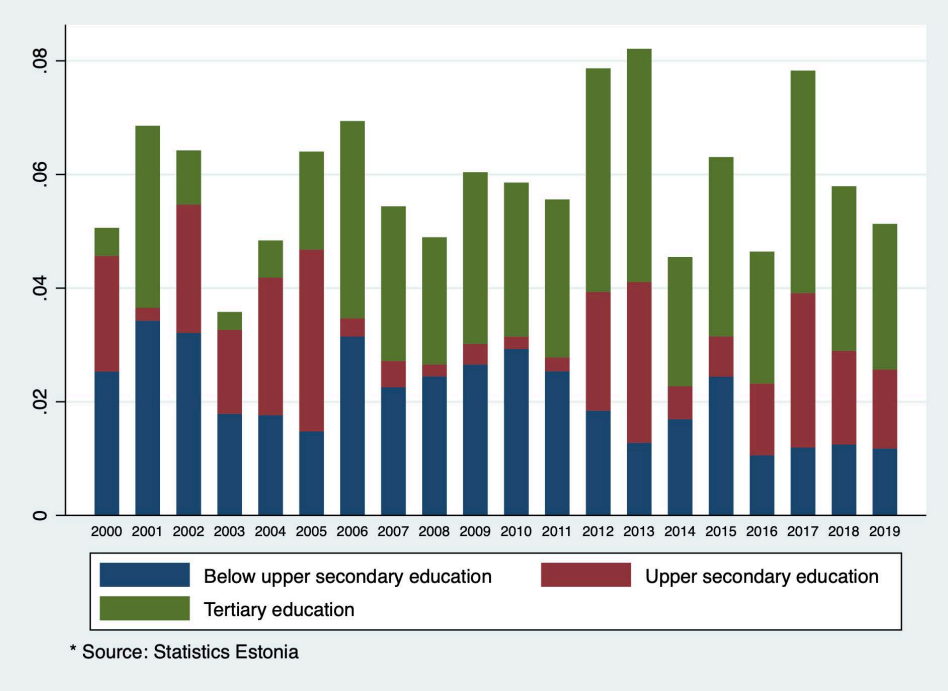


Figure 11: Structure and dynamics of the Duncan Employment Segregation Index for educational attainment levels by nationality. Groups: Estonian native speakers and Russian speakers for the years 2000–2019

According to Statistics Estonia (2020), the number of schools with instruction in Estonian decreased from 566 in 2000 to 436 in 2018, a fall of 23%, while the number of schools with instruction in Russian fell from 100 in 2000 to just 19 in 2018, a fall of 81%.

This could reflect demographic, migration, and structural problems, though it is also the case that a bilingual school system was introduced in the transition period as the immersion programme, and the number of such schools more than tripled from 19 in 2000 to 64 in 2018. The share of all basic school pupils with a mother tongue other than Estonian in schools where the language of instruction is Estonian is constantly growing, but mostly through an increase in the proportion of students in immersion classes, while the proportion directly in Estonian schools has diminished a little (see the analysis and overview of HTM, 2015 for the years 2005–2014). Vihalemm and Hogan-Brun (2013) show that there were both good and bad sides to the immersion programmes.

Thus the problems of segregation in schools still exist in Estonia. An additional and compounding disadvantage is the high concentration of Russian speakers in the north-east of Estonia who have limited opportunities and motivation to learn Estonian.

3. The main research hypotheses and contribution of the study

Integrating the labour market better and including disadvantaged groups and language minorities in it is one of the key challenges for advanced economies where populations witnessing low natural birth rates are ageing and declining. Transition and post-transition economies in particular need to sustain their labour resources for economic growth and convergence. The current evidence has shown that including and integrating ethnic and language minorities socio-economically has been a challenge and that the disparities have persisted between their economic outcomes and those of the native populations. Against this background, the research in this thesis aims to cast light on the differences and the reasons why the labour market outcomes observed vary across individuals depending on their ethnic and language backgrounds.

The thesis contributes to the literature of applied economics of language in two main ways. Firstly, the research in it addresses the particular linguistic context in Estonia, a European Union member country that hosts a homogeneous, sizeable Russian-speaking community, which is a legacy from the 50 years of the soviet era and which is about a quarter of Estonian population. A variety of datasets are used (ELFS, IMS, AES and PIAAC), and the novel integration of these sources of information to investigate how labour market outcomes in Estonia are related to language skills leads to new empirical insights. These insights led the author of the thesis to realise the crucial role of socio-economic, institutional and linguistic contextual factors that have a major imprint on whether there are returns to language and how large they are, and on what the economic outcomes in the labour market are. Given this, the second main contribution of the thesis is its conceptualisation of the contextual factors that determine the effect of language in the labour market. The third article finds empirical evidence by conducting a meta-regression and finds support for the importance of the contextual factors and study design elements that are largely responsible for the heterogeneous results on language returns in the literature on the subject.

The empirical studies in Paper I and Paper II address the research questions below and test the related hypotheses both in combination and in isolation:

- Do language skills in Estonian, Russian and English affect wages and the probability of unemployment across ethnic and gender lines for individuals working or seeking work in Estonia?
- Do language skills affect wages and the probability of unemployment differently or similarly for the Estonian and Russian language subgroups and across the gender lines, and if so how?
- Is the observed wage gap between native and non-native Estonian-speaking men attributable to or explained by the human capital characteristics observed, including cognitive skills?
- How do knowledge of languages and labour market relationships change across time?
- The analysis incorporates several individual-level background characteristics such as age, education, marital status, number of children, being in an inter-ethnic household, industry, occupation, and the region where the individual resides, and controls for how they affect wages and unemployment in both the Estonian and Russian language subgroups and across gender lines.

Paper I contributes by showing how Estonian and Russian as the two local languages, and English as the most powerful cross-border business language, relate to the employment status and wage of individuals residing and working in Estonia. This paper observes trends from 1989 until 2012. The study found strong ethnic occupational segregation for Russian-speaking men, at least for those who had attained no higher than secondary education. The Russian-speaking men were predominantly working in a Russian language environment, while the same was not true for the Russian speaking women. This result suggests that Russian-speaking men are mostly in jobs with little need to communicate and that the occupational segregation by gender reinforces the linguistic segregation of Russian men.

Paper II extends the first paper in that it investigates why language skills have no effect on wages for Russian-speaking men. To do this, it studies the wage differentials of native Estonian-speaking and Russian-speaking men while controlling for the impact of cognitive skills in literacy, numeracy, and problem-solving. This research was restricted to the male population only to mitigate the problems of endogenous self-selection into employment.

Paper III conducts a meta-regression analysis on how language skills affect labour incomes depending on socio-economic, institutional and language-related contextual factors, controlling for the study design and the estimation method. The meta-regression approach controls for how the global power of the language (PLI or Power Language Index) and linguistic diversity (LDI or Linguistic Diversity Index) affect the wage returns to knowing the local state language, a significant local minority language, or the business language. The study also controls for the impact of environmental factors such as globalisation (KOF Globalisation Index), economic structure (the unemployment rate and GDP per capita) and urbanisation (urban population statistics) that affect language-related wage returns. To the best of the author's knowledge, Paper III is the first meta-regression study to evaluate and synthesise the heterogeneous empirical literature that estimates the language returns in labour market outcomes. The main research questions and the hypotheses for testing in Paper III were:

- whether the socio-economic and linguistic context affects the estimated returns to language retrieved from the primary empirical studies;
- whether the study design affects the estimated returns to language retrieved from the primary empirical studies.

All three research papers in the thesis contribute to the literature on applied economics of language, with the first two papers providing primary empirical evidence and the third paper addressing the underlying heterogeneities in primary studies with a novel meta-regression analysis. Section 5, Synthesis and concept, generalises the research findings of the thesis into a holistic concept for the determinants that affect returns to language in the labour market.

4. Data and methodology

Table 1 presents a summary table for the research papers in this thesis.

Table 1: Summary table of the research papers in the thesis

Aim and context of the research	**Study design: Data and method**	**Main results**
Paper I: Bormann, Ridala & Toomet, 2019. Language Skills in an Ethnically Segmented Labour market: Estonia 1989–2012. *International Journal of Manpower*		
Effect of proficiency in Russian or Estonian and English on the earnings and probability of unemployment for Estonian and Russian men and women in Estonia	Data: ELFS Integration Monitor Survey (2008)[6] Method: OLS and LPM	*Wage* Russian language: medium income premium; Estonian: no income premium for Russian men and substantial premium for Russian women; English: large income premium for everyone *Unemployment* only Estonian language matters
Paper II: Ridala & Toomet, 2019. Wage gap in Ethnically Segmented Labour Market: The Role of Cognitive Skills. *Eastern European Economics*		
Wage gap between ethnic Estonian and Russian men in Estonia	Data: ELFS and PIAAC Method: Oaxaca-Blinder decomposition	Language skills have little impact on the differential; Cognitive skills have a moderate effect on the wage gap by reducing the unexplained difference from 20% to 16%
Paper III: Ridala, 2021. Language skills and labour market returns: A meta-regression analysis. *Language Problems and Language Planning*		
Effect on returns to (1) English/other business language, (2) state/local language, and (3) significant minority language, for the impact of environmental factors such as globalisation and migration trends, multilingual societies, labour market policies and regulations, economic structure, and urbanisation for OECD countries or OECD key partner countries like China, India, and South Africa	Data: 61 primary studies Method: Meta-regression, random effects model.	The results provide evidence that knowing the more influential languages is associated with higher labour market rewards; linguistic diversity has a negative effect on the returns to business and minority languages, but mixed effects emerged for the state language; urbanisation has a negative effect, but one that is modest in size, on returns to the business language and the significant minority language, while the effect for returns to the local state language is missing; the KOF Globalisation index gave no significant results; the results for unemployment are highly mixed

[6] See (Vetik et al., 2008) for more detailed information.

Paper I and Paper II draw on the Estonian Labour Force Survey (ELFS), which is conducted by Statistics Estonia, the national statistical office. The methodology for the survey changed in 1999, so part of the research in Paper I was limited to the years 2000–2012. Since 2000 this survey has used a quarterly rotating panel and collects roughly 16,000 annual observations from about 4,000 individuals. Earlier waves starting from 1989 were also included for the long-term analysis within this paper though until 1999, the survey had an annual cross-sectional structure. The research in Paper II focused on the year 2011.

The data for Paper I were confined to those for men and women in their prime working age of 25 to 55; the official retirement age was 55 for women and 60 for men until 1994 followed by a gradual increase over subsequent years. Paper II considered only male workers aged 24–63, as the legal retirement age for men was 63 in 2011.

The ELFS gives information about ethnic background and allows Estonian and non-Estonian workers to be distinguished. However, as most non-Estonians use Russian as their first language, they are referred as 'Russians' in Papers I and II. Ethnic identity in Estonia is usually based on language, though this may differ in certain cases such as individuals born in multilingual families. The labour market outcome considered in Paper I and Paper II was the "last net salary for the main job". Paper I also considered the employment status as working or unemployed of individuals active in the labour market.

The ELFS presents self-reported information on language skills. The survey asks whether the respondent speaks the language at home, is able to write and speak, only speaks, or only understands the language. The empirical research compressed that information into a dummy variable that equals one if the individual is able to speak and write the language and zero otherwise.

Paper I used data from the Integration Monitor Survey (2008) for additional evidence on ethnic and gender segregation. The Integration Monitor Survey is conducted by Saar Poll and aims to collect information on ethnic relations in Estonia. The analysis in Paper I looked at two survey questions on ethnic segregation and on workplace communication.

The second paper supplements the ELFS with the Survey of Adult Skills, which was collected within the Programme for the International Assessment of Adult Competencies (PIAAC), run jointly by the Organisation for Economic Co-operation and Development (OECD) and the Estonian Ministry of Education and Research between August 2011 and March 2012. The PIAAC provides representative and internationally comparable data about the skills of adult populations, and is designed to measure several key cognitive skills and social skills in the workplace. The cognitive skills are measured by targeted skill tests and thus provide scalable measures of an individual's abilities. The central variables of interest for the second paper are literacy, numeracy, and the ability to solve problems in a technology-rich environment (PS-TRE). The PIAAC also allows the cognitive skills to be linked to a broad set of individual background characteristics including education, employment, immigrant status and family background. It does not contain questions on ethnic background, so for the purposes of the study the respondent's ethnicity is identified from their home language, which is their *first language, learned at home in childhood and still understood*.

Paper III employs a meta-regression analysis based on 61 empirical research papers published between 1998 and 2019 in the field of economics of language that estimate how wages depend on language skills in OECD countries or in OECD key partner countries like China, India and South Africa. Most of the underlying papers for the meta-regression analysis applied the Mincer-type earnings regression approach (Mincer, 1958), which

estimates the relationship between human capital variables and labour earnings, with the language skills of individuals included under human capital. The meta-dependent variable is the vector of estimates of the language effect from the underlying wage regressions in the primary studies. The meta-independent variables, or moderators, control for the cross-study confounders in study design and context, with the design factors varying across studies and the context factors varying across the countries observed and years. Six variables control for the cross-study context: (1) the unemployment rate in per cent of the active population in employment (source: World Bank); (2) real GDP per capita (source: World Bank, in current USD); (3) the Power Language Index (PLI)[7], which ranks the languages studied by their role geographically, economically, and in communication, knowledge and media, and diplomacy (Chan, 2016); (4) the Linguistic Diversity Index (LDI)[8], which captures the probability of two people randomly selected from the population speaking different native languages; (5) the KOF Globalisation Index (KOFGI)[9], which ranks countries on a scale of economic, social and political globalisation (Dreher, 2006); and finally (6) urbanisation, which is the percentage of the urban population within the total population in the World Bank statistics[10]. The study design variables control for whether the data structure is panel or cross-section, and for whether the estimator type used was ordinary least squares (OLS) or an instrumental variable (IV) estimator.

Paper I analyses the relationship between labour market outcomes and language skills using an ordinary linear regression. Linear probability models are applied to estimate the relationship between language skills and the probability of unemployment. The effect of language skills on wages is estimated by OLS to facilitate comparison with other studies, since Azam et al. (2013) and Di Paolo and Tansel (2015) for instance report only OLS results, and moreover find that the available proxy variables have little effect on the final results. Three models were estimated with various sets of control variables. Model 1 only includes the dummies for language skills and year. It describes the raw effect of language skills net of strong time trends in the data. Model 2 additionally includes the individual characteristics of age and education and the family descriptors for marital status, an indicator for inter-ethnic marriage, and having children aged up to 17, and also county of residence. Finally, Model 3 also adds industry, occupation and other workplace descriptors. Only the first two models are estimated for the probability of unemployment.

Paper II analyses the average wage gap between Estonians and Russian speakers using the popular Oaxaca-Blinder method to decompose the wage difference between those groups into two components, explained and unexplained differences. A number of models are estimated that differ in the types of explanatory variables they include. Model 1 only includes a constant, corresponding to the raw differentials in the data. Model 2 adds age, immigrant status as born outside Estonia or not, and family descriptors for being married or cohabiting, and having children aged up to 17. Model 3 additionally includes education descriptors, and Model 4 adds dummies for industry and occupation. Model 5 was estimated using ELFS data, which also include language skills, and Model 6 includes geographical controls in addition to language skills. In this way it was possible

[7] http://www.kailchan.ca/wp-content/uploads/2016/12/Kai-Chan_Power-Language-Index-full-report_2016_v2.pdf

[8] https://en.wikipedia.org/wiki/Linguistic_diversity_index#Rankings_by_country

[9] https://kof.ethz.ch/en/forecasts-and-indicators/indicators/kof-globalisation-index.html

[10] https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS

to assess which variables are related to the unexplained wage differential and see if the ELFS and the PIAAC corroborate each other. Only one version of each model used the ELFS data, while an amended version of each model was estimated for the PIAAC data, which include the three cognitive skills, both separately and all together. This shows whether introducing skills narrows the unexplained wage differential significantly, which would suggest that literacy, numeracy, and PS-TRE explain a part of the differential.

All the underlying articles of the current thesis use micro-data, which have certain data limitations. The ELFS data are missing strong instruments that would better correct for omitted variable bias and measurement errors, but the authors believe that the OLS results facilitate comparison with other studies and are satisfactory for inferring a linearly approximated relationship. A particular limitation is that although individual characteristics are able to explain a large part of the returns to English, the error term may still contain unobserved factors that correlate with English language skills and so may be a cause of ability bias. Another limitation is that the information on language skills is self-reported. Dustmann and Fabbri (2003) claim that self-reported language variables are likely to suffer from misclassification errors because individuals have a tendency to over or under-report their language skills.

A limitation for Paper II is the absence of explicit information on language skills in the PIAAC survey. In consequence the language used to complete the survey was taken instead, as the survey could be completed in Estonian, Russian or English. Admittedly this may misclassify a number of bilingual individuals, but as Toomet (2011) showed, there is only a low correlation between Estonian language skills and income. Another limitation is that unlike the ELFS, the PIAAC does not contain information on where the respondents reside. In consequence some of the model specifications did not control for residency in the metropolitan area or capital city, which is a variable with strong explanatory power. As Leping and Toomet (2008) show, this may lead the unexplained wage gap to be underestimated, so estimates in this paper rather reflect the lower bound.

The empirical contributions in Paper I and Paper II are built on cross-sectional survey data, which do not allow for any control for the unobserved heterogeneity of individuals or for the time dynamics and state dependencies in the way that a panel data structure does. Hence one of the central limitations has been the underlying cross-sectional data structure, as the estimates based on that structure may suffer from some omitted variable bias. Analysis based on cross-sectional surveys is, however, very common in the empirical literature on economics of language and authors with strong contextual knowledge of the field can infer the potential biases and guide their readers accordingly in drawing conclusions about the estimated results.

The highly heterogeneous study contexts and study designs in the empirical literature on economics of language may give rise to concerns about external validity. The meta-regression in Paper III conducts a systematic evaluation of the results from context-specific primary research papers that estimate the returns to language, and in doing so it addresses the external validity concerns by controlling for the effect and impact of a number of contextual and study design factors.

5. Synthesis and concept

This section aims to synthesise the key empirical findings from the three papers in the thesis. This synthesis is the foundation for a conceptual framework for a better understanding of whether and how far language skills generate labour market benefits for the individual. Figure 12 below presents the concept.

The literature on economics of language mostly considers individual characteristics and the workplace descriptors in its analysis of how language skills affect labour market outcomes. In a similar manner, the original empirical studies in Paper I and Paper II incorporated a number of individual-level characteristics as controls in the labour market outcome regressions.

Following from the literature synthesis and meta-regression on the literature of economics of language provided in Paper III, and the literature on labour market integration, the labour market effects of globalisation, and the economic value of reciprocal bilingualism, the author proposes a conceptual framework for a holistic understanding of the effect of second language skills on labour market outcomes (see Figure 12). While conventional individual and workplace characteristics are mostly accounted for in the research of economics of language, the role of environmental factors is less clear, and so Paper III of this thesis attempts to link them with the returns to language using the meta-regression analysis. Language as a tool for communication and cultural manifestation is hard to measure quantitatively, though the benefits of language have a pecuniary economic value. The relationship between language skills and the economic benefits that stem from knowledge of languages is endogenous. This means that language skills improve labour market opportunities, but employment and higher labour income in turn widen social networks and generate opportunities for enhancing language skills.

The economic value of a language depends crucially on its global power and influence. Chan (2016) introduces a concept that compares and ranks languages across multiple dimensions including geography, the economy, communication, knowledge, media and diplomacy. The global influence of languages is quantified in the Power Language Index (PLI). The case of Estonia provides an unusual example where the small titular language of the native population coexists with powerful languages spoken by the local Russian-speaking minority and the business community where English dominates. The meta-regression analysis highlights the importance of the influence of languages as measured with the PLI, and provides evidence that knowing the more influential business languages leads to a stronger increase in labour earnings.

The meta-analysis in the third paper shows that the KOF Globalisation Index, which reflects the trends of globalisation and migration together with labour market policy and regulations, was estimated with low precision and gave no significant results. This finding could be partially explained by the relatively strong positive correlation with other economic indicators of GDP per capita and unemployment, but from the other side it could hint at a tentative hypothesis that the effects of globalisation counterbalance each other, as some of them increase the returns to language skills and others of them reduce the returns. In this case, the globalisation of labour markets could reward knowledge of a second language by increasing the need and demand for language learning with better wage prospects and improved probability of employment, while at the same time there may be a high level of supply in language skills that would reduce the private returns from knowing widely-spoken languages. To confirm this tentative hypothesis the sample size for the meta-analysis needs to be enlarged.

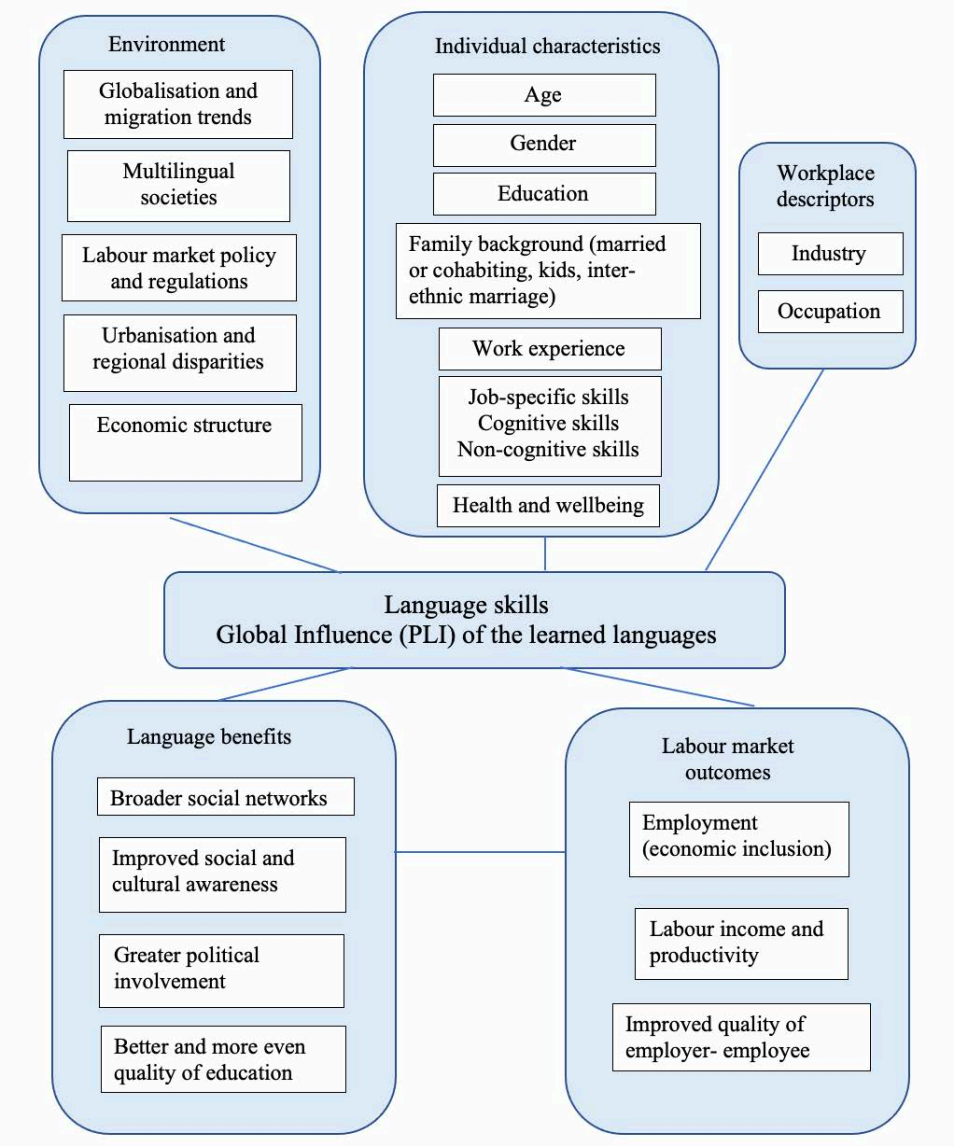


Figure 12: Concept of the effect of second language skills on labour market outcomes.
*Sourc*e: Compiled by the author

The results from Paper III on economic structure proxied by unemployment suggest that better language skills in the business language and in official state languages benefit workers in labour markets that have high unemployment, and this evidence is also noticeable in the Estonian transition context as reflected in Paper I. The effect upon returns to language skills may vary depending on the nature of structural imbalances, which depict the supply and demand of labour in the economy as indicated by unemployment. The results from the meta-regression in Paper III also suggest that immigrant populations gain from knowing the local state language or a common business language, particularly so in economies that witness high unemployment and potentially large structural imbalances. This meta-study also reveals that economies suffering from

high unemployment do not grant returns to the local minority language as a second spoken language.

In direct contrast to this, GDP per capita, which in general reflects the advancement and structural diversity of an economy, shows that higher production per capita is associated positively with command of the local minority language as a second language.

According to the contemporary literature on labour economics, wage growth depends on two key drivers, which are "human capital" and "job shopping" (Rubinstein & Weiss, 2006). Better language skills and broader social networks help to reduce information asymmetries between job seekers and employers, and this not only generates private benefits for the workers and employers, but also more generally reduces frictions and inefficiencies in the labour market. Grin (2003) claims that foreign language skills may lead to better job matching. This is because improving language skills increases human capital and so potentially increases labour productivity, which is manifested in higher wages. A wage premium could perhaps arise because foreign language skills as a rule signal higher productivity even if a worker's job profile does not directly require them to know foreign languages, and so they could be a signal for employers about a worker's unobserved cognitive ability and motivation (Stöhr, 2015). The positive correlation between bilingualism and cognitive skills has been noted by Adesope et al. (2010), Bialystok et al. (2009) and Costa and Sebastián-Gallés (2014), along with the implications of this for earnings. Employees with command of more than one language can be more productive than monolingual ones, usually because they can perform duties that monolingual employees cannot (Grenier, 1984).

Language skills are also endogenous to educational outcomes. Better command of language opens up access to higher quality education, while education in turn enhances knowledge of languages. This reinforcing complementarity of education and language skills leads to better employment prospects for the more educated, who are then more likely to be involved in lifelong learning. Language skills are also important for signalling purposes since they provide evidence of an individual's ability to learn new skills and so reflect that individual's potential productivity (Mavisakalyan, 2017).

The environment has an important role in generating incentives for language learning and in shaping the relationship between language skills and labour market outcomes. Living in a linguistically diverse environment, like many metropolitan areas, may be associated with a higher probability of knowing a second language. Both the multilingual society of a metropolitan area and the greater diversity and size of its labour markets promote incentives for language learning. Diverse labour markets reward language skills by offering improved prospects for employment and remuneration for work. Equally though, the sizable ethnic groups in metropolitan environments may establish their own micro-societies, which allow some members of the ethnic group to escape from the need to interact directly with the host country society and speak the language of the host country. The division of roles within these large ethnic micro-societies or language enclaves may reduce the incentive to learn the host country language, at least for some members of the language group.

The Linguistic Diversity Index (LDI), which describes multilingual societies and is incorporated into the third paper of the thesis, shows that diversity has a positive effect on the returns to knowing the local state language, but weakens the returns to the business language. Multilingual countries like Belgium, Switzerland, Luxembourg, Canada, or Latvia and Estonia apparently encourage the learning of local languages.

Urbanisation also shows a negative effect on returns to the business language, though one that is modest in size.

Another important environmental factor in the transmission of language skills into economic outcomes is the structure and the industry composition of the economy. The rewards for language skills may vary greatly across industrial sectors. Bormann et al. (2019) show that the effects of the second language on wages for Russian speakers in Estonia are largest in the mining and quarrying industry and in the accommodation and food service sector. The largest gains for native Estonians from skills in English as a second language emerge in the information and communications sector. Estonian men also have high returns to language in the construction sector and the accommodation and food service sector (Bormann et al., 2019).

The major determinant of the rewards of language skills is the occupation. Bormann et al. (2019) show that Estonians gain most from second language skills when they hold positions as managers, senior officials or legislators, or when they are employed in professional and technical occupations in the ISCO-08 categories. By contrast, low-skill occupations like cleaners and helpers, food preparation assistants, and sales assistants do not benefit from second language skills. The association for Russian speakers between occupation and returns to language skills was negative or was not significant. Leppik and Vihalemm (2015) conclude, and Bormann et al. (2019) confirm, that Russian-speaking men are prevalent in those jobs that are associated with limited career mobility. Russian women, in contrast, benefit from fluency in the second language when they hold managerial and professional positions, though less than Estonian women do.

The economic benefits from skills in the local state or business language may go beyond private returns and generate externalities or positive social returns. The diversity of human capital and the variety of skills grow, and this enriches and deepens labour markets. Socio-economic integration of native speakers and immigrant populations opens up mutual gains such as access to the human capital and resources of both groups (Besevegis & Pavlopoulos, 2008). Using the resources of both ethnic groups by improving inter-ethnic networks may lead to multilateral assistance for minorities and in turn be converted into economic benefits in the form of better employability, higher wages, and the creation of new jobs. Successful integration is when the law of one wage holds in any country when native and non-native workers are compared, meaning workers whose mother tongue is the main official language earn the same as those for whom this is not the case (Armstrong, 2015).

6. Final comments

Integrating labour markets better and including disadvantaged groups and language minorities is one of the key challenges for contemporary economies in sustaining their economic growth. The empirical evidence from Papers I and II has shown that ethnic and language minorities in Estonia have been included and integrated socio-economically only sluggishly and the disparities in economic outcomes relative to those of native populations have persisted.

After a decade as part of the EU, Estonia is still witnessing ethnic and social segregation despite the progress that has been made in education, in improving language skills, and in increasing the share of the population who hold Estonian and EU citizenship. The ethnic segregation seems to be less of an issue for women, as the first paper of the thesis concluded. This paper finds evidence that knowledge in Estonian is increasingly more important for labour market success of Russian speaking women. Unfortunately, it cannot be seen from the current data whether this is because segregation is decreasing or because the Estonian language has become increasingly important across the economy.

The meta-study from Paper III demonstrates that the socio-economic, institutional and ethnolinguistic environment plays an important role in generating the incentives for language learning and in shaping the relationship between language skills and labour market outcomes. The two primary empirical studies in the thesis (Paper I and Paper II), the meta-regression (Paper III) and the proposed concept (see Figure 12) together suggest lines of investigation that would address the main research limitations of the existing studies and offer better grounds for setting policy.

Empirical research on economics of language is mostly conducted using cross-sectional surveys, but there are several limitations to this. The cross-sectional data may be prone to external validity issues that arise from the particular time and context of the observation and study. Comparative studies of countries or regions allow some control of environmental background characteristics, but they do not address a very commonly noted limitation of the cross sections, which is that they do not allow unobserved individual heterogeneities and dynamic individual patterns such as individual state-dependencies to be controlled for. Another common limitation in the language studies is in how they measure language skills, which are often self-reported. The research in the current thesis has also employed surveys that use self-reported data for language skills, which may suffer from reduced reliability because of self-reporting bias. The meta-regression in Paper III of the thesis shows there to be a significant difference between estimates arrived at using ordinary regression methods and those from instrumental regression methods, which suggests that the measurement issues with language skills are non-negligible. The pervasiveness of the English language in economic and social interaction in the digital era makes fluency in English a necessary control for measuring an individual's economic and labour market success. This means it is of paramount importance to control for knowledge of English when estimating returns to local languages in contemporary labour markets.

To build on the current research given the limitations listed above, suggested further lines of study are:

- Data availability and structure should be improved so that reliable estimates on language returns in labour markets can be retrieved. Panel data structures would allow for better consistency of estimates. Additional dimensions, such as the measures of cognitive skills used in Paper II, would offer better instrumentation options for alleviating

the potential bias in self-reported language skills. It would also be worth extending the data sources to include job portals to investigate the link between language skills and labour supply and demand patterns.

- The explosive immigration from third countries into the EU has also been evident, though less intense, in Estonia and the other Baltic states. Unlike in most other EU countries, the new wave of immigrants in the Baltic countries is dominated by refugees from Ukraine, Belarus and other former states of the Soviet Union (Rudi, 2019). Comparing and contrasting the emerging labour market patterns from this new wave of immigration with those of the long-established Russian-speaking community would offer valuable insights and novel lines of investigation that could lead to a better understanding of the processes of social and economic integration.

List of Figures

List of Tables

References

Adamchik, V., Hyclak, T., & Sedlak, P. (2019a). Poland in an integrated European economy: Are foreign language skills valued by employers in the Polish labor market? *Journal of Transition Studies Review, 26*(1), 31–55.

Adamchik, V. A., Hyclak, T. J., Sedlak, P., & Taylor, V. (2019b). Wage returns to English proficiency in Poland. *Journal of Labor Research, 40*, 276–295. doi:10.1007/s12122-019-09291-2

Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research 8*(2), 207–245. doi:10.3102/0034654310368803

Aiyar, S., Barkbu, B., Batini, N., Berger, H., Detragiache, E., Dizioli, A., Ebeke, C., Lin, H., Kaltani, L., Sosa, S., Spilimbergo, A., Topalova, P. (2016). *The refugee surge in Europe: Economic challenges*. IMF Staff Discussion Note SDN/16/02, International Monetary Fund. Retrieved from https://www.imf.org/external/pubs/ft/sdn/2016/sdn1602.pdf

Aldashev, A., Gernandt, J., & Thomsen, S. L. (2009). Language usage, participation, employment and earnings: Evidence for foreigners in West Germany with multiple sources of selection. *Labour Economics, 16*(3), 330–341.

Altonji, J. G., & Blank, R. M. (1999). Race and gender in the labor market. In O. C. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics* (Vol. 3, pp. 3143–3259). Amsterdam: Elsevier Science Publishers B.V.

Angouri, J. (2014). Multilingualism in the workplace: Language practices in multilingual contexts. *Multilingua: Journal of Cross-Cultural and Interlanguage Communication, 33*(1-2), 1–9.

Aptekar, S. (2009). Contexts of exit in the migration of Russian speakers from the Baltic countries to Ireland. *Ethnicities, 9*(4), 507–526.

Armstrong, A. (2015). Equilibria and efficiency in bilingual labour market. *Journal of Economic Behavior & Organization, 112*(C), 204–220.

Azam, M. C., Chin, A., & Prakash, N. (2013). The returns to English-language skills in India. *Economic Development and Cultural Change, 61*(2), 335-367.

Becker, G. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy, 70*(5), 9–49.

Becker, G. (1964). *Human capital*. New York: National Bureau of Economic Research: Distributed by Columbia University Press.

Becker, G. S. (2010). *The economics of discrimination*. Chicago & London: University of Chicago Press.

Berman, E., Lang, K., & Siniver, E. (2003). Language-skill complementarity: returns to immigrant language acquisition. *Labour Economics, 10*(3), 265–290.

Besevegis, E., & Pavlopoulos, V. (2008). Acculturation patterns and adaptation of immigrants in Greece. In M. Finklestein, K. Dent-Brown (Eds.). *Psychosocial Stress in Immigrants and in Members of Minority Groups as a Factor of Terrorist Behavior* (pp. 23–34). Retrieved from http://ebooks.iospress.nl/volumearticle/24525

Bialystok, E., Craik, F. I., Green, D. W., & Gollan, T. H. (2009). Bilingual minds. *Psychological Science in the Public Interest, 10*(3), 89–129. doi:10.1177/1529100610387084

Black, D., Haviland, A., Sanders, S., & Taylor, L. (2006). Why do minority men earn less? A study of wage differentials among the highly educated. *The Review of Economics and Statistics, 88*(2), 300–313.

Blackaby, D. H., Leslie, D. G., Murphy, P. D., & O'Leary, N. C. (2005). Born in Britain: How are native ethnic minorities faring in the British labour market? *Economics Letters, 88*(3), 370–375.

Bleakley, H., & Chin, A. (2004). Language skills and earnings: Evidence from childhood immigrants. *Review of Economics and Statistics, 86*, 481–496.

Borjas, G. J., Freeman, R. B., & Katz, L. F. (1996). Searching for the effect of immigration on the labor market. *The American Economic Review, 86*(2), 246–251.

Bormann, S.-K., Ridala, S., & Toomet, O. (2019). Language skills in an ethnically segmented labour market: Estonia 1989-2012. *International Journal of Manpower*. doi: https://doi.org/10.1108/IJM-06-2017-0115

Borooah, V., & Mangan, J. (2009). Multiculturalism versus assimilation: Attitudes towards immigrants in western countries. *International Journal of Economic Sciences and Applied Research, 2*(2), 33–50.

Bound, J., & Freeman, R. B. (1992). What went wrong? The erosion of relative earnings and employment among young black men in the 1980s. *The Quarterly Journal of Economics, 107*(1), 201–232.

Bratsberg, B., & Ragan, J. (2002). The impact of host-country schooling on earnings: A study of male immigrants in the United States. *The Journal of Human Resources, 37*(1), 63–105. doi:10.2307/3069604

Breton, A. (1964). Economics of nationalism. *Journal of Political Economy, 72*(4), 376–386.

