

Understanding migration decisions and their reverberations

Inaugural-Dissertation
zur Erlangung des akademischen Grades einer Doktorin
der Wirtschafts- und Sozialwissenschaften
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von:
Finja Katharina Krüger, M.Sc.
aus Kiel

Kiel, April 2023

Erstbegutachtung:	Prof. Dr. Tobias Heidland
Zweitbegutachtung:	Prof. Dr. Dr. Ulrich Schmidt
Tag der mündlichen Prüfung:	22.09.2023

Acknowledgements

I want to thank everyone who played a part in making this dissertation happen. Without your continuous support, none of this would've been possible.

I'd like to thank Tobias Heidland. You always pushed me to believe in myself, even during moments of doubt. Your guidance, advice, and encouragement, especially during moments of panic, were a lifesaver. Your dedication, time, effort, and flexibility were crucial in making this dissertation possible. I wouldn't be where I am today without your help.

A huge thanks also goes to Ulrich Schmidt. Over six years ago, you introduced me to research. Knowing that you believed in me gave me the confidence in my capabilities, and without that, this dissertation wouldn't have been possible. You've been by my side, offering guidance and support all these years, and I can't express how much that means to me.

I want to give a big shout-out to my colleagues, who always had a listening ear, put up with my bursts of energy, and made my time at IfW so much fun. Thanks to the development and behavioral team, and a special thanks to Lena and Jan. You've been there for me every step of the way, encouraging and believing in me. I'll always remember the countless conversations, whether they were about research or life in general. I can't stress enough how much I owe to Malte. One thing I'm certain of is that this dissertation would have been impossible without you. Thank you for all our conversations, our after-work Kaffee und Kuchen, our lunch breaks, the laughter, the walks, the fun, and your open door. Your unconditional support got me through the tough times.

I am truly thankful to all my friends for distracting me from research when I needed it the most. A special thanks to Jasper for all those walks in Schrevenpark and your unwavering support, both personally and in research.

Above all, I want to thank my parents. You've been my biggest cheerleaders for almost 30 years, unconditionally supporting me with love and encouragement no matter what. You celebrate my successes and hold me up when I need it most. You've always reminded me that life is about so much more than academia. I've learned a ton from you—your curiosity for new things, your belief that everything will work out, and your endurance. Thank you for always being there for me, reassuring me, and for your love and guidance.

Contents

1	Introduction	1
2	Counting the Costs: The Impact of Heterogeneous Migration Costs on Female Mobility	9
2.1	Introduction	10
2.2	Literature	12
2.3	Hypothesis and study design	15
2.3.1	Family obligation costs	16
2.3.2	Gender discrimination costs	17
2.3.3	Leaving-family costs	18
2.3.4	Social norm costs	19
2.4	Data sampling and collection	20
2.5	Analysis and results	22
2.5.1	Summary Statistics	22
2.5.2	Empirical approach	23
2.5.3	Analysis	24
2.5.4	Index for indirect migration costs	31
2.5.5	PCA for indirect migration costs	34
2.6	Discussion and conclusion	37
	Appendix	44
2.A.1	Study area	44
2.A.2	Summary statistics	46
2.A.3	Analysis	46
2.A.3.1	PCA for indirect migration costs	49
2.A.4	Robustness checks	52
3	Forecasting Bilateral Refugee Flows with High-dimensional Data and Machine Learning Techniques	67
3.1	Introduction	68
3.2	Literature on Migration Forecasting	70
3.3	Data	71

3.3.1	Asylum Seeker Flows	71
3.3.2	Languages	72
3.3.3	Keywords	73
3.3.4	Google Trends Data	74
3.3.5	Classical Predictors	75
3.4	Forecasting Methodology and Performance Comparison	76
3.4.1	Elastic Net	78
3.4.2	Random Forest	78
3.4.3	Extreme Gradient Boosting Regression Tree	79
3.4.4	Factor Approach	79
3.4.5	Ensemble Forecasts	79
3.5	Results	80
3.6	Towards a Feasible Forecasting Approach for Refugee Flows	85
	Appendix	91
3.A.1	Literature	91
3.A.1.1	Trends in Approaches	91
3.A.1.2	Trends in Data Sources	92
3.A.2	Tables	98
3.A.3	Appendix on Methodology	105
3.A.4	Additional Results	106
4	Which Individual-Level Factors Explain Public Attitudes Toward Im-	
	migration? A Meta-Analysis	110
4.1	Introduction	111
4.2	Theoretical Arguments	113
4.3	Data and Methods	117
4.3.1	Study selection and data generation	117
4.3.2	The Meta-Analytical methodology	120
4.4	Results	121
4.4.1	Main results	121
4.4.2	Heterogeneity analysis	123
4.4.3	Study heterogeneity	125
4.5	Conclusion	127
	Appendix	134
4.A.1	List of journals	134
4.A.2	Study selection approach	139
4.A.3	List of measures of attitudes to immigration	151
4.A.4	List of independent variables	154
4.A.5	Meta-analytical approach	155

4.A.6 Forest plots	157
4.A.7 Publication bias	174
4.A.8 Bayesian model averaging	176
4.A.9 Description and Summary Statistics	186
5 Country, Culture or Competition - What Drives Attitudes toward Im-	
migrants in Sub-Saharan Africa?	191
5.1 Introduction	192
5.2 Drivers of attitudes toward immigrants: theory and evidence	194
5.3 Hypotheses, data & study design	197
5.3.1 Hypotheses	197
5.3.2 Sampling and data collection	198
5.3.3 Sample characteristics	200
5.4 Empirical approach	200
5.4.1 Conjoint experiment	200
5.4.2 Observational data	206
5.5 Results	210
5.6 Conclusion	222
Appendix	231
5.A.1 Background Uganda and Senegal	231
5.A.1.1 Study area	232
5.A.2 Figures	233
5.A.3 Tables	264
Erklärungen	280
Eigenständigkeitserklärung	281
Erklärung der Koautoren	282

Chapter 1

Introduction

Migration has become a defining feature of the contemporary global landscape. It is estimated that in 2020, there were 281 million international migrants worldwide, representing 3.6% of the world's population (International Organization for Migration, 2022). International migration (henceforth typically abbreviated simply as migration) is one of the most complex issues facing policymakers today, and understanding its many interconnected factors is essential for designing effective policies. The decision to migrate is a complex one that involves a wide range of economic, social, and political factors at both the individual and household levels. Once the decision to migrate is made, forecasting the flow of refugees and migrants becomes crucial for policy-makers and practitioners alike. Finally, the perception of host populations toward immigrants and refugees is an important aspect of the migration process that can shape the success or failure of integration policies and programs.

My dissertation delves into the complexities of the migration process, which involves various steps that policymakers need to consider when crafting efficient policies. To this end, I have written four papers that explore different aspects of migration. These papers cover a wide range of topics, including the decision to migrate, predicting refugee flows, and examining public attitudes toward immigrants, both generally and in Sub-Saharan African countries in particular. Furthermore, I employ four different methods across my four papers - survey-observational, machine learning, meta-analytic, and experimental - to gain a more complete understanding of the decision to migrate and its various factors. By taking this broad approach both thematically as well as methodology-wise, I can provide detailed insights into the complexities of migration and offer valuable guidance for policymakers seeking to design effective and evidence-based policies.

The decision to migrate has almost exclusively been studied from a male perspective, neglecting potential gender differences in the decision to migrate. However, data on migration patterns show that women have limited geographic mobility. For example, data from the Migration Data Portal (2023) indicates that in 2019, women accounted

for only 41.5% of international migrant workers. Therefore, understanding the gendered aspects of migration is crucial for policymakers. In my *first paper*, I provide a gendered perspective on the study of international migration intentions by examining potential structural disparities in direct and indirect costs of migration between males and females. According to neo-classical economic theory, individuals make migration decisions based on a cost-benefit analysis (Borjas, 1989). These costs can include the financial costs of moving as well as psychological costs, such as leaving behind friends and family (Todaro, 1980).

In my paper, I use primary observational survey data gathered from potential migrants from Uganda and Senegal to separate different cost factors according to gender. My data indicate substantial gender differences in migration intentions, with women having a 6.3% lower predicted probability of having migration intentions as compared to men. To explain this gendered migration intention gap, I investigate whether there are distinct valuations of the expected direct cost of migration for women, but I do not find any significant gender-based differences in this factor. To delve deeper, I introduce four novel indirect cost components, namely family obligation, gender discrimination, leaving-family, and social norm costs, and explore their impact on migration intentions for men and women.

Family obligation costs consider an individual's family situation, including marital status, number of children, and dependent adults in the household. Gender discrimination costs include both the perceived level of gender discrimination and the ideal situation regarding gender rights. Leaving-family costs consider the costs associated with the loss of influence in household decision-making regarding the future of the children when migrating. Lastly, social norm costs cover the expenses linked to breaking norms when migrating illegally and the opinions of an individual's family members regarding their decision to migrate. The findings reveal that indirect costs significantly affect migration intentions, with family obligation, leaving-family, and social norm costs being particularly important. Increased family obligation and social norm costs reduce migration intentions for both males and females. Leaving the family behind is a critical determinant only for women, along with certain aspects of social norm costs that are particularly relevant for women. Accounting for the four novel indirect cost factors reduces the gender migration intention gap by almost 45% and diminishes the effect of gender to insignificance.

The first paper's findings on factors that influence migration intentions for men and women can help policymakers to design targeted interventions that address the specific needs of different groups, improve migration outcomes and promote gender equality. Additionally, this can help to address gender-based discrimination and enhance the overall well-being of migrants and their families. For example, policies that reduce family obligation and social norm costs can encourage more women as well as men to migrate. Similarly, policies that provide support for families left behind can help alleviate the

concerns of women who are reluctant to migrate due to leaving-family costs.

Over the past decade, forecasting refugee flows has become a central issue across various disciplines due to increased political attention, improved data availability, and methodological and technological advancements. In particular, the 2013-2015 refugee crisis in the European Union, which saw over 1.2 million asylum applications at its peak in 2015, highlighted the importance of preparedness for governments and humanitarian organizations to ensure safe conditions for asylum seekers and refugees.

However, forecasting refugee flows presents unique challenges due to the circumstances of forced migration, such as the lack of data collection in conflict areas, making it difficult to obtain early signals about international population movements. Additionally, the occurrence of violent conflict, natural disasters, or economic crises tends to be unpredictable, further complicating predictions (Disney et al., 2015).

In response to these challenges, in my *second paper*, my co-authors and I focus on monthly refugee flow forecasting models for 150 origin countries to the EU27, using machine learning and high-dimensional data, including digital trace data from Google Trends. This paper offers three main contributions.

First, we construct a unique dataset that combines a range of Google Trends time series capturing internet searches for migration-related keywords at the origin with administrative records of asylum registries at the destination. We also included hundreds of predictor candidate variables capturing the occurrence of natural disasters and violent conflict at origin, as well as a wide range of socio-economic, labor, and monetary statistics in both origin and destination countries. This rich dataset allows us to exploit monthly frequency in the panel of bilateral refugee flow corridors.

Second, we compare the forecasting performance of different models and approaches that are commonly used in data science forecasting applications. The approach involves testing model performance in a forecasting-at-scale exercise where the models were assessed based on their average forecasting power over many migration corridors and across a range of different specifications that involve different compositions of information sets, including high-frequency Google Trends. This is different from existing work, which typically focuses on assessing corridor-specific customized forecasting approaches using a single algorithm (Carammia et al., 2022). With that approach, we can analyze the performance differences of forecasting models at both the corridor and aggregate levels.

Finally, we provided novel insights into the forecasting power of Google Trends data by systematically comparing forecast performance across specifications with and without these indices. Specifications using only Google Trends predictors produce useful forecasting results for large refugee flow corridors. This approach has the advantage of close-to-real-time availability independent of data publishing lags, which typically restrict the viability of short-term forecasts decisively.

In summary, the second paper offers a unique approach to refugee flow forecasting, using machine learning and high-dimensional data, including digital trace data from Google Trends. The work can contribute to the preparedness of governments and humanitarian organizations to ensure safe conditions for asylum seekers and refugees en route and upon arrival.

The perception of migrants among the local population upon migrants' arrival in the destination country plays a crucial role in determining their successful integration. However, immigration has become a major concern for voters and is often considered one of the main challenges for countries and political entities like the European Union (e.g. Commission (2019a, 2019b)). Public attitudes toward immigration have now become a political divide, especially after the "migration crisis" in 2015 (Hobolt, 2016). Consequently, scholars have shown a growing interest in understanding the factors behind individual differences in attitudes toward immigration. However, no firm conclusions have been drawn about the drivers of attitudes toward immigration. To address this gap, in my *third paper* my co-authors and I conduct a meta-analysis to assess the determinants of attitudes toward immigrants.

In this study, we examine articles published in the top thirty journals across economics, political science, sociology, psychology, and migration/ethnic/demographic studies between 2009 and 2019. Specifically, we look at individual factors that drive preferences toward migration policy (e.g., preferred levels of immigration) and views about immigrants' contribution to society. We focus on eight individual-level indicators, including age, gender, education, income, occupational and unemployment status, respondents' minority background, and the type of area (urban versus rural) they live in. We select these individual independent variables based on previous reviews of public opinion on migration and focus only on widely-used determinants (see e.g. Hainmueller and Hopkins (2014)).

The meta-analysis reveals that two individual-level characteristics are most significantly associated with attitudes toward immigration: education (positively) and age (negatively). On the one hand, educated individuals consistently hold more positive attitudes toward immigration, and education's effect is not only due to its connection to social class but also represents a value-based cleavage. On the other hand, older respondents hold significantly more anti-immigration attitudes than younger respondents. Positive attitudes toward immigration are also correlated with an individual's economic standing. Higher-income levels and high-skill occupations lead to more positive attitudes toward immigration. Living in urban areas is also positively associated with pro-immigration attitudes because urban residents have more contact with migrants and potentially self-select into cities. Additionally, based on Bayesian model averaging, we demonstrate that omitting certain variables in the analysis can lead to wrongly attributed effects on other explanatory variables. For instance, the effect of education, income, and age may vary

depending on whether other individual characteristics, such as minority background and gender, are considered.

Finally, paper number 3 shows that attitudes toward immigration policy and attitudes toward immigrants' contribution are not necessarily explained by the same individual characteristics. Therefore, it is important to consider different factors when addressing and attempting to change attitudes toward immigration. The findings provide a unique and insightful perspective on attitudes toward immigration, with significant implications for both researchers and policymakers.

In my *fourth paper*, my colleagues and I delve deeper into the topic of attitudes toward immigrants by studying attitudes toward immigrants in general and specific immigrant groups, such as Chinese immigrants, in Sub-Saharan Africa. As an important destination region for international migrants, Africa currently hosts about 10% of the world's international migrants, of which over 20% come from outside the continent (Achieng et al., 2020). Within the continent, Sub-Saharan Africa is the region that attracts the majority of immigrants, with Western and Eastern Africa hosting almost 60% of them. With strong population growth and increasing regional integration, immigrant populations in Africa are projected to grow significantly in the future (OECD, 2021), making it the most dynamic region in terms of migration in the next decades.

However, despite this growing trend, the scientific literature on attitudes toward immigrants has primarily focused on high-income countries, particularly Europe and the USA, as seen in studies such as Ajzenman et al. (2022), Alesina et al. (2023), Card et al. (2012), Hainmueller and Hopkins (2015), and Valentino et al. (2019). As a result, we know very little about how attitudes toward immigrants are determined in Sub-Saharan Africa and other developing countries, leaving policymakers from these regions lacking in quantitative evidence on the impact of migration on their citizens' attitudes. To address this gap, this paper aims to study attitudes toward immigrants with original experimental and observational survey data from Uganda and Senegal. The conjoint experiment reveals that sociotropic concerns, both economic and cultural, are the main drivers of preferences for immigrants, with egocentric economic concerns and power concerns playing a relatively minor role. Respondents strongly prefer immigrants who are willing to integrate and have high-paying jobs, regardless of respondents' skill levels. Moreover, respondents prefer African immigrants over non-African immigrants. Furthermore, our survey data reinforces the importance of sociotropic concerns and reveals that although Chinese immigrants are viewed somewhat more negatively than the average immigrant, attitudes toward both groups are shaped by the same mechanisms. Although power concerns are generally large in absolute terms, they have no explanatory power for attitudes toward immigrants.

This paper contributes to the literature in three ways. Firstly, we investigate whether

drivers found influential in developed countries also apply to Sub-Saharan Africa. Secondly, we extend the three well-established theoretical drivers of public attitudes - sociotropic economic concerns, egocentric economic concerns, and sociotropic cultural concerns - to include a fourth driver: concerns about foreign influence, which we term power concerns. Given the region's history of colonialism and dependence, power concerns may be particularly relevant in Sub-Saharan Africa, as opposed to China's official non-interference policy (Dreher et al., 2018). Finally, we differentiate between different immigrant groups, with a particular focus on attitudes toward Chinese immigrants, a group that has grown significantly in recent years and is perceived controversially by local populations and policymakers alike (Park, 2022).

The fourth paper's findings provide policymakers with evidence-based support for easing migration restrictions, for example, within the African Union. They also suggest that when aiming to reduce tensions and xenophobia, communicating the positive impacts of immigration on society at large and promoting immigrants' willingness to integrate into host societies are worthwhile mechanisms to target.

Taken together, these papers provide a comprehensive examination of various aspects of international migration. By studying the decision to migrate, forecasting refugee flows, analyzing the factors affecting attitudes to immigration, and investigating the reception of immigrants in host countries, I contribute to a better understanding of the complexities and nuances of international migration. My findings have important policy implications for managing migration flows, addressing gender disparities in migration intentions, and improving attitudes toward immigrants.

Bibliography

- Achieng, M., El Fadil, A., Adepoju, A., Fumagalli, C., & Nyabola, N. (2020). *Africa migration report. Challenging the narrative*. International Organization for Migration, Addis Ababa.
- Ajzenman, N., Aksoy, C. G., & Guriev, S. (2022). Exposure to transit migration: Public attitudes and entrepreneurship. *Journal of Development Economics*, 102899.
- Alesina, A., Miano, A., & Stantcheva, S. (2023). Immigration and redistribution. *The Review of Economic Studies*, 90(1), 1–39.
- Borjas, G. J. (1989). Economic theory and international migration. *International Migration Review*, 23(3), 457–485.
- Carammia, M., Iacus, S. M., & Wilkin, T. (2022). Forecasting asylum-related migration flows with machine learning and data at scale. *Scientific Reports*, 12(1), 1457.
- Card, D., Dustmann, C., & Preston, I. (2012). Immigration, wages, and compositional amenities. *Journal of the European Economic Association*, 10(1), 78–119.
- Commission, E. (2019a). *Standard Eurobarometer 92 - Autumn 2019*.
- Commission, E. (2019b). *Standard Eurobarometer 95 - September 2021*.
- Disney, G., Wiśniowski, A., Forster, J. J., Smith, P. W., & Bijak, J. (2015). Evaluation of existing migration forecasting methods and models. *ESRC Centre for Population Change, University of Southampton*.
- Dreher, A., Fuchs, A., Parks, B., Strange, A. M., & Tierney, M. J. (2018). Apples and dragon fruits: The determinants of aid and other forms of state financing from China to Africa. *International Studies Quarterly*, 62(1), 182–194.
- Hainmueller, J., & Hopkins, D. J. (2014). Public attitudes toward immigration. *Annual Review of Political Science*, 17, 225–249.
- Hainmueller, J., & Hopkins, D. J. (2015). The hidden American immigration consensus: A conjoint analysis of attitudes toward immigrants. *American Journal of Political Science*, 59(3), 529–548.
- Hobolt, S. B. (2016). The Brexit vote: A divided nation, a divided continent. *Journal of European Public Policy*, 23(9), 1259–1277.
- International Organization for Migration. (2022). *World migration report 2022*. IOM.

- Migration Data Portal. (2023). Gender and migration [Downloaded from <https://www.migrationdataportal.org/themes/gender-and-migration>] on 07 March 2023].
- OECD. (2021). *International migration outlook 2021*. Organisation for Economic Co-operation and Development Publishing.
- Park, Y. J. (2022). Forever foreign? Is there a future for Chinese people in Africa? *Journal of Ethnic and Migration Studies*, 48(4), 894–912.
- Todaro, M. (1980). Internal migration in developing countries: A survey. In *Population and economic change in developing countries* (pp. 361–402). University of Chicago Press.
- Valentino, N. A., Soroka, S. N., Iyengar, S., Aalberg, T., Duch, R., Fraile, M., Hahn, K. S., Hansen, K. M., Harell, A., Helbling, M., Jackman, S. D., & Kobayashi, T. (2019). Economic and cultural drivers of immigrant support worldwide. *British Journal of Political Science*, 49(4), 1201–1226.

Chapter 2

Counting the Costs: The Impact of Heterogeneous Migration Costs on Female Mobility

Finja Krüger

Abstract

Although the gender gap in international migration is a well-documented phenomenon, gender has been largely ignored by economists who study migration. This paper provides a gendered perspective to the study of international migration intentions by examining the structural disparities in direct and indirect costs between men and women. Using primary survey data gathered from 1,944 potential migrants from Uganda and Senegal, this study introduces four novel indirect cost components— family obligation, gender discrimination, leaving-family, and social norm costs— and explores how they affect men’s and women’s migration intentions. The results indicate a substantial gender gap in migration intentions, with women having a 6.3% lower predicted probability of having migration intentions as compared to men holding several important factors known to be relevant for migration intentions constant. I do not find any gender-based differences in the expected direct costs of migration. Indirect costs significantly affect migration intentions, with family obligation, leaving-family, and social norm costs being particularly influential. Importantly, while family obligation and social norm costs are important for both genders, leaving the family behind is a relevant determinant only for women. Certain aspects of social norm costs are also particularly relevant to women. Accounting for these indirect cost factors reduces the gender migration intention gap by almost 45% and turns the gender coefficient statistically insignificant.

Acknowledgment: Funding from the Mercator Foundation, project "MEDAM II" (19002900), is acknowledged.

2.1 Introduction

Although the gender gap in international migration has significantly decreased over time, it remains a well-documented phenomenon that persists to this day (Abel & Cohen, 2022). For instance, data from the Migration Data Portal (2023) indicate that in 2019, women accounted for only 41.5% of international migrant workers. However, the topic of gender in international migration has been largely overlooked by economists, despite the increasing attention paid to gender issues in almost all other areas of applied economics (e.g., Bertrand et al. (2010), Blau and Kahn (2017), Budig and England (2001), Croson and Gneezy (2009), Duflo (2012), and Goldin and Katz (2002)).

My paper offers a gendered perspective to the study of international migration decision-making and seeks to answer the question of how gender-specific migration costs affect migration intentions. While neo-classical economic theory posits that individuals base their migration decisions on a cost-benefit analysis (Borjas, 1989), this study delves deeper by examining the potential structural disparities in direct and indirect costs between men and women. Through this investigation, I aim to determine whether these differences can account for variations in the migration aspirations of men and women. By considering the gendered nuances of migration costs, this paper provides a more comprehensive understanding of the complex decision-making processes that underlie international migration.

This study makes a valuable contribution to the ongoing discourse on migration costs by introducing four novel indirect cost components. Firstly, family obligation costs are considered, which encompasses information on an individual's family situation, including their marital status, the number of children, and dependent adults in the household. Secondly, gender discrimination costs are factored in, which include both the perceived level of gender discrimination and the ideal situation regarding gender rights. Thirdly, leaving-family costs are also considered, which assess the costs associated with the loss of influence in household decision-making regarding the future of the children when migrating. Lastly, social norm costs are incorporated, which cover the expenses linked to breaking the norms when migrating illegally and the opinions of the individual's family members regarding the respondents' decision to migrate. These new cost components serve to expand our understanding of the multifaceted and complex nature of migration costs.

Furthermore, this study adds value to the investigation of the gender migration gap by systematically investigating the gendered effects of migration costs. More specifically, this research examines the structural differences in expected direct migration costs as well as indirect costs between men and women and investigates whether these differences can explain the gender gap in migration aspirations. By combining the gendered cost components, I introduce, the study aims to provide a comprehensive analysis of the

decision-making process behind migration, shedding light on the complex factors that influence migration decisions and their gendered implications.

This study draws on comprehensive primary survey data gathered from 1,944 potential migrants from Uganda and Senegal. Using this data, the analysis examines the various factors that contribute to indirect costs as well as their gendered impact on migration intentions. To achieve this goal, two distinct approaches are employed to combine the various aspects of indirect costs. Firstly, the study draws upon existing literature to manually construct indices for indirect costs by combining different survey questions. Secondly, a Principal Component Analysis is used to compute the different factors involved. By employing these complementary approaches, this research aims to provide a nuanced and comprehensive understanding of the complex role that indirect costs play in shaping gendered migration patterns.

The sample of this study reveals a substantial gender difference in migration intentions, with women having a 6.3% lower predicted probability of having migration intentions, even when holding several important factors known to be relevant for migration intentions constant (Migali & Scipioni, 2019). Incorporating four different indirect cost components in my analysis reveals that family obligation, leaving-family, and social norm costs significantly affect migration intentions. While gender differences in the expected cost of migration are insignificant, indirect costs have different impacts on men and women.

Although family obligation and social norm costs are important for both genders, leaving the family behind is a critical determinant only for women. Moreover, my analysis, including the PCA, indicates that certain aspects of social norm costs are particularly relevant to women. By accounting for the indirect cost factors and their heterogeneity, a large part of the gender gap in migration intentions can be explained. Specifically, including the four different indirect cost factors in the analysis reduces the difference in the predicted probability of having migration intentions between men and women to 4.0% for the manually constructed indices and 3.5% for the principal components, indicating that nearly 45% of the gender-based differences in migration intentions can be explained. Furthermore, the impact of gender is diminished to insignificance.

The paper is organized as follows. Section 2.2 provides a discussion of the relevant theory and existing evidence upon which the analysis is built. The hypotheses and study design are presented in Section 2.3, followed by a description of data sampling and collection in Section 2.4. Section 2.5 covers the summary statistics and empirical approach, leading to the presentation of the results. Finally, Section 2.6 presents the conclusion.

2.2 Literature

According to neo-classical economic theory, individuals make decisions about migration based on a cost-benefit analysis (Borjas, 1989). The costs of migration include the financial costs of moving, as well as psychological costs like the emotional and social costs of leaving behind friends and family (Todaro, 1980). The benefits of migration include better job opportunities, higher wages, and improved living conditions. To make an informed decision about whether or not to migrate, individuals must weigh the costs and benefits against each other. This cost-benefit analysis can be influenced by a variety of factors, such as the individual's education level, language skills, and job experience.

One factor which is highly understudied in the literature is the influence of gender, such as gender roles, norms, and expectations. Some studies focus solely on men (Bah, Batista, et al., 2018; Munshi & Rosenzweig, 2016; Shrestha, 2020). In fact, much of the existing literature disregards women altogether, incorrectly assuming that they do not migrate for work or education. These studies often implicitly consider women mainly as tied movers (for an overview of family migration research see Cooke (2008)) or vulnerable forced migrants. Other studies assume that migration costs are generally homogeneous across genders (Dustmann et al., 2023), which is highly unlikely to be true.

The literature discussing the "feminization of migration", which assesses female migration as an individual, independent act, started to grow in the 1980s (e.g. Lauby and Stark (1988)).¹ Many of these studies document the limited geographic mobility of women. However, there exists no established explanation for this phenomenon. Some research shows that women may face additional barriers to migration, such as discrimination, exploitation, and gender-based violence, which can affect their ability to migrate (O'Neil et al., 2017). In addition, studies have found that women's decision to migrate is often influenced by their family responsibilities, such as caring for children or elderly relatives (Hidrobo et al., 2022). This study takes a novel approach by integrating various direct and indirect cost factors and assesses the gendered impact on migration intentions.

There is little research that assesses the direct costs of migration. One exception are the KNOMAD-ILO Migration and Recruitment Costs Surveys, which collect data on the financial and non-financial expenses that migrant workers experience as they seek employment opportunities in foreign countries (World Bank, 2017). Their findings show that the direct cost of migration varies by gender (Table 2.1). Women who migrate with their families tend to have the lowest migration costs. However, when they migrate alone, their migration costs are higher than those of single men.² This can be attributed

¹The concept of "feminization of migration" is controversial as women have comprised almost half of all migrants for several decades (Jolly et al., 2005). Nevertheless, some regions have observed a rise in female migration, and the term "feminization of migration" can also refer to changes in migration patterns, where more women are migrating independently, such as for job opportunities, instead of joining male family members.

