



EQUALITAS Working Paper No. 7

Is there a Gender Bias in the Use of Foreign Languages in Europe?

Victor Ginsburgh

Juan Prieto-Rodriguez

2012

Is there a Gender Bias in the Use of Foreign Languages in Europe?

Victor Ginsburgh
ECARES, Brussels and CORE, Louvain-la-Neuve
Av. F.D., Roosevelt, 39, 1050 Bruxelles, Belgium
vginsbur@ulb.ac.be

Juan Prieto-Rodriguez
Department of Economics, University of Oviedo
Facultad de Economía y Empresa
Av. Del Cristo s/n, 33006 Oviedo, Spain
Tel +(34) 985103768
Fax +(34) 985104871

Abstract

The process of globalization has given incentives to workers (essentially white collars) to learn foreign languages and break linguistic barriers. The purpose of this paper is to evaluate the effect of the most important foreign languages used at the workplace on wages of men and women in Northern and Southern Europe and to study whether there are gender differences related to this human capital. We found an important gender bias linked to discrimination, or at least the existence of a glass ceiling, under different forms in France, Italy, Spain and Denmark, but less so in other Northern countries.

JEL Codes: J16, J31, J70, Z10

Keywords: Foreign language; gender wage differentials; wage structure; human capital

1. Introduction

The process of globalization has forced or given incentives to workers (essentially white collars) to learn foreign languages and break linguistic barriers. In most cases, Europeans have chosen English as a second language, and (a poor version of) English has become the *lingua franca* in Europe. In 2006, 37 percent of the EU population speak it in a more than “basic” way; German and French come next with 22 and 17 percent.¹ There are nevertheless many countries, including large ones such as Spain, France, Italy and Poland where less than 25 percent of the population speak the *lingua franca*. This situation is changing fast now, since among the youngest generations, the number of English speakers has become much larger.² Worldwide, English is spoken by more than 1.5 billion people (Crystal 2001), and is probably the language that is most often used in international contacts and trade.

But English is not the only language that has an economic impact on international trade. This is shown by Melitz (2008), for instance, who uses two measures of linguistic distances³ between trading partners and tries to estimate their separate effect on foreign trade. “Open-circuit communication” (OCC) requires the language to be either official or widely spoken (at least 20 percent of the population knows the language). Spanish, for instance, will be an OCC between Argentina (where 92 percent of the population knows Spanish) and Mexico (where this percentage is 88). Likewise, Arabic will be an OCC between Mauritania (38 percent) and Iraq (58 percent).⁴ Melitz identifies fifteen such languages. A “direct communication” (DC) language is any language common (that is, spoken by at least four percent in each country) in a pair of countries. The rationale for introducing this second measure is based on Melitz's

¹ These numbers include native speakers.

² See Table 2 in Fidrmuc, Ginsburgh and Weber (2009).

³ The linguistic distance between two countries is equal to zero if they share the same language (Germany and Austria for example), and to one if they cannot communicate at all. Intermediate cases are possible; Italians and French can understand each other. For details, see Ginsburgh and Weber (2011, chapter 3).

⁴ Numbers between brackets are those that appear in Table A1 of Melitz (2008).

claim that any linguistic distance measure works in explaining the intensity of trades, but not all of them tell the same story. Noting that all OCCs are also DC languages but OCCs may have a larger impact on international trade, he suggests distinguishing channels through which the effect takes place, and separates therefore OCCs' additional effects that may depend on translation (which can be produced as long as there are enough people who can provide it in both countries) and DCs' effects which enable traders to communicate directly. His estimation results point to the following conclusions. "Direct communication" has obviously a large positive effect on trades: A ten percent increase in the probability that two citizens, one in country *A*, the other in *B*, speak the same language increases their trades by ten percent. European⁵ OCCs also contribute, but with an additional effect. However, and interestingly enough, Melitz shows that English is not more effective than other European languages in promoting trade once the percentage of the population that knows it is controlled for.

Several papers analyse the effect that mastering the language of the destination country has for immigrants.⁶ However, this strand of the literature usually focuses on the economic value of just one language, the one that is spoken in the destination country. Ginsburgh and Prieto-Rodriguez (2011) examine the importance of returns on *several* languages used at the working place by male native workers in nine countries of the European Union. In all nine countries, language proficiency has a positive effect on earnings, but results are heterogeneous. There are differences between Northern (Austria, Denmark, Finland and Germany) and Southern Europe (France, Greece, Italy, Portugal and Spain). In Northern Europe, English is used at the workplace by some 20 to 50 percent of the individuals, and benefits from larger (log) returns than other languages that are less spoken, but also less often used at the workplace. Maybe there are linguistic alternatives to English in Northern Europe

⁵ Other OCC languages such as Arabic, Chinese, Hindi, Malay, Farsi and Turkish have no effect on trade.

