



Universidad del Rosario
Facultad de Economía

A Dissertation in Candidacy for the Degree of Doctor of Philosophy in
Economics

Three Essays on Human Capital

By:

Mauricio Giovanni Valencia Amaya

Dissertation Directors:

PhD. Darwin Cortés – Universidad del Rosario

PhD. Darío Maldonado – Universidad de los Andes

Bogotá D.C., Colombia

October 2018

Acknowledgments

I would like to express my gratitude to my advisors, Prof. Darwin Cortés and Prof. Darío Maldonado, for their guidance, patience, and motivation. I also want to thank the Department of Economics of Universidad del Rosario for their support. I thank the participants of the XIX and XX LACEA Annual Meetings for their helpful comments on my presentations. A special thanks to Karen Chávez and Alexis Munari for their suggestions on one of the chapters of this thesis. Last but not the least, I would like to acknowledge the unconditional support and love of my parents: Luz Alba and José Rubiel.

Introduction

The theory of human capital, proposed in the 1950s and 1960s by Schultz, Mincer, and Becker, brought to the attention of development economics the importance of education and accumulation of knowledge. From this theoretical point of view, education and training can be considered as an investment that would render benefits in the future, both to the individual who invested in her human capital —micro level— and to the country where she resides —macro level—. In fact, the theory states that a great part of the differences in wages are due to differences in the productivity of individuals, but productivity is itself determined by previous investments made by them in education and training (Cahu and Zylberberg, 2004). Human capital is then a key element in the development of nations; it enhances the welfare and the choices available to people (Nahapiet, 2011), but at the same time it fosters economic growth (Barro, 2001). As Becker states:

The 21st century is clearly placing much greater emphasis than ever before on the importance of knowledge and information to the development of both countries and individuals (...) This means that it is more important than ever for both individuals and for nations to acquire knowledge, skills, and the experience to know how to acquire additional information (Becker, 2011, p. xv).

The concept of human capital comprises not only investments in education and on-the-job training, but also investments in health, as individuals pursue lowering their mortality rates and improving their quality of life (Becker, 2007). The relationship between health and human capital is straightforward: higher expectations on a longer and healthier lifespan would incentive greater investments in education and training, since returns on these investments would be higher (Becker, 2007). The Universal Declaration of Human Rights states, for instance, that both education and health are rights that governments must guarantee and protect. In many countries, citizens are required to attain a minimum level of education and are also encouraged to invest in the acquisition of new skills, knowledges, and capabilities. Many governments provide free healthcare, firms worldwide invest in health research and development, and individuals strive to decrease their mortality rates. That is, there are incentives for a society to invest in building and improving the human capital of its population through education, training, and better health practices, with the hope that this will render benefits, in terms of higher productivity.

However, the decision of an individual of how much to invest in human capital or whether to invest at all depends on many factors that influence their attitudes towards acquiring it as well as the expected returns of that investment. External shocks, such as natural disasters, epidemics, pandemics, or disease outbreaks; influencers, such as parents, peers, or education agents; or the socioeconomic, cultural, and political environment, these are all factors that would propel or deter the accumulation of human capital.

Take, for instance, the case of a sudden disease outbreak. This event will affect human capital accumulation through two possible channels: First, if individuals of school-age get sick, they might be absent from school during the convalescence period, not to mention that a disease might also deteriorate the cognitive capabilities of the patients. Second, even if individuals of school-age do not get sick, but parents, teachers, and/or classmates do, this might slow down or even impede their learning process. The social interactions of students are then assumed to be important for human capital accumulation and these might be affected not only by external shocks but also by their parents' decisions, such as, for example, naming them with uncommon names. Names convey information on the bearer's background and they might induce discriminatory behavior of teachers and employers or become a peers' motive for bullying behavior. Both discrimination and bullying might prevent the proper acquisition of human capital and demotivate further investments on it. Now, parents influence on their children's human capital go beyond their naming practices. Parents pass on their values, traditions, and culture and these, in turn, influence their children's attitudes, behaviors, and expectations. Consider, for example, the case of the

second-generation immigrants, who have acquired their human capital in the host country. The traditional human capital perspective states that the market should value equally the same skills, knowledge, and capabilities of two individuals born in the same country no matter where their parents were born. But, what happens to the human capital returns of the second-generation immigrants when their inherited values are not shared with the people from the host country and/or when these values mismatch the labor market prospects? This PhD dissertation deals precisely with some of above issues and their effects on human capital.

The first paper, “Health Shocks and Human Capital Accumulation: Impacts of the 2010 Dengue Epidemic in Colombia on Standardized Test Scores”, written in co-authorship with Darío Maldonado and Darwin Cortés, explores the causal effect of an unexpected dengue outbreak in Colombia on students’ test scores in a national standardized test. The relationship between health and education has been a growing topic of research (Eide and Showalter, 2011). Yet, most of these studies have focused on the effects of in utero and early childhood health shocks on different academic outcomes. Much less attention has been paid to the effects of shocks occurring during late childhood and adolescence. However, investments during adolescence can be as important for adult human capital outcomes as those made earlier in life (Lundberg and Wuermli, 2012). Furthermore, the intellectual capacity of teenagers can either increase or decrease (Ramsden et al., 2011) and their cognitive skills can still be strengthened by schooling instruction (Carlsson et al., 2015). On top of that, negative shocks and lack of investments during this life stage may even cancel the positive effects of previous investments and prevent a smooth transition to adulthood (Lundberg and Wuermli, 2012). This paper contributes then to this body of the literature. Its results show that sudden health shocks might negatively affect the accumulation of human capital as proxied by the performance in standardized test scores.

The second paper, “The Academic Burden of Carrying an Atypical Name: The Case of Antioquia, Colombia”, also in co-authorship with Darío Maldonado and Darwin Cortés, discusses the influence of atypical names on students’ academic outcomes. Usually, a name serves two purposes: to differentiate but also to categorize the name bearer (Mateos 2014). Carrying a distinctive name might help employers, peers, and teachers infer certain characteristics of its bearer, and so it might as well become a channel through which discrimination and/or bullying behaviors operate. Different research has studied the impacts on labor outcomes of bearing names linked to minority groups (Fryer and Levitt, 2004; Bertrand and Mullainathan 2004; Cotton et al. 2008; Watson et al. 2011). Although not conclusive, their results point to a possible negative effect of carrying distinctive names on wages and employment opportunities. If discriminatory practices against atypical name bearers happen also in the classroom, as found by Figlio (2005), the accumulation of human capital for this group might be at risk. The results from this paper suggest that bearing atypical names decrease students’ test scores, although the results are significant only for boys and for students from full-day institutions. The paper explores bullying behavior, gender differences in social support, and some psychological models, such as the big-fish-little-pond effect, as possible explanations.

Finally, the third paper, “The Cultural Distance of Parental Country of Birth and the Second-Generation Immigrants’ Wage Assimilation: The Case of Sweden”, proposes a cultural distance index to analyze the effect of inherited work-related values on a central question in the immigration literature: The economic assimilation of second-generation immigrants. One of the most important factors explaining the earnings gap between immigrants and natives is the geographical source of their human capital (Friedberg, 2000). The children of the first-generation immigrants, who were born in the receptor country, are supposed to acquire country-specific skills and country-specific knowledge that should improve their economic performance and close the gap in labor market outcomes, *vis-à-vis* natives. But, different empirical studies show mixed results, in terms of the assimilation degree of the second-generation immigrants (Abramitzky

et al., 2014; Riphahn, 2003; Blau et al., 2013; Blau, 2015; Portes et al., 2005). The results from this paper show that, compared with natives, having a culturally-different background negatively affects the second-generation immigrants' wages. The impact is greater the more culturally-distant their background is.

References

- ABRAMITZKY, Ran; BOUSTAN, Leah Platt; ERIKSSON, Katherine (2014). "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration", *Journal of Political Economy*, Vol. 122, No. 3, pp. 467-506.
- BLAU, Francine (2015). "Immigrants and gender roles: assimilation vs. culture", *IZA Journal of Migration*, Vol. 4, No. 1, pp. 1-21.
- BLAU, Francine; KAHN, Lawrence; LIU, Albert Yung-Hsu; and PAPPS, Kerry (2013). "The transmission of women's fertility, human capital, and work orientation across immigrant generations", *Journal of Population Economics*, Vol. 26, No. 2, pp. 405-435.
- CAHUC, Pierre and ZYLBERBERG, Andre (2004). *Labor Economics*, MIT Pres., Cambridge, Mass.
- CARLSSON, Magnus; DAHL, Gordon B.; ÖCKERT, Björn; and ROTH, Dan-Olof (2015). "The Effect of Schooling on Cognitive Skills", *The Review of Economics and Statistics*, Vol. XCVII, No. 3, pp. 533-547.
- COTTON, John; O'NEILL, Bonnie; and GRIFFIN, Andrea (2008). "The "Name Game": Affective and Hiring Reactions to First Names." *Journal of Managerial Psychology*, Volume 23, No. 1, pp. 18-39.
- FRIEDBERG, Rachel M. (2000). "You Can't Take It with You? Immigrant Assimilation and the Portability of Human Capital", *Journal of Labor Economics*, Vol. 18, No. 2, pp. 221-251.
- BECKER, Gary (2007). "Health as Human Capital: Synthesis and Extensions", *Oxford Economic Papers*, No. 59, pp. 379-410.
- BECKER, Gary (2011). "Foreword", in: Burton-Jones, Alan and Spender, J.-C. (eds.), *The Oxford Handbook of Human Capital*, Oxford University Press, Oxford.
- BERTRAND, Marianne and MULLAINATHAN, Sendhil (2004). "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *The American Economic Review*, Vol. 94, No. 4, pp. 991-1013.
- EIDE, Eric and SHOWALTER, Mark (2011). "Estimating the relation between health and education: What do we know and what do we need to know?", *Economics of Education Review*, No. 30, pp. 778-791.
- FRYER, Roland and LEVITT, Steven (2004). "The Causes and Consequences of Distinctively Black Names." *The Quarterly Journal of Economics*, Vol. CXIX, No. 3, pp. 767-805.
- LUNDBERG, Mattias and WUERMLI, Alice (2012). *Children and Youth in Crisis. Protecting and Promoting Human Development in Times of Economic Shocks*, The World Bank, Washington D.C.
- MATEOS, Pablo. (2014). "How We Got Our Names: Identity in Personal Names." In: Pablo Mateos, *Names, Ethnicity and Populations. Tracing Identity in Space*, ed. Springer, pp. 29-47.
- NAHAPIET, Janine (2011). "A Social Perspective: Exploring the Links between Human Capital and Social Capital", in: Burton-Jones, Alan and Spender, J.-C. (eds.), *The Oxford Handbook of Human Capital*, Oxford University Press, Oxford.

PORTES, Alejandro; FERNÁNDEZ-KELLY, Patricia; HALLER, William (2005). "Segmented assimilation on the ground: The new second generation in early adulthood", *Ethnic and Racial Studies*, Vol. 28, No. 6, pp. 1000-1040.

WATSON, Stevie; APPIAH, Osei; and THORNTON, Corliss (2011). "The Effect of Name on Pre-Interview Impressions and Occupational Stereotypes: The Case of Black Sales Job Applicants." *Journal of Applied Social Psychology*, Vol. 41, No. 10, pp. 2405–2420.

Paper 1

Health Shocks and Human Capital Accumulation: Impacts of the 2010 Dengue Epidemic in Colombia on Standardized Test Scores

Darwin Cortés, Darío Maldonado, and Mauricio G. Valencia-Amaya

Health Shocks and Human Capital Accumulation: Impacts of the 2010 Dengue Epidemic in Colombia on Standardized Test Scores

Darwin Cortés, Darío Maldonado, and Mauricio G. Valencia-Amaya¹

Abstract

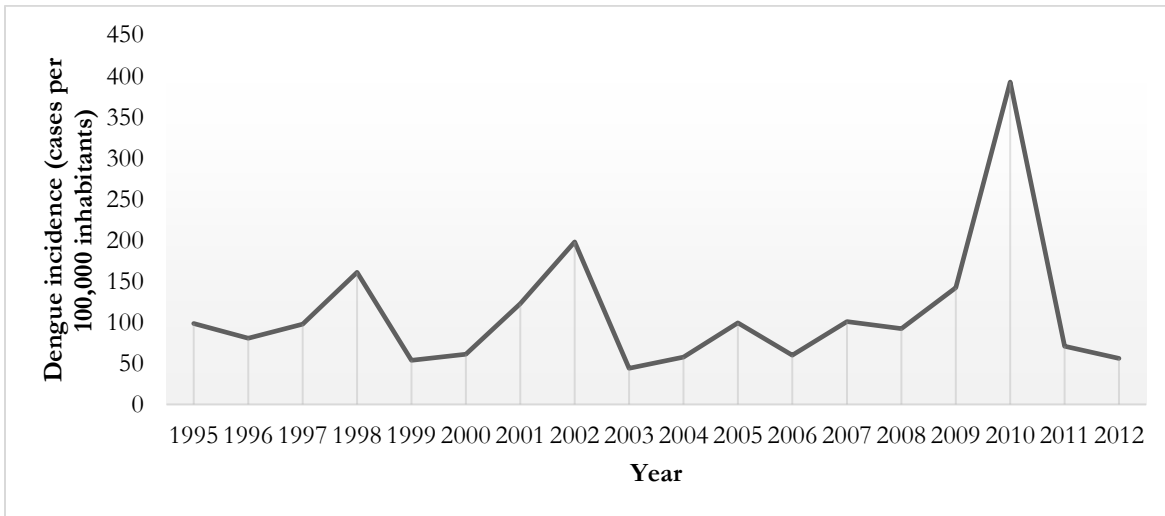
This paper analyses the effects of a sudden disease outbreak on human capital outcomes. It studies the causal impact of an unexpected dengue epidemic that occurred in Colombia in 2010 on the national Saber 11 test scores. The results from a difference-in-differences design indicate that the dengue outbreak reduced the average test scores of the 2011-2013 period. The short-lived negative impact decreased reading test scores more than math test scores. The findings show differential effects depending on the prevalent occupation of the infected people; but, in general, there was a higher impact on municipalities where most of the patients were students. Finally, the sudden dengue shock decreased the average test scores of students from middle- and lower-income households, but it did not affect the scores of the richer students.

I. Introduction

This paper investigates the impact of the sudden dengue outbreak of 2010 in Colombia on students' test scores in the national standardized test "Saber 11" during the period 2011-2013. In 2010 the worst dengue epidemic in a decade hit Colombia (see Graph 1). The number of people infected per 100,000 inhabitants increased 176 percent with respect to the previous year and 300 percent with respect to the average of the previous 10 years. The dengue outbreak was broad in geographical scope, reaching regions with historical low incidence rates, but also intensifying its magnitude in endemic regions. The monetary costs of this epidemic event have been addressed by the literature (Castro, 2013; Castañeda-Orjuela et al., 2012; Bello et al., 2011); however, there is a lack of research on the non-monetary cost of this disease, especially in terms of human capital, such as students' academic outcomes.

¹ Darwin Cortés: Associated professor, Department of Economics, Universidad del Rosario, Bogotá, Colombia. Darío Maldonado: Associated professor, School of Government Alberto Lleras Camargo, Universidad de los Andes, Bogotá, Colombia. Mauricio G. Valencia-Amaya: Ph.D. student, Department of Economics, Universidad del Rosario, Bogotá, Colombia.

Graph 1. Dengue Incidence in Colombia, 1995-2012 (Cases per 100,000 Inhabitants).



Source: Authors' calculations, based on Sivigila data.

Even though the incidence of vector-borne diseases in Colombia is high —dengue and malaria are endemic infectious diseases in the country—, to the best of our knowledge, research on the causal relations between these diseases and academic outcomes remains unexplored for the Colombian case. A research by Barron, Gamboa, and Rodríguez-Lesmes (2018) analyzed the impact of the 2010 dengue outbreak on school attendance and test scores in the national standardized test Saber 11 at the municipality level using OLS. In general, their results suggest that the outbreak did not have a significant impact on test scores; however, the outbreak seems to have caused a reduction on attendance in areas infected with severe dengue. Our approach differs from that of Barron, Gamboa, and Rodríguez-Lesmes (2018) in at least three aspects: (1) we use individual level data, which allows to account for the direct impacts of the disease; (2) we center our analysis on the effects of a real sudden shock, by focusing on municipalities that reported positive cases of dengue infections in 2010, but had zero cases in 2008 and 2009, and comparing them with municipalities that reported zero cases of dengue during the period 2008-2010, (3) we restrict the analysis to municipalities located at or below 1,200 meters above the sea level, in order to guarantee a positive probability of acquiring dengue.

Overall, the results indicate that the sudden dengue outbreak of 2010 decreased the 2011-2013 average Saber 11 tests scores. The effect of the shock was short-lived, and its impact was stronger on reading and on municipalities that reported most of its dengue-infected people as students. For these municipalities there was a large negative effect on math. For municipalities that reported most of the dengue-infected cases as non-students the effect was larger on reading. Finally, the general decrease in the test scores was stronger for students from middle- and lower-income households. The rest of this paper is organized as follows: section II presents the characterization of the dengue disease in 2010; section III reviews some of the literature on the relationship between health conditions and human capital outcomes; section IV introduces the methodology; section V show the main results; section VI discusses some of the main findings; and finally, the last section concludes.

II. Dengue Characterization

Dengue fever is a vector-borne disease caused by a virus, which is transmitted from human to human by mosquitos of the *Aedes* type. The bite of an infected mosquito causes high fever episodes lasting for up to two weeks. Other symptoms include: Headaches, anorexia, nausea, skin rash, joint and muscles pain, hemorrhage, and low WBC (White Blood Cell) count. Malaise, illness, and weakness might follow for quite a few weeks after the severe symptoms are gone (U.S. Department of Health and Human Services, 2009).

In Colombia, the “Instituto Nacional de Salud” (INS) classifies dengue into dengue and severe dengue (dengue hemorrhagic fever). According to the INS, dengue accounted for 94 percent of the 2010 cases, whereas severe dengue represented six percent of the cases. During 2010, the mortality incidence of this vector-borne disease was 0.2 percent. The dengue outbreak of 2010 affected mostly youngsters and young adults. In fact, the average age of the affected population was 25 years old. A closer look at the age distribution shows that 50 percent of the reported cases concentrated in people younger than 20 years old (25 percent correspond to people younger than 10 and 25 percent to people between 10 and 20 years old). The disease affected almost equally males and females: 51 and 49 percent, respectively. Moreover, the dengue outbreak of 2010 concentrated in urban areas (85 percent), a result that goes in line with the dengue factsheet of the World Health Organization, which states that dengue is presented mostly in urban and semi-urban areas (WHO, 2016, July). Climate variability has been linked to changes in the incidence of vector-borne diseases in Colombia (De la Mata and Valencia, 2014).

III. Literature Review

The health conditions of school children have been recognized by the literature as an important determinant of educational outcomes (Wolfe, 1985; Pollitt, 1990; Behrman, 1996; Sabia, 2007; Glewwe and Miguel, 2008; Belot and James, 2011). Even before a child is born, her future academic success can be affected by maternal health conditions during pregnancy (Daniele and Ostuni, 2013). Different studies highlight the negative effect of early life exposure to infectious diseases on human capital outcomes later in life, as well as the long-term positive effects of being exposed to disease-eradication campaigns during this life stage. For instance, early childhood exposure to malaria has been found to deteriorate late adulthood health conditions (Chang et al., 2014; Hong, 2013), to increase the rates of poverty (Barreca, 2010), and to lower educational attainment (Barreca, 2010; Burlando, 2015). Similarly, being exposed to anti-malaria campaigns early in life has been related to lower work disability levels (Hong, 2013), and to higher income levels (Bleakley, 2010; Mora, 2014), higher educational attainment (Lucas, 2010; Mora, 2014; Percoco, 2013), and higher literacy rates (Lucas, 2010). In some cases, the effects of improving early life conditions are greater on females than on males (Percoco, 2013; Bhalotra and Venkataramani, 2013), and in some others the effects are conditional on pre-campaign local characteristics, such as the existence of child labor markets or the importance given to school attendance (Mora, 2014). Still, some studies find modest impacts of malaria eradication campaigns on human capital variables or even no effects at all (Cutler et al., 2010).

Experimental and quasi-experimental designs during late childhood confirm some of the above-mentioned results. An experimental study carried out in Sri Lanka for nine months found that malaria prevention campaigns had a positive impact on school attainment (Fernando et al., 2006) and that being repeatedly exposed to this infectious disease increases the chances of lower performance at school (on both tailored tests and school tests) (Fernando et al., 2003). In the case of Sub-Saharan Africa countries, a quasi-experimental study on the effect of a major anti-malaria campaign on primary school children

educational outcomes found that exposure to the campaign increased both school attainment and school grades, and decreased schooling delay in more than 60 percent of the 22 countries where it was implemented (Kuecken, Thuilliez, and Valfort, 2013) (countries with the highest GNI per capita were the more successful). Regarding the effect of vector-borne diseases on schooling outcomes, different research conclude that these diseases have a negative impact on educational outcomes (Roland and Scorzafave, 2011; Suaya et al., 2009). For example, in a multi-country study on the direct and indirect cost of dengue in eight countries, students, who were diagnosed with dengue and had to be hospitalized, lost on average 5.6 days of school (Suaya et al., 2009).

IV. Methodology

A. Identification Strategy: Difference-in-Differences Estimation

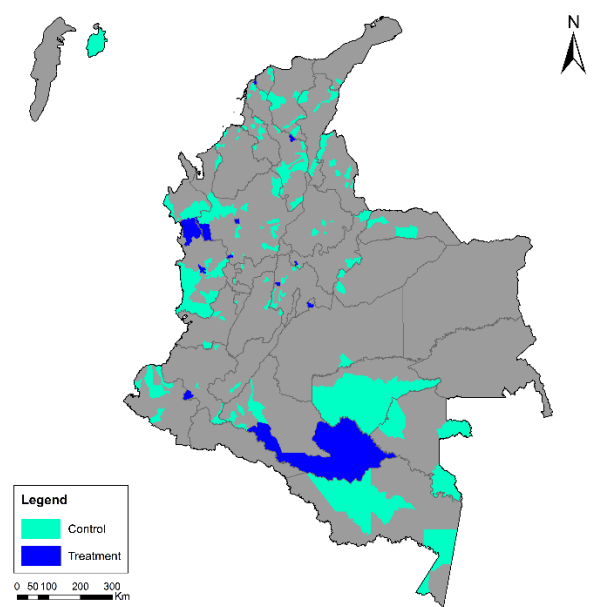
To identify the causal effect of the sudden dengue outbreak of 2010 on the students' Saber 11 tests scores, this paper implements as the main identification strategy a difference-in-differences approach with repeated cross sections, exploiting the unexpectedly spurt in dengue cases in 2010 (the time dimension) and its variation across municipalities (the regional dimension). The pre-treatment period goes from 2008 to 2010 and the post treatment period from 2011 to 2013. Equation (1) presents the baseline model specification:

$$Score_{ijt} = \alpha + \gamma Post_t + \delta Treat_j * Post_t + \Gamma X_i + \theta_t + \theta_j + \theta_s + \varepsilon_{it} \quad (1),$$

where $Score$ represents the Saber 11 test score of student i living in municipality j in the year t ; $Treat$ is a dummy variable that equals 1 if municipality j is considered treated; $Post$ is a dummy variable that takes the value 1 for the post-treatment period (2011-2013); θ_t , θ_j , and θ_s are year, municipality, and school fixed effects; \mathbf{X} are other student and student's households-varying characteristics.

A municipality is considered treated if it reported zero dengue cases per 100,000 inhabitants in the years 2008 and 2009 but reported positive cases in the year 2010. Municipalities that reported zero dengue cases per 100,000 inhabitants in the years 2008, 2009, and 2010 are used as controls. Now, with the purpose of guaranteeing a positive probability of being treated, we considered only municipalities located at or below 1,200 meters above the sea level (m.a.s.l.). The mosquito vector does not usually live at altitudes higher than 1,200 m.a.s.l.; in fact, close to 86 percent of all dengue cases in Colombia occurred in municipalities located at or below this altitude, according to Sivigila information. Under this definition, there are 16 municipalities in the treatment group and 146 municipalities in the control group (see Map 1).

Map 1. Treatment and Control Municipalities.



Source: Authors' calculations, based on Sivigila data.