²This also holds in Senegal which, jointly with Uganda, constitutes the setting for the empirical part

to the increased risk of violence and trafficking that women face during migration, against which they need to financially safeguard themselves. Additionally, the higher costs for women could be viewed as a "pink tax" on direct migration expenses for female migrants. Furthermore, some countries put a restriction on female mobility such that a married woman cannot obtain a passport in the same way as a married man or by restricting women from taking a job without their husband's permission (World Bank, 2020).

Table 2.1: Migration cost in constant USD equivalent of total expenditure

		Men	Women	p.overall
Overall	Married	1172.10	443.68	0.003
	Single	896.03	1128.82	0.593
Observations		1899	503	
		Male	Female	
Senegal to Italy	Married	1269.36	1205.53	
	Single	1514.38	2149.02	
Observations		72	9	

Note: Based on data from World Bank, 2017

The migration decision can be influenced by various factors that contribute to the indirect costs of migration. Specifically, the indirect costs of migration increase for individuals with families as they must consider not only their own needs but also the needs of their families. Moreover, the gendered responsibilities within the household can affect the costs of men and women differently. I refer to these additional costs as family obligation costs. Accounting for gender and marital status, Gubhaju and De Jong (2009) find that compared to other marital status categories, married women exhibit the lowest likelihood of migrating. The findings lend support to the proposition that married women in South Africa face mobility constraints and encounter difficulties in making migration decisions due to their responsibilities as primary caregivers for children and elderly family members. These constraints may be attributed to limitations imposed on them by male members of the household. Hughes et al. (2020) results from Nepal indicate that gendered family norms, which assign breadwinning to men and caregiving to women, may oversimplify the complex ways in which individuals fulfill familial responsibilities. Men who prioritize family obligations tend to migrate internationally, particularly to India, even if it means sacrificing some economic benefits for the sake of being closer to family. On the other hand, women who prioritize family needs are less likely to migrate. However, if women have the responsibility of caring for their adult parents, they are more likely to migrate to India but not to other international or domestic destinations. Therefore, it is important to account for family dependencies as well as the evaluation of family commitment when

of my paper. Uganda is not covered in the KNOMAD-ILO data.

assessing family obligation costs.

Gender discrimination has been a key topic in the discussion of female migration costs. The literature primarily views gender discrimination as a push factor encouraging women to migrate. The evidence supporting this assumption is mixed (Hugo, 2000) since women may also be prevented from leaving due to restrictions on the roles assigned to them (Zachariah et al., 2001). When examining this relationship, the majority of the literature depends on macro data. For example, Ferrant and Tuccio (2015) present evidence that higher levels of discriminatory institutions in origin countries impede the likelihood of female migration. Categorizing countries either as patriarchal or matrifocal, Massey et al. (2006) find lower female-to-male migration ratios in patriarchal societies. In matrifocal countries, on the other hand, migration rates of women are way higher, sometimes even surpassing those of men. Differentiating between low and high-income countries, Gutmann et al. (2023) find that enhancing economic or political rights for women can decrease the gender migration gap in the poorest countries while empowering women in richer countries increases the gap due to reduced benefits of migration.

Ruyssen and Salomone (2018) use microdata to evaluate the impact of gender discrimination on international migration. Their study offers compelling evidence supporting the hypothesis that gender discrimination intensifies women's migration aspirations. However, it does not influence their actual migration behavior which is affected by more traditional factors like income or networks.

When women migrate, this affects the ones who remain behind, especially dependent people like children and the elderly, which in turn can affect the migration decision (Böhme et al., 2015; Stöhr, 2015). This adds another cost component to migration which I call leaving-family costs. For example, research indicates that children with migrant mothers are more prone to falling behind in their education than those with migrant fathers (Cortes, 2015). Additionally, research shows that when women migrate, less investment is spent on the human capital of the children left behind. Households headed by the husband of a woman who migrated, tend to spend significantly less on education as compared to households where the men migrated (Pfeiffer & Taylor, 2007). Especially, the school performance of girls may be influenced by the absence of their mothers, as they often have to shoulder more household chores. According to McKenzie and Rapoport (2011), girls aged 16-18 in migrant households are more likely to report that their primary activity was housework, as compared to those in non-migrant households. This in turn affects their schooling attendance and attainment. According to Duflo (2003), women tend to invest more in girls than in boys, resulting in girls being more affected by the migration of women in their lives. Furthermore, the emigration of highly skilled women tends to negatively affect three crucial indicators in education and health - infant mortality, under-5 mortality, and gender-specific secondary school enrollment rates (Dumont et al., 2007).

Another critical decision that can shape the future of children, and one where migrating individuals may lose the ability to make critical choices, is the appropriate age for marriage. A qualitative study carried out in Kenya, Senegal, Uganda, and Zambia reveals that there is a strong association between formal education and postponement of marriage (Petroni et al., 2017). In Uganda, girls are often mainly responsible for household chores instead of receiving an education. This reinforces gender inequality, as women become financially dependent on men. Consequently, this may result in an increased incidence of child marriage, particularly among girls from low-income families.

Migration is not always authorized, and irregular migration adds another cost component to the decision to migrate. These costs are referred to as social norm costs. Ryo (2013) explores the factors that influence an individual’s decision to migrate illegally across international borders. The author discusses how personal values, cultural norms, economic opportunities, and the perceived risks and rewards associated with unauthorized migration all contribute to the decision-making process. The author finds that the intent to migrate illegally is significantly influenced by individuals’ overall legal attitudes, their moral stance on violating immigration law, their perception of the legitimacy of authorities, and their social norms regarding border crossing. Therefore, a thorough understanding of an individual’s underlying values and norms is essential to gaining a complete understanding of their decision to migrate illegally.

2.3 Hypothesis and study design

In this paper, I expand the previously established assumptions regarding the cost components of migration by combining different indirect cost factors. Additionally, I contribute to the investigation of gender disparities in migration decision-making thereby helping to explain part of the gender migration gap. The upcoming sections formulate hypotheses regarding individual cost components and illustrate the corresponding study design.³

Drawing from existing literature, I differentiate between two components, namely direct costs and indirect costs. Direct costs include expenses incurred during the actual travel. As the migration costs of married individuals are often unclear due to the inclusion of their families, my hypothesis focuses on the migration costs of singles. Specifically, based on data from World Bank (2017), I hypothesize:

Hypothesis 1 *Single women have higher expected direct migration costs than single men.*

In my sample, I observe individuals in their home countries. This implies that direct costs of migration are *expected* costs, which is different from the World Bank’s estimates

³All hypotheses have been pre-registered. The pre-analysis plan has been submitted to OSF on 02 November 2021.

of *actual* costs that were discussed above. Studying the willingness to migrate based on expected direct costs has the advantage that assessments will not be biased or informed by the experience of migration of a selected sample. As a proxy for expected direct migration costs, I ask for the expected average migration costs of Ugandan/Senegalese migrants to Gulf and European countries.

Indirect costs refer to the non-monetary or intangible costs that individuals or households may face when deciding to migrate. These costs are not directly related to the physical transportation of the migrants, but rather to the social, cultural, and psychological impacts of migration. I assume, based on the literature, that the indirect costs comprise the following subcomponents that impact male and female migrants differently:

- Family obligation costs
- Gender discrimination costs
- Leaving-family costs
- Social norm costs

2.3.1 Family obligation costs

It is plausible that compared to married family members with children, single individuals find migration considerably easier. When migrating as a family, the decision becomes more complex and costlier, as it affects more people. Individuals must consider the logistics of bringing their family and children along or making arrangements for their care while away. This increases the indirect costs of migration. Due to their familial responsibilities as the caregiver of the family, women bear a higher cost of family obligation as compared to men who are perceived as the breadwinner. This results in less migration for women and more migration for men. Therefore, I hypothesize:

Hypothesis 2 *Family obligation costs decrease migration intentions for women.*

To test this hypothesis, I asked the following question regarding family bonds and dependencies, taking also the valuation of family commitment into account.

Is the respondent married	yes/no
How many children (biological and adopted) do you have?	number
How many children do you provide care for?	number
How many elderly and sick family members do you provide care for?	number
Overall, which do you think is better for most people in [country] today—to put individual needs first or to put family needs first?	individual, about the same, family

2.3.2 Gender discrimination costs

Gender discrimination costs capture the costs of women experiencing discrimination in their home country. This may push them to seek better opportunities and conditions in places where gender discrimination is less prevalent. Conversely, when women are limited in their roles and opportunities, this may prevent them from leaving their home country in search of more favorable conditions. Ruysen and Salomone (2018) distinguish between the impact of gender discrimination on migration intentions as well as migration plans. They find that gender discrimination is a key factor that can increase the desire to migrate. However, they highlight that whether or not these intentions are translated into actual migration depends on several factors, such as income, social networks, and familial obligations. Given the focus on migration intentions rather than concrete plans in this analysis, I hypothesize:

Hypothesis 3 *Gender discrimination costs increase migration intentions for women.*

I draw inspiration from the work of Ruysen and Salomone (2018) on gender discrimination and migration intentions. Their seminal contribution to the literature lies in their innovative use of a subjective, individual measure of gender discrimination. Unlike previous studies, Ruysen and her colleague capture the first-hand experiences of women in facing gender biases, both societal and familial. To derive their measure, the authors use questions from the Gallup World Poll, which I also covered in my data collection. They are as follows:

Do you believe women in [country] are treated with respect and dignity, or not? 1=very much to 6=not at all

In this country, are men and women treated fairly at work, or not? 1=very much to 6=not at all

In your opinion, is domestic violence a serious problem for [country] or not?" 1=very much to 6=not at all

Ruyssen and Salomone (2018) find that perceived gender discrimination has a negative and significant association with economic equality in the workplace, while there exists a significant positive correlation with the severity of domestic violence. I furthermore added two questions asking about the ideal situation on gender equality:

In general, is it important that men and women have the same opportunities at work 1=very much to 6=not at all

In general, is it important that women are treated with respect and dignity? 1=very much to 6=not at all

Ruyssen and Salomone (2018) primarily focus on the first question about respect and dignity due to the geographical limitation of the other questions being asked. They show that this measure is highly correlated to the existing (aggregate) measures, i.e. the Cingranelli-Richards (CIRI) Human Rights Dataset (2014) indicator of women's economic (Wecon), the OECD Social Institutions and Gender Index (SIGI) and the Country Policy and Institutional Assessments (CPIA) equality index from the World Bank.

2.3.3 Leaving-family costs

Migration and the resulting separation from family creates an opportunity cost, in which migrants may forfeit their ability to raise their children and make decisions regarding the future of them. Thus, it is necessary to include the leaving-family costs as another indirect cost component when assessing the overall costs of migration. This measure seeks to capture the expenses associated with not being involved in the family's decision-making process related to one's children. I consider preferences on education and marriage age to be a robust predictor of the leaving-family cost. A woman who highly values the education of her children and especially her daughters as well as the marriage age of her daughters is therefore expected to experience higher migration costs. Therefore, I hypothesize:

Hypothesis 4 *Leaving-family costs decrease migration intentions for women.*

To test this hypothesis, I ask respondents a variety of questions about the optimal level of human capital of their children as well as questions on the marriage age of girls.

In general, what is the minimum level of education you would find acceptable for your son(s) to reach? 1=primary, 3=postsecondary to

In general, what is the minimum level of education you would find acceptable for your daughter(s) to reach? 1=primary, 3=postsecondary to

In general, do you agree or disagree that the girls in your households should get the same level of education as the boys? 1= agree very strongly to 5= disagree very strongly

What is the minimum age at which you will agree to marry off your daughters? number

In your opinion, how much do you agree with the following statement: Early marriage of girls should be banned by law. 1=totally agree to 6=totally disagree

2.3.4 Social norm costs

In general, individuals may choose to migrate through regular or irregular channels. It is plausible that engaging in irregular migration incurs a cost associated with violating social norms. Individuals who exhibit a strong adherence to the law are therefore expected to be less likely to migrate. There exists research indicating that there may be gender differences in morality, but findings are mixed and can vary depending on the context and culture being studied (Jaffee & Hyde, 2000). As that paper discusses, some studies suggest that women are more likely to prioritize care and fairness, whereas men are more likely to prioritize justice and rights. However, other studies have found no significant gender differences in moral reasoning. Therefore, the gendered perspective will be neglected and I only hypothesize:

Hypothesis 5 *Social norm costs decrease migration intentions.*

Following Ryo (2013), I ask the following survey items to capture the costs of migrating illegally. However, while the former study only includes the male perspective,

I can investigate a gendered perspective on morality in the decision to migrate.

In your opinion, how much do you agree with the following statements?

Disobeying the law is sometimes justified.	1=totally agree to 6=totally disagree
--	--

It is okay to disobey the law when you disagree with it.	1=totally agree to 6=totally disagree
--	--

It is okay to enter another country without papers if you need money to support one's family.	1=totally agree to 6=totally disagree
---	--

Other countries have the right to limit the number of immigrants who can enter the country.	1=totally agree to 6=totally disagree
---	--

If you were to go to live or work in a richer country with papers, do you think your family would...?	1=completely approve to 6=completely disapprove
---	--

If you were to go to live or work in a richer country without papers, do you think your family would...?	1=completely approve to 6=completely disapprove
--	--

2.4 Data sampling and collection

This paper is part of a bigger research project of the Mercator Dialogue on Asylum and Migration for which we gathered primary survey data at the individual level in Uganda and Senegal. While these two countries do not represent the entire Sub-Saharan African region, they serve as valuable case studies from relevant areas and share common features with other countries in the region. Our analysis focused on Senegal because it is one of the most important countries of origin for both regular and irregular migrants to Europe, especially from Sub-Saharan Africa. The percentage of Western African migrants in Europe has grown significantly from 12% in the mid-1990s to 19% in the mid-2020s. This trend is driven by factors such as common languages and residual colonial ties (United Nations, Department of Economic and Social Affairs, Population Division, 2020). To understand differences in migration preferences, we also examined Uganda, which shares

some migration-related characteristics with Senegal such as a common language with a former colonizer and widespread dissatisfaction in the population that may act as push factors. However, Uganda has distinct migration patterns compared to Senegal, with a lower share of emigration and very few irregular migrants but a rising trend of legal migration to Gulf countries.

Our survey of 1,204 individuals in Uganda took place in October and November 2021, while the survey of 1,500 individuals in Senegal was conducted between February and April 2022. Our sampling strategy aimed to capture a representative sample of the mobile population in each country. In collaboration with our local partners, we selected a representative sample of enumeration areas, based on census districts. These enumeration areas were somewhat clustered in different regions in each country for logistic reasons. Hence, not every province of the country is included in the sample but it should still be representative of the differences in the different countries. In Uganda, we interviewed individuals from 36 enumeration areas located in the districts of Kampala, Mbale, Gulu, Wakiso, Masaka, and Mbarara (see Appendix Figure 2.A1). In Senegal, we conducted interviews with individuals from 60 enumeration areas located in the districts of Dakar, Diourbel, Matam, Saint Louis, Tambacounda, and Ziguinchor (see Appendix Figure 2.A2).

In each enumeration area, we then conducted a complete household listing. The listing gives us an overview of the basic demographic characteristics of all households and household members in an enumeration area (typically about 100-120 households). Then, we randomly sampled individuals aged between 18 and 40 years for the main interview. This age group constitutes around 30% of the population in both Uganda and Senegal (World Bank, 2022). We achieved a response rate of 99% for the main interview in Uganda and 75% in Senegal. Based on the listing data, we do not find systematic differences between participants of the survey and those who refused or could not be found. Each survey, including this questionnaire and experiments used in other papers, lasted between 90 to 150 minutes. We collected data on household and individual-level socioeconomic characteristics, individual migration intentions, and attitudes toward immigrants.

Out of a total of 2,704 respondents, 760 participants have to be dropped from the analysis due to missing answers in relevant variables (249 from Uganda and 511 from Senegal).⁴ The sample covers both rural and urban areas. Overall women comprised 56.1% of the sample because of a skewed gender ratio in rural areas, especially in Uganda.

⁴The working sample is representative of the overall sample since the missing values do not follow a pattern, see Appendix 2.A1.

2.5 Analysis and results

2.5.1 Summary Statistics

Table 2.2 summarizes the respondents' socio-demographic characteristics and highlights the differences between male and female respondents for the working sample of 1,944 participants. The average participant is 27.8 years old, with a slightly higher age for women. A significant proportion of the sample (44.2%) has only primary education and below, primarily driven by female respondents (39.2% of men and 48.1% of women). 39.8% and 16.0% have at least secondary education and tertiary education or vocational training, respectively. Notably, the female sample is less educated than the male sample (41.1% of men and 38.8% of women have at least some secondary education, while 19.7% of men and 13.1% of women have tertiary education or training). The average monthly household income is \$145, while the average monthly income per household member is \$39.5. 41% of respondents report having family members, relatives, or friends who live in another country. 58% of respondents are married and the average number of children per participant is 1.79. Both measures are significantly higher for female as compared to male participants. 73% of the respondents would like to move to another country given the opportunity and these migration intentions are significantly higher for men. These gender differences in migration intentions are the focal point of this study.

Table 2.2: Summary statistics and gender differences for the working sample

	N	Gender differences		
		Male	Female	p.overall
	1944	853	1091	
SEN	989 (50.9%)	480 (56.3%)	509 (46.7%)	
UGA	955 (49.1%)	373 (43.7%)	582 (53.3%)	
Age	27.8 (6.59)	27.6 (6.51)	27.9 (6.66)	0.318
Education:				<0.001
Completed primary and below	859 (44.2%)	334 (39.2%)	525 (48.1%)	
At least some secondary	774 (39.8%)	351 (41.1%)	423 (38.8%)	
Tertiary or training	311 (16.0%)	168 (19.7%)	143 (13.1%)	
HH income in \$	145 (587)	152 (330)	139 (728)	0.606
PP HH income in \$	39.5 (150)	42.4 (93.7)	37.2 (183)	0.423
Contact abroad	0.41 (0.49)	0.47 (0.50)	0.37 (0.48)	<0.001
Married	0.58 (0.49)	0.49 (0.50)	0.65 (0.48)	<0.001
Number of children	1.79 (2.02)	1.30 (1.93)	2.18 (2.01)	<0.001
Migration intentions	0.73 (0.45)	0.76 (0.43)	0.70 (0.46)	0.002

Note: The marital status variable was assessed via the household listing. NA is re-coded as not married. Results are robust to this re-coding, compare Table 2.6 and Table 2.A19.

2.5.2 Empirical approach

The analysis investigates how different cost factors affect migration intentions. Since the outcome variable is binary, a logistic model is applied.⁵ In the subsequent analysis, different cost factors explaining migration intentions are successively added to the regression. Specifically, I create an index for each indirect cost factor by using the answers given for each respective factor's set of questions. To achieve this, I sum the standardized answers while assigning equal weight to each of them. As a result, I obtain a single index for the four cost factors, which I can subsequently incorporate into my analysis. One major limitation of this method is that it assumes an equal contribution of each question to the indirect cost factor. To address this issue, I proceed by conducting a PCA, which allows me to relax this assumption.

PCA is a dimension reduction technique that identifies a new set of variables, called principal components, which are linear combinations of the original variables while retaining most of the variability in the data. In this context, the aim is to identify principal components that can serve as representative measures of the indirect costs of migration. As PCA can only manage numerical variables, in the following analysis, all ordinal data is transformed into numeric data. PCA is based on the covariance matrix of the data, which is a symmetric matrix X that describes the degree of the linear relationship between the standardized variables (the matrix X with n observations and p variables). Given the matrix X with dimensions $n \times p$, the covariance matrix C is defined as follows: $C = (1/(n - 1)) * X' X$ where X' is the transpose of X . The principal components are extracted by projecting the data onto the eigenvectors. The eigenvectors represent the axes in the p -dimensional space that capture the most variability in the data, while the eigenvalues quantify the amount of variability captured by each eigenvector. The principal components are given by: $PC = X * S$, where S is the matrix of eigenvectors of C , ordered in decreasing order of their corresponding eigenvalues. The proportion of variability explained by each principal component can be computed as the ratio of its eigenvalue to the total sum of eigenvalues. Deciding on the number of principal components is essential. I follow Kaiser's criterion, according to which one should retain only those principal components with eigenvalues greater than one (Yeomans & Golder, 1982). This way the amount of variance extracted by each component is at least as large as the variance of one observed variable. Finally, it is important to note that the results cannot be causally interpreted. However, this study is the first to decompose various indirect factors associated with migration costs, based on gender, and offers valuable insights into the gender disparity observed in migration intentions. The findings of this study also open up new avenues for future research in this field.

⁵The results are robust to running a Probit and OLS model, see Appendix Tables 2.A11, 2.A12, 2.A13 and 2.A14.

2.5.3 Analysis

To study migration intentions and variation according to gender, I start by running a logistic regression model with the dependent binary migration intentions variable on a gender as well as a country dummy. The results indicate that being a woman is significantly negatively associated with the likelihood of having migration intentions (see Appendix Table 2.A2). Respondents from Uganda are, however, significantly more likely to express migration intentions⁶. As logistic regression coefficients have limited interpretability, I am relying on the average marginal effect (AME) derived from the model. Analysis of the AME reveals a significant gender gap in migration decision-making. The predicted probability of having migration intentions for women is 8.0% lower as compared to their male counterparts. In the subsequent analysis, I successively add different potential cost factors to the regression analysis and investigate whether they can explain part of the gendered migration intention gap.

Direct migration costs

In the first step of exploring the gendered migration intention gap, the direct costs of migration are considered. Table 2.3 presents the average expected migration cost of Ugandan/Senegalese migrants differentiated by gender and marital status. Remarkably, the findings contradict the expected outcome of Hypothesis 1. In our study population, unmarried women exhibit the lowest predicted migration costs, in stark contrast to the actual migration costs reported in World Bank (2017), where unmarried women have among the highest costs.

Table 2.3: Expected direct cost of migration in \$

		Men	Women	p.overall
European country	Married	2888.76	2142.46	<0.001
	Single	3029.59	2088.26	<0.001
Observations		706	805	
Gulf country	Married	1353.38	1028.59	0.020
	Single	1313.32	917.09	0.002
Observations		648	781	

As previously mentioned, the analysis considers the expected migration expenses rather than the actual ones. Hence, the disparities in anticipated migration expenses

⁶The coefficient for female is -0.422 and for Uganda 0.982. Both variables are significant at the 0.01 confidence level, see Appendix Table 2.A2

could be attributed to an increased awareness of the actual migration costs that arise with an increase in migration probabilities. When considering migration to Europe in a logistic regression, the expected cost coefficient is positive and statistically significant (Table 2.4 Model 1 and 2). This supports the argument that individuals who are more inclined to express their intent to migrate anticipate greater costs associated with migration in some contexts. For migration to the Gulf, there is no significant correlation between migration intentions and the expected cost of migration.

The difference in significance may be explained by the contrasting migration patterns to the two regions. Unauthorized migration is more common among individuals migrating to European countries, whereas legal labor migration sponsored by employers predominates in the Gulf. Under this sponsorship arrangement, sponsors bear the cost of travel and provide housing, and in return, they exert control over the migrant's employment and mobility (Robinson, 2022). As a result the financial burden of the journey on the migrant is reduced, potentially diminishing the significance of such expenses.

Table 2.4: Regression results including expected migration costs

	<i>Dependent variable: Migration intentions</i>			
	(1)	(2)	(3)	(4)
Female	-0.401*** (0.130)	-1.053 (0.872)	-0.296** (0.134)	-1.058 (0.840)
Married	-0.352*** (0.130)	0.978 (0.868)	-0.338** (0.133)	0.905 (0.858)
Uganda	1.042*** (0.137)	1.056*** (0.138)	0.976*** (0.136)	0.995*** (0.136)
Log exp. migcost European country	0.150** (0.061)	0.218* (0.114)		
Log exp. migcost Gulf			0.054 (0.063)	0.114 (0.122)
Female:log exp. migcost European country		0.089 (0.117)		
Married:log exp. migcost European country		-0.183 (0.118)		
Female:log exp. migcost Gulf				0.116 (0.127)
Married:log exp. migcost Gulf				-0.191 (0.131)
Constant	-0.039 (0.490)	-0.523 (0.866)	0.661 (0.456)	0.267 (0.822)
Observations	1,511	1,511	1,429	1,429
Log Likelihood	-787.158	-785.824	-737.128	-735.838

Note: *p<0.1; **p<0.05; ***p<0.01, This table is based on a logit model and reports the coefficients, dependent variable: dummy equal to 1 when respondent has migration intentions

The interaction between the gender dummy and the migration costs variable is not significant (Columns 2 and 4). Thus, there is no indication that female migrants have distinct evaluations of migration costs and the incorporation of the direct cost variable falls short in explaining the variations in the gendered migration intentions.⁷ Thus, in the subsequent sections, I shift the analysis toward the indirect costs associated with migration.⁸

⁷The large effect size of the female coefficient in Column 2 and 4 when introducing the interaction effects seem to be due to multicollinearity. Hence, I do not explore these coefficients further.

⁸In Appendix 2.A3, I have included the expected income differential in the regression analysis. Despite this addition, the gender-based differences in migration intentions remain unexplained.

Indirect cost of migration

As a first step in analyzing indirect costs, I control for several important factors that are generally known to mainly indirectly affect the costs of migration (Migali & Scipioni, 2019). These factors included age, education, and the logarithm of per person household income⁹. Additionally, I included a variable indicating whether the respondent has any family members, relatives, or friends who live in another country (referred to as "Contact abroad"). By including this variable, I can factor in the influence of networks abroad on migration intentions.

When controlling for potential confounding factors that might drive the gender migration intention gap, significant gender differences in intentions to migrate persists (Table 2.5). Women as compared to men have a 6.3% lower predicted probability of having migration intentions, holding all other factors constant and including district fixed effects (see Appendix Table 2.A4 Model 2).

⁹Age is included in a linear form; see Appendix 2.A3 for a visualization of the relationship between migration intentions and age.

Table 2.5: Regression results baseline model

	<i>Dependent variable: Migration intentions</i>		
	(1)	(2)	(3)
Female	-0.400*** (0.112)	-0.354*** (0.115)	-0.325*** (0.123)
Age	-0.038*** (0.008)	-0.040*** (0.008)	-0.041*** (0.009)
At least some secondary education	0.482*** (0.127)	0.515*** (0.135)	0.458*** (0.146)
Tertiary education or training	0.295* (0.168)	0.334* (0.175)	0.210 (0.191)
Log PP HH income in \$	-0.061* (0.032)	-0.072** (0.033)	-0.052 (0.036)
Contact abroad	0.173 (0.113)	0.198* (0.116)	0.171 (0.125)
Constant	1.703*** (0.265)	1.757*** (0.345)	2.850*** (0.838)
Fixed effects	Country	District	EA
Observations	1,944	1,944	1,944
Log Likelihood	-1,063.860	-1,038.513	-988.120

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. This table is based on a logit model and reports the coefficients, for version reporting average marginal effects, see Appendix Table 2.A4, dependent variable: dummy equal to 1 when respondent has migration intentions, reference category: completed primary education and below

To expand the analysis and obtain a more nuanced understanding of the factors affecting migration intentions, I incorporate information on the family situation into the regression. This includes the marital status of the respondent as well as their parenthood status. The marital status is represented as a binary variable indicating whether the respondent is married or not. To capture the effects of parenthood, I include different measures in the regression, such as the number of children, the logarithm of the number of children, and a binary variable indicating whether the respondent has any children (Table 2.6 Models 2 to 4). It is possible that the number of children one provides care for, rather than the number of children as such, has an impact on migration decisions. Children one provides care for are typically younger and require more care and support from their parents, which can make migration decisions more complex and challenging. For example, parents with dependent children may be more reluctant to migrate due to concerns about disrupting their children's education or social networks. Additionally, the number of dependent children may affect the economic resources available to the

family, which can influence the feasibility of migration. To account for this, Models 5 to 7 of Table 2.6 include a variable on the number of children the respondent provides care for, the logarithm of that number, as well as a dummy variable indicating whether the respondent provides care for any children, respectively. Models 8 to 11 of Table 2.6 include the different variables on parenthood as well as a dummy of having children (the respondent provides care for).

Based on the results of Table 2.6, I include the logarithm of the number of children in the subsequent analysis (Column 3). This variable has the greatest impact in decreasing the absolute size of the gender gap coefficient and is the most significant variable, with a p-value of 0.012. Equivalently, for the family cost, I use the logarithm of the number of elderly and sick family members the respondent provides care for.