⁶ See, for instance, Bleakley and Chin (2004), Bratsberg, Ragan and Nasir (2002), Chiswick and Miller (1995), Dustmann and Van Soest (2002), Dustmann (1999) and Leslie and Lindley (2001). In this paper, however, we concentrate on native workers only.

but they are much less attractive than English. In Southern Europe, languages that are less known than English may also get larger rewards. And this is indeed the case for German in France, French in Greece and Spain, French and German in Italy, French and Spanish in Portugal. Four countries in this list (France, Italy, Portugal and Spain) share Romance languages. Though English is also the most used language by firms, any other Romance language, especially French, is an alternative to English. This dichotomy between the two parts of Europe is consistent with Melitz's (2008) empirical conclusion that, in general, English is not the only effective European language in promoting trade.

There is a surprising lack of papers that examine the gender impact of foreign languages on wages of workers. Most papers are focused on male workers (Ginsburgh and Prieto, 2011 or Bratsberg, Ragan and Nasir, 2002), merely use a dummy to control for a possible gender effect (Bleakley and Chin, 2004 and Gonzalez, 2005), or estimate independent equations for female and male workers without exploring all the implications of gender (Dustmann and Van Soest, 2002).

The purpose of this paper is to evaluate the effect of the most important foreign languages used at the workplace on wages of men and women, both at the mean and along the wage distribution, in six European countries: Denmark, Finland, France, Germany, Italy and Spain.⁷ We find that the foreign language premium is present for both genders. However, the point estimate of the mean is always smaller for women, though the differences are hardly statistically significantly different from 0, with the exception of Germany (at the 10 percent probability level). The story is different along the wage distribution. In Finland and Germany, there are no significant gender differences of the language premium along the wage distribution, and the premium varies somewhat, but stays close to the mean, both for men and

⁷ Sweden is included in the European Community Household Panel survey that we use, but the questions on languages are not included.

women. This is almost the case in Denmark, though the return for males slowly rises with wages and the difference between men and women becomes significant in the higher wage quantiles. This is more marked in France, Italy and Spain, where premia differ across genders. In France returns on foreign language used at job are stable along all the wage distribution for both men and women but we find a statistically significant gender difference of some 20 percent at all quantiles. A different pattern prevails in Italy and Spain, where returns on foreign languages increase substantially for men (but not for women) in the upper quantiles leading to a significant and quite large gender difference at the upper quantiles. Discrimination, or at least the existence of a glass ceiling linked to this specific type of human capital, is thus present under different forms in France, Italy, Spain and Denmark although much less so in other North European countries such as Germany or Finland.

The paper is organized as follows. Section 2 discusses the model and the econometric issues in estimating its parameters using instrumental variables. Section 3 turns to the results obtained through usual IV and IV quantile regression methods, and discusses their economic relevance. Conclusions are drawn in Section 4.

2. The Model: Estimation and Data Used

Our purpose is to estimate the effect of language knowledge (and use at the workplace) on earnings. The standard (language-augmented) Mincer-type equation can be specified as:

$$(1) \quad \ln w_i = x_i\beta + D_i\gamma + u_i$$

for individuals $i = 1, 2, \dots, N$, where the vector β and the scalar γ are parameters, w_i represents the wage rate, x_i is a vector of exogenous variables and u_i is a random error. The

vector x_i contains the following control variables: two dummy variables that represent higher and secondary education (the omitted variable being primary or no education, respectively); the number of years of job tenure and its square; the number of years of potential experience (number of years spent working after schooling) and its square. Note that this specification does not include sector or occupation as controls. If knowing foreign languages is a prerequisite to securing a job in occupations such as international trade or international finance, then one would observe high wages in these occupations, but these would be due to language ability since workers with the same qualification but who do not know foreign languages would be working in other sectors or occupations. Therefore, controlling for occupation will underestimate the real value of language usage at the workplace.⁸

D_i is usually a dummy variable that takes the value 1 if individual i reports being proficient in a specific language, and 0 otherwise, as in Galasi (2003) or Williams (2006). Here we use a different specification: D_i stands for the share of the population that *ignores* the language used by the worker at his workplace. This so-called “disenfranchisement rate,” varies between 1 if nobody (in a given country) knows the language and 0 if everyone knows it.⁹ This is a measure of the scarcity of the supply of the language in a given country. Therefore, when an individual mentions a language used at her job, the disenfranchisement rate measures not only her knowledge (the supply side), but also what the firm needs, that is, the demand side and, therefore, we deal with the matching of workers to jobs. It may happen that a language with a high disenfranchisement rate is also one for which there are few opportunities to find a job where this language is requested, but if the worker finds a matching firm, his reward may nevertheless be important. This will have an influence on the estimation procedure since our definition of the language variable is not a pure human capital variable. Table 1 lists the

⁸ Albrecht et al. (2003, p 171) find that when they account for occupation the gender gap at the top of the distribution falls substantially. But they also argue, “that including occupation [...] is really another way of showing the glass ceiling effect, which manifests itself partly through occupational segregation.”

⁹ See Ginsburgh and Weber (2005).

values of disenfranchisement rates for the five main languages spoken by workers in the six countries of interest.