Potential Confounding Variables

To improve the precision and validity of the estimations, this paper controls for municipality, school, households, and students' characteristics. Municipality and school characteristics are represented by dummy variables that capture the fixed influence of these places on students' test scores. Students' characteristics include age, age squared, a dummy for appropriate grade-related age (being 15 or 16 years old), and a binary indicator for being male. Household characteristics include dummies for: social strata, income levels, mother education levels, father education levels, ownership of household's appliances, internet access at home; and a dummy for overcrowding (equal to 1 if the number of people per dormitory was equal or greater than 2.5).

B. Data

The information on dengue incidence is provided by “Sistema Nacional de Vigilancia en Salud Pública” (National Public Health Surveillance System) Sivigila of the “Instituto Nacional de Salud” (National Health Institute). Since 1998 Sivigila oversees monitoring the dynamics of health-related events from a specific list of priority diseases for public health surveillance. At the end of 2006 Sivigila implemented specific protocols to collect information on priority diseases at the individual level. Each time a health provider receives a case, it is obliged and in charge of notifying the event to the health authorities following a well-defined guideline for compilation and notification, specific to each of the priority diseases. The information in the Sivigila database includes, among others, the disease code, disease name, notification date, notification week, medical consultation date, hospital name, patient's age, gender, occupation, type of health insurance, ethnic group, municipality and department of residence, municipality and department of origin, and area.

The ICFES database contains the test results from the examination Saber 11, a standardized national test applied to high school Colombian students prior to graduation. The test is developed by the Instituto Colombiano para el Fomento de la Educación Superior —ICFES (Colombian Institute for the

Promotion of Higher Education). The test assesses the academic skills of grade 11 high school students and it is administered twice a year, according to the academic year of the school; however, for most of the institutions the academic year starts in late January or early February and ends in mid or late November; this calendar is known as “calendar A”. Some universities require the test results as part of their application processes; the results also serve as a quality indicator that allows comparing the country’s high schools’ performance. The Saber 11 test has two components: a common core, which evaluates the students’ knowledge in eight (8) different subjects: language (Spanish), mathematics, biology, chemistry, physics, philosophy, social science, and foreign language (English); and a flexible core, which allows students to choose one subject out of the six available options, divided in four in-depth subjects: language, mathematics, biology, or social science, and two interdisciplinary subjects: environment or violence and society.

This paper uses the Saber 11 (calendar A) database for the period 2008-2013. Specifically, it uses the following information from the database: test results, student characteristics, and household characteristics. Test results include the test scores for language, mathematics, as well as the total score, which is calculated as the arithmetic mean of the common core components (language, mathematics, biology, chemistry, physics, philosophy, social science, and English). Student characteristics include: date of birth; mother education and father education, which ranks the student parents’ level of education on an ascending scale from 0 to 10 (where 0 means “none education” and 10 “graduate studies”); and sex (male or female).

Household characteristics include: social stratum (the social stratification by law, ranging from 1 to 6 — 1 indicating the lowest and 6 the highest, with each strata sharing similar socioeconomic characteristics; a few students classified as a strata 8 —not stratified, were omitted from the database); monthly household income, ranging from 1 to 7, according to the number of minimum wages earned by the household unit on a monthly basis (1: less than 1 minimum wage; 2: between 1 and 2; 3: between 2 and 3; 4: between 3 and 5; 5: between 5 and 7; 6: between 7 and 10; 7: more than 10); number of bedrooms and number of people living in the house; and, household’s belongings (car, computer, DVD player) and services installed at home (Internet connection, cable TV) (some variables in the original database specified the number of objects the household had, these were recoded to “having at least one”). Table 1 presents the summary statistics of this dataset.

Table 1. Summary Statistics.

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Male	207,857	0.460	0.498	0	1
Age	207,857	17.564	3.526	14	71
Age (15-16 years old)	207,857	0.412	0.492	0	1
Mother Education Levels	207,857				
2 (lowest)	207,857	0.031	0.174	0	1
3	207,857	0.170	0.376	0	1
4	207,857	0.182	0.386	0	1
5	207,857	0.176	0.380	0	1
6	207,857	0.255	0.436	0	1
7	207,857	0.016	0.125	0	1
8	207,857	0.052	0.221	0	1

9	207,857	0.016	0.127	0	1
10	207,857	0.084	0.278	0	1
11	207,857	0.014	0.116	0	1
12	207,857	0.004	0.063	0	1
13 (highest)	207,857	0.000	0.002	0	1
Father Education Levels	207,857				
2 (lowest)	207,857	0.048	0.213	0	1
3	207,857	0.187	0.390	0	1
4	207,857	0.179	0.383	0	1
5	207,857	0.152	0.359	0	1
6	207,857	0.241	0.427	0	1
7	207,857	0.014	0.115	0	1
8	207,857	0.046	0.210	0	1
9	207,857	0.017	0.130	0	1
10	207,857	0.093	0.290	0	1
11	207,857	0.015	0.121	0	1
12 (highest)	207,857	0.009	0.094	0	1
Socioeconomic Strata	207,857				
1 (lowest)	207,857	0.566	0.496	0	1
2	207,857	0.256	0.436	0	1
3	207,857	0.124	0.329	0	1
4	207,857	0.036	0.187	0	1
5	207,857	0.012	0.108	0	1
6 (highest)	207,857	0.007	0.083	0	1
Household Income Levels	207,857				
1 (lowest)	207,857	0.386	0.487	0	1
2	207,857	0.418	0.493	0	1
3	207,857	0.111	0.314	0	1
4	207,857	0.053	0.223	0	1
5	207,857	0.017	0.130	0	1
6	207,857	0.008	0.091	0	1
7 (highest)	207,857	0.007	0.081	0	1
Overcrowded Household	207,857	0.275	0.446	0	1
Household Vehicle Ownership	207,857	0.129	0.335	0	1
Household Computer Ownership	207,857	0.376	0.484	0	1
Household DVD Ownership	207,857	0.608	0.488	0	1
Household Cable TV Ownership	207,857	0.476	0.499	0	1
Household Internet Access	207,857	0.276	0.447	0	1

Source: Authors' calculations based on ICFES and Sivigila.

V. Results

A. Benchmark

Table 2 presents the estimation results of equations (1) using pooled OLS with clustered standard errors at the municipality level. The sudden dengue outbreak, as measured by the variable *Treat*, decreased average Saber 11 test scores during the post-treatment period (2011-2013) by 0.110 standard deviations. By subject, the unexpected outbreak reduced reading tests scores by 0.179 standard deviations, which is similar in magnitude to the effect of age (0.192 standard deviations). The disease outbreak did not affect math tests scores, at least when considering the whole post-treatment period.

Table 2. Diff-in-Diff: Effects of the 2010 Dengue Outbreak on Saber 11 Test Scores, 2011-2013.

Model		Dependent Variable: Saber 11 Z-Score			
		(i)	(ii)	(iii)	(iv)
Average	Treat*Post	-0.0844** (0.037)	-0.108** (0.040)	-109** (0.048)	-0.110** (0.045)
	Observations	207,857	207,857	207,857	207,857
	R-squared	0.005	0.077	0.355	0.391
Reading	Treat*Post	-0.164*** (0.043)	-0.181*** (0.042)	-0.178*** (0.049)	-0.179*** (0.048)
	Observations	207,962	207,962	207,962	207,962
	R-squared	0.003	0.055	0.184	0.207
Math	Treat*Post	0.000388 (0.038)	-0.0179 (0.042)	-0.0151 (0.048)	-0.0158 (0.047)
	Observations	207,962	207,962	207,962	207,962
	R-squared	0.005	0.05	0.201	0.233
Year fixed effects			X	X	X
Municipality fixed effects			X	X	X
School fixed effects				X	X
Controls					X

Notes: (a) *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; (b) standard errors in parenthesis are clustered at the municipality level; (c) *Treat* is a dummy variable equal to 1 if a student was living in a municipality for which the number of dengue cases per 100,000 inhabitants was zero in the years 2008 and 2009, but positive in the year 2010; (d) the coefficient *Treat*Post* is the estimated parameter δ of equation (1), which is the difference-in-differences estimation of the effect of the sudden dengue outbreak of 2010 on the average test score of the 2011-2013 period; (e) dengue cases include the total number of dengue, severe dengue, and dengue mortality; (f) control variables include age, age squared, and dummy variables for: male, being 15 or 16 years old, social strata, income levels, mother education levels, father education levels, ownership of household's appliances, internet access at home, and overcrowding; (g) all scale variables were converted into z-scores.

B. Robustness Tests

One of the most important identifying assumptions of the difference-in-differences strategy is the parallel-trend assumption, which guarantees that in the absence of treatment, the average change in the outcome of the treatment group would have been the same as the average change in the outcome of the control group. So, even though treatment and control groups differ in different characteristics (see Table 3), the common-trends assumption is satisfied in our case: both treatment and control groups show similar trends in the average Saber 11 test scores during the pre-treatment period (2008-2010) (see Graph 2). These results are further confirmed by Graph 3, which shows the regression coefficients of the treatment variable interacted with the time dummies. As expected, the coefficients for the pre-treatment period are insignificant and close to zero. Finally, we test on common pre-dynamics, following Mora and Reggio (2012, 2014). The null hypothesis of this test is that both treatment and control groups have

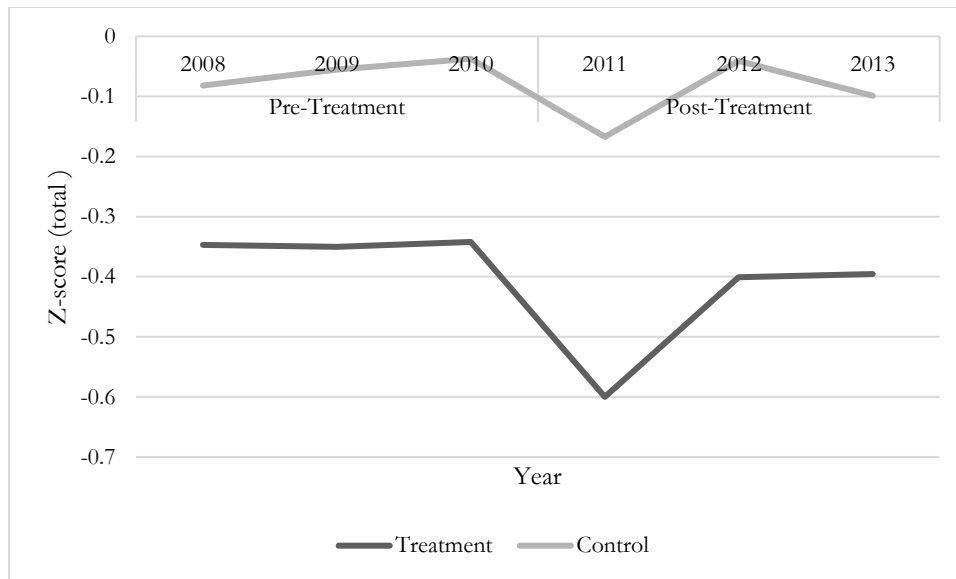
similar dynamics in the outcome variable during the pre-treatment period. Since the null is not rejected (p-value: 0.253) (see Table 4), we can be confident that the common-trends assumption is fulfilled.

Table 3. Summary Statistics: Treatment and Control Groups.

Variable	Control	Treatment	Difference
Altitude	216.20	381.27	-165***
Area	410.55	1450.49	-1039***
Temperature	27.38	26.19	1.186***
Age	17.57	18.45	-0.878***
Age (15-16)	0.39	0.29	0.101***
Mother's Education	4.41	3.50	0.915***
Father's Education	4.38	3.38	0.995***
Social Strata	1.73	1.20	0.533***
Income	1.95	1.53	0.426***
Sisben	1.87	1.27	0.605***
Male	0.46	0.49	-0.03***
Urban	0.81	0.64	0.173***
Work	0.09	0.11	-0.026***
Car	0.13	0.04	0.085***
Computer	0.35	0.16	0.192***
DVD	0.60	0.46	0.14***
Internet	0.26	0.07	0.185***

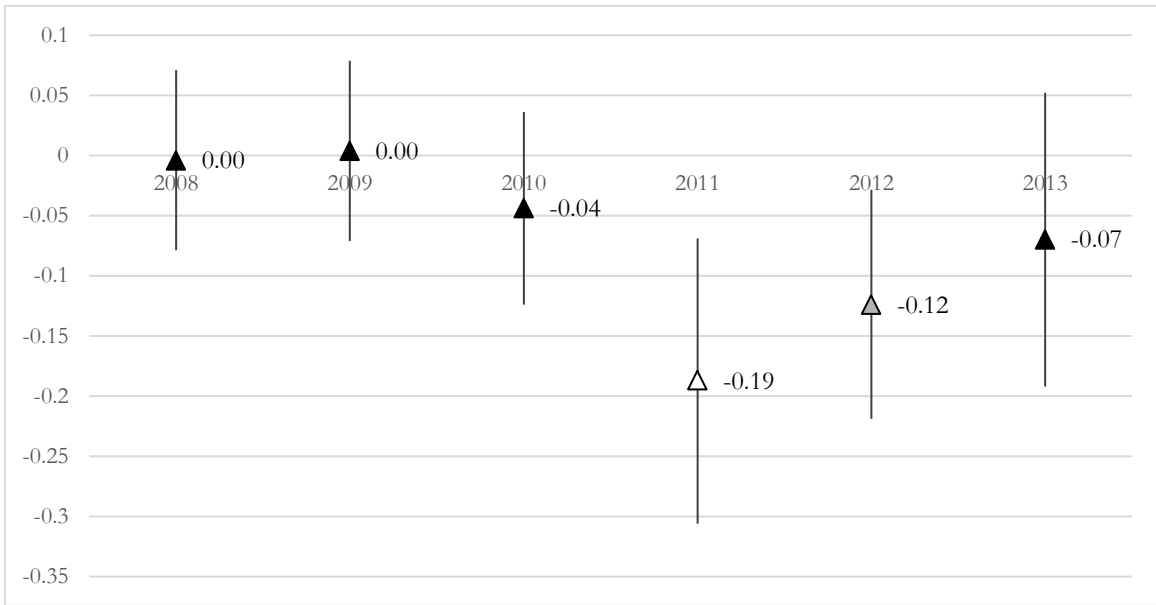
Source: Authors' calculations based on ICFES and Sivigila.

Graph 2. Graphic Inspection of the Common-Trends Assumption: Average Test Scores (Z-Score) for Treatment and Control Groups, 2008-2013.



Source: Authors' calculations, based on ICFES data.

Graph 3. Regression Coefficients of Treatment Variable Interacted with Time Dummies.



Source: Authors' calculations, based on ICFES data.

Note: The color of each triangle indicates the significance level of the estimated coefficient, as follows: White for $p\text{-value} < 0.01$, light grey for $p\text{-value} < 0.05$, dark grey for $p\text{-value} < 0.1$, and black for $p\text{-value} \geq 0.1$.

Table 4. Mora and Reggio Common Pre-Dynamics Test.

Unconditional Fully Flexible Model					
Output: Total Score			Number of obs.	207,857	
Sample Period: 2008-2013			H0: Common Pre-dynamics	1.988	
Treatment Period: 2011-2013			p-value	0.3701	
Pre-treatment (q) / Post-Treatment (s)	s=2011	s=2012	s=2013	H0: q=q-1	H0: s=s-1
q=2008	-0.144 (0.032)	-0.068 (0.028)	0.013 (0.028)		22.014 [0.000]
q=2009	-0.112 (0.052)	-0.006 (0.077)	0.107 (0.106)	-0.031 [0.305]	9.946 [0.007]
q=2010	-0.084 (0.097)	0.080 (0.232)	0.279 (0.424)	-0.029 [0.603]	1.373 [0.503]

Notes: (1) Common-trend assumption test based on Mora and Reggio (2012) and Mora and Reggio (2014); (2) Robust Standard Errors in parenthesis; (3) p-values in brackets.

As for the possible confounders, the results presented in Table 2 are robust and consistent to three different specifications. The first controls for year and municipality fixed characteristics (column ii); the second adds school fixed effects (column iii); and the third additionally accounts for students' and households' time-varying characteristics (column iv). Since the results from the three specifications are almost identical in magnitude and run in the same direction, we can be certain that our results reflect the impact of the unexpected dengue outbreak on test scores.

It is important to note that the size of the *Treat*Post* coefficient increases when municipality fixed effects are added to the simple difference-in-differences estimation. This would imply that the estimation of the coefficient, without such fixed effects, might be biased; that is, the fixed characteristics of municipalities

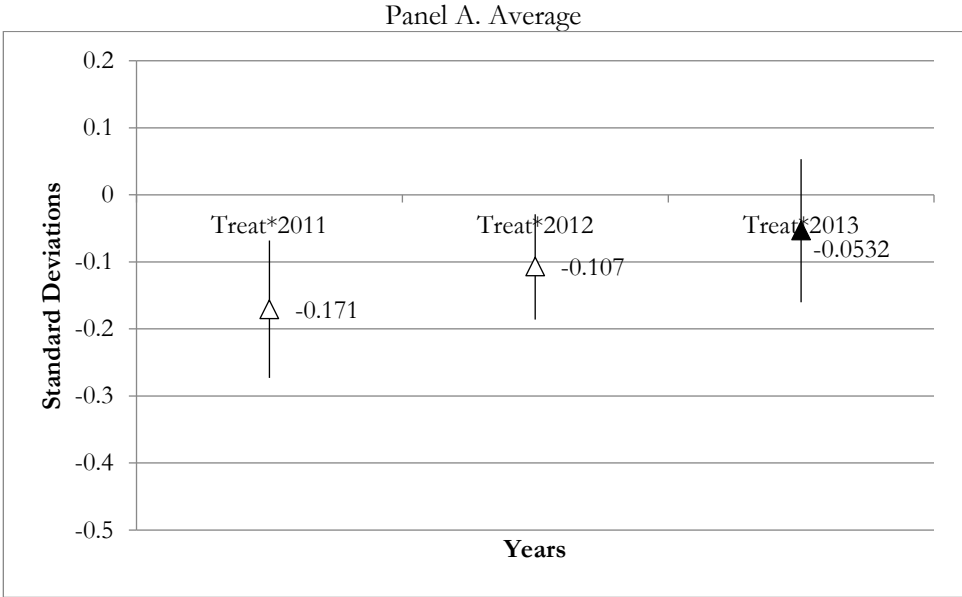
help explain both the test scores and the fact that a municipality experienced positive dengue cases in 2010, after having had zero cases in 2008 and 2009. However, the estimator would be downward biased. For instance, treated municipalities might share certain characteristics that make them more prone to host a dengue outbreak and at the same time these traits might be related to lower test scores. It can also be the case that these common features both prevents dengue epidemics to happen and increases test scores. Therefore, the effect of the omitted variables on the treatment indicator and on the outcome seems to follow opposite directions, and so not including these fixed effects would underestimate the impact of the dengue outbreak.

C. Heterogeneous Effects

1. Duration

To account for the duration of the sudden dengue impact on the test scores, Graph 4 reports the time-heterogeneous effects for the years 2011, 2012, and 2013. This specification controls for year, municipality, and school fixed effects, as well as for time-varying students’ and households’ characteristics. In general, the effect was short-lived. The abrupt epidemic outburst lowered average test scores in 2011, but the magnitude of the effect attenuated in 2012, and finally disappeared in 2013. That is, the negative effect of the disease was larger for students who were in the 10th grade at the time of the shock than for those who were in the 9th grade and had apparently no long-term effect on those who were in the 8th grade (panel A). By subject, the dengue shock lowered both reading and math tests scores in 2011, but the negative impact was almost three times larger for reading than for math (panels B and C). The impact on reading persisted throughout the post-treatment period, though its magnitude declined gradually from year to year: 60 percent from 2011 to 2012 and almost 70 percent from 2011 to 2013. The size of the effect on math disappeared drastically in the year 2012.

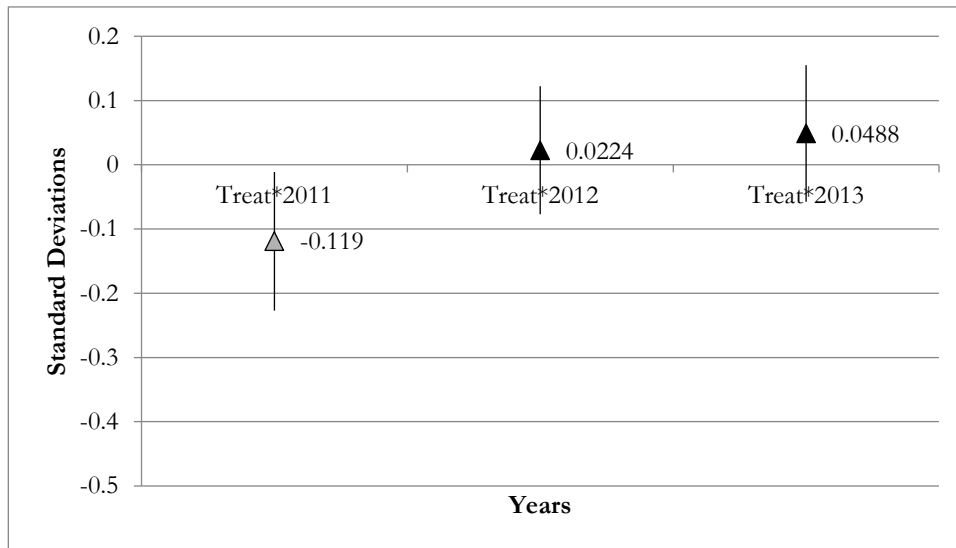
Graph 4. Diff-in-Diff: 95% Confidence Interval for the Effect of the 2010 Dengue Outbreak on Saber 11 Test Scores, 2011-2013.



Panel B. Reading



Panel C. Math



Notes: (a) *Treat* is a dummy variable equal to 1 if a student was living in a municipality for which the number of dengue cases per 100,000 inhabitants was zero in the years 2008 and 2009, but positive in the year 2010; (b) the coefficients *Treat*2011*, *Treat*2012*, and *Treat*2013* are the estimated parameters δ_{2011} , δ_{2012} , and δ_{2013} of equation (1), which are the difference-in-differences estimations of the effect of the sudden dengue outbreak of 2010 on the test scores of the years 2011, 2012, and 2013; (c) the color of each triangle indicates the significance level of the estimated coefficient, as follows: White for $p\text{-value} < 0.01$, light grey for $p\text{-value} < 0.05$, dark grey for $p\text{-value} < 0.1$, and black for $p\text{-value} \geq 0.1$; (d) dengue cases include the total number of dengue, severe dengue, and dengue mortality; (e) control variables include age, age squared, and dummy variables for: male, being 15 or 16 years old, social strata, income levels, mother education levels, father education levels, ownership of household's appliances, internet access at home, and overcrowding; (f) all scale variables were converted into z-scores.

2. Short-Run Analysis

This section presents the heterogeneous effects of the 2010 dengue outbreak on the Saber 11 test scores by occupation and income levels. The analysis focuses on the impact in the year 2011, as the effect concentrated in this year, as the previous section showed.

Table 5 presents the estimation of equation (1) for the year 2011 that accounts for the occupational composition of the people infected with dengue, based on the International Standard Classification of Occupations (ISCO-88). Treated municipalities were divided into three mutually exclusive groups: (a) municipalities that classified most of the dengue-infected population as students, (b) municipalities that classified most of the dengue-infected population as non-students and reported professionals among them (physical, mathematical and engineering science professionals; life science and health professionals; teaching professionals; and other professionals), and (c) municipalities that classified most of the dengue-infected population as non-students and reported no professionals among the infected. The results show that the unexpected dengue outbreak of 2010 reduced students' average test scores of 2011 for all groups; however, the size of the effect was more than twice as large for test-takers who lived in treated municipalities where most of the dengue-infected people were students. By subject, the sudden outburst reduced reading test scores of all groups, but the effect was larger for the “mostly non-students with professionals” group. In the case of math scores, the dengue epidemics of 2010 decreased the “mostly students” group scores by 0.493 standard deviations and the “mostly non-students without professionals” group scores by 0.117 (although barely significant); the sudden health shock had no effect on the “mostly non-students with professionals” group scores.

Table 5. Diff-in-Diff: Effect of the 2010 Dengue Outbreak on the 2011 Saber 11 Test Scores, by Occupation.