Breton, A. (1978). Nationalism and language policies. *Canadian Journal of Economics, 11*(4), 656–668.

Bretton, H. (1976). Political science, language, and politics. In W. O'Barr & J. O'Barr (Eds.), *Language and Politics*. The Hague: Mouton.

Brubaker, R. (1996). *Nationalism reframed: Nationhood and the national question in the new Europe*. Cambridge: Cambridge University Press.

Budría, S., Martinez de Ibarreta, C., & Swedberg, P. (2017). The impact of host language proficiency across the immigrants' earning distribution in Spain. *IZA Journal of Development and Migration, 7*(1), 12.

Budría, S., & Swedberg, P. (2012). *The Impact of language proficiency on immigrants' earnings in Spain*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2170645

Cappellari, L., & Di Paolo, A. (2015). *Bilingual schooling and earnings: Evidence from a language-in-education reform*. Research report 9431, Bonn: Institute of Labor Economics. Retrieved from Bonn: https://www.econstor.eu/bitstream/10419/124941/1/dp9431.pdf

Carliner, G. (1981). Wage differences by language group and the market for language skills in Canada. *Journal of Human Resources, 16*(3), 384–399.

Carlsson, M., & Rooth, D.-O. (2011). *Employer attitudes, the marginal employer and the ethnic wage gap*. Research report 6227, Bonn: Institute of Labor Economics. Retrieved from https://www.iza.org/publications/dp/6227/employer-attitudes-the-marginal-employer-and-the-ethnic-wage-gap

Carnevale, A. P., Fry, R. A., & Lowell, B. L. (2001). Understanding, speaking, reading, writing, and earnings in the immigrant labor market. *American Economic Review, 91*(2), 159–162. doi:10.1257/aer.91.2.159

Casale, D., & Posel, D. (2011). English language proficiency and earnings in a developing country: the case of South Africa. *The Journal of Socio-Economics, 40*(4), 385–393.

Castells, M. (1996). *The rise of the network society, the information age: Economy, society and culture* (Vol. I). Cambridge, MA; Oxford, UK: Blackwell.

Chan, K. L. (2016). Power Language Index. Which are the world's most influential languages? Retrieved from http://www.kailchan.ca/wp-content/uploads/2016/12/Kai-Chan_Power-Language-Index-full-report_2016_v2.pdf

Chiswick, B. R. (1977). Sons of immigrants: Are they at an earnings disadvantage? *American Economic Review, 67*(1), 376–380.

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

Chiswick, B. R. (1998). Hebrew language usage: determinants and effects on earnings among immigrants in Israel. *Journal of Population Economics, 11*(2), 253–271.

Chiswick, B. R. (2008). *The economics of language: An introduction and overview*. Research report 3568, Bonn: Institute of Labor Economics. Retrieved from http://repec.iza.org/dp3568.pdf

Chiswick, B. R., & Larsen, N. (2015). Russian Jewish immigrants in the United States: The adjustment of their English language proficiency and earnings in the American Community Survey. *Contemporary Jewry, 35*(3), 191–209.

Chiswick, B. R., & Miller, P. W. (1995). The endogeneity between language and earnings: International Analyses. *Journal of Labor Economics, 13*(2), 246–288.

Chiswick, B. R., & Miller, P. W. (2002a). *Do enclaves matter in immigrant adjustment?* Research report 449, Bonn: Institute of Labor Economics. Retrieved from https://www.econstor.eu/bitstream/10419/21507/1/dp449.pdf

Chiswick, B. R., & Miller, P. W. (2002b). Immigrant earnings: language skills, linguistic concentrations and the business cycle. *Journal of Population Economics, 15*, 31–57. doi:10.1007/PL00003838

Chiswick, B. R., & Miller, P. W. (2010). Occupational language requirements and the value of English in the US labor market. *Journal of Population Economics, 23*(1), 353–372.

Chiswick, B. R., & Miller, P. W. (2016). *Does bilingualism among the native born pay?* Research report 9791, Bonn: Institute of Labor Economics. Research report 9791. Retrieved from https://www.econstor.eu/bitstream/10419/21507/1/dp449.pdf

Chiswick, B. R., & Repetto, G. (2001). Immigrant adjustment in Israel: Literacy and fluency in Hebrew and earnings. In S. Djajic (Ed.), *International Migration: Trends, Policy and Economic Impact* (pp. 204–228). New York: Routledge.

Christofides, L. N., & Swidinsky, R. (2010). The economic returns to the knowledge and use of a second official language: English in Quebec and French in the Rest-of-Canada. *Canadian Public Policy, 36*(2), 137–158.

Clark, K., & Drinkwater, S. (2009). Dynamics and diversity: ethnic employment differences in England and Wales, 1991-2001. *Research in Labour Economics, 29*, 299–333.

Cohen-Goldner, S., & Eckstein, Z. (2008). Labor mobility of immigrants: Training, experience, language, and opportunities. *International Economic Review, 49*(3), 837–872.

Cornwell, K., & Inder, B. (2008). Language and labour markets in South Africa. *Journal of African Economies, 17*(3), 490–525.

Costa, A., & Sebastián-Gallés, N. (2014). How does the bilingual experience sculpt the brain? *Nature Reviews Neuroscience, 15*(5), 336–345. doi:10.1038/nrn3709

Di Paolo, A., & Raymond, J. (2012). Language knowledge and earnings in Catalonia. *Journal of Applied Economics, 15*(1), 89–118.

Di Paolo, A., & Tansel, A. (2015). Returns to foreign language skills in a developing country: The case of Turkey. *The Journal of Development Studies, 51*(4), 407–421. doi:10.1080/00220388.2015.1019482

Di Paolo, A., & Tansel, A. (2019). English skills, labour market status and earnings of Turkish women. *Empirica, 46*(4), 669–690.

Dreher, A. (2006). Does globalization affect growth? Evidence from a new Index of Globalization. *Applied Economics, 38*(10), 1091–1110.

Drinkwater, S. J., & O'Leary, N. C. (1997). Unemployment in Wales: Does language matter? *Regional Studies, 31*(6), 583–591.

Drydakis, N. (2012). Ethnic identity and immigrants' wages in Greece. International Journal of Intercultural Relations, 36(3), 389–402.

Duchêne, A. (2008). *Ideologies across nations: The construction of linguistic minorities at the United Nations*. New York: De Gruyter.

Duncan, A., & Mavisakalyan, A. (2015). Russian language skills and employment in the Former Soviet Union. *Economics of Transition, 23*(3), 625–656.

Duncan, O. D., & Duncan, B. (1955). A methodological analysis of segregation indexes. *American Sociological Review, 20*(2), 210–217. doi:DOI: 10.2307/2088328

Dustmann, C., & Fabbri, F. (2003). Language proficiency and labour market performance of immigrants in the UK. *The Economic Journal, 113*(489), 695–717.

Dustmann, C., & Theodoropoulos, N. (2010). Ethnic minority immigrants and their children in Britain. *Oxford Economic Papers, 62*(2), 209–233.

Dustmann, C., & van Soest, A. (2001). Language fluency and earnings: estimations with misspecified indicators. *The Review of Economics and Statistics, 83*(4), 663–674.

Dustmann, C., & Van Soest, A. (2002). Language and the earnings of immigrants. *Industrial and Labor Relations Review, 55*(3), 473–492.

Eamets, R., & Tiwari, A. K. (2019). Minimum wage in Estonia and its impact on employment and wage distribution. *CESifo DICE Report, 16*(4), 37−43. Retrieved from www.ifo.de/DocDL/dice-report-2018-4-eamets-tiwari-january.pdf

Egger, P. H., & Lassmann, A. (2012). The language effect in international trade: A meta-analysis. *Economics Letters, 116*(2), 221–224.

Elliott, R. J. R., & Lindley, J. K. (2008). Immigrant wage differentials, ethnicity and occupational segregation. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 171*(3), 645–671.

Ethnologue. (2017). Ethnologue: Languages of the World. Twentieth Edition (SIL International).

ETUI (2017). Labour market reforms in Estonia: background summary. Retrieved from https://www.etui.org/covid-social-impact/estonia/labour-market-reforms-in-estonia-background-summary

Fabo, B., Beblavý, M., & Lenaerts, K. (2017). The importance of foreign language skills in the labour markets of central and Eastern Europe: assessment based on data from online job portals. *Empirica, 44*(3), 487–508.

Ferraro, S., Meriküll, J., & Staehr, K. (2018). Minimum wages and the wage distribution in Estonia. *Applied Economics, 50*(49), 5253–5268.

Fishman, J. A. (1967). *Readings in the sociology of language*. The Hague, Paris: Mouton.

Frankel, J., & Rose, A. (2002). An estimate of the effect of currencies on trade and income. *Quarterly Journal of Economics, 117*(2), 437–466.

Friedman, M., & Kuznets, S. (1945). *Income from independent professional practice*. New York: National Bureau of Economic Research.

Galasi, P. (2003). *Estimating wage equations for Hungarian higher-education graduates*. Budapest Working Paper on the Labour Market No. BWP 2003/4, Institute of Economics, Hungarian Academy of Sciences. Retrieved from http://dx.doi.org/10.2139/ssrn.465882

Gazzola, M., Grin, F., & Wickström, B.-A. (2016). A concise bibliography of language economics. In M. Gazzola & B.-A. Wickström (Eds.), *The Economics of Language Policy*. Cambridge, MA: The MITT Press.

Gazzola, M., & Wickström, B.-A. (Eds.) (2016). *The economics of language policy*. Cambridge, MA: The MITT Press.

Gil, J. (2010). The double danger of English as a global language. *English Today, 26*(1), 51–56. doi:10.1017/S0266078409990575

Ginsburgh, V., & Moreno-Ternero, J. D. (2018). Compensation schemes for learning a lingua franca in the European Union. *World Economy, 41*(7), 1775–1789. doi:10.1111/twec.12644

Ginsburgh, V., & Prieto Rodriguez, J. (2013). Is there a gender bias in the use of foreign languages in Europe? *Kyklos, 66*(4), 552–566.

Ginsburgh, V. A., & Prieto-Rodriguez, J. (2011). Returns to foreign languages of native workers in the European Union. *Industrial and Labor Relations Review, 64*(3), 599–617. Retrieved from www.jstor.org/stable/41149481

Ginsburgh, V., & Weber, S. (Eds.) (2016). *The Palgrave Handbook of Economics and Language.* London: Palgrave Macmillan UK.

Ginsburgh, V., & Weber, S. (2020). The economics of language. *Journal of Economic Literature, 58*(2), 348–404. doi:10.1257/jel.20191316

Gonzalez, L. (2005). Nonparametric bounds on the returns to language skills. *Journal of Applied Econometrics, 20*(6), 771–975. doi:10.1002/jae.795

Grenier, G. (1984). The effects of language acquisition on the wage of Hispanic-American males. *Journal of Human Resources, 19*(1), 35–52.

Grenier, G. (1987). Earnings by language group in Quebec in 1980 and emigration from Quebec between 1976 and 1981. *Canadian Journal of Economics, 20*(4), 774–791.

Grenier, G. (2015). *The value of language skills*. IZA World of Labor. doi: 10.15185/izawol.205. Retrieved from https://wol.iza.org/articles/economic-value-of-language-skills/long

Grenier, G., & Vaillancourt, F. (1983). An economic perspective on learning a second language. *Journal of Multilingual and Multicultural Development, 4*(6), 471–483.

Grin, F. (1996a). Economic approaches to language and language planning: an Introduction. *International Journal of the Sociology of Language*, 121, 1–16.

Grin, F. (1996b). The economics of language: Survey, assessment, and prospects. *International Journal of the Sociology of Language, 121*(1), 17–44.

Grin, F. (2003). Language planning and economics. *Current Issues in Language Planning, 4*(1), 1–66.

Grin, F. (2016). Fifty years of economics in language policy: Critical assessment and priorities In M. Gazzola & B.-A. Wickström (Eds.), *The Economics of Language Policy*. Cambridge: The MIT Press.

Grin, F., & Sfreddo, C. (1998). Language-based earnings differentials on the Swiss labour market: is Italian a liability? *International Journal of Manpower, 19*(7), 520–532.

Guven, C., & Islam, A. (2015). Age at migration, language proficiency, and socioeconomic outcomes: Evidence from Australia. *Demography, 52*, 513–542.

Hall, M., & Farkas, G. (2008). Does human capital raise earnings for immigrants in the low-skill labor market? *Demography, 45*(3), 619–639.

Hazans, M. (2007). Looking for the workforce: the elderly, discouraged workers, minorities, and students in the Baltic labour markets. *Empirica, 34*, 319–349.

Head, K., & Mayer, T. (2014). Gravity equations: workhorse, toolkit and cookbook. In Gopinath, G., Helpman, E., Rogoff, K. (Eds.). *Handbook of International Economics* (vol. 4, 131–195). Retrieved from https://doi.org/10.1016/B978-0-444-54314-1.00003-3

Helemäe, J., & Saar, E. (2015). Estonia: Visible inequalities, silenced class relations. *East European Politics and Societies and Cultures, 29*(3), 565–576. doi:10.1177/0888325415604907

Hellerstein, J., & Neumark, D. (2003). Ethnicity, language, and workplace segregation: Evidence from a new matched employer-employee data set. *Annales d'Economie et de Statistique, 71/72*, 19–78. doi:10.2307/20079047

Hocevar, T. (1975). Equilibria in linguistic minority markets. *Kyklos, 28*(2), 337–357.

HTM. (2015). *Analüüs ja ettepanekud eesti keele õppe tõhustamiseks põhikoolis*. Retrieved from https://www.hm.ee/sites/default/files/analuus_ja_ettepanekud_eesti_keele_oppe_tohustamiseks_pohikoolis.pdf

Hwang, S.-S., Xi, J., & Cao, J. (2010). The conditional relationship between English language proficiency and earnings among US immigrants. *Ethnic and Racial Studies, 33*(9), 1620–1647. doi:10.1080/01419871003642375

Islam, N. (1995). Growth empirics: a panel data approach. *Quarterly Journal of Economics, 110*, 1127–1170.

Isphording, I. (2013). *Returns to local and foreign language skills: causal evidence from Spain*. Retrieved from https://EconPapers.repec.org/RePEc:zbw:rwirep:398

Isphording, I., & Sinning, M. (2012). *The returns to language skills in the US labor marke.*, Research report 7080, Bonn: Institute of Labor Economics. Retrieved from https://www.iza.org/publications/dp/7080/the-returns-to-language-skills-in-the-us-labor-market

Jaumotte, F., Koloskova, K., & Saxena, S. (2016). *Impact of migration on income levels in advanced economies*. Spillover note, International Monetary Fund. Retrieved from https://www.imf.org/en/Publications/Spillover-Notes/Issues/2016/12/31/Impact-of-Migration-on-Income-Levels-in-Advanced-Economies-44343

Kahanec, M. (2007). *Ethnic competition and specialization*. IZA Discussion Papers no. 3167, Institute for the Study of Labor (IZA), Retrieved from https://ideas.repec.org/p/iza/izadps/dp3167.html

Kaljund, A. L. (2018). Orientalism against empire: The paradox of postcoloniality in Estonia. *Anthropological Quarterly, 91*(2), 749–770.

Kim, J. (2003). Education, English language proficiency, and earnings of male immigrants in the US labor market. *Journal of Business & Economics Research, 1*(3), 17–26.

Kim, M., Liu, A. H., Tuxhorn, K.-L., & Brown, D. S. (2015). Lingua mercatoria: Language and foreign direct investment. *International Studies Quarterly, 59*(2), 330–343.

Klamer, A. (2003) A pragmatic view on values in economics, *Journal of Economic Methodology, 10*(2), 191–212.

Klein, C. (2004). *La valorisation des compétences linguistiques : importance du sexe et/ou du statut professionnel?* Paper presented at the 11e Journées d'étude sur les données longitudinales, Dijon.

Korts, K. (2009). Inter-ethnic attitudes and contacts between ethnic groups in Estonia. *Journal of Baltic Studies, 40*(1), 121–137.

Kroncke, C., & Smith, K. (1999). The wage effects of ethnicity in Estonia. *Economics of Transition, 7*(1), 179–199.

Lamberton, D. M. (2002). *The economics of language*. Vol. 150, International Library of Critical Writings in Economics series. Cheltenham: Edward Elgar Publishing.

Lang, K. (1986). A language theory of discrimination. *The Quarterly Journal of Economics, 101*(2), 363–382.

Lehmer, F., & Ludsteck, J. (2011). The immigrant wage gap in Germany: Are East Europeans worse off? *International Migration Review, 45*(4), 872–906.

Leping, K.-O., & Toomet, O. (2008). Emerging ethnic wage gap: Estonia during political and economic transition. *Journal of Comparative Economics, 36*(4), 599–619.

Leppik, M. (2020). *The segmented integration and mediated transnationalism of Estonian Russian-speaking populations.* Institute of Social Studies, Tartu: University of Tartu Press.

Leppik, M., & Vihalemm, T. (2015). The paradox of national language acquisition: Russian speakers' labor market positions in Estonia. *Journal of Baltic Studies, 46*(4), 471–496.

Leslie, D., & Lindley, J. (2001). The impact of language ability on employment and earnings of Britain's ethnic communities. *Economica, 68*(272), 587–606.

Levanon, A. (2014). Who succeeds as an immigrant? Effects of ethnic community resources and external conditions on earnings attainment. *Research in Social Stratification and Mobility, 36*, 13–29.

Levinsohn, J. (2007). Globalization and the returns to speaking English in South Africa. In A. Harrison (Ed.), *Globalization and Poverty* (pp. 629–646): University of Chicago Press.

Lewis, E. G. (2011). *Immigrant-native substitutability: The role of language ability*. Research report 17609. Cambridge: NBER Working Paper Series. Retrieved from https://www.nber.org/papers/w17609

Lindemann, K., & Saar, E. (2011). Ethnic inequalities in education. In J. Helemae & R. Võõrmann (Eds.), *The Russian second generation in Tallinn and Kohtla-Järve: The TIES study in Estonia* (pp. 59–91). Amsterdam: Amsterdam University Press.

Lindley, J. (2002). The English language fluency and earnings of ethnic minorities in Britain. *Scottish Journal of Political Economy, 49*(4), 467–487.

Longhi, S., Nicoletti, C., & Platt, L. (2013). Explained and unexplained wage gaps across the main ethno-religious groups in Great Britain. *Oxford Economic Papers, 65*(2), 471–493.

Lucas, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics, 22*(1), 3–42. doi:10.1016/0304-3932(88)90168-7

Mankiw, N. G., Romer, D., & Weil, D. N. (1992). A contribution to the empirics of economic growth. *Quarterly Journal of Economics, 107*(2), 407–437.

Marschak, J. (1965). The economics of language. *Behavioral Science, 10*(2), 135–140.

Masso, J., & Krillo, K. (2010). Estonia, Latvia and Lithuania: Minimum wages in a context of migration and labour shortages. In D. Vaughan-Whitehead (Ed.), *The Minimum Wage Revisited in the Enlarged EU* (pp. 113–151). UK: Edward Elgar Publishing.

Masso, J., Meriküll, J. & Vahter, P. (2020). *The role of firms in the gender wage gap*. University of Tartu - Faculty of Economics and Business Administration Working Paper Series 120. Retrieved from https://ideas.repec.org/p/mtk/febawb/120.html

Mavisakalyan, A. (2017). Returns to language skills in transition economies. *IZA World of Labor, 416*. doi:10.15185/izawol.416

May, S. (2016). Globalization, language(s), and mobility. In M. Gazzola & B.-A. Wickstrom (Eds.), *The economics of language policy* (pp. 383–401). Boston: MITT Press.

McManus, W., Gould, W., & Welch, F. (1983). Earnings of Hispanic men: The role of English language proficiency. *Journal of Labor Economics, 1*(2), 101–130.

McManus, W. (1990). Labor market effects of language enclaves: Hispanic men in the United States. *The Journal of Human Resources, 25*(2), 228–252. doi:10.2307/145755

Melitz, J. (2008). Language and foreign trade. *European Economic Review, 52*(4), 667–699.

Melitz, J., & Toubal, F. (2014). Native language, spoken language, translation and trade. *Journal of International Economics, 93*(2), 351–363.

Meriküll, J., Kukk, M., & Rõõm, T. (2020). *What explains the gender gap in wealth? Evidence from administrative data*. National Bureau of Economic Research (NBER) Working Paper Series. Retrieved from https://www.nber.org/papers/w26920

Meriküll, J., & Mõtsmees, P. (2017). Do you get what you ask? The gender gap in desired and realised wages. *International Journal of Manpower, 38*(6), 893–908.

Mincer, J. (1958). Investment in human capital and personal income distribution. *Journal of Political Economy, 66*(4), 281–302. doi:10.1086/258055

Mincer, J. (1974). *Schooling, experience and earnings.* New York: Columbia University Press.

Miranda, A., & Zhu, Y. (2013a). English deficiency and the native-immigrant wage gap. *Economics Letters, 118*(1), 38–41.

Miranda, A., & Zhu, Y. (2013b). The causal effect of deficiency at English on female immigrants' labor market outcomes in the UK. *IZA Discussion Paper No. 7841.*

Modood, T. (2005). A Defence of multiculturalism. *Soundings: a journal of politics and culture, 29*(1), 62–71. doi:10.3898/136266205820466869

Mufwene, S. S. (2010). Globalization and the spread of English: What does it mean to be anglophone? *English Today, 26*(1), 57–59. doi:10.1017/S0266078409990605

Mägi, K. (2018). *Ethnic residential segregation and integration of the Russian-speaking population in Estonia.* Dissertationes Geographicae Universitatis Tartuensis, 69. University of Tartu Press.

Männasoo, K., Ridala, S. (2021). Big language minority and small language majority: language skills and unemployment in Latvia. [unpublished]

OECD (2013). OECD skills outlook 2013: First results from the Survey of Adult Skills. Paris: Publishing. doi:10.1787/9789264204256-en

OECD (2018). *Education Policy Outlook 2018: Putting Student Learning at the Centre*. Retrieved from https://www.hm.ee/sites/default/files/epo2018_country_snapshots.pdf

Pajumets, M. (2019). *Pathways to citizenship for third-country nationals in the EU Member States: Estonian national report*. Tallinn: European Migration Network Focussed study 2019. Retrieved from https://ec.europa.eu/home-affairs/sites/homeaffairs/files/08_estonia_citizenship_study_en.pdf

Pettai, V. A. (2006). Explaining ethnic politics in the Baltic states: Reviewing the triadic nexus model. *Journal of Baltic Studies, 37*(1), 24–36.

Pohla, T., Kaldur, K., & Kivistik, K. (2016). *Eesti tööturu rahvuslik ja keeleline jaotus: meta-analüüs*. Retrieved from: https://www.ibs.ee/wp-content/uploads/Eesti-t%C3%B6%C3%B6turu-rahvuslik-ja-keeleline-jaotus.pdf

Pool, J. (1972). National development and language diversity. In *Advances in the Sociology of Language* (Vol. II). The Hague: Mouton.

Raynauld, A., & Marion, G. (1972). Une analyse économique de la disparité inter-ethnique des revenus. *Revue économique, 23*(1), 1–19.

Razin, A. (1972). Investment in human capital and economic growth. *Metroeconomica, 24*(2), 101–116.

Reimers, C. (1983). Labor market discrimination against Hispanic and Black men. *Review of Economics and Statistics 65*(4), 570–579.

Reimers, C. (1985). A comparative analysis of the wages of Hispanics, Blacks and non-Hispanic Whites. In G. Borjas & M. Tienda (Eds.), *Hispanics in the U.S. Labor Market* (pp. 27–76). Orlando: Academic Press.

Reinholde, I. (2004). Challenges for Latvian public administration in the European integration process. In A. Dimitrova (Ed.), *Driven to change: The European Union's enlargement viewed from the East.* (pp. 163–178). Manchester: Manchester University Press.

Rendon, S. (2007). The Catalan premium: Language and employment in Catalonia. *Journal of Population Economics, 20*(3), 669–686.

Ridala, S., & Toomet, O.-S. (2019). Wage gap in ethnically segmented labour market: The role of cognitive skills. *Eastern European Economics, 57*(1), 20–30. doi:10.1080/00128775.2018.1524713

Ridala, S. (2021). Language skills and labour market returns: A meta-regression analysis. *Language Problems and Language Planning*, 44(2), 200–241. doi: https://doi.org/10.1075/lplp.20012.rid

Rubinstein, A. (2000). *Economics and language: Five essays*. Cambridge: Cambridge University Press.

Rubinstein, Y., & Weiss, Y. (2006). Post schooling wage growth: Investment, search and learnin. In E. Hanushek & F. Welch (Eds.), *Handbook of the Economics of Education* (Vol. 1, pp. 1–67). Elsevier.

Rudi, H. (2019). *Eestis võib inglise keelt oskamata poes või restoranis hätta jääda*. Retrieved from https://www.err.ee/957812/eestis-voib-poes-voi-soogikohas-sattuda-umbkeelse-teenindaja-otsa

Rõõm, T. (2003). *Reservation wages in Estonia*. Retrieved from https://www.eestipank.ee/en/publication/working-papers/2003/12003-marit-room-reservation-wages-estonia

Rõõm, T. (2007). *Haridus ja tööturg Eestis*. Eesti Panga Toimetised (12). Retrieved from https://www.eestipank.ee/publikatsioon/toimetised/2007/122007-tairi-room-haridus-ja-tooturg-eestis

Saarela, J., & Finnäs, F. (2003). Unemployment and native language: The Finnish case. *The Journal of Socio-Economics, 32*(1), 59–80.

Saiz, A., & Zoido, E. (2005). Listening to what the world says: Bilingualism and earnings in the United States. *The Review of Economics and Statistics, 87*(3), 523–538.

Schafgans, M. A. (1998). Ethnic wage differences in Malaysia: Parametric and semiparametric estimation of the Chinese–Malay wage gap. *Journal of Applied Econometrics, 13*(5), 481–504.

Schultz, T. W. (1962). Investment in human beings. *Special Supplement to Journal of Political Economy, 70*(5), 1–157.

Shields, M. A., & Price, S. W. (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.

Stöhr, T. (2015). The returns to occupational foreign language use: Evidence from Germany. *Labour Economics, 32*(C), 86–98.

Tainer, E. (1988). English language proficiency and earnings among foreign born men. *Journal of Human Resources, 23*(1), 108–122.

Tammaru, T. (1997). Venelaste kohanemise regionaalsed erisused ja kohanemist mõjutavad tegurid *Akadeemia, 9*(11), 2283–2302.

Tammaru, T. (1999). *Venelased Eestis: ränne ja kohanemine*. Tallinn: Siseakadeemia kirjastus.

Tammaru, T., & Kulu, H. (2003). The ethnic minorities of Estonia: Changing size, location, and composition. *Eurasian Geography and Economics, 44*(2), 105–120.

Toomet, O. (2011). Learn English, not the local language! Ethnic Russians in the Baltic States. *American Economic Review, 101*(3), 526–531.

Trejo, S. J. (2001). *Intergenerational progress of Mexican-origin workers in the U.S. labor market*. IZA Institute of Labor Economics, DP No. 377. Retrieved from https://www.iza.org/publications/dp/377/intergenerational-progress-of-mexican-origin-workers-in-the-us-labor-market

Uzawa, H. (1965). Optimum technical change in an aggregative model of economic growth. *International Economic Review, 6*(1), 18–31.

Vaillancourt, F. (1980). *Differences in earnings by language groups in Quebec , 1970: An economic analysis.* International Center for Research on Bilingualism. Retrieved from Quebec: https://eric.ed.gov/?id=ED196262

Veebel, V., & Loik, R. (2012). Estonia: Life in post-communist Eastern Europe after EU membership. In D. O'Beachain (Ed.), *Life in Post-Communist Europe after EU Membership* (pp. 163-184). London: Routledge, Taylor & Francis Group.

Vetik, R., Lauristin, M., Helemäe, J., Korts, K., Kruusvall, J., Nimmerfeldt, G., . . . Vihalemm, T. (2008). *Integration Monitoring of the Estonian society*. Integratsiooni Sihtasutus ja Rahvastikuministri Büroo. Retrieved from: https://www.integratsioon.ee/sites/default/files/196_207.pdf

Vihalemm, T., & Hogan-Brun, G. (2013). Language policies and practices across the Baltic: processes, challenges and prospects. *European Journal of Applied Linguistics, 1*(1), 55–82. doi:10.1515/eujal-2013-0004

Walsh, J. R. (1935). Capital concept applied to man. *The Quarterly Journal of Economics, 49*(2), 255–285.

Wang, H., Smyth, R., & Cheng, Z. (2017). The economic returns to proficiency in English in China. *China Economic Review, 43*(C), 91–104.

Williams, D. R. (2011). Multiple language usage and earnings in western Europe. *International Journal of Manpower, 32*(4), 372–393.

Yao, Y., & van Ours, J. C. (2015). Language skills and labor market performance of immigrants in the Netherlands. *Labour Economics, 34*, 76–85.

Zhang, W., & Grenier, G. (2013). How can language be linked to economics? A survey of two strands of research. *Language Problems and Language Planning, 37*(3), 203–226.

Zhen, Y. (2015). English Proficiency and Earnings of Foreign-Born Immigrants in the USA from 1980 to 2000: The Effect of Minority-Language Enclave. *Forum for Social Economics, 45*(4), 329–349. doi: 10.1080/07360932.2015.1026920

Acknowledgements

I would like sincerely to thank:

- Professor Kadri Männasoo, the thesis supervisor, for her support and valuable comments and her help with the thesis, especially paper III and Introduction of the thesis;
- My co-supervisor, Ott-Siim Toomet, for his inspiration on the research topic, help with data collection and feedback, valuable discussions throughout, and co-authoring of the first two papers;
- My co-supervisor, Associate Professor Ants Aasma, for his support and for helping me to focus on writing the thesis;
- Co-author Sven Kristjan Bormann for his valuable contribution to our paper and support during the PhD journey;
- All the funding institutions that have provided support during my studies: Tallinn University of Technology; Estonian Research Council, project ETF8627; Doctoral School in Economics and Innovation, supported by the European Union European Regional Development Fund (Tallinn University of Technology ASTRA project "TTÜ Development Program 2016–2022," project code: 2014-2020.4.01.16-0032); the Erasmus Programme of the European Union Project No 611059-EPP-1-2019-1-EE-EPPJMO-MODULE; the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 952574; European Economic Area (EEA) Financial Mechanism 2014-2021, Baltic Research Programme Project No S-BMT-21-8 (LT08-2-LMT-K-01-073); SA Archimedes DoRa T6 grant;
- The participants in seminars: professors, staff and doctoral students at Tallinn University of Technology for their useful suggestions and help in improving the papers;
- The participants in the seminars of Real Estate and Construction Management and Centre for Banking and Finance in KTH for their valuable comments for the first paper and the Department of Industrial Economics and Management in KTH, especially Vladimir Kutcherov, for giving me space for work and valuable comments for my second paper of the thesis;
- Robin Hazlehurst, who did excellent language editing of Paper III and Introduction of the thesis;
- My mother, who encourages me, but unfortunately left us too early and with who we shared four years fighting against cancer;
- My family, my father and especially my husband Märt, my daughters Anna-Liisa and Eveliina and my son Emil for their support during my doctoral studies;
- All my friends who supported me during my doctoral studies.

Abstract
Essays on language skills and labour market outcomes

The current thesis *Essays on language skills and labour market outcomes* integrates three self-contained essays from the field of economics of language. In the context of the globalisation of the world economy and labour markets, increased migration and diversity in cultures, languages and native identities, there is growing interest in understanding the mechanisms that could improve the labour market outcomes of individuals through their language skills, especially in a second language. Language skills are an important determinant of labour market outcomes not only for immigrant populations or only for linguistic or national minorities, but for all individuals who gain in their professional endeavours from knowing a second language. Language skills are an integral part of human capital and these skills improve both the individual's outlook for employment and their productivity when they hold a job that requires them to know a second language. This thesis provides both original empirical evidence and a meta-regression synthesis of the existing empirical literature about the role that language skills play in labour market outcomes for both language majority and minority populations in a given socio-linguistic, economic and institutional context.

The first essay *Language Skills in an Ethnically Segmented Labour market: Estonia 1989–2012* conducts a comprehensive study that estimates how language skills impact labour market outcomes in Estonia, a multilingual society with two major local languages, Estonian and Russian. The empirical analysis uses data from the Estonian Labour Force Survey (ELFS). The unusual historical and cultural background of Estonia means that the Estonian context allows the impact of the language skills to be investigated in multiple dimensions. Firstly, the study looks at how knowing Estonian, the local state language, affects the labour market outcomes of the mostly Russian-speaking minority. Secondly, it studies how knowing Russian, a major second language in Estonia and a powerful language in the world, affects native Estonian workers in the local labour market. Thirdly, it estimates how knowing English, the most powerful global language and a leading business language in the world, affects the labour market outcomes of both native Estonians and Russian speakers in Estonia. The results are quite mixed. Fluency in Estonian is related to a reduction in unemployment of approximately 5 percentage points for men, but to virtually no income premium for them. Hence the study calls for further investigation of why Russian men do not receive a wage premium from fluency in Estonian, but only reduce their risk of unemployment. In contrast, Russian-speaking women who are fluent in Estonian not only reduce their risk of unemployment, but also earn approximately 10% more. English skills are associated with a substantial income premium for both language groups and genders but are virtually unrelated to the probability of employment. Russian skills are associated with higher wages for native Estonians, but this wage premium has been dropping over time. Knowledge of Russian and the probability of employment remains quite unrelated for native Estonians. The study contributes by providing new evidence on the role of language skills for both genders and documents the dynamics of the effects of the return to language in the highly linguistically diverse labour market in Estonia.

The second essay, *Wage Gap in an Ethnically Segmented Labour market: The role of Cognitive Skills*, extends the understanding of the income disparities between native Estonian men and men from the local Russian-speaking minority in Estonia. Using data from the ELFS and from the Survey of Adult Skills, the latter collected within the

Programme for the International Assessment of Adult Competencies (PIAAC), the paper analyses the wage differentials between men from the Russian-speaking minority and native Estonian men. The Oaxaca-Blinder decomposition tests whether the wage gap observed between native and non-native Estonian-speaking men is related to commonly observed skills and human capital characteristics or not. The findings show that even once several measures of cognitive skills for literacy, numeracy, and problem-solving in technology-rich environments (PS-TRE) are introduced on top of customary human capital controls such as education and work experience, 80% of the wage differentials remain unexplained. This essay concludes that the most likely explanation for the unexplained wage gap is the height of the barriers to labour market entry, combined with high levels of professional and social segregation of Russian speakers and native Estonians. This research is unique in that it contributes to understanding the role of cognitive skills in the wage equations in the context of the highly linguistically diverse labour market in Estonia.

The third essay is titled *Language skills and labour market returns: A meta-regression analysis*. The article carries out a meta-regression analysis that studies the link between the returns to language skills as estimated by the bulk of the relevant empirical literature, and the socio-economic, institutional and ethnolinguistic contextual factors and study design factors that may have an effect on estimates of language returns. The meta-regression analysis separated three types of study. Firstly, studies that investigated the language returns from skills in the local state language for immigrant or other national minority populations in the country. Secondly, studies that investigated the language returns of native speakers to knowing the second largest local language in linguistically diverse countries or in countries with two or more official state languages. Thirdly and finally, studies of the returns to skills in a global professional language, mainly English, in countries where these languages are non-native. The meta-regression evidence, along with the socio-linguistic and economic contextual controls and controls for primary study data structure and estimation method, shows that knowing a language that is more influential according to the language power index (LPI) is associated with higher labour market rewards. Linguistic diversity and a multilingual environment decrease the returns to skills in a professional language such as English, but have little effect on returns to a significant local minority language and to the local official state language. Urbanisation has a modest negative effect on the returns to language skills in English or other business languages. The country's income level measured by GDP per capita increases the returns to skills in a local significant minority language or a second state language. A high unemployment rate increases the return to knowing English or another business language as well as the returns to knowing the local state language. In general, the article contributes by casting light on the existing empirical evidence that investigates the relationship between language skills and labour market outcomes in diverse socio-linguistic, economic and institutional contextual settings while controlling for the main study design factors, such as estimation method and underlying data structure. To the best of the knowledge of the author, this is the first meta-study to evaluate and synthesise the heterogeneous empirical literature that estimates how language returns affect labour market outcomes.

All three essays contribute to a new knowledge and understanding of the role of language skills in labour market outcomes. The first two articles provide original empirical evidence for the labour market outcomes as related to the skills individuals in English, Estonian as a state language, and Russian as a major minority language in

Estonia, a small European member state with 1.32 million inhabitants hosting a sizeable Russian-speaking minority that constitutes more than 25% of its total population. The third article carries out a novel meta-study research that extends beyond specific socio-linguistic, economic and institutional contexts and generates new synthesised evidence on language returns across heterogeneous labour markets and studies that investigate them.