Table 2.6: Regression results including marital status and parenthood

	<i>Dependent variable: Migration intentions</i>										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Female	-0.332*** (0.117)	-0.276** (0.119)	-0.245** (0.122)	-0.263** (0.121)	-0.294** (0.118)	-0.264** (0.120)	-0.270** (0.121)	-0.250** (0.122)	-0.248** (0.122)	-0.270** (0.121)	-0.267** (0.121)
Married	-0.160 (0.122)	-0.097 (0.125)	-0.062 (0.128)	-0.082 (0.128)	-0.131 (0.123)	-0.105 (0.125)	-0.107 (0.125)	-0.068 (0.128)	-0.066 (0.129)	-0.109 (0.125)	-0.108 (0.125)
Age	-0.036*** (0.009)	-0.023** (0.010)	-0.023** (0.010)	-0.028*** (0.010)	-0.028*** (0.010)	-0.028*** (0.010)	-0.030*** (0.009)	-0.021** (0.010)	-0.021** (0.011)	-0.027*** (0.010)	-0.027*** (0.010)
At least some secondary education	0.501*** (0.136)	0.459*** (0.137)	0.450*** (0.137)	0.466*** (0.137)	0.476*** (0.136)	0.473*** (0.136)	0.482*** (0.136)	0.448*** (0.137)	0.448*** (0.137)	0.472*** (0.136)	0.472*** (0.136)
Tertiary education or training	0.315* (0.176)	0.242 (0.178)	0.230 (0.179)	0.255 (0.178)	0.292* (0.176)	0.283 (0.176)	0.288 (0.176)	0.225 (0.179)	0.227 (0.179)	0.284 (0.176)	0.283 (0.176)
Log PP HH income in \$	-0.072** (0.033)	-0.076** (0.033)	-0.071** (0.033)	-0.068** (0.033)	-0.075** (0.033)	-0.074** (0.033)	-0.072** (0.033)	-0.073** (0.034)	-0.073** (0.034)	-0.074** (0.033)	-0.075** (0.033)
Contact abroad	0.195* (0.116)	0.195* (0.116)	0.190 (0.116)	0.189 (0.116)	0.203* (0.116)	0.200* (0.116)	0.196* (0.116)	0.191 (0.117)	0.192* (0.117)	0.201* (0.116)	0.201* (0.116)
Number kids		-0.083** (0.033)						-0.065* (0.038)			
Log number kids			-0.118** (0.047)						-0.188 (0.119)		
Kids				-0.309** (0.151)				-0.171 (0.172)	0.246 (0.381)		
Care kids number					-0.056** (0.027)					-0.039 (0.032)	
Log care kids number						-0.090** (0.040)					-0.127 (0.109)
Care kids							-0.270** (0.138)			-0.158 (0.167)	0.134 (0.377)
Constant	1.730*** (0.347)	1.454*** (0.364)	1.219*** (0.402)	1.609*** (0.352)	1.583*** (0.353)	1.406*** (0.375)	1.680*** (0.348)	1.447*** (0.365)	1.012** (0.515)	1.600*** (0.354)	1.300*** (0.478)
Fixed effects	District	District	District	District	District	District	District	District	District	District	District
Observations	1,944	1,944	1,944	1,944	1,944	1,944	1,944	1,944	1,944	1,944	1,944
Log Likelihood	-1,037.656	-1,034.580	-1,034.483	-1,035.538	-1,035.453	-1,035.110	-1,035.717	-1,034.087	-1,034.275	-1,035.005	-1,035.047

Note: *p<0.1; **p<0.05; ***p<0.01, This table is based on a logit model and reports the coefficients, dependent variable: dummy equal to 1 when respondent has migration intentions, reference category: completed primary education and below

2.5.4 Index for indirect migration costs

The subsequent step involves systematically examining the 20 questions related to indirect cost factors as part of the manually-built indices for family obligation, gender discrimination, leaving-family, and social norm costs.

Models 1 to 6 of Table 2.7, which differ in the incorporation of interaction terms between the gender dummy and the indirect cost factors, as well as the inclusion of fixed effects (country, district, enumeration area (EA)), demonstrate that the inclusion of all indirect cost factors leads to a decrease in the coefficient of the gender dummy on migration intentions. Furthermore, when including an interaction term and accounting for district or enumeration area fixed effects, the female coefficient is insignificant (Models 5 and 6). The predicted probability of having migration intentions is only 4.0% lower for women as compared to men controlling for district fixed effects (Appendix Tables 2.A5 Model 5) (compared to 6.3% in the prior model).

The impact of family obligation costs on migration intentions is significant and impacts men and women equally. It indicates that individuals who prioritize family over their individual needs, including those who are married, have more children, and are caregivers for elderly or sick family members, are less likely to express their intention to migrate.

Although it was expected that the gender discrimination variable would have an insignificant impact in the models without interactions, it is surprising that the interaction between the female variable and the gender discrimination costs variable is also insignificant. These findings suggest that women's migration intentions are not correlated with their perception of feeling treated with disrespect or without dignity, unfairly at work, or their perception of domestic violence as a serious issue - at least in the countries studied in this paper. These results hold up even when only using the question on feeling treated without respect and dignity, as conducted by Ruysen and Salomone (2018), and when constructing the index solely based on questions about the actual situation of gender discrimination as compared to the preferred situation (see Appendix 2.A15). Examining the analysis exclusively focusing on women, overall I also do not observe a significant impact of the gender discrimination variable on migration intentions (see Appendix 2.A16). As a result, the general pattern identified by Ruysen and Salomone (2018) may not apply to all countries, or at least not with the same level of statistical significance.¹⁰

¹⁰To investigate this further, I replicate the analysis conducted by Ruysen and Salomone (2018) and present the findings in Appendix Table 2.A17. Using Gallup data, I restrict the sample to women aged 18 to 40, which is consistent with my data. Columns 1 to 4 display the results for the worldwide sample, for Africa, and for Uganda and Senegal. Although the regression coefficient for gender discrimination for the global sample differs somewhat from those of Ruysen and Salomone (2018) due to my use of different survey waves, they are similar in terms of effect size. Notably, gender discrimination has a weaker impact in Africa, particularly in Uganda and Senegal, as compared to the worldwide sample (see Appendix Table 2.A18 for AMEs). Especially, while feeling treated with disrespect and without dignity increases the probability of having migration intentions by 5.2% for the worldwide sample, the average marginal effect is only 1.1% in Uganda and Senegal. Additionally, when restricting the analysis to the

In Models 1 to 3, the leaving-family cost variable has a positive and significant coefficient. However, upon including an interaction term with the female dummy in Models 4 to 6, it becomes evident that the interaction with gender is the primary driver of the results. This suggests that the impact of this component is highly heterogeneous and relevant only to women. Specifically, women who prioritize a higher level of education for both their sons and daughters, value education equality more, and prefer a higher marriage age for their daughters, as well as a ban on child marriage, have higher migration intentions.

This finding contradicts Hypothesis 4, which posits that leaving-family costs reduce migration intentions because women may have limited decision-making power regarding their children's future when migrating. One possible explanation is that women who plan to migrate intend to take their children with them, either to ensure that they retain decision-making power over their children's future or to provide them with better opportunities abroad. During the survey, we asked respondents with migration intentions who would accompany them if they migrated. 31.0% of women indicated they would migrate alone, whereas 45.8% planned to migrate with either all of their children or just their younger/older children. The correlation coefficient between the index of leaving-family costs and the desire to take children along is -0.012. This finding lacks evidence supporting the notion that women with greater leaving-family costs tend to migrate with their children.

The results support the hypothesis that the social norm costs variable has a significant impact on reducing migration intentions, equally affecting both men and women. This variable estimates the extent of adherence to the law, the acknowledgment of other countries' right to limit the number of migrants, and the expected attitudes of the respondent's family toward migration, regardless of its legality.¹¹

countries within the working sample, the results in the main specification are not statistically significant, possibly due to a smaller effect size and reduced sample size. These findings suggest that the conclusions drawn by Ruyssen and Salomone (2018) are context-dependent and may not hold to the same extent in Uganda and Senegal as they do globally. Finally, Column 5 presents the results of the same analysis with some variations in the measurement of certain variables using my primary data. The results show that feeling treated without respect and dignity decreases the predicted probability of women having migration intentions by 1.4%, holding all other variables constant.

¹¹For robustness of results see Appendix Table 2.A20.

Table 2.7: Regression results for the indices

	<i>Dependent variable: Migration intentions</i>					
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.291** (0.118)	-0.232* (0.121)	-0.199 (0.128)	-0.255** (0.118)	-0.195 (0.122)	-0.166 (0.129)
Age	-0.026*** (0.009)	-0.026*** (0.010)	-0.029*** (0.010)	-0.026*** (0.009)	-0.026*** (0.010)	-0.029*** (0.010)
At least some secondary education	0.399*** (0.131)	0.452*** (0.139)	0.414*** (0.149)	0.409*** (0.132)	0.464*** (0.139)	0.427*** (0.149)
Tertiary education or training	0.178 (0.174)	0.230 (0.180)	0.130 (0.195)	0.195 (0.175)	0.245 (0.181)	0.147 (0.196)
Log PP HH income in \$	-0.070** (0.033)	-0.079** (0.034)	-0.062* (0.037)	-0.076** (0.033)	-0.086** (0.035)	-0.069* (0.038)
Contact abroad	0.108 (0.115)	0.145 (0.118)	0.148 (0.127)	0.101 (0.116)	0.143 (0.119)	0.148 (0.128)
Family obligation	-0.276** (0.109)	-0.302*** (0.113)	-0.288** (0.120)	-0.253 (0.154)	-0.292* (0.158)	-0.277* (0.167)
Gender discrimination	-0.119 (0.101)	-0.162 (0.104)	-0.170 (0.111)	-0.106 (0.149)	-0.142 (0.152)	-0.118 (0.160)
Leaving family	0.252*** (0.093)	0.247** (0.097)	0.184* (0.103)	-0.034 (0.142)	-0.034 (0.144)	-0.135 (0.154)
Social norm	-0.626*** (0.118)	-0.655*** (0.122)	-0.644*** (0.129)	-0.542*** (0.186)	-0.512*** (0.189)	-0.474** (0.200)
Female:family obligation				-0.022 (0.194)	-0.006 (0.199)	-0.002 (0.212)
Female:gender discrimination				-0.031 (0.201)	-0.049 (0.204)	-0.106 (0.214)
Female:leaving family				0.499*** (0.183)	0.496*** (0.186)	0.561*** (0.199)
Female:social norm				-0.128 (0.241)	-0.229 (0.246)	-0.272 (0.261)
Constant	1.465*** (0.294)	1.393*** (0.376)	2.433*** (0.852)	1.428*** (0.298)	1.352*** (0.379)	2.421*** (0.855)
Fixed effects	Country	District	EA	Country	District	EA
Observations	1,944	1,944	1,944	1,944	1,944	1,944
Log Likelihood	-1,041.727	-1,015.148	-969.496	-1,037.799	-1,011.120	-964.774

Note: *p<0.1; **p<0.05; ***p<0.01, This table is based on a logit model and reports the coefficients, for version reporting average marginal effects, see Appendix Table 2.A5, dependent variable: dummy equal to 1 when respondent has migration intentions, reference category: completed primary education and below

2.5.5 PCA for indirect migration costs

In the following section, I lift the assumption of the equal contribution of all questions to the indirect cost measures and conduct a PCA to generate the components. For family obligation costs, a PCA using four principal components can capture all the variance in the variables (see Appendix Figure 2.A4). I incorporate the first principal component in the analysis, which is primarily determined by the logarithm of the number of children and the marital status, with a smaller contribution (12.26%) by the number of elderly and sick family members for whom care is provided (see Appendix Table 2.A6). Furthermore, the subsequent analysis includes the first two principal components of gender discrimination costs which account for over 60% of the total variance in the original data (Appendix Figure 2.A5). The first principal component is primarily determined by variables related to the optimal level of female equality (Appendix Table 2.A7). Specifically, respondents were asked to rate the importance of gender equality in the workplace and whether women are treated with respect and dignity. These questions explain over 90% of the contribution to the first principal component. The second principal component assesses respondents' experiences of gender discrimination by asking about fair treatment in the workplace and respect for women, as well as the issue of domestic violence to a lesser extent. The leaving-family costs' first principal component includes variables associated with the minimum acceptable level of education for children, the opinion on child marriage, and the importance of providing girls with equal education opportunities as boys (Appendix Table 2.A8). The second component assesses opinions on the appropriate age for marriage and considers the importance of girls receiving equal education opportunities, albeit to a lesser extent.

The analysis includes the first three principal components of the social norm costs variables that collectively account for over 65% of the total variance in the data (see Appendix Figure 2.A7). The first component relates to following the law (even when disagreeing with it) and the approval to enter another country illegally to support one's family (see Appendix Table 2.A9). The second component is associated with the respondent's family's opinion on working illegally in a richer country and the respondent's stance on other countries limiting legal immigration. Notably, 18% of the component pertains to whether it is acceptable to enter another country without legal papers for economic reasons. The third component is predominantly related to the respondent's family's opinion on legally working in a richer country. It is important to note that while all other components indicate an increase in indirect costs with an increase in the component, this component exhibits the opposite relationship. Specifically, an increase in the third principal component of the social norm costs indicates higher approval from the respondent's family for them to legally work in a richer country.

The principal component-based regression results without interactions align with the

manually constructed index (see Table as shown in Table 2.8 Columns 1 to 3). The findings suggest that family obligations have a significant negative impact on migration intentions, while the gender discrimination coefficient is insignificant. The study further reveals that leaving-family costs have a significant positive impact on migration intentions, while social norms costs have a significant negative influence, keeping the reversed effect of the third principal component in mind. Regarding the gender interaction coefficient, Models 4 to 6 reveal that leaving-family costs are primarily relevant to women, as in the index. However, in contrast to the index, the principal components also show a gendered relationship with the third social norm costs component.

Compared to the index, the use of principal components allows for a more detailed analysis of which individual variables are driving the observed correlations. The findings of the study indicate that family obligation costs significantly reduce migration intentions for both men and women. This implies that individuals who are married and have a higher number of dependents, including elderly and sick family members that require care, are less likely to express their intention to migrate.

As in the index, leaving-family costs have a gendered association and are primarily relevant to women. However, not all variables related to leaving-family costs reduce migration intentions for women; rather, they are the variables associated with the minimum acceptable level of education for children and the importance of providing equal educational opportunities for girls and boys that significantly influence migration intentions.

The second principal component of the social norm costs significantly reduces migration intentions for both men and women. This component primarily covers the respondent's family's opinion on working illegally in a richer country. Moreover, respondents' stance on other countries limiting legal immigration and entering another country without legal papers for economic reasons also contributes notably to the component. The positive coefficient of approval from the respondent's family for them to legally work in a richer country (the third component of social norm costs) is significant for both men and women but has an additional distinct effect for women. Overall, the inclusion of the principal components as well as interactions with gender reduces the gender migration intention gap to 3.5% (see Appendix Table 2.A10 Model 5).¹²

¹²For robustness of results see Appendix Table 2.A21.

Table 2.8: Regression results for the PCA

	<i>Dependent variable: Migration intentions</i>					
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.259**	-0.196	-0.151	-0.233*	-0.172	-0.137
	(0.122)	(0.125)	(0.133)	(0.123)	(0.127)	(0.134)
Age	-0.024**	-0.026**	-0.028***	-0.024**	-0.026**	-0.028**
	(0.010)	(0.010)	(0.011)	(0.010)	(0.010)	(0.011)
At least some secondary education	0.367***	0.411***	0.404***	0.364***	0.408***	0.404***
	(0.135)	(0.142)	(0.152)	(0.136)	(0.143)	(0.153)
Tertiary education or training	0.198	0.251	0.191	0.209	0.258	0.195
	(0.179)	(0.185)	(0.199)	(0.179)	(0.185)	(0.200)
Log PP HH income in \$	-0.070**	-0.084**	-0.065*	-0.078**	-0.090**	-0.074*
	(0.033)	(0.035)	(0.038)	(0.034)	(0.036)	(0.039)
Contact abroad	0.112	0.149	0.153	0.109	0.155	0.170
	(0.117)	(0.120)	(0.129)	(0.118)	(0.121)	(0.130)
Family obligation_1	-0.124**	-0.124**	-0.108*	-0.140*	-0.154**	-0.149*
	(0.057)	(0.058)	(0.062)	(0.075)	(0.078)	(0.082)
Gender discrimination_1	0.065	0.048	0.065	0.005	-0.007	0.035
	(0.047)	(0.048)	(0.052)	(0.067)	(0.069)	(0.073)
Gender discrimination_2	-0.027	-0.054	-0.055	0.019	-0.007	-0.008
	(0.046)	(0.048)	(0.050)	(0.072)	(0.073)	(0.076)
Leaving family_1	0.076*	0.077*	0.061	-0.035	-0.025	-0.055
	(0.040)	(0.042)	(0.044)	(0.064)	(0.065)	(0.068)
Leaving family_2	0.138**	0.106*	0.063	0.098	0.067	0.024
	(0.055)	(0.056)	(0.060)	(0.072)	(0.073)	(0.077)
Social norm_1	-0.048	-0.039	-0.025	-0.003	0.015	0.023
	(0.045)	(0.046)	(0.049)	(0.065)	(0.066)	(0.069)
Social norm_2	-0.097*	-0.087	-0.089	-0.192**	-0.183**	-0.169*
	(0.054)	(0.057)	(0.060)	(0.081)	(0.083)	(0.088)
Social norm_3	0.381***	0.425***	0.446***	0.195**	0.216**	0.217**
	(0.052)	(0.055)	(0.059)	(0.088)	(0.090)	(0.097)
Female:family obligation_1				0.050	0.074	0.093
				(0.095)	(0.097)	(0.104)
Female:gender discrimination_1				0.103	0.093	0.049
				(0.095)	(0.097)	(0.103)
Female:gender discrimination_2				-0.076	-0.082	-0.085
				(0.094)	(0.095)	(0.099)
Female:leaving family_1				0.188**	0.172**	0.198**
				(0.083)	(0.084)	(0.089)
Female:leaving family_2				0.103	0.102	0.112
				(0.099)	(0.102)	(0.108)
Female:social norm_1				-0.084	-0.102	-0.088
				(0.087)	(0.089)	(0.094)
Female:social norm_2				0.160	0.164	0.139
				(0.106)	(0.109)	(0.115)
Female:social norm_3				0.275**	0.312***	0.349***
				(0.110)	(0.114)	(0.122)

Constant	1.413*** (0.318)	1.405*** (0.401)	2.778*** (0.904)	1.400*** (0.324)	1.384*** (0.407)	2.670*** (0.894)
Fixed effects	Country	District	EA	Country	District	EA
Observations	1,944	1,944	1,944	1,944	1,944	1,944
Log Likelihood	-1,022.688	-994.389	-950.124	-1,014.332	-985.495	-940.894

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. This table is based on a logit model and reports the coefficients, for version reporting average marginal effects, see Appendix Table 2.A10, dependent variable: dummy equal to 1 when respondent has migration intentions, reference category: completed primary education and below

2.6 Discussion and conclusion

While there exists vast literature on drivers of migration, systematic knowledge about gender-specific factors is lacking. However, being able to understand gendered differences in migration is relevant for policymakers. By examining how gender impacts the decision to migrate, policymakers can develop focused interventions to improve migration outcomes and promote gender equality, while also addressing gender-based discrimination and enhancing the overall well-being of migrants and their families.

To shed light on gender differences in the decision to migrate, my study examines the expected costs of migration and various indirect cost factors that may impact migration decision-making and assesses how they affect men and women respectively. I do not find any significant gender differences in the expected direct cost of migration. To assess the impact of indirect costs, I focus on four particular cost components - family obligation, gender discrimination, leaving-family, and social norm costs - and construct indicators for each using both theoretical knowledge to create an index and principal component analysis.

Both of these approaches indicate that family obligation costs reduce migration intentions for both men and women. While family ties potentially complicate the decision to migrate for both genders, Hypothesis 2 proposed that these costs would be more salient for women, given the societal expectations placed on them as primary caregivers in the family. As a result of these gender norms, women are often left behind to take care of the family while men, perceived as the primary breadwinners, migrate and remit funds back home (De Jong, 2000). The lack of evidence supporting Hypothesis 2 could be attributed to the possibility that the survey questions employed in this study did not effectively capture the intended concept. Future research should examine the gender roles of individuals by assessing their views on who is responsible for providing care and who is responsible for generating income in the household.

Despite Hypothesis 3 stating that gender discrimination motivates women to consider migration, my findings suggest otherwise. In particular, I do not find any significant correlation between gender discrimination and migration intentions for women. This

implies that the general pattern found by Ruysen and Salomone (2018) does not hold in all countries, at least at the same level of statistical significance. The reason for this may be that context is important for whether women are considering migrating in response to gender discrimination. In some contexts discrimination may not lead to women considering migration while in others other drivers of migration may be more important, hence making discrimination insignificant. Further research is required to fully comprehend the interplay between discrimination and migration in different settings.

The leaving-family variable shows a positive and significant relationship with migration intentions among women in both models. Surprisingly, while Hypothesis 4 predicts that women with higher leaving-family costs have lower migration intentions, the opposite holds. The PCA analysis reveals that the questions regarding the educational attainment of children are the main drivers of these results. While the leaving-family cost variable aimed to capture the cost associated with the loss of decision-making power over the future of children due to migration, the variables used in the models may capture a different aspect. The findings suggest that women who prioritize education equality and aspire for higher education for both their sons and daughters are more likely to have migration intentions. This may indicate that they view migration as a means to make education more feasible for their children since secondary and tertiary education is not free in Uganda and Senegal (Petroni et al., 2017). An alternative explanation could be that the questions used do not capture the costs associated with leaving family behind but instead reflect internalized gender roles. Respondents who prioritize their daughters' education as highly as their sons' and encourage them to stay in school longer are more likely to have a non-traditional and gender-equal mindset. Therefore, the results may hint that more empowered women have higher migration intentions. However, further investigations are needed to confirm this interpretation.

Hypothesis 5 suggests that social norm costs decrease migration intentions for both men and women. However, this study highlights that this is only partially true. A principal component heavily influenced by the attitudes of the respondent's family toward working illegally in a richer country and the respondents' stance on restrictions to legal immigration and illegal entry into another country has a significant impact on migration intentions for both genders. While the family's opinion on legally working in a wealthier country holds significance for both men and women, there exists an additional effect specifically relevant to women. This highlights the existence of distinct differences between the cost of illegality, which is relevant for both genders, and the impact of the family's opinion on legal migration, which is particularly relevant to women. While this study offers valuable insights, further research is required to explore the drivers of social norm differences between men and women.

The incorporation of all four indirect cost factors in the analysis, while considering gender heterogeneity, reduces the gender gap in migration intentions by almost 45%. Ad-

ditionally, the effect of the gender coefficient is diminished to insignificance. While this effect cannot be causally interpreted it still provides valuable insights into the observed gender disparity in migration intentions by decomposing the various indirect factors associated with migration costs, based on gender. Additionally, it is important to note that the scope of this paper is limited to investigating migration intentions and does not include the actual migration decision. Therefore, the gender gap observed in migration intentions may not fully account for differences in actual migration behavior. However, research has consistently shown that migration intentions are strong predictors of future migration behavior (Tjaden et al., 2019), suggesting that the factors driving a person's willingness to migrate also influence their decision to do so (Van Dalen et al., 2005). One can also apply the same logic as in Chort (2014) and consider migration intentions as rational, meaning that they are correlated with the same factors that determine actual migration decisions. As such, it can be assumed that the implications of my findings can also provide insights into actual migration decisions. Furthermore, women exhibit a lower tendency than men to put their migration intentions into action (Chort, 2014). This suggests that the gender migration intention gap identified in this study may be a conservative estimate of the actual gender migration gap, and the results could potentially offer even greater explanatory power.

Understanding the factors that drive gendered migration patterns can inform targeted policies and interventions aimed at promoting gender equality and enhancing the well-being of migrants and their families. For example, my findings suggest the need for policies that address family obligation costs, such as offering financial incentives to encourage families to migrate together. Programs that aim to strengthen family bonds, such as family reunification schemes or social protection programs, could also be implemented to mitigate the negative effects of family obligations on migration intentions. Additionally, the study highlights the importance of addressing social norms that disproportionately affect women's migration intention. Efforts should be made to address social norms that discourage legal migration. This could be achieved by raising awareness about the benefits of legal migration and promoting a culture of respect for the law. Programs that provide information on legal migration channels and offer support for migrants during the migration process could also be implemented. By identifying these specific factors and their differential impact on men and women, policymakers can develop more effective and equitable policies to support migrant populations.

Bibliography

- Abel, G. J., & Cohen, J. E. (2022). Bilateral international migration flow estimates updated and refined by sex. *Scientific Data*, 9(1), 173.
- Bah, T. L., Batista, C., et al. (2018). Understanding willingness to migrate illegally: Evidence from a lab in the field experiment. *Nova School of Business and Economics Working Paper*, (1803).
- Bertrand, M., Goldin, C., & Katz, L. F. (2010). Dynamics of the gender gap for young professionals in the financial and corporate sectors. *American Economic Journal: Applied Economics*, 2(3), 228–255.
- Blau, F. D., & Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, 55(3), 789–865.
- Böhme, M. H., Persian, R., & Stöhr, T. (2015). Alone but better off? Adult child migration and health of elderly parents in Moldova. *Journal of Health Economics*, 39, 211–227.
- Borjas, G. J. (1989). Economic theory and international migration. *International Migration Review*, 23(3), 457–485.
- Budig, M. J., & England, P. (2001). The wage penalty for motherhood. *American Sociological Review*, 66(2), 204–225.
- Chort, I. (2014). Mexican migrants to the us: What do unrealized migration intentions tell us about gender inequalities? *World Development*, 59, 535–552.
- Cooke, T. J. (2008). Migration in a family way. *Population, Space and Place*, 14(4), 255–265.
- Cortes, P. (2015). The feminization of international migration and its effects on the children left behind: Evidence from the Philippines. *World Development*, 65, 62–78.
- Croson, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic literature*, 47(2), 448–474.
- De Jong, G. F. (2000). Expectations, gender, and norms in migration decision-making. *Population Studies*, 54(3), 307–319.
- Duflo, E. (2003). Grandmothers and granddaughters: Old-age pensions and intrahousehold allocation in South Africa. *The World Bank Economic Review*, 17(1), 1–25.

- Dufflo, E. (2012). Women empowerment and economic development. *Journal of Economic literature*, 50(4), 1051–1079.
- Dumont, J.-C., Martin, J. P., & Spielvogel, G. (2007). Women on the move: The neglected gender dimension of the brain drain. *IZA Discussion Paper*, (2920).
- Dustmann, C., Fasani, F., Meng, X., & Minale, L. (2023). Risk attitudes and household migration decisions. *Journal of Human Resources*, 58(1), 112–145.
- Ferrant, G., & Tuccio, M. (2015). South–south migration and discrimination against women in social institutions: A two-way relationship. *World Development*, 72, 240–254.
- Goldin, C., & Katz, L. F. (2002). The power of the pill: Oral contraceptives and women’s career and marriage decisions. *Journal of Political Economy*, 110(4), 730–770.
- Gubhaju, B., & De Jong, G. F. (2009). Individual versus household migration decision rules: Gender and marital status differences in intentions to migrate in South Africa. *International Migration*, 47(1), 31–61.
- Gutmann, J., Marchal, L., & Simsek, B. (2023). Women’s rights and the gender migration gap. *ILE Working Paper Series*, (10222).
- Hidrobo, M., Mueller, V., & Roy, S. (2022). Cash transfers, migration, and gender norms. *American Journal of Agricultural Economics*, 104(2), 550–568.
- Hughes, C., Bhandari, P., Young-DeMarco, L., Swindle, J., Thornton, A., & Williams, N. (2020). Family obligation attitudes, gender, and migration. *International Journal of Sociology*, 50(4), 237–264.
- Hugo, G. (2000). *Migration and women’s empowerment*. Oxford University Press.
- Jaffee, S., & Hyde, J. S. (2000). Gender differences in moral orientation: A meta-analysis. *Psychological Bulletin*, 126(5), 703.
- Jolly, S., Reeves, H., & Piper, N. (2005). Gender and migration: Overview report. *Trabajo y empleo*.
- Lauby, J., & Stark, O. (1988). Individual migration as a family strategy: Young women in the Philippines. *Population Studies*, 42(3), 473–486.
- Massey, D. S., Fischer, M. J., & Capoferro, C. (2006). International migration and gender in Latin America: A comparative analysis. *International Migration*, 44(5), 63–91.
- McKenzie, D., & Rapoport, H. (2011). Can migration reduce educational attainment? Evidence from Mexico. *Journal of Population Economics*, 24, 1331–1358.
- Migali, S., & Scipioni, M. (2019). Who’s about to leave? A global survey of aspirations and intentions to migrate. *International Migration*, 57(5), 181–200.
- Migration Data Portal. (2023). Gender and migration [Downloaded from [https://www.migrationdataportal.org/themes/gender-and-migration] on 07 March 2023].
- Munshi, K., & Rosenzweig, M. (2016). Networks and misallocation: Insurance, migration, and the rural-urban wage gap. *American Economic Review*, 106(01), 46–98.