Dependent Variable: Saber 11 Z-Score				
Occupation Prevalence among the Infected Population	Subject	Average	Reading	Math
Mostly Students	Treat*2011	-0.393*** (0.035)	-0.233*** (0.068)	-0.493*** (0.035)
Mostly Non-students (Includes Professionals)	Treat*2011	-0.176*** (0.065)	-0.372*** (0.090)	-0.0787 (0.065)
Mostly Non-students (Different from Professionals)	Treat*2011	-0.149** (0.067)	-0.282*** (0.101)	-0.117* (0.064)
Observations		207,857	207,962	207,962
R-squared		0.391	0.207	0.233
Year fixed effects		X	X	X
Municipality fixed effects		X	X	X
School fixed effects		X	X	X
Controls		X	X	X

Notes: (a) *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; (b) standard errors in parenthesis are clustered at the municipality level; (c) *Treat* is a dummy variable equal to 1 if a student was living in a municipality for which the number of dengue cases per 100,000 inhabitants was zero in the years 2008 and 2009, but positive in the year 2010; (d) the coefficient *Treat*2011* is the estimated parameter δ_{2011} of equation (1), which is the difference-in-differences estimation of the effect of the sudden dengue outbreak of 2010 on the average test score of the year 2011; (e) dengue cases include the total number of dengue, severe dengue, and dengue mortality; (f) control variables include age, age squared, and dummy variables for: male, being 15 or 16 years old, social strata, income levels, mother education levels, father education levels, ownership of household's appliances, internet access at home, and overcrowding; (g) all scale variables were converted into z-scores; (h) occupations were classified according to the International Standard Classification of Occupations (ISCO-88).

Table 6 presents the results by income level. The 2010 dengue outbreak hit only students from lower- and middle-income households, but the impact was stronger on the latter group. Students were grouped into three different income levels, depending on the number of minimum wages per month earned by the student's household: lower range (less than three), middle range (between three and seven), and higher range (more than seven). The overall negative impact on the average score for the lower- and middle-income range was 0.149 and 0.264, respectively, but there was no significant impact on students from

higher-income households. This conclusion applies also to the analysis by subject: the sudden health shock decreased reading and math scores for students from poorer and middle-income families, had no significant impact on the scores of students from richer families, but, comparatively, it struck harder on the middle-class.

Table 6. Diff-in-Diff: Effect of the 2010 Dengue Outbreak on the 2011 Saber 11 Test Scores, by Students' Household Income Level.

		Dependent Variable: Saber 11 Z-Score		
Household Income level	Subject	Average	Reading	Math
Lower Range	Treat*2011	-0.149*** (0.0536)	-0.306*** (0.0885)	-0.115** (0.0549)
Middle Range	Treat*2011	-0.264*** (0.0591)	-0.389*** (0.0989)	-0.234*** (0.0621)
Higher Range	Treat*2011	0.0013 (0.152)	-0.131 (0.14)	0.232 (0.181)
Observations		207,857	207,962	207,962
R-squared		0.390	0.207	0.233
Year fixed effects		X	X	X
Municipality fixed effects		X	X	X
School fixed effects		X	X	X
Controls		X	X	X

Notes: (a) *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; (b) standard errors in parenthesis are clustered at the municipality level; (c) *Treat* is a dummy variable equal to 1 if a student was living in a municipality for which the number of dengue cases per 100,000 inhabitants was zero in the years 2008 and 2009, but positive in the year 2010; (d) the coefficient *Treat*2011* is the estimated parameter δ_{2011} of equation (1), which is the difference-in-difference estimation of the effect of the sudden dengue outbreak of 2010 on the test scores of the year 2011; (e) dengue cases include the total number of dengue, severe dengue, and dengue mortality; (f) control variables include age, age squared, and dummy variables for: male, urban area, being 15 or 16 years old, social strata, Sisben levels, mother education levels, father education levels, employment status of the student, ownership of household's appliances, internet access at home, and overcrowding; (g) all scale variables were converted into z-scores; (h) household income level 1: less than three minimum wages per month, household income level 2: between three and seven minimum wages per month, household income level 3: more than seven minimum wages per month. Number of municipalities in treatment and control groups by income levels: income level 1: treatment: 16, control: 146; income level 2: treatment: 16, control: 146; income level 3: treatment: 16, control: 145.

VI. Discussion

Even though the causal relationship between health and education has been a growing topic of research (Eide and Showalter, 2011), most of these studies have focused on the effects of in utero and early childhood health shocks, but much less attention has been paid to the effects on educational outcomes of health shocks during late childhood and adolescence. The results from this research contribute to this latter body of the literature. Investments during adolescence can be as important for adult human capital outcomes as those made earlier in life (Lundberg and Wuermli, 2012). Moreover, the intellectual capacity of teenagers, measured with verbal and non-verbal intelligent quotients (IQ), is malleable; that is, it can either increase or decrease (Ramsden et al., 2011) and their cognitive skills can still be strengthened by schooling instruction (Carlsson et al., 2015). On top of that, negative shocks and lack of investments during this lifetime period may even cancel the positive effects of investments made during previous life stages and prevent a smooth transition to adulthood (Lundberg and Wuermli, 2012).

The results from this paper suggest that the sudden dengue outbreak of 2010 decreased students' Saber 11 tests scores of the 2011-2013 period. The impact was long-lasting for reading, though it declined

gradually; whereas, for math it was a “one-shot” effect, as it did not have a duration. This section discusses these findings in the light of the relevant literature.

The relative importance of the social component in the reading and math learning processes might help explain the difference in the size and duration of the impact by subject. Social interactions between students and their parents, peers, and teachers are important in the learning process of reading (Bloome, 1985), so interruptions in these interactions may have a negative impact on the development of their language skills (Steelman and Doby, 1983). For instance, a decrease in parents’ health might affect their ability not only to provide physical and monetary resources, but also to devote time and effort to school-related activities; parental illness might have then a negative impact on children’s academic performance (Andrews and Logan, 2010; Dhanaraj, 2016a, 2016b). Students’ peers exert also an important influence on academic outcomes (Sacerdote, 2001; Hanushek et al., 2003), especially on the ones related to verbal skills (Zimmerman, 2003). Finally, given the importance of teachers for students’ academic achievement (Rikvin, Hanushek, and Kain, 2005), their absence in the classroom is linked to a decrease in students’ test scores (Miller, Murnane, and Willett, 2008a and 2008b); that is, the deterioration of teachers’ health or illness of members of their families can lead to lower students’ achievement (Das et al., 2007).

As for students themselves, one of the direct consequences of getting sick is missing school days, and one of the short-term impacts of students’ absenteeism is lower performance in test scores (Kamhöfer et al., 2015). Different studies conclude that school absence reduces students’ math test scores (Aucejo and Romano, 2014; Goodman, 2014; Paredes and Ugarte, 2011), whereas spending more time in school usually increases reading scores (Fitzpatrick, Grissmer, and Hastedt, 2011; Carlsson et al., 2015). The results from this paper indicate also that the effect of the dengue outbreak of 2010 was larger for students from middle-income households than for students from lower-income households. This finding is supported by Jackson (2009), who states that having a high social status does not seem to prevent the negative consequences of health shocks. In her study, for example, the negative effect of a health shock on educational attainment was larger for white teenagers than for their black or Hispanic counterparts.

Conclusions

This paper studied the impact of the sudden dengue outbreak of 2010 in Colombia on students’ test scores in the national standardized test “Saber 11” during the period 2011-2013. The unexpected dengue outbreak of 2010 decreased the 2011-2013 average Saber 11 tests scores of students who lived in municipalities that reported positive cases of dengue in 2010, after having reported zero cases of dengue per 100,000 inhabitants in 2008 and 2009, in comparison with students who lived in municipalities that reported zero cases of dengue from 2008 to 2010. The effect of the dengue outbreak was short-lived; it concentrated in 2011, decreased in 2012, and disappeared by 2013. The impact was larger for reading, for which it persisted until 2013, than for math, for which it lasted only until 2011. The magnitude of the effect was stronger for municipalities that reported most of the dengue-infected people as students. For these municipalities, there was also a large and negative impact on math test scores. But, when most of the dengue-infected people were non-students, the dengue shock had a bigger impact on reading. By income level, students from middle- and lower-income households suffered the most, in terms of the magnitude of the effect on their test scores.

References

ANDREWS, Rodney J. and LOGAN, Trevon D. (2010). “Family Health, Children’s Own Health, and Test Score Gaps”, *The American Economic Review*, Vol. 100, No. 2, pp. 195-199.

- AUCEJO, Esteban and ROMANO, Teresa (2014). “Assessing the Effect of School Days and Absences on Test Score Performance”, *CEP Discussion Papers*, No. 1302.
- BARRECA, Alan I. (2010). “The Long-Term Economic Impact of in Utero and Postnatal Exposure to Malaria”, *The Journal of Human Resources*, Vol. 45, No. 4, pp. 865-892.
- BARRON, Kai; GAMBOA, Luis F.; and RODRIGUEZ-LESMESS, Paul (2018). “Behavioural Response to a Sudden Health Risk: Dengue and Educational Outcomes in Colombia”, *The Journal of Development Studies*, DOI: 10.1080/00220388.2018.1425392.
- BHALOTRA, Sonia R. and Venkataramani, Atheendar (2013). “Cognitive Development and Infectious Disease: Gender Differences in Investments and Outcomes”, *IZA Discussion Paper*, No. 7833.
- BEHRMAN, Jere R. (1996). “The Impact of Health and Nutrition on Education”, *The World Bank Research Observer*, Vol. 11, No. 1, pp. 23-37.
- BELLO, Sandra; DÍAZ, Ernesto; MALAGÓN-ROJAS, Jeandran; ROMERO, Martín; and SALAZAR, Valentina (2011). “Medición del impacto económico del dengue en Colombia: una aproximación a los costos médicos directos en el periodo 2000-2010”, *Biomédica*, Vol. 31, Suplemento 3, I Tomo, Memorias XX Congreso Latinoamericano de Parasitología, pp. 110-113.
- BELOT, Michèle and JAMES, Jonathan (2011). “Healthy School Meals and Educational Outcomes”, *Journal of Health Economics*, Vol. 30, No. 3, pp. 489–504.
- BLEAKLEY, Hoyt (2010). “Malaria Eradication in the Americas: A Retrospective Analysis of Childhood Exposure”, *American Economic Journal: Applied Economics*, Vol. 2, No. 2, pp. 1-45.
- BLOOME, David (1985). “Reading as a Social Process”, *Language Arts*, Vol. 62, No. 2, pp. 134-142.
- BURLANDO, Alfredo (2015). “The Disease Environment, Schooling, and Development Outcomes: Evidence from Ethiopia”, *The Journal of Development Studies*, Vol. 51, No. 12, pp. 1563–1584.
- CARLSSON, Magnus; DAHL, Gordon B.; ÖCKERT, Björn; and ROTH, Dan-Olof (2015). “The Effect of Schooling on Cognitive Skills”, *The Review of Economics and Statistics*, Vol. XCVII, No. 3, pp. 533-547.
- CASTAÑEDA-ORJUELA, Carlos; DÍAZ, Hernando; ALVIS-GUZMAN, Nelson; OLARTE, Andrés; RODRÍGUEZ, Heidy; CAMARGO, Guido; and DE LA HOZ-RESTREPO, Fernando (2012). “Burden of Disease and Economic Impact of Dengue and Severe Dengue in Colombia, 2011”, *Value in Health Regional Issues*, Vol. 1, No. 2, pp. 123-128.
- CASTRO, Raúl (2013). Costos del dengue en Colombia (2010-2012): Impacto en salud pública. Mimeo.
- CHANG, Simon; FLEISHER, Belton; KIM, Seonghoon; and LIU, Shi-Yung (2014). “Long-term Health Effects of Malaria Exposure around Birth: Evidence from Colonial Taiwan”, *Economic Development and Cultural Change*, Vol. 62, No. 3, pp. 519-536.
- CUTLER, David; FUNG, Winnie; KREMER, Michael; SINGHAL, Monica; and VOGL, Tom (2010). “Early-life Malaria Exposure and Adult Outcomes: Evidence from Malaria Eradication in India”, *American Economic Journal: Applied Economics*, Vol. 2, No. 2, pp. 72-94.
- DANIELE, Vittorio and OSTUNI, Nicola (2013). “The burden of disease and the IQ of nations”, *Learning and Individual Differences*, No. 28, pp. 109–118.

- DAS, Jishnu; DERCON, Stefan; HABYARIMANA, James, and KRISHNAN, Pramila (2007). "Teacher Shocks and Student Learning", *Journal of Human Resources*, Vol. 42, No. 4, pp. 820-862.
- DE LA MATA, Dolores and VALENCIA-AMAYA, Mauricio-G. (2014). "The Health Impacts of Severe Climate Shocks in Colombia", *IDB Working Paper Series*, No. IDB-WP-498.
- DHANARAJ, Sowmya (2016a). "Economic vulnerability to health shocks and coping strategies: evidence from Andhra Pradesh, India", *Health Policy and Planning*, Vol. 0, No. 0, pp. 1–10.
- DHANARAJ, Sowmya (2016b). "Effects of parental health shocks on children's schooling: Evidence from Andhra Pradesh, India", *International Journal of Educational Development*, No. 49, pp. 115–125.
- EIDE, Eric and SHOWALTER, Mark (2011). "Estimating the relation between health and education: What do we know and what do we need to know?", *Economics of Education Review*, No. 30, pp. 778– 791.
- FERNANDO, S. D.; GUNAWARDENA, D. M.; BANDARA, M. R. S. S.; DE SILVA, D.; CARTER, R.; MENDIS, K. N.; and WICKREMASINGHE, A. R. (2003). "The Impact of Repeated Malaria Attacks on the School Performance of Children", *American Journal of Tropical Medicine and Hygiene*, No. 6, Vol. 69, pp. 582-588.
- FERNANDO, Deepika; DE SILVA, Damani; CARTER, Richard; MENDIS, Kamini N.; and WICKREMASINGHE, Rajitha (2006). "A Randomized, Double-Blind, Placebo-Controlled, Clinical Trial of the Impact of Malaria Prevention on the Educational Attainment of School Children", *American Journal of Tropical Medicine and Hygiene*, No. 3, Vol. 74, pp. 386-393.
- FITZPATRICK, Maria; GRISSMER, David; HASTEDT, Sarah (2011). "What a difference a day makes: Estimating daily learning gains during kindergarten and first grade using a natural experiment", *Economics of Education Review*, No. 30, pp. 269–279.
- GLEWWE, Paul and MIGUEL, Edward (2007). "The Impact of Child Health and Nutrition on Education in Less Developed Countries", in: SCHULTZ, T. Paul and STRAUSS, John A. (eds.) *Handbook of Development Economics*, Elsevier, Ed. 1, Vol. 4, No. 5.
- GOODMAN, Joshua (2014). "Flaking Out: Student Absences and Snow Days as Disruptions of Instructional Time", *NBER Working Paper Series*, No. 20221.
- HANUSHEK, Eric; KAIN, John; MARKMAN, Jacob; and RIVKIN, Steven (2003). "Does Peer Ability Affect Student Achievement?", *Journal of Applied Econometrics*, Vol. 18, No. 5, pp. 527-544.
- HONG, Sok Chul (2013). "Malaria: An Early Indicator of Later Disease and Work Level", *Journal of Health Economics*, Vol. 32, No. 3, pp. 612-632.
- JACKSON, Margot (2009). "Understanding links between adolescent health and educational attainment", *Demography*, Vol. 46, No. 4, pp. 671-694.
- KAMHÖFER, Daniel A.; CATTAN, Sarah; KARLSSON, Martin; NILSSON, Therese (2015). "The Effects of Sickness Absence in School on Educational Achievements, Mortality and Income", *Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015: Ökonomische Entwicklung - Theorie und Politik - Session: Modelling Causal Effects*, No. F08-V3.

- KUECKEN, Maria; THUILLIEZ, Josselin; and VALFORT, Marie-Anne (2013). “Does malaria control impact education? A study of the Global Fund in Africa”, *Documents de travail du Centre d’Economie de la Sorbonne*, No. 2013.75.
- LUCAS, Adrienne (2010). “Malaria Eradication and Educational Attainment: Evidence from Paraguay and Sri Lanka”, *American Economic Journal: Applied Economics*, Vol. 2, No. 2, pp. 46-71.
- LUNDBERG, Mattias and WUERMLI, Alice (2012). *Children and Youth in Crisis. Protecting and Promoting Human Development in Times of Economic Shocks*, The World Bank, Washington D.C.
- MILLER, Raegen; MURNANE, Richard; and WILLETT, John (2008a). Do Teacher Absences Impact Student Achievement? Longitudinal Evidence from One Urban School District, *Educational Evaluation and Policy Analysis*, Vol. 30, No. 2, pp. 181-200.
- MILLER, Raegen; MURNANE, Richard; and WILLETT, John (2008b). “Do worker absences affect productivity? The case of teachers”, *International Labour Review*, Vol. 147, No. 1, pp. 71-89.
- MORA, Claudio Alberto (2014). “Can Benefits from Malaria Eradication be Increased? Evidence from the Costa Rican Malaria Campaigns”, Job Market Paper, Pontificia Universidad Católica de Chile.
- MORA, Ricardo and REGGIO, Iliana (2014). “dqd: A command for treatment effect estimation under alternative assumptions”, *Economics Working Papers*, No. we1407, Universidad Carlos III, Departamento de Economía.
- MORA, Ricardo and REGGIO, Iliana (2012). “Treatment effect identification using alternative parallel assumptions”, *Economics Working Papers*, No. we1233, Universidad Carlos III, Departamento de Economía.
- PAREDES, Ricardo D. and UGARTE, Gabriel A. (2011). “Should Students Be Allowed to Miss?”, *The Journal of Educational Research*, Vol. 104, No. 3, pp. 194-201.
- PERCOCO, Marco (2013). “The Fight Against Disease: Malaria and Economic Development in Italian Regions”, *Economic Geography*, Vol. 89, No. 2, pp. 105-125.
- POLLITT, Ernesto (1990). *Malnutrition and infection in the classroom*, Unesco, 198p.
- RAMSDEN, Sue; RICHARDSON, Fiona M.; JOSSE, Goulven; THOMAS, Michael S. C.; ELLIS, Caroline; SHAKESHAFT, Clare; SEGHER, Mohamed L.; and PRICE, Cathy J. (2011). “Verbal and non-verbal intelligence changes in the teenage brain”, *Nature*, Vol. 479, No. 7371, pp. 113-116.
- RIVKIN, Steven; HANUSHEK, Eric; and KAIN, John (2005). “Teachers, Schools, and Academic Achievement”, *Econometrica*, Vol. 73, No. 2, pp. 417-458.
- ROLAND, D. A.; SCORZAFAVE, L. G. (2011). “Dengue Outbreaks and the Impact on School Proficiency”, in: *VI International Conference - Making Policy a Health Equity Building Process*, 2011, Cartagena de Indias.
- SACERDOTE, Bruce (2001). “Peer Effects with Random Assignment: Results for Dartmouth Roommates”, *The Quarterly Journal of Economics*, Vol. 116, No. 2, pp. 681-704.
- SABIA, Joseph (2007). “The Effect of Body Weight on Adolescent Academic Performance”, *Southern Economic Journal*, Vol. 73, No. 4, pp. 871-900.

STEELMAN, Lala Carr and DOBY, John T. (1983). "Family Size and Birth order as Factors on the IQ Performance of Black and white Children", *Sociology of Education*, Vol. 56, No. 2, pp. 101-109.

SUAYA, Jose A.; SHEPARD, Donald S.; SIQUEIRA, João B.; MARTELLI, Celina T.; LUM, Lucy C. S.; HUAT TAN, Lian; KONGSIN, Sukhontha; JIAMTON, Sukhum; GARRIDO, Fátima; MONTOYA, Romeo; ARMIEN, Blas; HUY, Rekol; CASTILLO, Leticia; CARAM, Mariana; SAH, Binod K.; SUGHAYYAR, Rana; TYO, Karen R.; and HALSTEAD, Scott B. (2009). "Cost of Dengue Cases in Eight Countries in the Americas and Asia: A Prospective Study", *American Journal of Tropical Medicine and Hygiene*, No. 5, Vol. 80, pp. 846-855.

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES (2009). *Dengue and Dengue Hemorrhagic Fever: Information for Health Care Practitioners*, Centers for Disease Control and Prevention, retrieved from: https://www.cdc.gov/dengue/resources/denguedhf-information-for-health-care-practitioners_2009.pdf.

World Health Organization (WHO) (2016, July). *Dengue and Severe Dengue Fact Sheet*, retrieved from: <http://www.who.int/mediacentre/factsheets/fs117/en>.

WOLFE, Barbara L. (1985). "The Influence of Health on School Outcomes: A Multivariate Approach", *Medical Care*, Vol. 23, No. 10, pp. 1127-1138.

ZIMMERMAN, David J. (2003). "Peer Effects in Academic Outcomes: Evidence from a Natural Experiment", *Review of Economics and Statistics*, Vol. 85, No. 1, pp. 9-23.

Paper 2

The Academic Burden of Carrying an Atypical Name: The Case of Antioquia, Colombia

Darwin Cortés, Darío Maldonado, and Mauricio G. Valencia-Amaya

The Academic Burden of Carrying an Atypical Name: The Case of Antioquia, Colombia

Darwin Cortés, Darío Maldonado, and Mauricio G. Valencia-Amaya²

Abstract

This paper studies the effect of carrying an atypical name on students' test scores, using a unique dataset for Antioquia, Colombia — “Olimpiadas del Conocimiento” test—. Following an education production function approach, the results show that atypical name bearers scored 0.1 standard deviations less in both math and reading. The impact was significant only for boys and for students from full-day institutions. Bullying behavior is explored as a yet-to-test channel.

I. Introduction

This paper analyzes the relationship between atypical names and the performance of students in standardized tests. Fryer and Levitt (2004) and Bertrand and Mullainathan (2004) brought to the international academic spotlight the subject of the influence of names on socioeconomic outcomes, specifically on labor outcomes, although finding opposite conclusions. In general, there is still an ongoing debate on whether specific names do really influence labor market opportunities or not. Gaviria et al. (2007) studied the Colombian case, finding that carrying atypical names imposes a cost in terms of lower earnings in the labor market (over 10%). This paper extends Gaviria's et al. (2007) research by focusing on the barely explored subject of the relationship between atypical names and academic outcomes. To the best of our knowledge, there are only few international studies dealing with this subject (mostly by Figlio, 2005a, 2005b). Up to date, there is no research for a developing country on this topic.

Our results show that atypical names have a negative effect on math and reading scores, and that carrying such kind of names has a higher academic cost for boys, for students from full-day schools, as well as for high tuition fees pupils. These findings contribute to the academic debate on the influence of names on human capital outcomes and shed some light on the possibility of discriminatory practices in the classroom. The rest of the paper is structured as follows: section II presents the literature review; section III presents the methodology; section IV highlights the main results; section VI discusses these results; and the last section concludes.

II. Literature Review

This section presents a review of the factors behind the choice of a child's name, as well of the different impacts on labor and academic outcomes of carrying a name that is linked to a minority group.

A. Name Practices

As a rule, the purpose of carrying a name is twofold: to differentiate but also to categorize the name bearer (Mateos, 2014a). Name practices are the result of the interaction of different socio-cultural, socioeconomic, religious, geographic, and linguistic dimensions, and so naming practices follow patterns and “fashion waves” (Mateos, 2014a, 2014b). Parents are not only influenced by their own preferences

² Darwin Cortés: Associated professor, Department of Economics, Universidad del Rosario, Bogotá, Colombia. Darío Maldonado: Associated professor, School of Government Alberto Lleras Camargo, Universidad de los Andes, Bogotá, Colombia. Mauricio G. Valencia-Amaya: Ph.D. student, Department of Economics, Universidad del Rosario, Bogotá, Colombia.

but also by society preferences at a time, their social class, their national origins, and the naming decisions of other parents (Mateos, 2014b; Head and Mayer, 2007).

For instance, Bloothoof and Onland (2011) categorize the name practices in the Netherlands during the last two decades of the 20th century based on two independent aspects: (1) education and family income, and (2) parents' religiosity and conservative behavior vs. trendy-like behavior. Their results indicate that less educated and poorer parents tend to name their children with non-Dutch names, probably influenced by the Anglo-American media (see Disdier et al., 2010 for the French case); whereas more educated and richer parents tend to stick to traditional Dutch-names. But, independent of education and income, the authors claim that parents' fashion awareness seem to play also an important role in the decision-making process of naming a child.