[Table 1 approximately here]

The impact of languages on wages is represented by a unique (endogenous) scalar variable D_i (or a polynomial of D_i). Dealing with several endogenous variables may lead to inefficient estimation (especially since some of the disenfranchisement variables are, on average, close to one—only a small percentage of Finnish workers speak Italian, for example) and the difficulty of finding enough instruments. The scalar nature of D_i will also be useful in our quantile instrumental variable estimation procedure. The expected value for parameter γ in (1) is positive, which implies that the larger the disenfranchisement rate, the higher the expected return on the language. This formulation has the advantage that all foreign languages are subsumed by a unique variable, while in every country, the effect of a language can easily be retrieved by multiplying γ by the disenfranchisement rate of the language.

Equation (1) is subject to unobserved heterogeneity similar to the one faced in estimating the returns on education. Indeed, both education and earnings may depend on unobservable individual skills and talent. Some right hand-side variables, in our case D_i , will be correlated with the error term, leading to biased ordinary least squares estimates. The solution is to use instruments.

Dustmann and Van Soest (2001 and 2002) discuss a second issue, related to misclassified language indicators. In panel (or cross-section) data, language ability is usually self-reported by the worker, and is affected by two types of errors: a purely random error, that is independent over time and an error that is time-persistent, since an individual may have the

same tendency to over- or under-report over time.¹⁰ To deal with the time-independent measurement error, Dustmann and Van Soest suggest using leads and lags of self-reported language fluency as instruments for current fluency. They note, however, that this does not eliminate time-persistent errors. These are probably of limited importance in our case, since the language reported is the one used at the workplace, and the answer can hopefully be considered more objective. Hence, we expect estimates of languages returns to be unbiased if this kind of problem is corrected by the use of lagged instruments. This leads us using $D_{i,t-1}$ as instrument.

We also expect higher wages to be linked to better occupations that often require managing in more than one language. People involved in the international strategy of a company should be able to use non-native languages in their job. Therefore, the use of foreign languages could be considered a necessary condition for promotion to some of the best positions. If this is true, including occupation dummies would lead to underestimate language returns and estimating the effects at the mean of the wage distribution may not be sufficient to capture the whole picture. Quantile regression could therefore be more appropriate, since it allows studying how returns vary at different points of the distribution of earnings. This has the additional advantage of allowing a better control of workers heterogeneity even if, as is the case in our GMM estimations, occupation is not included in the estimations. The reason is that unobserved characteristics will tend to be more similar around a specific wage quantile and, thus, occupational differences will tend to be smaller. Here, we use Chernozhukov and Hansen's (2004, 2005 and 2006) instrumental variable quantile regression estimator.

However, foreign languages knowledge is not a sufficient condition to reach top occupations. Not all workers who are competent in foreign languages will occupy top jobs. Beyond the

¹⁰ Dustmann and Van Soest (2001) formulate a model that allows isolating the two types of errors if panel data are used to estimate the returns equation.

relative scarcity of this skill, this may be due to gender differences in promotion policies. According to the glass ceiling hypothesis, the probability of being promoted, despite using more than one language in the workplace, would be much lower among women, all other observable characteristics being equal. In conclusion, if the wage increases associated with the use of foreign languages in the workplace are lower for women, especially at the very top of the wage distribution, this would be empirical evidence of confirming the glass ceiling hypothesis.¹¹

The database consists of the European Community Household Panel (ECHP), which contains information on panels of individuals in 15 European countries from 1994 to 2001 and which was also used in Ginsburgh and Prieto (2011). This information is homogenous across countries since the surveys were coordinated by EUROSTAT, although the sample sizes vary across countries and years, and worse for our purpose, the questions on foreign language practices change over time (or are not introduced in the survey). The surveys contain the socio-economic characteristics of individuals older than 16, grouped by households; they include personal characteristics, family structure, current employment, education and training, labour status, wages, family income from sources other than wages and salaries, region of residence, and languages (main and second) used at the main job. Since the panel is consistent after 2000, we decided to use the waves 2000 and 2001 only. We could have run panel techniques, but preferred to use the information contained in the 2000 survey to use $D_{i,t-1}$, the lagged value of disenfranchisement, to control, as mentioned earlier, for non-persistent time components of the error term.¹²

¹¹ Quantile regression has been used several times to analyze wage differences across genders to test for the glass ceiling (and the stick floor) hypothesis. See, for instance, Albrecht et al. (2003), Arulampalam et al. (2007), Gardeazabal and Ugidos (2005) and Machado and Mata (2005).

¹² Since we also run IV quantile regressions, this allows us to compare directly the results of IV and quantile estimation.

We include only natives, even if they do not have the citizenship of the country where they were born. Since most of them should have attended the national schooling system (or domestic international schools), we assume that they know the official language of the country. Table 2 contains information on the number of observations in each country, as well as on the share of individuals who report needing one or several among the main European languages, that is English, French, German, Italian, or Spanish,¹³ in addition to the official language of the country.