Lühikokkuvõte
Esseed keeleoskusest ja tööturu väljunditest

Käesolev väitekiri "Esseed keeleoskusest ja tööturu väljunditest" ühendab kolm iseseisvat teadusartiklit (esseed) keeleökonoomika valdkonnast, mis on 2019. ja 2020. aastal avaldatud rahvusvahelistes eelretsenseeritud *Web of Science* ja *Scopus* indekseeritud teadusajakirjades ning mis vastavad Eesti Teaduse Infosüsteemi teaduspublikatsioonide kõrgeimale kategooriale (ETIS 1.1).

Maailmamajanduse ja tööturgude globaliseerumise, suurenenud rände, peaaegu kõigi kõrgelt arenenud majandusega riikide rahvastiku vananemise ning kultuuri, keele ja identiteedi suureneva põimumise tingimustes kasvab huvi tööturgude toimimist parandavate mehhanismide senisest parema mõistmise vastu. Keeleoskus on inimkapitali lahutamatu osa ja avaldab olulist mõju indiviidi edukusele tööturul. Antud väitekirja eesmärgiks on analüüsida põhirahvuse (riigikeele rääkijad) ning rahvusvähemuse (riigikeelt emakeelena mitterääkijad) elanikkonna keeleoskuse rolli ja mõju tööturu väljunditele – hõive staatusele ja palga kujunemisele. Nimetatud probleemi algupärane empiiriline uurimistöö esimeses ja teises essees tugineb Eesti kui mitmekeelse ühiskonna ja tööturu andmete ja konteksti uurimisele. Kolmas essee tugineb arvestatavale hulgale rahvusvaheliselt avaldatud empiirilisele ja kontekstilt heterogeensele teaduskirjandusele, mis uurib keeleoskuse ja tööturu väljundite seoseid erinevates riikides ja keelekeskkondades.

Eesti unikaalsus keeleuurimuste läbiviimiseks tuleneb Eesti lähiajaloost endise Nõukogude Liidu vabariigina 1940–1991, mis tõi kaasa olulise venekeelse rahvus- ja keelevähemuse tekke ning riigi *de facto* kahekeelseks muutumise. Kuigi seoses Eesti taasiseseisvumisega 1991. aastal kehtestati ainsa riigikeelena eesti keel, on valitsev olukord paljus sarnane mitut riigikeelt omavate riikidega nagu näiteks Kanada, kus kõrvuti elavad prantsuse keelt kõnelev vähemus ning inglise keelt kõnelev enamus. Eesti kujutab endast eriomast sotsiaal-etnilist, keele, kultuuri ja institutsionaalse kujunemise keskkonda, mis paljuski erineb tavapärasest immigratsiooni- ja integratsioonikirjanduses käsitletavast kontekstist. Kirjanduses on valdavalt käsitlemist leidnud immigratsioonikontekstid, mida iseloomustab põliselanikkonnaga võrreldes suhteliselt väiksearvuline ning keelelt ja kultuurilt heterogeenne immigrantide esimene või teine põlvkond. Eestis on seevastu tegemist valdavalt teise ja kolmanda põlvkonna suurearvulise ning homogeense vene keelt rääkiva elanikkonnaga, mis moodustab üle veerandi Eesti kogurahvastikust ning ületab põlisrahvuse, eestlaste, osakaalu mitmetes Eesti piirkondades, eeskätt Venemaaga piirnevas Kirde- ja Ida-Eestis. Vene keelt kõnelev elanikkond valdavalt ei tunneta oma staatust immigrantidena ning nad peavad Eestit oma kodumaaks vaatamata tugevatele säilinud nii kultuurilistele kui perekondlikele sidemetele naaberriigi Venemaaga. Teisalt väärtustab vene keelne elanikkond enda identiteeti ja emakeelt kui mõjukat rahvuste vahelise suhtluse keelt endise Nõukogude Liidu aladel.

Esimeses essees "Keeleoskus etniliselt segmenteeritud tööturul: Eesti 1989–2012" viiakse Eesti tööjõu-uuringutele (ELFS) tuginedes läbi põhjalik uuring keeleoskuse mõjust tööturu väljunditele Eesti mitmekeelses ühiskonnas. Eesti kontekst võimaldab uurida keeleoskuse mõju kolmes dimensioonis. Esiteks, riigikeele oskuse roll teise keelena vene-keelse rahvusvähemuse jaoks Eesti tööturul. Teiseks, vene keele oskuse roll eestlaste kui riigikeele kõnelejate jaoks tööturul arvestades vene keele mõjukust nii Eestis kui iseäranis endise Nõukogude Liidu aladel (globaalse keeleindeksi mõjukuselt kuulub vene

keel kümne rahvusvaheliselt enim räägitava keele hulka). Kolmandaks, inglise keele kui ärikeele roll Eesti tööturul nii eestlastest põlisrahvuse kui kohalike vene keelt emakeelena kõnelejate jaoks. Artiklis kasutatakse harilikku vähimruutude meetodit, samas teaduskirjanduses on rakendatud ka instrumenteeritud vähimruutude meetodit, et korrigeerida mitte-vaadeldavatest muutujatest tingitud hinnangunihkeid ja mõõtmisvigu. Otsus hariliku vähimruutude meetodi kasuks langetati kahel põhjusel. Esiteks, olemasolevate andmete jaoks polnud võimalik leida sobivaid instrumente ning teiseks, valitud meetod hõlbustab võrdlust teiste sama hüpoteesi kontrollivate, uuringutega. Tulemused tööturu väljundite seostest keeleoskusega on väga eripalgelised sõltudes nii rahvusest kui soost. Vene keelt kõnelevate meeste puhul on eesti keele kui riigikeele oskus seotud umbes 5 protsendipunkti võrra madalama töötusega, kuid sissetulekupreemia praktiliselt puudub. Antud leid annab põhjust edasiseks uurimistööks, leidmaks selgust, miks eesti keele oskus, mis küll kahandab töötuks jäämise tõenäosust, ei anna vene keelt kõnelevatele meestele palgalisa. Seevastu vene keelt kõnelevate naiste puhul ei vähenda eesti keele oskus mitte üksnes töötuks jäämise tõenäousust, vaid toob kaasa ka umbes 10 protsendipunkti võrra kõrgema töötasu. Inglise keele oskus tõstab mõlema rahvusgrupi sissetulekut nii meeste kui naiste puhul, kuid seos töötusega puudub. Vene keele oskus tõstis nii eesti rahvusest meeste kui ka naiste palka, ent palgapreemia suurus on vaatlusperioodil kahanenud ning vene keele oskusel sisuliselt puudub seos tööpuudusega. Tänaseni napib uuringuid, mis analüüsiksid, kuidas mõjutab vähemusrahvuse keele oskus põlisrahvuse esindajate töötuse määra ning palga kujunemist. Käesolev uurimus lisab empiirilisi tulemusi antud küsimuse kohta, käsitledes keeloskuse mõju nii põlisrahvusest meeste kui ka naiste tööturu väljunditele. Seejuures on oluline märkida, et uuring analüüsib keeleoskuse mõju tööturu väljunditele dünaamiliselt, tuues välja trende mõju suunas ja suuruses.

Teise essee “Palgalõhe etniliselt segmenteeritud tööturul: kognitiivsete oskuste roll” eesmärgiks on uurida põlis- ja vähemusrahvusest meeste sissetulekute erinevusi ning nende põhjuseid Eestis. Tuginedes tööjõu-uuringu ELFS (Estonian Labour Force Survey) andmetele ja täiskasvanute pädevuste rahvusvahelise hindamise programmi (Programme for the International Assessment of Adult Competencies PIAAC) raames kogutud täiskasvanute oskuste uuringu andmetele analüüsitakse artiklis eesti ja vene keelt emakeelena kõnelevate meeste palgaerinevuste tegureid Oaxaca-Blinderi dekompositsioonimeetodi abil. Oaxaca-Blinderi palgalõhe dekomponeerimine võimaldab mõõta tegureid, mis palgalõhet kirjeldavad, tuues seejuures välja mitte-selgitatava, ehk võimaliku diskrimineerimisega seotud palgaerinevuse suuruse eesti ja vene emakeelega meeste vahel. Tulemused näitavad, et vaatamata mitmete kognitiivsete oskuste kontrollimisele: kirjaoskus, arvutamisoskus ja probleemide lahendamine tehnoloogiarikastes keskkondades (PS-TRE), jääb 80% palgalõhest eesti ja vene emakeelega meeste vahel seletamatuks. Antud essees jõutakse järeldusele, et selgitamata palgalõhe on kõige tõenäolisemalt seotud venekeelsete inimeste tööturule sisenemise barjääridega, madala sotsiaal-kultuurilise integratsioonitasemega ning põlisrahvuse ja venekeelse vähemusrahvuse kõrge ametialase ning sotsiaalse segregatsiooniga. Kuna Eestile eriomaseid sotsiaal-etnilisi ja institutsionaalseid tingimusi käsitlevaid uurimusi napib, siis antud essee lisab uudseid ja ainulaadseid leide valdkonna teaduskirjandusse.

Kolmanda essee pealkiri on “Keeleoskus ja tööturu väljundid: meta-regressioonanalüüs”. Artikkel keskendub meta-regressioonanalüüsile ning hindab tööturu väljundite, ehk palga kujunemise seost sotsiaalmajanduslike, institutsionaalsete

ja etno-lingvistiliste teguritega. Meta-regressioonanalüüs jaotab varasemad uuringud kolme gruppi. Esiteks, uuringud, mis keskenduvad immigrantide või etnilise vähemusrahvuse riigikeele oskuse ning tööturu väljundite vahelise seose hindamisele. Teiseks, uuringud, mis keskenduvad põlisrahvuse (riigikeele rääkijate) vähemuskeele oskuse seosele palgaga. Kolmandaks, uuringud, mis mõõdavad ärikeele, ehk valdavalt inglise keele, oskuse mõju palga kujunemisele. Meta-regressioonuuring tugineb 61 varasema empiirilise artikli tulemustele, mis on publitseeritud ajavahemikus 1998–2019 ning baseeruvad Majanduskoostöö ja Arengu Organisatsiooni (OECD) riikide ja OECD võtmepartnerite (nagu Hiina, India ja Lõuna Aafrika) andmetel. Meta-uuringu sõltuvaks muutujaks on keeleoskuse mõju, ehk palgapreemia suurust peegeldav koefitsient varasemates uuringutes. Sõltumatuteks muutujateks on mitmed sotsiaalset, majanduslikku ja keelekonteksti kirjeldavad tegurid, sealhulgas uuritava keele globaalne mõjukus (PLI, "Power Language Index"), keelekeskkonna mitmekesisus ehk riigi mitmekeelsus (LDI, "Linguistic Diversity Index"), globaliseerumine ("KOF Globalisation Index"), linnastumine (Urbanisation rate), töötuse määr ja riigi jõukus ehk sisemajanduse koguprodukt ühe elaniku kohta.

Meta-regressioonanalüüs näitas, et suuremat globaalset mõju omava keele oskus on seotud kõrgema palgapreemiaga. Samuti selgus, et mida mitme-keelsem on riik, seda väiksem on ärikeele oskusega, ehk enamasti inglise keele oskusega, kaasnev tasulisa. Samas ei too keeleline mitmekesisus kaasa kohaliku vähemuskeele oskusega kaasnevat palgapreemiat. Riigikeele oskuse eest saadav preemia on aga positiivselt mõjutatud keelelisest mitmekesisusest, st mida suurem on keeleline mitmekesisus, seda suurem on riigikeele oskusega kaasnev palgapreemia. Veel selgus meta-uuringust, et kõrgem linnastumine mõjutab inglise keele kui ärikeele oskusest saadavat tasupreemiat negatiivselt, kuid nimetatud seose mõju ulatus on tagasihoidlik. Riigi jõukus (SKP elaniku kohta) tõstab vähemusrahvuse keele oskusest saadavat palgapreemiat. Seevastu kõrge tööpuudus tõstab palgapreemiat riigikeele oskajate ning inglise keele kui ärikeele valdajate jaoks.

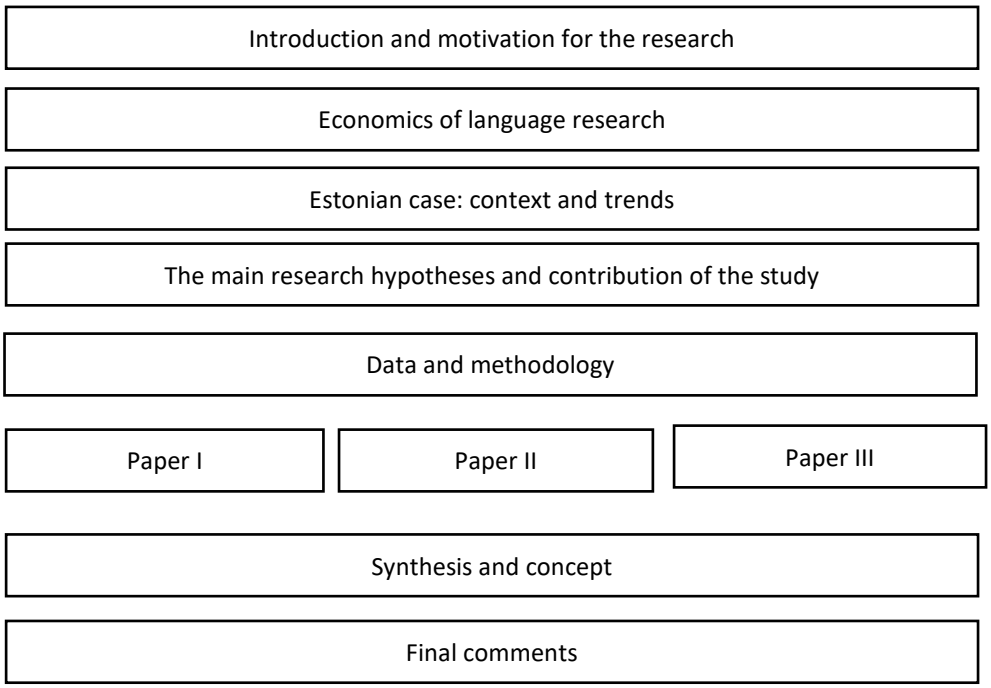
Autorile teadaolevalt on antud meta-regressioonanalüüs esimene püüdlus kirjeldada kontekstuaalsete tegurite, ehk sotsiaalmajanduslike, etnilis-lingvistiliste ning institutsionaalsete faktorite mõju keeleoskuse ja palga kujunemise seosele, kontrollides aluseks olevate uuringute hindamismeetodit ning andmestruktuuri.

Kõik kolm esseed annavad arvestatava panuse uue teadmise loomisse keeleoskuse rollist tööturu väljunditele. Kaks algupärast empiirilist uuringut indiviidide keeleoskusest ning tööturu väljunditest keskenduvad eestlastest põlisrahvuse esindajate ning venekeelsete vähemusrahvuse esindajate kõrvutamisele, uurides keeleoskuse mõju sissetulekutele Euroopa Liidu ühes väikseimas, 1.32 miljonilise elanikkonnaga riigis, mis on koduks suurearvulisele vene emakeelega kogukonnale, mis moodustab üle veerandi riigi elanikkonnast. Kolmas artikkel viib läbi ainulaadse ning esmakordse meta-regressioonuuringu, mis keskendub varasema heterogeense, riigi- ja keelekonteksti- spetsiifilise kirjanduse statistilisele sünteesile tuvastamaks sotsiaal-lingvistiliste, majanduslike ja institutsionaalsete tegurite rolli keeleoskuse ja palga vahelise seose suunale ja tugevusele.

Appendices

Appendix 1. Graphical outline of the study

The current thesis encompasses three separate research papers published by internationally recognized publishers and indexed in Scopus and/or Web of Science databases. The composition of the thesis aims at building a link between individual publications and serves as an umbrella to provide a broader context for the topic of interest – language skills and labour market outcomes.



Appendix 2. Paper I

LANUAGE SKILLS IN AN ETHNICALLY SEGMENTED LABOUR MARKET: ESTONIA 1989-2012

Publication:

Bormann, S.-K.; Ridala, S.; Toomet, O.-S., 2019. Language Skills in an Ethnically Segmented Labour market: Estonia 1989-2012. *International Journal of Manpower*, vol. 40, no. 2, pp. 304–327. DOI: https://doi.org/10.1108/IJM-06-2017-0115 (ETIS 1.1).

Copyright © Emerald Publishing Limited.

Draft as conference proceedengs:

Bormann, S.-K.; Ridala, S.; Toomet, O.-S. 2015. Language Skills in an Ethnically Segmented Labor market: Estonia 1989-2012. Fourth World Conference SOLE – EALE 2015: 4th Joint meeting of SOLE – EALE, Montreal, Canada, 26-28 June 2015. The Society of Labor Economists, 1–16.

Language skills in an ethnically segmented labour market: Estonia 1989 – 2012

Sven-Kristjan Bormann
School of Economics and Business Administration, University of Tartu, Tartu, Estonia

Svetlana Ridala
School of Business and Governance, Tallinn University of Technology, Tallinn, Estonia and

Ott-Siim Toomet
Information School, University of Washington, Seattle, Washington, USA

Abstract

Purpose – The purpose of this paper is to analyse the relationship between skills in the Estonian, Russian and English language, and labour market outcomes in Estonia, a linguistically divided country.

Design/methodology/approach – The authors use the Estonian Labour Force Surveys 1992–2012. The authors rely on multivariate linear regression models to document the relationship between language skills and labour market outcomes.

Findings – Estonian language knowledge (for ethnic Russians) are important determinants of unemployment. Wage, in contrary, is closely related to English skills. Ethnic Russian men do not earn any premium from speaking Estonian, while women, fluent in Estonian earn approximately 10 per cent more. For ethnic Estonians, Russian fluency is associated with a similar income gain.

Research limitations/implications – Due to the observational nature of the data, the effects reported in this study are not causal effects. As a second limitation, the self-reported language skills data may be imprecise and hence the effects the authors report may be too small.

Practical implications – The results stress the role of workplace segregation, both along gender and ethnic lines, in determining the individual labour market experience.

Originality/value – The authors provide a comprehensive overview of the effects of language skills in a rapidly developing labour market in a linguistically divided economy. The authors analyse several languages with different legal status and document long-term trends in the effects.

Keywords: Ethnicity, Estonia, Wage gap, Unemployment, Language skills
JEL classification: J15, J31, J71

This article is © Emerald Publishing Limited and permission has been granted for this version to appear here https://digikogu.taltech.ee/et/Item/94656268-5c7e-4b78-b5df-bf088dea3560. Emerald does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Emerald Publishing Limited. DOI: https://doi.org/10.1108/IJM-06-2017-0115

1 Introduction

Numerous studies document that language skills are related to better labour market outcomes. Speaking different languages may make workers genuinely more productive, either by facilitating communication at the workplace or by making it possible to perform tasks that require language knowledge. Enhanced productivity, in turn, should be reflected both in better wage and in improved employment prospects.

The bulk of the existing literature analyses the effect of knowledge of the local majority language for immigrants. It is well known that language skills are associated with an income premium between 5 and 35 percent (Bleakley & Chin, 2004; Chiswick & Miller, 2010, 1995, 2002, 2015; Dustmann & Fabbri, 2003; Leslie & Lindley, 2001; Shields & Price, 2002). For instance, Chiswick (2008) argues that "proficient" speakers typically earn about 15% more than "not proficient" speakers. The effect works largely through improving educational attainment (Bleakley & Chin, 2004; Rooth & Saarela, 2007). Language skills are also complementary to other types of human capital (Chiswick & Miller, 2007). Immigrants who are fluent in the local ethnic majority language also suffer substantially less from unemployment (Dustmann & Fabbri, 2003; Shields & Price, 2002).

However, in the case of less-than-perfect ethnic integration, the "local" language in ethnic enclaves may be the minority language instead, in this way limiting the usefulness of the majority language skills (Chiswick & Miller, 2002; Hwang, Xi, & Cao, 2010) in particular in context of ethnic segregation and limited access to upper-end jobs (Toomet, 2011). However, the existing evidence suggests that the majority language is an important production factor in the ethnic enclaves as well (Clark & Drinkwater, 2000; Zhou & Logan, 1989).

There is also an increasing body of literature focusing on other languages than the local main business language. These studies generally find that English is a significant predictor of better income (Azam, Chin, & Prakash, 2013; Casale & Posel, 2011; Fabo, Beblavý, & Lenaerts, 2017; Isphording, 2013; Lang & Siniver, 2009; Toomet, 2011; Williams, 2011). The role of other languages, in particular in a multi-lingual economy, is less clear. For instance, Drinkwater and O'Leary (1997) finds a positive effect for Welsh speaking workers in Wales and Armstrong (2015) shows that bilingual native French-speakers enjoy a significant premium both inside and outside of Quebec. François Grin (1998), in contrast, shows that Italian language skills in Swiss non-Italian speaking regions are associated with a lower wage. It remains unclear if this is a true effect or just an omitted variable bias.

Few previous studies have provided a comprehensive analysis of a labour market from the viewpoint of language skills. Bellante and Kogut (1998); Chiswick and Miller (2010, 1995, 2002, 2007) analyse the effects of dominant language fluency for earnings among immigrants in the United States, Chiswick and Miller (1995) also do this for Canada, Australia and Israel. Leslie and Lindley (2001) analyse both employment and income but focus on a single year and English language only. Yao and van Ours (2015) only analyse the role of Dutch language on labour market performance in the Netherlands.

The literature on Central Eastern Europe (CEE) and former Soviet Union countries is rather limited. We know that the different development trajectories have led to disparities both favouring the Russian-speaking minority in Ukraine (Constant, Kahanec, & Zimmermann, 2011) and hindering them in Estonia (Leping & Toomet, 2008). Kahanec and Zaiceva (2009) show that non-citizens suffer less favourable outcomes, in particular, ethnic Russians in Estonia and Latvia. Duncan and Mavisakalyan (2015) show that Russian language proficiency

is valuable in the former USSR republic in the Caucasus (Armenia, Azerbaijan and Georgia) by increasing employment likelihood between six and nine percentage points. Toomet (2011), however, fails to find any income premium from the majority language fluency in Estonia and Latvia. This outcome is corroborated by Lindemann (2013) who shows that Estonian fluency is associated with faster job finding rate while it does not explain the lower occupational prestige of Russian speakers. A number of these studies have to overcome severe data limitations, such as lack of fine-grained information on ethnic background in EU-wide surveys (Kahanec, Zaiceva, & Zimmermann, 2010).

In this paper, we analyse the relationship between both majority and minority language skills, and English for both wage and unemployment. We analyse Estonia, a former Soviet republic and now a rapidly growing and transforming Easter European country. It is characterized by a stark linguistic division where the Russian-speaking community forms almost 30% of the population. We use the Estonian Labour Force Survey, one of the higher-quality long-term datasets, well suitable for our purpose. We provide long–term trends from the last years of the planned economy in 1989 till the post-recession time in 2012.

We find that the Estonian language, the ethnic majority and official language of the country, is associated with a very small albeit slowly increasing wage premium over the years. Interestingly, this is true only for women, Russian men enjoy virtually no premium from the majority language skills. Russian skills, on the other hand, are clearly related to higher wage but the relationship is slowly fading. English fluency was associated with a very large wage premium during the early years of economic transition, later its importance declined but remained sizeable. In contrast, the sole language that plays a role in determining unemployment probability is Estonian. Its fluency is associated with approximately 5 percentage points lower probability of being unemployed. These results suggest that the Estonian labour market is segmented both along gender and the along ethnic background, with men more likely occupied in less communication-intense jobs.

The paper continues as follows: The next section provides a brief background of the Estonian society and the labour market. In the following section, we discuss the data and provide some descriptive statistics. In Section 4 we describe the econometric approach, sections 5 and 6 present the results for unemployment probability and wages. The two last sections, 7 and 8, are devoted to an additional analysis of labour market segregation, and concluding discussion.

2 Background: Ethnic Groups and Languages in Estonia

This section briefly reviews the institutional and economic background in Estonia, a country of 1.3M inhabitants, focusing on the role of ethnic groups and languages. See Leping and Toomet (2008) and Lindemann and Saar (2011) for more detailed information.

Before the Second World War, the country was ethnically rather homogeneous with 90 percent of the population being ethnic Estonians. After the War, the Russian-speaking population grew rapidly through migration and reached 40% by 1989. The migrants largely settled in Ida-Virumaa, a mineral-rich industrial area in North-East of the country, and in the metropolitan area near the capital city Tallinn. The Russian-speakers are still very unevenly distributed.

Such an influx of a mainly Russian-speaking population led to two *de facto* official languages by 1970s. The society was largely divided across language, and the division was

also reflected at workplace (Leppik & Vihalemm, 2015). Certain sectors in the economy, such as the armed forces, railways and the merchant fleet were dominated by Russian-speaking workers while more local industries, such as agriculture and public administration, were dominated by ethnic Estonians. Both communities followed the media in their own respective language, the universities taught many subjects in both Russian and Estonian, and service sector was largely bilingual. However, the bilingualism was mainly one-sided as the bulk of Estonians spoke Russian, but not the other way around. A particular institution in parts of the Soviet Union was a segregated school system, where Russian schools existed along the schools teaching in the local language. Hence, a common education system, one of the central integration mechanisms in the developed world, was (and still largely is) missing in Estonia. The widening use of Russian also caused increasing concerns about the future of the Estonian language, in turn leading to an unwillingness to use Russian by the ethnic Estonians. This further strengthened the language-based segregation.

When the central Soviet authority weakened during the late 1980s, the Estonian-speaking community was quick to organize and established a nation-state in 1991 where Estonian was granted status as the sole official language. The above-mentioned "one-sided bilingualism" started turning around. The universities quickly moved to Estonian-only curriculum, the Estonian language became a compulsory subject in Russian schools while teaching Russian in Estonian-language schools declined. These trends have continued until today, according to HTM (2015), one quarter of the students from non-Estonian-speaking families studied in Estonian schools. As a result, the younger generation of ethnic Estonians is increasingly less fluent in Russian while the ethnic Russians are steadily getting better in Estonian.

The Estonian economy, a part of the Soviet economy up to 1991, successfully re-oriented to the Western markets and has been growing rapidly since the mid-1990s. Figure 1 shows that it grew averaged at approximately 7% annually before the financial crisis in 2008, afterwards the figure has been around 3%. As a by-product of the re-adjustment, English skills become extremely valuable while the supply was still rather limited during the early years of the transition.

The economic and political reforms led to substantial disparities in income and education opportunities in the early 1990s (Leping & Toomet, 2008; Lindemann & Saar, 2011). Inter-ethnic contacts are largely limited to superficial encounters at work and in public space, the number of close private ties remain limited (Korts, 2009). The separate worlds are also reflected in media which may present quite different and occasionally antagonistic viewpoints depending on the language (Korts & Kõuts, 2002).

3 Data

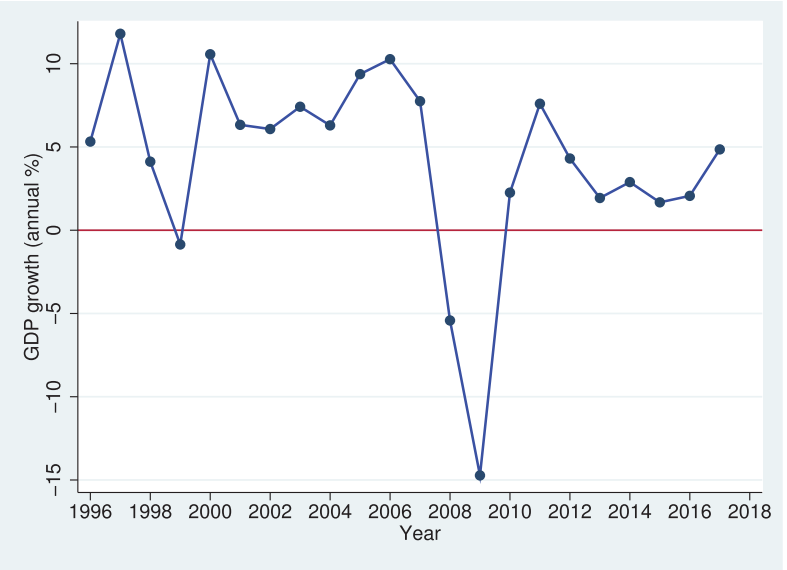
3.1 Variables and Sample Selection

Our main analysis relies on the Estonian Labour Force Survey (ELFS), first conducted in 1995 by the Estonian Statistical Office. We limit ourselves to the years 2000–2012 in the main analysis, motivated by a methodology break: Until 1999 the survey was conducted as an annual cross-section and later as a quarterly rotating panel. The survey contains data on approximately 4000 distinct individuals annually, however, due to its rotating nature, we have roughly 16,000 annual observations for each year since 2000. For the long-term analysis, we also use the earlier waves. These waves also include a retrospective section that contains

income up to several years back.

We limit ourselves to men and women in the main working age, between 25 and 55 years old. Further, we only analyse individuals who are either working or actively looking for a job. The dataset allows us to control for standard personal characteristics and human capital variables, such as age, education and family status. We discuss the most important variables below.

Figure 1: Annual GDP growth in Estonia



Source: World Bank - World Development Indicators

In all waves of the survey, respondents are asked about their "ethnic nationality". Typically, individuals have a single ethnic identity. The answers are coded as "Estonian" or "non–Estonian". As most of those who do not consider themselves Estonian use Russian as their primary language, we refer to this group as "Russians".

The labour market outcomes we analyse are income, the "last salary on the main job", and labour force status (working or unemployed).

The survey includes self-reported language skills on languages used at home and all other languages they understand. The respondents are given some guidance in assessing their skills, they can report either no knowledge, understanding only, speaking but not writing, writing, or using the language at home. We consolidate this information into a dummy variable, equal to one if one speaks or writes, or uses that language at home, and zero otherwise. Hence, we disregard passive understanding only.

For further evidence on ethnic and gender segregation at the workplace, we use data from 2008 Integration Monitor Survey. The survey was conducted by SaarPoll and focuses on ethnic relations in Estonia. The sample size is 1500. 500 of the respondents consider themselves not ethnic Estonians. We analyse two questions: "which is the language of communication at your workplace?" and "which languages do you need at work". The first question addresses

the ethnic segregation while the second focuses on the need for communication.

3.2 Descriptive Statistics

The average wages, language skills and other explanatory variables together with the relative frequencies for the period 2000–2012 are presented in Table 1. We report the data separately for ethnic nationalities (Estonians and Russians) and language skills (Estonian, Russian and English).

The table indicates that there are substantial disparities between Estonians and Russians in wage and unemployment. For instance, the average monthly salary[1] of Estonian males who are not fluent in English is 532 € while for Russians it is only 452 €. However, we also see that Russians who speak English earn approximately 40% more than those who do not (452 € versus 326 € monthly for men and 307 € versus 221 € for women). Estonian skills are associated with lower income premium, 15% for men (378 € versus 329 € monthly) and 30% for women (260 € versus 203 €). Analogously, English language skills are more closely associated with income for Estonians compared to Russian language skills. In contrast, unemployment shows a different picture. For ethnic Russians, it is Estonian language fluency that is more closely associated with low unemployment, those fluent in Estonian have 11.2 percent unemployment rate while for those not fluent it is 17.3 percent. The corresponding numbers for English are 11.2 and 15.1. For Estonians, English is more closely associated with unemployment than Russian: Estonians fluent in English have on average 5.0 percent unemployment while those who are not fluent have 8.8 percent. The difference related to Russian fluency is only 1.6 percentage points. However, these correlations are not easy to interpret as besides of the strong time trends, English skills are also related to education, metropolitan residence, and general ability.

The language fluency variables clearly show that ethnic identity is closely related to language use: more than 99% of Estonians report being fluent in Estonian, the figure for Russians is similar. We also see that Estonians are better at English than Russians, and they are also better at Russian than vice versa.

Average age reflects the trend of younger generations being more fluent in English and Estonian while older Estonians being better in Russian. The family status–related variables correspond to the respective age distribution. Not surprisingly, inter-ethnic households include a disproportionally large share of Russian-speaking Estonians and Estonian-speaking Russians. The average experience follows the same pattern as age. Public sector jobs are associated with a disproportionally large share of Estonian fluency among Russian workers, probably reflecting the formal language requirements and little room for ethnic segregation in that sector.

In conclusion, the descriptive analysis suggests that for Russians, English skills are a more important driver for income than Estonian skills while the opposite is true for unemployment. The labour market outcomes of Estonians are substantially better correlated with English skills than with Russian skills.

[1] When interpreting these figures, one should keep in mind that during this period, 2000-2012, the average monthly salary grow by 180%, from 314 € to 887 € while the language knowledge changed rapidly as well. However, the ethnic composition of the workforce remained roughly constant.

Table 1: Averages and relative frequencies of selected explanatory variables 2000 – 2012

	Estonians				Russians			
Variable	RUS 0	RUS 1	ENG 0	ENG 1	EST 0	EST 1	ENG 0	ENG 1
Wage, men (EUR)	543.86	537.22	438.53	738.63	401.59	489.96	388.91	660.39
Wage, women (EUR)	345.51	376.98	301.13	488.83	238.25	338.62	266.87	435.60
Unemployment (%)	8.74	7.13	8.83	4.99	17.30	11.18	15.06	11.23
EST 1 (%)	99.97	99.58	99.50	99.95	0.00	100.00	40.84	80.97
RUS 1 (%)	0.00	100.00	77.35	82.86	99.90	99.04	99.65	98.77
ENG 1 (%)	29.47	37.16	0.00	100.00	6.70	30.69	0.00	100.00
College Degree (%)	15.61	23.44	10.77	41.38	12.66	28.64	13.93	49.16
Married (%)	40.09	55.71	54.11	49.53	63.54	60.78	63.38	56.99
Inter-ethnic household (%)	1.34	5.20	4.79	3.70	5.07	19.46	11.97	12.17
Age	37.05	41.83	42.39	38.05	42.14	40.54	42.15	37.83
Kids (number)	1.08	1.00	1.00	1.00	0.71	0.76	0.75	0.68
Residence County								
Metropolitan (%)	17.74	22.58	14.03	35.25	35.43	56.37	38.82	75.50
North-East (%)	0.64	3.14	3.46	1.10	53.84	15.85	40.90	11.54
Workplace								
Tenure (years)	6.17	7.33	7.45	6.40	8.19	7.23	8.08	5.99
Public sector (%)	20.03	24.81	21.79	27.55	11.62	21.67	16.28	18.14
Primary sector (%)	14.10	9.80	13.00	6.43	9.91	7.46	8.82	8.11
Secondary sector (%)	31.02	27.20	32.63	19.41	47.21	29.72	41.54	24.07
Tertiary sector (%)	26.17	30.12	26.52	34.46	27.17	33.96	28.37	40.96
No. of observations	14,943	57,270	46,525	25,688	13,805	12,830	21,773	4,862

Notes: Workplace–related variables are reported for employed individuals, other variables for the complete workforce. Sample is limited to persons 25-55 year old. RUS 0 refers to individuals who do not speak Russian and RUS 1 to those who do. Analogously for English (ENG) and Estonian (EST). Primary sector = agriculture, fishing, mining; secondary sector = manufacturing, construction, electricity; tertiary (service) sector = wholesale, hotels, transport, financial intermediation, real estate; public sector = public administration, education, health

4 Method

We analyse the relationship between labour market outcomes and language skills using the ordinary linear regression. Several related studies use instrumental variables to correct for omitted variable bias and measurement errors (Bleakley & Chin, 2004; Chiswick & Miller, 1995). We refrain from doing that. Our decision is mainly data-driven, there are no suitable instruments in our data,[2] but it also facilitates comparison with other studies. For instance, Azam et al. (2013) and Paolo and Tansel (2015) only report OLS results, and, even more, find that the available proxy variables have little influence on the final results.

We use linear probability models (LPM) to estimate the relationship between language

[2] We run an IV regression using age at immigration to instrument Estonian skills for those Russians who were in the country at 1990. The results are not statistically significant at any conventional level.

skills and unemployment, mainly in order to avoid unnecessary parametric assumptions in the corresponding binary choice models.[3] Similarly, we estimate the effect of language skills on wage by OLS.

We describe the outcome (log wage and unemployment status) y of individual i at year t as

$$y_{it} = \boldsymbol{\alpha}' \boldsymbol{X}_{it} + \boldsymbol{\beta}' \boldsymbol{L}_{it} + \eta_i + \upsilon_t + \epsilon_{it}, \tag{1}$$

where $\boldsymbol{L}$ is the vector of language skill descriptors. $\boldsymbol{X}$ are the other individual characteristics. η_i is the individual-specific effect, υ_t are the year dummies and ϵ_{it} is the idiosyncratic error term. As we use panel data, we cluster the standard errors on individuals. When analysing the long-term time trend, we use the same specification (without year fixed effects) to estimate separate models for each individual year.

We estimate three models that differ in the control variables. The first specification only includes the language skills and year dummies. It describes the "raw" effect of language skills, net of strong time trends in our data. The second specification additionally includes individual characteristics (age, education) and family descriptors (marital status, an indicator for inter-ethnic marriage, having kids up to 17 years old) and county of residence. Finally, the third specification also adds industry, occupation, and other workplace descriptors. Note that for the unemployment probability, we only estimate the first two models.