B. Names and Labor Market Outcomes

Although some studies, such as Fryer and Levitt (2004) find no negative effects of having a distinctive name on labor market outcomes (see Martínez, 2013 for the Spanish case), once controlling for specific birth factors, some other studies go so far as to recommend, based on their conclusions, that names should be left off the curriculum vitae (Bertrand and Mullainathan, 2004; Cotton et al., 2008; Watson et al., 2011). This section reviews some of the recent literature on the names and labor outcomes.

Arai and Skogman-Thoursie (2009) analyze the impact on the labor earnings of immigrants of changing inherited surnames for neutral or Swedish-sound surnames. They conclude that by doing so, these immigrants increased their annual earnings, and therefore there is evidence of discrimination in the Swedish labor market based on the country of origin; that is, employers seem to have a more negative attitude towards non-Swedish-sound surnames (see Rubinstein and Brenner, 2014 for the Sephardic/Ashkenazi Jews case). A similar result was found by Biavaschi et al. (2013) for the American case. The authors study a random sample of immigrants to the United States during the first decades of the 20th century, specifically, those who by 1930 had completed their naturalization process. They find that those who changed their given names for popular American names experienced a significant increase in their labor earnings, having a positive effect on the upward mobility of these immigrants.

However, it seems that having an American name is not sufficient for minorities to succeed in the U.S. labor market. According to Bertrand and Mullainathan (2004) chances of getting a callback for a job interview increases by having a white-sounding name, as opposed to having an African-American name. Moreover, at least in legal careers in the U.S., having a masculine moniker seems to be correlated with more successful careers for women, according to the results of Coffey and McLaughlin (2009) who empirically tested the Portia hypothesis. Finally, Kalist and Lee (2009) find a positive correlation between unpopular names and delinquency, independent of ethnic background. The authors suggest that low socioeconomic backgrounds, like having an uneducated mother, who is more likely to give their children unpopular names (Lieberson and Bell, 1992; Aura and Hess, 2004, cited by Kalist and Lee, 2009), and inappropriate home environment, like living in nontraditional households, may be the link behind the relationship between name and crime. Although no causal relationship is established, they put forward the possibility of a casual relation in cases when having an unpopular name leads to job market discrimination.

C. Names Perceptions as a Possible Channel

Cotton et al. (2008) study the perceptions, affective reactions, and hireability of unique/common and ethnic-related first names. Their results suggest not only that more unique first names vis-à-vis common

first names are least liked by the evaluated respondents, as they are related to least favorable characteristics, but they are also least likely to be hypothetically hired. What is more, these findings do not seem to be necessarily associated with its ethnicity, but with the uniqueness of it, and so there may be a certain bias in people’s perception of a name, in terms of preconceived stereotypes, that might trigger negative reactions towards their bearers. Gebauer et al. (2012) come to similar conclusions, when comparing online-daters popularity based on their given names: online-daters with negative names receive fewer visits to their dating profiles. The authors suggest that negative names affect the name bearer’s interpersonal relations, and therefore they have a harmful effect on different social outcomes, such as school performance, since the negative impression of a given name may lead to discrimination by peers and teachers.

Finally, by sending identical fictional emails of prospective PhD students with randomly assigned names to staff of an American university, Milkman et al. (2014) find that women and minorities comparatively received less replies than Caucasian males, suggesting that ethnicity or gender identifiers, such as names, may become obstacles at some organizational entry points. This may even lower the academic opportunities of specific-name bearers, as will be addressed in the next section.

D. Impacts of Names on Academic Performance

The literature points out also a possible negative impact of certain students’ name characteristics on their academic performance. For instance, Figlio (2005a) finds that boys having girl-sounding names tend to misbehave more compared with their counterparts, as well as with their previous behavior during primary school. In other related paper (Figlio, 2005b), the same author suggests that teachers may treat students differentially based on the characteristics of their name: Teachers presumably infer unobserved traits of the contribution of parents on the education of their children and may discriminate against students with names denoting a low socioeconomic background, by expecting academically less from them. This chain of events may lead to worse cognitive performance and lower test scores of students with certain name characteristics. Therefore, since teachers’ stereotypes influence students’ academic performance (Ambady et al., 2001, cited by Figlio, 2005b), their selective treatment seems to lead to a self-fulfilling prophecy.

However, Foster (2008) claims that having a particular name, such as a black name, is a byproduct of segregation and poverty, and not the other way around. In fact, she finds little support for the possible effect of racial identifiability of first and last names on undergraduate grades and so argues that “*any raw association between racially identifiable names and outcomes is the by-product of a third-party cause of both*”. Yet, according to the author, it is possible that name-based effects may have already set before starting university studies.

III. Methodology

A. Identification Strategy

Consider an education production function of the form (Hanushek, 1979):

$$Y_{it} = f(X_{it}, P_{it}, S_{it}, T_{it}, I_i) \quad (1)$$

where the student i academic outcome Y (i.e. test scores) at time t is a function of the cumulative influence of his/her family (X) since birth and up to time t ; the cumulative influences of his/her peers (P), school (S), and teachers (T) since the first year of school and up to time t ; and his/her innate abilities (I). Now, when the cumulative effect of inputs in an education production function is not observable, the regression coefficients of equation (1) will be biased on observable characteristics (omitted-variable bias).

A solution to this issue is to include a lagged academic outcome Y_{it-1} , to account for cumulative inputs that explain the outcome at time t but that are not available in the database.

Accordingly, our approach relies on the following equation:

$$Y_{i2013} = \beta_0 + \beta_1 \textit{Atypical_Name}_i + \beta_2 Y_{i2012} + \beta_3 X_{i2013} + \theta_s + \varepsilon_{it} \quad (2)$$

Where Y_{i2013} is the z-score test result of the student i in the “Olimpiadas del Conocimiento” test in the year 2013 (reading and math); $\textit{Atypical_Name}_i$ is a dummy equal to 1 if the student i has no other namesakes in either his/her first name or middle name among all registered students in the Department of Antioquia in the year 2013; Y_{i2012} is the z-score test result of the student i in the “Olimpiadas del Conocimiento” test in the year 2012 (reading and math); X_{i2013} is a vector of the student’s i control variables that includes sex (dummy equal to 1 if the student is male), grade repeater (dummy equal to 1 if the student is a grade repeater), and underage (dummy equal to 1 if the student is younger than 18 years old); θ_s is a dummy variable that captures school effects; and ε_{it} is an error term. Table 1 presents the descriptive statistics of the variable $\textit{Atypical_Name}_i$.

Table 1. Descriptive Statistics of the “Atypical Name” Variable.

Variable	Observations	Participation
Atypical Name	296	
By Sex		
<i>Males</i>	99	0.33
<i>Females</i>	197	0.67
By School day		
<i>Full day</i>	77	0.26
<i>Morning</i>	129	0.44
<i>Afternoon</i>	83	0.28

Source: Authors’ calculations based on “Olimpiadas del Conocimiento” test.

It is important to note that the “Olimpiadas del Conocimiento” test is unbiased with respect to the variable $\textit{Atypical_Name}_i$. This is a standardized test that is graded by a computer and so the results are expected to measure the cumulative performance of students, according to the education production function. Therefore, we can think of the impact of atypical names on scores to be free from the bias of an external grader.

Now, given that the name structure is fixed throughout time, the effect of having an atypical name is already reflected in Y_{i2012} , that is the test score of the previous year. The effect of $\textit{Atypical_Name}_i$ on Y_{i2013} will reflect then the additional impact of carrying such type of name on tests scores from 2012 to 2013. Adolescence is a transitional stage from childhood to adulthood, during which the individual constantly experiences physical, emotional, and psychological changes. Is because of these ongoing changes that we can expect our variable of interest to still influence the academic performance of students.

B. Data

This paper uses the results of a standardized test called “Olimpiadas del Conocimiento”. This is a unique dataset relating full names and test scores in math and reading for the years 2012 and 2013. The dataset contains information on the results of a regional test applied by the Antioquia governorship to public and private high school students (grades 10th and 11th) of 124 Antioquia’s municipalities, excluding the

capital city Medellín, for which an independent test is applied. The test measures mainly quantitative and verbal skills, and it is divided into three sections: (1) mathematics —20 questions on algebra, trigonometry, pre-calculus, sets, and statistics—, (2) reading —20 questions on syntax, semantics, and pragmatics—, and (3) a flexible component —eight questions on social science and natural science—. The test is given in Spanish. For the mathematics and reading sections, the final score ranges from 0 to 100, as each correct answer is awarded 5 points; for the flexible component, the final score ranges from 0 to 40. In 2012, 49,347 students from 477 different schools took the test; whereas, in 2013, these figures were 69,742 and 665, respectively. Out of the 69,742 students who took the test in 2013, 16,489 students took it also in 2012. This unique setting allows us to control for students’ test score from the previous year (2012), but also reduces our sample size. To test whether the remaining sample is representative of Antioquia Department, we conducted a t-test to compare the mean of one of our control variables (sex) with that of the group of students from Antioquia (excluding Medellín) who took the ICFES “Saber 11” national test in 2013. According to the t-test result of Table 2, we cannot reject the null hypothesis that the mean difference between the two groups is zero. Table 3 presents the descriptive statistics of “Olimpiadas del Conocimiento” dataset.

Table 2. T-test of the Mean Difference in the Gender Structure between “Olimpiadas del Conocimiento” and ICFES “Saber 11” Groups.

Groups	Observations	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
Saber 11	36,696	0.429	0.003	0.495	0.424	0.434
Olimpiadas del Conocimiento	14,335	0.432	0.004	0.495	0.424	0.440
Combined	51,031	0.430	0.002	0.495	0.426	0.434
Difference		-0.003	0.005		-0.013	0.006
Ho: diff = 0	Ha: diff < 0		Ha: diff != 0		Ha: diff > 0	
DF = 51,029	Pr(T < t) = 0.245		Pr(T > t) = 0.491		Pr(T > t) = 0.7542	

Source: Authors’ calculations based on “Olimpiadas del Conocimiento” and ICFES “Saber 11” tests.

Table 3. “Olimpiadas del Conocimiento”: Descriptive Statistics.

Variable	Obs.	Mean	Std. Dev.	Min.	Max.	
Male	16,455	0.45	0.50	0.00	1.00	
Grade repeater	16,455	0.09	0.29	0.00	1.00	
Under-age (younger than 18 years old)	16,455	0.91	0.29	0.00	1.00	
Math (2013)	General	16,455	37.00	14.06	0.00	100.00
	Males	7,388	39.16	14.59	0.00	95.00
	Females	9,067	35.25	13.36	0.00	100.00
Math (2012)	General	16,455	35.08	13.27	5.00	100.00
	Males	7,388	36.20	13.64	5.00	100.00
	Females	9,067	34.17	12.88	5.00	100.00
Reading (2013)	General	16,455	44.38	14.42	0.00	95.00
	Males	7,388	44.52	14.47	0.00	95.00
	Females	9,067	44.26	14.37	0.00	90.00
Reading (2012)	General	16,455	42.16	14.30	5.00	95.00
	Males	7,388	41.29	14.37	5.00	90.00
	Females	9,067	42.86	14.20	5.00	95.00

Source: Authors’ calculations based on “Olimpiadas del Conocimiento” test.

IV. Results

A. Baseline

Tables 4 presents the estimation results of equation (2) using ordinary least squares (OLS) with standard errors clustered at the school level. Controls are added in a stepwise way. According to the results, having an atypical name has a significant and negative impact on math and reading scores. The size of the effect is similar on both subjects (around 0.1 standard deviations). The results are robust to three different specifications (columns i, ii, and iii): The first is the simple model; the second controls for covariates related to characteristics of the student, such as sex, age, and grade repetition; and the third, additionally to these controls, includes the test scores of the previous year and accounts for school fixed effects.

Table 4. Effect of Atypical Names on the “Olimpiadas del Conocimiento” Test Scores, Regression Results: OLS.

Model	(i)		(ii)		(iii)		(iv)	
Dependent Variable: 2013 z-score	Math	Reading	Math	Reading	Math	Reading	Math	Reading
Atypical Name (dummy)	-0.179*** (0.055)	-0.219*** (0.058)	-0.140*** (0.053)	-0.207*** (0.057)	-0.0934* (0.050)	-0.119*** (0.056)	-0.098** (0.050)	-0.111** (0.053)
R-sqr	0.001	0.001	0.033	0.019	0.214	0.152	0.256	0.225
Number of observations	16,455	16,455	16,455	16,455	16,455	16,455	16,455	16,455
Controls			X	X	X	X	X	X
2012 z-score					X	X	X	X
School fixed effects							X	X

Notes: (a) *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; (b) standard errors in parenthesis are clustered at the school level; (c) the coefficients for “Atypical Name” is the estimated parameter β_1 of equation (2); (d) “Atypical Name” is a dummy variable equal to 1 if a given name has no other namesakes in the list of all students enrolled at Antioquia Department in 2013; (e) control variables include dummy variables for male, under-age students (younger than 18 years old), and grade repeaters; (f) all scale variables were converted into z-scores.

B. Heterogeneous Effects

This section explores the existence of heterogeneous effects along the dimensions of sex and school day. Table 5 reports the estimation results by sex. In general, the results show that atypical names have a strong, negative, and significant effect on males’ scores in both test components, but not on females’. The size of the effect on male’s reading scores is almost twice as large as that on math scores. Finally, Table 6 shows the heterogeneous effects by the length of the school day. The results state that the impact of having an atypical name is significant only for students of the full day length, for whom the negative effect is around 0.3 standard deviations.

Table 5. Effect of Atypical Names on the “Olimpiadas del Conocimiento” Test Scores, by Sex, Regression Results: OLS.

Sex	Dependent Variable: 2013 z-score	Math	Reading
Males	Atypical Name (dummy)	-0.193** (0.087)	-0.332*** (0.085)
Females	Atypical Name (dummy)	-0.044 (0.060)	0.000 (0.070)
R-sqr		0.292	0.225
Number of observations		16,455	16,455
Controls		X	X
School fixed effects		X	X
2012 z-score		X	X

Notes: (a) *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; (b) standard errors in parenthesis are clustered at the school level; (c) the coefficients for “Atypical Name” is the estimated parameter β_1 of equation (2); (d) “Atypical Name” is a dummy variable equal to 1 if a given name has no other namesakes in the list of all students enrolled at Antioquia Department in 2013; (e) control variables include dummy variables for male, under-age students (younger than 18 years old), and grade repeaters; (f) all scale variables were converted into z-scores.

Table 6. Effect of Atypical Names on the “Olimpiadas del Conocimiento” Test Scores, by School Day, Regression Results: OLS.

School Day	Dependent Variable: 2013 z-score	Math	Reading
Full Day	Atypical Name (dummy)	-0.345*** (0.082)	-0.289** (0.117)
Morning	Atypical Name (dummy)	0.024 (0.069)	-0.006 (0.068)
Afternoon	Atypical Name (dummy)	-0.043 (0.101)	-0.092 (0.104)
R-sqr		0.257	0.225
Number of observations		16,455	16,455
Controls		X	X
2012 z-score		X	X
School fixed effects		X	X

Notes: (a) *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; (b) standard errors in parenthesis are clustered at the school level; (c) the coefficients for “Unique Name” is the estimated parameter β_1 of equation (2); (d) “Unique Name” is a dummy variable equal to 1 if a given name has no other namesakes in the list of all students enrolled at Antioquia Department in 2013; (e) control variables include dummy variables for male, under-age students (younger than 18 years old), and grade repeaters; (f) all scale variables were converted into z-scores.

C. Robustness Check

Table 7 presents the estimation results of equation (2) modifying the definition of $Atypical_Name_i$. In this case, this categorical variable takes three possible values: 0 if the student i has a common name; 1 if the student i has a unique name —meaning not having a namesake in both his/her first name and middle name among all registered students in the Department of Antioquia in the year 2013—; and 2 if the student i has a combination of a unique name and a common name in either his/her first name or middle name. The results show that among atypical names bearers, those who have at least one common name perform better in the test, vis-à-vis the group that has no common names. In comparison with the results

from Table 5, the size of the β_1 coefficient is bigger for males without common names in both reading and math; for males with at least one common name the size is smaller, barely significant for reading and non-significant for math. For females, the coefficients remain close to zero and non-significant.

Table 7. Effect of Atypical Names on the “Olimpiadas del Conocimiento” Test Scores, Atypical Names with and without Common Names, by Sex, Regression Results: OLS.

Sex	Dependent Variable: 2013 z-score	Reading	Math
Male	Atypical Name (without common names)	-0.398*** (0.118)	-0.283** (0.128)
	Atypical Names (with at least one common name)	-0.236* (0.124)	-0.0522 (0.109)
Female	Atypical Name (without common names)	0.0164 (0.086)	-0.0445 (0.072)
	Atypical Names (with at least one common name)	-0.0347 (0.108)	-0.0712 (0.106)
R-sqr		0.226	0.256
Number of observations		16,455	16,455
Controls		X	X
2012 z-score		X	X
School fixed effects		X	X

Notes: (a) *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; (b) standard errors in parenthesis are clustered at the school level; (c) the coefficients for “Atypical Name” is the estimated parameter β_1 of equation (2); (d) “Atypical Name” is a dummy variable equal to 1 if the student i has a unique name —meaning not having a namesake in both his/her first name and middle name among all registered students in the Department of Antioquia in the year 2013— and 2 if the student i has a combination of a unique name and a common name in either his/her first name or middle name; (e) control variables include dummy variables for male, under-age students (younger than 18 years old), and grade repeaters; (f) all scale variables were converted into z-scores.

Table 8 presents another modification of the variable $Atypical_Name_i$ of equation (2). Instead of using it to indicate the presence of a unique name, we use it to test the effect of having a unique surname, which a priori should not possibly influence students’ test scores. As expected, we found no significant effects of this modified name structure on either math or reading scores—for both males and females.

Table 8. Effect of Atypical Surnames on the “Olimpiadas del Conocimiento” Test Scores, by Sex, Regression Results: OLS.

Dependent Variable: 2013 z-score	Math	Reading	Math		Reading	
			Males	Females	Males	Females
Atypical Surname (dummy)	-0.106 (0.137)	-0.089 (0.121)	-0.123 (0.219)	-0.086 (0.158)	-0.126 (0.168)	-0.046 (0.169)
R-sqr	0.256	0.225	0.235		0.204	
Number of observations	16,455	16,455	16,455		16,455	
Controls	X	X	X	X	X	X
2012 z-score	X	X	X	X	X	X
School fixed effects	X	X	X	X	X	X

Notes: (a) *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; (b) standard errors in parenthesis are clustered at the school level; (c) the coefficients for “Atypical Surname” is the estimated parameter β_1 of equation (2); (d) “Atypical Surname” is a dummy variable equal to 1 if a surname has no other namesakes in the list of all students enrolled

at Antioquia Department in 2013; (e) control variables include dummy variables for male, under-age students (younger than 18 years old), and grade repeaters; (f) all scale variables were converted into z-scores.

V. Discussion

The findings presented in the previous session showed that carrying atypical names seems to have a negative effect on the math and reading test scores. However, the effect is statistically significant only for boys. A possible explanation for this result might be related to bullying. Students with atypical names constitute a minority —only 1,8 percent of the people in our dataset have such type of name— and, in general, minorities suffer more from bullying (Zych, Ortega-Ruiz and Del Rey, 2015). Now, carrying unique names might affect students' popularity —being liked by peers—, which seems to be a buffer against this behavior (Pellegrini, 2004); and so, there is even a higher probability for this group of students to be victims of bullying. In fact, students who are bullied are less popular than their counterparts (Sanders, 2004), which reinforce the vicious circle. Name-calling is also one of the most common forms of direct verbal aggression (Al-Bitar et al., 2013). In terms of gender differences, the literature states that boys are more likely to experience bullying than girls (Ma, 2004; Pellegrini, 2004; Ponzo, 2013; Al-Bitar et al., 2013; Sullivan, Cleary, and Sullivan, 2004; Zych, Ortega-Ruiz, and Del Rey, 2015). But not only that, compared with girls, males are more likely to be both the bullies and the victims, as boys use this behavior as a strategy to gain or maintain a social status (Marsh et al., 2004).

Overall, bullying has a negative impact on students' academic outcomes (Konishi et al., 2010; Ponzo, 2013; Popp et al., 2014; van der Werf, 2014). Its prevalence is not only negatively related to academic achievement at the school level (Lacey and Cornell, 2013), but also at the individual level, as interpersonal relations within the school affect students' academic performance (Konishi et al., 2010). It has been found, for example, that bullying lowers math, reading, and science tests scores —with a larger impact on reading than on math— and that its negative effect persists as students get older (Ponzo, 2013).

However, differences in how men and women rely on social support networks can deepen or attenuate the negative impact of this behavior. In general, social support helps lessen its negative effect on internalizing distress (Davidson and Demaray, 2007), and family relationships mediates between stressors and psychopathologies (Grant et al., 2005). Now, women's social networks are larger and more diversified than men's. For instance, social support is more effective in lowering risk for depression in women than in men (Kendler, Myers, and Prescott, 2005). In this sense, a strong and diverse social support network, together with low involvement in bullying behavior, might help girls buffer the negative effect of bullying on test scores.

The results showed also that the academic cost of having atypical names was higher for students from high tuition fees schools. One possible explanation is related to the big-fish-little-pond effect model. This effect, which is explained, among others, by the social comparison theory, states that students with similar skills will have lower/higher academic self-concepts when attending high/low average ability schools (Marsh et al., 2008). The effect is captured by how students compare themselves with their reference group (Huguet et al., 2009, cited by Seaton, Marsh, and Craven, 2010). The big-fish-little-pond effect has been tested and confirmed by different studies within educational psychology (Zeidner and Schleyer, 1998; Marsh and Hau, 2003; Seaton, Marsh, and Craven, 2010), but also applied to other fields; for example, Goldsmith (2004) concludes that black and Latino students' beliefs are less optimistic in segregated-white schools, in comparison with segregated-minority schools. In our case, the prevalence of atypical names is almost 12 times lower in high tuition fees schools than in low tuition fees schools. This

means that students with atypical names will be more notorious in institutions where most of their classmates carry common names and this, in turn, can trigger the big-fish-little-pond effect.

The impact of atypical names is significant only for students attending full-day schools. The coefficients for half-day schools (morning and afternoon) are close to zero and statistically insignificant. The intuition behind this result is supported by the literature on the negative effects of exposure to aggressive classrooms. For example, students who are exposed to multiple years of classroom aggression develop more hostile behaviors than students with less exposure to this kind of environment (Thomas and Bierman, 2006). Also, the duration and intensity of bullying are related to a higher probability of suffering from depression in adulthood (Lund et al., 2008). Moreover, the higher the frequency of bullying at school, the higher the incidence of physical and mental health issues, which are negatively related to short- and long-term human capital outcomes (Due et al., 2005).

Finally, although not explored in this paper, there exists also the possibility of a teachers' discriminatory behavior. As argued by Figlio (2005b), students' names carry information on their socioeconomic and ethnic background, and this in turn can be used by teachers and peers as a discriminating tool. Different studies have found that teachers' expectations affect students' performance —known as the Pygmalion effect— (Rosenthal, 2002; Friedrich et al., 2015). Teachers' perceptions of observable characteristics of students influence also their interactions, reactions, and focus of attention (Morris, 2005). What is more, different studies suggest that teachers' grading practices might be gender bias, which might affect students' learning outcomes, motivation, and career decisions (Terrier, 2014; Lavy and Sand, 2015). Therefore, it is possible that students carrying atypical names be treated differently by their teachers, and if grading practices are also found to be name-bias, then the achievement gap in education for already disadvantaged students can be exacerbated.

Conclusion

In this paper we studied the effect of atypical names on academic performance. We used the “Olimpiadas del Conocimiento” test scores, a test applied to high school students (grades 10th and 11th) of Antioquia Department (Colombia) in the years 2012 and 2013. Our results indicate that, even after controlling for students' and schools' fixed characteristics, atypical names have a negative effect on students' math ($p < 0.05$) and reading ($p < 0.05$) tests scores. We found that carrying atypical names decreased boys' test scores, in both math ($p < 0.05$) and reading ($p < 0.01$) but did not affect girls'. The impact was larger for students from high tuition fees schools ($p < 0.01$ for both subjects) —the effect on low fee schools was barely significant ($p < 0.1$)—, as well as for students from full-day institutions ($p < 0.01$ for math and $p < 0.05$ for reading). There were no significant effects for half-day schools. We discussed bullying behavior and gender differences in social support as possible explanations for the differential effect on boys and girls. We argued that psychological models, such as the big-fish-little-pond effect, might help explain some of the other heterogeneous results. Finally, as supported by the literature, it is possible that atypical names might elicit negative perceptions on their bearers or on their backgrounds, generating a discriminatory behavior from teachers and/or peers, which would affect their knowledge acquisition and, therefore, their academic outcomes and future career prospects. We propose the development of experimental studies and further research, to identify possible channels of transmission.