- O'Neil, T., Fleury, A., & Foresti, M. (2017). Gender equality, migration and the 2030 agenda for sustainable development. *Louise Arbour, UN Special Representative for International Migration*, 45.
- Petroni, S., Steinhaus, M., Fenn, N. S., Stoebenau, K., & Gregowski, A. (2017). New findings on child marriage in Sub-Saharan Africa. *Annals of Global Health*, 83(5-6), 781–790.
- Pfeiffer, L., & Taylor, J. E. (2007). Gender and the impacts of international migration: Evidence from rural Mexico. In *The international migration of women* (Morrison, A. R., Schiff, M. and Sjöblom, M., pp. 99–124). World Bank Publications.
- Robinson, K. (2022). What is the Kafala system? [Downloaded from [<https://www.cfr.org/backgrounders/what-kafala-system>] on 06 April 2023]. *Council on foreign relations*.
- Ruyssen, I., & Salomone, S. (2018). Female migration: A way out of discrimination? *Journal of Development Economics*, 130, 224–241.
- Ryo, E. (2013). Deciding to cross: Norms and economics of unauthorized migration. *American Sociological Review*, 78(4), 574–603.
- Shrestha, M. (2020). Get rich or die tryin': Perceived earnings, perceived mortality rates, and migration decisions of potential work migrants from Nepal. *The World Bank Economic Review*, 34(1), 1–27.
- Stöhr, T. (2015). Siblings' interaction in migration decisions: Who provides for the elderly left behind? *Journal of Population Economics*, 28, 593–629.
- Tjaden, J., Auer, D., & Laczko, F. (2019). Linking migration intentions with flows: Evidence and potential use. *International Migration*, 57(1), 36–57.
- Todaro, M. (1980). Internal migration in developing countries: A survey. In *Population and economic change in developing countries* (pp. 361–402). University of Chicago Press.
- United Nations, Department of Economic and Social Affairs, Population Division. (2020). International migrant stock [Downloaded from [<https://www.un.org/development/desa/pd/content/internationalmigrant-stock>] on 30 March 2023].
- Van Dalen, H. P., Groenewold, G., & Schoorl, J. J. (2005). Out of Africa: What drives the pressure to emigrate? *Journal of Population Economics*, 18, 741–778.
- World Bank. (2017). KNOMAD-ILO Migration Costs Surveys 2016 [Downloaded from [<https://microdata.worldbank.org/index.php/catalog/2944>] on 08 March 2023].
- World Bank. (2020). *Women, business and the law 2020*. International Bank for Reconstruction; Development.
- World Bank. (2022). World development indicators [Downloaded from [<https://databank.worldbank.org/source/world-development-indicators>] on 02 December 2022].

Yeomans, K. A., & Golder, P. A. (1982). The Guttman-Kaiser criterion as a predictor of the number of common factors. *The Statistician*, 221–229.

Zachariah, K. C., Mathew, E. T., & Rajan, S. I. (2001). Impact of migration on Kerala's economy and society. *International Migration*, 39(1), 63–87.

Appendix

2.A.1 Study area

Figure 2.A1: Interview locations in Uganda

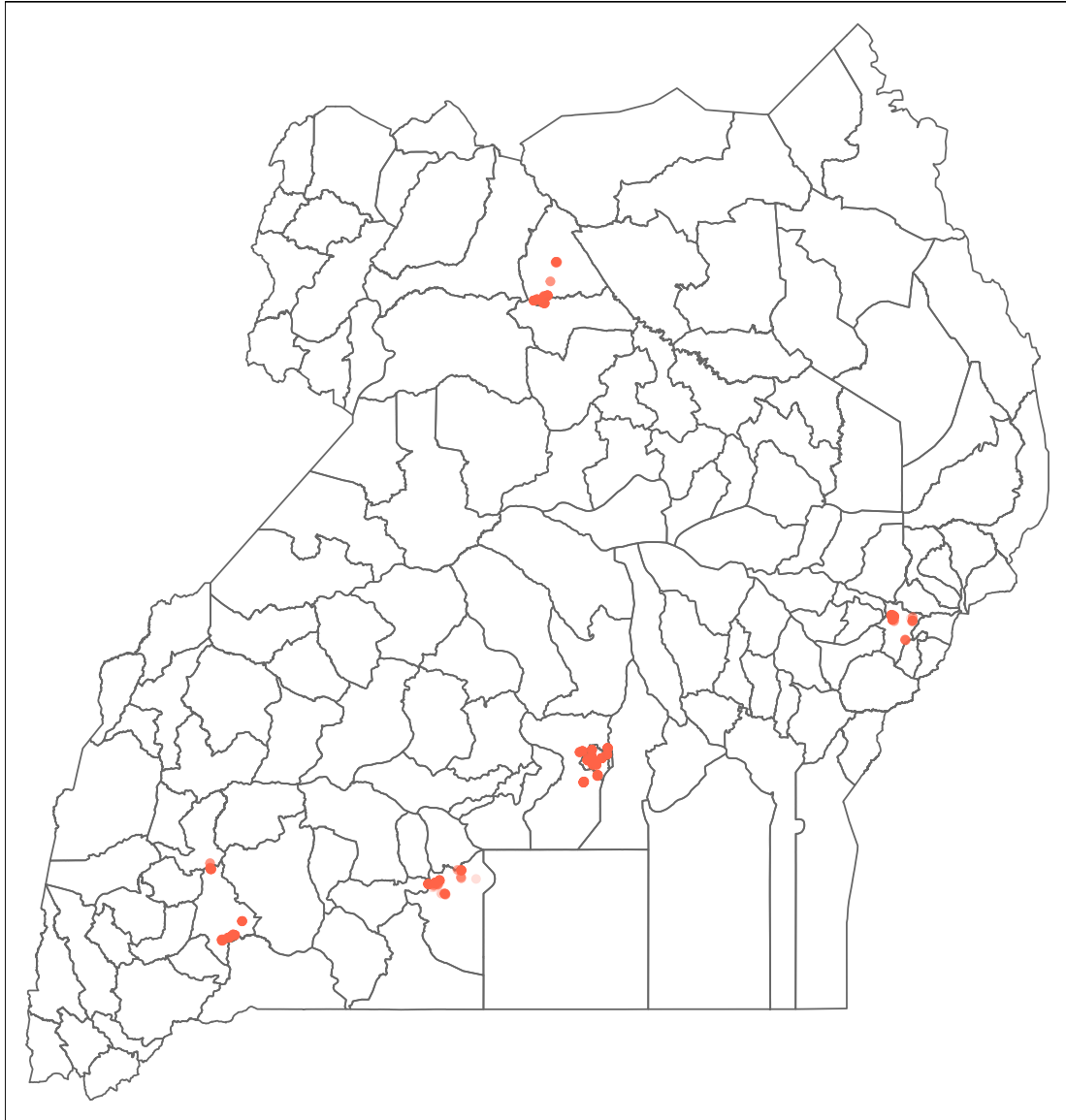
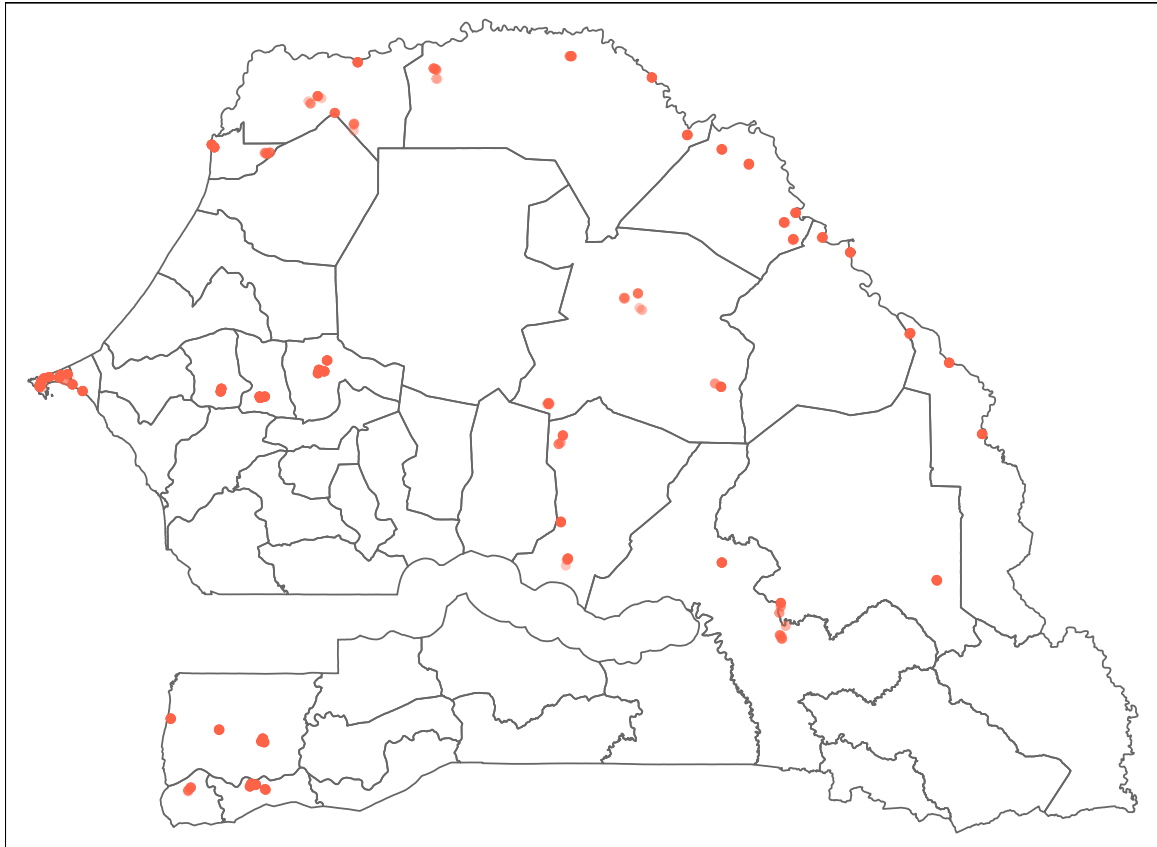


Figure 2.A2: Interview locations in Senegal



2.A.2 Summary statistics

Table 2.A1: Summary statistics and gender differences for the full sample

		N	Gender differences		
			Male	Female	p.overall
		2704	1142	1562	
SEN	1500 (55.5%)		693 (60.7%)	807 (51.7%)	
UGA	1204 (44.5%)		449 (39.3%)	755 (48.3%)	
Age	27.7 (6.67)	2704	27.5 (6.59)	27.9 (6.72)	0.086
Education:		2702			<0.001
Completed primary and below	1250 (46.3%)		483 (42.3%)	767 (49.1%)	
At least some secondary	1033 (38.2%)		447 (39.2%)	586 (37.5%)	
Tertiary or training	419 (15.5%)		211 (18.5%)	208 (13.3%)	
HH income in \$	201 (2374)	2429	174 (500)	222 (3111)	0.573
PP HH income in \$	62.1 (602)	2429	55.0 (182)	67.5 (781)	0.565
Contact abroad	0.40 (0.49)	2698	0.45 (0.50)	0.35 (0.48)	<0.001
Married	0.59 (0.49)	2464	0.49 (0.50)	0.67 (0.47)	<0.001
Number of children	1.80 (2.04)	2678	1.24 (1.88)	2.20 (2.06)	<0.001
Migration intentions	0.71 (0.45)	2686	0.75 (0.44)	0.69 (0.46)	0.001

Note: Summary statistics and gender differences for the full sample

2.A.3 Analysis

Table 2.A2: Regression results basic model

	<i>Dependent variable: Migration intentions</i>	
	(1)	AME
Female	-0.422*** (0.107)	-0.080
Uganda	0.982*** (0.108)	0.187
Constant	0.790*** (0.088)	
Observations	1,944	
Log Likelihood	-1,092.794	

Note: *p<0.1; **p<0.05; ***p<0.01, This table is based on a logit model and reports the coefficients and the AMEs, dependent variable: dummy equal to 1 when respondent has migration intentions

Table 2.A3: Regression results including expected migration costs and expected income differential

	<i>Dependent variable: Migration intentions</i>			
	(1)	(2)	(3)	(4)
Female	-0.363*** (0.135)	-1.002 (0.958)	-0.329** (0.140)	-1.193 (0.918)
Married	-0.355*** (0.134)	0.803 (0.940)	-0.337** (0.139)	1.049 (0.927)
Log exp. migcost European country	0.127** (0.064)	0.189 (0.119)		
Log income differential European country	0.083* (0.045)	0.074 (0.073)		
Log exp. migcost Gulf country			0.007 (0.066)	0.069 (0.129)
Log income differential Gulf country			0.067* (0.038)	0.069 (0.062)
Female:log exp. migcost European country		0.106 (0.127)		
Female:log income differential European country		-0.021 (0.095)		
Married:log exp. migcost European country		-0.192 (0.128)		
Married:log income differential European country		0.036 (0.095)		
Female:log exp. migcost Gulf country				0.127 (0.134)
Female:log income differential Gulf country				0.005 (0.080)
Married:log exp. migcost Gulf country				-0.200 (0.139)
Married:log income differential Gulf country				-0.014 (0.080)
Uganda	1.071*** (0.143)	1.081*** (0.144)	1.005*** (0.143)	1.026*** (0.144)
Constant	-0.418 (0.540)	-0.810 (0.950)	0.595 (0.499)	0.184 (0.891)
Observations	1,446	1,446	1,361	1,361
Log Likelihood	-742.360	-741.059	-690.919	-689.487

Note: *p<0.1; **p<0.05; ***p<0.01, This table is based on a logit model and reports the coefficients, dependent variable: dummy equal to 1 when respondent has migration intentions, reference category: completed primary education and below

Figure 2.A3: Relationship between age and migration intentions

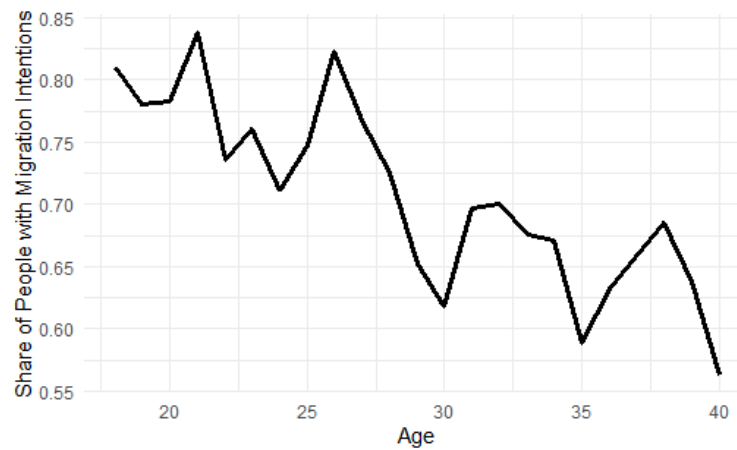


Table 2.A4: AMEs for baseline model

	(1)	(2)	(3)
Female	-0.073	-0.063	-0.055
Age	-0.007	-0.007	-0.007
At least some secondary education	0.089	0.093	0.078
Tertiary education or training	0.057	0.062	0.037
Log PP HH income in \$	-0.011	-0.013	-0.009
Contact abroad	0.032	0.035	0.029
Fixed effects	Country	District	EA

Note: This table is based on a logit model and reports the average marginal effects, for version reporting coefficients, standard errors and p-values, see Table 2.5, reference category: completed primary education and below

Table 2.A5: AMEs for the indices

	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.052	-0.040	-0.033	-0.052	-0.040	-0.034
Age	-0.005	-0.005	-0.005	-0.005	-0.004	-0.005
At least some secondary education	0.072	0.079	0.069	0.073	0.081	0.070
Tertiary education or training	0.033	0.042	0.023	0.036	0.044	0.025
Log PP HH income in \$	-0.013	-0.014	-0.010	-0.014	-0.015	-0.011
Contact abroad	0.019	0.025	0.024	0.018	0.025	0.024
Family obligation	-0.049	-0.052	-0.047	-0.047	-0.051	-0.045
Gender discrimination	-0.021	-0.028	-0.028	-0.022	-0.029	-0.029
Leaving family	0.045	0.043	0.030	0.046	0.044	0.031
Social norm	-0.112	-0.114	-0.106	-0.109	-0.111	-0.103
Interactions	No	No	No	Yes	Yes	Yes
Fixed effects	Country	District	EA	Country	District	EA

Note: This table is based on a logit model and reports the average marginal effects, for version reporting coefficients, standard errors and p-values, see Table 2.7, reference category: completed primary education and below

2.A.3.1 PCA for indirect migration costs

Figure 2.A4: Scree plot for family obligation costs

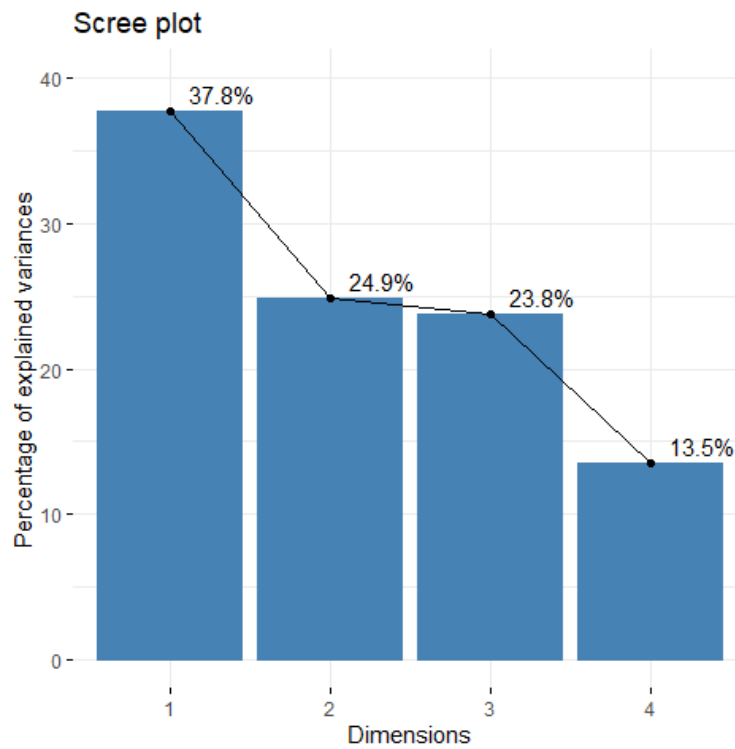


Table 2.A6: Contribution of a variable to a given principal component for family obligation costs in %

	PC1	PC2	PC3	PC4
Log number kids	47.05	0.10	0.62	52.22
Log care elderly number	12.26	0.92	81.56	5.26
Married	37.84	7.87	13.72	40.57
Individualist	2.85	91.10	4.10	1.95

Note: Individualist assesses whether respondents prefer to put family or individual needs first.

Figure 2.A5: Scree plot for gender discrimination costs

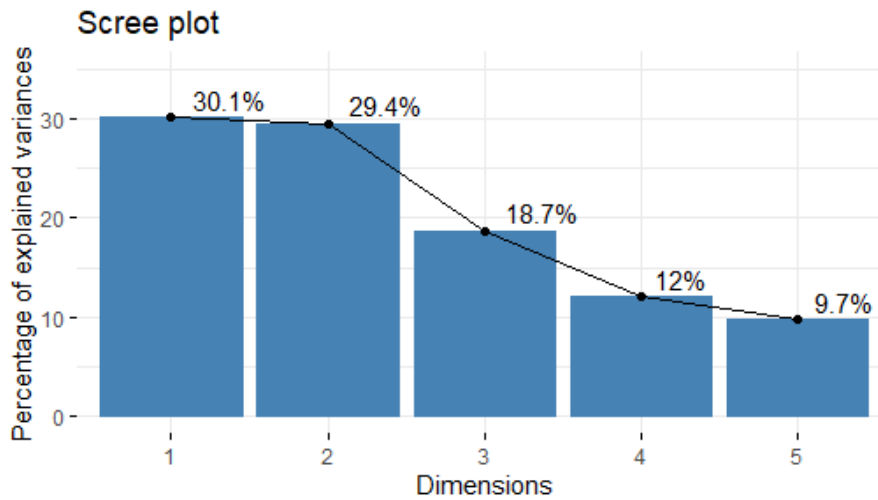


Table 2.A7: Contribution of a variable to a given principal component for gender discrimination costs in %

	PC1	PC2	PC3	PC4	PC5
Women disrespect	0.68	47.09	1.04	33.24	17.95
Work treat	1.25	43.93	7.50	29.97	17.34
Domestic violence	4.46	8.44	84.99	2.10	0.01
Women equality	47.45	0.37	1.35	18.14	32.69
Women disrespect2	46.16	0.16	5.11	16.55	32.01

Note: Women disrespect assesses whether respondents feel that women are treated with respect and dignity while Women disrespect2 assesses whether it's important that women are treated that way. Work treat asks about fair treatment at work and Women equality assesses the importance of equal opportunities at work.

Figure 2.A6: Scree plot for leaving-family costs

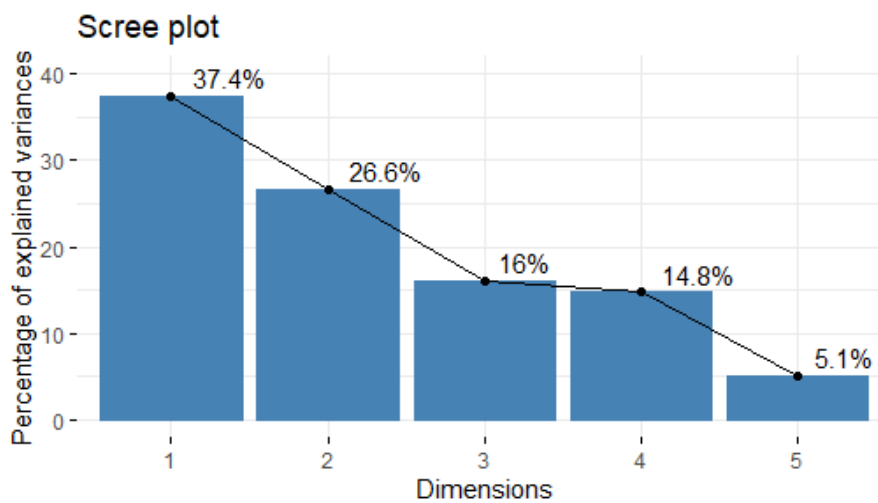


Table 2.A8: Contribution of a variable to a given principal component for leaving-family costs in %

	PC1	PC2	PC3	PC4	PC5
Marry age	4.24	30.51	53.01	11.98	0.26
Early marriage	1.75	40.45	2.77	55.01	0.02
Educ sons	37.44	10.31	1.43	4.20	46.61
Educ daughters	42.65	4.68	0.41	0.50	51.77
Educ equality	13.92	14.05	42.38	28.31	1.34

Figure 2.A7: Scree plot for social norm costs

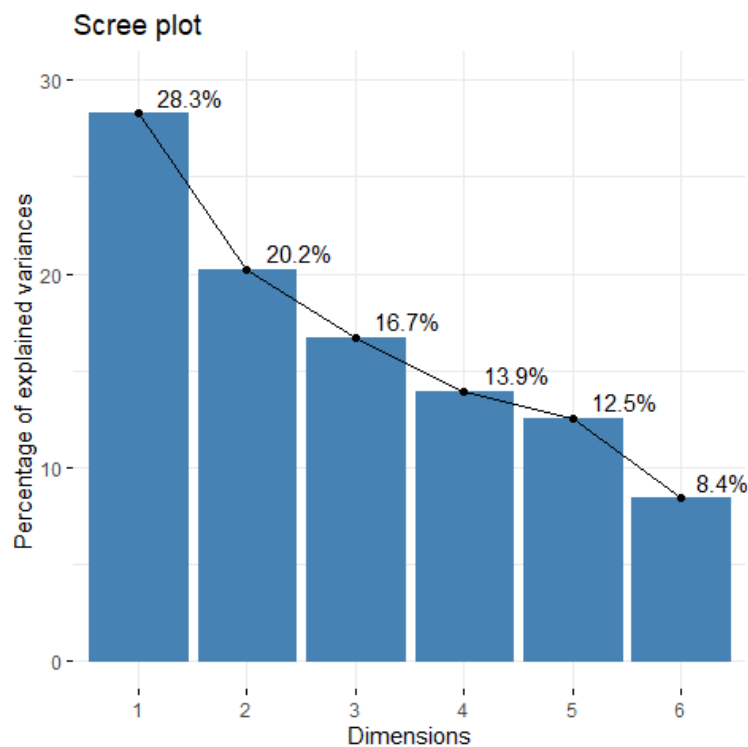


Table 2.A9: Contribution of a variable to a given principal component for social norm costs in %

	PC1	PC2	PC3	PC4	PC5	PC6
Law disobey	35.42	1.87	1.48	14.84	1.71	44.67
Law disobey2	39.20	0.04	1.90	5.76	0.03	53.07
Law illegal entry	15.08	18.23	5.91	14.40	44.55	1.83
Law entry limit	7.92	27.06	0.46	58.47	6.04	0.05
Famop legalmig	0.75	2.97	90.24	1.65	4.10	0.28
Famop illegalmig	1.63	49.82	0.01	4.87	43.57	0.09

Note: Law disobey(2) captures respondents' opinions on disobeying the law (when disagreeing with it). Law illegal entry assesses respondents' stance on entering another country without papers and Law entry limit on other countries' right to limit illegal migration. Famop (il)legalmig captures the families' opinion on working in a richer country (illegally).