5 Results

5.1 Unemployment

The central results for unemployment regressions are presented in Table 2[4]. The table clearly indicates that Estonian language skills are associated with an approximately 5 percentage points lower unemployment for both Russian men and women. We also see that the effect is not influenced in any major way by the individual characteristics. For men it is virtually unchanged while for women the estimate falls by 30%, from 6.5 to 4.5 percentage points.

Neither Russian nor English language skills show a comparably strong effect. Knowledge of Russian is not associated with less unemployment for Estonian men, and is even related to a slightly higher unemployment for Estonian women. The latter outcome may be related to omitted variable bias, Russian knowledge was widespread in the Soviet era and may not have been closely associated with favourable unobserved characteristics, however, it has been increasingly associated with age through the period of analysis. English skills are related to less unemployment both for Estonian men and women, but the effect is small. For Russians, the most important language is clearly Estonian. Those who are fluent in that language are approximately 5 percentage point less likely unemployed, the effect changes only a little when introducing the individual controls. In contrast, the effect of English is close to zero and not statistically significant. As in the case of Estonian, individual characteristics change the estimates only a little.

In summary, these results suggest that Estonian is the only language that matters for the employment prospects. Despite widespread bilingualism in the service sector, and increasing

[3] The results of corresponding probit models are virtually identical to the LPM results.

[4] We restrict both models to have the same number of observations as requested by a reviewer. The unrestricted results are provided in the Appendix (Table A1).

Table 2: Estimation results for unemployment.

	Men		Women	
	1	2	1	2
Estonians:				
Russian	−0.026***	−0.001	−0.009*	0.010**
	(0.006)	(0.006)	(0.005)	(0.005)
English	−0.047***	−0.019***	−0.024***	−0.008*
	(0.004)	(0.005)	(0.004)	(0.004)
No. of observations	36,132	36,132	36,015	36,015
R^2	0.022	0.047	0.012	0.036
Russians:				
Estonian	−0.052***	−0.052***	−0.065***	−0.045***
	(0.010)	(0.011)	(0.009)	(0.010)
English	−0.020	0.006	−0.014	−0.003
	(0.012)	(0.014)	(0.011)	(0.012)
No. of observations	12,942	12,942	13,674	13,674
R^2	0.033	0.066	0.023	0.043
Year dummies	✓	✓	✓	✓
Indiv. charact.		✓		✓

Notes: Standard errors (clustered on individuals) in parentheses. Individual characteristics are age, age^2, education, marriage, number of children and inter-ethnic household. The sample period is fromyear 2000 to year 2012. Sample is limited to persons 25–55 years old. *,**,***Significant at 10, 5 and 1 per cent levels, respectively

global importance of English, these languages are not closely associated with unemployment probability.

5.2 Wage

Table 3 presents the central wage regression results.[5] The table confirms the descriptive evidence that wage is positively associated with fluency in all analysed languages with English showing the largest effect. The only exception is Estonian language for Russian men where we fail to find any positive effect for any of the specifications. This finding qualitatively repeats the results of Toomet (2011). However, for women we observe a sizeable positive effect between 6 and 16 percentage points, suggesting that Toomet (2011) results are not valid for the female labour market. The effect decreases somewhat (from 0.167 to 0.114) when including the individual characteristics into the model, suggesting that the premium is partly caused by education and potentially also by unobserved individual performance. Adding observable workplace characteristics lowers the effects further to 0.066, suggesting that part of the language skill premium is realized through workplace selection.

[5] As above, we restrict both models to have the same number of observations as requested by a reviewer. The unrestricted table is in the appendix Table A2.

Table 3: Estimation results for log wage

	Men			Women		
	1	2	3	1	2	3
Estonians:						
Russian	0.112***	0.078***	0.043***	0.094***	0.039***	0.016*
	(0.015)	(0.015)	(0.014)	(0.011)	(0.011)	(0.009)
English	0.343***	0.158***	0.108***	0.311***	0.130***	0.075***
	(0.013)	(0.015)	(0.013)	(0.010)	(0.010)	(0.009)
No. of observations	21,785	21,785	21,785	26,673	26,644	26,449
R^2	0.730	0.760	0.801	0.769	0.810	0.830
Russians:						
Estonian	0.020	−0.013	−0.009	0.167***	0.114***	0.066***
	(0.018)	(0.019)	(0.018)	(0.014)	(0.014)	(0.013)
English	0.248***	0.152***	0.103***	0.223***	0.089***	0.054***
	(0.026)	(0.027)	(0.024)	(0.021)	(0.022)	(0.018)
No. of observations	7,978	7,978	7,978	9,697	9,697	9,697
R^2	0.731	0.746	0.794	0.794	0.809	0.854
Year dummies	✓	✓	✓	✓	✓	✓
Indiv. charact.		✓	✓		✓	✓
Workplace charact.			✓			✓

Notes: Standard errors (clustered on individuals) in parentheses. Individual characteristics are age, age^2, education, marriage, number of children, inter-ethnic household. Workplace characteristics are industry, occupation, ownership type of company, number of workers in company, working in the public sector, experience in company. The sample period is from year 2000 to year 2012. Sample is limited to persons 25-55 year old.*,**,***Significant at 10, 5 and 1 per cent levels, respectively

We also see that knowledge of Russian language is associated with a substantial income premium of 10% for both Estonian men and women. Individual and workplace controls lower this figure substantially implying that both, ability bias and workplace mobility, play a role.

English language fluency is associated with a very large wage premium, up to 34 log points, for all the analysed groups. The number falls substantially when introducing the individual characteristics into the model (model 2), additional workplace characteristics have less effect. However, it remains noticeably more important than Russian for Estonians in all specifications. The estimates are rather similar for Estonian and Russian men but little lower for Russian women than for Estonian women.

In conclusion, the wage regressions suggest that English is the language which is most closely associated with income for men. The effect of Russian language fluency is minor and that of Estonian language fluency non-existent. For women, all three languages are associated with higher wage but the Russian language to a smaller extent. This result is in a stark contrast to the unemployment regressions and we return to this question in the discussion below.

6 Extensions

6.1 A Long-Run View

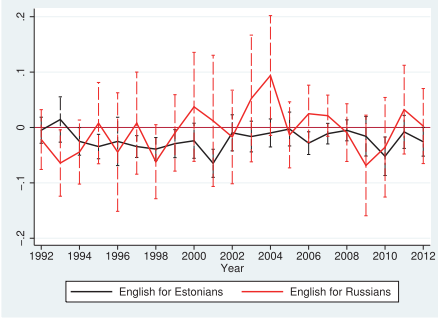
As the Estonian economy went through turbulent times and rapid changes during the 1990s, it is instructive to analyse the long-term trends in the estimated effect size. Here we look at the yearly estimates from 1992 to 2012. The reader should keep in mind that while the pre-2000 waves include retrospective income, they do not include retrospective language skills. In this way, the results before 1995 are partly based on extrapolation, assuming an individual's language skills did not vary before that date.

We report the results for model 2, i.e. models that include year dummies and individual controls, but not workplace characteristics.[6]

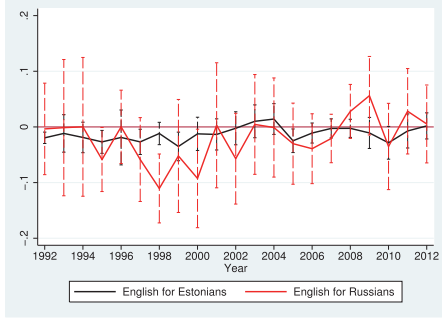
Figure 2: The effect of language fluency on unemployment, 1992–2012

(a) Non-native languages for men

(b) Non-native languages for women



(c) English for Men



(d) English for Women

Notes: Control variables are age, age^2, education, marriage, number of children, inter-ethnic household; 25-55 year olds only.

The effect of language fluency on unemployment is shown in Figure 2, and on wage in

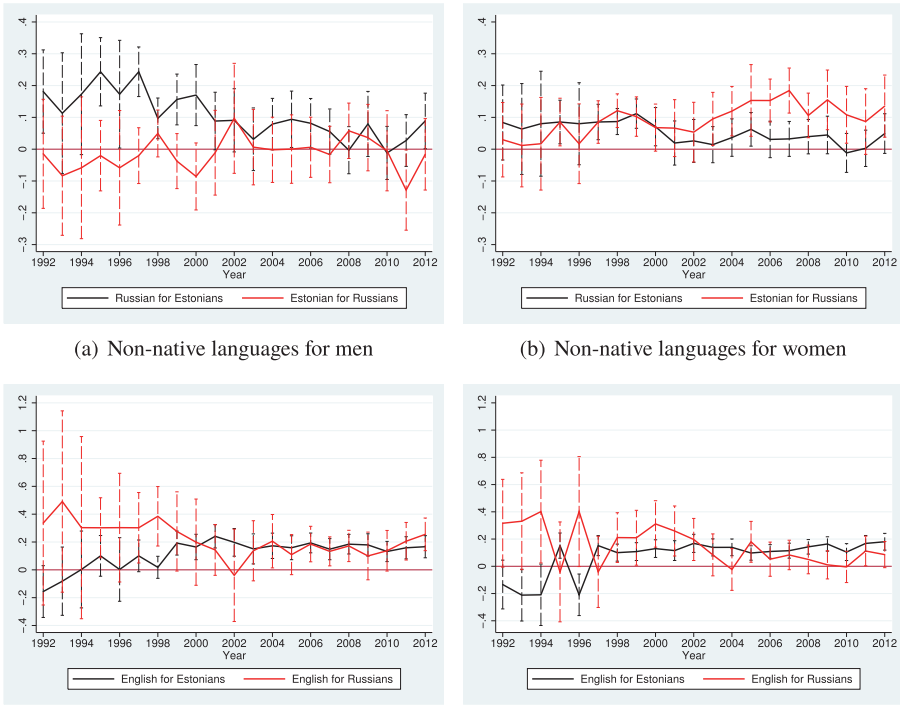
[6] The results for models with workplace data (model 3 for wage) look qualitatively similar. More controls in the model reduce the effect size but does not qualitatively change its temporal pattern.

Figure 3.

The results reveal a few interesting facts. First, the effects are relatively stable over time, indicating that neither the economic environment nor the data quality have changed rapidly. Estonian language fluency has been associated with less unemployment for both Russian men and women since around the year 2000, in the 1990s it had little effect. Russian language fluency, in contrast, had never had an impact on Estonians' unemployment probability.

For Estonian men, fluency in English is associated with less unemployment through the whole period of observation whereas fluency in Russian has no effect. Only a few of the results are pointwise statistically significant though. For Estonian women, both languages are associated with less unemployment in the 1990s but not later. The picture is different for Russians. For men, we cannot see any substantial effect of fluency in English whereas fluency in Estonian is clearly associated with less unemployment from the late 1990s on. For women, the picture is mostly similar, just the effect of Estonian language fluency seems to be somewhat delayed compared to that for men.

Figure 3: The effect of language fluency on wage, 1992–2012



(a) Non-native languages for men

(b) Non-native languages for women

(c) English for Men

(d) English for Women

Notes: Control variables are age, age^2, education, marriage, number of children, inter-ethnic household; 25-55 year olds only.

The wage effects show a similar smooth long-term trend. For Estonians, we see a slowly

fading effect of Russian fluency whereas English becomes important in the late 1990s. For Russian men, fluency in Estonian has never been significantly associated with a wage premium. The results are remarkably stable over the whole period, indicating that the peculiar labour market institutions have been persistent for two decades. However, for women we can see that Estonian language fluency turned important in the late 1990s, suggesting that there is either less workplace segregation or another kind of necessity to know the official language at the jobs commonly taken by women. The importance of English seems to be falling over time for Russians, possibly due to the increasing supply of these skills.

6.2 Does the Effect Differ by Age?

As age is a central determinant of how major economic and political reforms influence individual careers, we also analyse the time trends by age groups (Table 5). We group the observation period into 4 sub-periods: early transition (1989–1995); first economic growth (1996–1999); the economic boom, EU accession (2000–2007); and the recession (2008–2012). We use the same econometric framework as described in section 4, the results presented here are based on the specification 2.

For ethnic Estonians, knowledge of Russian (Table 5, upper panel) has become less important after its peak in the late 1990s. Notably, the association between wage and language skills are small in the young groups, the age groups where knowledge of that language declining rapidly. For older age groups, the estimates are larger but not statistically significant since the recession period. English was not important during the early years of transition but has been very closely associated with better pay since 1996. Its importance seems to fall for the young and middle male groups while the oldest men (45-55) and the youngest women (25-35) show a growing trend. This result potentially reflects both the changing supply and demand for English.

Now we turn to ethnic Russians (lower panel in the Table 5). The association between Estonian skills and income shows no statistically significant results at all. This is the most unexpected outcome in our analysis, and we'll return to this below in Section 7. The effect of Estonian language skills was initially small for Russian women as well, but since the year 2000, we observe a sizeable gain of approximately 10% for all subgroups.

The positive outcomes for women for the second decade in our analysis suggests that the previously segregated "Estonian" and "Russian" economies are slowly becoming a more integrated one where ethnic Russians, at least women, with good Estonian command get better-paid jobs than those who cannot speak the language. In comparison, the estimates for the English language skills trend downward over time for all age groups for both men and for women. A likely cause is the increasing abundance of these skills over time.

We repeat the exercise for the unemployment models. However, due to relatively small groups, most of the results are not statistically significant and we do not report those here.

Table 4: The trend over time across the age groups

Estonian speakers. Dependent variable: log wage						
	Men			Women		
years/age:	25-35	36-45	46-55	25-35	36-45	46-55
Russian:						
1989-1995	0.064	0.247***	0.187***	0.117*	0.033	0.058
	(0.069)	(0.068)	(0.064)	(0.053)	(0.057)	(0.056)
1996-1999	0.116**	0.161**	0.190***	0.059*	0.114***	0.131***
	(0.042)	(0.049)	(0.054)	(0.030)	(0.030)	(0.033)
2000-2007	0.054*	0.157***	0.093*	0.041	0.022	0.042
	(0.026)	(0.037)	(0.037)	(0.024)	(0.021)	(0.026)
2008-2012	0.041	0.032	0.088	0.032	0.012	0.061
	(0.032)	(0.049)	(0.069)	(0.030)	(0.026)	(0.033)
English:						
1989-1995	0.097	-0.008	0.066	0.070	0.042	0.101
	(0.064)	(0.065)	(0.083)	(0.049)	(0.048)	(0.078)
1996-1999	0.190***	0.276***	0.137*	0.113***	0.101***	0.101**
	(0.036)	(0.049)	(0.054)	(0.031)	(0.028)	(0.038)
2000-2007	0.166***	0.158***	0.225***	0.159***	0.073***	0.164***
	(0.027)	(0.036)	(0.041)	(0.024)	(0.021)	(0.025)
2008-2012	0.088*	0.179***	0.236***	0.166***	0.123***	0.145***
	(0.039)	(0.035)	(0.043)	(0.037)	(0.024)	(0.029)
Russian speakers. Dependent variable: log wage						
	Men			Women		
years/age:	25-35	36-45	46-55	25-35	36-45	46-55
Estonian:						
1989-1995	-0.027	-0.113	-0.029	0.134*	0.073	-0.039
	(0.094)	(0.077)	(0.079)	(0.056)	(0.052)	(0.059)
1996-1999	-0.023	0.001	0.049	0.049	0.034	0.071
	(0.050)	(0.048)	(0.060)	(0.045)	(0.035)	(0.037)
2000-2007	-0.038	-0.034	0.042	0.099**	0.089**	0.122***
	(0.038)	(0.049)	(0.043)	(0.033)	(0.031)	(0.028)
2008-2012	-0.004	-0.034	0.002	0.145**	0.123**	0.103**
	(0.055)	(0.050)	(0.047)	(0.051)	(0.044)	(0.038)
English:						
1989-1995	0.325*	0.362*	0.473*	0.191	0.096	-0.096
	(0.160)	(0.151)	(0.216)	(0.104)	(0.085)	(0.353)
1996-1999	0.324**	0.224*	0.129	0.241**	0.048	0.281*
	(0.101)	(0.103)	(0.132)	(0.078)	(0.091)	(0.120)
2000-2007	0.142**	0.107	0.169*	0.127**	0.104	0.127*
	(0.051)	(0.097)	(0.072)	(0.049)	(0.060)	(0.057)
2008-2012	0.140*	0.212***	0.138	0.029	0.105*	0.040
	(0.056)	(0.057)	(0.072)	(0.049)	(0.050)	(0.055)

Notes: Standard errors (clustered on individuals) in parentheses. Individual characteristics are age, age^2, education, marriage, number of children, inter-ethnic household. *,**,***Significant at 10, 5 and 1 per cent levels, respectively

7 Additional Evidence: Education and Workplace

The most surprising result in our analysis so far – the fact that Russian men earn virtually no Estonian language premium while women do – warrants a deeper look at the data. We split the average wage for the Russian male wage-earners along three dimensions: Estonian language skills (EST 0 and EST 1), residence in Tallinn metro area (Tallinn and Other), and education (< HS, HS, College Degree). The metropolitan area is one of the two major regions where the Russian-speaking population is concentrated, and its booming economy may be very different from the other such region, the North-East (Ida-Virumaa), and the rest of the country. In order to correct for the strong wage and language skill trends through the observation period, we do not present the average wage, but instead the respective OLS residuals where we explain the log wage by year fixed effects and include no other explanatory variables (Table 7).

The table shows that the picture is in fact complex. We can see that there is little to no gain from Estonian skills for the largest group of workers, those with a high school degree. Even more, outside of the metropolitan area, the lowest and highest education groups have in fact negative language premium. However, these two groups enjoy quite a noticeable premium in the Tallinn region. The opposite-signed effects in for these two education groups suggests that a selective internal migration may play a role.

Table 5: Average wage and time-trend corrected residual across location and education groups for Russian men

			Log Wage		
Location	Education	No. of observations	EST 0	EST 1	Premium
Tallinn	< HS	399	-0.20	-0.01	0.19
Tallinn	HS	2,164	0.07	0.04	-0.03
Tallinn	College Degree	610	0.15	0.30	0.15
Other	< HS	501	-0.12	-0.12	-0.01
Other	HS	2,658	-0.10	-0.07	0.03
Other	College Degree	429	0.17	0.11	-0.06

Notes: Log Wage refers to regression residual where log wage is explained by year fixed effects. EST 0 and EST 1 are mean values for those who cannot, and who can, speak Estonian; Premium = EST 1 − EST 0. Education level < HS refers to less than high school degree, HS is the high school degree as the highest completed education

We can conclude from the Table 7 that our most surprising finding, no positive effect of Estonian language skills for Russian men, is primarily driven by the largest group, workers with a high school degree. It also suggests that highly educated workers in the metropolitan area enjoy, in fact, a substantial language premium. However, this group is not very large.

The previous results can be explained by a combination of two hypotheses: first, Russian–speaking men, at least those with high school degree as their highest completed education level, are working primarily in a Russian language environment (ethnic segregation) while this is not true for women; and second, Russian-speaking men are working in jobs with little communication requirements (gender segregation). To get some additional evidence on these

Table 6: Language environment and language need at workplace

	men	women
main communication language Estonian	15.0	22.2
Need Estonian at work	53.0	61.9
Need English at work	26.5	11.9
No. of observations	132	134

Note: Percentage of respondents who respond affirmatively

Table 7: Language environment and language need at workplace: regression results

Dependent variable:	Need Estonian	Use Estonian	Need English
woman	0.262**	0.162*	-0.045
	0.117	0.095	0.192
college education	0.201**	0.132**	0.133*
	0.083	0.067	0.079
region Tallinn	-0.276***	-0.463***	0.036
	0.083	0.067	0.080
region North-East	-0.527***	-0.473***	0.022
	0.091	0.074	0.088
No. of observations	228	228	228
R^2	0.2598	0.2831	0.108

Note: Standard errors are in parentheses. Control variables include: marital status, education, occupation and age. *,**,***Significant at 10, 5 and 1 per cent levels, respectively

hypotheses, we use data from the 2008 Integration Monitor Survey. We focus on two questions: "which is the language of communication at your workplace?" and "which languages do you need at work". The first question addresses the ethnic segregation while the second focuses on the need of the language for communication.

Table 6 shows the proportion of Russian respondents who use Estonian as the main communication language at work, and who need Estonian and English at work. As we see, women are more likely to use Estonian while the opposite is true for English. Table 7 presents the corresponding regression results where we estimate the likelihood of having the need for and using the language at work while controlling for gender and other individual characteristics, including education and geographic location.

The table indicates that Russian women need Estonian 26 percentage points more at work than Russian men. The table also suggests that Russian women are 16 percentage points more likely to be working in an Estonian-speaking environment when controlling for other factors. Both figures are statistically significant (at the 5% and 10 % level respectively), although the standard errors are large. We can also see that college education is associated with much more Estonian need and usage, but in regions where large amounts of Russian speakers live, the capital Tallinn and the North-East, Estonian is less needed. In contrast, no gender difference

is visible for English needs.

In summary, this dataset supports both of the hypotheses outlined above: Russian women are working in a less segregated environment where Estonian is the main communication language and they need Estonian at work for other reasons as well. Gender differences in English usage are not statistically significant.

8 Discussion and Concluding Remarks

We document that all analysed language skills are associated with unemployment and wage, however in different ways. Our most intriguing observation is that the largest group of Russian men, high school graduates, earn virtually no income premium if fluent in Estonian, the majority and official language of the country, while ability to speak the language is related to a substantially lower unemployment rate. Our analysis supports Yao and van Ours (2015), and Lindemann (2013) findings that the labour market is segmented between ethnic and gender lines. However, we find that ethnic segregation is less prevalent in the male labour market. This suggests for the majority of Russian-speaking men, Estonian language gives a distinct advantage by facilitating hiring by Estonian-dominated employers. We also confirm Leppik and Vihalemm (2015) conclusion that these jobs are associated with little career mobility. There is also some anecdotal evidence suggesting that offices are typically manned by ethnic Estonians while manufacturing units have a large share of Russian workers. One can speculate that this outcome is related to a substantial population share with negative attitudes toward Russians (Korts, 2009). On a more positive tone, we find that this may not be the case for highly educated workers.

The ethnic segregation seems to be less of an issue for women. Unfortunately, neither Toomet (2011) nor Leppik and Vihalemm (2015) analyse men and women separately. We find some evidence that female labour market is less segregated along ethnic lines, and women are more frequently employed in jobs in which they need Estonian language more, such as working in a direct contact with customers. We also find evidence that over time Estonian is becoming more important in the female labour market. Unfortunately, the current data does not let us assess whether this is due to decreasing segregation, or economy–wide rise of the importance of Estonian language. No such trend is visible for men suggesting that 20 years —almost a generation— after establishing a nation-state and market economy, the labour market integration is still not complete.

The fact that the Russian language is not associated with any better employment prospects is in stark contrast to the results by Duncan and Mavisakalyan (2015). This is potentially related to the sharp re-orientation to Western markets and nation-state related political reforms Estonia experienced in the early 1990s. Rapid withdrawal from the former USSR economy has resulted in a laudable economic growth, but may have led to increased ethnic disparities, compared to the economies that moved in a slower pace, such as Ukraine (Constant et al., 2011) and the former Caucasus republics (Duncan & Mavisakalyan, 2015).

Our other results are more in line with other studies. English language fluency is related to a substantial income premium but to virtually no effect on unemployment. This is similar to findings by Ginsburgh and Prieto-Rodriguez (2011) and Fabo et al. (2017) who show that English is more important at better-paid jobs. In contrast, our outcomes suggest that English has little use at unstable jobs, at these jobs where workers frequently move between employment and unemployment. Note also that in case of English, we are able to explain a

substantial part of the raw effect by individual characteristics, suggesting that the true causal effect may be smaller than our estimates.

The study has two main limitations. Although we believe that our estimates are good indicators of a causal relationship, we cannot be sure this is the case as our data do not contain strong enough instruments. In particular, as individual characteristics are able to explain a large part of English returns, we suppose part of the remaining effect may also be related to an ability bias. Another weak point in the current data are the self-reported language skills. As language test data becomes available, we may be able to decrease the measurement error in the future.

References

Armstrong, A. (2015). Equilibria and efficiency in bilingual labour markets. *Journal of Economic Behavior & Organization*, *112*(C), 204 - 220. Retrieved from http://www.sciencedirect.com/science/article/pii/S0167268115000359

Azam, M., Chin, A., & Prakash, N. (2013). The returns to English-language skills in India. *Economic Development and Cultural Change*, *61*(2), 335-367. Retrieved from http://www.jstor.org/stable/10.1086/668277

Bellante, D., & Kogut, C. A. (1998). Language ability, us labor market experience and the earnings of immigrants. *International Journal of Manpower*, *19*(5), 319-330. Retrieved from http://dx.doi.org/10.1108/01437729810221995 doi: 10.1108/01437729810221995

Bleakley, H., & Chin, A. (2004). Language skills and earnings: Evidence from childhood immigrants. *Review of Economics and Statistics*, *86*(2), 481-496.

Casale, D., & Posel, D. (2011). English language proficiency and earnings in a developing country: The case of south africa. *The Journal of Socio-Economics*, *40*(4), 385-393. Retrieved from http://EconPapers.repec.org/RePEc:eee:soceco:v:40:y:2011:i:4:p:385-393

Chiswick, B. R. (2008, 6). *The economics of language: An introduction and overview* (Discussion Paper No. 3568). IZA, P.O. Box 7240, 53072 Bonn, Germany: IZA.

Chiswick, B. R., & Miller, P. (2010, January). Occupational language requirements and the value of english in the us labor market. *Journal of Population Economics*, *23*(1), 353-372. Retrieved from http://ideas.repec.org/a/spr/jopoec/v23y2010i1p353-372.html

Chiswick, B. R., & Miller, P. W. (1995, April). The endogeneity between language and earnings: International analyses. *Journal of Labor Economics*, *13*(2), 246-88. Retrieved from http://ideas.repec.org/a/ucp/jlabec/v13y1995i2p246-88.html

Chiswick, B. R., & Miller, P. W. (2002). Immigrant earnings: Language skills, linguistic concentrations and the business cycle. *Journal of Population Economics*, *15*, 31-57. doi: 10.1007/PL00003838

Chiswick, B. R., & Miller, P. W. (2007, 1). *Matching language proficiency to occupation: The effect on immigrants' earnings* (Discussion Paper No. 2587). IZA, P.O. Box 7240, 53072 Bonn, Germany: IZA.

Chiswick, B. R., & Miller, P. W. (2015). Chapter 5 - international migration and the economics of language. In B. R. Chiswick & P. W. Miller (Eds.), *Handbook of the economics of*

international migration (Vol. 1, p. 211 - 269). North-Holland. Retrieved from http://www.sciencedirect.com/science/article/pii/B9780444537645000050

Clark, K., & Drinkwater, S. (2000). Pushed out or pulled in? self-employment among ethnic minorities in England and Wales. *Labour Economics*, *7*, 603-628.

Constant, A. F., Kahanec, M., & Zimmermann, K. F. (2011, oct). The russian-ukrainian earnings divide. *Economics of Transition*, *20*(1), 1–35. doi: 10.1111/j.1468-0351.2011.00428.x

Drinkwater, S. J., & O'Leary, N. C. (1997). Unemployment in wales: Does language matter? *Regional Studies*, *31*(6), 583-591. Retrieved from http://ideas.repec.org/a/taf/regstd/v31y1997i6p583-591.html

Duncan, A., & Mavisakalyan, A. (2015). Russian language skills and employment in the former soviet union. *Economics of Transition*, *23*(3), 625-656. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/ecot.12075

Dustmann, C., & Fabbri, F. (2003). Language proficiency and labour market performance of immigrants in the UK. *The Economic Journal*, *113*(489), 695–717. Retrieved from http://dx.doi.org/10.1111/1468-0297.t01-1-00151

Fabo, B., Beblavý, M., & Lenaerts, K. (2017, Aug 01). The importance of foreign language skills in the labour markets of central and eastern europe: assessment based on data from online job portals. *Empirica*, *44*(3), 487–508. Retrieved from https://doi.org/10.1007/s10663-017-9374-6

François Grin, C. S. (1998). Language-based earnings differentials on the Swiss labour market: is italian a liability? *International Journal of Manpower*, *19*(7), 520-532. Retrieved from http://dx.doi.org/10.1108/01437729810237196

Ginsburgh, V. A., & Prieto-Rodriguez, J. (2011). Returns to foreign languages of native workers in the European Union. *Industrial and Labor Relations Review*, *64*(3), 599-618. Retrieved from http://www.jstor.org/stable/41149481

HTM. (2015). *Haridus- ja Teadusministeeriumi aasta-analüüs.* (Tech. Rep.). Retrieved from https://www.hm.ee/sites/default/files/aastaanalyys2015_0.pdf

Hwang, S.-S., Xi, J., & Cao, Y. (2010). The conditional relationship between English language proficiency and earnings among US immigrants. *Ethnic and Racial Studies*, *33*(9), 1-28.

Isphording, I. (2013). *Returns to local and foreign language skills: Causal evidence from spain* (Ruhr Economic Papers No. 398). Essen: RWI Essen. Retrieved from http://hdl.handle.net/10419/68284

Kahanec, M., & Zaiceva, A. (2009). Labor market outcomes of immigrants and noncitizens in the eu: An east-west comparison. *International Journal of Manpower*, *30*(1/2), 97-115. Retrieved from https://doi.org/10.1108/01437720910948429

Kahanec, M., Zaiceva, A., & Zimmermann, K. F. (2010, December). *Ethnic Minorities in the European Union: An Overview* (IZA Discussion Papers No. 5397). Institute for the Study of Labor (IZA). Retrieved from https://ideas.repec.org/p/iza/izadps/dp5397.html

Korts, K. (2009, 3). Inter-ethnic attitudes and contacts between ethnic groups in Estonia. *Journal of Baltic Studies*, *40*(1), 121–137.

Korts, K., & Kõuts, R. (2002). Inter-ethnic communication in Estonian media (in Estonian). In A. Valk (Ed.), *Estonia and Estonians in comparative perspective (in estonian)* (pp. 234–245). Tiigi 78, Tartu 50410, Estonia: Tartu University Press.

Lang, K., & Siniver, E. (2009, November). The Return to English in a Non-English Speaking Country: Russian Immigrants and Native Israelis in Israel. *The B.E. Journal of Economic Analysis & Policy*, *9*(1), 1-30. Retrieved from https://ideas.repec.org/a/bpj/bejeap/v9y2009i1n50.html

Leping, K.-O., & Toomet, O. (2008, December). Emerging ethnic wage gap: Estonia during political and economic transition. *Journal of Comparative Economics*, *36*(4), 599-619. Retrieved from http://ideas.repec.org/a/eee/jcecon/v36y2008i4p599-619.html

Leppik, M., & Vihalemm, T. (2015). The paradox of national language acquisition: Russian speakers' labor market positions in estonia. *Journal of Baltic Studies*, *46*(4), 1-26. Retrieved from http://dx.doi.org/10.1080/01629778.2015.1105833

Leslie, D., & Lindley, J. (2001, 11). The impact of language ability on employment and earnings of Britain's ethnic communities. *Economica*, *68*(272), 587-606.

Lindemann, K. (2013). The effects of ethnicity, language skills, and spatial segregation on labour market entry success in estonia. *European Sociological Review*, *30*(1). Retrieved from http://esr.oxfordjournals.org/content/early/2013/06/26/esr.jct020.abstract

Lindemann, K., & Saar, E. (2011). Ethnic inequalities in education. In R. Vetik & J. Helemäe (Eds.), *The Russian second generation in Tallinn and Kohtla-Järve* (pp. 59–91). Amsterdam University Press.

Paolo, A. D., & Tansel, A. (2015). Returns to foreign language skills in a developing country: The case of turkey. *The Journal of Development Studies*, *51*(4), 407-421. Retrieved from http://dx.doi.org/10.1080/00220388.2015.1019482

Rooth, D.-O., & Saarela, J. (2007). Native language and immigrant labour market outcomes: An alternative approach to measuring the returns for language skills. *International Migration & Integration*, *8*(2), 207-221.

Shields, M. A., & Price, S. W. (2002). The English language fluency and occupational success of ethnic minority immigrant men living in English metropolitan areas. *Journal of Population Economics*, *15*(1), 137-160. Retrieved from https://ideas.repec.org/a/spr/jopoec/v15y2002i1p137-160.html

Toomet, O. (2011). Learn english, not the local language! ethnic russians in the baltic states. *American Economic Review*, *101*(3), 526-31. Retrieved from http://ideas.repec.org/a/aea/aecrev/v101y2011i3p526-31.html

Williams, D. R. (2011). Multiple language usage and earnings in Western Europe. *International Journal of Manpower*, *32*(4), 372-393. Retrieved from https://ideas.repec.org/a/eme/ijmpps/v32y2011i4p372-393.html

Yao, Y., & van Ours, J. C. (2015). Language skills and labor market performance of immigrants in the netherlands. *Labour Economics*, *34*, 76 - 85. Retrieved from http://www.sciencedirect.com/science/article/pii/S0927537115000317

Zhou, M., & Logan, J. R. (1989). Returns on human capital in ethic enclaves: New york city's chinatown. *American Sociological Review*, *54*(5), 809–820. Retrieved from http://www.jstor.org/stable/2117755

A Additional tables

A.1 Different sample sizes

Table A1: Estimation results for unemployment.