[Table 2 approximately here]

It is worth noting that the shares of workers of both genders, who use foreign languages at the workplace are more than twice larger in the two Nordic countries. However, there is no clear pattern with regard to the observed differences between genders. According to the t-test reported in Table 2, the gap between the male and female shares of workers using a foreign language at work is not significantly different from zero in Finland and Italy although their levels are three times larger in Finland. In Denmark, France and Germany, a significant larger percentage of male workers use more than one language at job. Finally, only in Spain, do women use foreign language significantly more often than men, according to the estimated t-test. Among those who use foreign languages, the average disenfranchisement rate in the sample is always very slightly larger for women, but the difference is statistically significant in France and Italy only.

Hence, there is some empirical evidence that female workers especially in romance language countries, specialize in foreign languages. Spain is the only analysed country with a share of workers using foreign languages at the workplace larger for female workers than for males;

¹³ If more than one foreign language is used at the job, we assign the highest disenfranchisement rate (that is the one of the “rarest” language) to this variable.

meanwhile France and Italy are the only two countries where the disenfranchisement rate is significantly higher for women than for men.

3. Results

We use two distinct econometric techniques to estimate the returns on foreign languages. We first estimate Equation (1) using the Generalized Method of Moments (GMM), which is efficient in the case of heteroskedastic errors. This will give us the returns at the mean of the wage distribution. Then we turn to estimate the same returns at different points of the conditional distribution of log wages. Since workers in the same quantile of the conditional wage distribution can be expected to have similar unobserved characteristics (in particular, occupation), quantile regression may help to control for unobserved heterogeneity without over-controlling for occupation or sector. Increasing returns along quantiles could be a signal of a positive correlation between languages and unobserved heterogeneity. Although standard quantile regression could limit this endogeneity problem, it is not clear that it will completely remove it. Research by Chernozhukov and Hansen (2004, 2005, 2006) has extended the quantile regression approach to deal explicitly with endogeneity. They propose an instrumental variable quantile estimator which is naturally robust to weak identification and that we use here, with the same specification and instruments as in the GMM regressions.

Instrumental variable (IV) results are described in Table 3. The upper part of the table gives the results after instrumentation. The variable of interest is linguistic disenfranchisement. In Denmark, Finland and Germany, the relation between log returns is quadratic (strictly concave and in each case (with the exception of women in Finland) both coefficients are highly significantly different from zero. In the three other (romance language) countries the

relation is linear, and highly significant so as well.¹⁴ In addition to the disenfranchisement rate, the regressions also include a small number of standard control variables, such as education (positive and significant influence), years of tenure (positive and significant influence, concave function in some cases), and potential experience (significantly concave curve in all cases, except for women in Finland). The lower part of the table presents the regression coefficient of $D_{i,2001}$ on $D_{i,2000}$ (and other exogenous variables) in the regression in which $D_{i,2001}$ is instrumented (first stage of IV estimations). The R-squared illustrate that, except for Germany where the number of switches between 2000 and 2001 is the smallest, the differences in reporting language use between the two waves of the survey are quite important.

[Table 3 approximately here]

One important question is whether the returns inferred from these estimated coefficients are statistically different between genders. In Table 3, we include a Wald test to check whether the estimated disenfranchisement parameters are equal across genders. For the quadratic specification, this test follows a χ^2 distribution with two degrees of freedom. According to the results of the tests, gender differences are statistically significant only for Germany (at the 7.5 percent level). Given the values of the disenfranchisement variable D_i for German workers, this implies that German men get a higher return than women when they use English at the workplace but the difference is smaller when they use other foreign languages.¹⁵

Table 4 shows the returns by language. These are obtained by multiplying the regression coefficients picked up by “disenfranchisement” and “disenfranchisement squared,” by the

¹⁴ We also tried the quadratic specification, but came out non significant. Results are available upon request.

¹⁵ In fact, according to the correspondent t-test, the difference between the average (log) hourly wages of German men and women who use foreign language other than English is not significant. However, these workers represent less than 1 per cent of the country’s sample; thus their economic relevance can be considered limited.

disenfranchisement rates reported in Table 1. These should be taken with care, since in some countries, the number of workers who are observed speaking some languages (German in Finland for example) is quite small. As expected, the estimated returns on English are smaller in Denmark and Finland (where English is quite well known) than in France, Germany and Spain. However, they are also modest in Italy. In France, the returns on German and Spanish are larger than those on English, and so are those on French and German in Italy and Spain. Moreover, the number of workers who use these languages at the workplace, although smaller than those who use English, is not negligible. This is consistent with the findings by Melitz (2008) that English is not always the only communication language used in international transactions.

[Table 4 approximately here]

Returns obtained by women are systematically smaller than those obtained by men with the exception of German in Finland (but the number of speakers of German is quite small to draw solid inferences).¹⁶ However, the differences are usually small and, as discussed above, not statistically significant. Thus, despite the significant differences in the shares of female and male workers using foreign languages at the workplace in Denmark, France, Spain and Germany, estimation results at the mean show little evidence of returns between genders being different.