Table 2.A10: AMEs for the PCA

	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.045	-0.033	-0.024	-0.048	-0.035	-0.027
Age	-0.004	-0.004	-0.004	-0.004	-0.004	-0.004
At least some secondary education	0.064	0.070	0.065	0.063	0.069	0.065
Tertiary education or training	0.036	0.044	0.032	0.038	0.045	0.032
Log PP HH income in \$	-0.012	-0.014	-0.010	-0.013	-0.015	-0.012
Contact abroad	0.019	0.025	0.025	0.019	0.026	0.027
Family obligation_1	-0.022	-0.021	-0.017	-0.019	-0.019	-0.015
Gender discrimination_1	0.011	0.008	0.010	0.011	0.008	0.010
Gender discrimination_2	-0.005	-0.009	-0.009	-0.004	-0.009	-0.009
Leaving family_1	0.013	0.013	0.010	0.013	0.012	0.009
Leaving family_2	0.024	0.018	0.010	0.027	0.021	0.014
Social norm_1	-0.008	-0.007	-0.004	-0.009	-0.007	-0.004
Social norm_2	-0.017	-0.015	-0.014	-0.017	-0.015	-0.014
Social norm_3	0.066	0.072	0.071	0.061	0.066	0.066
Interactions	No	No	No	Yes	Yes	Yes
Fixed effects	Country	District	EA	Country	District	EA

Note: This table is based on a logit model and reports the average marginal effects, for version reporting coefficients, standard errors and p-values, see Table 2.8, reference category: completed primary education and below

2.A.4 Robustness checks

Table 2.A11: Regression results for the indices using probit model

	<i>Dependent variable: Migration intentions</i>					
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.166** (0.069)	-0.131* (0.071)	-0.111 (0.075)	-0.154** (0.069)	-0.118* (0.071)	-0.099 (0.075)
Age	-0.015*** (0.006)	-0.016*** (0.006)	-0.018*** (0.006)	-0.015*** (0.006)	-0.015*** (0.006)	-0.018*** (0.006)
At least some secondary education	0.232*** (0.077)	0.268*** (0.081)	0.244*** (0.087)	0.236*** (0.077)	0.272*** (0.081)	0.251*** (0.087)
Tertiary education or training	0.120 (0.102)	0.153 (0.105)	0.101 (0.114)	0.130 (0.102)	0.161 (0.105)	0.114 (0.114)
Log PP HH income in \$	-0.041** (0.019)	-0.048** (0.020)	-0.038* (0.022)	-0.045** (0.020)	-0.051** (0.020)	-0.041* (0.022)
Contact abroad	0.062 (0.068)	0.089 (0.070)	0.085 (0.074)	0.060 (0.068)	0.088 (0.070)	0.086 (0.074)
Family obligation	-0.171*** (0.064)	-0.182*** (0.066)	-0.170** (0.069)	-0.158* (0.089)	-0.180** (0.092)	-0.164* (0.096)
Gender discrimination	-0.069 (0.060)	-0.092 (0.062)	-0.087 (0.065)	-0.066 (0.088)	-0.086 (0.090)	-0.060 (0.094)
Leaving family	0.148*** (0.056)	0.151*** (0.058)	0.113* (0.061)	-0.017 (0.083)	-0.012 (0.084)	-0.062 (0.089)
Social norm	-0.366*** (0.070)	-0.386*** (0.071)	-0.380*** (0.075)	-0.320*** (0.108)	-0.306*** (0.110)	-0.276** (0.116)
Female:family obligation				-0.013 (0.113)	0.004 (0.115)	-0.001 (0.122)
Female:gender discrimination				-0.012 (0.119)	-0.019 (0.121)	-0.056 (0.126)
Female:leaving family				0.295*** (0.108)	0.291*** (0.110)	0.316*** (0.115)
Female:social norm				-0.069 (0.141)	-0.127 (0.144)	-0.165 (0.152)
Constant	0.873*** (0.174)	0.836*** (0.222)	1.448*** (0.475)	0.856*** (0.176)	0.817*** (0.224)	1.458*** (0.478)
Fixed effects	Country	District	EA	Country	District	EA
Observations	1,944	1,944	1,944	1,944	1,944	1,944
Log Likelihood	-1,042.119	-1,014.817	-968.924	-1,038.224	-1,010.935	-964.523

Note: *p<0.1; **p<0.05; ***p<0.01, This table is based on a logit model and reports the coefficients, dependent variable: dummy equal to 1 when respondent has migration intentions, reference category: completed primary education and below

Table 2.A12: Regression results for the indices using OLS

	<i>Dependent variable: Migration intentions</i>					
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.051**	-0.040*	-0.033	-0.052**	-0.041**	-0.035
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
Age	-0.005***	-0.005***	-0.005***	-0.005***	-0.005***	-0.005***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
At least some secondary education	0.075***	0.082***	0.069***	0.076***	0.083***	0.071***
	(0.023)	(0.024)	(0.025)	(0.023)	(0.024)	(0.025)
Tertiary education or training	0.041	0.049	0.029	0.044	0.052*	0.034
	(0.031)	(0.031)	(0.033)	(0.031)	(0.031)	(0.033)
Log PP HH income in \$	-0.013**	-0.016**	-0.011*	-0.015**	-0.017***	-0.012*
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Contact abroad	0.020	0.027	0.024	0.018	0.027	0.024
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
Family obligation	-0.048**	-0.051***	-0.045**	-0.041	-0.047*	-0.041
	(0.019)	(0.019)	(0.020)	(0.027)	(0.027)	(0.027)
Gender discrimination	-0.021	-0.029	-0.029	-0.021	-0.028	-0.022
	(0.019)	(0.019)	(0.019)	(0.027)	(0.027)	(0.027)
Leaving family	0.048***	0.049***	0.036**	-0.009	-0.006	-0.022
	(0.017)	(0.018)	(0.018)	(0.025)	(0.025)	(0.025)
Social norm	-0.117***	-0.118***	-0.108***	-0.095***	-0.087***	-0.075**
	(0.021)	(0.021)	(0.022)	(0.032)	(0.032)	(0.033)
Female:family obligation				-0.008	-0.003	-0.003
				(0.034)	(0.033)	(0.034)
Female:gender discrimination				-0.002	-0.004	-0.017
				(0.037)	(0.036)	(0.037)
Female:leaving family				0.107***	0.104***	0.110***
				(0.033)	(0.033)	(0.034)
Female:social norm				-0.036	-0.051	-0.056
				(0.042)	(0.042)	(0.043)
Constant	0.799***	0.787***	0.959***	0.797***	0.785***	0.960***
	(0.054)	(0.068)	(0.129)	(0.054)	(0.068)	(0.129)
Fixed effects	Country	District	EA	Country	District	EA
Observations	1,944	1,944	1,944	1,944	1,944	1,944
Adjusted R ²	0.094	0.114	0.119	0.098	0.117	0.123
F Statistic	19.416***	12.875***	3.509***	15.049***	11.317***	3.512***
	df= 11;	df = 21;	df = 105;	df = 15;	df = 25;	df = 109;
	1932	1922	1838	19289	1918	1834

Note: *p<0.1; **p<0.05; ***p<0.01, This table is based on a logit model and reports the coefficients, dependent variable: dummy equal to 1 when respondent has migration intentions, reference category: completed primary education and below

Table 2.A13: Regression results for the PCA using Probit model

	<i>Dependent variable: Migration intentions</i>					
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.146**	-0.110	-0.083	-0.143**	-0.107	-0.084
	(0.071)	(0.073)	(0.077)	(0.071)	(0.073)	(0.077)
Age	-0.014**	-0.016***	-0.017***	-0.014**	-0.016**	-0.017***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
At least some secondary education	0.207***	0.237***	0.235***	0.203**	0.233***	0.234***
	(0.079)	(0.082)	(0.088)	(0.079)	(0.083)	(0.089)
Tertiary education or training	0.129	0.163	0.137	0.135	0.169	0.144
	(0.104)	(0.107)	(0.116)	(0.104)	(0.108)	(0.116)
Log PP HH income in \$	-0.040**	-0.049**	-0.039*	-0.045**	-0.053**	-0.044**
	(0.020)	(0.021)	(0.022)	(0.020)	(0.021)	(0.023)
Contact abroad	0.062	0.085	0.081	0.062	0.090	0.091
	(0.068)	(0.070)	(0.075)	(0.069)	(0.071)	(0.075)
Family obligation_1	-0.075**	-0.074**	-0.064*	-0.085*	-0.093**	-0.088*
	(0.033)	(0.034)	(0.036)	(0.044)	(0.045)	(0.048)
Gender discrimination_1	0.038	0.030	0.043	0.004	-0.002	0.024
	(0.028)	(0.028)	(0.030)	(0.039)	(0.040)	(0.042)
Gender discrimination_2	-0.014	-0.029	-0.028	0.010	-0.005	-0.002
	(0.027)	(0.028)	(0.029)	(0.042)	(0.043)	(0.044)
Leaving family_1	0.045*	0.047*	0.037	-0.015	-0.010	-0.024
	(0.024)	(0.025)	(0.026)	(0.037)	(0.037)	(0.039)
Leaving family_2	0.083**	0.065*	0.038	0.055	0.039	0.014
	(0.032)	(0.033)	(0.035)	(0.043)	(0.043)	(0.045)
Social norm_1	-0.030	-0.026	-0.016	-0.006	0.006	0.015
	(0.026)	(0.027)	(0.028)	(0.038)	(0.038)	(0.040)
Social norm_2	-0.056*	-0.053	-0.055	-0.111**	-0.109**	-0.097*
	(0.032)	(0.033)	(0.035)	(0.047)	(0.048)	(0.050)
social norm_3	0.230***	0.255***	0.265***	0.116**	0.129**	0.129**
	(0.031)	(0.032)	(0.035)	(0.052)	(0.054)	(0.057)
Female:family obligation_1				0.031	0.049	0.056
				(0.055)	(0.056)	(0.060)
Female:gender discrimination_1				0.061	0.055	0.030
				(0.056)	(0.057)	(0.060)
Female:gender discrimination_2				-0.041	-0.044	-0.047
				(0.055)	(0.055)	(0.058)
Female:leaving family_1				0.105**	0.097**	0.108**
				(0.048)	(0.049)	(0.051)
Female:leaving family_2				0.067	0.068	0.069
				(0.059)	(0.060)	(0.063)
Female:social norm_1				-0.046	-0.061	-0.058
				(0.051)	(0.052)	(0.054)
female:social norm_2				0.095	0.096	0.075
				(0.062)	(0.063)	(0.067)
Female:social norm_3				0.169***	0.190***	0.207***
				(0.066)	(0.067)	(0.071)

Constant	0.844*** (0.187)	0.851*** (0.236)	1.570*** (0.494)	0.844*** (0.190)	0.846*** (0.239)	1.564*** (0.495)
Fixed effects	Country	District	EA	Country	District	EA
Observations	1,944	1,944	1,944	1,944	1,944	1,944
Log Likelihood	-1,023.281	-994.314	-949.501	-1,014.857	-985.235	-940.483

Note: *p<0.1; **p<0.05; ***p<0.01, This table is based on a logit model and reports the coefficients, dependent variable: dummy equal to 1 when respondent has migration intentions, reference category: completed primary education and below

Table 2.A14: Regression results for the PCA using OLS

	<i>Dependent variable: Migration intentions</i>					
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.044**	-0.031	-0.024	-0.049**	-0.036*	-0.029
	(0.021)	(0.021)	(0.022)	(0.021)	(0.021)	(0.022)
Age	-0.004**	-0.005***	-0.005***	-0.004**	-0.005**	-0.005**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
At least some secondary education	0.066***	0.071***	0.065***	0.065***	0.071***	0.066***
	(0.024)	(0.024)	(0.025)	(0.024)	(0.024)	(0.025)
Tertiary education or training	0.042	0.051*	0.038	0.044	0.053*	0.041
	(0.031)	(0.031)	(0.033)	(0.031)	(0.031)	(0.033)
Log PP HH income in \$	-0.013**	-0.016***	-0.011*	-0.015**	-0.017***	-0.013**
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Contact abroad	0.021	0.027	0.025	0.020	0.028	0.026
	(0.020)	(0.021)	(0.021)	(0.020)	(0.020)	(0.021)
Family obligation_1	-0.023**	-0.022**	-0.018*	-0.025*	-0.027**	-0.024*
	(0.010)	(0.010)	(0.010)	(0.013)	(0.013)	(0.013)
Gender discrimination_1	0.012	0.009	0.012	0.001	-0.001	0.006
	(0.008)	(0.008)	(0.009)	(0.012)	(0.012)	(0.012)
Gender discrimination_2	-0.004	-0.010	-0.011	0.003	-0.003	-0.003
	(0.008)	(0.008)	(0.008)	(0.012)	(0.012)	(0.013)
Leaving family_1	0.014**	0.015**	0.012	-0.005	-0.003	-0.007
	(0.007)	(0.007)	(0.007)	(0.010)	(0.010)	(0.011)
Leaving family_2	0.028***	0.022**	0.014	0.017	0.011	0.004
	(0.010)	(0.010)	(0.010)	(0.013)	(0.013)	(0.013)
Social norm_1	-0.010	-0.008	-0.005	-0.0001	0.004	0.005
	(0.008)	(0.008)	(0.008)	(0.011)	(0.011)	(0.011)
Social norm_2	-0.017*	-0.015	-0.015	-0.033**	-0.031**	-0.026*
	(0.010)	(0.010)	(0.010)	(0.014)	(0.014)	(0.014)
Social norm_3	0.079***	0.085***	0.082***	0.039**	0.043***	0.038**
	(0.010)	(0.010)	(0.010)	(0.016)	(0.016)	(0.016)
Female:family obligation_1				0.009	0.014	0.014
				(0.016)	(0.016)	(0.017)
Female:gender discrimination_1				0.019	0.016	0.009
				(0.017)	(0.016)	(0.017)
Female:gender discrimination_2				-0.013	-0.013	-0.013
				(0.016)	(0.016)	(0.016)
Female:leaving family_1				0.034**	0.031**	0.035**
				(0.014)	(0.014)	(0.014)
Female:leaving family_2				0.025	0.025	0.024
				(0.018)	(0.018)	(0.018)
Female:social norm_1				-0.019	-0.023	-0.019
				(0.015)	(0.015)	(0.015)
Female:social norm_2				0.029	0.028	0.020
				(0.018)	(0.018)	(0.019)
Female:social norm_3				0.060***	0.063***	0.065***
				(0.020)	(0.020)	(0.021)

Constant	0.782*** (0.057)	0.779*** (0.070)	0.981*** (0.130)	0.786*** (0.057)	0.782*** (0.070)	0.978*** (0.130)
Fixed effects	Country	District	EA	Country	District	EA
Observations	1,944	1,944	1,944	1,944	1,944	1,944
Adjusted R ²	0.115	0.136	0.139	0.122	0.143	0.146
F Statistic	17.894*** df = 15; 1928	13.265*** df = 25; 1918	3.889*** df = 109; 1834	12.726*** df = 23; 1920	10.836*** df = 33; 1910	3.839*** df = 117; 1826

Note: *p<0.1; **p<0.05; ***p<0.01, This table is based on a logit model and reports the coefficients, dependent variable: dummy equal to 1 when respondent has migration intentions, reference category: completed primary education and below

Table 2.A15: Regression results alternative gender discrimination measures

	<i>Dependent variable: Migration intentions</i>					
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.254** (0.120)	-0.182 (0.123)	-0.154 (0.130)	-0.251** (0.119)	-0.189 (0.122)	-0.153 (0.129)
Age	-0.025*** (0.009)	-0.025*** (0.010)	-0.029*** (0.010)	-0.025*** (0.009)	-0.025*** (0.010)	-0.028*** (0.010)
At least some secondary education	0.409*** (0.132)	0.468*** (0.139)	0.429*** (0.149)	0.414*** (0.132)	0.469*** (0.139)	0.433*** (0.150)
Tertiary education or training	0.192 (0.175)	0.249 (0.181)	0.152 (0.196)	0.199 (0.175)	0.250 (0.181)	0.155 (0.196)
Log PP HH income in \$	-0.077** (0.033)	-0.088** (0.035)	-0.070* (0.038)	-0.078** (0.033)	-0.088** (0.035)	-0.071* (0.038)
Contact abroad	0.098 (0.116)	0.139 (0.119)	0.140 (0.128)	0.101 (0.116)	0.144 (0.119)	0.147 (0.128)
Family obligation	-0.254* (0.154)	-0.291* (0.158)	-0.278* (0.167)	-0.253* (0.154)	-0.294* (0.158)	-0.276* (0.167)
Women disrespect	-0.001 (0.087)	-0.033 (0.088)	-0.029 (0.092)			
Gender discrimination real				-0.055 (0.124)	-0.084 (0.125)	-0.115 (0.131)
Leaving family	-0.022 (0.141)	-0.018 (0.143)	-0.121 (0.154)	-0.026 (0.141)	-0.022 (0.143)	-0.127 (0.154)
Social norm	-0.534*** (0.186)	-0.503*** (0.189)	-0.467** (0.200)	-0.534*** (0.186)	-0.501*** (0.189)	-0.466** (0.200)
Female:family obligation	-0.021 (0.194)	-0.011 (0.199)	0.002 (0.212)	-0.021 (0.194)	-0.005 (0.199)	0.002 (0.212)
Female:women disrespect	-0.058 (0.111)	-0.089 (0.113)	-0.090 (0.118)			
Female:gender discrimination real				-0.055 (0.160)	-0.058 (0.161)	-0.057 (0.169)
Female:leaving family	0.497*** (0.183)	0.494*** (0.186)	0.567*** (0.198)	0.505*** (0.183)	0.504*** (0.186)	0.577*** (0.199)
Female:social norm	-0.129 (0.240)	-0.233 (0.246)	-0.271 (0.260)	-0.126 (0.240)	-0.223 (0.245)	-0.259 (0.260)
Constant	1.431*** (0.298)	1.357*** (0.380)	2.377*** (0.854)	1.415*** (0.298)	1.337*** (0.379)	2.371*** (0.854)
Fixed effects	Country	District	EA	Country	District	EA
Observations	1,944	1,944	1,944	1,944	1,944	1,944
Log Likelihood	-1,038.186	-1,010.929	-964.877	-1,037.899	-1,011.374	-964.666

Note: *p<0.1; **p<0.05; ***p<0.01, This table is based on a logit model and reports the coefficients, dependent variable: dummy equal to 1 when respondent has migration intentions, reference category: completed primary education and below

Table 2.A16: Regression results female sample

	<i>Dependent variable: Migration intentions</i>					
	(1)	(2)	(3)	(4)	(5)	(6)
Age	-0.017 (0.012)	-0.017 (0.012)	-0.017 (0.013)	-0.018 (0.012)	-0.020 (0.013)	-0.019 (0.014)
At least some secondary education	0.511*** (0.169)	0.492*** (0.178)	0.530*** (0.201)	0.439** (0.175)	0.410** (0.185)	0.460** (0.207)
Tertiary education or training	0.397 (0.249)	0.377 (0.256)	0.385 (0.285)	0.393 (0.257)	0.373 (0.264)	0.419 (0.292)
Log PP HH income in \$	-0.074* (0.043)	-0.094** (0.044)	-0.049 (0.051)	-0.075* (0.044)	-0.099** (0.046)	-0.052 (0.053)
Contact abroad	0.155 (0.155)	0.190 (0.158)	0.198 (0.178)	0.169 (0.158)	0.204 (0.162)	0.253 (0.183)
Family obligation	-0.287** (0.144)	-0.290* (0.149)	-0.296* (0.168)			
Gender discrimination	-0.140 (0.139)	-0.180 (0.143)	-0.172 (0.160)			
Leaving family	0.440*** (0.124)	0.411*** (0.132)	0.340** (0.146)			
Social norm	-0.671*** (0.154)	-0.736*** (0.161)	-0.780*** (0.183)			
Family obligation_1				-0.095 (0.076)	-0.078 (0.078)	-0.071 (0.089)
Gender discrimination_1				0.113* (0.068)	0.087 (0.070)	0.118 (0.079)
Gender discrimination_2				-0.054 (0.061)	-0.072 (0.064)	-0.066 (0.070)
Leaving family_1				0.146*** (0.054)	0.132** (0.057)	0.113* (0.063)
Leaving family_2				0.208*** (0.080)	0.165** (0.083)	0.110 (0.094)
Social norm_1				-0.085 (0.061)	-0.086 (0.064)	-0.062 (0.074)
Social norm_2				-0.034 (0.073)	-0.016 (0.077)	-0.023 (0.087)
Social norm_3				0.467*** (0.067)	0.521*** (0.071)	0.582*** (0.083)
Constant	0.839** (0.348)	1.025** (0.460)	17.608 (1,743.698)	0.942** (0.369)	1.213** (0.487)	18.047 (1,672.639)
Fixed effects	Country	District	EA	Country	District	EA
Observations	1,091	1,091	1,091	1,091	1,091	1,091
Log Likelihood	-600.692	-582.253	-525.077	-581.065	-561.087	-506.409

Note: *p<0.1; **p<0.05; ***p<0.01, This table is based on a logit model and reports the coefficients, dependent variable: dummy equal to 1 when respondent has migration intentions, reference category: completed primary education and below

Table 2.A17: Replication of Ruysen and Salomone (2018) coefficients

	<i>Dependent variable: Migration intentions</i>				
	Gallup data				Survey data
	World	Africa	Uganda and Senegal	Uganda and Senegal	Uganda and Senegal
Gender discrimination	0.188*** (0.00)	0.137*** (0.01)	0.031 (0.03)	0.043* (0.03)	-0.043 (0.088)
Married	-0.170*** (0.00)	-0.204*** (0.01)	-0.285*** (0.03)	-0.286*** (0.03)	-0.127 (0.103)
High skilled	0.074*** (0.00)	0.053*** (0.01)	-0.190** (0.09)	-0.190** (0.09)	0.153 (0.152)
Number of children	-0.018*** (0.00)	-0.025*** (0.00)	-0.056*** (0.01)	-0.054*** (0.01)	-0.036 (0.026)
Urban	0.124*** (0.00)	0.134*** (0.01)	0.019 (0.03)	0.021 (0.03)	0.026 (0.092)
Employed	0.006* (0.00)	0.010 (0.01)	0.095*** (0.03)	0.093*** (0.03)	0.084 (0.091)
Log HH income pc	0.019*** (0.00)	0.034*** (0.00)	0.010 (0.01)	0.004 (0.01)	-0.003 (0.027)
HH size	0.010*** (0.00)	0.023*** (0.00)	0.041*** (0.01)	0.040*** (0.01)	0.021* (0.012)
Network abroad	0.399*** (0.01)	0.684*** (0.03)	0.764*** (0.11)	0.767*** (0.11)	0.169* (0.094)
Age	-0.014*** (0.00)	-0.015*** (0.00)	-0.014*** (0.00)	-0.014*** (0.00)	-0.009 (0.008)
Daily Experience Index	-0.004*** (0.00)	-0.003*** (0.00)	-0.002*** (0.00)		
Constant	-0.198*** (0.02)	-0.384*** (0.03)	0.135 (0.11)	0.014 (0.11)	0.229 (0.312)
Fixed effects	Country	Country	Country	Country	Country
Observations	790,808	212,709	12,656	12,656	973
Log Likelihood	-390017.355	-120290.867	-8150.963	-8158.442	-557.832

Note: *p<0.1; **p<0.05; ***p<0.01, This table is based on a probit model and reports the coefficients, for version reporting average marginal effects, see Appendix Table 2.A18, dependent variable: dummy equal to 1 when respondent has migration intentions

Table 2.A18: Replication of Ruysen and Salomone (2018) AMEs

	<i>Dependent variable: Migration intentions</i>				
	Gallup data				Survey data
	World	Africa	Uganda and Senegal	Uganda and Senegal	Uganda and Senegal
Gender discrimination	0.052	0.044	0.011	0.016	-0.014
Married	-0.170	-0.204	-0.285	-0.286	-0.041
High skilled	0.020	0.017	-0.070	-0.070	0.050
Number of children	-0.005	-0.008	-0.020	-0.020	-0.012
Urban	0.034	0.043	0.007	0.008	0.008
Employed	0.002	0.003	0.035	0.034	0.027
Log HH income pc	0.005	0.011	0.004	0.001	-0.001
HH size	0.003	0.007	0.015	0.015	0.007
Network abroad	0.111	0.219	0.281	0.283	0.055
Age	-0.004	-0.005	-0.005	-0.005	-0,003
Daily Experience Index	-0.001	-0.001	-0.001		
Fixed effects	Country	Country	Country	Country	Country

Note: This table is based on a probit model and reports the average marginal effects, for version reporting coefficients, standard errors and p-values, see Table 2.A17

Table 2.A19: Regression results including marital status and parenthood without NAs for marital status

	<i>Dependent variable: Migration intentions</i>										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Female	-0.332*** (0.119)	-0.274** (0.122)	-0.246** (0.124)	-0.265** (0.124)	-0.292** (0.121)	-0.267** (0.123)	-0.277** (0.124)	-0.252** (0.124)	-0.250** (0.124)	-0.277** (0.124)	-0.273** (0.124)
Married	-0.202 (0.127)	-0.131 (0.130)	-0.096 (0.134)	-0.120 (0.134)	-0.169 (0.128)	-0.145 (0.130)	-0.151 (0.131)	-0.104 (0.134)	-0.101 (0.134)	-0.155 (0.131)	-0.152 (0.131)
Age	-0.035*** (0.009)	-0.022** (0.010)	-0.023** (0.011)	-0.028*** (0.010)	-0.028*** (0.010)	-0.028*** (0.010)	-0.030*** (0.010)	-0.021* (0.011)	-0.021** (0.011)	-0.027*** (0.010)	-0.027*** (0.010)
At least some secondary education	0.497*** (0.139)	0.452*** (0.141)	0.447*** (0.141)	0.463*** (0.141)	0.470*** (0.140)	0.472*** (0.140)	0.481*** (0.140)	0.444*** (0.141)	0.444*** (0.141)	0.468*** (0.140)	0.469*** (0.140)
Tertiary education or training	0.297 (0.180)	0.222 (0.183)	0.214 (0.184)	0.239 (0.183)	0.275 (0.181)	0.270 (0.181)	0.275 (0.181)	0.207 (0.184)	0.211 (0.184)	0.271 (0.181)	0.271 (0.181)
Log PP HH income in \$	-0.077** (0.034)	-0.081** (0.034)	-0.076** (0.034)	-0.073** (0.034)	-0.080** (0.034)	-0.078** (0.034)	-0.077** (0.034)	-0.079** (0.034)	-0.079** (0.034)	-0.079** (0.034)	-0.080** (0.034)
Contact abroad	0.191 (0.118)	0.192 (0.118)	0.187 (0.118)	0.185 (0.118)	0.201* (0.118)	0.196* (0.118)	0.192 (0.118)	0.189 (0.119)	0.189 (0.119)	0.199* (0.118)	0.199* (0.118)
Number kids		-0.087*** (0.034)					-0.071* (0.039)				
Log number kids			-0.119** (0.048)						-0.201* (0.121)		
Kids				-0.306** (0.154)				-0.153 (0.176)	0.286 (0.389)		
Care kids number					-0.059** (0.027)					-0.048 (0.033)	
Log care kids number						-0.085** (0.041)					-0.151 (0.112)
Care kids							-0.240* (0.141)			-0.099 (0.171)	0.243 (0.385)
Constant	1.760*** (0.353)	1.474*** (0.370)	1.251*** (0.408)	1.645*** (0.359)	1.609*** (0.360)	1.455*** (0.382)	1.715*** (0.355)	1.470*** (0.371)	1.009* (0.523)	1.618*** (0.360)	1.264*** (0.486)
Fixed effects	District	District	District	District	District	District	District	District	District	District	District
Observations	1,840	1,840	1,840	1,840	1,840	1,840	1,840	1,840	1,840	1,840	1,840
Log Likelihood	-989.146	-985.834	-986.041	-987.159	-986.786	-986.973	-987.689	-985.456	-985.769	-986.620	-986.773

Note: *p<0.1; **p<0.05; ***p<0.01, This table is based on a logit model and reports the coefficients, dependent variable: dummy equal to 1 when respondent has migration intentions, reference category: completed primary education and below

Table 2.A20: Regression results for the single indices

	<i>Dependent variable: Migration intentions</i>			
	(1)	(2)	(3)	(4)
Female	-0.281** (0.118)	-0.351*** (0.116)	-0.347*** (0.116)	-0.268** (0.118)
Age	-0.028*** (0.010)	-0.040*** (0.008)	-0.040*** (0.009)	-0.036*** (0.009)
At least some secondary education	0.470*** (0.136)	0.518*** (0.135)	0.489*** (0.136)	0.530*** (0.136)
Tertiary education or training	0.259 (0.177)	0.342* (0.175)	0.311* (0.176)	0.333* (0.176)
Log PP HH income in \$	-0.072** (0.034)	-0.074** (0.033)	-0.075** (0.034)	-0.081** (0.034)
Contact abroad	0.199* (0.116)	0.201* (0.116)	0.179 (0.117)	0.156 (0.118)
Family obligation	-0.300* (0.157)			
Gender discrimination		-0.128 (0.149)		
Leaving family			-0.037 (0.141)	
Social norm				-0.523*** (0.188)
Female:family obligation	-0.044 (0.196)			
Female:gender discrimination		-0.041 (0.199)		
Female:leaving family			0.528*** (0.183)	
Female:social norm cost				-0.200 (0.243)
Constant	1.384*** (0.372)	1.774*** (0.346)	1.785*** (0.348)	1.649*** (0.349)
Fixed effects	District	District	District	District
Observations	1,944	1,944	1,944	1,944
Log Likelihood	-1,034.209	-1,037.395	-1,030.647	-1,023.627

Note: *p<0.1; **p<0.05; ***p<0.01, This table is based on a logit model and reports the coefficients, for version reporting average marginal effects, see Appendix Table 2.A5, dependent variable: dummy equal to 1 when respondent has migration intentions, reference category: completed primary education and below

Table 2.A21: Regression results for the single principal components

	<i>Dependent variable: Migration intentions</i>			
	(1)	(2)	(3)	(4)
Female	-0.268** (0.120)	-0.350*** (0.117)	-0.367*** (0.118)	-0.212* (0.119)
Age	-0.025** (0.010)	-0.040*** (0.009)	-0.040*** (0.009)	-0.037*** (0.009)
At least some secondary education	0.459*** (0.137)	0.520*** (0.135)	0.473*** (0.137)	0.467*** (0.140)
Tertiary education or training	0.248 (0.178)	0.344** (0.175)	0.314* (0.177)	0.326* (0.180)
Log PP HH income in \$	-0.071** (0.034)	-0.077** (0.034)	-0.072** (0.034)	-0.087** (0.035)
Contact abroad	0.195* (0.116)	0.199* (0.116)	0.167 (0.117)	0.176 (0.120)
Family obligation_1	-0.169** (0.076)			
Gender discrimination_1		0.004 (0.064)		
Gender discrimination_2		-0.018 (0.072)		
Leaving family_1			-0.048 (0.064)	
Leaving_family_2			0.093 (0.070)	
Social norm_1				0.019 (0.065)
Social norm_2				-0.198** (0.080)
Social norm_3				0.249*** (0.087)
Female:family obligation_1	0.016 (0.094)			
Female:gender discrimination_1		-0.042 (0.087)		
Female:gender discrimination_2		-0.044 (0.091)		
Female:leaving family_1			0.205** (0.082)	
Female:leaving family_2			0.157* (0.094)	
Female:social norm_1				-0.108 (0.087)
Female:social norm_2				0.181* (0.104)
Female:social norm_3				0.296*** (0.110)
Constant	1.287*** (0.385)	1.769*** (0.347)	1.875*** (0.352)	1.670*** (0.361)

Fixed effects	District	District	District	District
Observations	1,944	1,944	1,944	1,944
Log Likelihood	-1,034.448	-1,037.705	-1,027.314	-994.669

Note: *p<0.1; **p<0.05; ***p<0.01, This table is based on a logit model and reports the coefficients, for version reporting average marginal effects, see Appendix Table 2.A10, dependent variable: dummy equal to 1 when respondent has migration intentions, reference category: completed primary education and below

Chapter 3

Forecasting Bilateral Refugee Flows with High-dimensional Data and Machine Learning Techniques

Konstantin Boss, Andre Gröger, Tobias Heidland, Finja Krüger & Conghan Zheng

Abstract

We develop monthly refugee flow forecasting models for 150 origin countries to the EU27, using machine learning and high-dimensional data, including digital trace data from Google Trends. Comparing different models and forecasting horizons and validating them out-of-sample, we find that an ensemble forecast combining Random Forest and Extreme Gradient Boosting algorithms consistently outperforms forecast horizons between 3 to 12 months. For large refugee flow corridors, this holds in a parsimonious model exclusively based on Google Trends variables, which has the advantage of close-to-real-time availability. We provide practical recommendations about how our approach can enable ahead-of-period refugee forecasting applications.