References

AL-BITAR, Zaid B.; AL-OMARI, Iyad K.; SONBOL, Hawazen N.; AL-AHMAD, Hazem T.; and CUNNINGHAM, Susan J. (2013). “Bullying among Jordanian schoolchildren, its effects on school

performance, and the contribution of general physical and dentofacial features.” *American Journal of Orthodontics and Dentofacial Orthopedics*, Vol. 144, No. 6, pp. 872-878.

ANTONUCCI, Toni C. and AKIYAMA, Hiroko (1987). “An Examination of Sex Differences in Social Support Among Older Men and Women.” *Sex Roles*, Vol. 17, No. 11/12, pp. 737-749.

ARAI, Mahmood and SKOGMAN-THOURSIE, Peter (2009). “Renouncing Personal Names: An Empirical Examination of Surname Change and Earnings.” *Journal of Labor Economics*, Vol. 27, No. 1, pp. 127-147.

ALDRIN, Emilia (2011). *Namnval som social handling. Val av förnamn och samtal om förnamn bland föräldrar i Göteborg 2007-2009*. Uppsala Universitet, Namn och Samhälle 24, 297 pp.

BERTRAND, Marianne and MULLAINATHAN, Sendhil (2004). “Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination.” *The American Economic Review*, Vol. 94, No. 4, pp. 991-1013.

BIAVASCHI, Costanza; GIULIETTI, Corrado; SIDDIQUE, Zahra (2013). “The Economic Payoff of Name Americanization.” *IZA Discussion Paper*, No. 7725.

BLOOTHOOFT, Gerrit and ONLAND, David (2011). “Socioeconomic Determinants of First Names.” *Names*, Vol. 59, No. 1, pp. 25-41.

COFFEY, Bentley and MCLAUGHLIN, Patrick (2009). “Do Masculine Names Help Female Lawyers Become Judges? Evidence from South Carolina.” *American Law and Economics Review*, Vol. 11, No. 1, pp. 112-133.

COLLADO, M. Dolores; ORTUÑO-ORTÍN, Ignacio; and ROMEU, Andrés (2008). “Surnames and Social Status in Spain.” *Investigaciones Económicas*, Vol. XXXII, No. 3, pp. 259-287.

COTTON, John; O’NEILL, Bonnie; and GRIFFIN, Andrea (2008). “The “Name Game”: Affective and Hiring Reactions to First Names.” *Journal of Managerial Psychology*, Volume 23, No. 1, pp. 18-39.

DAVIDSON, Lisa M. and DEMARAY, Michelle K. (2007). “Social Support as a Moderator Between Victimization and Internalizing-Externalizing Distress from Bullying.” *School Psychology Review*, Vol. 36, No. 3, pp. 383-405.

DISDIER, Anna-Célia; HEAD, Keith; and MAYER, Thierry (2010). “Exposure to foreign media and changes in cultural traits: Evidence from naming patterns in France.” *Journal of International Economics*, Vol. 80, No. 2, pp. 226-238.

DUE, Pernille; HOLSTEIN, Bjørn E.; LYNCH, John; DIDERICHSEN, Finn; GABHAIN, Saoirse Nic; SCHEIDT, Peter; and CURRIE, Candace (2005). “Bullying and symptoms among school-aged children: international comparative cross-sectional study in 28 countries.” *The European Journal of Public Health*, Vol. 15, No. 2, pp. 128-132.

FIGLIO, David N. (2005a). “Boys Named Sue: Disruptive Children and Their Peers.” *NBER Working Paper Series*, No. 11277.

FIGLIO, David N. (2005b). “Names, Expectations, and the Black-White Test Score Gap.” *NBER Working Paper Series*, No. 11195.

- FOSTER, Gigi (2008). "Names will never hurt me: Racially distinct names and identity in the undergraduate classroom." *Social Science Research*, No. 37, pp. 934–952.
- FRIEDRICH, Alena; FLUNGER, Barbara; NAGENGAST, Benjamin; and JONKMANN, Kathrin (2015). "Pygmalion effects in the classroom: Teacher expectancy effects on students' math achievement." *Contemporary Educational Psychology*, No. 41, pp. 1-12.
- FRYER, Roland and LEVITT, Steven (2004). "The Causes and Consequences of Distinctively Black Names." *The Quarterly Journal of Economics*, Vol. CXIX, No. 3, pp. 767-805.
- GAVIRIA, Alejandro; MEDINA, Carlos; y PALAU, María del Pilar (2007). "Las consecuencias económicas de un nombre atípico. El caso colombiano." *Documento CEDE* 2007-09.
- GEBAUER, Jochen; LEARY, Mark; and NEBERICH, Wiebke (2012). "Unfortunate First Names: Effects of Name-Based Relational Devaluation and Interpersonal Neglect." *Social Psychological and Personality Science*, Vol. 3, No. 5, pp. 590-596.
- GRANT, Kathryn; COMPAS, Bruce E.; THURM, Audrey E.; MCMAHON, Susan D.; GIPSON, Polly Y.; CAMPBELL, Amanda J.; KROCHOCK, Kimberly; and WESTERHOLM, Robert I. (2006). "Stressors and child and adolescent psychopathology: Evidence of moderating and mediating effects." *Clinical Psychology Review*, No. 26, pp. 257– 283.
- HANUSHEK, Eric A. (1979). "Conceptual and Empirical Issues in the Estimation of Educational Production Functions." *The Journal of Human Resources*, Vol. 14, No. 3, pp. 351-388.
- HEAD, Keith and MAYER, Thierry (2007). "Detection of local interactions from the spatial pattern of names in France." *Journal of Regional Science*, No. 48, No. 1, pp. 67-95.
- KALIST, David and LEE, Daniel (2009). "First Names and Crime: Does Unpopularity Spell Trouble?" *Social Science Quarterly*, Vol. 90, No. 1, pp. 39-49.
- KENDLER, Kenneth; MYERS, John; and PRESCOTT, Carol A. (2005). "Sex Differences in the Relationship Between Social Support and Risk for Major Depression: A Longitudinal Study of Opposite-Sex Twin Pairs." *American Journal of Psychiatry*, No. 162, pp. 250-256.
- KONISHI, Chiaki; HYMEL, Shelley; ZUMBO, Bruno D.; and LI, Zhen (2010). "Do School Bullying and Student–Teacher Relationships Matter for Academic Achievement? A Multilevel Analysis." *Canadian Journal of School Psychology*, Vol. 25, No. 1, pp. 19-39.
- LACEY, Anna and CORNELL, Dewey (2013). "The Impact of Teasing and Bullying on Schoolwide Academic Performance." *Journal of Applied School Psychology*, Vol. 29, No. 3, pp. 262-283.
- LAVY, Victor and SAND, Edith (2015). "On the Origins of Gender Human Capital Gaps: Short And Long-Term Consequences Of Teachers' Stereotypical Biases." *NBER Working Paper Series*, No. w20909.
- LUND, Rikke; NIELSEN, Karoline; HANSEN, Ditte; KRIEGBAUM, Margit; MOLBO, Drude; DUE, Pernille; and CHRISTENSEN, Ulla (2008). "Exposure to bullying at school and depression in adulthood: A study of Danish men born in 1953." *European Journal of Public Health*, Vol. 19, No. 1, pp. 111–116.
- MA, Xin (2004). "Who are the victims?" In: Sanders, Cheryl E. and Phye, Gary D. (eds.), *Bullying: Implications for the Classroom*, Elsevier Academic Press, London, pp. 20-35.

- MARSH, Herbert W.; HAU, Kit-Tai (2003). "Big-Fish--Little-Pond effect on academic self-concept: A cross-cultural (26-country) test of the negative effects of academically selective schools." *American Psychologist*, Vol. 58, No. 5, pp. 364-376.
- MARSH, Herbert W.; PARADA, Roberto H.; CRAVEN, Rhonda G.; and FINGER, Linda (2004). "In the looking glass: a reciprocal effect model elucidating the complex nature of bullying, psychological determinants, and the central role of self-concept." In: Sanders, Cheryl E. and Phye, Gary D. (eds.), *Bullying: Implications for the Classroom*, Elsevier Academic Press, London, pp. 63-110.
- MARSH, Herbert W.; SEATON, Marjorie; TRAUTWEIN, Ulrich; LÜDTKE, Oliver; HAU, K. T.; O'MARA, Alison J.; and CRAVEN, Rhonda G. (2008). "The Big-fish--little-pond-effect Stands Up to Critical Scrutiny: Implications for Theory, Methodology, and Future Research." *Educational Psychology Review*, Vol. 20, No. 3, pp. 319-350.
- MARTÍNEZ, Jose A. (2013). "Do names matter? The influence of names on perception about professionals in Spain." *Economics and Business Letters*, Vol. 2, No. 2, pp. 66-74.
- MATEOS, Pablo (2014a). "How We Got Our Names: Identity in Personal Names." In: Pablo Mateos, *Names, Ethnicity and Populations. Tracing Identity in Space*, ed. Springer, pp. 29-47.
- MATEOS, Pablo (2014b). "Forenames and Social Stratification." in: Pablo Mateos, *Names, Ethnicity and Populations. Tracing Identity in Space*, ed. Springer, pp. 81-114.
- MILKMAN, Katherine; AKINOLA, Modupe; and CHUGH, Dolly (2014). "What Happens Before? A Field Experiment Exploring How Pay and Representation Differentially Shape Bias on the Pathway into Organizations." Electronic copy available at: <http://ssrn.com/abstract=2063742>.
- MORRIS, Edward W. (2005). "From "Middle Class" to "Trailer Trash:" Teachers' Perceptions of White Students in a Predominately Minority School." *Sociology of Education*, Vol. 78, No. 2, pp. 99-121.
- PELLEGRINI, A.D. (2004). "Bullying during middle school years." In: Sanders, Cheryl E. and Phye, Gary D. (eds.), *Bullying: Implications for the Classroom*, Elsevier Academic Press, London, pp. 177-202.
- PONZO, Michela (2013). "Does bullying reduce educational achievement? An evaluation using matching estimators." *Journal of Policy Modeling*, No. 35, pp. 1057-1078.
- POPP, Ann Marie; PEGUERO, Anthony A.; DAY, Kristin R.; and KAHLE, Lindsay L. (2014). "Gender, Bullying Victimization, and Education." *Violence and Victims*, Vol. 29, No. 5, pp. 843-856.
- ROSENTHAL, Robert (2002). "The Pygmalion Effect and Its Mediating Mechanisms." In: Joshua Aronson (ed.), *Improving Academic Achievement. Impact of Psychological Factors in Education*, Academic Press, pp. 25-36.
- RUBINSTEIN, Yona and BRENNER, Dror (2014). "Pride and Prejudice: Using Ethnic-Sounding Names and Inter-Ethnic Marriages to Identify Labor Market Discrimination." *Review of Economic Studies*, Vol. 81, No. 1, pp. 389-425.
- SANDERS, Cheryl E. (2004). "What is bullying?" In: Sanders, Cheryl E. and Phye, Gary D. (eds.), *Bullying: Implications for the Classroom*, Elsevier Academic Press, London, pp. 2-19.

SEATON, Marjorie; MARSH, Hebert W.; and CRAVEN, Rhonda G. (2010). "Big-Fish-Little-Pond Effect: Generalizability and Moderation—Two Sides of the Same Coin." *American Educational Research Journal*, Vol. 47, No. 2, pp. 390-433.

SULLIVAN, Keith; CLEARY, Mark; and SULLIVAN, Ginny (2004). *Bullying in Secondary Schools. What it Looks Like and How to Manage it*. Paul Chapman Publishing, London, 272 pp.

TERRIER, Camille (2014). "A helping hand for girls? Gender bias in marks and its effect on student progress." *IPP Policy Brief*, No. 14.

THOMAS, Duane E. and BIERMAN, Karen L. (2006). "The impact of classroom aggression on the development of aggressive behavior problems in children." *Development and Psychopathology*, Vol. 18, No. 2, pp. 471–487.

WATSON, Stevie; APPIAH, Osei; and THORNTON, Corliss (2011). "The Effect of Name on Pre-Interview Impressions and Occupational Stereotypes: The Case of Black Sales Job Applicants." *Journal of Applied Social Psychology*, Vol. 41, No. 10, pp. 2405–2420.

ZEIDNER, Moshe and SCHLEYER, Esther Jane (1998). "The Big-Fish–Little-Pond Effect for Academic Self-Concept, Test Anxiety, and School Grades in Gifted Children." *Contemporary Educational Psychology*, No. 24, pp. 305–329.

ZYCH, Izabela; ORTEGA-RUIZ, Rosario, DEL REY, Rosario (2015). "Systematic review of theoretical studies on bullying and cyberbullying: Facts, knowledge, prevention, and intervention." *Aggression and Violent Behavior*, No. 23, pp. 1–21.

Paper 3

The Cultural Distance of Parental Country of Birth and the Second-Generation Immigrants' Wage Assimilation: The Case of Sweden

Mauricio G. Valencia-Amaya

Paper 3. The Cultural Distance of Parental Country of Birth and the Second-Generation Immigrants' Wage Assimilation: The Case of Sweden

Mauricio G. Valencia-Amaya

Abstract

This paper analyzes the effect on wages of the cultural distance, in terms of work-related values, between Sweden and the country of birth of the parents of the second-generation immigrants. It uses a unique Swedish longitudinal database of the population born in 1974 and their labor market outcomes between 1996 and 2006. As a proxy of the work-related values of a country, this paper proposes a cultural distance index based on the six Geert Hofstede cultural dimensions. Compared with natives with parents of non-immigrant origin, the results indicate that second-generation immigrants earn less when having a culturally-distant background. Being a woman, working in the public sector, having a higher education degree, having a Swedish mother or father, time, or having a culturally-closer background weakens or even eliminates this negative impact.

“Culture matters for economic behavior” (Blau, 2015)

I. Introduction

This paper studies the effect of the cultural distance of parents' country of origin, in terms of work-related values, on the wage assimilation of the second-generation immigrants. It uses a unique longitudinal dataset of the population born in 1974, who were living in Sweden between 1996 and 2006. A central question in the immigration literature is the economic and cultural assimilation of second-generation immigrants. The source of the human capital—whether it was acquired locally or not—is considered one of the most important factors explaining the earnings gap between immigrants and natives (Friedberg, 2000). The children of immigrants born in the receptor country are supposed to acquire country-specific skills, such as proficiency of the native language, cultural and social skills (country-specific knowledge), local labor market experience, connections with natives, and local education, all of which might improve their economic performance and close the gap in labor market outcomes, vis-à-vis natives.

However, different empirical studies show mixed results, in terms of the assimilation degree of the second generation of immigrants. In some cases, the performance of the local-born children of immigrants still lags that of natives (Abramitzky et al., 2014; Riphahn, 2003), which means that the initial wage gap of the first generation persists into the second (Abramitzky et al., 2014). Possible reasons include a lack of educational assimilation—school attendance and degree completion—of the second generation compared with their native peers (Riphahn, 2003) and/or discrimination practices in the labor market (Midtbøen, 2016). In some other cases, the performance of the second-generation immigrants converges towards native levels in both cultural and economic outcomes, such as labor supply, fertility, and education attainment, narrowing the gaps with the time spent in the country (Blau et al., 2013; Blau, 2015). Yet, other studies argue for a mixed experience, also known as segmented assimilation, where a part of the second-generation immigrants converges and surpasses the labor and education outcomes of the first generation and the other lags, especially those from unprivileged backgrounds, in terms of their

incorporation mode, parents' human capital, and family type, (Portes et al., 2005) or those whose parents were born in specific countries (Algan et al., 2010; Tran and Valdez, 2017).

Through their decisions, education or socioeconomic background, parents seem to play an important role in the success or failure stories of the second-generation immigrants. For example, there is historical evidence for the second-generation immigrants in the U.S. of performing better in the labor market and closing the gap with natives at a faster pace, when their parents opted for local names instead of foreign names (Abramitzky et al., 2016). Research studies for The Netherlands (Van Ours et al., 2003) and Switzerland (Bauer and Riphahn, 2007) have found that conditional on the education of their parents most of the second generation attain similar academic outcomes compared to natives, showing that their educational attainment is correlated with their parents' educational levels. Moreover, in the Dutch case, the social origins of the immigrants —coming from poor families, for instance—, together with the differences in human capital variables, seem to explain the so-called “ethnic penalty” of the second-generation in the labor market (Gracia et al., 2016). Even mothers' fluency in the local language seems to have an important role in the cultural assimilation process of the second-generation, at least in the case of the Turkish immigrants across Western Europe (Ali and Fokkema, 2015).

Most important, parents transmit their values, norms, preferences, and beliefs to their children. This cultural transmission across and within generations gears social learning that helps preserve group coherence, incentives cooperation and leads to both cultural continuity and cultural change (Phalet and Schönplflug, 2001b; Schönplflug, 2001; Bisin and Verdier, 2008). Furthermore, cultural transmission influences everyday behavior, supports major life decisions, shapes preferences and attitudes, and affects job performance (Schönplflug, 2001; Bisin and Verdier, 2008). First-generation immigrant parents have then an important role in closing or widening the socioeconomic gaps of their children, especially when it is possible that their families develop bilinear cultural orientations (Nauck, 2001b). They not only influence the second-generation immigrants' language retention/acquisition, their career paths, and their social and identification assimilation (Nauck, 2001), but also pass on their values to children at an early stage in life, and these values tend to persist over time, especially the religious ones (Min, Silverstein, and Lendon, 2012). Parents also pass along their cultural capital, which is found to be of importance in the acculturation process of immigrant families (Nauck, 2001). The cultural transmission weights on the socioeconomic investment decisions of the second-generation, as their adaptation depends on the acquisition of the values of achievement and collectivism from the first-generation (Phalet and Schönplflug, 2001).

The failure of the ‘melting pot’ prediction, the endurance of ‘ethnic capital’, and the attachment of many immigrant groups to their cultural traits (Bisin and Verdier, 2008) lead to the question of what happens when the culture, in terms of work-related values, transmitted from the first- to the second-generation immigrants mismatch the expected set of values of the host country? Is there a cost for the second-generation in terms of labor market outcomes for carrying behaviors inherited from their parents' culture vis-à-vis the host country culture? This paper uses the case of Sweden to address these questions. There is a lack of research on the relationship between the divergence of the work-related values transmitted to the second-generation from the values of the host country —cultural distance— and their assimilation in the labor market. One of the few to deal with this subject is Lundborg (2013), who studies the labor market integration of refugee immigrants in Sweden. The author finds that refugees from culturally-distant countries remain unemployed more days than those from less-distant countries. However, the study never defines what “culturally-distant” / “less-distant” countries are and does not deal with the labor market performance of the second generation. This paper defines cultural distance and offers a

methodology to assess it. It also contributes to the literature related to the socioeconomic assimilation of the second-generation immigrants and sheds light in the understanding of the factors behind their wage differentials with respect to natives, which is important for the design of immigration policies aimed at facilitating the assimilation process of immigrants and their future generations.

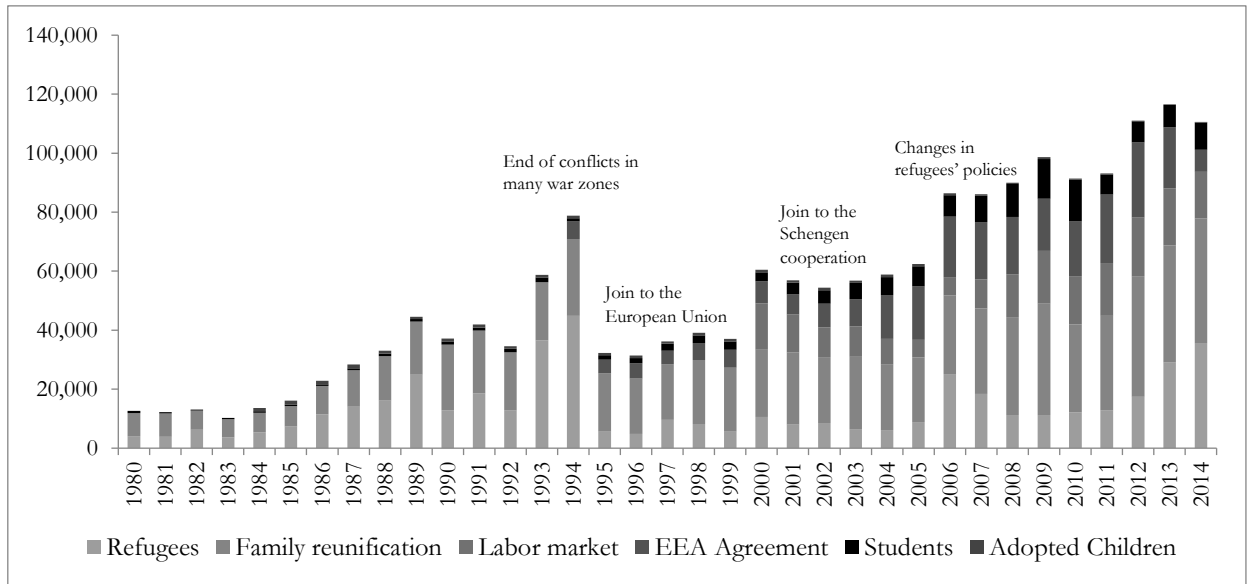
The results show evidence that the cultural background of the second-generation immigrants carries a cost in terms of lower wages in the Swedish labor market. The cost is higher the more culturally-distant the country of origin of the parents is. Factors such as being a female, having a job in the public sector, being highly educated, having a Swedish mother/father, time, or having a culturally-closer background diminish or even eliminate this negative impact. This paper is structured as follows: Section 2 describes the immigration patterns in Sweden since the Second World War and reviews the literature on the assimilation of the second-generation immigrants in Sweden; Section 3 presents the methodology; Section 4 shows the main results; and Section 5 discusses these results and concludes.

II. Immigration Patterns and the Second-Generation in Sweden: An Overview

A. A Brief History of Recent Immigration in Sweden

After being a net exporter of migrants for a long time, during the World War II Sweden became a net importer, as refugees, mostly from neighbor countries, started to arrive to the country. After the Second World War, people continued migrating to Sweden, but this time for a reason other than war: labor. During the nineteen sixties, both organized and spontaneous workers from places such as Turkey, Yugoslavia, Italy, and Greece swelled the Swedish labor force and sowed the seeds for future migration patterns. However, by the end of that decade, the country began to regulate the entrance of new labor immigrants according to the current needs of the Swedish market. The only exceptions in this norm were: (1) Nordic citizens, (2) refugees, and (3) close relatives of earlier migrants. Those exceptions became, then, the new engines behind the immigration flows in the period afterwards. During the nineteen eighties, there was an increase in the number of asylum seekers, especially from the Middle East, Somalia, and Eastern Europe. The first years of the nineteen nineties brought geopolitical changes that ended conflicts in many war zones around the globe (except for the former Yugoslavia) and so the number of asylum seekers decreased drastically, particularly during the second half of the decade. The join of Sweden to the European Union in 1995 and to the Schengen cooperation in 2001 brought about an increase in the number of immigrants from European countries. In 1997 the country hardened the policy of family reunification. In 2005-2006 the Swedish government reassessed refused entry cases of asylum seekers and changed the legislation towards unaccompanied minors. These policies lead to an increase in the number of permits granted to refugees. Another important policy change was introduced in 2011, when universities started to charge a fee to students from outside the European Union (Migrationsverket, 2015). Graph 1 shows the number of residence permits granted by the Swedish government between 1980 and 2014.

Graph 1. Sweden: Residence Permits Granted, 1980-2014.



Source: Migrationsverket (2015); own calculations.