For female workers, the effects of foreign language proficiency on earnings of GMM and IVQ estimations are often very similar. In most cases, using foreign languages generates significant wage premia (20 to 50 percent), except for Danish women. Also, the females' wage premium is fairly stable over the entire distribution of wages, which implies that the

¹⁶ Returns on German in Denmark are negative due to the very negative coefficient for disenfranchisement squared, and the quite large number of speakers of German.

returns are similar for all quantiles without rising at the upper end of the distribution, as is the case of men. In fact, in five of the six countries there is a smooth declining trend for the returns of women in the highest quantile (detailed figures can be obtained from the authors).

For men, the returns on foreign language are significantly rising in the upper quantiles of the wage distribution in Denmark. This is also true for Italy and Spain. In these Southern countries, the access of men to the most exclusive occupations seems to be linked with their knowledge of languages, maybe because these countries experience the highest English disenfranchisement rates and restrictions on language training are very important. In the other countries the pattern is quite stable over the wage distribution but usually the language premium for men is larger than for women (detailed figures can be obtained from the authors).

We now look at how the difference in returns between men and women varies at different points of the conditional distribution of (the log of) wages. Figure 1 summarizes these differences by country using the results of instrumental variable quantile (IVQ) regressions (Chernozhukov and Hansen 2004, 2005, 2006). These are run by dividing the wage distribution into 19 quantiles, both for men and women.¹⁷ We assume that the estimations for men and women are independent and compute accordingly a 95 percent confidence interval for their difference.¹⁸ The results are given in Figure 1, where the continuous line represents the gender difference in the language premium, which is picked up by the disenfranchisement rate along the distribution of wages. The shaded regions represent the 95 percent confidence intervals around the gender difference; they show that the estimated coefficients are significantly different at some points of the wage distribution in all countries with the exception of Finland and Germany.

¹⁷ Only the linear term of the disenfranchisement rate is included.

¹⁸ The standard deviation of the difference is therefore simply equal to the square root of the sum of the variances.

[Figure 1 approximately here]

Figure 1 also shows that the returns on foreign languages are higher for men than for women in certain areas of the wage distribution but the reverse happens only in one of the bottom quantiles in Spain. In Northern Europe, the gender differences are hardly significant: only Denmark exhibits significant differences at the very top of the distribution.

The differences are more important in France and Spain where returns for men are larger in the middle range of the wage distribution as well. In France, they do not seem to be related to the glass ceiling hypothesis since the wage gap does not increase in the top of the wage distribution and is similar to the (significant) difference observed in other quantiles. In Spain, in the upper quantiles men can more than double their wages if they use a foreign language with a high disenfranchisement rate, while returns for women remain more or less stable with an estimated 27.5 percent premium for English. Hence, in this country gender differences are significant in the central quantiles of the wage distribution but they become very important at the top. In Italy and Denmark these differences are also significant at the top of the distribution, but less important. For instance, the expected difference in the English premium for Danish men and women at the very top quantiles is 10 percent; it rises to 50 percent in Italy but it is three times larger in Spain. Both observations lead to suspect that some discrimination against women is going on, or that men do not accept these higher positions without being paid more, while women do. Since this happens especially at the very top of the income distribution, the findings may give support the glass ceiling hypothesis in Italy and Spain.

There are thus some differences between genders in France, Italy and Spain and less marked in Denmark. Again, it is hard to infer whether this is correlated with the stylized facts mentioned in the introduction concerning the differences in the returns on languages between Northern Europe, where English is the *lingua franca*, and Southern Europe which also uses other languages than English. But it is remarkable that as we move south, gender differences become more important.

4. Conclusions

Our paper shows that languages other than the national tongue used at the workplace lead to important returns on wages both for native male and female workers. As shown in many papers, this is also so for (usually male) immigrants who learn the national language of the country to which they move, but the returns they generate by knowing the national language are not as large. This finding is not unexpected since there is less market pressure to pay higher wages to immigrants who learn and speak the national language, than to nationals.

We use two techniques to evaluate these returns: Instrumental variables regression, which estimate the returns at the mean of the wage distribution, and instrumental variables quantile regression to estimate the returns at several quantiles of this conditional distribution. This made it possible to isolate two types of differences between men and women. In France, women with the same language abilities than men earn systematically (and statistically significantly) less than men, in all quantiles. In Italy and Spain the same phenomenon is present, but in Italy only at the upper quantiles. In Germany and Finland, none of the two types of gender differences could be detected. In Denmark, despite the significant increase in the upper quantiles of returns for men, gender differences associated to foreign languages are not very large. Additionally, we find that in romance language countries, women seem to

specialize in languages more than men: Spain is the only country with a share of workers using foreign languages that is larger for women than for men; meanwhile France and Italy are the only two countries where the disenfranchisement rate is significantly higher for women than men. This indicates the possible existence of wage discrimination in all three Southern countries, since two workers using the same foreign language at their job will receive different wage premia depending on their gender. We also detected a glass ceiling for female workers in Spain and Italy and to some extent in Denmark.