Acknowledgment: Funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 882986 is acknowledged.

3.1 Introduction

In 2013-2015, the European Union experienced a substantial increase in refugee flows, with more than 1.2 million asylum applications registered at its peak in 2015 alone.¹ The experience of 2015 has triggered widespread political debate and stressed the importance of preparedness for governments and humanitarian organizations to ensure safe conditions for asylum seekers and refugees en route and upon arrival. That increased political attention, combined with improved data availability and methodological and technological advances over the past decade, has moved the objective of migration forecasting to the center of scientific attention across different disciplines. Forecasting refugee flows is especially challenging due to the particular circumstances of forced migration, such as the lack of media reporting or data collection in conflict areas, making it particularly difficult to obtain early signals about international population movements. Additionally, violent conflict, natural disasters, or economic crises tend to occur unexpectedly and infrequently, which further complicates prediction (Disney et al., 2015).

In this context, Böhme et al. (2020) have demonstrated that migration-related Google Trends Indices (GTI) recorded in migrant origin countries hold additional *in-sample* predictive power over classical predictors when explaining bilateral migration flows to the OECD in a yearly gravity-type regression framework.² Building on that proof-of-concept paper, we create a realistic policy prediction framework (Kleinberg et al., 2015) - simulating the information set available to the analyst including Google Trends - and assess the *out-of-sample* performance of different forecasting model specifications using bilateral high-frequency flows of asylum seekers to the European Union. Google Trends allows studying keyword-specific search behavior at an aggregate, anonymous level to follow trends by country over time. The platform enables us to track the search interest in certain keywords such as "passport" or specific (destination) country names in migrant countries of origin over time, thus learning about interest in migration in general, and certain bilateral corridors in particular.

This paper provides three main contributions. First, we construct a data set of asylum seeker registrations by applicant nationality of unprecedented frequency and scope featuring a real-life practitioner's information set to forecast asylum seeker arrivals at destination. To this end, we combine a range of Google Trends time series capturing internet searches for migration-related keywords at origin with administrative records of asylum registries at destination as well as hundreds of predictor candidate variables cap-

¹That peak constituted an increase of nearly 130% compared to the previous year alone. Recent figures show that asylum registrations have once again reached similar levels in 2022 (Eurostat, 2022).

²Most people who perform online searches for information, knowingly or unknowingly, use the Google search engine and their queries are thus captured by Google Trends. Google has a market share of well over 95 percent of the search engine market in the vast majority of countries of the world. China is the major exception with most searches being done on Baidu.

turing the occurrence of natural disasters and violent conflict at origin and a wide range of socio-economic, labor and monetary statistics in both origin and destination countries. The richness of this dataset goes clearly beyond those that are typically used in the migration literature and allows us to exploit monthly frequency in the panel of bilateral refugee flow corridors.³ Second, we compare the forecasting performance of different models and approaches that are common in data science forecasting applications. We test model performance in a forecasting-at-scale exercise where the models are assessed based on their average forecasting power over many migration corridors and across a range of different specifications that involve different compositions of information sets, including high-frequency Google Trends. In other words, our objective is to identify the best aggregate performing forecasting model specification, composed of a particular combination of available predictor candidates, including Google Trends. This is inherently different from existing work, which has focused on assessing corridor-specific customized forecasting approaches using a single algorithm (Carammia et al., 2022). Our approach allows us to analyze model performance differences at the corridor- as well as at the aggregate level. Third, we provide novel insights into the forecasting power of Google Trends data by systematically comparing forecast performance across specifications with and without these indices, as well as assessing their performance in specifications using only Google Trends predictors. Given that Google Trends are available at up to daily frequency and close to real-time, the latter specification, if well-performing, would hold particular attractiveness for the analyst as it is independent of data publishing lags, which typically restrict the viability of short-term forecasts decisively.

Our results show that the ensemble forecast composed of the Random Forest (RF) and Extreme Gradient Boosting (XGBoost) algorithms based on a host of classical push- and pull-factor in combination with around 200 unique GTI predictors consistently outperforms alternative machine learning methods such as Elastic Net and Factor Approaches. The ensemble forecasting model also performs better than our evaluation benchmark - the Random Walk⁴ - for forecasting horizons of 3, 6, and 12 months forecasts out-of-sample. We also show that these positive results carry over to a more parsimonious ensemble forecasting model exclusively based on GTI predictors for the group of large refugee flow corridors of high policy-relevance. This simplified model is available on a close to real-time basis, independently of the data availability and publishing lags of classical predictor variables. The results thus suggest that our approach may be successfully exploited in refugee forecasting tools to create ahead-of-period predictions for forecasting horizons of between 3 and 12 months.

³By corridor we mean a single origin-destination relationship, such as migration from Afghanistan to Germany.

⁴The Random Walk prediction is a "naive" forecasting technique in which the forecast for the next period is equivalent to the value of the dependent variable of the previous period.

3.2 Literature on Migration Forecasting

The literature on international migration is rich and multi-disciplinary. Here, we briefly review a selection of the studies most closely related to our research. For an overview of the models used and common data sources, see Appendix 3.A.1.1 and 3.A.1.2. Since refugee migration flows depend on many social, economic, and political factors, forecasting them is a highly complex and complicated undertaking. Shock events such as violent conflicts, economic crises, or policy changes make accurate prediction a difficult task. Furthermore, forecasts may be subject to large prediction errors because the data quality of migration and refugee flows is often poor (Disney et al., 2015). A related problem is the absence of a universal theory to explain the push- and pull factors of migration and refugee flows, which could provide a guideline regarding the choice of variables to include in such forecasting exercise (Bijak et al., 2019). For these reasons, the migration forecasting literature has focused largely on developing early warning systems aimed at i) predicting short-term fluctuations or ii) on in-sample exercises to gauge the importance of different push- and pull factors empirically. For example, Napierała et al. (2021) develop an early warning model using high-frequency (i.e., weekly and monthly) data of asylum application registrations in several European countries. Their model is inspired by statistical control theory and generates alerts whenever they detect that a threshold in the number of asylum applications is passed. Shellman and Stewart (2007) build an early warning model predicting civil violence, poor economic conditions, and foreign interventions, all factors leading to an increase in forced migration. The model is then applied to the case of Haitian citizens fleeing to the US on a weekly basis. While the model performs well in the specific case under consideration, the paper does not investigate out-of-sample performance in other corridors.

In a paper closely related to ours, Carammia et al. (2022) develop an early warning and forecasting pipeline using data from Google Trends applied to monthly asylum data for EU destination countries to forecast arrivals. While their research objective and important data sources coincide with those used in our paper, their approach differs significantly from ours. Most importantly, the authors assess a *corridor*-specific forecasting approach using a single algorithm (here: Elastic Net). They apply a three-step procedure that involves an early warning step that selects input variables for the forecasting steps. The latter consists in first estimating an Adaptive Elastic Net on the selected variables and, second, using standard time-series models to forecast the explanatory variables retained by the Elastic Net as new data inputs. This is different from our approach which leaves the predictor selection task fully to the respective forecasting model, thus keeping all potential information available for prediction *ex-ante* and reducing uncertainty arising from introducing additional modeling steps. Consequently, our ambition is to

assess the forecasting performance at scale across many corridors as opposed to creating a corridor-specific model. In addition, they use Google Trends bundled in “topics”, which are clusters of keywords. Such clusters leave no control over the choice of keywords and cannot be combined with destination country names as they are pre-specified by Google. Using customized keywords allows us to construct migration-related “bilateral” search terms tailored to the specific origin-destination relation such as “visa Germany”. This approach yields corridor-time-specific variation in Google Trends, which likely increases the predictive power of our GTI information set. Lastly, the forecasting performance in their paper is evaluated against an ARIMA(1,0,1) model in terms of percentage errors for a narrowly selected time period between 30 April 2017 to 1 September 2019 - a period which was characterized by stationary asylum lodgings that fluctuated around their mean values for many migration corridors. However, such stationary models are not appropriate when modeling trending episodes in the data, which is precisely one of our objectives. The important spikes in migration flows during the refugee migration episode of the years 2015/2016 are, therefore, left out of their assessment. We take a different approach and include all data available from our sources and cover the maximum time period available between January 2008 and April 2021. This implies that we are testing our models on the challenging task of predicting all large forced migration events during that time period across all corridors, including the major refugee spells of the years 2015/2016.

3.3 Data

To test the out-of-sample performance of the previous forecasting models, we rely on a panel data set of bilateral asylum seeker flows between 150 origin countries to the EU27 destination countries with monthly frequency. We combine this panel with a wide range of potential predictor variables, composed of almost 200 migration-related Google Trends variables and “classical” push- and pull-factor predictor variables. Our primary outcome variable is the monthly number of asylum applicants registered in an EU member state by country of origin of the applicant as provided by Eurostat. For the collection of Google Trends data, we follow Böhme et al. (2020) and use a list of migration-related search terms that we collect in several languages. In what follows, we describe the data extraction and construction procedure and provide brief descriptive statistics.

3.3.1 Asylum Seeker Flows

The composition of the panels we use in the forecasting exercise is dictated by the availability of the target variable, i.e., asylum seeker flows. The Eurostat database contains asylum applications from individuals of international origins to the EU27 plus

the United Kingdom and Norway. The first observations date back to January 2008.⁵ We drop countries with populations of less than 100,000 from the sample. Moreover, we discard origin countries without monthly migration flow data.

In particular, the variable we use to capture refugee flows is “asylum applicants” as registered in a member state by country of origin of the applicant. This variable is recorded as a continuous count variable, rounded to five. In the Eurostat database, an asylum applicant is defined as a person who has submitted an application for international protection or was included in such an application as a family member during the reference period. By definition, asylum-seeking individuals differ from other types of people on the move (who may be engaged in voluntary movement). While they may cross a border *irregularly* before submitting their application for international protection, asylum seeker movement is considered a *forced displacement*, or involuntary.⁶ Given the original aim of the ITFLOWS project to assist policymakers and civil society in managing “mixed migration” flows to the EU, including both irregular migration and refugee flows, we consider the number of asylum applications the most suitable outcome variable for this prediction exercise. Also, given the continuous nature of this variable, it is more challenging to predict, compared to alternative options such as a binary indicator for an increase or decrease in the number of applicants.

3.3.2 Languages

To determine the languages for which we want to extract Google Trends data, we proceed in the following way. First, we collect all country names that were defined as either “origin countries,” “transit countries,” “potential additional origin countries,” and “watchlist origins” within the ITFLOWS project. Second, we cross-reference the languages spoken in these countries with the languages associated with each country in the database of Melitz and Toubal (2014). While there are many more languages spoken in these countries, we focus on languages that are defined as “official languages”. Their definition of “official languages” requires a language to be the official language as per the

⁵Since the refugee data provided by The European Border and Coast Guard Agency (Frontex) aggregates border-crossing flows by routes rather than by destination country, we choose the Eurostat data source, which enables us to introduce a bilateral dimension of refugee flows from a specific origin to a specific destination. See, <https://frontex.europa.eu/we-know/migratory-map/>, accessed February 26, 2023.

⁶The terms “migrant”, “refugee” and “unrecognized refugee” are understood as follows: In its widest scope and for some authors, the term “migrant” also includes asylum seekers and refugees. However, we emphasize that the two latter categories are regulated by additional instruments and are given specific guarantees in international and European law. Individuals who have been formally granted refugee status are understood as “recognized refugees.” Any person (asylum applicant, irregular migrant or not) who meets the eligibility criteria, but have not applied or applied and have not yet been granted asylum by a state is understood as a “non-recognized refugee” (to be distinguished from unsuccessful asylum applicants); and both recognized and unrecognized are refugees.

CIA World Factbook in at least two countries. This approach results in the following languages being considered in this paper: English, French, Arabic, Spanish, Portuguese, Turkish, Farsi, Pashto, Hausa, and Fulfulde.⁷

The origin countries are then sorted into different categories by the languages spoken in each country. To associate each origin country to one or more languages, we use the database provided in Melitz and Toubal (2014). We do not discriminate between spoken, common, and official languages. If a language is referenced as either of the three, we associate the origin country with that language. Moreover, those countries that are not in the Melitz and Toubal (2014) database are added to the English origins list unless an official language from our set can be identified from other sources, such as the CIA World Factbook. As shown in Appendix Table 3.A2, the categorization by language leads to 86 English, 30 French, 27 Arabic, 24 Spanish, 9 Portuguese, 7 Turkish, 6 Farsi, 6 Fulfulde, 2 Hausa, and 2 Pashto origin countries.

3.3.3 Keywords

The selection of keywords is crucial for our approach, as we need to cover a set of search terms that refugees will likely look up before (or during) their trip. To cover a broad and relevant set, the keyword selection strategy rests on two pillars. First, we follow Böhme et al. (2020) by extracting several terms related to each of the two base keywords “immigration” and “economic” through the website *Semantic Link*. This service checks the English version of the *Wikipedia* encyclopedia for the most common co-occurrences between the given keywords and words that appear at least 1,000 times in the database. By choosing the initial terms “immigration” and “economic”, we aim to cover two main fields of search terms related to migration intent. Of the 200 keywords generated in this way, we discard 27 already included in the set of 67 keywords in Böhme et al. (2020). Then, we manually exclude 20 terms from the list which are seemingly ambiguous or irrelevant. These are mostly numbers, names, and acronyms. In total, from this first exercise, we obtain a list of 153 keywords.

Second, we feed a list of keywords consisting of the 153 terms from the *Semantic Link* exercise and the 40 unique keywords remaining from Böhme et al. (2020) for a total of 193 keywords into the Google Trends API to fetch the so-called *related queries* in each of the 86 English-speaking countries of origin. The *related queries* category outputs the Google searches similar to the keyword the user provided to the API, which occurs

⁷We widen the set for Arabic by including also the Algerian, Moroccan, Egyptian, Saudi Arabian, and Hassaniya versions, as well as the Afghan type of Farsi called Dari. Comparing some results, there were no substantial gains from tailoring the languages to the specific dialects, which is why we focus simply using standard Arabic and Farsi in the particular cases mentioned.

most often in a given geographic area. For example, queries related to the search term “visa” are “application”, “authorization”, or “lottery”. By including such related queries, we aim to complement the *Semantic Link* list with frequently searched terms related to our two categories of “immigration” and “economic” conditions. The output from this exercise is then manually scraped for the most relevant new terms, which yields 37 additional terms. Manual selection is necessary since many terms are not strictly related to human migration, such as the word “server migration” which refers to an information technology process.

Combining the lists obtained from the two exercises and the unique terms in Böhme et al. (2020) results in a list of 230 keywords. Since, in many cases, we have singular and plural forms (e.g., “migrant” and “migrants”) or semantically similar terms (e.g., “application” and “applicant”), we combine these terms using the search operator “+”. This operator ensures that the query delivers all searches for all terms connected by it. For example, a Google Trends search for “migrant+migrants” will provide data for a combination of the interest over time in both terms. By combining terms in this fashion, we obtain a final list of 192 keyword expressions. The full list is reported in Appendix Table 3.A3.

This English list is the starting point for translating all terms into the selected languages. In the case of the Roman languages in our set, we adapt the translation always to include male/female adjustments and accented/not-accented versions when applicable. Again, we add these additional spellings using the “+” operator to keep the different lists per language of equal lengths and comparable.

3.3.4 Google Trends Data

An advantage of our Google Trends customized search term approach is that we can exploit bilateral variation in search intensities in origin countries concerning a specific destination country. To obtain these, we interact the 192 keywords with each of the EU27 country names (see Appendix Table 3.A4) to obtain bilateral search terms. That yields a combination of topical keywords related to migration in relation to a specific destination country, e.g., “visa Germany” or “embassy France”. For each origin country, we add these 5,184 (192×27) bilateral terms to the list of keywords. This list also contains the original list of 192 topical keywords and the individual EU27 destination country names separately. Examples are search terms like “visa” or “unemployment” for topical keywords and “France” for EU27 destination country names. The final list has 5,403 entries and is the same length for each language considered.

We extract the Google Trends data for each of the terms specified on the final list of keywords through an API provided by Google. We proceed by language groups to account for the fact that people likely use Google searches in their native language. We collect this data at a monthly frequency between January 2004 and April 2021, the maximum time period available at the time of data construction. A noteworthy measurement issue with Google Trends data is that the API returns zero values when the search volume falls below an undeclared threshold. The resulting positive values of the index reflect what Google terms “interest over time” and are calculated as a fraction of searches for the given term relative to the total number of searches over a month in the same country. In the second step, the data for the entire period is automatically scaled to values between 0 and 100 such that 100 corresponds to the highest relative search intensity in a month measured over the entire period.⁸ Table 3.A3 lists all 192 search terms and provides descriptive statistics for each resulting time series.

3.3.5 Classical Predictors

We also collect a host of monthly economic indicators for each origin country and each destination country in our sample. These variables correspond to more “classical” predictors as used in the economics literature on migration and are supposed to capture various migration push- and pull factors, as listed in Appendix Table 3.A5 and 3.A6. We have an abundance of variables to choose from in the macroeconomic and agricultural domains, with over 300 variables collected for each. In addition, we have over 100 variables that sufficiently cover conflict and disaster data, political variables, labor force data, and short-term business data.

The first set of “classical” predictors of asylum seeker flows are indicators capturing socio-economic dimensions on both sides of the migration corridor. For example, macroeconomic and financial indicators at the origin country level, such as GDP growth or Consumer Price Index, provide proxies of the current economic situation during a specific month of observation and may provide signals on push factors. In contrast, the same variables at the destination level constitute proxies for the attractiveness of the EU27 country and may thus capture pull factors. For monthly macroeconomic statistics, we combine sources including the Statistical Office of the European Union (Eurostat), the International Labour Organization (ILO), and the International Monetary Fund (IMF).

⁸For example, suppose there are only three months in the sample, and the fractions of searches for the keyword “visa” relative to the entire search activity in a given country are 0.1, 0.2, and 0.5. Then 100 would be assigned to the third month, 40 to the second, and 20 to the first.

An essential push factor is human-caused or natural shocks in the origin country. Adverse environmental shocks and political conflicts may decrease income and cause social instability. Political events and indicators from the Rulers, Elections, and Irregular Governance Dataset (REIGN), political violence indicators from Armed Conflict Location and Event Data Project (ACLED), and disaster indicators from International Disaster Database (EM-DAT) are merged into the monthly panel. Event entries are aggregated into monthly counts.

3.4 Forecasting Methodology and Performance Comparison

Migration data is often marked by changing dynamics and fluctuations in the set of relevant predictors. Rossi (2021) discusses the effect of such instabilities on forecasting methods and performances and suggests that Big Data approaches may help to improve forecasts. In particular, she suggests considering three different approaches: First, aggregating, then forecasting as in principal component models. Second, forecasting while aggregating as in regularization models. Third, forecasting and then aggregating by computing ensemble predictions across different models. Based on these suggestions, we select a range of empirical models and combinations thereof to forecast migration flows that have been used in the migration or violent conflict forecasting literature and then rank them by their forecasting performance. The models we compare are the Elastic Net (Carammia et al., 2022), the Random Forest (Mueller & Rauh, 2022), as well as the Extreme Gradient Boosted Regression Tree and factor models. To assess the forecasting performance of these approaches, we compare the forecasts to those obtained from our benchmark model, which is obtained by using the last available data point as the forecast at all horizons. We refer to this benchmark as the Random Walk (RW) forecast. For a discussion of the weaknesses of our approach, see Appendix 3.A.3.

The model comparison is carried out individually over all possible bilateral relationships of countries of origin and EU27 destinations in our monthly data set. The variable to be forecast is the number of *asylum applications* by individuals of nationality from a country of origin o in a country of destination d , which is the best proxy of refugee flows available to us for the member states of the European Union. The information set used to make the forecast is denoted by X_{od} and contains different combinations of migration push- and pull factors such as, e.g., socio-economic characteristics in origin- and destination countries, migration-related GTI variables, and the lagged dependent variables. To compare the performance of different model compositions in terms of the information set, for each algorithm, we perform three sets of forecasts: first, motivated

by the migration literature in economics using bilateral flow data in a gravity-type regression specification, we specify the predictors to be only the “classical” push- and pull factors. Second, based on the first specification, we add our GTI data to the vector of predictors. Third, we perform the forecast using only GTI variables without any “classical” predictors. The latter specification is particularly appealing for the analyst as it involves only one data source (Google Trends) available at up to daily frequency starting in 2004 and with a maximum publishing lag of 24 hours. That provides obvious advantages for the viability of conducting forecasts close to real-time.

We conduct a moving-window pseudo-out-of-sample exercise to assess the accuracy of the out-of-sample forecasts. That is, we use a window of constant size $w = 50$ to train each model and predict the number of *asylum seekers* at different forecast horizons, namely $h = 1, 3, 6, 12$ months ahead. The choice of the window size is arbitrary, but experimenting with different sizes has shown that longer training sets do not necessarily improve performance or change the ranking of models. Importantly, given that the sample starts in January 2008, this choice implies that all models are asked to forecast the large migration change in 2015, which is a novelty in the migration prediction literature. We obtain all forecasts using a direct approach. This requires offsetting the dependent and independent data by the forecast horizon for training and then feeding the latest available independent data to the trained model for prediction. Although we do not use real-time data for training or evaluation, this is a proper out-of-sample forecast. Moreover, note that the explanatory data contains only the lag corresponding to h and no deeper lags that would add potentially useful information. We check lag-augmentation as a robustness exercise and find limited use for it.

Motivated by our objective, we select three forecasting models that can deal with variable selection, which is a crucial criterion given our rich data set. These are the Elastic Net, the Random Forest, and the Extreme Gradient Boosted Regression Tree (XGBoost). While the Elastic Net is a linear model, the latter two can accommodate non-linearities. For performance comparison purposes, we rely on the Theil ratio (T), which we compute as the ratio of the root mean squared error ($RMSE$) statistics of the candidate model over the Random Walk, as follows:

$$T = \frac{RMSE_X}{RMSE_{RW}}$$

with X representing the respective model under investigation.⁹ A value larger than 1

⁹The RMSE is calculated as: $\sqrt{(\frac{1}{n}) \sum_{t=1}^n (\hat{y}_{t|h} - y_t)^2}$, with h being the respective forecasting horizon.

thus indicates that on average over all migration corridors of the selected sample, the Random Walk outperformed the specific model. A value smaller than 1 implies that the candidate model outperforms in terms of forecasting performance over the forecasts using the Random Walk. In what follows, we briefly discuss the forecasting models under investigation.

3.4.1 Elastic Net

The Elastic Net (EN) is a linear regression model which allows for parameter shrinkage according to the following penalized regression:

$$\hat{\beta}_{EN} = \arg \min_{\beta} |y - X\beta|^2 + \alpha\lambda|\beta| + (1 - \alpha)\lambda|\beta|^2$$

Parameter regularization is performed against the L^1 (LASSO) and the L^2 (Ridge) penalties which ensure model selection through the LASSO part and allow for some control for multicollinearity through the Ridge part. We set the weight on each of the two to $\alpha = 0.5$ as in Carammia et al. (2022) and select the regularization parameter λ as the one yielding the smallest MSE in training.¹⁰

3.4.2 Random Forest

The Random Forest (RF) is an ensemble prediction algorithm that relies on decision trees (Breiman, 2001). In each decision tree, the algorithm finds the optimal variable to split the input data, thus creating a node. The metric for computing the optimal split for our regression problem is the MSE. Nodes in the tree are grown using bootstrapped data sets ("bagging") obtained from the original inputs. While individual decision trees can be prone to bias, averaging across many individual trees can produce predictions that are less driven by idiosyncratic error. Three hyperparameters have to be set to run the analysis. Firstly, the number of trees, which we set to 10,000 to ensure that the many variables in our data set are used with high enough frequency in the regression trees. Secondly, the number of randomly chosen features by each tree m . Reducing the number of features chosen in each tree reduces both correlation and strength. Strength is defined as the error rate of each tree. Breiman (2001) shows that the forest error rate depends negatively on the correlation between any two trees and positively on the strength. We follow the rule-of-thumb recommendation to set $m = p/3$ where p is the number of variables. Thirdly, the depth of each tree represents the number of splits of each tree in the forest. We only require the standard minimum of five terminal nodes but allow the trees to be more complex at the cost of computational speed.

¹⁰The errors are assumed to follow a Poisson distribution.

3.4.3 Extreme Gradient Boosting Regression Tree

While the RF is an extension of decision trees by bagging, the Extreme Gradient Boosting Regression Tree (XG) advances the model by boosting it (Friedman, 2001). Building upon weak learners (in our case, the decision trees), the model calls the original method repeatedly, using a different subset of the data each time. By sequentially fitting over the residuals (MSE) of the previous weak learner, the model finally aggregates the results of the steps into one strong learner. While the sequential setup increases computational time, the model's prediction often has higher accuracy. For extreme gradient boosting, the first hyperparameter that needs to be set is the number of rounds B the model is run, which we keep at a high value of 10,000 to ensure convergence of the loss function at the risk of potentially overfitting the sample. The second hyperparameter is the learning rate, $\eta \in (0, 1)$, which controls the contribution of each past tree to the current approximation. By decreasing the learning rate, the model becomes more robust to overfitting but requires more computation time. We choose a low value of $\eta = 0.3$ to counter the overfitting problem we may obtain from the high number of rounds.

3.4.4 Factor Approach

Finally, we use a factor approach to forecast the following model for asylum applicants.

$$y_{t+h} = c + \beta' f_t + \epsilon_{t+h}$$

Kim and Swanson (2018) show that in the context of economic forecasting, such models can outperform simple benchmarks. Stock and Watson (2002) show that the factors f_t can be consistently estimated using the Principal Component (PC) estimator. After extracting PCs from the set of predictors, the model is estimated assuming a Poisson error distribution.

3.4.5 Ensemble Forecasts

Ensemble forecasts are constructed by equally weighting the forecast made by each model in the given combination and summing up. Suppose, for example, the Random Forest predicts 100 asylum seekers, and the XG Boost method suggests 120 asylum seekers. Given that there are two models in the combination, the equally weighted ensemble forecast would be $1/2(RF + XG) = 110$. Such ensemble forecasts can sometimes improve upon single model forecasts (Kim & Swanson, 2018).