	Men		Women	
	1	2	1	2
Estonians:				
Russian	−0.026***	−0.010	−0.009*	0.010**
	(0.006)	(0.006)	(0.005)	(0.005)
English	−0.047***	−0.019***	−0.024***	−0.008*
	(0.004)	(0.005)	(0.004)	(0.004)
No. of observations	36,160	36,132	36,050	36,015
R^2	0.022	0.047	0.012	0.036
Russians:				
Estonian	−0.052***	−0.052***	−0.065***	−0.045***
	(0.010)	(0.011)	(0.009)	(0.010)
English	−0.020	0.006	−0.015	−0.003
	(0.012)	(0.014)	(0.011)	(0.012)
No. of observations	12,946	12,942	13,689	13,674
R^2	0.033	0.066	0.023	0.043
Year dummies	✓	✓	✓	✓
Indiv. charact.		✓		✓

Note: Standard errors (clustered on individuals) in parentheses. Individual characteristics are age, age^2, education, marriage, number of children, inter-ethnic household. The sample period is from year 2000 to year 2012. Sample is limited to persons 25-55 year old.*,**,***Significant at 10, 5 and 1 per cent levels, respectively

Table A2: Estimation results for log wage

	Men			Women		
	1	2	3	1	2	3
Estonians:						
Russian	0.107***	0.072***	0.045***	0.096***	0.041***	0.016***
	(0.015)	(0.015)	(0.013)	(0.011)	(0.011)	(0.009)
English	0.343***	0.162***	0.108***	0.311***	0.131***	0.075***
	(0.013)	(0.015)	(0.013)	(0.010)	(0.010)	(0.009)
No. of observations	22,290	22,274	21,785	26,673	26,644	26,449
R^2	0.726	0.755	0.801	0.769	0.810	0.850
Russians:						
Estonian	0.018	−0.014	−0.009	0.167***	0.114***	0.066***
	(0.018)	(0.019)	(0.018)	(0.014)	(0.014)	(0.012)
English	0.249***	0.155***	0.103***	0.223***	0.089***	0.054***
	(0.026)	(0.027)	(0.025)	(0.021)	(0.022)	(0.020)
No. of observations	8,234	8,233	7,978	9,904	9,895	9,697
R^2	0.731	0.746	0.795	0.794	0.809	0.855
Year dummies	✓	✓	✓	✓	✓	✓
Indiv. charact.		✓	✓		✓	✓
Workplace charact.			✓			✓

Notes: Standard errors (clustered on individuals) in parentheses. Individual characteristics are age, age^2, education, marriage, number of children, inter-ethnic household. Workplace characteristics are industry, occupation, ownership type of company, number of workers in company, working in the public sector, experience in company. The sample period is from year 2000 to year 2012. Sample is limited to persons 25-55 year old. *, *** Significant at 10 and 1 per cent levels, respectively

A.2 All coefficients table

Table A3: Dependent variable: unemployment

	Estonian men		Estonian women		Russian men		Russian women	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Fluent in Russian	-0.026***	-0.010	-0.009	0.010*				
	(0.006)	(0.006)	(0.005)	(0.005)				
Fluent in Estonian					-0.052***	-0.052***	-0.065***	-0.045***
					(0.010)	(0.011)	(0.009)	(0.010)
Fluent in English	-0.047***	-0.019***	-0.024***	-0.008	-0.020	0.006	-0.015	-0.003
	(0.004)	(0.005)	(0.004)	(0.004)	(0.012)	(0.014)	(0.011)	(0.012)
Age		0.001		-0.017***		0.004		-0.004
		(0.003)		(0.003)		(0.006)		(0.006)
Age^2		0.000		0.000***		-0.000		0.000
		(0.000)		(0.000)		(0.000)		(0.000)
<=basic		0.061***		0.068***		0.053***		0.097***
		(0.007)		(0.009)		(0.015)		(0.020)
College		-0.028***		-0.036***		-0.053***		-0.032**
		(0.005)		(0.004)		(0.012)		(0.011)
Married		-0.054***		-0.023***		-0.118***		-0.045***
		(0.005)		(0.004)		(0.012)		(0.009)
Number of Children		-0.008***		0.014***		-0.002		0.018**
		(0.002)		(0.002)		(0.006)		(0.007)
Harju county & Tallinn		0.045***		0.060***		0.118**		0.094*
		(0.013)		(0.013)		(0.039)		(0.048)
Ida-Viru county		0.084***		0.091***		0.131***		0.129**
		(0.020)		(0.020)		(0.039)		(0.048)
Rest of Estonia		0.065***		0.081***		0.143***		0.079
		(0.013)		(0.013)		(0.041)		(0.048)
Interethnic Household		0.039***		0.020		-0.014		0.001
		(0.012)		(0.011)		(0.015)		(0.013)
Constant	0.163***	0.072	0.106***	0.357***	0.203***	0.067	0.225***	0.221
	(0.009)	(0.057)	(0.007)	(0.054)	(0.013)	(0.118)	(0.014)	(0.126)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	36,160	36,132	36,050	36,015	12,946	12,942	13,689	13,674
R^2	0.022	0.047	0.012	0.035	0.032	0.065	0.022	0.041

Notes: Standard errors (clustered on individuals) in parentheses.
*, **, *** Significant at 10, 5 and 1 per cent

Table A4: Dependent variable log wage: Estonian men & Estonian women

	Estonian men			Estonian women		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Fluent in Russian	0.107***	0.072***	0.044**	0.096***	0.041***	0.016
	(0.015)	(0.015)	(0.013)	(0.011)	(0.011)	(0.009)
Fluent in English	0.343***	0.162***	0.108***	0.311***	0.131***	0.075***
	(0.013)	(0.015)	(0.013)	(0.010)	(0.010)	(0.009)
Age		0.023**	0.020**		0.045***	0.030***
		(0.007)	(0.006)		(0.006)	(0.005)
Age^2		-0.000***	-0.000***		-0.001***	-0.000***
		(0.000)	(0.000)		(0.000)	(0.000)
Education						
<=basic		-0.147***	-0.084***		-0.163***	-0.083***
		(0.014)	(0.013)		(0.014)	(0.013)
College		0.269***	0.215***		0.360***	0.228***
		(0.017)	(0.017)		(0.011)	(0.011)
Married		0.127***	0.082***		-0.005	-0.012
		(0.012)	(0.011)		(0.009)	(0.008)
Number of Children		0.021***	0.015**		-0.035***	-0.023***
		(0.005)	(0.005)		(0.005)	(0.004)
Harju county & Tallinn		-0.649***	-0.547***		-0.967***	-1.093***
		(0.078)	(0.100)		(0.105)	(0.130)
IdaViru		-0.852***	-0.691***		-1.265***	-1.298***
		(0.084)	(0.104)		(0.109)	(0.132)
Rest of Estonia		-0.858***	-0.680***		-1.216***	-1.263***
		(0.078)	(0.100)		(0.105)	(0.130)
Interethnic Household		-0.047*	-0.054*		-0.064**	-0.047*
		(0.023)	(0.021)		(0.023)	(0.020)
Industry						
B			0.146*			0.001
			(0.074)			(0.113)
C			0.050			-0.008
			(0.029)			(0.034)
D			0.075***			0.046
			(0.020)			(0.029)
E			0.141***			0.157**
			(0.036)			(0.048)
F			0.278***			0.053
			(0.022)			(0.043)
G			0.145***			0.025
			(0.024)			(0.028)
H			0.153***			0.048
			(0.035)			(0.033)
I			0.265***			0.033
			(0.024)			(0.033)

J	0.190**	0.083
	(0.063)	(0.043)
K	0.075*	0.049
	(0.032)	(0.034)
L	0.150***	0.124***
	(0.033)	(0.034)
M	-0.020	-0.053
	(0.040)	(0.032)
N	-0.034	-0.003
	(0.049)	(0.031)
O	0.050	-0.036
	(0.033)	(0.033)
P	-0.139*	-0.126**
	(0.071)	(0.038)
Q	0.014	0.005
	(0.103)	(0.047)
R	-0.057	-0.103*
	(0.102)	(0.044)
S	-0.087	-0.039
	(0.111)	(0.088)
U		0.049
		(0.145)
Occupation		
1	0.187***	0.211**
	(0.032)	(0.067)
2	0.034	0.119
	(0.034)	(0.066)
3	-0.019	-0.009
	(0.032)	(0.066)
4	-0.146***	-0.160*
	(0.039)	(0.067)
5	-0.251***	-0.228***
	(0.034)	(0.067)
6	-0.084	-0.084
	(0.049)	(0.074)
7	-0.072*	-0.151*
	(0.033)	(0.069)
8	-0.099**	-0.225***
	(0.032)	(0.067)
9	-0.374***	-0.378***
	(0.036)	(0.067)
Ownership form		
20	-0.161***	-0.078***
	(0.026)	(0.014)
30	-0.068**	0.021
	(0.026)	(0.023)

40			0.224***			0.185***
			(0.028)			(0.026)
50			0.369**			0.133
			(0.140)			(0.078)
90			0.881*			0.385***
			(0.406)			(0.084)
*n*Workers			0.043***			0.046***
			(0.003)			(0.002)
PublicSector			-0.038			-0.006
			(0.029)			(0.022)
ExperienceInCompany			0.002			0.011***
			(0.002)			(0.002)
ExperienceInCompany^2			-0.000			-0.000***
			(0.000)			(0.000)
Constant	7.863***	8.351***	8.122***	7.610***	8.009***	8.266***
	(0.021)	(0.153)	(0.160)	(0.016)	(0.154)	(0.183)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	22,290	22,274	21,785	26,673	26,644	26,449
R^2	0.726	0.755	0.801	0.769	0.810	0.850

Notes: Standard errors (clustered on individuals) in parentheses.

*,**,*** Significant at 10, 5 and 1 per cent.

Explanations of the industry and occupation categories are after the last table

Table A5: Dependent variable log wage: Russian men & Russian women

	Russian men			Russian women		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Fluent in Estonian	0.018	-0.014	-0.009	0.167***	0.114***	0.066***
	(0.018)	(0.019)	(0.018)	(0.014)	(0.014)	(0.013)
Fluent in English	0.249***	0.155***	0.103***	0.223***	0.089***	0.054**
	(0.026)	(0.027)	(0.025)	(0.021)	(0.022)	(0.018)
Age		0.024*	0.012		0.014	0.008
		(0.010)	(0.009)		(0.008)	(0.007)
Age^2		-0.000**	-0.000		-0.000*	-0.000
		(0.000)	(0.000)		(0.000)	(0.000)
Education						
<=basic		-0.084***	-0.051*		-0.034	0.017
		(0.024)	(0.022)		(0.023)	(0.020)
College		0.181***	0.100***		0.272***	0.125***
		(0.027)	(0.026)		(0.019)	(0.018)
Married		0.139***	0.102***		-0.021	-0.032**
		(0.021)	(0.018)		(0.013)	(0.011)
Number of Children		0.009	0.011		-0.007	-0.003
		(0.011)	(0.009)		(0.009)	(0.008)
Harju county & Tallinn		-0.722***	-0.603**		-1.222***	-1.298***

	(0.184)	(0.216)	(0.159)	(0.117)
Ida-Viru county	-0.824***	-0.778***	-1.338***	-1.463***
	(0.184)	(0.215)	(0.159)	(0.118)
Rest of Estonia	-0.813***	-0.690**	-1.290***	-1.372***
	(0.185)	(0.216)	(0.159)	(0.118)
Interethnic Household	0.062*	0.067*	0.019	0.020
	(0.031)	(0.026)	(0.023)	(0.019)
Industry				
B		0.328**		0.162
		(0.104)		(0.099)
C		0.221**		-0.013
		(0.071)		(0.066)
D		0.098		-0.031
		(0.067)		(0.060)
E		0.118		0.033
		(0.072)		(0.071)
F		0.277***		0.065
		(0.069)		(0.083)
G		0.067		-0.037
		(0.071)		(0.062)
H		0.259**		-0.042
		(0.082)		(0.064)
I		0.240***		0.037
		(0.070)		(0.064)
J		0.447**		0.181
		(0.145)		(0.093)
K		0.045		-0.142*
		(0.074)		(0.068)
L		0.099		0.018
		(0.088)		(0.080)
M		-0.049		-0.120
		(0.094)		(0.066)
N		-0.069		-0.070
		(0.098)		(0.065)
O		-0.024		-0.076
		(0.087)		(0.066)
P		-0.353		-0.185*
		(0.330)		(0.089)
Q		-0.045		0.019
		(0.192)		(0.075)
R		0.054		-0.240
		(0.263)		(0.195)
S		0.095		0.044
		(0.189)		(0.123)
Occupation				
1		-0.005		0.089

			(0.075)			(0.062)
2			-0.137			0.074
			(0.074)			(0.058)
3			-0.215**			-0.099
			(0.071)			(0.053)
4			-0.296***			-0.208***
			(0.081)			(0.056)
5			-0.462***			-0.310***
			(0.067)			(0.059)
6			-0.403**			-0.138
			(0.127)			(0.094)
7			-0.307***			-0.197***
			(0.069)			(0.057)
8			-0.298***			-0.238***
			(0.069)			(0.055)
9			-0.548***			-0.440***
			(0.071)			(0.054)
Ownership form						
20			-0.179***			-0.028
			(0.038)			(0.026)
30			-0.117***			-0.041
			(0.026)			(0.027)
40			0.020			0.064*
			(0.030)			(0.030)
50			-0.069			0.213*
			(0.139)			(0.095)
90			-0.001			-0.138
			(0.191)			(0.471)
*n*Workers			0.032***			0.030***
			(0.004)			(0.003)
PublicSector			-0.021			0.011
			(0.041)			(0.032)
ExperienceInCompany			0.011***			0.013***
			(0.003)			(0.002)
ExperienceInCompany^2			-0.000*			-0.000***
			(0.000)			(0.000)
Constant	7.918***	8.208***	8.430***	7.535***	8.570***	8.925***
	(0.023)	(0.268)	(0.300)	(0.018)	(0.218)	(0.202)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	8,234	8,233	7,977	9,904	9,895	9,697
R^2	0.731	0.745	0.793	0.793	0.809	0.854

Notes: Standard errors (clustered on individuals) in parentheses.

*,**,*** Significant at 10, 5 and 1 per cent.

Explanations of the industry and occupation categories are below this table

Industry categories (according to NACE Rev.2): A – agriculture, forestry and fishing; B –

mining and quarrying; C – manufacturing; D – electricity, gas, steam and air conditioning supply; E – water supply; sewerage, waste management and remediation activities; F – construction; G – wholesale and retail trade; repair of motor vehicles and motorcycles; H – transportation and storage; I – accommodation and food service activities; J – information and communication; K – financial and insurance activities; L – real estate activities; M – professional, scientific and technical activities; N – administrative and support service activities; O – public administration and defence; compulsory social security; P – education; Q – human health and social work activities; R – arts, entertainment and recreation; S – other service activities; T – activities of households as employers; undifferentiated goods and services producing activities of households for own use; U – activities of extraterritorial organisations and bodies

Occupation categories according to ISCO-08: 1 – managers; 2 – professional; 3 – technicians and associate professionals; 4 – clerical support workers; 5 – service and sales workers; 6 – skilled agricultural, forestry and fishery workers; 7 – craft and related trades workers; 8 – plant and machine operators, and assemblers; 9 – elementary occupations; 10 – armed forces occupations end

Corresponding author

Sven-Kristjan Bormann can be contacted at: bormann@ut.ee

Appendix 3. Paper II

WAGE GAP IN AN ETHNICALLY SEGMENTED LABOR MARKET: THE ROLE OF COGNITIVE SKILLS

Publication:

Ridala, S.; Toomet, O.-S., 2019. Wage Gap in an Ethnically Segmented Labor market: The role of Cognitive Skills. *Eastern European Economics*, vol. 57, no. 1, pp. 20–30. DOI: https://doi.org/10.1080/00128775.2018.1524713 (ETIS 1.1).

Acknowledgement: The Version of Record of this manuscript has been published and is available in Eastern European Economics 2019 © Taylor & Francis Group, LLC. doi.org/10.1080/00128775.2018.1524713

Draft as conference proceedengs:

Ridala, S.; Toomet, O.-S. 2015. *Wage Gap in an Ethnically Segmented Labor market: The role of Cognitive Skills.* The 7th international conference "Economic Challenges in Enlarged Europe", Conference Proceedings, 14–16. June 2015, Tallinn, Estonia.

Eastern European Economics, 57: 20–30, 2019
Copyright © Taylor & Francis Group, LLC
ISSN: 0012-8775 print / 1557-9298 online
DOI: https://doi.org/10.1080/00128775.2018.1524713

Routledge
Taylor & Francis Group

Check for updates

Wage Gap in an Ethnically Segmented Labor Market: The Role of Cognitive Skills

Svetlana Ridala

School of Economics and Business Administration, Tallinn University of Technology, Tallinn, Estonia

Ott Toomet

Information School, University of Washington, Seattle, WA, USA

We analyzed the ethnic wage gap in Estonia that hosts a large Russian-speaking minority population. We used the Oaxaca-Blinder decomposition to test whether the observed wage gap among the working population is related to commonly unobserved skills. Data from the Estonian Labor Force Survey and the Programme for the International Assessment of Adult Competencies (PIAAC) have been used. We find that the three PIAAC skills combination lowered the estimated unexplained wage differential by approximately 20%. We conclude that the unexplained wage gap is most likely related to entry barriers, combined with low-level segregation and segregated social networks.

Keywords: Cognitive skills, Estonia, ethnicity, PIAAC 2013, wage gap

JEL Classification: J15, J31, J71

Numerous ethnic minority groups have lower wages. The possible explanations for this observation range from lower qualifications, lower language skills (Dustmann and Fabbri 2003; Dustmann and Theodoropoulos 2010), and weaker family background (Black et al. 2006) to asymmetric effects caused by macroeconomic shocks (Borjas, Freeman, and Katz 1996; Bound and Freeman 1992). Nevertheless, even if controlling for all the available information, minorities still tend to earn less than the majority population (see Kahanec 2007). The examples include blacks and middle eastern workers in Europe (Blackaby et al. 2005; Clark and Drinkwater 2009; Lehmer and Ludsteck 2011; Longhi, Nicoletti, and Platt 2013; Miranda and Zhu 2013), ethnic Russians in the Baltic states (Leping and Toomet 2008), and blacks in the southern United States (Altonji and Blank 1999; Black

Svetlana Ridala obtained her M.Sc. in Mathematics in 2003 from Tallinn University, Estonia. Since 2009, she has been a doctoral student in Economics at Tallinn School of Economics and Business Administration at Tallinn University of Technology, Estonia.

Ott Toomet obtained his Ph.D. in Economics in 2005 from Aarhus University, Denmark. Since 2017, he has been a lecturer at the Information School, University of Washington, Seattle, WA.

Correspondence should be addressed to Svetlana Ridala, School of Economics and Business Administration, Tallinn University of Technology, Akadeemia tee 3, Tallinn 12618, Estonia. E-mail: svetlana.ridala@gmail.com

et al. 2006). There are also examples of different methodologies that bear on the wages of immigrants/foreigners, like analyses of reservation wages for non-Estonians and Estonians in Estonia (Rõõm 2003), stochastic earnings frontier estimations for immigrants and natives in Germany (Lang 2005), estimations using a formal model of job-search, as well as an econometric methodology for assessing information utilized during the search process for immigrants in the United States (Daneshvary et al. 1992).

Traditionally, studies that analyze unexplained wage differential rely on survey data that include little information on human capital besides formal education (Albouy 2008; Altonji and Blank 1999; Aptekar 2009; Okumura and Usui 2016). More recently, an increasing number of authors have used different measures of cognitive skills and personality traits (Carneiro, Heckman, and Masterov 2005; Fortin 2008; Hertz, Tilly, and Massagli 2001; Nordman, Sarr, and Sharma 2015). These studies conclude that cognitive skills play a role in explaining the differences. For instance, Fryer (2011) finds that accounting for cognitive skills leads to a dramatic fall in the white-black income gap, and calls the "problem of the skill gap" the problem for the twenty-first century (p. 925).

This article complements the still small body of literature by using the rigorous Programme for International Assessment of Adult Competencies (PIAAC) test data on three dimensions of cognitive skills: literacy, numeracy, and problem solving in technology-rich environments (referred below as PS-TRE). Unlike Carneiro, Heckman, and Masterov (2005), the PIAAC skills are focused on common work-related tasks and are measured at the same time as income. We focus on prime working-age men and interethnic wage differential.

Our results indicate that the PIAAC direct human capital measures have only a moderate effect on the unexplained wage differential—without skill controls, the Estonian majority men earn approximately 20% more compared to the Russian-speaking minority men; skill controls lower this difference to 16%.

This article is divided into six sections. The next section provides a brief background. The third section describes the datasets and variables and provides summary statistics. The fourth section describes methodology, empirical results are described in the fifth section, and the last section concludes.

BACKGROUND

Estonians, the ethnic majority in Estonia, form almost 70% of the population of the country (Tammur, Noorkõiv, and Plakk 2014). They are speakers of Estonian, a language that belongs to the Finno-Ugric language family, and have been living in the same region for generations. The largest minority groups in the country are Russians (25.2%), Ukrainians (1.7%), and Belarussians (1%). Because these groups commonly speak Russian as their first language and share a relatively similar cultural background, they are often colloquially referred to as "Russians." In this article, we follow this convention and refer to the minorities as Russians.

Before the Second World War, Estonia was ethnically relatively homogeneous, populated mainly by ethnic Estonians, who constituted 88% of the population. After the war and incorporation into the USSR, the country became a destination for substantial immigration from Russia and other parts of the Soviet Union. The immigration was encouraged by the

Soviet authorities and was related to rapid industrialization as newly opened mines, power plants, and factories needed a lot of labor.

Most of the newcomers settled in the capital Tallinn and in the industrial northeastern part of the country, and by 1989 their share had reached 40% of the population. As the immigrants were typically Russian speakers, this led to two *de facto* official languages by the 1970s. This trend was strengthened by the fact that immigrants were working in a small number of industries, and tended to live in a geographically limited area. Russian was also the working language in many sectors of the economy and society. Analogously, the media and education were (and still are) segregated with Russians following Russian-language media and attending schools with Russian as the instruction language, whereas Estonians almost exclusively prefer Estonian language media and schools. As a counter-reaction and a way of passive resistance, Estonians developed a binary distinction between their own "Estonian" and the alien "Russian" people, with the respective racial hierarchy where Estonians stood clearly above Russians (Kaljund 2018).

The situation changed during the late 1980s. After becoming independent in 1991, the importance of the Estonian language grew rapidly as the new constitution stated that the main goal of the Estonian State is to preserve the Estonian language and culture. The symbolic and political power of the two main ethnic groups changed their place with Estonians obtaining the leading position in society. This was reflected by Estonians developing an advantage by ascending to administrative and professional jobs, with Russians being overrepresented in blue-collar jobs. The country still maintains its language-based segregated schools. In this way, one of the most important institutions for building social ties between nationalities is missing.

The relationship between the two main ethnic groups has mostly been "normal" though periodically somewhat tense. The race-based binary distinction between "own" and "alien" groups is still widespread (Kaljund 2018). There are no significant political parties or movements that strive for a confrontation between the nationalities and voting is typically based on political preferences, not nationality. There is no direct evidence of ethnic discrimination. However, substantial income disparities between Estonians and Russians arose around the time of independence and economic restructuring (Leping and Toomet 2008). Dissatisfaction with the social and economic position has led many Russian-speaking youngsters to immigrate to Western countries (Aptekar 2009), albeit the Estonian-speaking younger generation has also witnessed significant emigration.

DATA AND DESCRIPTIVE STATISTICS

Data Sources

The empirical analysis is based on two different datasets—the Estonian Labor Force Survey (ELFS) and the Programme for the International Assessment of Adult Competencies (PIAAC) for Estonia. Both datasets allow us to control for standard personal characteristics and human capital variables, such as age, education, and family status.

The ELSF is conducted as a quarterly rotating panel where approximately 4,000 distinct individuals a year are surveyed. However, due to its rotating nature, there are roughly

16,000 annual observations. We focus on the surveys from 2011. The ELFS includes information that is commonly found in labor force surveys, including demographic characteristics, wage, and workplace data. The PIAAC data was collected by the OECD and the Estonian Ministry of Education and Research between August 2011 and March 2012. It provides representative and internationally comparable data about skills of the adult populations, designed to measure several key cognitive workplace and social-participation relevant skills. The interviews were conducted in either Estonian or Russian. Our central variables of interest are literacy, numeracy, and the ability to solve problems in a technology-rich environment (PS-TRE). Literacy is measured as the ability to understand and use information from written texts in a variety of contexts. Numeracy is the ability to use and communicate mathematical information and ideas—for instance, the ability to interpret graphs and tables. PS-TRE covers the specific class of problems people deal with when using information and communication technologies (ICT). Technology skills measure typical everyday computer-related tasks, such as Internet search and e-mail handling (see OECD 2013, for further information).

Variables and Sample Selection

We limited ourselves to employed males 24–63 years old to lower the number of students and retirees in our sample (the legal retirement age was 63 years for men in 2011). This gave us 5,537 observations for ELFS and 2,762 for PIAAC. Below, we briefly discuss the most important variables.

Information on the ethnic background for ELFS is based on a question about the respondents' *ethnic nationality*, which is present in all waves of the survey. In most cases, this means the individual ethnic identity. Usually, the identification is language based, although it may differ in certain circumstances such as individuals born in multilingual families. This variable only allows us to distinguish between Estonian and non-Estonian workers. However, as most non-Estonians use Russian as their first language, we refer to them also as "Russians." Information on the ethnic background for the PIAAC data is based on the respondents' *first language, learned at home in childhood and still understood*, due to the absence of the nationality question.

For both datasets, we use the monthly wage at the main job as the income variable. The ELFS reports "last net salary at the main job," while PIAAC gives the corresponding gross salary.

The ELFS includes self-reported information on language skills by asking whether one speaks the language at home, is able to write and speak, speak, or only understands the language. We collapse this information into a dummy, equal to zero for languages one cannot speak (but can potentially understand), and one for language one speaks and potentially writes. Unfortunately, PIAAC does not contain language skills information. Hence, we use the survey completion language instead (it was possible to use Estonian, Russian, or English). Admittedly, this may misclassify a number of bilingual individuals, but we know that Estonian language skills are not closely related to income (Toomet 2011).

Unlike ELFS, PIAAC does not include geographic information. Hence, we have to rely on econometric models that do not control for residency in the metropolitan area, a variable that

carries rather strong explanatory power. As Leping and Toomet (2008) show, this leads to underestimation of the unexplained wage gap, so our estimates rather reflect the lower bound.

Descriptive Statistics

The averages for the most important variables and the corresponding relative frequency for ELFS and PIAAC are presented in Table 1. We report the data separately for Estonians and

TABLE 1.
Descriptive Statistics.

	ELFS		*PIAAC*					
			Total		*Low skill*		*High skill*	
Variable	*Estonians*	*Russians*	*Estonians*	*Russians*	*Estonians*	*Russians*	*Estonians*	*Russians*
Wage (EUR)	789.56	625.36	1242.46	1001.62	1114.52	927.45	1322.01	1109.39
Literacy			280.37	262.31	243.76	239.12.	308.55	299.18
Numeracy			281.65	267.44	243.61	243.52	310.92	305.47
PS-TRE			280.25	270.22	240.62	242.71	303.01	301.87
Age	41.55	43.23	41.98	42.87	44.61	44.59	39.95	40.12
Aged 24–35	0.34	0.30	0.34	0.34	0.25	0.28	0.41	0.42
Aged 36–50	0.37	0.31	0.39	0.35	0.40	0.35	0.38	0.35
Aged 51–63	0.24	0.32	0.27	0.35	0.37	0.37	0.21	0.23
Living with a partner	0.40	0.56	0.74	0.69	0.77	0.72	0.77	0.74
Number of children	0.71	0.50	0.88	0.80	0.81	0.73	0.93	0.90
Education								
Less than high school	0.20	0.17	0.15	0.12	0.26	0.16	0.05	0.04
High school	0.57	0.62	0.47	0.47	0.50	0.51	0.44	0.40
College degree	0.23	0.21	0.38	0.41	0.24	0.32	0.50	0.55
Location								
Capital area	0.32	0.56						
North-East	0.03	0.32						
Occupation								
Professional	0.13	0.05	0.13	0.08	0.04	0.04	0.20	0.15
Skilled	0.24	0.13	0.40	0.22	0.20	0.13	0.54	0.37
Unskilled	0.51	0.55	0.43	0.52	0.54	0.55	0.34	0.47
Industry								
Public sector	0.16	0.08	0.17	0.10	0.11	0.07	0.21	0.15
Primary sector	0.07	0.04	0.07	0.04	0.09	0.05	0.05	0.02
Secondary sector	0.33	0.40	0.32	0.37	0.33	0.36	0.31	0.37
Tertiary sector	0.22	0.21	0.24	0.21	0.18	0.18	0.28	0.27
No. of observations	4110	1427	1975	708	896	445	1079	263

Notes: PIAAC reports gross wage, ELFS net wage. All variables are weighted and reported for the complete workforce. Primary sector = Agriculture, Fishing, Mining; Secondary sector = Manufacturing, Construction, Electricity; Tertiary (service) sector = Wholesale, Hotels, Transport, Financial intermediation, Real estate; Public sector = Public administration, Education, Health. Skilled = Managers, Professionals, Technicians and Associate Professionals; Unskilled: Clerical support Workers, Skilled Agricultural, Forestry and Fishery Workers, Craft and Related Trades Workers, Plant and Machine Operators, and Assemblers, Elementary Occupations. Skills median = 275.9696.

Russians. For the PIAAC data, we additionally split the sample into "high skilled" and "low skilled" samples where we have split the individual average value of the three PIAAC skill scores (literacy, numeracy, PS-TRE) along its median.

We see that there are no substantial differences in individual characteristics across these ethnic groups in both datasets, although a number of average values differ noticeably between the data sources. Both datasets confirm that, on average, Russian men earn approximately 20% less than Estonian men. Figure 1 shows that for datasets for both groups, the wage distribution is well approximated by a log-normal distribution. The distribution for Estonians is clearly shifted upward, and has a slightly larger spread as the difference is more pronounced at the upper end of the distribution.

The industrial distribution differs more across ethnicity. The Estonians are more likely to work in agriculture, public administration, and trade; the non-Estonians in mining, energy, manufacturing, and transport. The reason may be related to the regional settlement pattern. The ethnic groups differ by employment across different occupations as well. While Estonians are overrepresented among managerial, professional, and agricultural occupations, there are many more Russians in craft and related trades.

ESTIMATION: OAXACA-BLINDER DECOMPOSITION

We use the popular Oaxaca-Blinder method to decompose the average wage difference between Estonians (E) and Russians (R) into two, explained and unexplained, components (Blinder 1973; Oaxaca 1973).

Assuming y_i^g, the log wage of individual i from ethnic group $g \in \{E, R\}$, can be written as

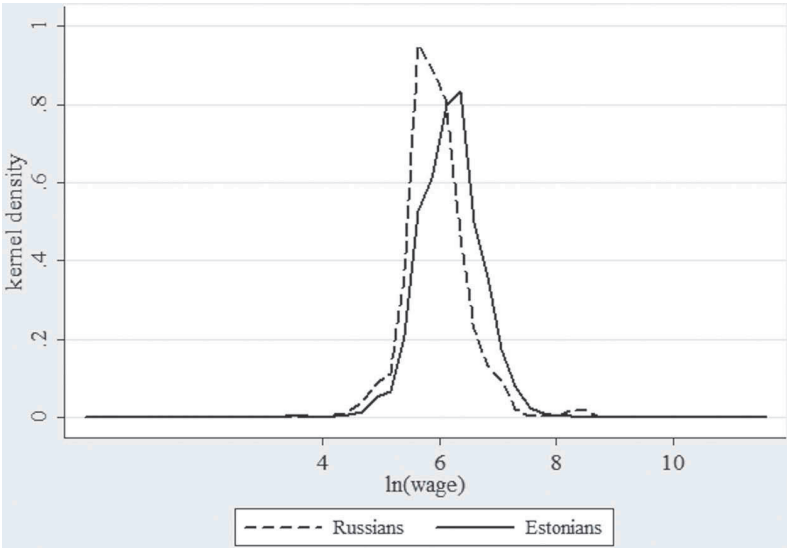
$$y_i^g = \boldsymbol{\beta}^{g\prime} \boldsymbol{X}_i^g + \boldsymbol{\gamma}^{g\prime} \boldsymbol{Z}_i^g + \epsilon_i^g, \quad E(\epsilon_i^g) = 0 \tag{1}$$

where $\boldsymbol{X}^g$ contains the explanatory variables that are common to both groups, such as age and skills, and $\boldsymbol{Z}^g$ is a vector, containing the group-specific variables. We only control for Estonian language skills for Russians (and the other way around), as virtually all Estonians are fluent in Estonian and all Russians fluent in Russian. Finally, $\boldsymbol{\beta}^g$ and $\boldsymbol{\gamma}^g$ are the corresponding parameters, and ϵ_i^g is the random error.

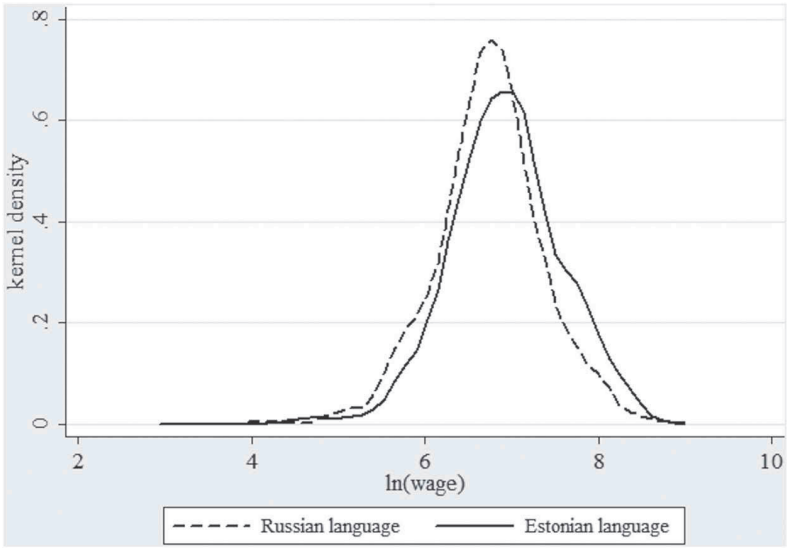
Let the upper bar denote the sample average of the corresponding variable and hat the estimated parameter value. The difference between Russians' average log wage, $\bar{y}^R$ and Estonians' average log wage $\bar{y}^E$ can be decomposed by adding and subtracting the term $\hat{\boldsymbol{\beta}}^{E\prime}\overline{\boldsymbol{X}}^R$ as follows:

$$\bar{y}^R - \bar{y}^E = \hat{\boldsymbol{\beta}}^{E\prime}\left(\overline{\boldsymbol{X}}^R - \overline{\boldsymbol{X}}^E\right) + \left(\hat{\boldsymbol{\beta}}^{R\prime} - \hat{\boldsymbol{\beta}}^{E\prime}\right)\overline{\boldsymbol{X}}^R + \left(\hat{\boldsymbol{\gamma}}^{R\prime}\overline{\boldsymbol{Z}}^R - \hat{\boldsymbol{\gamma}}^{E\prime}\overline{\boldsymbol{Z}}^E\right) \equiv \Delta_x + \Delta_{\boldsymbol{\beta}} + \Delta_z. \tag{2}$$

The first component, Δ_x, captures the mean wage difference, caused by differences in common individual characteristics, such as age or education; Δ_β captures the unexplained difference, caused by different valuations of common skills; and finally, Δ_z describes the difference, caused by explanatory variables that are not present or not well-defined for the other group.[1] The standard errors for all components can be calculated by the delta method. Note that the decomposition can also be performed using $\overline{\boldsymbol{X}}^E$ instead of $\overline{\boldsymbol{X}}^R$ in the term Δ_β, we use the



(a) ELFS



(b) PIAAC

FIGURE 1. Density of Log Wages for ELFS and PIAAC Datasets.

minority-specific explanatory variables $\overline{X}^R$ here. This specification answers the question: what would the wage of Russian workers be, given their current characteristics, if their characteristics were valued similarly to the Estonian-speaking workers.

We estimate a number of models that differ with respect to the types of explanatory variables included (cf. Leping and Toomet 2008). Model 1 only includes a constant, corresponding to the raw differentials in the data. Model 2 adds age, immigrant status (born outside Estonia), and family descriptors (married or cohabiting, having children up to 17 years old). Model 3 additionally includes education descriptors, and Model 4 the dummies for industry and occupation. For ELFS data, we are also able to estimate Models 5 (also includes language skills) and 6 (includes geographical controls in addition to language skills). In this way, we can assess which variables are related to the unexplained wage differential, and see if ELFS and PIAAC corroborate each other. For ELFS data we rely on just one version of each model, while for PIAAC data we also estimate an amended version of every model that includes the three cognitive skills, both separately and all three together. This permits us to assess whether the unexplained wage differential will significantly narrow by introducing skills, suggesting that literacy, numeracy, and PS-TRE explain a part of the differential. We use sampling weights provided in the data for all the estimations.

RESULTS

The decomposition results (unexplained differential Δ_β) are given in Table 2. Negative values mean that Russians' observed characteristics are undervalued compared to the corresponding characteristics of Estonians. Model 1 (raw differentials) indicates that Russian men earn approximately $e^{-0.200} = 0.819$ of the average wage of Estonians. The results for ELFS and PIAAC (without skills) are remarkably similar. However, the other PIAAC models indicate that

TABLE 2.
Unexplained Wage Differential in Favor of Ethnic Russians (Δ_β).

		PIAAC				
Model	*ELFS*	*no skills*	*ICT*	*literacy*	*numeracy*	*3 skills*
1	−0.194***	−0.203***	−0.173***	−0.168***	−0.148***	−0.204***
constant	(0.027)	(0.038)	(0.046)	(0.043)	(0.041)	(0.053)
2	−0.202***	−0.201***	−0.177***	−0.177***	−0.156***	−0.196***
+ age, immigrant, family	(0.028)	(0.045)	(0.056)	(0.049)	(0.049)	(0.063)
3	−0.196***	−0.213***	−0.188***	−0.196***	−0.174***	−0.205***
+ education	(0.027)	(0.044)	(0.056)	(0.048)	(0.048)	(0.064)
4	−0.149***	−0.168***	−0.172***	−0.150***	−0.141***	−0.184***
+ industry, occupation	(0.027)	(0.048)	(0.057)	(0.048)	(0.049)	(0.064)
5	−0.157***					
+ language skills	(0.030)					
6	−0.215***					
+ regional descriptors	(0.043)					
No. of obs.	5537	2762	2762	2762	2762	2762

Notes: Linearized standard errors in parentheses. Significance levels: *10%, **5%, ***1%. Each following model includes all the explanatory variables of the previous models, plus the indicated variables.

Russians earn in fact $e^{-0.150} = 0.861$ of the salary of Estonians with comparable skills. On average, Russians men have lower skills.

Introducing different controls may increase or decrease the unexplained differential. For instance, both industry and occupation noticeably lower the differential (for ELFS data) suggesting part of the differential stems from better-paid industries. Regional controls will increase the gap substantially in absolute value. Leping and Toomet (2008) show that this is related to a larger-than-average differential in the Tallinn metropolitan area. Language skills have little impact on the differential. However, the standard errors are rather large and it is not obvious if these differences are statistically significant.

Comparison of ELFS and PIAAC models indicates that these provide consistent results given the confidence intervals.[2] Most importantly, the PIAAC models that include the cognitive skills consistently explain approximately 0.02–0.04 units, or 10–20% more of the wage differential. This figure is broadly constant across the specifications and skills included in the model. However, it seems that skills carry less weight in Model 4, the model that includes both industry and occupation. Apparently, higher skills are a way to enter the better-paid industries. However, as standard errors are rather large, this evidence remains somewhat speculative. Unfortunately, we cannot estimate Model 6 that includes geographic controls on PIAAC data.

We also performed pooled decomposition (see Jann 2008). The inference is qualitatively the same (results available upon request).