B. The Second-generation Immigrants in Sweden

In general, the degree of integration of immigrants and therefore the acquirement of country-specific skills depends on the amount of time spent in the host country, but also on the age at migration. Arriving at a young age increases the chances of integrating with the locals, in terms of higher probabilities of marrying, working with, or living close to a native (Åslund, Böhlmark, and Skans, 2015). Now, regarding the second-generation immigrants in Sweden, the literature points to the persistence of wage differentials. The second-generation immigrants, vis-à-vis children of natives, have lower probabilities of finding a job in the Swedish market (Behrenz, Hammarstedt and Månsson, 2007). Even after controlling for human capital traits, having an immigrant background, compared with having a Swedish background, is related to a higher probability of being unemployed as well as with a lower annual wage income (Behtoui, 2004). Having a Swedish father is also related to better labor-market outcomes than having a Swedish mother, which might be a sign of labor-market discrimination practices (Behtoui, 2004). At the school level, the literature has also found differences in academic outcomes and choices between children of Swedes and children of immigrants. For instance, there is some evidence that the school performance of children of immigrants from non-European countries is lower than that of their counterparts from European countries (Jonsson and Rudolphi, 2011). However, among those who enroll in upper secondary education their choices are more oriented towards academic studies and less oriented towards vocational studies; whereas children of Europeans and Swedes show similar preferences for these two options (Jonsson and Rudolphi, 2011).

III. Methodology

A. The Geert Hofstede Cultural Dimensions

This paper uses the Geert Hofstede cultural dimensions as a proxy of the work-related values of a country (Hofstede, 1984). These cultural dimensions stemmed initially from two global studies on employee values conducted by IBM between 1967 and 1973. The survey collected close to 117,000 questionnaires, comprising 88,000 different local respondents from 71 different countries where the company had

subsidiaries at the time (Minkow, 2013). Out of the 38 occupations surveyed, seven existed in all countries. Hofstede used the information from this subsample for his seminal research, which included four dimensions: power distance, individualism versus collectivism, masculinity versus femininity, and uncertainty avoidance. Subsequent studies, using updated versions of the questionnaire, widened the number of countries and allowed the inclusion of two additional dimensions: long term orientation versus short term normative orientation and indulgence versus restraint (Grove, 2015; Minkow, 2013). Hofstede's theory has been applied in different academic fields, such as international business, international economics, cross-cultural psychology, international management, ethics, and structure of language, among others (Minkow, 2013). The cultural dimensions might help understand the influence of values on the behavior of people in the work place; they also provide a framework to group and differentiate cultures (McNett et al., 2005). This paper uses the 2010 version of Hofstede's cultural dimensions. The Appendix presents the questionnaire used to build these dimensions.

The six Geert Hofstede cultural dimensions are:

Power Distance Index (PDI): It shows the degree to which a society accepts a hierarchical structure—higher values— versus a more equal distribution of power—lower values—. It measures the ability of a culture to accept power differences within an organization (McNett et al., 2005) and represents inequality defined from below (Moskowitz, 2009). Societies with smaller values in the index are more prone to minimize inequalities, to stand up for equal treatment of group members, and to reach decision in a democratically way (Jackson and Hogg, 2010c). On the contrary, societies with higher values in the PDI expect groups to be unequal, with stiff hierarchies, which assign power to its members based on age, sex, or position of authority, for instance (Jackson and Hogg, 2010c). Higher values in the index are highly correlated with countries with lower social mobility, more rigid social classes, and higher income inequality (Jackson and Hogg, 2010a).

Individualism versus Collectivism (IDV): It denotes the preferences for a society in which individuals depend less on others—higher values—, as opposed to a more cohesive society in which individuals are expected to care not only for themselves but for others (relatives or members of a common group)— lower values—. This index measures the degree to which a society values individual success over group welfare (McNett et al., 2005). Higher values indicate that the country considers success as a personal achievement and disregards any collective contributions to it; whereas, lower values point to societies in which collective well-being and success steps over any individual selfishness (McNett et al., 2005). Individualist societies value independence and stress freedom of choice, self-sufficiency, individual goals, and personal responsibility (Jackson and Hogg, 2010c). Collectivist societies, on the other hand, emphasize common goals and value harmony between the members of a group, reciprocity, obedience, and dependence (Jackson and Hogg, 2010c).

Masculinity versus Femininity (MAS): It relates to a competitive society with preferences for material success, achievement and confidence—higher values— versus a more consensus- and cooperative-oriented society with preferences for quality of life and caring for the less privileged—lower values—. It answers the question of how countries manage the duality of sexes (Rodriguez, 2009). Masculine societies stress assertive values, tend to classify jobs based on gender, and await women to be away from the labor force; feminine societies, on the contrary, stress supportive values, put less emphasis on performance, and expect women to participate in the labor market (McNett et al., 2005). In feminine countries men and women tend to share the same values of modesty and caring; whereas in masculine countries,

although considered a shared value, men prioritize competitiveness more than women (Moskowitz, 2009).

Uncertainty Avoidance Index (UAI): It deals with the degree to which societies accept uncertainty. Higher values denote intolerance towards unorthodox ideas; whereas lower values denote a more open attitude towards different codes of conduct. The legal system, technology, and religion of a country determines how people deal with and tolerate uncertainty (Jackson and Hogg, 2010c), as well as their level of comfort in novel situations (Moskowitz, 2009). This index considers three variables: the willingness to break rules, the predilection for employment stability, and the level of anxiety of a society (McNett et al., 2005). Higher values denote societies with stricter regulations, long-term employment opportunities, and a greater need for stability, all of which translate into fewer decisions made by the individuals, fear of difference, and resistance to change (McNett et al., 2005; Jackson and Hogg, 2010c; Jackson and Hogg, 2010d; Ramoglou, 2009). Lower values, on the other hand, represent societies with fewer regulations, more tolerance to work change, and the sense that people feel less threatened by uncertainty; individuals in this societies make more decisions and adopt innovative ideas more easily (McNett et al., 2005; Jackson and Hogg, 2010c; Jackson and Hogg, 2010d).

Long-Term Orientation versus Short-Term Normative Orientation (LTO): It refers to the predisposition of societies to invest in the present to achieve long-term goals —higher values— versus societies that encourage the preservation of traditions and established norms —lower values—. Long-term oriented societies value persistence and frugality, expecting these values to bear fruit in the future, and adapt easily to changing scenarios; while, short-term oriented societies value respect for traditions and attach to values linked to the past and/or present that do not promote adaptability to change (McNett et al., 2005; Grove, 2015).

Indulgence versus Restraint (IND): It stands for the degree of permissiveness of a society towards the gratification of human needs. Higher values denote more freedom to enjoy life, whereas lower values denote more repression and stricter social norms. Indulgent societies values leisure more than hard work and thrift and tolerate less deviations from established norms; whereas restricted societies put more weigh on hard work and thrift than on leisure time and are more relaxed towards deviations from norms (Minkow, 2013). Higher values of this index denote countries that allow free gratification, and their citizens feel they have control over their lives; lower values refer to countries that control human impulses through stricter social norms, and their citizens sense that self-gratification should be suppressed or limited (Bareiss, 2014; Grove, 2015).

The range of the score of each dimension runs from 0 to 100.

B. Cultural Distance

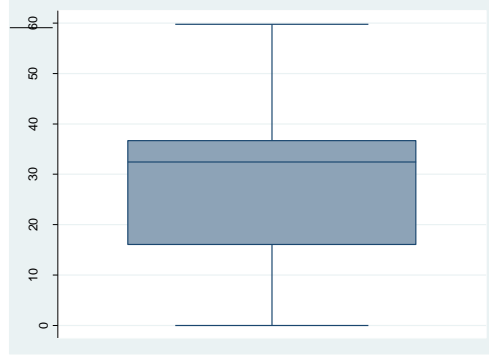
This paper introduces a Cultural Distance (CD) measure based on the Geert Hofstede dimensions. It calculates the average cultural distance between Sweden and any other country. The indicator is constructed as follows:

$$CD_i = \frac{1}{2} \left\{ \left[\frac{1}{6} \sum_{k=1}^6 |(GH_{k_{Sweden}} - GH_{k_j})| \right]_m + \left[\frac{1}{6} \sum_{k=1}^6 |(GH_{k_{Sweden}} - GH_{k_j})| \right]_f \right\} \quad (1),$$

where, CD_i denotes the absolute average difference of the aggregated Geert Hofstede dimensions (k) between Sweden and the country of birth (j) of individual (i)'s mother (m) and father (f). The CD value for an individual (i) whose both parents were born in Sweden is zero.

Now, based on the results of the CD , this paper constructs a Cultural Distance Index (CDI). The goal of this index is to group the second-generation immigrants, using the median of the distribution of the CD , according to the distance from Sweden of their cultural background. Graph 2 shows the box plot of CD .

Graph 2. Box Plot of the Cultural Distance (CD).



Source: Own calculations.

The continuous variable CD is then categorized into three groups, according to the following rules:

$$CDI_i = \begin{cases} (i) \text{ belongs to Group 1 if } CD_i = 0 \text{ (Swedes with both parents born in Sweden)} \\ (i) \text{ belongs to Group 2 if } 0 < CD_i \leq \mu_{1/2}(CD) \text{ (individuals with a culturally-closer background)} \\ (i) \text{ belongs to Group 3 if } CD_i > \mu_{1/2}(CD) \text{ (individuals with a culturally-distant background)} \end{cases}$$

where $\mu_{1/2}(CD)$ is the median value of the distribution of CD . Group 1 contains all individuals born in Sweden whose parents were also born in the country. Now, based on the average distant of the countries of birth of their parents with respect to Sweden, groups 2 and 3 comprise individuals born in Sweden, but with a culturally-closer background (Group 2) or with a culturally-distant background (Group 3). The number of individuals in each group is: 91,326 (Group 1); 14,906 (Group 2); 4,325 (Group 3).

Theoretically, individuals with a background culturally-closer to the Swedish culture might differ in their labor market outcomes from individuals with a culturally-distant background. Then, categorizing the variable CD allows this paper to compare the performance over time of these two groups between them and vis-à-vis Swedes with a non-foreign background.

C. Identification Strategy

The basic empirical specification is given by equation (2):

$$\begin{aligned} \text{Log}(rWage_{it}) = & \beta_0 + \beta_1 \text{Cultural_Distance}_i + \beta_2 \mathbf{X}_{it} + \beta_3 \text{Log}(rWage_{i(t-1)}) + \theta_t + \theta_m + \theta_n \\ & + \varepsilon_{it} \end{aligned} \quad (2),$$

where, $rWage$ represents the real wage; i and t denote, respectively, individual and year indices; $\text{Cultural_Distance}_i$ indicates either the Cultural Distance CD_i of equation (1) or the Cultural Distance Index CDI_i , as defined in the previous section; \mathbf{X} is a vector of individual's characteristics; θ_t , θ_m , and

θ_n refer to year, municipality, and industry fixed effects. This model specification controls then for the influence of aggregate trends (year fixed effects), as well as for the unobserved heterogeneity at both the municipality (municipality fixed effects) and the industry (industry fixed effects) levels.

The coefficient of interest is β_1 , which estimates the effect on average real wages of the cultural distance inherited from the parents. It measures the impact for the second-generation immigrants of having a culturally-distant background, compared to natives whose both parents were born in Sweden. According to the literature reviewed in the introduction section, β_1 is expected to have a negative sign; that is, the greater the distance between the work-related values inherited by the second-generation immigrants and those prevalent in the host country, the greater the cost in terms of average wages. To identify the effect of *Cultural_Distance_i* on the average real wages, this paper takes advantage of both the structure of the data —longitudinal information— and the availability of administrative records on previous wages. By controlling for previous wages, for example, this study accounts for the market valuation of non-observable characteristics, such as innate abilities and noncognitive skills.

D. Data

This paper uses the Longitudinal Database on Education, Income, and Employment (LOUISE) of the National Bureau of Statistics of Sweden. This database includes registered records from 1990 onwards on all the Swedish population aged 16 and older. The information is collected from different institutions, such as the Register-based Labor Market Statistics (RAMS), education registers, Total Population Register (RTB), Income and Property Register (IoF), the Population and Housing Census, and the AMS Applicant Records (Händel). The information in the database is updated every year. Particularly, this paper uses the records of the population born in 1974 and their labor outcomes between 1996 and 2006, when they were between 22 and 32 years old. The database comprises 133,011 individuals, of which 106,453 are Swedes and 26,558 are immigrants.

The dependent variable real wage (*rWage*) from Equation (1) was built using the total gross income during the year, according to the tax authorities registers from the employers (in hundred kronor). The variable of interest, cultural distance (*CDI*), was built using the individual's *country of birth*, the *mother's country of birth*, and the *father's country of birth*. Additionally, the following variables from the database were used as controls: *residential coordinates* (used to build an internal migration dummy), *gender* (dummy for male), *child 0_3* (dummy for having children between 0 and 3 years old in the household), *child 4_6* (dummy for having children between 4 and 6 years old in the household), *child 7_10* (dummy for having children between 7 and 10 years old in the household), *child 11_15* (dummy for having children between 11 and 15 years old in the household), *child 16_17* (dummy for having children between 16 and 17 years old in the household), *child 18plus* (dummy for having children older than 18 years old in the household), *sector code* (sectoral classification, used to build a private sector dummy), *firm branch* (firm branch of activity, according to the first two digits of the Swedish Standard Industrial Classification, SNI2002), *education level* (education, highest completed, used to build dummies for seven education levels: pre-school education; pre-primary education; secondary education; less than two years of post-secondary education; two years or longer of post-secondary education; graduate education; and unknown), *academic discipline* (field of study, used to build dummies for ten academic disciplines: general education; pedagogy and teacher training; humanities and art; social science, law, trade, and administration; natural sciences, mathematics, and data; technology and manufacturing; agricultural, forestry, and animal health care; healthcare and social care; services; and unknown); and *Germanic language background* (dummy for having at least one parent born in a country whose language has Germanic roots), which is included following one of the

main conclusions in Dustmann, Machin, and Schönberg (2010): language spoken at home matters the most for the academic achievement of ethnic minority pupils in England vis-à-vis white British students. Table 1 presents the summary statistics of the dataset.

Table 1. Summary Statistics (Mean Values by Group)

Variable	Group 1: <i>Swedes with both parents born in Sweden</i>	Group 2: <i>Swedes with a culturally-closer background</i>	Group 3: <i>Swedes with a culturally-distant background</i>
Number of Individuals	91,326	14,906	4,325
Log of Real Wage	5.28	4.80	3.66
Male (Dummy)	0.51	0.51	0.52
Internal Migration (Dummy)	0.18	0.18	0.16
Children between 0 and 3 years old dwell in the household (Dummy)	0.22	0.21	0.23
Children between 4 and 6 years old dwell in the household (Dummy)	0.09	0.09	0.11
Children between 7 and 10 years old dwell in the household (Dummy)	0.05	0.06	0.08
Children between 11 and 15 years old dwell in the household (Dummy)	0.02	0.03	0.06
Children between 16 and 17 years old dwell in the household (Dummy)	0.01	0.01	0.04
Children older than 18 years old dwell in the household (Dummy)	0.14	0.16	0.27
Private Sector (Dummy)	0.65	0.68	0.72
Education Level 1: Pre-School (Dummy)	0.00	0.00	0.02
Education Level 2: Pre-Primary (Dummy)	0.08	0.12	0.13
Education Level 3: Secondary (Dummy)	0.54	0.55	0.57
Education Level 4: Less Than Two Years of Post-Secondary (Dummy)	0.12	0.11	0.10
Education Level 5: Two Years or Longer of Post-Secondary (Dummy)	0.25	0.21	0.17
Education Level 6: Graduate (Dummy)	0.00	0.00	0.00
Education Level 7: Unknown (Dummy)	0.00	0.00	0.01
Academic Discipline 1: General Education (Dummy)	0.17	0.22	0.30
Academic Discipline 2: Pedagogy and Teacher Training (Dummy)	0.06	0.04	0.02
Academic Discipline 3: Humanities and Art (Dummy)	0.05	0.06	0.05
Academic Discipline 4: Social Science, Law, Trade, and Administration (Dummy)	0.19	0.18	0.19
Academic Discipline 5: Natural Sciences, Mathematics, and Data (Dummy)	0.04	0.04	0.04
Academic Discipline 6: Technology and Manufacturing (Dummy)	0.25	0.23	0.19
Academic Discipline 7: Agricultural, Forestry, and Animal Health Care (Dummy)	0.02	0.01	0.00
Academic Discipline 8: Healthcare and Social Care (Dummy)	0.11	0.09	0.09
Academic Discipline 9: Services (Dummy)	0.05	0.05	0.04
Academic Discipline 10: Unknown (Dummy)	0.07	0.07	0.07
Germanic Language Background (Dummy)	0.00	0.75	0.04

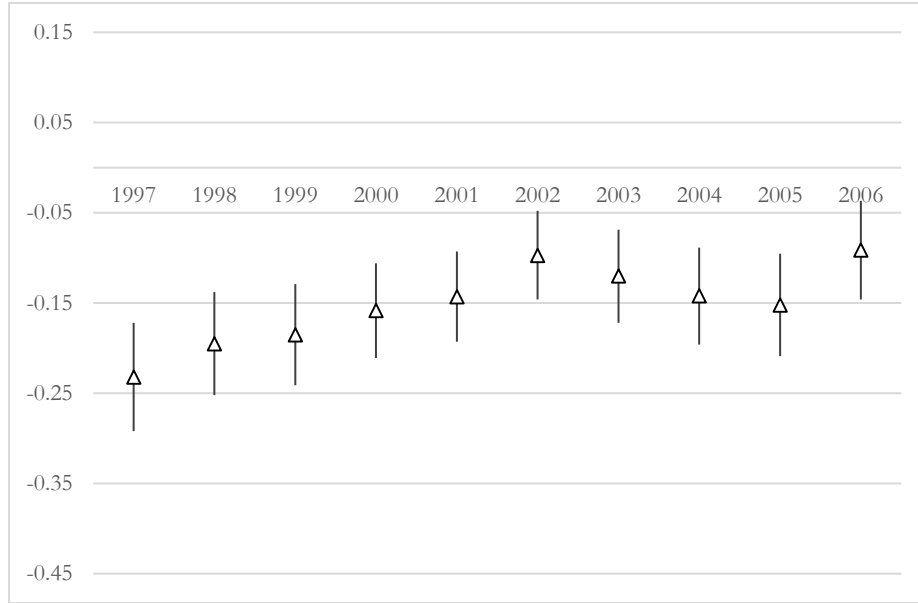
Source: LOUISE; own calculations.

IV. Empirical Results

A. Baseline

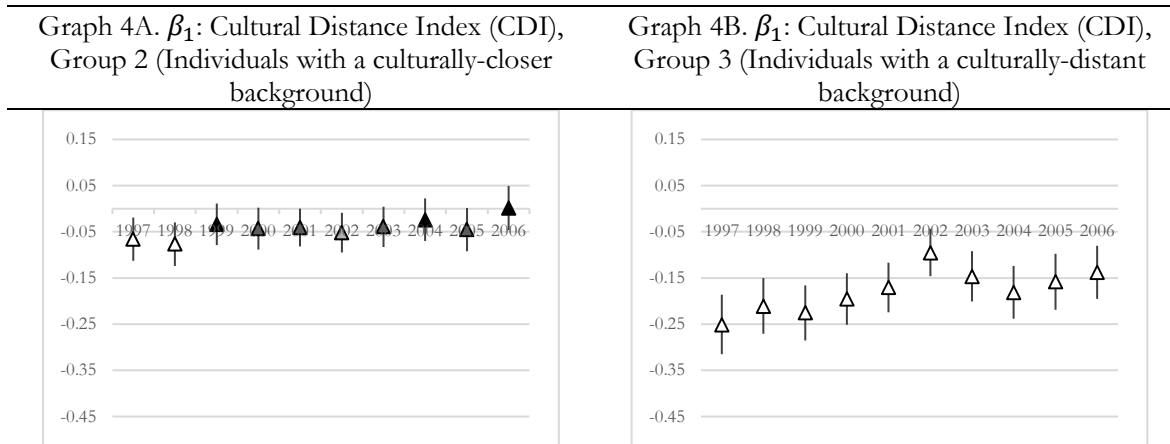
Even after being born in Sweden—which implies having acquired country-specific skills and having a Swedish education—there is a significant negative effect on the real wages of the second-generation resulting from having a foreign background, but this penalty decreases over time. This is shown in Graph 3. The results indicate also that the negative effect is bigger, the greater the cultural distance between Sweden and the country of birth of parents of the second-generation immigrants (see Graph 4). Individuals with a culturally-closer background (Graph 4A) earn on average 6.38 percent less than their local counterparts at the beginning of the period. This difference is statistically significant only for the years 1997 and 1998 (at a 0.01 level). For the latter part of the period, the difference decreases and loses significance. In contrast, individuals with a culturally-distant background (Graph 4B) earn on average 22.20 percent less in 1997, but even though this difference decreases towards 12.89 percent in 2006, it remains statistically significant at a 0.01 level during the whole period. To get an idea of the magnitude, the negative effect of being female is 9.48 percent in 1997 and 16.18 percent in 2006.

Graph 3. Cultural Distance Background (CD_i) and Real Wages, 1997-2006.



Notes: (a) Triangles represent the OLS estimate of β_1 of Equation (2), using the variable CD_i of Equation (1) as the definition of $Cultural_Distance_i$. (b) The color of each triangle indicates the significance level of the estimated coefficient, based on robust standard errors, as follows: White for $p\text{-value} < 0.01$, light grey for $p\text{-value} < 0.05$, dark grey for $p\text{-value} < 0.1$, and black for $p\text{-value} \geq 0.1$. (c) The high of each bar indicates the 95% confidence interval. (d) Controls include the real wage of the previous year, industry fixed effects, electoral districts/municipalities subdivisions fixed effects (aka SAMS: Small Areas for Market Statistics), and dummy variables for: male, education levels (pre-school education; pre-primary education; secondary education; less than two years of post-secondary education; two years or longer of post-secondary education; graduate education; and unknown), academic disciplines (general education; pedagogy and teacher training; humanities and art; social science, law, trade, and administration; natural sciences, mathematics, and data; technology and manufacturing; agricultural, forestry, and animal health care; healthcare and social care; services; and unknown), working in the private sector, internal migration, outmigration, the presence of children by age range living in the household (0-3, 4-6, 7-10, 11-15, 16-17, 18+), and having a Germanic language background. (e) Equation (2) is corrected for selection bias using the Heckman model.

Graph 4. Cultural Distance Background (CDI_i) and Real Wages, 1997-2006.