3.5 Results

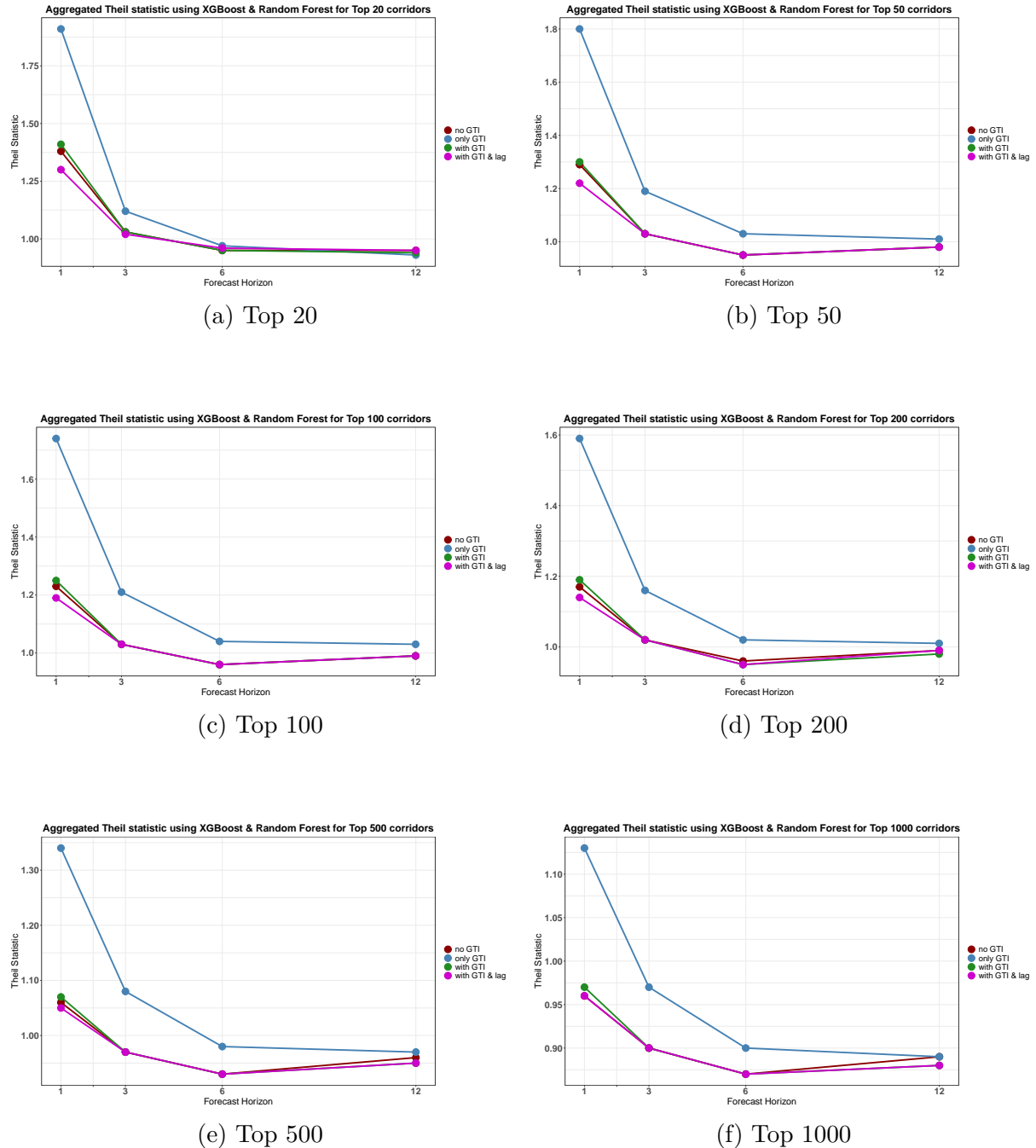
We start the discussion of our forecasting exercise by focusing on the best-performing model, the ensemble forecasting model composed of the Random Forest and XGBoost algorithms. The main results from the ensemble model are summarized in Figure 3.1.¹¹ The figures depict the Theil ratio for six subsamples of bilateral corridors of different importance for refugee flows, ranging from the top 20 most important corridors, in terms of total asylum seeker numbers over the 2008 to 2020 period, to the top 1000 corridors.¹² Each panel (a)-(f) reports results from a different subsample. In each panel, the horizontal axis represents the four forecasting horizons of 1, 3, 6, and 12 months and there are four lines corresponding to different specifications of the information set used in each forecast (i.e., no GTI, only GTI, with GTI, and with GTI and lagged dependent variable). The vertical axis, on the other hand, reflects the Theil ratio of the RF-XG ensemble model against the benchmark forecast based on the RW.

We start by describing the general findings that emerge from comparing the out-of-sample forecasting performances across the six subsamples. A recurrent pattern in all panels is that the shape of the lines is convex with an L-shape, indicating that the performance of the RF-XG ensemble model is generally better for longer forecasting horizons. In particular, performance for the one-month forecasting horizon is relatively poor, as reflected by the Theil statistics larger than one, especially in panels (a) through (e), indicating that the RW outperforms, on average, in the case of short-term forecasts in the top 500 group. We believe that this is due to two factors: First, the RW “forecast” may be relatively more powerful in the very short term in which the temporal proximity to the last month is small, and hence the number of asylum seekers of the last period may be a good approximation for that of the current period. Second, the predictive power of our candidate models hinges on the fact that classical predictors and/or digital trace data measured at the origin during the previous time period (i.e., last month) are predictive of registrations of asylum seekers from that origin at the destination. Considering the top 20 corridors, it may easily take several months for irregular migrants and refugees to arrive at their desired destination and register for asylum. If this is the case, we expect our approach to have low predictive power in the very short term.

¹¹We report results based on the pure RF and XGBoost models in Appendix Figures 3.A1 and 3.A2, respectively. The general results for all models are reported in Appendix Table 3.A7.

¹²In the case of the top 20 corridors of asylum seeker flows to the group of EU27 destination countries this includes (in order of magnitude of total registered asylum cases by nationality of origin over the period 2008-2020): Syria-Germany [719,650], Afghanistan-Germany [265,215], Iraq-Germany [249,040], Syria-Sweden [130,265], Serbia-Germany [124,090], Nigeria-Italy [121,255], Venezuela-Spain [110,795], Albania-Germany [95,040], Iran-Germany [91,430], Syria-Greece [84,330], Pakistan-Italy [78,440], Syria-Hungary [78,355], Colombia-Spain [75,860], Eritrea-Germany [73,995], Afghanistan-Hungary [73,885], Russia-Poland [72,990], Afghanistan-Greece [71,680], Afghanistan-Austria [71,420], Afghanistan-Sweden [69,265], Syria-Austria [68,500].

Figure 3.1: Main forecasting results from the ensemble model composed of the Random Forest and XGBoost algorithms for six subsamples of bilateral corridors of different importance for refugee flows



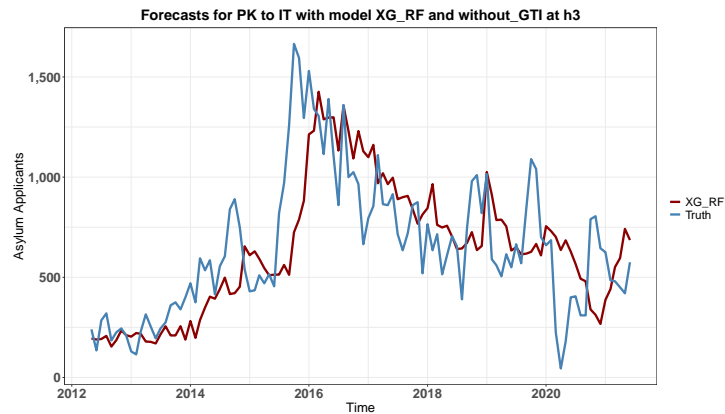
Note: Main forecasting results from the ensemble model composed of the Random Forest and XGBoost algorithms for six subsamples of bilateral corridors of different importance for refugee flows (top 20 - top 1000 corridors in terms of total asylum seeker numbers over the 2008 to 2020 period) over four forecasting horizons (1, 3, 6, and 12 months) and four specifications of the information set (no GTI, only GTI, with GTI, and with GTI and lagged dependent variable). The curves in each panel depict the Theil ratio of the ensemble model against the benchmark forecast based on the RW. Source: Author calculations.

A second general pattern relates to the relative performance of the four different compositions of the information set, depicted by each line. The GTI-only model (blue lines) generally performs relatively worse than the other specifications. This model shows the worst performance over all horizons for panels (b) through (e), indicating that, despite its attractiveness due to the ease of implementation, the GTI-only model is not a silver bullet for refugee flow forecasting. In other words, classical predictors seem to carry important signals that cannot be substituted entirely by digital trace data. Furthermore, for the remaining three models composed of only classical predictors (brown line), classical plus GTI (green line), and classical plus GTI and lagged dependent variable (purple line), the performance differences are often minor, indicating that the GTI variables provide limited additional predictive power over the (large) set of classical predictors used. Next, we analyze the relative performance of the ensemble model in comparison to the RW.

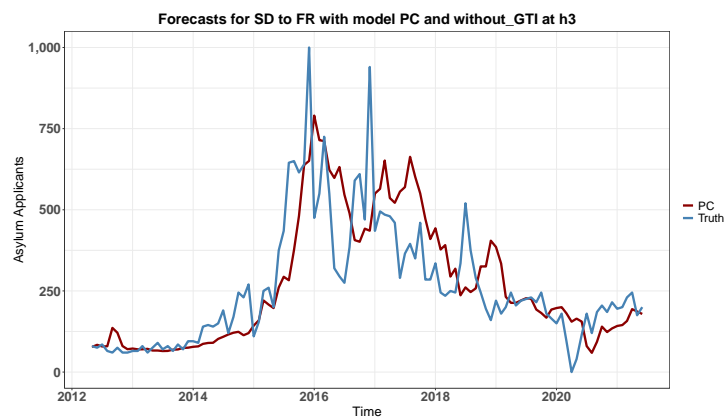
In all subsamples, we observe that the Theil statistic drops below one for any model including classical predictors at longer horizons, indicating that the ensemble model outperforms the RW. The precise horizon from which the ensemble model is superior to the RW in terms of forecasting performance varies between 6 and 1 months horizons for the top 20 to the top 1000 corridor subsamples, respectively. Interestingly, the magnitude of performance gains increases from high-importance to low-importance corridors. One of the reasons behind this pattern is that there tend to be fewer large shocks in refugee flows in low-importance corridors such that those time series are stationary, which, in turn, can be forecasted more accurately by the models. It is important to emphasize that the lines in Figure 3.1 represent the simple average of Theil statistics for each sample, respectively. This implies that the results for specific corridors may be significantly better or worse within the respective subsample.

We now turn to the role of our GTI predictors for the forecasting exercise. Focusing on the historically largest corridors for asylum seeker flows in absolute numbers in panel (a), we find that the GTI-only model outperforms the RW at horizons of six months and above. Further, the GTI-only model also performs equally well at $h = 6$ and even slightly better than any other specifications of the information set at $h = 12$. Since the curves represent average results across all corridors included in the sample, it is clear that the GTI-only model performs very well for some of these corridors. Due to the obvious implementation advantages of a forecasting model based on a single data source, this underlines the prospects of the GTI approach for selected policy prediction applications.

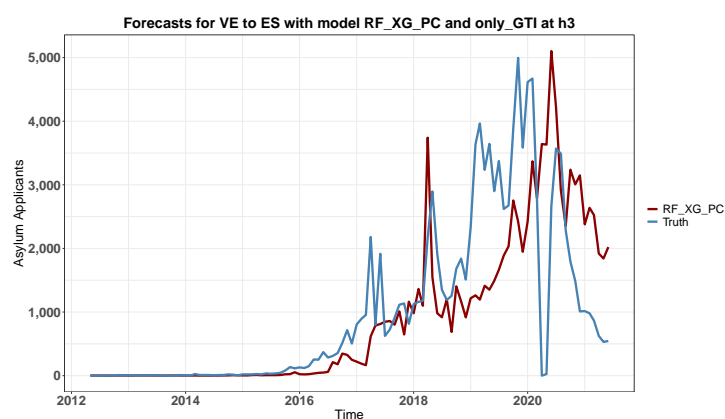
Figure 3.2: Time series plots comparing recorded asylum seeker flows (truth) to selected out-of-sample forecasts



(a) Forecast using RF-XG ensemble model in no-GTI specification for the corridor Pakistan to Italy at 3 month horizon



(b) Forecast using PC model in no-GTI specification for the corridor Sudan to France at 3 month horizon



(c) Forecast using RF-XG-PC ensemble model in only-GTI specification for the corridor Venezuela to Spain at 3 month horizon

We now turn to the discussion of the performance of the Elastic Net and Factor approach using principal components, as reported in Appendix Table 3.A7, which belong to the group of generalized linear regression models. Such models can extrapolate trends into the future, making them suitable for migration forecasting as shown, for example, in Carammia et al. (2022) for the case of EN. Given the large number of forecasts across corridors, horizons, and pseudo-samples we have to carry out, such forecasts must be appropriately automated. We follow their approach by setting the weight on the LASSO and Ridge component of the Elastic Net equal to $\alpha = 0.5$. The regularization penalty λ is chosen on a grid of 100 potential values on a log scale where the last value regularizes all parameters to zero. We select the λ associated with the smallest in-sample MSE. Regarding the principal components, a crucial choice is the number of principal components to use for the model. Typically, the variation in large data sets of socio-economic variables can be described by a small set of principal components that are useful for forecasting (Stock & Watson, 2002). While statistical criteria are available to determine the (in-sample) number of principal components to be used (Bai & Ng, 2013), this would render the exercise extremely time-consuming. Instead, we use a simple threshold and include principal components if they explain more than 5% of the dataset's variation.

These choices for the principal components and elastic net forecasts work well in many forecasts. However, in some cases, both models can grossly overfit the data or simply fail due to a lack of variation in the dependent variable. In such cases, the forecast errors can be nearly infinitely large, which then affects the aggregate error statistics accordingly. We observe this phenomenon in our results for these models. While some of the overestimates are obvious and could potentially be corrected by a human forecaster in the loop, this problem can be critical in the case of full automation of the forecasting process. Since the error sizes depend on the migration corridor at hand, it is not possible (and not prudent) to introduce arbitrary cutoffs for excluding certain values from the evaluation of the model. Despite the very poor relative performance across our sample, we emphasize that on certain migration corridors with good variation in the dependent variable and with largely stationary behavior, both the Elastic Net and the principal components estimator yield valuable forecasts.

In addition to the relative performance measures compared to the RW as reflected by the Theil statistics, we also report absolute forecasts for selected corridors in Figure 3.2. The three panels report time series plots comparing the recorded asylum seeker flows, depicted by the blue lines (i.e., ground truth data), to the selected out-of-sample forecasts for specific forecasting models and horizons, represented by the red lines. The examples have been selected from the group of top 50 corridors conditional on outperforming the

RW (i.e., from those with a Theil statistic below one).¹³

Panel (a) depicts forecasts based on our best-performing RF-XG ensemble model in the no-GTI specification for the corridor from Pakistan to Italy with a 3-month horizon. A visual inspection shows that the model performs relatively well in this corridor, as reflected by the high correlation between the blue and the red line. Reassuringly, the model predicts both pronounced increases and decreases relatively well, for example, the sudden increase in flows during the 2015/2016 refugee wave. However, the inspection of the ground truth data around the beginning of the COVID-19 pandemic in early 2020 also shows that the model performs worse in the face of sudden disruptive shocks. The related shutdowns during the pandemic led to a general halt of human migration and refugee flows, as can be seen by the sudden drop in the blue lines across all panels in the first and second quarters of 2020. This is precisely the period when larger forecasting errors can be observed and the red line reflects that the model would have overpredicted refugee arrivals during that period (this applies to all three models depicted). Panel (b) depicts forecasts using the PC model in the no-GTI specification for the corridor from Sudan to France with a 3-month horizon. Note that this is an example of the PC model outperforming the RW in the specific corridor despite an extremely large average Theil statistic in the top 50 sample. Panel (c) depicts forecasts based on the RF-XG-PC ensemble model in the only-GTI specification for the corridor from Venezuela to Spain with a three month horizon. Here, the forecasting errors are generally negative, implying that the model overpredicts, especially starting in the year 2019. An extreme overprediction occurs during the pandemic when flows drop to zero and the model continues to overpredict thereafter. Despite these prediction errors, this example shows that the only-GTI approach can work relatively well for the top 20 corridors of high policy relevance. Overall the figures provide evidence of the good absolute forecasting performance of selected models and compositions of the information set, in the context of specific refugee flow corridors.

3.6 Towards a Feasible Forecasting Approach for Refugee Flows

This paper evaluates the practical feasibility of building a refugee flow forecasting model combining high-dimensional data with machine learning techniques. We implement an out-of-sample forecasting model as suggested by Böhme et al. (2020). We extract migration-related Google Trends time series that can be used as predictors for bilateral refugee flows and combine them with asylum seeker flows and an extensive range of

¹³Among the Top 50 corridors there are a total of 1436 corridor/model/information set/horizon combinations where the model outperforms the Random Walk. Among these, 103 come from the Elastic Net, Factor Models, or their ensemble. The examples presented here are chosen purely for illustration.

classical predictor variables with monthly frequency to construct a high-dimensional data set for performance testing. We have built a bilateral refugee panel database capturing flows between more than 150 origin countries and the group of EU27 destination countries with monthly frequency. We then evaluate the out-of-sample performance of different statistical models against a naive forecasting procedure constituted by the Random Walk, which tends to be “hard to beat” in forecasting exercises.

The results are encouraging in the following sense. Among the machine learning models we analyze, the ensemble forecast composed of the Random Forest and XGBoost models shows the best average performance in our context. In particular, we have presented evidence that this model consistently outperforms the Random Walk for forecasting horizons of 3, 6, and 12 months forecasts out-of-sample, depending on the sample of corridors analyzed. Our results also show that when comparing the predictive power of the Google Trends predictors in the same specification to those of hundreds of “classical” predictors capturing different types of push- and pull factors, performance gains from the GTI predictors are marginal on aggregate. This underlines that digital trace data is not a silver bullet for refugee flow forecasting if the forecaster includes a large vector of predictor candidates proxying for a diverse set of migration push- and pull factors.

Yet, focusing on the subsample of the top-20 corridors in terms of aggregate refugee flows between 2008 and 2020, we have shown that the positive results carry over to a specification of the ensemble model exclusively based on GTI predictors. The latter has the practical benefit of being available on a close to real-time basis, independently of the data availability and publishing lags of “classical” predictor candidates such as GDP growth or consumer price indices. In this selected sample of high policy relevance, the GTI-only specification outperforms all other models on average, including the Random Walk, for forecasting horizons of 6 and 12 months. This average performance improves further when focusing on specific corridors within the group of top 20 corridors. Our results imply that the Google Trends indices we extracted offer positive predictive power for such corridors, which could potentially be exploited for a refugee forecasting tool customized to selected corridors of particularly large flows.

For policy applications, we thus recommend customizing the selection of forecasting models according to the maximum performance within a specific refugee corridor, following our approach. We have provided average results and selected examples that represent high-performance combinations of the forecasting model, the information set, and the forecasting horizon. For origin countries with poor push-factor data availability, we suggest including Google Trends data, particularly in cases where the forecast horizon is longer than three months. Despite the wide information set used in this exercise, our

findings also reflect the fundamental limitations of the forecasting exercise: anticipating large and sudden shocks, such as the COVID-19 pandemic, which lead to an unexpected halt in migration flows, is impossible. Our approach is not exempted from that limitation and the selected examples show this clearly for the forecasts during the year 2020 and beyond.

It is important to note that the performance statistics we present are averaged over the respective bilateral refugee flow corridors. As discussed, there is heterogeneity in the models' predictive performance across corridors. In other words, the models tested work better for some corridors and worse for others. In other words, for specific corridors, the single best-performing model may be different than the one suggested by the average results. This heterogeneity should be made transparent whenever forecasting results are published or used to inform policymakers and migration management agencies. Another disclaimer is that this exercise's only available benchmark for predictive performance is data from previous years. There is no guarantee that the proposed approach will successfully predict future flows. Anyone applying the method should investigate the approach critically in their specific use case before scaling it up. Yet, overall, our results show that the approach does help enable refugee forecasting applications with ahead-of-period predictions.

Bibliography

- Adema, J. A. H., & Guha, M. (2022). Following the online trail of Ukrainian refugees through Google Trends. *CESifo Forum*, 23(04), 62–66.
- Alexander, M., Polimis, K., & Zagheni, E. (2019). The impact of Hurricane Maria on out-migration from Puerto Rico: Evidence from Facebook data. *Population and Development Review*, 617–630.
- Alexander, M., Polimis, K., & Zagheni, E. (2020). Combining social media and survey data to nowcast migrant stocks in the United States. *Population Research and Policy Review*, 1–28.
- Alvarez-Plata, P., Brücker, H., & Siliverstovs, B. (2003). *Potential migration from Central and Eastern Europe into the EU-15: An update*. European Commission, Directorate-General for Employment; Social Affairs.
- Bai, J., & Ng, S. (2013). Principal components estimation and identification of static factors. *Journal of Econometrics*, 176(1), 18–29.
- Bijak, J., Disney, G., Findlay, A. M., Forster, J. J., Smith, P. W., & Wiśniowski, A. (2019). Assessing time series models for forecasting international migration: Lessons from the United Kingdom. *Journal of Forecasting*, 38(5), 470–487.
- Böhme, M. H., Gröger, A., & Stöhr, T. (2020). Searching for a better life: Predicting international migration with online search keywords. *Journal of Development Economics*, 142, 102347.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
- Burzynski, M., Deuster, C., & Docquier, F. (2020). Geography of skills and global inequality. *Journal of Development Economics*, 142, 102333.
- Carammia, M., Iacus, S. M., & Wilkin, T. (2022). Forecasting asylum-related migration flows with machine learning and data at scale. *Scientific Reports*, 12(1), 1457.
- Dao, T., Docquier, F., Maurel, M., & Schaus, P. (2018). Global migration in the 20th and 21st centuries: The unstoppable force of demography. *Review of World Economics*, 157, 417–449.
- Disney, G., Wiśniowski, A., Forster, J. J., Smith, P. W., & Bijak, J. (2015). Evaluation of existing migration forecasting methods and models. *ESRC Centre for Population Change, University of Southampton*.

- Dustmann, C., Casanova, M., Fertig, M., Preston, I., & Schmidt, C. M. (2003). *The impact of EU enlargement on migration flows*. Research Development; Statistics Directorate, Home Office.
- Eurostat. (2022). *Annual asylum statistics*. Retrieved March 18, 2022, from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Annual_asylum_statistics
- Fertig, M., & Schmidt, C. M. (2005). *Aggregate-level migration studies as a tool for forecasting future migration streams*. Routledge.
- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 1189–1232.
- Golenvaux, N., Alvarez, P. G., Kiossou, H. S., & Schaus, P. (2020). An LSTM approach to forecast migration using Google Trends. *arXiv preprint arXiv:2005.09902*.
- Hanson, G., & McIntosh, C. (2016). Is the Mediterranean the new Rio Grande? US and EU immigration pressures in the long run. *Journal of Economic Perspectives*, 30(4), 57–82.
- Hausmann, R., Hinz, J., & Yildirim, M. A. (2018). Measuring Venezuelan emigration with Twitter. *Kiel Working Paper*, (2106).
- Kim, H. H., & Swanson, N. R. (2018). Mining big data using parsimonious factor, machine learning, variable selection and shrinkage methods. *International Journal of Forecasting*, 34(2), 339–354.
- Kleinberg, J., Ludwig, J., Mullainathan, S., & Obermeyer, Z. (2015). Prediction policy problems. *American Economic Review*, 105(5), 491–95.
- Melitz, J., & Toubal, F. (2014). Native language, spoken language, translation and trade. *Journal of International Economics*, 93(2), 351–363.
- Mueller, H., & Rauh, C. (2022). The hard problem of prediction for conflict prevention. *Journal of the European Economic Association*, 20(6), 2440–2467.
- Napierała, J., Hilton, J., Forster, J. J., Carammia, M., & Bijak, J. (2021). Toward an early warning system for monitoring asylum-related migration flows in Europe. *International Migration Review*, 56(1), 33–62.
- Palotti, J., Adler, N., Morales-Guzman, A., Villaveces, J., Sekara, V., Garcia Herranz, M., Al-Asad, M., & Weber, I. (2020). Monitoring of the Venezuelan exodus through Facebook’s advertising platform. *Plos one*, 15(2), e0229175.
- Rossi, B. (2021). Forecasting in the presence of instabilities: How we know whether models predict well and how to improve them. *Journal of Economic Literature*, 59(4), 1135–90.
- Sardoschau, S. (2020). *The future of migration to Germany: Assessing methods in migration forecasting*. DeZIM-Institut.
- Shellman, S. M., & Stewart, B. M. (2007). Predicting risk factors associated with forced migration: An early warning model of Haitian flight. *Civil Wars*, 9(2), 174–199.

- Spyratos, S., Vespe, M., Natale, F., Weber, I., Zagheni, E., & Rango, M. (2018). Migration data using social media: A European perspective. *Publications Office of the European Union*.
- Stock, J. H., & Watson, M. W. (2002). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, 97(460), 1167–1179.
- Suleimenova, D., Bell, D., & Groen, D. (2017). A generalized simulation development approach for predicting refugee destinations. *Scientific reports*, 7(1), 1–13.
- Wanner, P. (2021). How well can we estimate immigration trends using Google data? *Quality & Quantity*, 55(4), 1181–1202.
- Zagheni, E., Garimella, V. R. K., Weber, I., & State, B. (2014). Inferring international and internal migration patterns from Twitter data. *Proceedings of the 23rd international conference on world wide web*, 439–444.
- Zagheni, E., & Weber, I. (2012). You are where you e-mail: Using e-mail data to estimate international migration rates. *Proceedings of the 4th annual ACM web science conference*, 348–351.
- Zagheni, E., Weber, I., & Gummadi, K. (2017). Leveraging Facebook’s advertising platform to monitor stocks of migrants. *Population and Development Review*, 721–734.

Appendix

3.A.1 Literature

3.A.1.1 Trends in Approaches

Traditionally, the most common approach to predicting migration flows has been to apply **time series models** (Disney et al., 2015). In the most basic specifications, time series models predict future migration based only on past migration, using autoregressive models. For example, Dustmann et al. (2003) apply this method to predict immigration to the United Kingdom and Germany following the EU enlargement in 2004. However, such simple models disregard factors other than past migration that determine refugee flows. Thus, they can only provide reliable predictions as long as these factors do not vary, such that migration flows remain stable over time. Indeed, the migration flows estimated by Dustmann et al. (2003) and other authors modeling migration after the EU enlargements in the 2000s (e.g., Alvarez-Plata et al. (2003) and Fertig and Schmidt (2005)) turned out to be erroneous and lacking power. To remedy this issue at least partly, it is possible to include additional control variables, such as unemployment rates, into autoregressive distributed lag models (Bijak et al., 2019).

Gravity-models add an important layer of complexity. They allow for the incorporation of factors from both the sending and the receiving countries as well as the bilateral relationship, thus going far beyond time series models, which tend to rely entirely on variations within one country. Gravity models use differences between countries and variation over time in these differences to predict future migration. Due to past data limitations, the popularity of gravity models has been on the rise only recently. For example, Hanson and McIntosh (2016) estimate the impact of geographical distance, the presence of shared borders and languages, former colonial ties, and GDP differentials on future immigration to several European countries. However, their results are only partly plausible because the model relies on average relationships and is not tailored to a specific origin or destination country. For Germany, Hanson's and McIntosh's model predicts that the stock of people born in foreign countries will be close to zero in 2020 and turn negative by 2030. Despite such examples of gravity models giving unrealistic results due to the inability to sufficiently tailor them to each country, the advantages of gravity models compared to time series models should not be underestimated. In addition to the characteristics mentioned above, they also allow the modeling of future migration flows to one country based on the experiences of other countries. Hence, scenario modeling becomes possible even for cases with a lack of data availability.

A third approach to predict medium-term migration flows is **theory-based structural modeling**. These models are primarily used to predict longer-term migration, as

the factors they consider gain relevance only in the long run. Dao et al. (2018) develop a structural migration model in which migration depends on income differentials between countries. In turn, these income differentials are spurred by differences in education. Burzynski et al. (2020) refine this model such that it considers differences in educational and labor costs, as well as countries' consumption levels. However, while these theory-based studies have the potential to model relevant drivers of migration, caution is also warranted when interpreting these findings. For example, Burzynski et al. (2020) forecast that about a million immigrants from Mexico would be present in Germany by 2020, which, as Sardoschau (2020) points out, is a considerable overestimation. Again, this points out the conflict between a model's aims (here: understanding long-run forces) and short- or medium-run predictions.

3.A.1.2 Trends in Data Sources

Migration predictions rely predominantly on official data (e.g. immigrant registries) and if explanatory variables are used, they typically come from national statistics or cross-national databases such as those of the World Bank.

In recent years, a new source of data has become available and is increasingly used in migration forecasting: Digital trace data, i.e., the traces individuals leave behind from online behavior. Examples include internet searches, locations tied to services such as email accounts, and social media data (see Table 3.A.1.2).

The main drawback of such data is that they are not representative of the overall population as there exists a self-selection of users into these specialized services. The generalization of the forecasts depends on who uses a service, e.g., in most countries, Instagram covers a far lower share of the population above the age of 50 than Facebook, which in turn is less representative of the population than the users of Google search. Google search, with a global market share of about 92 percent of all internet users who search online globally (the outlier being China with only 3 percent, which strongly decreases the global average), has a particularly broad user base. Yet, not everyone has access to the internet in the first place, with intersectional differences or disadvantages along lines of race, gender, ethnicity, and class, among others. Furthermore, as Böhme et al. (2020) summarize, searching for something online implies some interest in a topic but not necessarily that it will lead to behavior. Our literature survey shows that the majority of studies using "big data"-based approaches to predict migration flows use Google Trends and Facebook Ads Manager (see Table 3.A.1.2). These studies typically do not forecast future migration, except in the very short term ("nowcasting"). Mostly, they develop models and test if they can predict actual migration flows (i.e. the present).