References

- Albrecht, J., Björklund, A., & Vroman, S. (2003). Is there a glass ceiling in Sweden? *Journal of Labor Economics*, 21(1), 145-177.
- Arulampalam, W., Alison, L. B., & Bryan, M. L. (2007). Is there a glass ceiling over Europe? Exploring the gender pay gap across the wage distribution. *Industrial and Labor Relations Review*, 60(2), 163-186.
- Bleakley, H., & Chin, A. (2004). Language skills and earnings: Evidence from childhood immigrants. *The Review of Economics and Statistics*, 86(2), 481-496.
- Bratsberg, B., Ragan, J. F., Jr., & Nasir, Z. M. (2002). The effect of naturalization on wage growth: A panel study of young male immigrants. *Journal of Labor Economics*, 20(3), 568-597.
- Crystal, D. (2001). *A dictionary of language*. Chicago: University of Chicago Press.
- Chernozhukov, V., & Hansen, C. (2004). The Effects of 401(k) Participation on the Wealth Distribution: An Instrumental Quantile Regression Analysis. *The Review of Economics and Statistics*, 86(3), 735-751.
- Chernozhukov, V., & Hansen, C. (2005). An IV Model of Quantile Treatment Effects. *Econometrica*, 73(1), 245-261.
- Chernozhukov, V., & Hansen, C. (2006). Instrumental quantile regression inference for structural and treatment effect models. *Journal of Econometrics*, 132(2), 491-525.
- Chiswick, B. R., & Miller, P. W. (1995). The endogeneity between language and earnings: international analyses. *Journal of Labor Economics*, 13(2), 246-288.
- Dustmann, C. (1999). Temporary Migration, Human Capital, and Language Fluency of Migrants. *Scandinavian Journal of Economics*, 101(2), 297-314, doi:10.1111/1467-9442.00158.
- Dustmann, C., & van Soest, A. (2001). Language fluency and earnings: Estimation with misclassified language 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.
- Fidrmuc, J., Ginsburgh, V., & Weber, S. (2009). Voting on the choice of core languages in the European Union. *European Journal of Political Economy*, 25(1), 56-62, doi:10.1016/j.ejpoleco.2008.07.002.
- Galasi, P. (2003). Estimating wage equations for Hungarian higher-education graduates. *Budapest Working Papers on the Labour Market, Institute of Economics, Hungarian Academy of Sciences*(0304).
- Gardeazabal, J., & Ugidos, A. (2005). Gender wage discrimination at quantiles. *Journal of Population Economics*, 18(1), 165-179.
- Ginsburgh, V., & Prieto-Rodriguez, J. (2011). Returns to foreign languages of native workers in the European Union. *Industrial & Labor Relations Review*, 64(3), 599-618.
- Ginsburgh, V., & Weber, S. (2005). Language disenfranchisement in the European Union. *Journal of Common Market Studies*, 43(2), 273-286, doi:10.1111/j.0021-9886.2005.00555.x.
- Ginsburgh, V., & Weber, S. (2011). *How many languages do we need?: The economics of linguistic diversity*. Princeton, NJ: Princeton University Press.
- Gonzalez, L. (2005). Nonparametric bounds on the returns to language skills. *Journal of Applied Econometrics*, 20(6), 771-795.
- Leslie, D., & Lindley, J. (2001). The impact of language ability on employment and earnings of Britain's ethnic communities. *Economica*, 68(272), 587-606.

- Machado, J. A. F., & Mata, J. (2005). Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of Applied Econometrics*, 20(4), 445-465, doi:10.1002/jae.788.
- Melitz, J. (2008). Language and foreign trade. *European Economic Review*, 52(4), 667-699.
- Williams, D. R. (2006). The economic returns to multiple language usage in Western Europe. *IRISS Working Paper Series*, 2006-07.

Table 1 Disenfranchisement Rates (in %)

	Denmark	Finland	France	Germany	Italy	Spain
English	25	39	58	46	61	64
French	95	99	-	84	71	81
German	63	93	92	-	96	98
Italian	100	100	95	99	-	98
Spanish	98	99	85	98	97	-

Source: Ginsburgh and Weber (2005, p. 279).