DISCUSSION AND CONCLUSION

We analyzed the unexplained wage differential between ethnic Estonian and Russian men in the Estonian labor market using both ELFS and PIAAC data. Both datasets suggest that Estonians earn approximately 20% more on average, confirming the earlier results by Leping and Toomet (2008). When including direct controls for cognitive skills, present in the PIAAC dataset, the unexplained differential falls by approximately 10–20%, or 0.02–0.04 log units. This is true even for models that include education level, and industry and occupation of the workplace, although the standard errors are large. This outcome suggests that the human capital descriptors in common surveys, such as ELFS, are not as good as the dedicated measures of cognitive skills, available in specialized datasets, such as PIAAC.

However, despite introducing several dimensions of cognitive skills, 80% of the differential is left unexplained. As PIAAC skill measures are based on tests, dedicated to accurately measure the workplace-relevant skills, we do not believe that such a large part of the unexplained differential is related to left-out measures of cognitive skills. More plausible, the difference is related to noncognitive skills or social skills, and discrimination in the form of entry barriers combined with low-level segregation.

The fact that Russian men have lower skills on average also suggests that the education provided by Russian-language schools lags behind what is provided in schools where Estonian is the language of instruction.

It would be interesting to extend this study by inclusion of noncognitive skills. Unfortunately, these skills are hard to measure and rarely available in sizable datasets.

NOTES

1. While the first two terms are standard in this literature, the third term has been used less often. It facilitates interpretation of the two other terms in case of important control variables that cannot be used for both groups at the same time. Leaving such variables out would arguably change the constant and possibly introduce a missing variable bias into the other estimates.
2. As ELFS and PIAAC samples are independent, one can directly compare the corresponding estimates. Comparison of different models estimated on the same sample requires separate tests, beyond the scope of this article.

ACKNOWLEDGMENTS

We thank participants of the conference "Economic Challenges in Enlarged Europe" 2015, and participants of the Doctoral Summer School in Economics and Innovation in Estonia for helpful comments and suggestions.

REFERENCES

Albouy, D. 2008, November. "The Wage Gap between Francophones and Anglophones: A Canadian Perspective, 1970–2000." *Canadian Journal of Economics* 41 (4):1211–38. doi:10.1111/j.1540-5982.2008.00501.x.

Altonji, J. G., and R. M. Blank. 1999. "Race and Gender in the Labor Market." In *Handbook of Labor Economics*, edited by O. C. Ashenfelter, and D. Card, vol. 3, Chapter 48, 3143–259. Amsterdam: Elsevier Science Publishers B.V.

Aptekar, S. 2009. "Contexts of Exit in the Migration of Russian Speakers from the Baltic Countries to Ireland." *Ethnicities* 9 (4):507–26. doi:10.1177/1468796809345433.

Black, D., A. Haviland, S. Sanders, and L. Taylor. 2006, May. "Why Do Minority Men Earn Less? A Study of Wage Differentials among the Highly Educated." *The Review of Economics and Statistics* 88 (2):300–13. doi:10.1162/rest.88.2.300.

Blackaby, D., D. Leslie, P. Murphy, and N. O'Leary. 2005. "Born in Britain: How are Native Ethnic Minorities Faring in the British Labour Market?" *Economics Letters* 88 (3):370–5. doi:10.1016/j.econlet.2005.03.008.

Blinder, A. S. 1973. "Wage Discrimination: Reduced Form and Structural Estimates." *The Journal of Human Resources* 8 (4):436–55. doi:10.2307/144855.

Borjas, G. J., R. B. Freeman, and L. F. Katz. 1996. "Searching for the Effect of Immigration on the Labor Market." *The American Economic Review* 86 (2):246–51. http://www.jstor.org/stable/2118131.

Bound, J., and R. B. Freeman. 1992. "What Went Wrong? the Erosion of Relative Earnings and Employment among Young Black Men in the 1980s." *The Quarterly Journal of Economics* 107 (1):201–32. doi:10.2307/2118327.

Carneiro, P., J. J. Heckman, and D. V. Masterov. 2005. "Labor Market Discrimination and Racial Differences in Premarket Factors." *The Journal of Law and Economics* 48 (1):1–39. doi:10.1086/426878.

Clark, K., and S. Drinkwater. 2009. "Dynamics and Diversity: Ethnic Employment Differences in England and Wales, 1991–2001." *Research in Labour Economics* 29:299–333.

Daneshvary, N., H. W. Herzog, R. A. Hofler, and A. M. Schlottmann. 1992. "Job Search and Immigrant Assimilation: An Earnings Frontier Approach." *The Review of Economics and Statistics* 74 (3):482–92. doi:10.2307/2109493.

Dustmann, C., and F. Fabbri. 2003. "Language Proficiency and Labour Market Performance of Immigrants in the UK." *The Economic Journal* 113 (489):695–717. doi:10.1111/1468-0297.t01-1-00151.

Dustmann, C., and N. Theodoropoulos. 2010. "Ethnic Minority Immigrants and Their Children in Britain." *Oxford Economic Papers* 62 (2):209–33. doi:10.1093/oep/gpq004.

Fortin, N. 2008. "The Gender Wage Gap among Young Adults in the United States: The Importance of Money Vs. People." *Journal of Human Resources* 43:886–920. doi:10.1353/jhr.2008.0006.

Fryer, R. G. 2011. "Chapter 10 - Racial Inequality in the 21st Century: The Declining Significance of Discrimination." In *Handbook of Labor Economics*, Vol. 4 *of Handbook of Labor Economics*, edited by D. Card, and O. Ashenfelter, 855–971. Amsterdam: Elsevier.

Hertz, T., C. Tilly, and M. P. Massagli. 2001. “Linking the Multi-City Study’s Household and Employer Surveys to Test for Race and Gender Effects in Hiring and Wage Setting.” In *Urban Inequality: Evidence from Four Cities*, edited by A. O’Connor, C. B. Tilly, and L. D. Bobo, 407–43. New York: Russell Sage Foundation.

Jann, B. 2008. “The Blinder-Oaxaca Decomposition for Linear Regression Models.” *Stata Journal* 8 (4):453–79. http://www.stata-journal.com/article.html?article=st0151.

Kahanec, M. 2007, November. “Ethnic Competition and Specialization.” IZA Discussion Papers 3167, Institute for the Study of Labor (IZA). http://ideas.repec.org/p/iza/izadps/dp3167.html.

Kaljund, A. L. 2018. “Orientalism against Empire: The Paradox of Postcoloniality in Estonia.” *Anthropological Quarterly* 91 (2):749–70. doi:10.1353/anq.2018.0032.

Lang, G. 2005. “The Difference between Wages and Wage Potentials: Earnings Disadvantages of Immigrants in Germany.” *The Journal of Economic Inequality* 3 (1):21–42. doi:10.1007/s10888-004-7581-4.

Lehmer, F., and J. Ludsteck. 2011. “The Immigrant Wage Gap in Germany: Are East Europeans Worse Off?” *International Migration Review* 45 (4):872–906. doi:10.1111/j.1747-7379.2011.00871.x.

Leping, K.-O., and O. Toomet. 2008. “Emerging Ethnic Wage Gap: Estonia during Political and Economic Transition.” *Journal of Comparative Economics* 36 (4):599–619. doi:10.1016/j.jce.2008.08.002.

Longhi, S., C. Nicoletti, and L. Platt. 2013. “Explained and Unexplained Wage Gaps across the Main Ethno-Religious Groups in Great Britain.” *Oxford Economic Papers* 65 (2):471–93. doi:10.1093/oep/gps025.

Miranda, A., and Y. Zhu. 2013. “English Deficiency and the Native-Immigrant Wage Gap.” *Economics Letters* 118 (1):38–41. doi:10.1016/j.econlet.2012.09.007.

Nordman, C., L. Sarr, and S. Sharma 2015. “Cognitive, Non-Cognitive Skills and Gender Wage Gaps: Evidence from Linked Employer-Employee Data in Bangladesh.” Working Papers DT/2015/19, DIAL (Developpement Institutions et Mondialisation). http://EconPapers.repec.org/RePEc:dia:wpaper:dt201519.

Oaxaca, R. 1973. “Male-Female Wage Differentials in Urban Labor Markets.” *International Economic Review* 14 (3):693–709. doi:10.2307/2525981.

OECD. 2013. *OECD Skills Outlook 2013: First Results from the Survey of Adult Skills*. Paris: Publishing. doi:10.1787/9789264204256-en.

Okumura, T., and E. Usui 2016, January. “Intergenerational Transmission of Skills and Differences in Labor Market Outcomes for Blacks and Whites.” IZA Discussion Papers 6227, Institute for the Study of Labor (IZA). https://www.econstor.eu/bitstream/10419/130345/1/dp9662.pdf.

Rõõm, M. 2003. “Reservation Wages in Estonia.” Working Paper January, Archive at WUSTL. https://core.ac.uk/download/pdf/7033105.pdf.

Tammur, A., K. Noorkõiv, and M. Plakk. 2014. *Population. Eesti Statistika Aastaraamat. 2014. Statistical Yearbook of Estonia*, 49–66. http://www.stat.ee/72571.

Toomet, O. 2011. “Learn English, Not the Local Language! Ethnic Russians in the Baltic States.” *American Economic Review* 101 (3):526–31. doi:10.1257/aer.101.3.526.

Appendix 4. Paper III

LANGUAGE SKILLS AND LABOUR MARKET RETURNS: A META-REGRESSION ANALYSIS.

Publication:

Ridala, S. (2020). Language skills and labour market returns: A meta-regression analysis. *Language Problems and Language Planning*, vol. 44, no. 2, pp. 200–241. DOI https://doi.org/10.1075/lplp.20012.rid (ETIS 1.1)

Reprint from

"Language skills and labour market returns. A meta-regression analysis", Ridala, S., pp. 200–241 in "Language Problems and Language Planning", Volume 44:2 - 2020.

Published online and expected in print: January 2021.

Copyright © Published by John Benjamins Publishing Company.

Amsterdam/Philadelphia.

Acknowledgement: The Version of Record of this manuscript has been published and is available in Language Problems and Language Planning 2021 © https://www.benjamins.com/catalog/lplp

Language skills and labour market returns

A meta-regression analysis

Svetlana Ridala
Tallinn University of Technology

Increased migration, global trade and the introduction of digital labour platforms call for a better understanding of the mechanisms that can enhance economic and labour market outcomes in the face of increasing disparities in culture, language and identity. The article contributes to the literature on labour market returns on language skills, which is very heterogeneous and context-specific, by carrying out a systematic review of that literature. The meta-regression analysis estimates the returns on language skills that come from socio-economic, institutional and ethnolinguistic factors along with controls for the study design. The meta-regression results for the labour market returns that stem from knowledge of the local state language, a significant local minority language or the business language provide evidence that knowing the more influential language is associated with higher labour market rewards. Linguistic diversity has a negative effect on the returns to skills in the business language but increases the returns to the local state language. Urbanisation has a negative, though quantitatively modest, effect on returns to language skills in the business language, whilst the GDP level increases the returns to minority language skills and unemployment increases returns to skills in the business language and the local state language.

Keywords: language returns, language skills, wage differentials, meta-regression analysis, ethnicity

1. Introduction

The globalisation of the world economy and labour markets has prompted interest in how language skills affect labour market outcomes. The ageing of the population in Europe and other advanced economies has triggered an expansion in migration flows and these trends call for different languages and minority populations to be better integrated and included in labour markets. The literature on Language Economics helps to create a better understanding of the relationship

https://doi.org/10.1075/lplp.20012.rid | Published online: 3 December 2020
Language Problems and Language Planning issn 0272-2690 | e-issn 1569-9889
© John Benjamins Publishing Company

between language skills and economic returns, and of the pecuniary benefits and disadvantages that can arise from ethnicity (Gazzola & Wickström, 2016; Ginsburgh & Weber, 2016; Grin, 2003).

Knowledge of languages is among the factors of human capital that contribute to an individual's productivity (Männasoo et al. 2018), social capital, and eventual labour market success. The literature on economic returns to language skills is diverse. The literature on labour market integration analyses whether and how much immigrants benefit from knowing the state language of their host country (Bleakley & Chin, 2004; Budría & Swedberg, 2012; Chiswick, 1998; Chiswick & Miller, 1995; Dustmann & Fabbri, 2003; Dustmann & van Soest, 2001). Globalisation has strengthened the role of English and other major languages as languages of international communication and business, and a number of empirical studies have investigated the impact of English as a business language in non-English-speaking countries (Azam et al., 2013; Casale & Posel, 2011; Di Paolo & Tansel, 2019; Fabo, Beblavý, & Lenaerts, 2017; Isphording, 2013; Stöhr, 2015; Wang & Cheng, 2017; Williams, 2011). The role and economic impact of other major international business languages have been studied by Saiz & Zoido (2005) among others. The empirical evidence on the role and economic impact in a country of the language that is second most commonly spoken there is quite limited, but there are prominent examples from strongly multilingual countries such as Estonia and Latvia (Bormann et al., 2019; Toomet, 2011), Canada (Christofides & Swidinsky, 2010), and Belgium and Luxembourg (Williams, 2011).

This paper contributes to the existing literature by providing a systematic review of the literature on language returns using a meta-regression analysis across three groups of primary studies on language returns. These groups study (1) returns on a local official or state language, (2) returns on a local significant minority language and (3) returns on business language.

Most of the literature in language economics is highly specific to its own context and is conducted with reference to a single country or to a particular ethnicity and language environment. This means there is only limited information on how environmental and context-specific factors more generally direct the effect that language skills have on economic and labour market outcomes. This paper presents a meta-study on how environmental and institutional factors facilitate or impede the transmission of the language effect on labour earnings. The meta-regression, a regression analysis where the coefficient estimates from existing studies form the dependent variable, identifies the role of contextual factors, including the power of language, socio-linguistic diversity and economic environment, and whether and how much language skills affect earnings.

This paper proceeds as follows: Section 2 gives a brief overview of the underlying literature on language skills and earnings; Section 3 describes the study

design and method; Section 4 presents the meta-regression results; and Section 5 discusses and concludes.

2. Literature on language skills and related earnings

A large number of empirical studies show that language skills have a positive impact on earnings. These studies are highly heterogeneous, but they may be divided into three broad categories: (1) labour market integration studies, which investigate the benefits of having skills in the language of the host country for immigrant or language-minority populations; (2) studies on the role of the second major local language and the labour market returns from knowing that language; (3) studies on the earnings premium that arises from skills in a global business language such as English.

The strand of literature that focuses on immigrant populations has found positive returns from knowing the language of the host country using data on male immigrant workers (Bleakley & Chin, 2004; Budría et al., 2017; Chiswick, 1998; Chiswick & Miller, 1995; Dustmann & van Soest, 2001), female immigrant workers (Miranda & Zhu, 2013a), or pooled samples of male and female immigrant workers (Bleakley & Chin, 2004; Di Paolo & Raymond, 2012; Dustmann & Fabbri, 2003).

There is an increasing body of literature studying the role played in the labour market outcomes of individuals by leading business languages rather than the ethnic majority language of their host country. This line of research is dominated by empirical studies on the impact of English as a business language in non-English-speaking countries (Azam et al., 2013; Casale & Posel, 2011; Di Paolo & Tansel, 2019; Fabo, Beblavý, & Lenaerts, 2017; Isphording, 2013; Stöhr, 2015; Wang et al., 2017; Williams, 2011). The role of business languages other than English has been analysed by Saiz & Zoido (2005), using a representative sample of US college graduates. They found that second language fluency gave a wage premium of 2% to 3%. Ginsburgh & Prieto-Rodriguez (2007) study the business language premium for male workers in nine member countries of the European Union. They report a language proficiency premium in earnings in all nine countries. Di Paolo & Tansel (2015) control for skills in six different foreign languages in a rapidly growing economy using a sample of Turkish male wage earners. They found significant rewards from the most powerful business languages, like English and Russian, and positive but small returns for skills in German and French, while knowledge of Arabic and Bulgarian did not give rise to any earnings premium in Turkey.

The role and economic value of skills in smaller languages is less clear though, especially in multilingual economies. Drinkwater & O'Leary (1997) found that unemployment is lower among bilingual Welsh-speakers than among monolingual English-speakers in the Welsh labour market. Saarela & Finnäs (2003) report that the unemployment rate of the Swedish-speaking minority in Finland is lower than that of the Finnish-speaking majority. They suggest that language proficiency and social integration are the probable reasons behind these differences. Rendon (2007) finds that the probability of being employed in Catalonia increases by between three and five percentage points for people who know how to read and speak Catalan, and by between two and six percentage points for those who can write in Catalan. For Canada, Christofides & Swidinsky (2010) find substantial, statistically significant rewards from the use of English as the second official language in Quebec[1] and insignificant effects from French in the rest of Canada. Armstrong (2015) contributes to the theoretical literature on second language acquisition, and he also shows that bilingual native French speakers enjoy a significant premium both inside and outside the province of Quebec.

Grin & Sfreddo (1998), in contrast, show that speaking Italian as a mother tongue in Swiss regions that are not Italian speaking is associated with a lower wage. The mother tongue does not seem to have any influence on earnings for women, as there is no significant difference between the earnings of women whose first language is German, French or Italian. Chiswick et al. (2000) find a negative association between skills in the indigenous non-dominant language and success in the labour market in Bolivia.

The body of language economics studies on Central and Eastern Europe and countries from the former Soviet Union is fairly limited. Returns to Russian skills vary across the post-communist transition economies with ethnic Russian speakers facing negative earning effects in Estonia (Leping & Toomet, 2008; Lindemann, 2013) and in Latvia (Kahanec & Zaiceva, 2009; Toomet, 2011), and positive earnings effects in Ukraine (Constant et al. 2012). Using data from Armenia, Azerbaijan and Georgia over 2008–2010, Duncan & Mavisakalyan (2015) document significant economic benefits from skills in the Russian language. They claim that skills in Russian increase an individual's probability of employment by about 6% for men and 9% for women.

The returns to second language skills vary across studies depending on the design and context of the study. Yao & van Ours (2015) for instance present an overview of previous studies on language skills and labour market performance over the period 1994 to 2013, showing that the language return estimates vary

1. Grenier (2019) and Castonguay (2019) note that the demographic share of the English-speaking minority in Quebec is increasing.

depending on the different estimation techniques used. According to Dustmann & Van Soest (2002), the effect of proficiency in German on earnings was 4% for women and 5% for men when ordinary least squares (OLS) estimates were used, but this increased to 12% for women and 14% for men when they used instrumental variable (IV) estimation, an econometric estimation technique that employs additional proxy variables, or instruments that are correlated with language skills but do not directly affect wages, in order to avoid reverse causality.[2] The fact that the instrumental variable estimate has a stronger effect than the OLS one suggests that language skills in wage models may indeed exhibit reverse causality or errors-in-variables. The example of how fluency in English affects earnings, estimated by Dustmann & Fabbri, (2003), shows a rise from 18% with OLS to 36% with a two-step least squares (2SLS) estimator, which is the most popular form of IV estimator.

The patterns of the language returns to unemployment are rather mixed. Donado (2017) claims, using a large European dataset, that proficiency in English or German as a second language reduces the probability of being unemployed for the native population, though not for immigrants, more than skills in French, Spanish, and Italian do. He also argues that females benefit more than males from learning foreign languages, while Bormann et al. (2019) show that this is not the case in the Baltic States. Gazzola & Mazzacani (2019) show that knowledge of English as a foreign language increased the probability of employment for men in Germany, Italy and Spain, while for women the effect was present only in Germany and Italy, but not in Spain.

Evidence from two of the Baltic States, Estonia and Latvia, adds to the research on the effects of majority and minority languages in countries with strongly multilingual societies. Bormann et al. (2019) and Ridala (2020) find different language benefit patterns in Estonia and Latvia though. In Latvia the relationship between second language skills and the probability of being unemployed is quite symmetrical for native Latvians and for Russian speakers, while in Estonia the relationship is asymmetric as fluency in the state language reduces the probability of being unemployed for the Russian-speaking minority, while fluency in Russian does not have any significant effect on unemployment for Estonians. Skills in English do not reduce the unemployment risk for Russian speakers, but

2. Instruments that have been used in instrumental variable estimations include the interaction term of age at arrival and the country of origin of immigrants, the percentage of host country language speakers in the neighbourhood, and the number of books at home (Bleakley & Chin, 2004, Miranda &Zhu 2013a; Miranda &Zhu 2013b). Budría & Swedberg (2012) also used a dummy variable for whether there was a child in the family who was fluent in the host country language. Dustmann and van Soest (2002), Dustmann & Fabbri, (2003) and Chiswick and Wang (2016) used the lags of the language skills variable as instruments.

they have a moderate positive effect on employment for native Estonians. The wage regressions show, however, that skills in English are strongly positively associated with labour income in the Estonian labour market. Fluency in Russian has an effect on earnings of about 5% for native speakers of Estonian. Fluency in Estonian has no effect on wages for Russian-speaking men but is significant for Russian-speaking women, who earn an 11% wage premium from their Estonian language skills.

Most of the evidence from the language economics literature confirms the positive effect of language skills on earnings; however, the estimates vary to a great degree depending on the language, study context and study design. The current paper presents a meta-regression analysis that systematically reviews existing comparable studies and in doing so contributes into better synthesis of the literature in the fields.

3. Method

3.1 Study design

Meta-analysis allows the estimated effects of interest from multiple homogenous studies to be combined into a single statistically generalised estimate (Sterne, 2009), and it allows the reasons for the variation in the effect across these primary studies to be explained. Meta-analysis improves on narrative literature surveys and complements them by using a systematic review carried out through statistical tests that control for the subjectivity bias in qualitative reviews.

This paper uses meta-analysis to investigate how the institutional and socioeconomic contexts of the various existing studies affect the estimated returns from knowing languages. Meta-regression relies on statistical generalisation and inference to assess how language skills affect earnings, while controlling for methodological heterogeneity in study design and for the heterogeneity of the population or the impact of contextual factors in a particular reference country in a given reference year. The contextual factors include globalisation and migration trends from the KOF Globalisation Index, the degree of language diversity from the Linguistic Diversity Index, economic structure and the level of development shown by the unemployment rate and GDP per capita, and urbanisation and regional disparities from urban population statistics.

The meta-regression analysis, or systematic review, proceeds in two steps. The first step defines the selection criteria, collects the primary articles and compiles the meta-data set, which contains structured information for examining the factors that affect the estimated labour returns to language fluency in the existing

literature. The second step sets up and estimates the meta-regression model. The primary studies for meta-regression analysis fall into three categories. The first category of studies estimates economic returns to knowledge of the local official or state language for ethnic minority and immigrant populations. These studies were conducted on local state languages in Estonia, the UK, the US, Australia, Canada, Israel, Germany, China and Norway among others (Aldashev et al., 2009; Berman et al., 2003; Bratsberg & Ragan, 2002; Chiswick et al., 2005; Chiswick & Miller, 2010; Hayfron, 2010; see the full list in the references, where studies included in the meta-analysis are marked by an asterisk (*) in front of the author's name). The second category consists of studies on economic returns for the native or majority language population from knowing and using the second most widely spoken local language (Bormann et al., 2019; Christofides & Swidinsky, 2010). The meta-regression analysis considers only second local languages that are spoken by at least 1% of the country's population, a list that includes Russian in Estonia, Catalan in Spain, French in Canada, Arabic in Turkey, German in Denmark and Italy, Italian in Luxembourg and France, and Spanish in France. The third category investigates the returns to proficiency in a major occupational language, such as English in non-English speaking countries including Estonia, Finland, Poland, India, Israel, Turkey, and South Africa among others (Azam et al., 2013; Casale & Posel, 2011; Di Paolo & Tansel, 2019; Saiz & Zoido, 2005; Stöhr, 2015). The meta-regression analysis considers only empirical papers in the field of Language Economics that contain the estimated main effect and its variance of language skills on wages, presented as the logarithm of hourly, weekly, monthly or annual earnings.

The selection of empirical papers for the meta-analysis is based on the principles of integrity and replicability. Clear selection criteria help to maintain these principles by reducing the subjective moderating influence of the researcher and by making the study replicable. The selection process discards papers that are incompatible in their data or their study design. To a certain degree, the choice of the papers lies in the subjective assessment of the authors and this raises concerns about the validity of the choice of papers.[3]

The meta-regression analysis uses empirical papers that apply a Mincer-type wage equation (Mincer, 1958), which estimates the logarithm of wages as a function of human capital variables such as education and work experience. The primary studies selected for the meta-regression augment the Mincer baseline wage estimation and include language skills as an additional control for human capital. The sample unit for the meta-regression is an individual-level estimated earnings equation from the primary studies. The coefficient estimates for the main effect

3. See, for example, the discussion between Hanushek (1998) and Krueger (2003).

of language skills or fluency on hourly, weekly, monthly or annual wage earnings constitute the dependent variable for the meta-regression. The explanatory or independent variables of the meta-regression are the primary study design characteristics along with socio-economic and institutional variables for the reference country. The study design characteristics include the size and gender composition of the sample and the econometric estimation method of ordinary pooled regression or instrumented regression. The socio-economic and institutional context variables include characteristics that vary between countries and study years, including language influence from the Power Language Index, language diversity from the Linguistic Diversity Index, globalisation from the KOF Globalisation Index, urbanisation from the urban population statistics, and economic structure and the level of income shown as the unemployment rate and GDP per capita. Appendix C contains the explanations and descriptive statistics for these explanatory variables or meta-covariates in the meta-regression.

The meta-study selection process used the JStor, ScienceDirect and Google Scholar search engines to retrieve 61 relevant research articles from which 41 (67%) were published in internationally recognised and peer-reviewed research journals indexed in WoS or Scopus, including the International Journal of Manpower, Industrial and Labor Relations Review, The Economic Journal, Labour Economics, Journal of Transition Studies Review, and Economic Development and Cultural Change. In addition, the sample of primary studies included research reports from the IZA (Institute of Labor Economics), publications by the National Bureau of Economic Research, and publications from scientific conferences (the full list of the publications is provided in the references; included studies are marked with an asterisk). Using only internationally published research does inevitably introduce a language bias through a strong over-representation of articles in English. Equally though, this selection criterion strengthens the academic credibility of the collection of primary studies underlying the meta-regression. This means the results of the current meta-study come from research addressed to the international research community and readership. The search was restricted to studies published in 1998 or later in order to capture the more recent socio-economic trends shaping labour markets in the 21st century. The reference countries for the primary studies were limited to OECD countries or to key partner countries of the OECD like China, India and South Africa. This restriction limits the scope of the study so that the results apply to modern labour markets. All the papers selected were checked for conformity in study design and structure so that they would allow comparative research on socio-economic and institutional factors that may have an effect on how language skills affect earnings. The search was conducted in the ScienceDirect, JStor and Google Scholar databases in March 2020 and it used broad keywords and phrases such as “returns

to language" and "economics of language" along with complementary keywords such as "human capital", "language", "earnings" and "labour market" and combinations of them. The final meta-regression sample shrank to 61 primary studies that in total contained 258 language return coefficients. This number of studies constitutes only a fraction of the keyword-related papers in the research databases, since only a few of them contained a compatible empirical analysis on language skills as a determinant of wages. A few of the selected studies report t-statistics instead of the standard error of the estimated language return coefficient, and in these cases the standard error was calculated using the "Cochrane handbook for systematic reviews of interventions" (2008, p. 175). Table B1 in the Appendix presents all the primary studies employed in the meta-regression analysis and provides descriptive statistics of the estimated coefficient for the main or first-order effects of language on earnings, presented in three categories of studies for (1) studies on the returns to the state or local language, (2) studies on the returns to a significant minority language and, (3) studies on the returns to English or another business language. Tables B2 and C1 in the Appendix describe the variables.

Before carrying out the meta-regression, the study applies Egger's test (Egger et al., 1997), a meta-analysis tool that helps to detect bias in the selection of publications and tests for the presence of small-study effects in the meta-regression sample. Small-study effects arise if non-significant findings from smaller studies remain unpublished, a bias that may distort the measurement of the true size of the effect. The null hypothesis of Egger's test is the absence of any small-study effect. Figure 1, the contour-funnel plot presents the coefficients for language skills for the three categories of primary studies. The contour funnels plot the sizes and standard errors of the effects in the primary studies. If there are no small-study effects, the dots are distributed symmetrically, but coefficients with smaller standard errors cluster at the top of the funnel and those with a large dispersion are located at the bottom. The coefficients for studies of English or another business language and for minority-language studies show no evidence of publication bias. The estimates of studies for local or state languages, however, have an indication of publication bias where the value for the Egger's test statistic is significant at 1.07 with a standard error of 0.19. Moreover, some of the estimates on returns to the state or local language tend to have much larger standard errors. This means the results from the studies of local or state languages should be interpreted with caution, since the sample bias suggests that non-significant results are underrepresented, which may give rise to overestimation of the meta-regression effects.

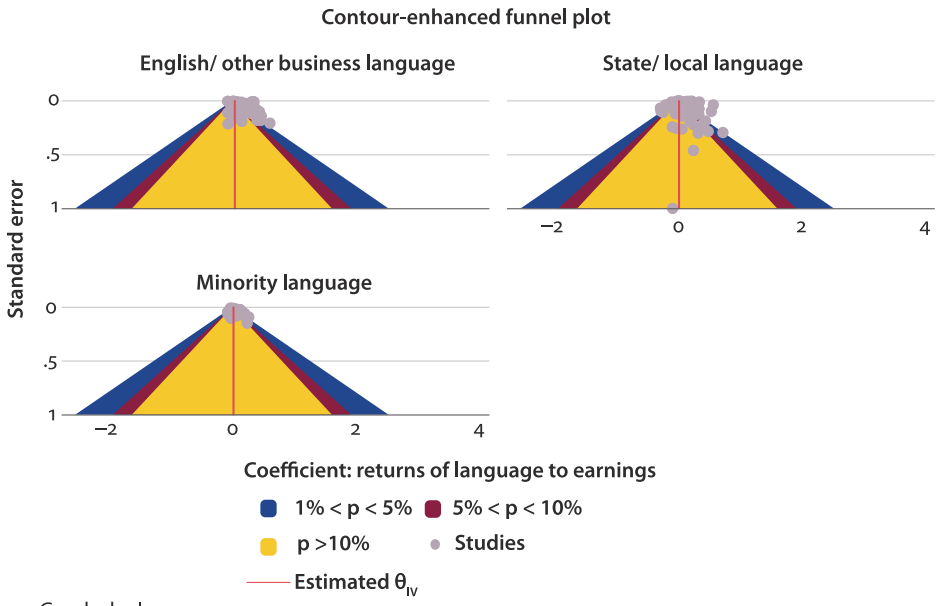


Figure 1. Contour-enhanced funnel plot of the coefficient against its standard error by three study category groups. *Source:* Author's calculations

The publication bias[4] is further tested for using the Duval & Tweedie (2000) nonparametric method. Their trim-and-fill test estimates the number of studies that could be missing in a meta-analysis and evaluates the impact of publication bias on the meta-regression results. The test results suggest that the small-study effect for English or other business languages should be corrected by adding one study to the group, the effect for state or local languages should be corrected by 20 studies, and that for minority languages by three. The resulting funnel plots with "trimmings" and "fillings" or "imputations" are presented in Figure 2. Begg (1997) and Duval (2005) are, however, critical about correcting publication bias with imputations and so the current analysis does not correct the meta-regression sample for potential publication bias and hence the results, in particular those for the local or state language returns, should be interpreted with caution.

The estimated coefficients of wage returns to language skills constitute the dependent variable for the meta-regression. The variation in the dependent variable is explained by controls in the study design and by socio-economic and institutional variables that vary by country and year. These are the level of unem-

4. The publication bias arises if the decision to publish depends on the statistical significance of the results of the study. Typically, significant findings are more likely to get published

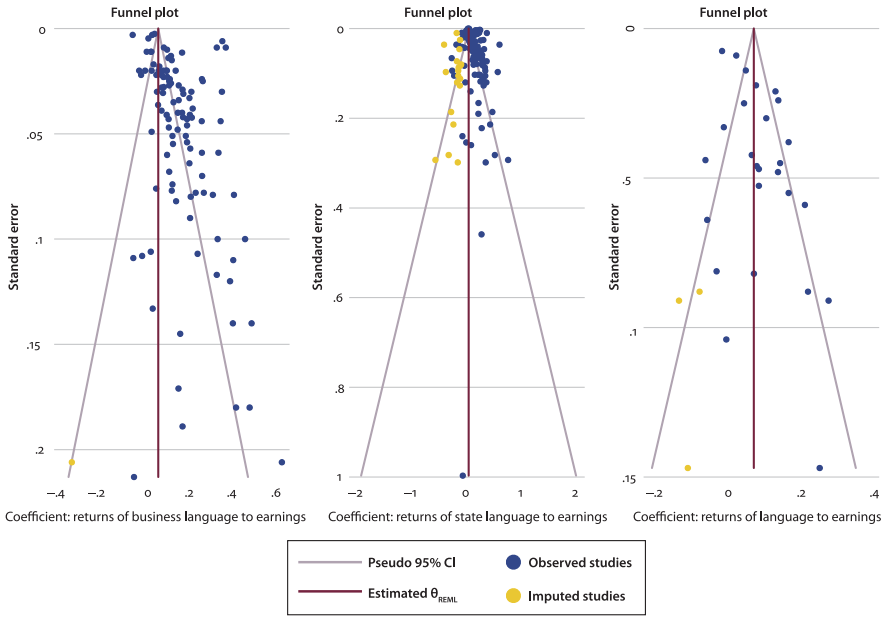


Figure 2. Trim-and-fill funnel plot of the coefficient against its standard error for different language groups. *Source:* Author's calculations

ployment, measured from the percentage of the active population in employment (source: World Bank); real GDP per capita (source: World Bank, in current US dollars); the Power Language Index (PLI),[5] which ranks languages by their influence across multiple dimensions including geography, economy, communication, knowledge and media, and diplomacy (Chan, 2016); the Linguistic Diversity Index (LDI) (Ethnologue, 2017), which ranks languages by the population of language speakers as a proportion of the total population; and the KOF Globalisation Index (KOFGI),[6] which ranks countries by their level of globalisation using economic, social and political criteria (Dreher, 2006, Savina et al., 2019). In addition, the meta-regression encompasses a measure of urbanisation, which is the percentage of urban residents in the total population as obtained from the statistics of the World Bank.[7,8]

5. http://www.kailchan.ca/wp-content/uploads/2016/12/Kai-Chan_Power-Language-Index-full-report_2016_v2.pdf
6. https://kof.ethz.ch/en/forecasts-and-indicators/indicators/kof-globalisation-index.html
7. https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS
8. The correlation statistics for these indicators are available on request from the author.

3.2 The meta-regression model

Meta-analysis rests on collective empirical evidence from the existing literature but also controls for the effect of differences in study design and context. Meta-regression is a particular form of meta-analysis that is designed for integrating and generalising results from numerous empirical studies that are conducted independently but pose a similar research question and apply a similar research design and methodology (Egger & Lassmann, 2012; Stanley, 2001). The current research integrates primary studies that estimate wage returns to human capital variables in the estimation framework introduced by Mincer (1958), and that investigate the effect of language skills on wages amongst other human capital variables. Figure 3 presents the concept of the study.

The meta-regression model takes the form:

$$\hat{\theta}_j = \sum_{k=1}^{M} a_k I_{kj} + \varepsilon_j \qquad j = 1, 2, \ldots, \mathrm{L},\ k = 1, 2, \ldots, \mathrm{M} \tag{1}$$

$\hat{\theta}_j$ represents the meta-dependent variable, which is the estimated coefficients or first-order effects from the impact of language fluency on log earnings from the primary study j amongst all the L primary studies included in the meta-analysis. I_{kj} are the k explanatory variables or meta-covariates (moderators) that control for variations in study design and context. The explanatory variables show in which direction and to what extent the study design and context factors explain the estimated returns to language skills in the primary studies. In other words, they capture the sign and strength of the relationship between the estimates of returns to language and the study-bound factors that may strengthen or weaken these estimates. ε_j are the error terms associated with the uncertainty and residual heterogeneity in the estimation of the primary study effects $\hat{\theta}_j$.