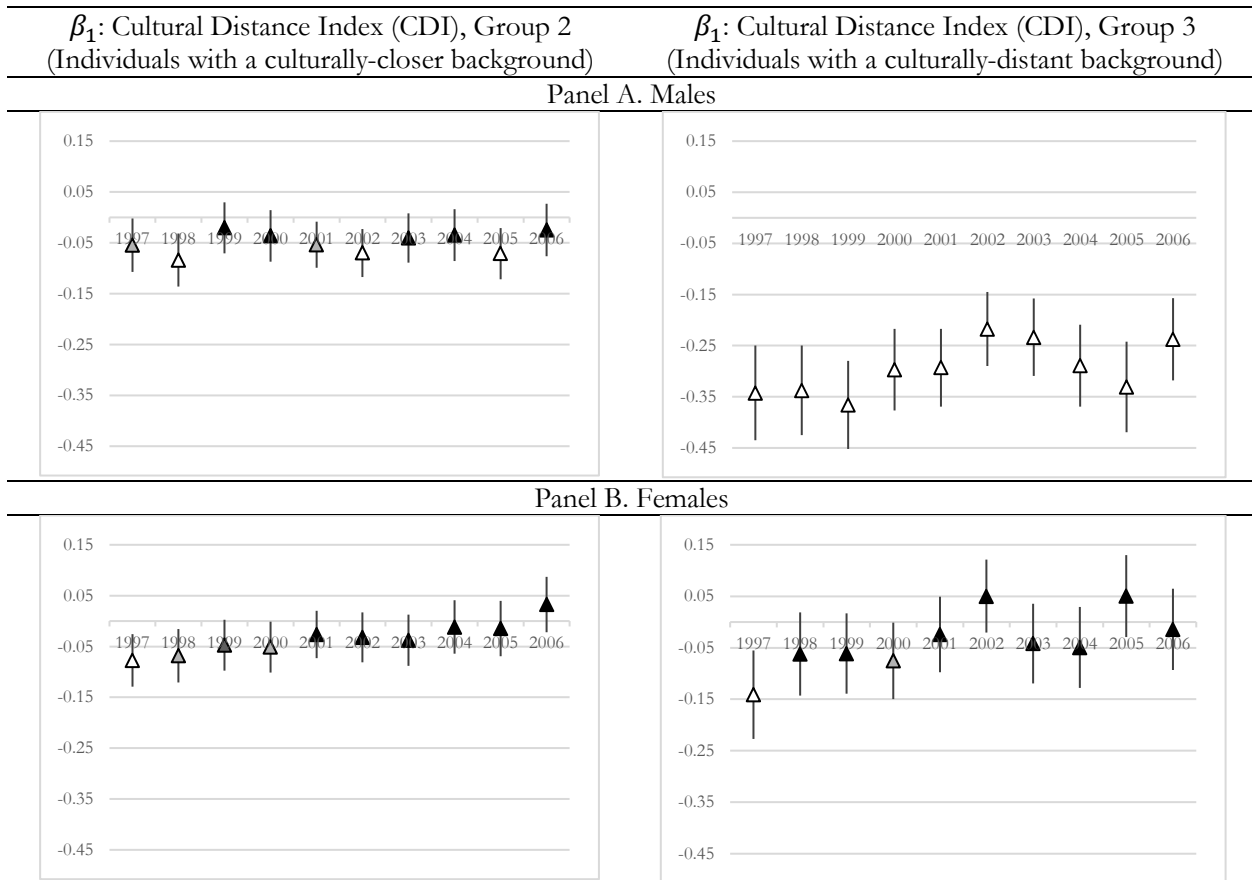


Notes: (a) Triangles represent the OLS estimate of β_1 of Equation (2), using the variable CDI_i as the definition of $Cultural_Distance_i$. (b) The variable CDI_i takes three possible values: 1 for Swedes whose parents were born in Sweden (this is the reference group), 2 for Swedes with a culturally-closer background (based on the country of birth of their parents), and 3 for Swedes with a culturally-distant background (based on the country of birth of their parents). (c) The color of each triangle indicates the significance level of the estimated coefficient, based on robust standard errors, as follows: White for $p\text{-value} < 0.01$, light grey for $p\text{-value} < 0.05$, dark grey for $p\text{-value} < 0.1$, and black for $p\text{-value} \geq 0.1$. (d) The high of each bar indicates the 95% confidence interval. (e) Controls include the real wage of the previous year, industry fixed effects, electoral districts/municipalities subdivisions fixed effects (aka SAMS: Small Areas for Market Statistics), and dummy variables for: male, education levels (pre-school education; pre-primary education; secondary education; less than two years of post-secondary education; two years or longer of post-secondary education; graduate education; and unknown), academic disciplines (general education; pedagogy and teacher training; humanities and art; social science, law, trade, and administration; natural sciences, mathematics, and data; technology and manufacturing; agricultural, forestry, and animal health care; healthcare and social care; services; and unknown), working in the private sector, internal migration, outmigration, the presence of children by age range living in the household (0-3, 4-6, 7-10, 11-15, 16-17, 18+), and having a Germanic language background. (f) Equation (2) is corrected for selection bias using the Heckman model.

B. Heterogeneous Effects

Graph 5 presents the estimation results of Equation (2) by sex. The results show that for the second-generation immigrants, the cost of having foreign parents is higher for males with a culturally-distant background. This group earns on average 29.04 percent less than their native counterparts in 1997. The wage gap for this group reduces to 21.18 percent in 2006. Males with a culturally-closer background performs better. Their wage gap with respect to Swedish males whose parents were born in Sweden is 5.32 percent in 1997 and a non-significant 2.47 percent in 2006. In the case of females, for both groups, the wage difference is significant only at the beginning of the period (at a 0.01 level), when the penalty was 7.46 percent for the culturally-closer group and 13.15 for the culturally-distant one. Since 2001 there are not significant wage differences between the second-generation immigrant females of any background and Swede females of local origin.

Graph 5. Cultural Distance Background and Real Wages: 1997-2006, by Sex.

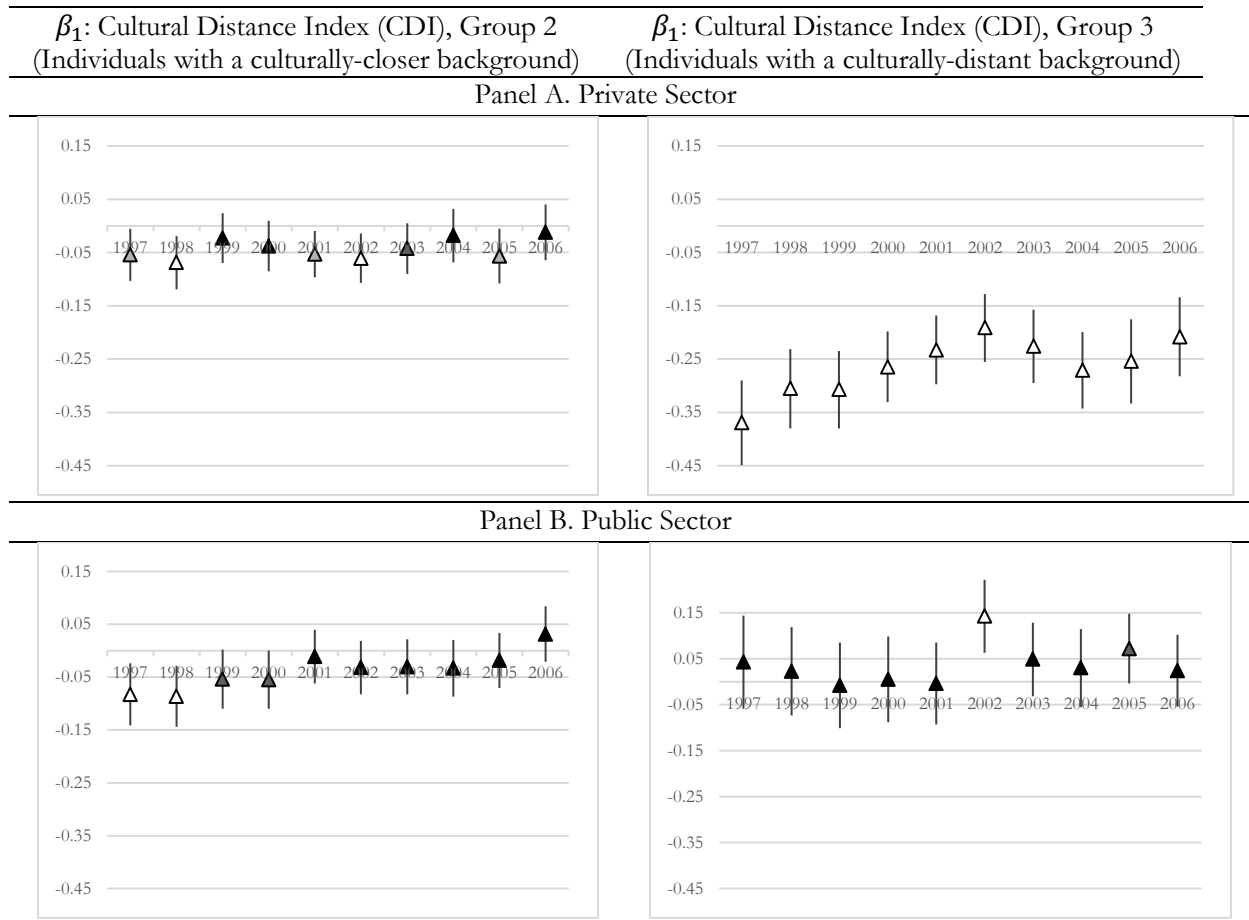


Notes: (a) Triangles represent the OLS estimate of β_1 of Equation (2), using the variable CDI_i as the definition of $Cultural_Distance_i$. (b) The variable CDI_i takes three possible values: 1 for Swedes whose parents were born in Sweden (this is the reference group), 2 for Swedes with a culturally-closer background (based on the country of birth of their parents), and 3 for Swedes with a culturally-distant background (based on the country of birth of their parents). (c) The color of each triangle indicates the significance level of the estimated coefficient, based on robust standard errors, as follows: White for p-value<0.01, light grey for p-value<0.05, dark grey for p-value<0.1, and black for p-value \geq 0.1. (d) The high of each bar indicates the 95% confidence interval. (e) Controls include the real wage of the previous year, industry fixed effects, electoral districts/municipalities subdivisions fixed effects (aka SAMS: Small Areas for Market Statistics), and dummy variables for: male, education levels (pre-school education; pre-primary education; secondary education; less than two years of post-secondary education; two years or longer of post-secondary education; graduate education; and unknown), academic disciplines (general education; pedagogy and teacher training; humanities and art; social science, law, trade, and administration; natural sciences, mathematics, and data; technology and manufacturing; agricultural, forestry, and animal health care; healthcare and social care; services; and unknown), working in the private sector, internal migration, outmigration, the presence of children by age range living in the household (0-3, 4-6, 7-10, 11-15, 16-17, 18+), and having a Germanic language background. (f) Equation (2) is corrected for selection bias using the Heckman model.

Graph 6 shows the estimation results of Equation (2) by employment sector. For the second-generation immigrants with a culturally-distant background working in the private sector is associated with a labor cost, in terms of lower wages, of 30.86 percent in 1997 and 18.78 in 2006. Having a culturally-closer background attenuates this gap. The difference decreases to 5.27 percent in 1997 and to a non-significant 1.20 in 2006. For the most part, the wages of the second-generation immigrants working in the public

sector, indistinct of the distance of their cultural background, are more like those of native Swedes of non-immigrant origin.

Graph 6. Cultural Distance Background and Real Wages: 1997-2006, by Employment Sector.

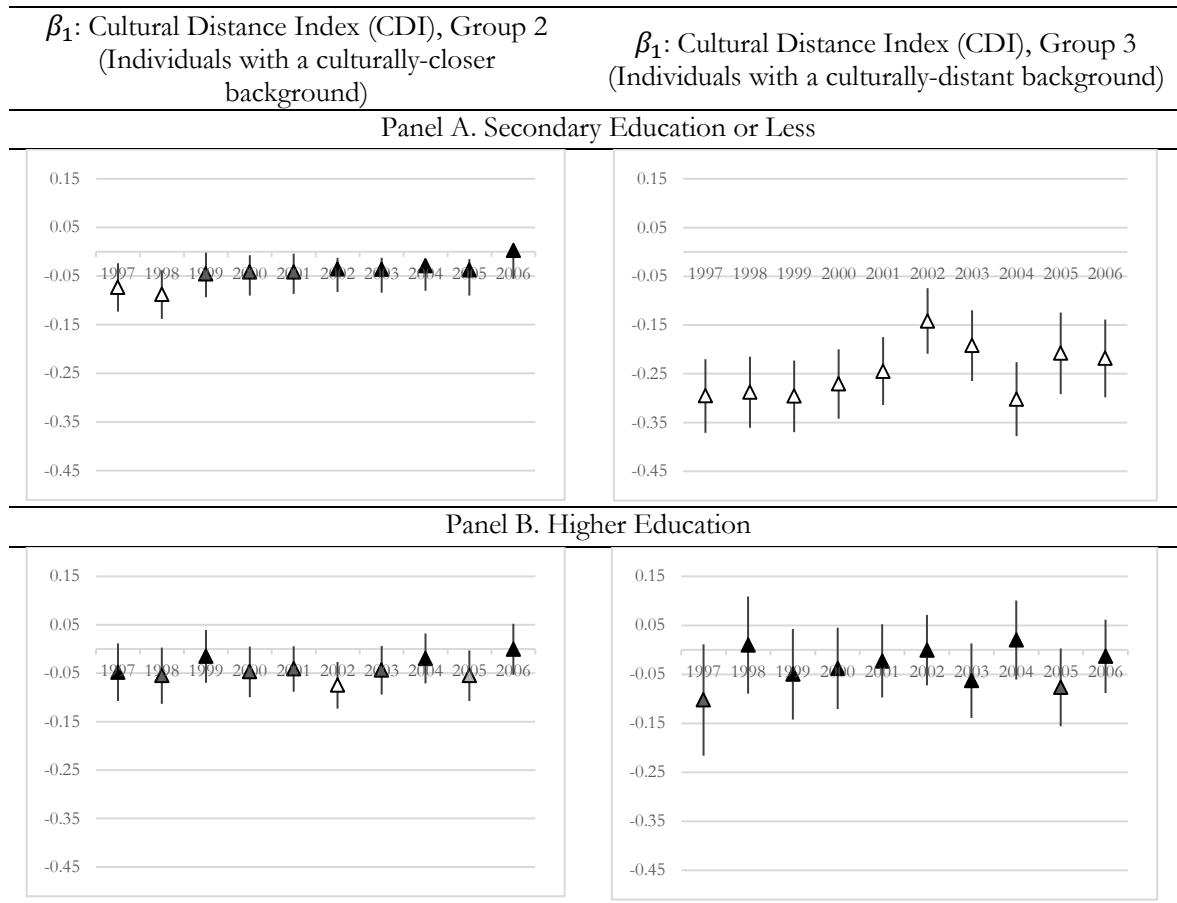


Notes: (a) Triangles represent the OLS estimate of β_1 of Equation (2), using the variable CDI_i as the definition of $Cultural_Distance_i$. (b) The variable CDI_i takes three possible values: 1 for Swedes whose parents were born in Sweden (this is the reference group), 2 for Swedes with a culturally-closer background (based on the country of birth of their parents), and 3 for Swedes with a culturally-distant background (based on the country of birth of their parents). (c) The color of each triangle indicates the significance level of the estimated coefficient, based on robust standard errors, as follows: White for p -value < 0.01 , light grey for p -value < 0.05 , dark grey for p -value < 0.1 , and black for p -value ≥ 0.1 . (d) The high of each bar indicates the 95% confidence interval. (e) Controls include the real wage of the previous year, industry fixed effects, electoral districts/municipalities subdivisions fixed effects (aka SAMS: Small Areas for Market Statistics), and dummy variables for: male, education levels (pre-school education; pre-primary education; secondary education; less than two years of post-secondary education; two years or longer of post-secondary education; graduate education; and unknown), academic disciplines (general education; pedagogy and teacher training; humanities and art; social science, law, trade, and administration; natural sciences, mathematics, and data; technology and manufacturing; agricultural, forestry, and animal health care; healthcare and social care; services; and unknown), working in the private sector, internal migration, outmigration, the presence of children by age range living in the household (0-3, 4-6, 7-10, 11-15, 16-17, 18+), and having a Germanic language background. (f) Equation (2) is corrected for selection bias using the Heckman model.

Graph 7 presents the estimation results of Equation (2) by education level. Education attenuates the labor cost of having a foreign origin. With respect to locals, achieving secondary education or less widens the

wage gap of the group with a culturally-distant background in 25.55 percent in 1997 and in 19.59 in 2006. For the group with a culturally-closer background, the wage difference is significant only in 1997 and 1998 (at a 0.01 level), with differences of 7.06 and 8.41 percent, respectively. Now, accomplishing higher education closes the wage gap of both groups with respect to natives of non-foreign origin, as shown in the Panel B of Graph 5.

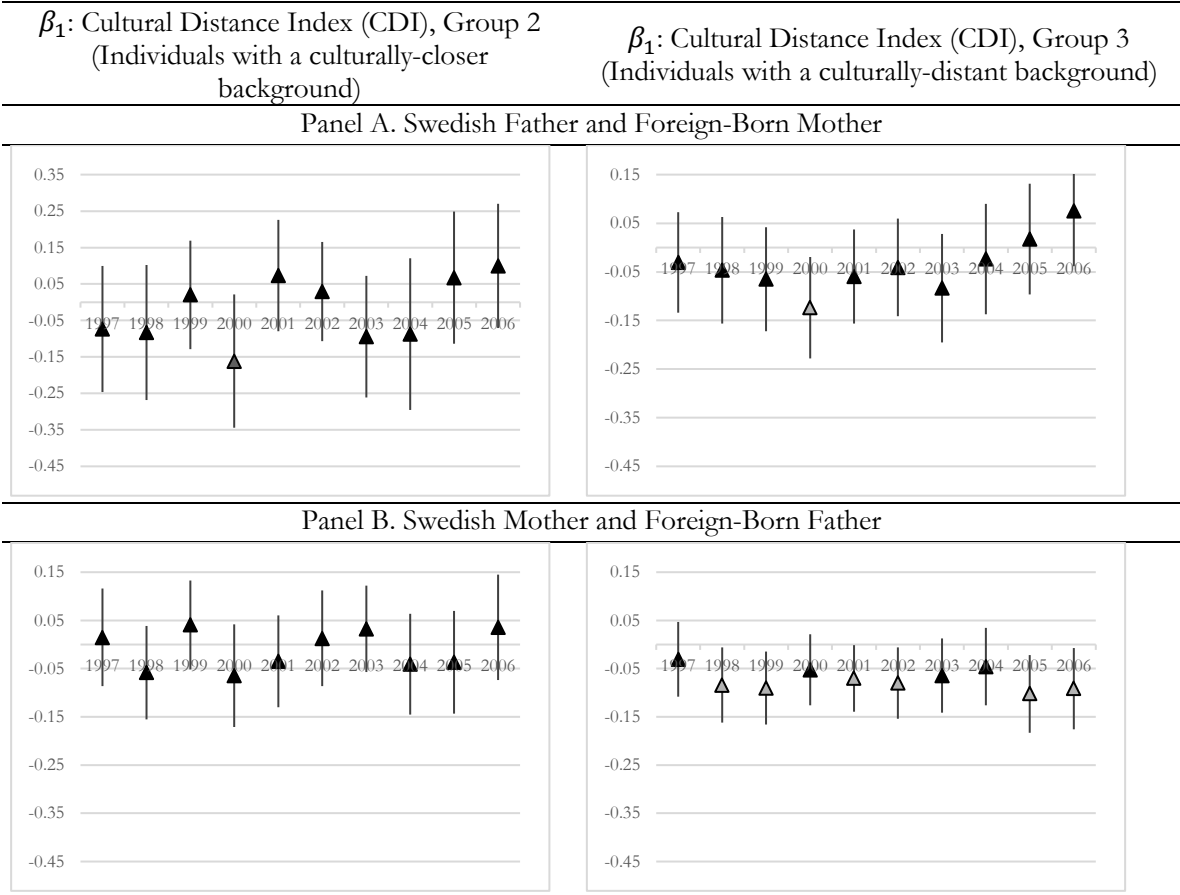
Graph 7. Cultural Distance Background and Real Wages: 1997-2006, by Education Level.



Notes: (a) Triangles represent the OLS estimate of β_1 of Equation (2), using the variable CDI_i as the definition of *Cultural Distance*_{*i*}. (b) The variable CDI_i takes three possible values: 1 for Swedes whose parents were born in Sweden (this is the reference group), 2 for Swedes with a culturally-closer background (based on the country of birth of their parents), and 3 for Swedes with a culturally-distant background (based on the country of birth of their parents). (c) The color of each triangle indicates the significance level of the estimated coefficient, based on robust standard errors, as follows: White for p-value<0.01, light grey for p-value<0.05, dark grey for p-value<0.1, and black for p-value≥0.1. (d) The high of each bar indicates the 95% confidence interval. (e) Controls include the real wage of the previous year, industry fixed effects, electoral districts/municipalities subdivisions fixed effects (aka SAMS: Small Areas for Market Statistics), and dummy variables for: male, education levels (pre-school education; pre-primary education; secondary education; less than two years of post-secondary education; two years or longer of post-secondary education; graduate education; and unknown), academic disciplines (general education; pedagogy and teacher training; humanities and art; social science, law, trade, and administration; natural sciences, mathematics, and data; technology and manufacturing; agricultural, forestry, and animal health care; healthcare and social care; services; and unknown), working in the private sector, internal migration, outmigration, the presence of children by age range living in the household (0-3, 4-6, 7-10, 11-15, 16-17, 18+), and having a Germanic language background. (f) Equation (2) is corrected for selection bias using the Heckman model.

Finally, Graph 8 presents the estimation results of Equation (2) considering mixed families, where one of the parents is Swedish and the other is foreign-born. In general, having a Swedish mother or father erases the effect on wages of having a different cultural background. For individuals with a Swedish Father and a foreign-born mother (Panel A of Graph 8), during the whole period, there are no significant differences (at a 0.01 level) between their wages of those of individuals with both parents born in Sweden. Now, having a Swedish mother and a foreign-born father (Panel B of Graph 8) closes the wage gap, but only when the foreign-born father comes from a culturally-closer country. When he comes from a culturally-distant country, the disappearance of the wage gap is less evident, as for some years there are significant wage differences (at a 0.05 level).

Graph 8. Cultural Distance Background and Real Wages: 1997-2006, Mixed Families.



Notes: (a) Triangles represent the OLS estimate of β_1 of Equation (2), using the variable CDI_i as the definition of *Cultural_Distance_i*. (b) The variable CDI_i takes three possible values: 1 for Swedes whose parents were born in Sweden (this is the reference group), 2 for Swedes with a culturally-closer background (based on the country of birth of their parents), and 3 for Swedes with a culturally-distant background (based on the country of birth of their parents). (c) The color of each triangle indicates the significance level of the estimated coefficient, based on robust standard errors, as follows: White for p-value<0.01, light grey for p-value<0.05, dark grey for p-value<0.1, and black for p-value≥0.1. (d) The high of each bar indicates the 95% confidence interval. (e) Controls include the real wage of the previous year, industry fixed effects, electoral districts/municipalities subdivisions fixed effects (aka SAMS: Small Areas for Market Statistics), and dummy variables for: male, education levels (pre-school education; pre-primary education; secondary education; less than two years of post-secondary education; two years or longer of post-secondary education; graduate education; and unknown), academic disciplines (general education; pedagogy and teacher training; humanities

and art; social science, law, trade, and administration; natural sciences, mathematics, and data; technology and manufacturing; agricultural, forestry, and animal health care; healthcare and social care; services; and unknown), working in the private sector, internal migration, outmigration, the presence of children by age range living in the household (0-3, 4-6, 7-10, 11-15, 16-17, 18+), and having a Germanic language background. (f) Equation (2) is corrected for selection bias using the Heckman model.

V. Discussion and Conclusion

This paper uses a proxy of culture based on the six cultural dimensions of Geert Hofstede, which measure the work-related values of a country. Children of first-generation immigrants are supposed to develop country-specific skills. For instance, they were born in the host country, studied and gained labor experience there, speak the local language, belong to local networks, have contact with natives, and consume local media. In this sense, based on the assumption that the source of their human capital is what explains any difference in labor market outcomes between immigrants and natives, the labor market results of the second-generation immigrants should not differ much from those of natives. Yet, the empirical evidence supports both success and failure stories and part of the literature argues that parents play a role in these outcomes. This paper explored this channel but from a new perspective: Parents pass along their cultural values to their children, which —once interiorized— will affect their behavior and influence their decisions. This paper argues that if the work-related values of the source country transmitted from the first to the second generation differ from the work-related values of the host country, the labor market will penalize this mismatch, in terms of lower wages.

Swedish culture is characterized by its moderation (*lagom*) and its “socially concerned individualism”, in which prevails not only the respectfulness of other people’s individualities —leading to a marked concern for equality—, but also the need to contribute to social welfare, reaching agreement, consensus, and avoiding conflicts (Barinaga, 1999). Not complying with this set of values might be penalized in the labor market. The results from this study suggest that the second-generation immigrants face a cost, in terms of lower wages, associated with their cultural background. What is more, this cost is bigger, the greater the cultural distance between Sweden and the country of origin of the first-generation immigrants. An average second-generation immigrant with a culturally-distant background earned 20.07 percent less in 1996 than natives with parents of non-foreign origin. This figure was 12.98 percent less in 2006. The heterogenous results show that different factors attenuate or even eliminate this penalty. One of these factors is being a female. A possible reason for this is that parents might vary the content and intensity in the transmission of values depending on the gender of their children (Phalet and Schönplflug, 2001), so certain values might be more strongly emphasized and interiorized by males than by females.

Having a higher education degree weakens also the negative effect of having a different cultural background. Once transmitted, some values are found to be a decreasing function of the level of education of the individual (Phalet and Schönplflug, 2001); therefore, higher levels of education might erode some of these inherited values. Another factor is working in the public sector. Some studies support the idea of differences in the work motivation between public and private sector employees —as in the value they place on public interests— (Aronsson, Bejerot and Härenstam, 1999; Svallfors, 2004). Also, in the case of Sweden, there are signs of more discriminatory practices in the private sector than in the public sector (Bihagen and Ohls, 2006; Ahmed, Andersson and Hammarstedt, 2013), and so it might be possible for larger work-values differences to be more penalized in the private sector. Finally, the acculturation process, defined as the cultural change that results from the influence of a dominant group (Berry, 1992), might explain the weakening over time of the negative influence of having a culturally-different background.