Table 2 Basic Data by Gender

	Denmark	Finland	France	Germany	Italy	Spain
Men (number)	1,029	928	1,649	2,383	2,214	2,401
% using a foreign language at work	61.4	48.4	24.8	17.8	14.5	12.5
Women (number)	980	871	1,399	1,893	1,500	1,463
% using a foreign language at work	51.1	51.9	20.7	14.9	16.4	15.7
Gender difference t-test	4.67	-1.49	2.68	2.59	-1.60	-2.67
Ho: diff = 0 Ha: diff \neq 0 (Prob > t)	0.000	0.137	0.007	0.009	0.109	0.008
Conditional on using foreign languages at work						
Men disenfranchisement rate (in %)	32.0	40.2	63.2	47.6	68.8	71.8
Women disenfranchisement rate (in %)	32.9	40.3	65.3	47.8	72.3	72.6
Gender difference t-test	-0.97	-0.20	-2.09	-0.20	-2.91	-0.77
Ho: diff = 0 Ha: diff \neq 0 (Prob > t)	0.331	0.843	0.037	0.844	0.004	0.443

Table 3 GMM Estimation Results

	Denmark		Finland		France		Germany		Italy		Spain	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
Second stage												
Disenfranchisement	1.4818*** (0.425)	0.5720*** (0.211)	0.9142*** (0.273)	0.3985 (0.276)	0.5057*** (0.058)	0.3040*** (0.068)	1.1604*** (0.135)	0.7301*** (0.148)	0.2927*** (0.054)	0.2123*** (0.042)	0.6077*** (0.118)	0.4197*** (0.086)
Disenfranchisement squared	-2.6389*** (0.844)	-1.0030** (0.406)	-0.8582* (0.501)	0.1281 (0.472)	--	--	-1.2883*** (0.205)	-0.5964*** (0.226)	--	--	--	--
Higher Education	0.3123*** (0.032)	0.2861*** (0.030)	0.2830*** (0.033)	0.2587*** (0.036)	0.3675*** (0.028)	0.4174*** (0.028)	0.2850*** (0.032)	0.3375*** (0.034)	0.5654*** (0.029)	0.5420*** (0.030)	0.3485*** (0.028)	0.4422*** (0.028)
Secondary Education	0.1148*** (0.032)	0.1307*** (0.029)	0.0615** (0.027)	0.0868*** (0.033)	0.0457 (0.042)	0.1108*** (0.040)	0.1120*** (0.027)	0.1386*** (0.030)	0.1720*** (0.014)	0.2611*** (0.019)	0.1330*** (0.020)	0.1776*** (0.025)
Tenure	0.0132** (0.006)	0.0181*** (0.004)	0.0002 (0.007)	0.0131** (0.006)	0.0275*** (0.005)	0.0233*** (0.006)	0.0148*** (0.005)	0.0303*** (0.005)	0.0117*** (0.003)	0.0095** (0.005)	0.0144*** (0.004)	0.0364*** (0.005)
Tenure sq. (x100)	-0.0488** (0.023)	-0.0606*** (0.014)	0.0294 (0.028)	-0.0430* (0.026)	-0.0360* (0.020)	-0.0047 (0.024)	-0.0014 (0.019)	-0.0684*** (0.022)	-0.0133 (0.013)	0.0085 (0.018)	0.0188 (0.018)	-0.0583*** (0.022)
Pot. Experience	0.0254*** (0.004)	0.0215*** (0.003)	0.0165*** (0.003)	0.0101*** (0.003)	0.0272*** (0.003)	0.0175*** (0.003)	0.0374*** (0.004)	0.0216*** (0.004)	0.0265*** (0.002)	0.0260*** (0.003)	0.0209*** (0.002)	0.0268*** (0.003)
Pot. exper. sq. (x100)	-0.0425*** (0.008)	-0.0378*** (0.007)	-0.0237*** (0.008)	-0.0097 (0.007)	-0.0494*** (0.007)	-0.0398*** (0.008)	-0.0727*** (0.008)	-0.0433*** (0.009)	-0.0447*** (0.005)	-0.0481*** (0.007)	-0.0351*** (0.004)	-0.0534*** (0.007)
Intercept	1.7236*** (0.063)	1.8029*** (0.043)	1.6454*** (0.047)	1.5597*** (0.053)	1.5335*** (0.030)	1.4976*** (0.034)	1.3786*** (0.052)	1.2666*** (0.057)	1.3053*** (0.023)	1.1824*** (0.033)	1.2468*** (0.022)	0.9543*** (0.030)
R-squared	0.2367	0.3083	0.2957	0.2435	0.3624	0.2764	0.2544	0.1919	0.4266	0.3793	0.4034	0.4466
Testing for the equality of the estimated disenfranchisement parameters between genders												
χ^2	3.89		2.08		2.25		5.16		1.18		1.29	
Prob > χ^2	0.1431		0.1817		0.1337		0.0757		0.2767		0.2564	
First step												
Disenfranchisement lagged	0.4900*** (0.037)	0.5676*** (0.040)	0.6114*** (0.033)	0.4901*** (0.037)	0.635*** (0.029)	0.654*** (0.023)	0.914*** (0.012)	0.931*** (0.007)	0.616*** (0.034)	0.722*** (0.0226)	0.449*** (0.043)	0.540*** (0.028)
R-squared	0.2442	0.3053	0.4785	0.4298	0.457	0.397	0.887	0.913	0.337	0.426	0.222	0.255
No. of observations	1029	980	928	871	1649	1399	2383	1893	2214	1500	2401	1463

The dependent variable is the (logged) wage rate. IV standard errors appear between brackets under the coefficients.