The random-effects meta-regression (RE model) allows for the residual heterogeneity stemming from the between-study variance that remains after controlling for the meta-covariates I_{kj}. In the RE model, the error term has two independent components, $\varepsilon_j = \eta_j + u_j$, where $\eta_j \sim N(0, \tau^2)$ captures the between-study component and $u_j \sim N(0, \sigma_j^2)$ is the within-study component. In this way, the RE model controls for unobserved heterogeneity such as the impact of the analytic rigour of the primary studies that is not fully captured by observed meta-covariates. The RE model treats the effect sizes as random and assumes that the collection of primary studies forms a random sample from a larger population of relevant studies. The between-study variance parameter, τ^2, is estimated with the restricted residual maximum likelihood (REML) estimator (see Appendix A for the details on the REML estimator). The fixed-effects meta-regression (FE

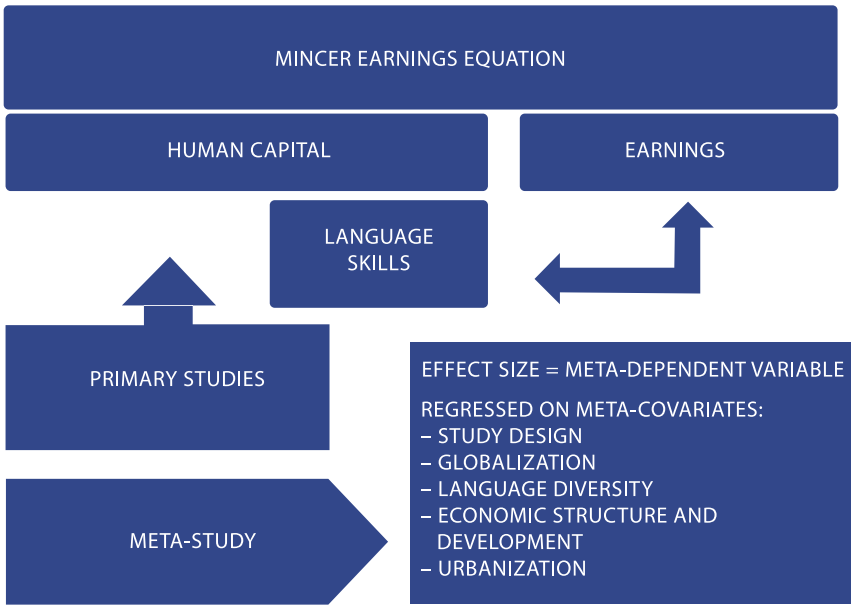


Figure 3. The study concept

model), by contrast, assumes that the observed meta-covariates or moderators capture all the between-study heterogeneity. Variations in the effects are attributed purely to random sampling error $u_j \sim N(0, \sigma_j^2)$. The FE model treats the within-study estimated variances, σ_j^2, as observed or known values that define the variance of the sampling error. That means the effect sizes are considered to be fixed, in the sense that the included primary studies define the entire population of interest. In principle, the FE model employs a linear weighted least squares estimator, with weights equal to the inverse of within-study variances, $1/\sigma_j^2$.

The restrictive assumptions of the FE model render the results more susceptible to Type I error. In other words, the FE model is more likely to reject the null hypothesis of zero meta-covariate coefficients, since it decomposes the effect size into two parts only, which are variation due to observed meta-covariates and variation due to random error. The RE model instead decomposes the effect size into three parts, which are variation explained by meta-covariates, between-study variation, and within-study variation. In consequence, the variation in the RE model is always larger than the variance in the FE model, or is at least equal to it. In the FE model the between-study heterogeneity is subsumed in the estimated coefficients, and this overestimates the variation captured by the meta-covariates. The size of that bias is reported under regression diagnostics as the residual heterogeneity statistic I^2 which captures the remaining between-study heterogeneity

present in the residuals after controlling for the variability in meta-covariates or moderators.

The current study estimates the RE model since the assumptions for the FE model are not satisfied and the primary studies under consideration do not constitute a finite population of relevant research. Table 1 presents the results from the RE model[9] along with regression diagnostics showing that a substantial percentage of the variability in language effect estimates is due to differences between the studies that are not explained by the meta-covariates. This evidence is captured by the high value of the residual heterogeneity statistic I^2 and it justifies the choice of the RE model over FE model for conducting the meta-regression.

4. Results of the meta-regression analysis on returns to language skills

Table 1 presents the random-effects meta-regression estimates[10] with their standard errors in parentheses, along with the regression diagnostics and between-study variance estimates τ^2 for the three categories of primary studies into the effects of English or another business language, the state or local language, and a minority language. The estimates for all the studies confirm there is a significant positive association between the Power Language Index (PLI) and the effect of language skills on earnings. The higher the PLI, or the global influence of the spoken language, is, the higher are the returns on earnings from fluency in that language. The degree of language power is particularly relevant for wages for languages spoken by a significant minority language group or if the language is declared to be the state or main local language. Toomet (2011) showed empirical evidence from Estonia and Latvia that a significant part of the large local Russian-speaking community prefers to gain skills in English rather than in Estonian or in Latvian, the small local state languages.

The Linguistic Diversity Index (LDI), which is very high in India, South Africa and Israel beyond Europe, and in Belgium and Switzerland within Europe, has a negative effect on the returns to knowledge of English or other business languages, but it has a positive effect on the returns from knowing the local state language or languages. This evidence suggests that language diversity encourages the learning of the local language, or even of multiple local languages, rather than just a single common, globally spoken business language.

9. The results of the FE model are available from the author upon request.

10. The results from the fixed-effects model are available from the author upon request.

Urbanisation has a qualitatively modest but significant negative effect on returns to the business language, whereas the effect from globalisation remains insignificant.[11] GDP per capita and the unemployment rate reflect the impact of economic advancement and structure on language returns. The meta-regression estimates show that GDP per capita increases the returns for a minority language and that unemployment increases the returns for English or another business language, and for the local or state language.

The meta-regression also controls for the design of the primary studies, such as the estimation method and the principles of sample formation. The ordinary least squares (OLS) estimates are lower than the instrumental variable (IV) estimates,[12] which are the reference category, and this implies a downward bias in the OLS estimates. The downward bias in the OLS is a common finding in the relevant literature and it has been explained by the presence of classical measurement error or attenuation bias (see Bleakley and Chin, 2004; Budría & Swedberg, 2012; Budría et al., 2017). Another possible explanation for why OLS estimates are biased downwards biased is endogenous selection, since OLS captures only the smaller variation in language skills within a group of wage earners and ignores the role that language skills play in selection into employment. A further source of bias might stem from the omission of individual heterogeneity and dynamic state dependence. Comparison of the panel estimates with the cross-section estimates, which are the reference category, shows however that the cross-section estimates are underestimated for the language returns on the state or local language. This implies that unobserved individual heterogeneity and state dynamics plays a more important role for language returns to the local or state language, whereas it has no significant effect on the English or business language or on the minority language when a number of other study-design and context variables are controlled for.

A categorical variable controlled whether the language effect on wages was estimated on a pooled sample of men and women, which was the reference category, or on a gender-homogeneous sample of men or women in isolation. The results imply that the language returns for wages have a stronger effect in pooled samples than in gender-homogeneous samples, and this is particularly evident for studies on the state or local language. This suggests that language returns for

11. The non-significant results for globalisation may partly be explained by the relatively strong association with economic development and structure captured by GDP level and unemployment, and this may impair the separate identification of the globalisation effect.

12. Instrumental variable (IV) estimates bear a broader connotation and include not only linear instrumental variable estimators (e.g. 2SLS), but also mixed distribution models of Heckman type that control for endogenous selection.

wages are driven by between-job variation, which is partly captured by gender because of gender job segregation. Unfortunately, separate controls for jobs are not feasible since they would overly fragment the sample given the limited number of primary studies available for the meta-regression.

Table 1. Meta-regression random-effects model estimates

	English/other business		State/local		Local minority	
Coefficient	**Coef.**	**Sth.Err.**	**Coef.**	**Sth.Err.**	**Coef.**	**Sth.Err.**
PLI	0.116***	0.034	0.180***	0.051	0.204***	0.071
LDI	−0.113**	0.054	0.330**	0.138	−0.077	0.096
KOFGI	0.081	0.066	−0.045	0.130	−0.200	0.195
Urban	−0.002*	0.001	0.0004	0.002	−0.005	0.003
GDP per capita	−0.001	0.001	−0.001	0.001	0.009*	0.005
Unemployment	0.003*	0.002	0.013***	0.005	−0.006	0.006
Ref: IV estimation						
OLS estimation	−0.124***	0.034	−0.110***	0.029	−0.100***	0.036
Ref: Cross-sectional						
Panel	−0.025	0.021	0.022**	0.033	0.030	0.034
Gender-pooled sample:						
Only men	−0.015	0.023	−0.083***	0.029	−0.033	0.035
Only women	−0.022	0.029	−0.053*	0.036	−0.043	0.034
Constant	0.301**	0.085	−0.041	0.154	0.438	0.337
No. of coefficients	108		124		26	
R-squared	36.46%		28.03%		76.55%	
τ^2	0.0062		0.0124		0.0010	
I2 (%)	94.38		99.83		40.65	
Prob > F	0.0000		0.0000		0.0000	

Notes: Significance levels:
* 10% ** 5% *** 1%
τ^2 stands for the between-study error variance.

In Appendix C, Figures C1, C2, C3 and C4 present bubble plots as a visualisation of a simple meta-regression relationship between the outcome variable and a single continuous meta-covariate or moderator variable.

5. Discussion and concluding remarks

The socio-economic, institutional and ethnolinguistic environment plays an important role in generating the incentives for language learning and in shaping the relationship between language skills and labour market outcomes. The meta-regression results for the labour market returns from knowing the business language, the local state language or a significant local minority language as a second language reveal a number of generalised contextual patterns.

The Power Language Index (PLI) has a strong positive effect on returns to language skills in general, but the effect was the strongest for the skills in the local minority languages and in local state languages, more so than for business languages. This means the economic value of a language depends crucially on its global power and influence and the meta-regression analysis provides evidence that knowing languages that are globally more influential is associated with higher labour market rewards.

The Linguistic Diversity Index (LDI) has a negative effect on the returns from knowing the business language, but it strengthens the returns for the local state language. The high linguistic diversity in countries such as India, South Africa, Israel, Belgium, Switzerland, Luxembourg, Canada, or Latvia and Estonia would seem to encourage the learning of local languages, and knowledge of local languages is appreciated in the labour markets of these countries.

Urbanisation has a negative effect on returns to the business language, though one that is modest in size. Metropolitan areas may be associated with a higher probability that knowledge of the second language is easily accessible, which may suppress private returns to employees from their language skills. On the other hand, metropolitan areas may encourage language learning, since the environment favours diversity and is notable for its sizeable and diverse labour markets. Equally though, the major ethnic groups in metropolitan environments may establish their own micro-societies that allow some members of the ethnic group to escape from the need to interact directly in the host country language. The division of roles within these large ethnic micro-societies or language enclaves may reduce the incentive for at least some of the language-group members to learn the host country language.

The KOF Globalisation Index was estimated with low precision and gave no significant results. This may in part be explained by the relatively strong positive correlation with other economic indicators of GDP per capita and unemployment, but it tentatively hints that there are counter-weighting factors to globalisation, some of which increase the returns to language knowledge and others of which reduce it. Globalised labour markets may reward language skills and offer improved prospects for employment and remuneration for work. On the

other hand, globalised labour markets have a high level of supply in language skills, which reduces the private returns from knowing widely-spoken languages. To corroborate this tentative hypothesis, the sample size for the meta-study would need to be enlarged.

Unemployment increased the returns from knowing the business language and the local or state language. Unemployment indicates structural problems such as imbalances in labour supply and demand in the economy and, depending on the nature of these structural imbalances, the effect upon returns to language skills may vary. The results from the current study suggest though that better language skills in the business language and in official state languages benefit workers in labour markets that have high unemployment. This may hint that the immigrant populations gain from knowing the local state language or a common business language if the economy is suffering from high unemployment. The meta-regression suggests that economies that are witnessing structural employment problems do not place high value on command of the local minority language as the second spoken language. On the contrary, a higher GDP level per capita, which reflects the advancement, depth and diversity of the economy, is associated positively with command of the local minority language as the second language. This suggests that economic growth and development leads to higher returns from knowing local languages, including premiums paid to speakers of the native state language who have a command of local minority languages.

In general, the economic benefits from language skills go beyond private returns and generate externalities and positive social returns. The variety of human capital diversity and skills grows, and this enriches and deepens labour markets. The socio-economic integration of native speakers and immigrant populations opens up mutual gains such as access to a wider and more diverse base of human capital (Besevegis & Pavlopoulos, 2008), and lays the ground for a fair and frictionless labour market applying the "law of one wage" (Amstrong, 2015). Using the resources of diverse ethnic groups by improving inter-ethnic networks may lead to multilateral gains, which in turn may be converted into economic benefits in the form of better employability, an increase in wages and the creation of new jobs.

Funding

This project has received support from the Erasmus Programme of the European Union Project No 611059-EPP-1-2019-1-EE-EPPJMO-MODULE, from Doctoral School in Economics and Innovation, supported by the European Union, European Regional Development Fund (Tallinn University of Technology ASTRA project "TTÜ Development Program 2016–2022", project code: 2014–2020.4.01.16-0032) and from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 952574.

Acknowledgements

The author would like to thank professor François Grin, professor Kadri Männasoo and two anonymous referees for their useful comments and suggestions on an earlier version of this paper. All remaining errors or omissions are mine.

References

*Adamchik, V., Hyclak, T., & Sedlak, P. (2019). Poland in an integrated European economy: Are foreign language skills valued by employers in the Polish labor market? *Journal of Transition Studies Review*, *26*(1), 31–55.

*Aldashev, A., Gernandt, J., & Thomsen, S. L. (2009). Language usage, participation, employment and earnings: Evidence for foreigners in West Germany with multiple sources of selection. *Labour Economics*, *16*(3), 330–341. https://doi.org/10.1016/j.labeco.2008.11.004

Armstrong, A. (2015). Equilibria and efficiency in bilingual labour market. *Journal of Economic Behavior & Organization*, *112*(C), 204–220. https://doi.org/10.1016/j.jebo.2015.01.011

*Azam, M. C., Chin, A., & Prakash, N. (2013). The returns to English-language skills in India. *Economic Development and Cultural Change*, *61*(2), 335–367. https://doi.org/10.1086/668277

Begg, C. B. (1997). Publication bias in meta-analysis: A Bayesian data-augmentation approach to account for issues exemplified in the passive smoking debate: Comment *Statistical Science*, *12*, 241–244.

*Berman, E., Lang, K., & Siniver, E. (2003). Language-skill complementarity: returns to immigrant language acquisition. *Labour Economics*, *10*(3), 265–290. https://doi.org/10.1016/S0927-5371(03)00015-0

Besevegis, E., & Pavlopoulos, V. (2008). Acculturation patterns and adaptation of immigrants in Greece. In (pp. 23–34): NATO Science for Peace and Security Series, E: Human and Societal Dynamics.

*Bleakley, H., & Chin, A. (2004). Language skills and earnings: Evidence from childhood Immigrants. *Review of Economics and Statistics*, *86*, 481–496. https://doi.org/10.1162/003465304323031067

*Bormann, S.-K., Ridala, S., & Toomet, O. (2019). Language skills in an ethnically segmented labour market: Estonia 1989–2012. *International Journal of Manpower*, *40*(2), 304–327. https://doi.org/10.1108/IJM-06-2017-0115

*Bratsberg, B., & Ragan, J. (2002). The impact of host-country schooling on earnings: A study of male immigrants in the United States. *The Journal of Human Resources*, *37*(1), 63–105. https://doi.org/10.2307/3069604

*Budría, S., & Swedberg, P. (2012). *The impact of language proficiency on immigrants' earnings in Spain*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2170645

Budría, S., Martinez de Ibarreta, C. & Swedberg, P. (2017). The impact of host language proficiency across the immigrants' earning distribution in Spain. *IZA J Develop Migration 7*, 12. https://doi.org/10.1186/s40176-017-0094-2

*Cappellari, L., & Di Paolo, A. (2015). *Bilingual schooling and earnings: Evidence from a language-in-education reform*. Research report 9431, Bonn: Forschungsinstitut zur Zukunft der Arbeit. Retrieved from: https://www.econstor.eu/bitstream/10419/124941/1/dp9431.pdf

*Carnevale, A. P., Fry, R. A., & Lowell, B. L. (2001). Understanding, speaking, reading, writing, and earnings in the immigrant labor market. *American Economic Review*, *91*(2), 159–162. https://doi.org/10.1257/aer.91.2.159

*Casale, D., & Posel, D. (2011). English language proficiency and earnings in a developing country: the case of South Africa. *The Journal of Socio-Economics*, *40*(4), 385–393. https://doi.org/10.1016/j.socec.2011.04.009

*Casey, T., & Dustmann, C. (2008). Intergenerational transmission of language capital and economic outcomes. *Journal of Human Resources*, *43*(3), 660–687. https://doi.org/10.1353/jhr.2008.0002

Castonguay, C. (2019). Quebec's new language dynamic. *Language Problems and Language Planning*, *43*(2), 113–134. https://doi.org/10.1075/lplp.00038.cas

Chan, K. L. (2016). Power Language Index. Which are the world's most influential languages? World Economic Forum on ASEAN. Retrieved from https://www.weforum.org/agenda/2016/12/these-are-the-most-powerful-languages-in-the-world

Chiswick, B. R. (1998). Hebrew language usage: determinants and effects on earnings among immigrants in Israel. *Journal of Population Economics*, *11*(2), 253–271. https://doi.org/10.1007/s001480050068

*Chiswick, B. R., & Larsen, N. (2015). Russian Jewish immigrants in the United States: The adjustment of their English language proficiency and earnings in the American Community Survey. *Contemporary Jewry*, *35*(3), 191–209. https://doi.org/10.1007/s12397-015-9137-2

*Chiswick, B. R., Lee, Y. L., & Miller, P. W. (2005). Immigrant earnings: A longitudinal analysis. *Review of Income and Wealth*, *51*(4), 485–503. https://doi.org/10.1111/j.1475-4991.2005.00165.x

Chiswick, B. R., & Miller, P. W. (1995). The endogeneity between language and earnings: international analyses. *Journal of Labor Economics*, *13*(2), 246–288. https://doi.org/10.1086/298374

*Chiswick, B. R., & Miller, P. W. (2002). *Do enclaves matter in immigrant adjustment?* Research report 449, Bonn: Forschungsinstitut zur Zukunft der Arbeit. Retrieved from: https://www.econstor.eu/bitstream/10419/21507/1/dp449.pdf

*Chiswick, B. R., & Miller, P. W. (2010). Occupational language requirements and the value of English in the US labor market. *Journal of Population Economics*, *23*(1), 353–372. https://doi.org/10.1007/s00148-008-0230-7

*Chiswick, B. R., & Miller, P. W. (2016). *Does bilingualism among the native born pay?* Research report 9791. Bonn: Forschungsinstitut zur Zukunft der Arbeit. Retrieved from: https://www.econstor.eu/bitstream/10419/21507/1/dp449.pdf

Chiswick, B.R., Patrinos, H., & Hurst, M.E. (2000). Indigenous language skills and the labor market in a developing economy: Bolivia. *Economic Development and Cultural Change*, *48*(2), 349–367. https://doi.org/10.1086/452462

*Chiswick, B.R., & Repetto, G. (2001). Immigrant adjustment in Israel: Literacy and fluency in Hebrew and earnings. In S. Djajic (Ed.), *International Migration: Trends, Policy and Economic Impact* (pp. 204–228). New York: Routledge. [Reprint in *The Economics of Language: International Analyses*, edited by Barry R. Chiswick and Paul W. Miller. New York: Routledge, 2007.]

*Chiswick, B.R., & Wang, Z. (2016). *Social contracts, Dutch language proficiency and immigrant economic perfotmance in the Netherlands: A longitudinal study*. Research report 9760. Bonn: Forschungsinstitut zur Zukunft der Arbeit. Retrieved from: http://ftp.iza.org/dp9760.pdf

*Christofides, L.N., & Swidinsky, R. (2010). The economic returns to the knowledge and use of a second official language: English in Quebec and French in the Rest-of-Canada. *Canadian Public Policy*, *36*(2), 137–158. https://doi.org/10.3138/cpp.36.2.137

Cochrane handbook for systematic reviews of interventions. (2008). (J. Higgins & S. Green. Eds.). England: John Wiley & Sons Ltd.

*Cohen-Goldner, S., & Eckstein, Z. (2008). Labor mobility of immigrants: Training, experience, language, and opportunities. *International Economic Review*, *49*(3), 837–872. https://doi.org/10.1111/j.1468-2354.2008.00499.x

Constant, A.F., Kahanec, M., & Zimmermann, K.F. (2012). The Russian-Ukrainian earnings divide. *Economics of Transition*, *20*(1), 1–35. https://doi.org/10.1111/j.1468-0351.2011.00428.x

*Cornwell, K., & Inder, B. (2008). Language and labour markets in South Africa. *Journal of African Economies*, *17*(3), 490–525. https://doi.org/10.1093/jae/ejm037

*Di Paolo, A., & Raymond, J. (2012). Language knowledge and earnings in Catalonia. *Journal of Applied Economics*, *15*(1), 89–118. https://doi.org/10.1016/S1514-0326(12)60005-1

*Di Paolo, A., & Tansel, A. (2015). Returns to foreign language skills in a developing country: The Case of Turkey. *The Journal of Development Studies*, *51*(4), 407–421. https://doi.org/10.1080/00220388.2015.1019482

*Di Paolo, A., & Tansel, A. (2019). *English skills, labour market status and earnings of Turkish women*. Retrieved from https://ideas.repec.org/p/iza/izadps/dp12160.html

Donado, A. (2017). Foreign languages and their impact on unemployment. *Labour*, *31*(3), 265–287. https://doi.org/10.1111/labr.12097

Dreher, A. (2006). Does globalization affect growth? Evidence from a new Index of Globalization. *Applied Economics*, *38*(10), 1091–1110. https://doi.org/10.1080/00036840500392078

Drinkwater, S.J., & O'Leary, N.C. (1997). Unemployment in Wales: Does language matter? *Regional Studies*, *31*(6), 583–591. https://doi.org/10.1080/00343409750131712

*Drydakis, N. (2012). Ethnic identity and immigrants' wages in Greece. *International Journal of Intercultural Relations*, *36*(3), 389–402. https://doi.org/10.1016/j.ijintrel.2011.09.002

Duncan, A., & Mavisakalyan, A. (2015). Russian language skills and employment in the Former Soviet Union. *Economics of Transition*, *23*(3), 625–656. https://doi.org/10.1111/ecot.12075

*Dustmann, C., & Fabbri, F. (2003). Language proficiency and labour market performance of immigrants in the UK. *The Economic Journal*, *113*(489), 695–717. https://doi.org/10.1111/1468-0297.t01-1-00151

*Dustmann, C., & van Soest, A. (2001). Language fluency and earnings: estimations with misspecified indicators. *The Review of Economics and Statistics*, *83*(4), 663–674. https://doi.org/10.1162/003465301753237740

*Dustmann, C., & Van Soest, A. (2002). Language and the earnings of immigrants. *Industrial and Labor Relations Review*, *55*(3), 473–492. https://doi.org/10.1177/001979390205500305

Duval, S. (2005). The trim and fill method. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments*. Chichester, UK: Wiley. https://doi.org/10.1002/0470870168.ch8

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*(2), 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x

Egger, M., Davey Smith, G. M. S. & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. In (Vol. *315*, pp. 629–634): BJM.

Egger, P. H., & Lassmann, A. (2012). The language effect in international trade: A meta-analysis. *Economics Letters*, *116*(2), 221–224. https://doi.org/10.1016/j.econlet.2012.02.018

Ethnologue. (2017). Ethnologue: Languages of the World. Twentieth Edition (SIL International).

Fabo, B., Beblavý, M., & Lenaerts, K. (2017). The importance of foreign language skills in the labour markets of central and Eastern Europe: assessment based on data from online job portals. *Empirica*, *44*(3), 487–508. https://doi.org/10.1007/s10663-017-9374-6

*Gao, W., & Smyth, R. (2011). Economic returns to speaking 'standard Mandarin' among migrants in China's urban labour market. *Economics of Education Review*, *30*(2), 342–352. https://doi.org/10.1016/j.econedurev.2010.11.002

Gazzola, M., & Mazzacani, D. (2019). Foreign language skills and employment status of European natives: evidence from Germany, Italy and Spain. *Empirica*, *46*, 713–740. https://doi.org/10.1007/s10663-019-09460-7

Gazzola, M., & Wickström, B.-A. (2016). *The Economics of Language Policy* (M. Gazzola & B.-A. Wickström. Eds.). Cambridge, MA: The MITT Press. https://doi.org/10.7551/mitpress/9780262034708.001.0001

Ginsburgh, V., & Weber, S. (2016). *The Palgrave Handbook of Economics and Language* (V. Ginsburgh & S. Weber. Eds.): Palgrave Macmillan UK.. https://doi.org/10.1007/978-1-137-32505-1

*Ginsburgh, V., & Prieto-Rodriguez, J. (2007). Returns to foreign languages of native workers in the EU. *Industrial and Labor Relations Review*, *64*(3), 599–617. https://doi.org/10.1177/001979391106400309

*Gonzalez, L. (2005). Nonparametric bounds on the returns to language skills. *Journal of Applied Econometrics*, *20*(6), 771–975. https://doi.org/10.1002/jae.795

Grenier, G. (2019). Quebec's language policy and economic globalization. *Language Problems and Language Planning*, *43*(2), 179–197. https://doi.org/10.1075/lplp.00041.gre

Grin, F. (2003). Language Planning and Economics. *Current Issues in Language Planning*, *4*(1), 1–66. https://doi.org/10.1080/14664200308668048

*Grin, F., & Sfreddo, C. (1998). Language-based earnings differentials on the Swiss labour market: is Italian a liability? *International Journal of Manpower*, *19*(7), 520–532. https://doi.org/10.1108/01437729810237196

*Guven, C., & Islam, A. (2015). Age at migration, language proficiency, and socioeconomic outcomes: evidence from Australia. *Demography*, *52*, 513–542. https://doi.org/10.1007/s13524-015-0373-6

*Hall, M. & Farkas, G. (2008). Does human capital raise earnings for immigrants in the low-skill labor market? *Demography*, *45*(3), 619–639. https://doi.org/10.1353/dem.0.0018

Hanushek, E.A. (1998). *The evidence on class size*, Occasional Paper 98-1, W. Allen Wallis Institute of Political Economy, University of Rochester. Retrieved from http://hanushek.stanford.edu/sites/default/files/publications/Hanushek%201998%20HouseTestimony%20Class%20Size.pdf

*Hayfron, J.E. (2010). Language training, language proficiency and earnings of immigrants in Norway. *Applied Economics*, *33*(15), 1971–1979. https://doi.org/10.1080/00036840010018630

*Hellerstein, J., & Neumark, D. (2003). Ethnicity, language, and workplace segregation: Evidence from a new matched employer-employee data set. *Annales d'Economie et de Statistique*, *71/72*, 19–78. https://doi.org/10.2307/20079047

*Henley, A., & Jones, R.E. (2005). Earnings and linguistic proficiency in a bilingual economy. *Manchester School*, *73*(3), 300–200. https://doi.org/10.1111/j.1467-9957.2005.00448.x

*Hwang, S.-S., Xi, J., & Cao, J. (2010). The conditional relationship between English language proficiency and earnings among US immigrants. *Ethnic and Racial Studies*, *33*(9), 1620–1647. https://doi.org/10.1080/01419871003642375

*Isphording, I. (2013). *Returns to local and foreign language skills: Causal evidence from Spain*. Retrieved from https://EconPapers.repec.org/RePEc:zbw:rwirep:398

Kahanec, M., & Zaiceva, A. (2009). Labor market outcomes of immigrants and non- citizens in the EU: An East-West comparison. *International Journal of Manpower*, *1/2*(30), 97–115. https://doi.org/10.1108/01437720910948429

*Kim, J. (2003). Education, English language proficiency, and earnings of male immigrants in the US labor market. *Journal of Business & Economics Research*, *1*(3), 17–26.

*Klein, C. (2004). *La valorisation des compétences linguistiques: importance du sexe et/ou du statut professionnel?* 11e Journées d'étude sur les données longitudinales, Dijon, 27–28 mai.

*Kroncke, C., & Smith, K. (1999). The wage effects of ethnicity in Estonia. *Economics of Transition*, *7*(1), 179–199. https://doi.org/10.1111/1468-0351.00009

Krueger, A. (2003). Economic Considerations and Class Size. *Economic Journal*, *113*(485), F34–F63. https://doi.org/10.1111/1468-0297.00098

*Lancee, B. (2010). The economic returns of immigrants' bonding and bridging social capital: The case of the Netherlands. *International Migration Review*, *44*(1), 202–226. https://doi.org/10.1111/j.1747-7379.2009.00803.x

*Lang, G., Siniver, E. (2006). *The return to English in a non-English speaking country: Russian immigrants and native Israelis in Israel*. NBER working paper series, Working Paper 12464.

*Lecker, T. (1997). Language usage and earnings among minorities. *Journal of Socioeconomics*, *26*(5), 525–532.

Leping, K.-O., & Toomet, O. (2008). Emerging ethnic wage gap: Estonia during political and economic transition. *Journal of Comparative Economics*, *36*(4), 599–619. https://doi.org/10.1016/j.jce.2008.08.002

*Levanon, A. (2014). Who succeeds as an immigrant? Effects of ethnic community resources and external conditions on earnings attainment. *Research in Social Stratification and Mobility*, *36*, 13–29. https://doi.org/10.1016/j.rssm.2013.11.002

*Levinsohn, J. (2007). Globalization and the returns to speaking English in South Africa. In A. Harrison (Ed.), *Globalization and Poverty* (pp. 629–646): University of Chicago Press. https://doi.org/10.7208/chicago/9780226318004.003.0016

*Lewis, E. G. (2011). *Immigrant-native substitutability: The role of language ability*. Research report 17609. Cambridge: NBER Working Paper Series. Retrieved from: https://www.nber.org/papers/w17609. https://doi.org/10.3386/w17609

Lindemann, K. (2013). The effects of ethnicity, language skills, and spatial segregation on labour market entry success in Estonia. *European Sociological Review*.

*Lindley, J. (2002). The English language fluency and earnings of ethnic minorities in Britain. *Scottish Journal of Political Economy*, *49*(4), 467–487. https://doi.org/10.1111/1467-9485.00242

Mincer, J. (1958). Investment in human capital and personal income distribution. *Journal of Political Economy*, *66*(4), 281–302. https://doi.org/10.1086/258055

*Miranda, A., & Zhu, Y. (2013a). The causal effect of deficiency at English on female immigrants' labor market outcomes in the UK. *IZA Discussion Paper No. 7841*.

*Miranda, A., & Zhu, Y. (2013b). English deficiency and the native-immigrant wage gap. *Economics Letters*, *118*(1), 38–41. https://doi.org/10.1016/j.econlet.2012.09.007

Männasoo, K., Hein, H., & Ruubel, R. (2018). "The contributions of human capital, R&D spending and convergence to total factor productivity growth", *Regional Studies 52* (12): 1598–1611. https://doi.org/10.1080/00343404.2018.1445848

Rendon, S. (2007). The Catalan premium: Language and employment in Catalonia. *Journal of Population Economics*, *20*(3), 669–686. https://doi.org/10.1007/s00148-005-0048-5

Ridala, S. (2020). Big language minority and small language majority: language skills and unemployment in Latvia. Manuscript submitted for publication.

Saarela, J., & Finnäs, F. (2003). Unemployment and native language: the Finnish case. *The Journal of Socio-Economics*, *32*(1), 59–80. https://doi.org/10.1016/S1053-5357(03)00007-6

*Saiz, A., & Zoido, E. (2005). Listening to what the world says: Bilingualism and earnings in the United States. *The Review of Economics and Statistics*, *87*(3), 523–538. https://doi.org/10.1162/0034653054638256

Savina, G., Haelg, F., Potrafke, N. & Sturm, J.-E. (2019): The KOF Globalisation Index – Revisited, Review of International Organizations, *14*(3), 543–574, https://doi.org/10.1007/s11558-019-09344-2

Stanley, T. D. (2001). Wheat from Chaff: Meta-Analysis as Quantitative Literature Review. *The Journal of Economic Perspectives*, *15*(3), 131–150. https://doi.org/10.1257/jep.15.3.131

Sterne, J. A. C. (2009). Meta-analysis on Stata: An updated collection from the Stata Journal. In. College Station: TX: Stata Press.

*Stöhr, T. (2015). The returns to occupational foreign language use: Evidence from Germany. *Labour Economics*, *32*(C), 86–98. https://doi.org/10.1016/j.labeco.2015.01.004

*Toomet, O. (2011). Learn English, not the local language! Ethnic Russians in the Baltic States. *American Economic Review*, *101*(3), 526–531. https://doi.org/10.1257/aer.101.3.526

*Trejo, S. J. (2001). *Intergenerational progress of Mexican-origin workers in the U.S. labor market*. Research report 377. Bonn: Forschungsinstitut zur Zukunft der Arbeit. Retrieved from: https://www.iza.org/publications/dp/377/intergenerational-progress-of-mexican-origin-workers-in-the-us-labor-market

*Wang, H., Smyth, R., & Cheng, Z. (2017). The economic returns to proficiency in English in China. *China Economic Review*, *43*(C), 91–104. https://doi.org/10.1016/j.chieco.2017.01.004

*Warman, C., Sweetman, A., & Goldmann, G. (2015). The portability of new immigrants' human capital: Language, education, and occupational skills. *Canadian Public Policy/Analyse de politiques*, *41*(S1), 64–79. https://doi.org/10.3138/cpp.2013-055

*Williams, D. R. (2011). Multiple language usage and earnings in western Europe. *International Journal of Manpower*, *32*(4), 372–393. https://doi.org/10.1108/01437721111148513
Yao, Y., & van Ours, J. C. (2015). Language skills and labor market performance of immigrants in the Netherlands. *Labour Economics*, *34*, 76–85. https://doi.org/10.1016/j.labeco.2015.03.005
*Zhen, Y. (2015). English proficiency and earnings of foreign-born immigrants in the USA from 1980 to 2000: The effect of minority-language Enclave. *Forum for Social Economics*, *45*(4), 329–349. https://doi.org/10.1080/07360932.2015.1026920
*Zibrowius, M. (2012). *Convergence or divergence? Immigrant wage assimilation patterns in Germany*. Research report 479. DIW Berlin: SOEP – Das Sozio-oekonomische Panel. Retrieved from: https://www.diw.de/documents/publikationen/73/diw_01.c.408060.de/diw_sp0479.pdf. https://doi.org/10.2139/ssrn.2142535
* An asterisk before the name of authors indicates that the study was included in the meta-regression analysis.

Appendix A. Restricted (residual) maximum likelihood (REML)

The restricted (residual) maximum likelihood (REML) estimator has the following log-likelihood functions.

$$\ln L_{REML}(\tau^2) = \ln L_{ML}(\tau^2) - \frac{1}{2}\ln\left\{\sum_{j=1}^{k}(\hat{\sigma}_j^2 + \tau^2)^{-1}\right\} + \frac{\ln(2\pi)}{2}$$

where the maximum likelihood log-likelihood function is:

$$\ln L_{ML}(\tau^2) = -\frac{K\ln(2\pi)}{2} - \frac{1}{2}\sum_{j=1}^{k}\ln(\hat{\sigma}_j^2 + \tau^2) - \frac{1}{2}\sum_{j=1}^{k}\frac{(\hat{\theta}_j - \hat{\theta}^*)^2}{\hat{\sigma}_j^2 + \tau^2}$$

where $\hat{\theta}^*$ is the weighted average of $\hat{\theta}_j$

$$\hat{\theta}^* = \frac{\sum_{j=1}^{k} w_j^* \hat{\theta}_j}{\sum_{j=1}^{k} w_j^*}$$

where the weights are $w_j^* = 1/(\hat{\sigma}_j^2 + \tau^2)$.