The results indicate too that by having a Swedish father or mother, the average wage of children of mixed couples born in Sweden is higher than that of children of foreign-born couples, also born in Sweden. There is evidence that marrying a Swedish improves the living standards of immigrants, in terms of income and employment (Dribe and Lundh, 2008), and that the social integration into the host country of children of mixed couples is greater than that of children of immigrant couples (Kalmijn, 2010). Having a Swedish father erases the wage gap between second-generation immigrants and natives. Whereas, having a Swedish mother diminishes the gap, but does not eliminate it, when the mother is born in a culturally-distant country. This finding shows that the social setting favors the predominance of the native parent's culture and that the cultural transmission from the immigrant parent might be lessened in this context.

In conclusion, the cultural distance, in terms of the work-related values identified by Geert Hofstede, between the country of origin of the parents of the second-generation immigrants and Sweden seems to negatively influence the average wage of the second generation. The greater the cultural distance, the larger the negative impact on wages. Being a woman, having a higher education degree, being a public-sector employee, or having a Swedish father or mother reduces or even cancels this impact. Time seems also to attenuate this negative effect.

References

- ABRAMITZKY, Ran; BOUSTAN, Leah Platt; ERIKSSON, Katherine (2014). "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration", *Journal of Political Economy*, Vol. 122, No. 3, pp. 467-506.
- ABRAMITZKY, Ran; BOUSTAN, Leah Platt; ERIKSSON, Katherine (2016). "Cultural Assimilation During the Age of Mass Migration", *NBER Working Paper*, No. 22381.
- AHMED, Ali; ANDERSSON, Lina; and HAMMARSTEDT, Mats (2013). "Sexual orientation and full-time monthly earnings, by public and private sector: evidence from Swedish register data", *Review of Economics of the Household*, Vol. 11, No. 1, pp. 83-108.
- ALGAN, Yann; DUSTMANN, Christian; Glitz, ALBRECHT; and MANNING, Alan (2010). "The Economic Situation of First and Second-Generation Immigrants in France, Germany and The United Kingdom", *The Economic Journal*, Vol. 120, No. 542, pp. 4-30.
- ALI, Syed and FOKKEMA, Tineke (2015). "The Importance of Peers: Assimilation Patterns among Second-generation Turkish Immigrants in Western Europe", *Journal of Ethnic and Migration Studies*, Vol. 41, No. 2, pp. 260-283.
- ARONSSON, Gunnar; BEJEROT, Eva; and HÄRENSTAM, Annika (1999). "Healthy Work: Ideal and Reality among Public and Private Employed Academics in Sweden", *Public Personnel Management*, Vol. 28 No. 2, pp. 197-215.
- ÅSLUND, Olof; BÖHLMARK, Anders; and SKANS, Oskar Nordström (2015). "Childhood and family experiences and the social integration of young migrants", *Labour Economics*, No. 35, pp. 135-144.
- BAREISS, Warren (2014). "Hofstede's Dimensions of Culture", *Encyclopedia of Health Communication*, SAGE Publications, Thousand Oaks, pp. 667-670.

BARINAGA, Ester (1999). "Swedishness through lagom Can words tell us anything about a culture?", *SSE/EFI Working Paper Series in Business Administration*, No. 1999:17.

BAUER, Philipp and RIPHAHN, Regina T. (2007). "Heterogeneity in the intergenerational transmission of educational attainment: evidence from Switzerland on natives and second-generation immigrants", *Journal of Population Economics*, Vol. 20, No. 1, pp. 121-148.

BEHRENTZ, Lars; HAMMARSTEDT, Mats; and MÅNSSON, Jonas (2007). "Second-Generation Immigrants in the Swedish Labour Market", *International Review of Applied Economics*, Vol. 21; No. 1, pp. 157-174.

BEHTOUI, Alireza (2004). "Unequal Opportunities for Young People with Immigrant Backgrounds in the Swedish Labour Market", *LABOUR: Review of Labour Economics and Industrial Relations*, Vol. 18, No. 4, pp. 633-660.

BERRY, J.W. (1992). "Acculturation and Adaptation in a New Society", *International Migration*, Vol. 30, pp. 69-85.

BIHAGEN, Erik and OHLS, Marita (2006). "The glass ceiling – where is it? Women's and men's career prospects in the private vs. the public sector in Sweden 1979–2000", *The Sociological Review*, Vol. 54, No. 1, pp. 20-47.

BISIN, Alberto and VERDIER, Thierry (2008). "Cultural Transmission", in: Durlauf, Steven N. and Blume, Lawrence E., *The New Palgrave Dictionary of Economics*, 2nd edition, Palgrave Macmillan (ed.).

BLAU, Francine (2015). "Immigrants and gender roles: assimilation vs. culture", *IZA Journal of Migration*, Vol. 4, No. 1, pp. 1-21.

BLAU, Francine; KAHN, Lawrence; LIU, Albert Yung-Hsu; and PAPPS, Kerry (2013). "The transmission of women's fertility, human capital, and work orientation across immigrant generations", *Journal of Population Economics*, Vol. 26, No. 2, pp. 405-435.

DRIBE, Martin and LUNDH, Christer (2008). "Intermarriage and Immigrant Integration in Sweden: An Exploratory Analysis", *Acta Sociologica*, Vol. 51, No. 4, pp. 329-354.

DUSTMANN, Christian; MACHIN, Stephen; and SCHÖNBERG, Uta (2010). "Ethnicity and Educational Achievement in Compulsory Schooling", *The Economic Journal*, Vol. 120, No. 546, pp. F272-F297.

FRIEDBERG, Rachel M. (2000). "You Can't Take It with You? Immigrant Assimilation and the Portability of Human Capital", *Journal of Labor Economics*, Vol. 18, No. 2, pp. 221-251.

GRACIA, Pablo; VÁZQUEZ-QUESADA, Lucía; and VAN DE WERFHORST, Herman G. (2016). "Ethnic penalties? The role of human capital and social origins in labour market outcomes of second-generation Moroccans and Turks in the Netherlands", *Journal of Ethnic and Migration Studies*, Vol. 42, No. 1, pp. 69-87.

GROVE, Cornelius N. (2015). "Value Dimensions: Hofstede", in: BENNETT, Janet M. (ed.), *The SAGE Encyclopedia of Intercultural Competence*, SAGE Publications, Thousand Oaks, pp. 842-845.

HOFSTEDE, Geert (1984). "Cultural dimensions in management and planning", *Asia Pacific Journal of Management*, Vol. 1, No. 2, pp. 81-99.

HOFSTEDE, Geert (2001). *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations*, second edition, SAGE Publications, Thousand Oaks.

JACKSON, Ronald L. II and HOGG, Michael A. (2010a). "Class Identity" in: JACKSON, Ronald L. II and HOGG, Michael A. (eds.), *Encyclopedia of Identity*, SAGE Publications, Thousand Oaks, pp. 92-94.

JACKSON, Ronald L. II and HOGG, Michael A. (2010b). "Collectivism/Individualism" in: JACKSON, Ronald L. II and HOGG, Michael A. (eds.), *Encyclopedia of Identity*, SAGE Publications, Thousand Oaks, pp. 101-102.

JACKSON, Ronald L. II and HOGG, Michael A. (2010c). "Dimensions of Cultural Variability" in: JACKSON, Ronald L. II and HOGG, Michael A. (eds.), *Encyclopedia of Identity*, SAGE Publications, Thousand Oaks, pp. 228-231.

JACKSON, Ronald L. II and HOGG, Michael A. (2010d). "Uncertainty Avoidance" in: JACKSON, Ronald L. II and HOGG, Michael A. (eds.), *Encyclopedia of Identity*, SAGE Publications, Thousand Oaks, pp. 852-851.

JONSSON, Jan O. and RUDOLPHI, Frida (2011). "Weak Performance—Strong Determination: School Achievement and Educational Choice among Children of Immigrants in Sweden", *European Sociological Review*, Vol. 27, No. 4, pp. 487–508.

KALMIJN, Matthijs (2010). "Consequences of Racial Inter-marriage for Children's Social Integration", *Sociological Perspectives*, Vol. 53, No. 2, pp. 271-286.

LUNDBORG, Per (2013). "Refugees' Employment Integration in Sweden: Cultural Distance and Labor Market Performance", *Review of International Economics*, Vol. 21, No. 2, pp. 219-232.

MCCNETT, Jeanne; LANE, Henry W.; MAZNEVSKI, Martha L.; MENDENHALL, Mark E.; and O'CONNELL, John (2005). "International Management", in: COOPER, Cary L. (ed.), *The Blackwell Encyclopedia of Management*, second edition, Blackwell Publishing.

MIGRATIONSVERKET (2015). "History of the Swedish Migration", online document, available in: <http://www.migrationsverket.se/English/About-the-Migration-Agency/Facts-and-statistics-/Facts-on-migration/History.html>, accessed on: May 28th, 2015.

MIDTBØEN, Arnfinn H. (2016). "Discrimination of the Second Generation: Evidence from a Field Experiment in Norway", *Journal of International Migration and Integration*, Vol. 17, No. 1, pp. 253-272.

MIN, Joohong; SILVERSTEIN, Merrill; and LENDON, Jessica P. (2012). "Intergenerational transmission of values over the family life course", *Advances in Life Course Research*, No. 17, pp. 112–120.

MINKOV, Michael (2013). "Cultural Dimensions Across Modern Nations", in: MINKOV, Michael, *Cross-Cultural Analysis: The Science and Art of Comparing the World's Modern Societies and Their Cultures*, SAGE Publications, Thousand Oaks, pp. 199-406.

MOSKOWITZ, Sanford (2009). "Hofstede's Five Dimensions of Culture", in: WANKEL, Charles (ed.), *Encyclopedia of Business in Today's World*, SAGE Publications, Thousand Oaks, pp. 817-818.

NAUCK, Bernhard (2001). "Intercultural Contact and Intergenerational Transmission in Immigrant Families", *Journal of Cross-Cultural Psychology*, Vol. 32, No. 2, pp. 159-173.

NAUCK, Bernhard (2001b). "Social Capital, Intergenerational Transmission and Intercultural Contact in Immigrant Families", *Journal of Comparative Family Studies*, Vol. 32, No. 4, pp. 465-488.

PHALET, Karen and SCHÖNPFLUG, Ute (2001). "Intergenerational Transmission in Turkish Immigrant Families: Parental Collectivism, Achievement Values and Gender Differences", *Journal of Comparative Family Studies*, Vol. 32, No. 4, pp. 489-504.

PHALET, Karen and SCHÖNPFLUG, Ute (2001b). "Intergenerational Transmission of Collectivism and Achievement Values in Two Acculturation Contexts: The Case of Turkish Families in Germany and Turkish and Moroccan Families in the Netherlands", *Journal of Cross-Cultural Psychology*, Vol. 32, No. 2, pp. 186-201.

PORTES, Alejandro; FERNÁNDEZ-KELLY, Patricia; HALLER, William (2005). "Segmented assimilation on the ground: The new second generation in early adulthood", *Ethnic and Racial Studies*, Vol. 28, No. 6, pp. 1000-1040.

RAMOGLIOU, Stratos E. (2009). "Uncertainty Avoidance" in: Wankel, Charles (ed.), *Encyclopedia of Business in Today's World*, SAGE Publications, Thousand Oaks, p. 1633.

RIPHAHN, Regina (2003). "Cohort effects in the educational attainment of second generation immigrants in Germany: An analysis of census data", *Journal of Population Economics*, Vol. 16, No. 4, pp. 711-737.

RODRIGUEZ, Jenny K. (2009). "Masculinity/Femininity", in: WANKEL, Charles (ed.), *Encyclopedia of Business in Today's World*, SAGE Publications, Thousand Oaks, pp. 1094-1095.

SCHÖNPFLUG, Ute (2001). "Intergenerational Transmission of Values: The Role of Transmission Belts", *Journal of Cross-Cultural Psychology*, Vol. 32, No. 2, pp. 174-185.

SVALLFORS, Stefan (2004). "Class, Attitudes and the Welfare State: Sweden in Comparative Perspective", *Social Policy & Administration*, Vol. 38, No. 2, pp. 119-138.

TRAN, Van C. and VALDEZ, Nicol M. (2017). "Second-Generation Decline or Advantage? Latino Assimilation in the Aftermath of the Great Recession", *International Migration Review*, Vol. 51, No. 1, pp. 155-190.

VAN OURS, Jan C. and VEENMAN, Justus (2003). "The educational attainment of second-generation immigrants in The Netherlands", *Journal of Population Economics*, Vol. 16, No. 4, pp. 739-753.

Appendix

Questions from the IBM Attitude Survey (Hofstede, 2001):

A Questions (used in all surveys, 1967- 73 or 1968- 73)

A1 Are you:

- I. Male (married)
2. Male (unmarried)
3. Female (married)
4. Female (unmarried)

A2 How long have you been employed by this company?

1. Less than one year
2. One year or longer, but less than three years
3. Three years or longer, but less than seven years
4. Seven years or longer, but less than fifteen years
5. Fifteen years or longer

A5-AI8. About your goals:

People differ in what is important to them in a job. In this section, we have listed a number of factors which people might want in their work. We are asking you to indicate how important each of these is to you.

In completing the following section, try to think of those factors which would be important to you in an ideal job; disregard the extent to which they are contained in your present job.

PLEASE NOTE: Although you may consider many of the factors listed as important, you should use the rating "of utmost importance" only for those items which are of the most importance to you.

With regard to each item, you will be answering the general question:

"HOW IMPORTANT IS IT TO YOU TO..." (Choose one answer for each line across)

How important is it to you to:

		Of utmost importance	Very important	Of moderate importance	Of Little importance	Of very little importance
A5	Have challenging work to do-work from which you can get a personal sense of accomplishment?	1	2	3	4	5
A6	Live in an area desirable to you and your family?	1	2	3	4	5
A7	Have an opportunity for high earnings?	1	2	3	4	5
A8	Work with people who cooperate well with one another?	1	2	3	4	5
A9	Have training opportunities (to improve your skills or to learn new skills)?	1	2	3	4	5
A10	Live in an area desirable to you and your family?	1	2	3	4	5
A11	Get the recognition you deserve when you do a good job?	1	2	3	4	5
A12	Have good physical working conditions (good ventilation and lighting, adequate work space, etc.)?	1	2	3	4	5

A1 3	Have considerable freedom to adopt your own approach to the job?	1	2	3	4	5
A1 4	Have the security that you will be able to work for your company as long as you want to?	1	2	3	4	5
A1 5	Have an opportunity for advancement to higher level jobs?	1	2	3	4	5
A1 6	Have a good working relationship with your manager?	1	2	3	4	5
A1 7	Fully use your skills and abilities on the job?	1	2	3	4	5
A1 8	Have a job which leaves you sufficient time for your personal or family life?	1	2	3	4	5

AI9-A32. About the satisfaction of your goals:

In the preceding questions, we asked you what you want in a job. Now, as compared to what you want, how satisfied are you at present with:

		Very satisfied	Satisfied	Neither satisfied nor dissatisfied	Dissatisfied	Very dissatisfied
A1 9	The challenge of the work you do—extent to which you can get a personal sense of accomplishment from it?	1	2	3	4	5
A2 0	The extent to which you live in an area desirable to you and your family?	1	2	3	4	5
A2 1	Your opportunity for high earnings in this company?	1	2	3	4	5
A2 2	The extent to which people you work with will cooperate with one another?	1	2	3	4	5
A2 3	Your training opportunities (to improve your skills or to learn new skills)?	1	2	3	4	5
A2 4	Your fringe benefits?	1	2	3	4	5
A2 5	The recognition you get when you do a good job?	1	2	3	4	5

A2 6	Your physical working conditions (good ventilation and lighting, adequate work space, etc.)?	1	2	3	4	5
A2 7	The freedom you have to adopt your own approach to the job?	1	2	3	4	5
A2 8	Your security that you will be able to work for this company as long as you want to?	1	2	3	4	5
A2 9	Your opportunity for advancement to higher level jobs?	1	2	3	4	5
A3 0	Your working relationship with your immediate manager?	1	2	3	4	5
A3 1	The extent to which you use your skills and abilities on your job?	1	2	3	4	5
A3 2	The extent to which your job leaves you sufficient time for your personal or family life?	1	2	3	4	5

A37 How often do you feel nervous or tense at work?

1. I always feel this way
2. Usually
3. Sometimes
4. Seldom
5. I never feel this way.

A43 How long do you think you will continue working for this company?

1. Two years at the most
2. From two to five years
3. More than five years (but I probably will leave before I retire)
4. Until I retire.

A48 If an employee did take a complaint to higher management, do you think he would suffer later on for doing this (such as getting a smaller salary increase, or getting the less desirable jobs in the department, etc.)?

1. Yes, the employee would definitely suffer later on for taking a complaint to higher management.
2. Yes, probably
3. No, probably not
4. No, the employee would definitely not suffer later on for taking a complaint to higher management.

A52 How often would you say your immediate manager is concerned about helping you get ahead?

1. Always
2. Usually
3. Sometimes
4. Seldom
5. Never

The descriptions below apply to four different types of managers. First, please read through these descriptions:

- Manager 1 Usually makes his/her decisions promptly and communicates them to his/her subordinates clearly and firmly. Expects them to carry out the decisions loyally and without raising difficulties.
- Manager 2 Usually makes his/her decisions promptly, but, before going ahead tries to explain them fully to his/her subordinates. Gives them the reasons for the decisions and answers whatever questions they may have.
- Manager 3 Usually consults with his/her subordinates before he/she reaches his/her decisions. Listens to their advice, considers it, and then announces his/her decision. He/she then expects all to work loyally to implement it whether or not it is in accordance with the advice they gave.
- Manager 4 (version 1967-69) Usually calls a meeting of his/her subordinates when there is an important decision to be made. Puts the problem before the group and invites discussion. Accepts the majority viewpoint as the decision.
(version 1970-73) Usually calls a meeting of his/her subordinates when there is an important decision to be made. Puts the problem before the group and tries to obtain consensus. If he/she obtains consensus, he/she accepts this as the decision. If consensus is impossible, he/she usually makes the decision him/herself.

A54 Now for the above types of manager, please mark the one which you would prefer to work under.

1. Manager 1
2. Manager 2
3. Manager 3
4. Manager 4

A551 And, to which one of the above four types of managers would you say your own manager most closely corresponds?

1. Manager 1
2. Manager 2
3. Manager 3
4. Manager 4
5. He does not correspond closely to any of them.

A56 How many years of formal school education did you complete?

1. 10 years or less
6. 15 years

- | | | | |
|----|----------|----|--------------------|
| 2. | 11 years | 7. | 16 years |
| 3. | 12 years | 8. | 17 years |
| 4. | 13 years | 9. | More than 17 years |
| 5. | 14 years | | |

A57 How old are you?

- | | | | |
|----|----------|----|------------|
| 1. | Under 20 | 5. | 35-39 |
| 2. | 20-24 | 6. | 40-49 |
| 3. | 25-29 | 7. | 50-59 |
| 4. | 30-34 | 8. | 60 or over |

A58 Considering everything, how would you rate your overall satisfaction in this company at the present time;

1. I am completely satisfied
2. Very satisfied
3. Satisfied
4. Neither satisfied nor dissatisfied
5. Dissatisfied
6. Very dissatisfied
7. I am completely dissatisfied

B Questions (used in all surveys, 1967-69 or 1968-69. but optional afterward)

B9 If you had a choice of promotion to either a managerial or a specialist position and these jobs were at the same salary level, which would appeal to you most? (You may already have been promoted in either direction, but just assume you could start again.)

1. I would have a strong preference for being a specialist
2. I would have some preference for being a specialist
3. It does not make any difference
4. I would have some preference for being a manager
5. I would have a strong preference for being a manager

B24 All in all, what is your personal feeling about working for a company which is primarily foreign owned?

1. All in all, I prefer it this way
2. It makes no difference to me one way or the other
3. I would prefer that it was not this way

B25 Suppose you quit this company. Do you think you would be able to get another job in your line of work at about the same income?

1. Yes, definitely
2. Yes, probably

3. No, probably not
 4. No, definitely not
- B39 How often would you say your immediate manager insists that rules and procedures are followed?
1. Always
 2. Usually
 3. Sometimes
 4. Seldom
 5. Never
- B44 How do you feel or think you would feel about working for a manager who is from a country other than your own?
1. In general, I would prefer to work for a manager of my own nationality
 2. Nationality would make no difference to me
 3. In general, I would prefer to work for a manager of a different nationality

How frequently, in your experience, do the following problems occur?

		Very frequently	Frequently	Sometimes	Seldom	Very seldom
B46	Employees being afraid to express disagreement with their managers	1	2	3	4	5
B47	Being unclear on what your duties and responsibilities are	1	2	3	4	5
B49	People above you getting involved in details of your job which should be left to you	1	2	3	4	5
B51	Some groups of employees looking down upon other groups of employees	1	2	3	4	5

B52-B61 About general beliefs:

Our company has employees in many countries and we are interested whether the personal opinions of employees differ from country to country. Listed below are a number of statements. These statements are not about the company as such, but rather about general issues in industry. Please indicate the extent to which you personally agree or disagree with each of these statements (mark one for each line across).

Remember: We want your own opinion (even though it may be different from that of others in your country).

		Strongly agree	Agree	Undecided	Disagree	Strongly disagree
B52	A corporation should have a major responsibility for the health and welfare of its employees and their immediate families	1	2	3	4	5
B53	Having interesting work to do is just as important to most people as having high earnings	1	2	3	4	5
B54	Competition among employees usually does more harm than good	1	2	3	4	5
B55	Employees lose respect for a manager who asks them for their advice before he makes a final decision	1	2	3	4	5
B56	Employees in industry should participate: more in the decisions made by management	1	2	3	4	5
B57	Decisions made by individuals are usually of higher quality than decisions made by groups	1	2	3	4	5
B58	A corporation should do as much as it can to help solve society's problems (poverty, discrimination, pollution, etc.)	1	2	3	4	5
B59	Staying with one company for a long time is usually the best way to get ahead in business	1	2	3	4	5
B60	Company rules should not be broken even when the employee thinks it is in the company's best interests	1	2	3	4	5
B61	Most employees in industry prefer to avoid responsibility, have little ambition, and want security above all	1	2	3	4	5

C Questions (discontinued in 1971 or before)

C1-C8. About your goals:

How important is it to you to (Mark one for each line across):

		Strongly agree	Agree	Undecided	Disagree	Strongly disagree

C1	Have the security that you will not be transferred to a less desirable job?	1	2	3	4	5
C20	Work in a department which is run efficiently?	1	2	3	4	5
C3	Have a job which allows you to make a real contribution to the success of your company?	1	2	3	4	5
C4	Work in a company which is regarded in your country as successful?	1	2	3	4	5
C5	Work in a company which stands in the forefront of modern technology?	1	2	3	4	5
C6	Work in a congenial and friendly atmosphere?	1	2	3	4	5
C7	Keep up-to-date with the technical developments relating to your work?	1	2	3	4	5
C8	Have a job on which there is a great deal of day-to-day learning?	1	2	3	4	5

C9-CI9. About general beliefs:

(Mark one for each line across)

		Of utmost importance	Very important	Of moderate importance	Of Little importance	Of very little importance
C9	A good manager gives his employees detailed and complete instructions as to the way they should do their jobs. He does not give them merely general directions and depends on them to work out the details	1	2	3	4	5
C10	Most companies have a genuine interest in the welfare of their employees	1	2	3	4	5
C11	In general, the better managers in a company are those who have been with the company the longest time	1	2	3	4	5
C12	There are few qualities in a man more admirable than dedication and loyalty to his company	1	2	3	4	5
C13	Most employees have an inherent dislike of work and will avoid it if they can	1	2	3	4	5
C14	Most employees want to make a real contribution to the success of their company	1	2	3	4	5
C15	For getting ahead in industry, knowing influential people is usually more important than ability	1	2	3	4	5
C16	By and large, companies change their policies and practices much too often	1	2	3	4	5
C17	A large corporation is generally a more desirable place to work than a small company	1	2	3	4	5
C18	Even if an employee may feel he deserves a salary increase, he should not ask his manager for it	1	2	3	4	5

C19	The private life of an employee is properly a matter of direct concern to his company	1	2	3	4	5
-----	---	---	---	---	---	---