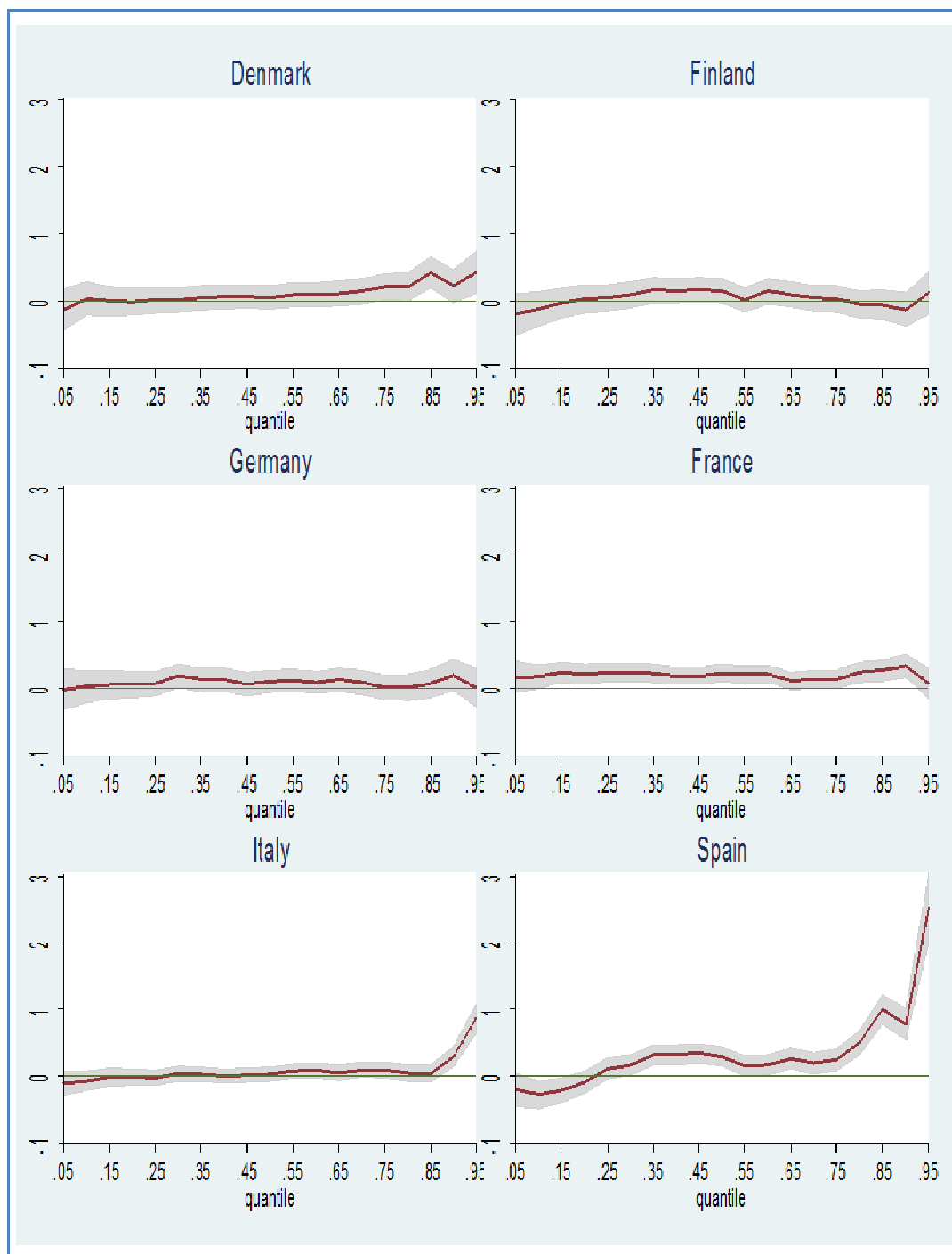
* p<0.10, ** p<0.05, *** p<0.01

Table 4 Returns on Languages Inferred From Disenfranchisement Rates

	Denmark		Finland		France		Germany		Italy		Spain	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
English	0.21 (520)	0.08 (405)	0.23 (439)	0.17 (441)	0.29 (344)	0.18 (225)	0.26 (411)	0.21 (272)	0.18 (218)	0.13 (132)	0.39 (189)	0.27 (136)
French	-	-	-	-	-	-	-	-	0.21 (44)	0.15 (49)	0.49 (85)	0.34 (70)
German	-0.11 (107)	-0.04 (86)	0.11 (9)	0.48 (10)	0.47 (44)	0.28 (42)	-	-	0.28 (52)	0.20 (62)	0.58 (27)	0.40 (23)
Spanish	-	-	-	-	0.43 (14)	0.26 (16)	-	-	-	-	-	-

Returns on languages for which the number of users at the workplace is smaller than 1% of the country's sample are not displayed. Numbers of workers are shown between brackets.

Figure 1
Gender differences in language returns



Note: The vertical axis is measured in hundred percent; 1, 2 and 3 = 100, 200 and 300 percent.