Appendix B. References and statistics on primary studies for the meta-regression

Table B1. Summary statistics of the studies included in meta-regression analysis by language groups

Selected studies on returns to English/other business language

Authors, publication year	Number of coefficients	Min	Max	Median	Standard deviation	Mean
Adamchik et al. (2019)	3	0.321	0.363	0.347	0.021	0.344
Azam et al. (2013)	2	0.223	0.345	0.284	0.086	0.284
Bormann et al. (2019)	4	0.054	0.108	0.089	0.025	0.085
Casale & Posel (2011)	1	0.398	0.398	0.398	.	0.398
Chiswick & Miller (2016)	7	−0.07	0.096	0	0.061	0.018
Chiswick & Repetto (2001)	1	0.160	0.160	0.160	.	0.160
Cohen-Goldner & Eckstein (2008)	2	−0.04	0.014	0.011	0.127	0.05
Cornwell & Inder (2008)	1	0.177	0.177	0.177	.	0.177
Di Paolo & Tansel (2015)	3	0.064	0.112	0.064	0.028	0.080
Di Paolo & Tansel (2019)	2	0.194	0.251	0.223	0.040	0.223
Drydakis (2012)	1	0.015	0.015	0.015	.	0.015
Ginsburgh & Rodriguez (2007)	9	0.139	0.401	0.252	0.084	0.258
Grin & Sfreddo (1998)	4	0.09	0.1	0.1	0.005	0.098
Isphording (2013)	18	−0.012	0.483	0.179	0.175	0.222
Klein (2004)	3	0.003	0.032	0.021	0.015	0.018
Kroncke & Smith (1999)	4	0.039	0.205	0.124	0.092	0.123
Lang & Siniver (2006)	1	0.013	0.013	0.013	.	0.013
Levinsohn (2007)	2	0.183	0.252	0.218	0.049	0.218
Saiz & Zoido (2005)	3	0.017	0.04	0.027	0.012	0.028

Authors, publication year	Number of coefficients	Min	Max	Median	Standard deviation	Mean
Stöhr (2015)	2	0.120	0.261	0.191	0.100	0.191
Toomet (2011)	2	0.124	0.624	0.384	0.339	0.384
Wang et al. (2017)	7	0.057	0.194	0.075	0.053	0.107
Williams (2011)	26	−0.067	0.321	0.124	0.097	0.120
Total	**108**	**−0.067**	**0.624**	**0.160**	**0.134**	**0.153**

Note: Languages: German, English, French, Greek, Italian, Japanese, Portuguese, Spanish; countries: Austria, Belgium, China, Denmark, Estonia, Finland, France, Germany, Greece, India, Ireland, Israel, Italy, Latvia, Luxembourg, Netherlands, Poland, Portugal, South Africa, Spain, Switzerland, Turkey, UK, US

Selected studies on returns to state/local language

Authors, publication year	Number of coefficients	Min	Max	Median	Standard deviation	Mean
Aldashev et al. (2009)	1	0.044	0.044	0.044	.	0.044
Berman et al. (2003)	1	0.065	0.065	0.065	.	0.065
Bleakley & Chin (2004)	2	0.222	0.334	0.278	0.079	0.278
Bormann et al. (2019)	2	−0.009	0.066	0.029	0.053	0.029
Bratsberg & Ragan (2002)	2	0.161	0.210	0.185	0.035	0.185
Budría & Swedberg (2012)	2	0.048	0.273	0.161	0.159	0.161
Carnevale et al. (2001)	1	0.144	0.144	0.144	.	0.144
Casey & Dustmann (2008)	4	−0.012	0.749	0.173	0.340	0.271
Chiswick & Larsen (2015)	4	0.234	0.300	0.245	0.030	0.256
Chiswick & Miller (2002)	2	0.148	0.592	0.37	0.314	0.37

Authors, publication year	Number of coefficients	Min	Max	Median	Standard deviation	Mean
Chiswick & Miller (2010)	4	0.013	0.041	0.024	0.013	0.025
Chiswick & Repetto (2001)	1	−0.094	−0.094	−0.094	.	−0.094
Chiswick & Wang (2016)	1	0.06	0.06	0.06	.	0.06
Chiswick et al. (2005)	3	0.194	0.235	0.199	0.022	0.209
Cohen-Goldner & Eckstein (2008)	2	0.1	0.11	0.105	0.007	0.105
Dustmann & Fabri (2003)	6	0.121	0.463	0.259	0.123	0.273
Dustmann & Van Soest (2001)	1	0.312	0.312	0.312	.	0.312
Dustmann & Van Soest (2002)	4	0.042	0.141	0.085	0.050	0.088
Gao & Smith (2011)	6	0.041	0.508	0.198	0.213	0.236
Gonzalez (2005)	1	0.19	0.19	0.19	.	0.19
Guven & Islam (2015)	1	0.281	0.281	0.281	.	0.281
Hall & Farkas (2008)	1	0.075	0.075	0.075	.	0.075
Hayfron (2010)	1	−0.075	−0.075	−0.075	.	−0.075
Hellerstein & Neumark (2003)	2	−0.019	0.003	−0.008	0.016	−0.008
Hwang et al. (2010)	1	0.100	0.100	0.100	.	0.100
Isphording (2013)	18	−0.081	0.328	0.190	0.121	0.153
Kim (2003)	2	0.173	0.266	0.220	0.066	0.220
Klein (2004)	9	−0.013	0.018	0.011	0.010	0.001
Lancee (2010)	1	0.115	0.115	0.115	.	0.115
Lang & Siniver (2006)	1	0.1	0.1	0.1	.	0.1

Authors, publication year	Number of coefficients	Min	Max	Median	Standard deviation	Mean
Lecker (1997)	3	0.132	0.199	0.151	0.035	0.161
Levanon (2014)	2	0.122	0.132	0.127	0.007	0.127
Lewis (2011)	2	0.149	0.209	0.179	0.042	0.179
Lindley (2002)	6	−0.267	0.334	0.188	0.218	0.139
Miranda &Zhu (2013a)	2	−0.275	−0.191	−0.233	0.060	−0.233
Miranda &Zhu (2013b)	2	−0.231	−0.158	−0.195	0.052	−0.195
Stöhr (2015)	1	0.097	0.097	0.097	.	0.097
Toomet (2011)	2	0.005	0.209	0.107	0.144	0.107
Trejo (2001)	4	−0.087	−0.042	−0.062	0.020	−0.063
Warman et al. (2015)	4	−0.046	0.56	0.329	0.252	0.293
Williams (2011)	5	0.048	0.292	0.204	0.099	0.196
Zhen (2015)	3	0.161	0.363	0.196	0.108	0.240
Zibrowius (2012)	1	0.015	0.015	0.015	.	0.015
Total	**124**	**−0.276**	**0.749**	**0.137**	**0.165**	**0.137**

Note: Languages: German, Dutch, English, Estonian, French, Hebrew, Luxembourgish, Latvian, Mandarin, Norwegian, Spanish; countries: Australia, Belgium, Canada, China, Estonia, Germany, Israel, Latvia, Luxembourg, Netherlands, Norway, Spain, UK, US

Selected studies on returns to significant minority language

Authors, publication year	Number of coefficients	Min	Max	Median	Standard deviation	Mean
Bormann et al. (2019)	2	0.016	0.043	0.030	0.020	0.030
Cappellari & Di Paolo (2015)	1	0.219	0.219	0.219	.	0.219
Chiswick & Repetto (2001)	1	−0.024	−0.024	−0.024	.	−0.024
Christofides & Swidinsky (2010)	4	−0.071	0.135	0.031	0.115	0.031
Di Paolo & Raymond (2012)	7	0.038	0.21	0.134	0.060	0.126

Authors, publication year	Number of coefficients	Min	Max	Median	Standard deviation	Mean
Di Paolo & Tansel (2015)	1	−0.019	−0.019	−0.019	.	−0.019
Grin & Sfreddo (1998)	4	0.06	0.14	0.08	0.346	0.09
Henley & Jones (2003)	1	−0.040	−0.040	−0.040	.	−0.040
Williams (2011)	5	−0.012	0.277	0.074	0.126	0.131
Total	**26**	**−0.071**	**0.277**	**0.063**	**0.094**	**0.073**

Note: Languages: Arabic, Catalan, German, English, French, Italian, Russian, Spanish, Welsh; countries: Canada, Denmark, Estonia, France, Israel, Italy, Luxembourg, Spain, Switzerland, Turkey, UK

Note: Some initial studies on language skills and labour market performance report the language coefficients from several models, including different controls for standard personal characteristics and human capital variables, such as age, education, and family status, and so only the coefficients from the model with the largest number of controls have been included.

Table B2. Sample descriptive table of primary studies

Authors, publication year	Language, (group nr.)	Country	sex
Adamchik et al. (2019)	English (1)	Poland	men, women, pooled
Aldashev et al. (2009)	German (2)	Germany	men
Azam et al. (2013)	English (1)	India	men, women
Berman et al. (2003)	Hebrew (2)	Israel	men
Bleakley & Chin (2004)	English (2)	US	pooled
Bormann et al. (2019)	English (1), Russian (3), Estonian (2)	Estonia	men, women

Authors, publication year	Language, (group nr.)	Country	sex
Bratsberg & Ragan (2002)	English (2)	US	men
Budría & Swedberg (2012)	Spanish (2)	Spain	men
Cappellari & Di Paolo (2015)	Catalan (3)	Spain	pooled
Carnevale et al. (2001)	English (2)	US	men
Casale & Posel (2011)	English (1)	South Africa	men
Casey & Dustmann (2008)	German (2)	Germany	men, women
Chiswick & Larsen (2015)	English (2)	US	men, women
Chiswick & Miller (2002)	English (2)	US	men
Chiswick & Miller (2010)	English (2)	US	men
Chiswick & Miller (2016)	Portuguese (1), French (1), Japanese (1), German (1), Spanish (1), Greek (1), Italian (1)	US	men
Chiswick & Repetto (2001)	English (1), Arabic (3), Hebrew (2)	Israel	men
Chiswick & Wang (2016)	Dutch (2)	Netherlands	men

Authors, publication year	Language, (group nr.)	Country	sex
Chiswick et al. (2005)	English (2)	Australia	men
Christofides & Swidinsky (2010)	French (3), English (3)	rest of Canada, Quebec	men, women
Cohen-Goldner & Eckstein (2008)	Hebrew (2), English (1)	Israel	men
Cornwell & Inder (2008)	English (1)	South Africa	pooled
Di Paolo & Raymond (2012)	Catalan (3)	Spain	men, women, pooled
Di Paolo & Tansel (2015)	English (1), French (1), German (1), Arabic (3)	Turkey	men
Di Paolo & Tansel (2019)	English (1)	Turkey	men, women
Drydakis (2012)	English (1)	Greece	pooled
Dustmann & Fabri (2003)	English (2)	UK	pooled
Dustmann & Van Soest (2001)	German (2)	Germany	men
Dustmann &Van Soest (2002)	German (2)	Germany	men, women
Gao & Smith (2011)	Mandarin (2)	China	men, women, pooled

Authors, publication year	Language, (group nr.)	Country	sex
Ginsburgh & Rodriguez (2007)	English (1)	Austria, Denmark, Finland, France, Germany, Greece, Italy, Portugal, Spain	pooled
Gonzalez (2005)	English (2)	US	pooled
Grin & Sfreddo (1998)	English (1), French (3), (Swiss) German (3)	Switzerland	men
Guven & Islam (2015)	English (2)	Australia	pooled
Hall & Farkas (2008)	English (2)	US	men
Hayfron (2010)	Norwegian (2)	Norway	men
Hellerstein & Neumark (2003)	English (2)	US	men
Henley & Jones (2003)	Welsh (3)	UK	pooled
Hwang et al. (2010)	English (2)	US	pooled
Isphording (2013)	Spanish (2), English (1), German (1), French (1)	Spain	men, women
Kim (2003)	English (2)	US	men
Klein (2004)	English (1), German (2), French (2), Luxembourgish (2)	Luxembourg	men, women, pooled
Kroncke & Smith (1999)	English (1)	Estonia	men
Lancee (2010)	Dutch (2)	Netherlands	pooled

Authors, publication year	Language, (group nr.)	Country	sex
Lang & Siniver (2006)	English (1), Hebrew (2)	Israel	pooled
Lecker (1997)	Hebrew (2)	Israel	men
Levanon (2014)	English (2)	US	men
Levinsohn (2007)	English (1)	South Africa	pooled
Lewis (2011)	English (2)	US	pooled
Lindley (2002)	English (2)	UK	men, women
Miranda &Zhu (2013a)	English (2)	UK	women
Miranda & Zhu (2013b)	English (2)	UK	men
Saiz & Zoido (2005)	Spanish (1), French (1), German (1)	US	pooled
Stöhr (2015)			
Toomet (2011)	English (1), Estonian (2)	Estonia	men
Toomet (2011)	English (1), Latvian (2)	Latvia	men
Trejo (2001)	English (2)	US	men
Wang et al. (2017)	English (1)	China	men, women, pooled
Warman et al. (2015)	English (2), French (2)	Canada	men, women

Authors, publication year	Language, (group nr.)	Country	sex
Williams (2011)	English (1), German (3)	Denmark	pooled
Williams (2011)	English (1), French (1), German (1)	Netherlands	pooled
Williams (2011)	English (1), French (2), German (2), Dutch (2)	Belgium	pooled
Williams (2011)	English (1), French (2), German (2), Italian (3)	Luxembourg	pooled
Williams (2011)	English (1), German, (1) Spanish (3), Italian (3)	France	pooled
Williams (2011)	French (1), German (1), Spanish (1)	UK	pooled
Williams (2011)	French (1), German (1)	Ireland	pooled
Williams (2011)	English (1), French (1), German (3)	Italy	pooled
Williams (2011)	English (1), French (1), German (1)	Greece	pooled
Williams (2011)	English (1), French (1)	Spain	pooled
Williams (2011)	English (1), French (1)	Portugal	pooled
Williams (2011)	English (1), French (1)	Austria	pooled
Williams (2011)	English (1), German (1)	Finland	pooled
Zhen (2015)	English (2)	US	pooled
Zibrowius (2012)	German (2)	Germany	men

Appendix C. Descriptive statistics and bubble plots on the meta-regression

Table C1. Description of the variables and summary statistics for the meta-regression

Name and description	Mean	Sth.Dev	Min	Max	N
Dependent variable definitions					
Estimated coefficients on returns to language skills to earnings Return from the knowledge of English/ other business language, state/ local language, minority language to the natural logarithm of hourly, weekly, monthly or annual earnings in the Mincer-type wage equation	0.138	(0.147)	−0.276	0.749	258
Independent variable definitions					
PLI Power Language Index (PLI) uses 20 indicators in five categories (1) geography, (2) economy, (3) communication, (4) knowledge and media and (5) diplomacy; and ranks languages on ascaleof 0–1 by their global influence. Reference: Chan (2016)	0.575	(0.339)	0.017	0.889	258
LDI The Linguistic Diversity Index (LDI) measures the diversity of languages spoken in a country. The scale ranges from 0 to 1. An index of 0 represents no linguistic diversity, meaning that everyone speaks the same language. An index of 1 represents total diversity, meaning that no two people speak the same language. No country has an index value of exactly 0 or Reference: Ethnologue, 2017	0.387	(0.190)	0.05	0.914	258
KOFGI	0.758	(0.182)	0	1	258

Name and description	Mean	Sth.Dev	Min	Max	N
KOF Globalisation Index (KOFGI) a composite index measuring globalisation for every country in the world along the economic, social and political dimensions. Originally measured on a scale of 1–100, but rescaled with linear transformation onto a scale of 1–0. Linear transformation formula: [KOF-min(KOF)]/[max(KOF)-min(KOF)] References: Savina et al., 2019, Dreher, 2006.					
Urban The urban population (% of total population), people living in urban areas as defined by national statistical offices. Data are collected and smoothed by the United Nations Population Division, https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS	75.31	(10.453)	29.2	97.1	258
GDP per capita In thousand USD. This indicator provides per capita values for gross domestic product (GDP) expressed in current international dollars converted by the purchasing power parity (PPP) conversion factor. GDP is the sum of the gross value added by all resident producers in the country plus any product taxes and minus any subsidies not included in the value of the products. The conversion factor is a spatial price deflator and currency converter that controls for price level differences between countries. Total population is a mid-year population based on the de facto definition of population, which counts all residents regardless of legal status or citizenship.	28.62	(12.477)	3.039	55.36	258

Name and description	Mean	Sth.Dev	Min	Max	N
Unemployment Unemployment (% of active population), the share of the labour force involuntarily not in employment. Source: International Labour Organization, ILOSTAT database.	7.686	(4.020)	0.56	30.2	258
IV/OLS Dummy variable = 1 if ordinary least squares (OLS) estimates and =0 if instrumental variable (IV) estimates	0.775	(0.418)	0	1	258
Panel/Cross-section Dummy variable = 1 if panel data structure and = 0 if cross-section	0.457	(0.499)	0	1	258
Pooled/Men/Women = 1 if pooled, = 2 if men and = 3 if women	1.806	(0.723)	1	3	258

Bubble plots on meta-regression

Figures C1, C2, C3 and C4 below plot the effect of the business, state and minority languages on earnings in relation to the variables for the Power Language Index (PLI), the Linguistic Diversity Index (LDI), the urban population as a percentage of the total population, and real GDP per capita, along with the fitted meta-regression line. The circles on the figures represent the estimates from each study, sized according to the precision of each estimate, as the inverse of its within-study variance, σ_i^2.

Figure C1 explains the relationship between the Power Language Index (PLI) and the returns of the business, state and minority languages for earnings, showing that knowing more powerful or influential languages leads to better income prospects. The fitted simple regression lines are upward sloped for all the three categories of language returns studied, meaning the returns to language skills are increasing in the Power Language Index (PLI) for all language groups. This implies that the returns from learning a powerful language are remunerative even if the native speakers of that language form a minority group in a given country.

The effects are relatively flat in the Linguistic Diversity Index for the business language and for the local or state language, but are decreasing (Figure C2) in the returns to the minority language.

Figure C3 shows that the degree of urbanisation is associated with lower returns to the business language, but no clear evidence arises for the minority or state or local languages. It appears that the widespread use of English and other business languages in areas of high population density actually reduces the income effect from knowing these widely spoken languages.

Figure C4 illustrates how returns to the minority language are higher in more affluent economies, but are lower for the business language. The evidence for the business language probably arises because of high-quality schooling, which guarantees a strong supply in knowl-

edge of the business languages. In less developed economies in contrast, knowledge of languages is scarce, and premiums are awarded to those who know business languages. The positive effect of the GDP level upon skills in minority languages may reflect the impact of the diversified markets in developed economies, which are better at utilising the specific skills of multilingual societies.

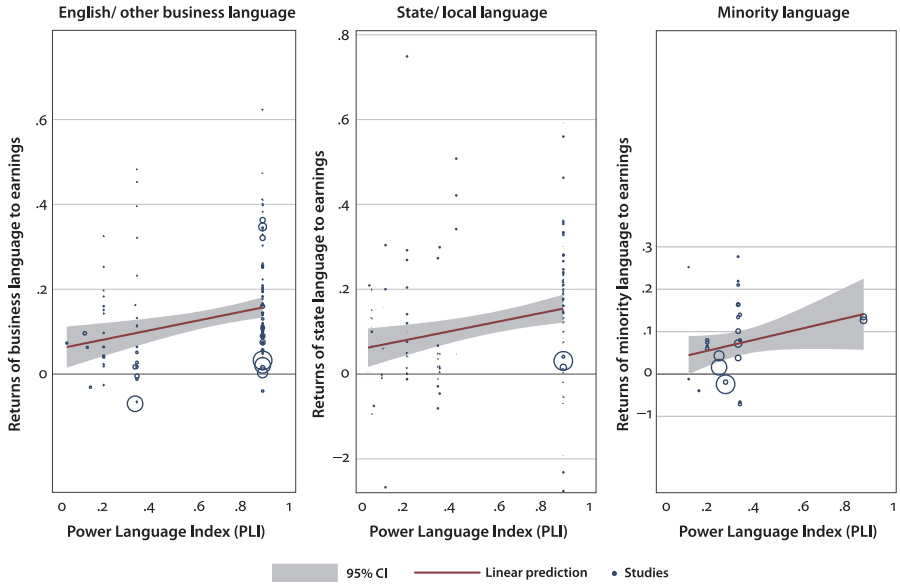


Figure C1. Bubble plot with fitted meta-regression line for the effect of English/other business, state/local and minority language on earning: Power Language Index (PLI)
Notes: Weights: Inverse-variance. The bubble sizes denote the within-study precision (as the inverse of within-study variance) of the coefficient estimates. *Source:* Author's calculations

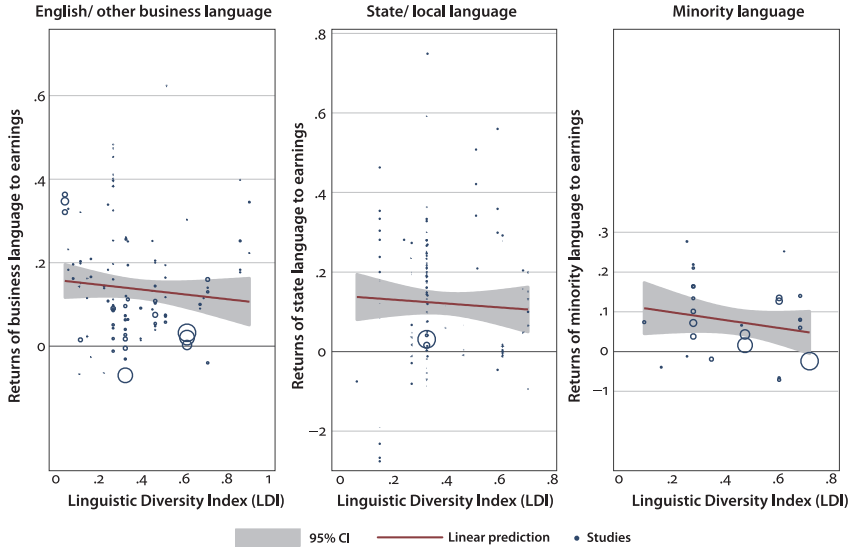


Figure C2. Bubble plot with fitted meta-regression line for the effect of English/other business, state/local and minority language on earnings: Linguistic Diversity Index (LDI)
Notes: Weights: Inverse-variance. The bubble sizes denote the within-study precision (as the inverse of within-study variance) of the coefficient estimates. *Source:* Author's calculations

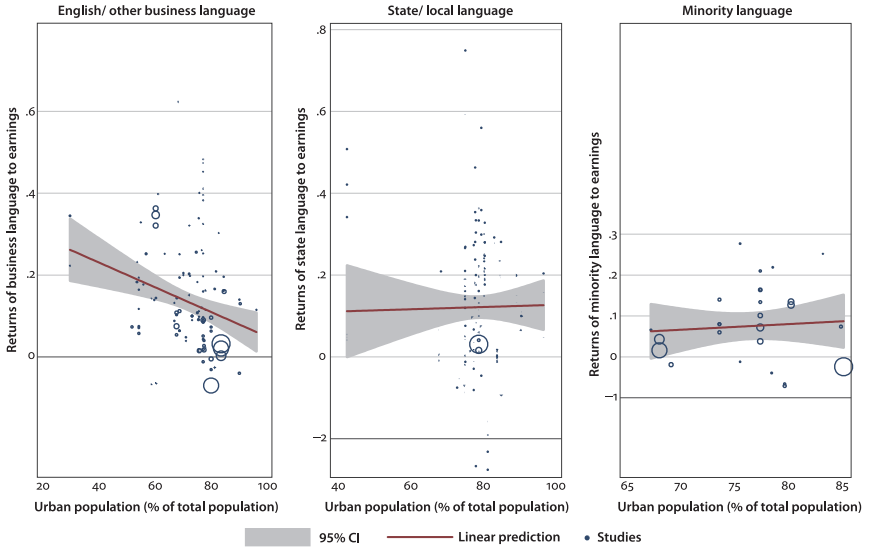


Figure C3. Bubble plot with fitted meta-regression line of the effect for English/other business, state/local and minority language on earnings: urban population as % of total population
Notes: Weights: Inverse-variance. The bubble sizes denote the within-study precision (as the inverse of within-study variance) of the coefficient estimates. *Source:* Author's calculations

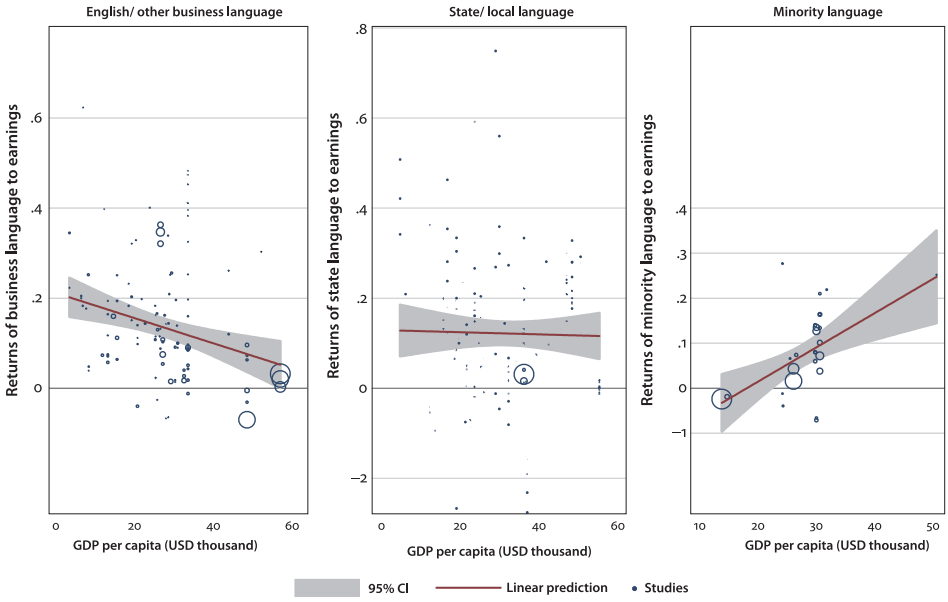


Figure C4. Bubble plot with fitted meta-regression line for the effect of English/other business, state/local and minority language on earnings: real GDP per capita
Notes: Weights: Inverse-variance. The bubble sizes denote the within-study precision (as the inverse of within-study variance) of the coefficient estimates. *Source:* Author's calculations

Резюме

Увеличение миграционных трендов, глобализация и внедрение цифровых платформ для рынков труда, требуют более глубокого понимания механизмов, улучшающих экономические результаты и результаты рынка труда в условиях растущего культурного, языкового и индивидуального неравенства. Статья вносит вклад в литературу об экономической отдаче от владения языком, которая является очень разнородной и контекстно-ориентированной, проводя систематический обзор этой литературы. Мета-регрессионный анализ оценивает отдачу от языковых навыков, обусловленную социально-экономическими, институциональными и этнолингвистическими факторами, включая проверку на дизайн исследований. Результаты мета-регрессии, о влиянии знания государственного языка, языка Значительного меньшинства и делового языка на доходы на рынке труда, свидетельствуют о том, что знание более влиятельных языков ассоциируется с более высоким вознаграждением на рынке труда. Языковое разнообразие оказывает негативное влияние на доход от знания делового языка, но в то же время увеличивает отдачу на знание местного государственного языка. Урбанизация оказывает отрицательное, хотя и скромное по размеру влияние на вознаграждение к деловому языку, в то время как уровень ВВП увеличивает отдачу на языки меньшинств, а безработица увеличивает отдачу на знание делового и государственного языков.

Resumo

Kreskinta migrado, tergloba komerco kaj la enkonduko de ciferecaj laboraj platformoj postulas pli bonan komprenon de la mekanismoj kiuj povas plifortigi ekonomiajn kaj labormerkatajn rezultojn fronte al kreskantaj malsamecoj de kulturo, lingvo kaj identeco. Per sistema trarigardo de la koncerna literaturo, la nuna artikolo kontribuas al la literaturo pri labormerkataj redonoj pro lingvaj kapabloj – temo tre heterogena kaj kuntekstospecifa. La metaregresa analizo taksas la redonojn pro lingvaj kapabloj fontantajn el sociekonomiaj, instituciaj kaj etnolingvistikaj faktoroj, kune kun kontroloj pro la studodezajno. La metaregresaj rezultoj pri labormerkataj redonoj rezultantaj el scipovo de la loka ŝtata lingvo, signifa loka minoritata lingvo aŭ la lingvo de komerco liveras atestaĵojn, ke scio de la pli influa lingvo asociiĝas kun pli altaj labormerkataj rekompencoj. Lingva diverseco negative efikas ĉe redonoj pro kapabloj en la komerca lingvo sed pligrandigas la redonojn pro kapabloj en la loka ŝtata lingvo. Urbiĝo havas negativan, kvankam kvante modestan, efikon je redonoj pri lingvaj kapabloj en la komerca lingvo, dum la nivelo de Malneta Enlanda Produkto (GDP) kreskigas la redonojn pro kapabloj en minoritata lingvo, kaj senlaboreco kreskigas la redonojn pro kapabloj en la komerca lingvo kaj la loka ŝtata lingvo.

JEL classification: J7, J31, Z13

Address for correspondence

Svetlana Ridala
Tallinn University of Technology
School of Business and Governance
Department of Economics and Finance
Akadeemia tee 3
Tallinn 12618
Estonia
svetlana.ridala@gmail.com

Curriculum vitae

Personal data

Name:	Svetlana Ridala
Date of birth:	28.03.1976
Place of birth:	Smolensk, Russia
Citizenship:	Estonian

Contact data

E-mail: svetlana.ridala@gmail.com

Education

Graduation	Institution	Field, degree
(2021)	TalTech	Economics, PhD
2003	Tallinn University	Mathematics, MSC
2000	Tallinn University	Mathematics and Physics, BSC
1997	Virumaa College	Mathematics, Physics and Informatics, Primary school teacher

Language competence

Language	Level
Russian	Native
Estonian	Fluent
English	Fluent
Swedish	Basic skills

Special courses

Period	Course	Educational organisation
July 2010	Panel Data Linear Analysis, Microeconometric Lab	Barcelona GSE
February 2011	"Monetary and Fiscal Policy in the EU"	Kiel Institute for the World Economy
July 2011	"Dynamic and Non-linear Panel Data Models"	Barcelona GSE
autumn 2014	seminars at Real Estate and Construction Management and Centre for Banking and Finance	KTH Royal Institute of Technology
autumn 2014	visiting researcher at the Department of Industrial Economics and Management	KTH Royal Institute of Technology

Professional employment

Period	Organisation, position
2018–…	Recommender OÜ, International business and partners
2018–…	Svetlana Ridala, Private investor
2015–2016	TalTech, Department of Economics and Finance, Lecturer (0.3)
2010–2015	TalTech, Department of Economics and Finance, Researcher
2008–2010	TalTech, Department of Economics and Finance, Lecturer

2004–2009	Tallinn School of Economics, Mathematics and statistics teacher
2002–2002	Paekivitoodete Tehase OÜ, Secretary
2001–2002	Lasnamäe District Administration, Senior assistant
2000–2001	Tallinna Sidekool, Informatics teacher

Publications and working papers

Ridala, S. (2021). Language skills and labour market returns: A meta-regression analysis.Language Problems and Language Planning, vol. 44, no. 2, pp. 200–241. DOI https://doi.org/10.1075/lplp.20012.rid (ETIS 1.1)

Bormann, S.-K.; Ridala, S.; Toomet, O.-S., 2019. Language Skills in an Ethnically Segmented Labour market: Estonia 1989-2012. *International Journal of Manpower*, vol. 40, no. 2, pp. 304–327. DOI: https://doi.org/10.1108/IJM-06-2017-0115 (ETIS 1.1)

Ridala, S.; Toomet, O.-S., 2019. Wage gap in Ethnically Segmented Labour Market: The Role of Cognitive Skills. *Eastern European Economics*, vol. 57, no. 1, pp. 20–30. DOI: https://doi.org/10.1080/00128775.2018.1524713 (ETIS 1.1)

Kornilov, S.; Ridala, S.; Aasma, A. (2016). Dynamics of Economic Adjustment under uncertainty: MS-VAR model for Baltic Dry Index. In 3rd International Multidisciplinary Scientific Conference on Social Sciences and Arts SGEM 2016. SGEM2016 Conference Proceedings, 24–31 August, 2016: 3rd International Scientific Conferences on Social Sciences & Arts S G E M 2016, 22 - 31 August, 2016, Varna (Bulgaria). Varna: International Multidisciplinary Scientific Conference on Social Sciences and Arts, 189−196. (Book 2; 5) (ETIS 3.1)

Bormann, S.-K.; Ridala, S.; Toomet, O. (2015). Language Skills in an Ethnically Segmented Labor market: Estonia 1989 – 2012. Fourth World Conference SOLE – EALE 2015: 4th Joint meeting of SOLE – EALE, Montreal, Canada, 26-28 June 2015. The Society of Labor Economists, 1−16. (ETIS 3.4)

Ridala, S.; Toomet, O. (2015). Wage gap in an Ethnically Segmented Labour Market: The Case of Estonia. The 7th international conference “Economic Challenges in Enlarged Europe”, Conference Proceedings, 14-16. June 2015: The 7th international conference "Economic Challenges in Enlarged Europe", Tallinn, 14-16. June 2015. Tallinn: Tallinn Technical University. (ETIS 3.4)

Ridala, S.; Aasma, A. (2012). Consumption in the Baltic States: myopia or liquidity con-straints? In: Proceedings of 30-th International Conference Mathematical Methods in Economics: Mathematical Methods in Economics 2012, Karvina, Czech Republic; 11-13 September 2012. (Eds.) Ramik, J; Stavarek, D. Karvina, Czech Republic; Silesian University, 763–768. (ETIS 3.1)

Elulookirjeldus

Isikuandmed

Nimi:	Svetlana Ridala
Sünniaeg:	28.03.1976
Sünnikoht:	Smolensk, Venemaa
Kodakondsus:	Eesti

Kontaktandmed

E-post: svetlana.ridala@gmail.com

Hariduskäik

Lõpetamine	Õppeasutus	Eriala, kraad
(2021)	TalTech	Majandusteooria, doktorikraad
2003	Tallinna Ülikool	Matemaatika, teadusmagister
2000	Tallinna Ülikool	Matemaatika ja Füüsika, loodusteaduse bakalaureus
1997	Virumaa Kõrgkool	Matemaatika, Füüsika, Informaatika, Põhikooli õpetaja

Keelteoskus

Keel	Tase
Vene keel	Emakeel
Eesti keel	Kõrgtase
Inglise keel	Kõrgtase
Rootsi keel	Algtase

Täiendusõpe

Õppimise aeg	Täiendusõpe	Täiendusõppe korraldaja nimetus
Juuli 2010	Paneelandmete lineaaranalüüs, Microeconometric Lab	Barcelona GSE, Hispaania
Veebr. 2011	Rahandus- ja Fiskaalpoliitika EL-s	Kiel Institute for the World Economy, Saksamaa
Juuli 2011	Dünaamilised ja mittelineaarsed paneeliandmed	Barcelona GSE, Hispaania
sügissemester 2014	seminarid Real Estate and Construction Management and Centre for Banking and Finance	KTH Royal Institute of Technology, Rootsi
sügissemester 2014	külalisteadur Department of Industrial Economics and Management	KTH Royal Institute of Technology, Rootsi

Teenistuskäik

Töötamise aeg	Tööandja, ametikoht
2018–…	Recommender OÜ, rahvusvaheline äri ja partnerid
2018–…	Svetlana Ridala, erainvestor
2015–2016	TalTech, Rahanduse ja majandusteooria instituut, külalislektor (0.3)
2010–2015	TalTech, Rahanduse ja majandusteooria instituut, teadur
2008–2010	TalTech, Rahanduse ja majandusteooria instituut, lektor
2004–2009	Tallinna Majanduskool, majandusmatemaatika ja statistika õpetaja
2002–2002	Paekivitoodete Tehase OÜ, sekretär
2001–2002	Lasnamäe Linnaosa Valitsus, Lasnamäe linnaosa vanema abi
2000–2001	Tallinna Sidekool, informaatika õpetaja

Publikatsioonid ja toimetised

Ridala, S. (2021). Keeleoskus ja tööturu väljundid: meta-regressioonanalüüs. Language Problems and Language Planning, vol. 44, no. 2, pp. 200–241. DOI https://doi.org/10.1075/lplp.20012.rid (ETIS 1.1)

Bormann, S.-K.; Ridala, S.; Toomet, O.-S., 2019. Keeleoskus etniliselt segmenteeritud tööturul: Eesti 1989-2012. *International Journal of Manpower*, vol. 40, no. 2, pp. 304–327. DOI: https://doi.org/10.1108/IJM-06-2017-0115 (ETIS 1.1)

Ridala, S.; Toomet, O.-S., 2019. Palgaerinevus etniliselt segmenteeritud tööturul: kognitiivsete oskuste roll. *Eastern European Economics*, vol. 57, no. 1, pp. 20–30. DOI: https://doi.org/10.1080/00128775.2018.1524713 (ETIS 1.1)

Kornilov, S.; Ridala, S.; Aasma, A. (2016). Majandusliku kohanemise dünaamika ebakindluse all: Baltic Dry Index MS-VAR mudel. 3rd International Multidisciplinary Scientific Conference on Social Sciences and Arts SGEM 2016. SGEM2016 Conference Proceedings, 24-31 August, Varna (Bulgaria), 189−196. (Book 2; 5) (ETIS 3.1)

Bormann, S.-K.; Ridala, S.; Toomet, O. (2015). Keeleoskus etniliselt segmenteeritud tööturul: Eesti 1989-2012. Fourth World Conference SOLE – EALE 2015: 4th Joint meeting of SOLE – EALE, Montreal, Canada, 26-28 June 2015. The Society of Labor Economists, 1−16. (ETIS 3.4)

Ridala, S.; Toomet, O. (2015). Palgaerinevus etniliselt segmenteeritud tööturul: kognitiivsete oskuste roll. The 7th international conference "Economic Challenges in Enlarged Europe", Conference Proceedings, 14-16. June 2015: The 7th international conference "Economic Challenges in Enlarged Europe", Tallinn, 14-16. June 2015. Tallinn: Tallinn Technical University. (ETIS 3.4)

Ridala, S.; Aasma, A. (2012). Tarbimine Balti riikides: müoopia või likviidsuspiirangud? In: Proceedings of 30-th International Conference Mathematical Methods in Economics: Mathematical Methods in Economics 2012, Karvina, Czech Republic; 11-13 September 2012. (Eds.) Ramik, J; Stavarek, D. Karvina, Czech Republic; Silesian University, 763–768. (ETIS 3.1)

ISSN 2585-6901 (PDF)
ISBN 978-9949-83-665-9 (PDF)