As summarized by Böhme et al. (2020), a few applications use internet metadata to approximate migration dynamics and patterns. For example, Zagheni et al. (2014)

predict migration flows with the help of geo-referenced Twitter data while Zagheni and Weber (2012) perform the task based on IP addresses.

Data	Study	Sample	Methods and findings
Bing Maps: to obtain locations of major settlements and routing information between the camps, conflict zones and other settlements; combined with UNHCR and conflict data	Suleimenova et al., 2017	Refugees from Burundi, Central African Republic and Mali to neighboring several countries	Agent-based modelling, using a generalized simulation development approach. Simulations predict more than 75% of the refugee destinations correctly after the first 112 days. The simulations are validated using actual data
Facebook: Ads Manager, through which advertisers can select subgroups (e.g., Puerto Ricans living in California) and get an estimate of the “potential reach” (monthly active users)	Zagheni et al., 2017	Expats from several countries in the US	Proof of concept that people classifying themselves as “expats” on Facebook are quite representative of actual immigrant stocks in the US. Also evaluate biases in the data and show how accounting for these biases leads to better estimates and predictions; discuss the challenges and opportunities ahead in this area of research.
Facebook: Ads Manager (see above)	Spyratos et al., 2018	Expats in 17 EU countries	Based on Zagheni et al., 2017 (see above), but accounting for bias in Facebook data. Estimated the number of expatriates in 17 EU countries based on the number of Facebook Network users who are classified by Facebook as “expats”. Methodology allowed for the timely capture of the increase of Venezuelan migrants in Spain (though some other estimates were implausible).
Facebook: Ads Manager (see above)	Alexander et al., 2019	Refugees from Puerto Rico to the US	Diff-in-Diff to obtain estimates of the percent change in migrants. Estimated that there was a 17 percent increase in the number of Puerto Rican migrants present in continental US over the period from October 2017 to January 2018 (in line with previous studies)

Data	Study	Sample	Methods and findings
Facebook: Ads Manager (see above); combined with survey data	Alexander et al., 2020	Migrants from Mexico, India and Germany to the US	Combination of Bayesian Hierarchical Model, time series, spatial and data models. Produce timely ‘nowcasts’ of migrant stocks. Compare projections with data from American Community Survey, show that the model outperforms alternatives that rely solely on either social media or survey data.
Facebook: Ads Manager (see above)	Palotti et al., 2020	Refugees and migrants from Venezuela to other Latin American countries	Estimated monthly active users from Ads Manager vs. estimates from Response for Venezuela, a high correlation between the two ($r = 0.99$)
Google Trends Index (GTI): time series of search intensities of the user’s choice of keywords, by (origin) country. The GTI can be restricted by geographical area, date, a set of predefined general search categories.	Böhme et al., 2020	Search behavior of an estimated 842 million speakers from 107 countries of origin in which at least one of the three selected languages (English, French, Spanish) is officially spoken; 34 OECD destination countries	Filter Google data for keywords related to “immigration” and economics. Estimate a range of fixed effects panel models (with a range of controls) using two different specific actions: a unilateral and a bilateral (gravity) model. Approach provides strong additional predictive power for international migration flows when compared to reference models. Evidence, based on survey data, that measures partly reflect genuine migration intentions.
Google Trends Index (GTI); see above	Golenvaux et al., 2020	Data from Böhme et al., 2020, see above	A long short-term memory (LSTM) approach (=machine learning), combined with Google Trends data. It outperforms two existing approaches in predicting the one-year ahead incoming international migration: the linear gravity model (from Böhme et al., 2020 and an artificial neural network model.

Data	Study	Sample	Methods and findings
Google Trends Index (GTI), see above	Wanner, 2021	Labor immigrants to Switzerland from Spain, Italy, France and Germany	Linear regression was used in order to measure the relationship between the number of searches (x) and the number of moves (y) that finally occurred. The results show that Google Trends can predict, to some extent, current and future (short-term) migration flows of adults arriving from Spain or Italy. However, the predictions appear not to be satisfactory for other flows (from France and Germany).
Google Trends Index (GTI), see above	Adema and Guha, 2022	Refugee flows from Ukraine to the European Union and their spatial distribution in EU host countries	Linear regression was used in order to measure the correlation between the predicted number of refugees using Google Trends and residence permits (x) and the actual number of registered refugees (y) recorded in the destination countries. The results show a strong positive association between predicted and actual values and a high coefficients of determination.
Twitter: Streaming API, which provides a 1% random sample of all geo-localized tweets at any given moment	Hausmann et al., 2018	Emigrant from Venezuela	Define as migrants those people who tweeted in Venezuela in spring 2017 and from another country in spring 2018. Estimate that up to 2,9 million Venezuelans have left the country in the past year.
Twitter: Streaming API (see above)	Zagheni et al., 2014	Emigrants from OECD countries	Diff-in-diff approach to reduce selection bias when inferring trends in out-migration rates for single countries. Methods can be used to predict turning points in migration trends, which are particularly relevant for migration forecasting. Geolocated Twitter data can substantially improve understanding of the relationships between internal and international migration.

Data	Study	Sample	Methods and findings
Yahoo: Self-reported age and gender of anonymized e-mail users were linked to the geographic locations (mapped from IP addresses) from where users sent e-mail messages over time (2009-2011).	Zagheni and Weber, 2012	Emigrants from 11 European countries	Estimates of age profiles of migration are qualitatively consistent with existing administrative data sources (Eurostat migration rates). Document the recent increase in human mobility and observe that female mobility has been increasing at a faster pace. Findings suggest that e-mail data may complement existing migration data.

Table 3.A1: Migration forecasting literature

3.A.2 Tables

Table 3.A2: List of Origin Countries by Language

	English		French	Spanish	Portuguese	Arabic	Turkish	Fulah	Hausa	Persian	Papasena
Albania	Japan	Serbia	Algeria	Argentina	Angola	Algeria	Armenia	Cameroon	Niger	Afghanistan	Afghanistan
Armenia	Kazakhstan	Sierra Leone	Andorra	Aruba	Brazil	Bahrain	Azerbaijan	Gambia	Nigeria	Bahrain	Pakistan
Australia	Kenya	Singapore	Benin	Belize	Cape Verde	Chad	Georgia	Mali		Iran	
Bahamas	Kiribati	Solomon Islands	Burkina Faso	Bolivia	Equatorial Guinea	Comoros	Iran	Mauritania		Qatar	
Bangladesh	Kyrgyzstan	Somalia	Burundi	Brazil	Guinea-Bissau	Djibouti	North Macedonia	Niger		Tajikistan	
Barbados	Laos	South Africa	Cameroon	Chile	Mozambique	Egypt	Turkey	Senegal		Uzbekistan	
Belarus	Lebanon	South Korea	Canada	Colombia	Paraguay	Eritrea	Turkmenistan				
Belize	Lesotho	South Sudan	Central African Republic	Costa Rica	São Tomé & Príncipe	Iraq					
Bhutan	Liberia	Sri Lanka	Chad	Cuba	Timor-Leste	Israel					
Bosnia & Herzegovina	Malawi	St. Lucia	Comoros	Dominican Republic		Jordan					
Botswana	Malaysia	St. Vincent & Grenadines	Congo - Brazzaville	Ecuador		Kuwait					
Brunei	Maldives	Suriname	Congo - Kinshasa	El Salvador		Lebanon					
Cambodia	Mauritius	Switzerland	Côte d'Ivoire	Equatorial Guinea		Libya					
Cameroon	Mexico	Taiwan	Djibouti	Guatemala		Mauritania					
Canada	Micronesia	Tanzania	Equatorial Guinea	Honduras		Morocco					
China	Moldova	Thailand	Gabon	Mexico		Niger					
Equatorial Guinea	Mongolia	Tonga	Haiti	Morocco		Oman					
Eswatini	Montenegro	Trinidad & Tobago	Lebanon	Nicaragua		Palestinian Territories					
Ethiopia	Myanmar	Turkey	Madagascar	Panama		Qatar					
Fiji	Nepal	Uganda	Mali	Paraguay		Saudi Arabia					
Gambia	New Zealand	Ukraine	Mauritius	Peru		South Sudan					
Ghana	Nigeria	United States	Morocco	United States		Sudan					
Grenada	Norway	Vanuatu	Niger	Uruguay		Syria					
Guyana	Pakistan	Vietnam	Rwanda	Venezuela		Tanzania					
Iceland	Papua New Guinea	Zambia	Senegal			Tunisia					
India	Philippines	Zimbabwe	Seychelles			United Arab Emirates					
Indonesia	Russia		Switzerland			Yemen					
Israel	Rwanda		Togo								
Jamaica	Samoa		Tunisia								

^a For each country, the most spoken language is chosen.

Table 3.A3: Descriptive Statistics of Google Trends Data

Search Term	Mean(U)	Std.Dev.(U)	Mean(B)	Std.Dev.(B)	Share of Non-zeros (U)	Share of Non-zeros (B)
advisers+advisors	4.62	13.07	0.02	0.91	0.23	0.00
agent	13.88	19.95	0.09	1.78	0.58	0.01
aliens	6.38	13.63	0.02	0.86	0.38	0.00
applicant+applicants+application+apply	19.08	22.72	0.37	3.67	0.70	0.02
appointment	10.99	18.34	0.11	2.07	0.49	0.01
arrival+arrivals	12.62	18.61	0.09	1.88	0.53	0.00
assimilate+assimilation	4.42	12.05	0.01	0.59	0.23	0.00
asylum	5.40	12.85	0.05	1.37	0.31	0.00
asylum seeker	0.70	5.16	0.00	0.48	0.04	0.00
austerity	2.83	9.95	0.00	0.32	0.16	0.00
bailout	2.78	9.42	0.01	0.53	0.22	0.00
benefit+benefits	22.60	23.77	0.17	2.47	0.71	0.01
bilateral	7.21	15.37	0.01	0.40	0.37	0.00
biometric	3.94	11.77	0.01	0.77	0.23	0.00
births	5.98	14.70	0.02	0.77	0.29	0.00
border controls+border control	0.71	4.88	0.01	0.67	0.04	0.00
bureau of immigration	0.47	3.88	0.00	0.33	0.04	0.00
business+businesses	19.08	22.30	0.34	3.48	0.68	0.02
card	20.79	22.23	0.37	3.75	0.79	0.02
certificate	15.50	21.59	0.13	2.16	0.63	0.01
check	20.15	22.61	0.23	2.99	0.70	0.01
checkpoint+checkpoints	3.68	10.86	0.01	0.49	0.23	0.00
citizen	11.48	18.36	0.18	2.66	0.54	0.01
citizenship+citizenships	7.21	14.83	0.19	2.67	0.40	0.01
compensation+compensations	7.97	16.23	0.05	1.42	0.40	0.00
competitiveness	3.43	10.79	0.00	0.35	0.25	0.00
consulate+consulates	7.09	14.77	0.23	2.86	0.39	0.02
contract+contracts	16.82	20.79	0.11	2.05	0.67	0.01
cooperation	6.81	13.88	0.02	0.93	0.47	0.00
crises+crisis	9.89	15.79	0.14	2.20	0.57	0.01
curtail	5.49	15.02	0.00	0.44	0.21	0.00
customs	9.91	16.12	0.07	1.60	0.54	0.00
cyclical	5.80	14.12	0.01	0.67	0.28	0.00
decentralization+decentralisation	2.12	8.06	0.00	0.33	0.19	0.00
decreased	5.32	13.62	0.01	0.79	0.24	0.00
deficits	1.59	7.82	0.00	0.31	0.09	0.00
democratization+democratisation	1.04	5.92	0.00	0.17	0.09	0.00
demographic+demography	4.78	11.92	0.03	1.01	0.34	0.00
department	13.54	19.23	0.15	2.30	0.67	0.01
deportation+deportations+deported	3.01	9.93	0.02	0.91	0.17	0.00
deregulation	1.09	6.26	0.00	0.24	0.09	0.00
detain+detained+detention	5.85	13.80	0.02	0.83	0.29	0.00
determinants	5.20	13.05	0.01	0.48	0.31	0.00
devaluation	2.18	8.69	0.00	0.27	0.16	0.00
diaspora	3.91	11.11	0.01	0.49	0.27	0.00
discriminate+discriminatory	2.41	9.40	0.00	0.40	0.12	0.00
disparities	2.15	9.00	0.00	0.40	0.11	0.00
diversification	3.55	11.30	0.00	0.44	0.22	0.00
diversity	7.70	15.76	0.04	1.25	0.44	0.00
documents	12.40	19.60	0.10	1.81	0.58	0.01
downturn	3.86	11.91	0.00	0.40	0.19	0.00
dual citizenship	1.82	7.96	0.06	1.40	0.11	0.00
dual nationality	0.72	4.61	0.00	0.15	0.06	0.00
earning+earnings	8.33	16.01	0.04	1.27	0.40	0.00

Table 3.A3: Descriptive Statistics of Google Trends Data (*continued*)

Search Term	Mean(U)	Std.Dev.(U)	Mean(B)	Std.Dev.(B)	Share of Non-zeros (U)	Share of Non-zeros (B)
economically	4.10	11.48	0.02	0.80	0.28	0.00
economist+economists	5.42	12.98	0.03	0.99	0.33	0.00
economy+economies	12.62	17.91	0.26	3.05	0.67	0.02
elites	3.71	11.35	0.01	0.63	0.21	0.00
embassy+embassies	14.44	18.95	0.59	4.55	0.72	0.05
emigrant+emigrants	2.88	9.55	0.01	0.68	0.18	0.00
emigrate+emigrated	3.88	10.71	0.03	1.03	0.23	0.00
emigration	4.19	11.60	0.05	1.44	0.30	0.00
employer+employers	8.38	16.62	0.05	1.46	0.40	0.00
employment	13.26	18.92	0.12	2.06	0.65	0.01
empowerment	5.76	14.17	0.00	0.30	0.32	0.00
enforcement+enforces	6.31	15.17	0.03	1.05	0.30	0.00
exclusion	4.42	12.32	0.01	0.69	0.26	0.00
exports	5.08	12.76	0.07	1.51	0.36	0.00
extension	12.46	19.01	0.07	1.54	0.57	0.00
foreigner+foreigners	9.87	16.43	0.14	2.22	0.51	0.01
form	22.82	23.94	0.25	3.16	0.74	0.02
GDP	8.40	15.88	0.21	2.72	0.46	0.01
geopolitical	2.36	9.03	0.00	0.41	0.16	0.00
globalisation+globalization	5.52	12.89	0.03	0.99	0.41	0.00
growth	14.13	19.19	0.11	2.01	0.65	0.01
H.R.+HR	13.84	21.51	0.06	1.45	0.55	0.00
hardship+hardships	4.38	12.36	0.02	0.93	0.25	0.00
hiring	9.06	17.54	0.07	1.80	0.41	0.00
homeland	9.04	16.75	0.03	1.06	0.44	0.00
ignoring	5.53	15.00	0.00	0.45	0.22	0.00
illegal+illegally	6.86	14.83	0.09	1.86	0.38	0.00
immigrant+immigrants	6.17	13.35	0.11	1.98	0.37	0.01
immigrate+immigrated	3.33	9.93	0.06	1.47	0.23	0.00
immigration	9.96	16.25	0.32	3.36	0.62	0.03
incentives	4.10	11.89	0.02	0.76	0.25	0.00
income+incomes	14.11	19.82	0.20	2.74	0.60	0.01
indentured	3.05	10.40	0.01	0.61	0.13	0.00
indicators	6.93	14.16	0.01	0.70	0.46	0.00
individualism	2.82	10.26	0.00	0.42	0.19	0.00
industrialisation+industrialization	2.50	9.14	0.03	1.01	0.19	0.00
industrialised+industrialized	0.99	5.74	0.01	0.64	0.07	0.00
inefficiency	1.27	6.78	0.00	0.20	0.09	0.00
inequalities+inequality	5.56	13.07	0.03	1.04	0.32	0.00
inflation	8.39	15.83	0.06	1.44	0.49	0.00
influx	4.21	12.64	0.00	0.46	0.19	0.00
instability	4.27	11.93	0.01	0.52	0.25	0.00
insurance	16.10	22.20	0.24	2.94	0.64	0.01
intermarriage	0.40	3.71	0.00	0.17	0.03	0.00
internship+internships	9.91	16.87	0.07	1.61	0.47	0.00
interview	15.53	19.94	0.10	1.88	0.61	0.01
job+jobs	27.88	25.32	0.76	5.11	0.81	0.05
labor+labour+laborers+labourers	14.84	19.13	0.12	2.11	0.66	0.01
layoff+layoffs	4.50	12.70	0.01	0.44	0.20	0.00
legalization+legalisation+legalisations+legalizations	2.57	8.57	0.01	0.61	0.18	0.00
liberalization+liberalisation	2.05	7.72	0.00	0.13	0.17	0.00
lottery	12.15	19.87	0.12	2.17	0.54	0.01
macro+macroeconomic	11.61	19.37	0.03	0.98	0.50	0.00
marriage	18.31	20.50	0.13	2.08	0.71	0.01

Table 3.A3: Descriptive Statistics of Google Trends Data (*continued*)

Search Term	Mean(U)	Std.Dev.(U)	Mean(B)	Std.Dev.(B)	Share of Non-zeros (U)	Share of Non-zeros (B)
migrant+migrants	4.33	11.25	0.04	1.26	0.30	0.00
migrate	5.48	13.81	0.07	1.63	0.30	0.00
migration	9.20	16.28	0.12	2.12	0.55	0.01
minimum	11.28	18.74	0.21	2.84	0.49	0.01
mismanagement	0.72	5.11	0.00	0.00	0.05	0.00
monetary	7.01	14.24	0.03	0.97	0.46	0.00
monopolies	1.73	8.33	0.00	0.35	0.10	0.00
multicultural+multiculturalism	2.37	8.67	0.01	0.58	0.16	0.00
nationality+nationalities	10.95	17.70	0.13	2.12	0.52	0.01
nationalization+nationalisation	1.63	7.25	0.00	0.19	0.12	0.00
naturalization+naturalisation+naturalisations+naturalizations	1.93	8.09	0.01	0.59	0.14	0.00
news	27.53	22.11	0.78	5.17	0.85	0.06
passport+passports	15.34	21.65	0.31	3.30	0.56	0.02
payroll+payrolls	9.72	17.60	0.04	1.27	0.43	0.00
pension+pensions	9.14	17.93	0.11	2.01	0.40	0.01
permit	10.78	18.50	0.14	2.36	0.50	0.01
pogroms	1.33	7.59	0.00	0.15	0.06	0.00
policies	8.82	15.77	0.07	1.67	0.52	0.00
polymakers	1.65	7.51	0.01	0.62	0.10	0.00
political asylum	1.08	6.17	0.01	0.68	0.07	0.00
political refugee	0.26	3.36	0.00	0.20	0.01	0.00
populate	3.83	11.75	0.00	0.11	0.21	0.00
privatization+privatisation	2.60	9.25	0.00	0.37	0.22	0.00
productivity	5.99	13.74	0.01	0.65	0.36	0.00
prosperity	6.05	14.39	0.02	0.86	0.29	0.00
quarantine	2.54	10.60	0.08	2.25	0.22	0.00
quota+quotas	8.18	16.29	0.02	0.94	0.39	0.00
recession+recessions	2.80	9.13	0.02	0.88	0.22	0.00
recruitment+recruitments	11.56	17.63	0.06	1.48	0.55	0.00
reforms	4.45	12.26	0.04	1.30	0.27	0.00
refugee+refugees	4.12	10.66	0.05	1.25	0.32	0.00
remuneration+remunerations	6.17	14.45	0.03	1.09	0.32	0.00
renewal	10.94	19.11	0.08	1.77	0.45	0.00
repatriation	2.36	9.67	0.00	0.43	0.13	0.00
required documents+required document	3.01	11.03	0.03	1.02	0.13	0.00
requirements	16.00	23.03	0.29	3.20	0.54	0.02
resettlement	1.68	7.53	0.00	0.34	0.12	0.00
restrict+restricting	4.21	11.55	0.01	0.71	0.26	0.00
restriction	6.08	14.23	0.03	1.05	0.34	0.00
restrictive	2.34	9.56	0.00	0.29	0.13	0.00
reunification	1.76	7.83	0.01	0.71	0.11	0.00
revitalization+revitalisation	1.44	7.08	0.00	0.31	0.09	0.00
salary+salaries	18.77	21.93	0.45	4.19	0.65	0.02
sanctions	5.60	13.38	0.03	1.01	0.30	0.00
Schengen	6.23	14.27	0.09	1.79	0.35	0.01
sectors	6.04	13.99	0.02	0.98	0.34	0.00
seekers	3.89	11.87	0.01	0.50	0.26	0.00
slump	8.06	15.65	0.02	0.77	0.34	0.00
smuggler+smugglers+smuggling	5.16	12.71	0.01	0.68	0.27	0.00
social security	6.64	14.25	0.05	1.29	0.41	0.00
sponsor	6.16	14.59	0.04	1.30	0.32	0.00
spouses	4.85	12.25	0.01	0.40	0.26	0.00
stabilisation+stabilization	4.88	12.12	0.00	0.36	0.28	0.00
stagnation	2.41	9.24	0.00	0.31	0.13	0.00

Table 3.A3: Descriptive Statistics of Google Trends Data (*continued*)

Search Term	Mean(U)	Std.Dev.(U)	Mean(B)	Std.Dev.(B)	Share of Non-zeros (U)	Share of Non-zeros (B)
stateless	1.54	7.63	0.00	0.33	0.12	0.00
status	17.73	21.55	0.25	2.97	0.66	0.02
stimulus	5.24	13.68	0.01	0.87	0.28	0.00
student visa	4.01	11.95	0.11	2.02	0.22	0.01
sufficiency	3.25	10.78	0.00	0.00	0.18	0.00
tariffs	5.83	13.68	0.01	0.64	0.35	0.00
tax+taxes	14.35	19.58	0.38	3.69	0.62	0.02
test	24.98	23.03	0.29	3.26	0.81	0.02
tightened+tightening	5.42	13.62	0.00	0.41	0.26	0.00
tourist+tourists	8.50	15.37	0.21	2.63	0.48	0.02
trafficked+trafficking	7.26	15.82	0.04	1.31	0.34	0.00
unauthorised+unauthorized	2.76	9.15	0.00	0.35	0.17	0.00
underdeveloped	1.82	8.09	0.00	0.39	0.12	0.00
undocumented	1.06	6.73	0.01	0.56	0.06	0.00
unemployment	7.10	14.01	0.10	1.90	0.48	0.01
union+unions	14.41	20.45	0.23	2.90	0.64	0.01
unskilled	1.46	7.87	0.00	0.32	0.07	0.00
unsustainable	1.45	7.94	0.00	0.00	0.06	0.00
vacancy+vacancies	11.91	18.84	0.08	1.70	0.53	0.01
viability	4.64	12.40	0.00	0.31	0.27	0.00
visa free	6.30	14.25	0.04	1.24	0.33	0.00
visa+visas	17.95	22.50	0.99	5.93	0.65	0.06
wage+wages	10.95	17.41	0.24	3.02	0.50	0.01
waiver+waivers	5.09	13.06	0.02	0.90	0.25	0.00
welfare	6.93	15.35	0.04	1.19	0.35	0.00
wellbeing	6.70	14.97	0.01	0.71	0.32	0.00
woes	1.62	7.85	0.00	0.18	0.10	0.00
work visa	4.67	12.70	0.16	2.45	0.25	0.01
worker	9.64	16.05	0.07	1.57	0.50	0.00
worsening	1.48	8.11	0.00	0.11	0.05	0.00
DESTINATION			7.23	14.27		0.46

Note: Search terms 1 to 192 are migration-related single keywords and keyword combinations. "U" and "B" in brackets indicate "unilateral" and "bilateral", respectively. Unilateral indices are the relative search intensities for each migration-related search term, the descriptive statistics are calculated over all origins and all months for each term. The bilateral index of each term is the search intensity for the query combining the term and each destination, the descriptive statistics are taken over all origin-destination binaries and all months. Search term 193 is the destination country name. The search intensities of it reflect migration intentions at the bilateral level, therefore its unilateral descriptive statistic is not available. ^a U: unilateral, B: bilateral, S.D.: standard deviation. ^b Search term 'DESTINATION' represents the 27 EU destinations country names used as search terms. ^c Google Trends data ranges from 0 to 100.

Table 3.A4: List of EU Destination Countries

Country	
Austria	Ireland
Belgium	Italy
Bulgaria	Lithuania
Cyprus	Luxembourg
Czechia	Latvia
Germany	Malta
Denmark	Netherlands
Estonia	Poland
Spain	Portugal
Finland	Romania
France	Sweden
Greece	Slovenia
Croatia	Slovakia
Hungary	

Table 3.A5: Descriptive statistics for selected classical predictor variables, monthly

Variable	Count	Mean	Std.Dev.	Min	Max
Unemployment rate, %	2941	6.19	2.31	2.10	19.46
Working-age population, million	2941	53.48	69.67	0.88	261.34
No. of Natural Disaster Events (destination)	1477	0.10	0.34	0	3
No. of Natural Disaster Events	13094	0.20	0.53	0	9
No. of Technological and Complex Disaster Events (destination)	1477	0.03	0.18	0	3
No. of Technological and Complex Disaster Events	13094	0.06	0.28	0	5

Table 3.A6: Data Collected and Databases

Type of data	Frequency	Number of Variables Collected	Data source	Data source description
Political violence events and fatalities	Fully-recorded	12	ACLED	Armed Conflict Location and Event Data Project
Disaster indicators	Fully-recorded	12	EMDAT	International Disaster Database
Leadership characteristics and election outcomes	Fully-recorded	19	REIGN	Rulers, Elections, and Irregular Governance Dataset
Election violence outcome indicators	Monthly	17	ELVI	Election Violence Events Dataset
Asylum and managed migration	Monthly	2	eurostat	Statistical Office of the European Union
Short-term business statistics	Monthly	22	eurostat	Statistical Office of the European Union
Agri-environmental indicators	Monthly	348	FAO	The Food and Agriculture Organization
Labor statistics	Monthly	49	ILO	International Labour Organization
Consumer prices	Monthly	4	ILO	International Labour Organization
Macroeconomic and financial indicators	Monthly	456	IMF	International Monetary Fund

3.A.3 Appendix on Methodology

We now discuss some difficulties with our approach. First, the panel data set we use is unbalanced. That means substantial numbers of missing observations in the explanatory variables for some but not all bilateral relationships. Hence, the information set available for forecasts is not always the same for each bilateral corridor. We do not perform any interpolation to replace missing values as this may add unnecessary error to the data, especially in the case of dummy variables or count variables such as disaster and conflict data. Second, due to the moving-window forecasting approach, some explanatory variables in the training window may contain missing observations. The R routines we use to train the models typically do not allow for missing observations. In such cases, the explanatory variable must be discarded, even if it contains only a single missing observation. Third, the testing data we use to make forecasts may contain missing explanatory variables. These are neither forecast outside the model nor interpolated from the preceding data. The affected variables are also discarded from the training set in such cases. Fourth, the dependent variable for some corridors exhibits very low variation as some bilateral relationships are not used for international migration at all or to a minimal extent. That may lead to very low-quality regression results as there is no variation to be explained. In the baseline case, this issue is ignored. Fifth, the models we use are all taken “off-the-shelf”, meaning that we do not fine-tune hyperparameters, which can be detrimental to forecasting performance. On the one hand, we do this for time reasons, as there are thousands of bilateral relationships. For each of these, we need to calculate over a hundred forecasts at different forecast horizons and with varying specifications. On the other hand, experimenting with different tuning strategies has yielded minimal performance gains at increasing time costs, which do not appear to warrant tuning at each forecasting point. Finally, the dependent variable in the forecast period may be missing. This lack of ground truth makes it impossible to compute errors. So, while the forecast can be recorded, the error cannot be used for forecast performance evaluation. That is a relatively minor issue due to the overall good coverage of the dependent variable.

For the computation of Theil’s U, it has to be noted that the Random Walk can make more forecasts than many of our models because, for certain training windows, the dependent variable may contain only a constant value, usually zero. The Random Walk prediction will then mechanically be zero again, whereas a proper regression has no variation to exploit and will not record any forecast. Hence, we restrict the error computation to those instances where all of the models and the Random Walk have made a proper forecast to not distort the evaluation sample in favor of models that forecast “simpler to forecast” periods.

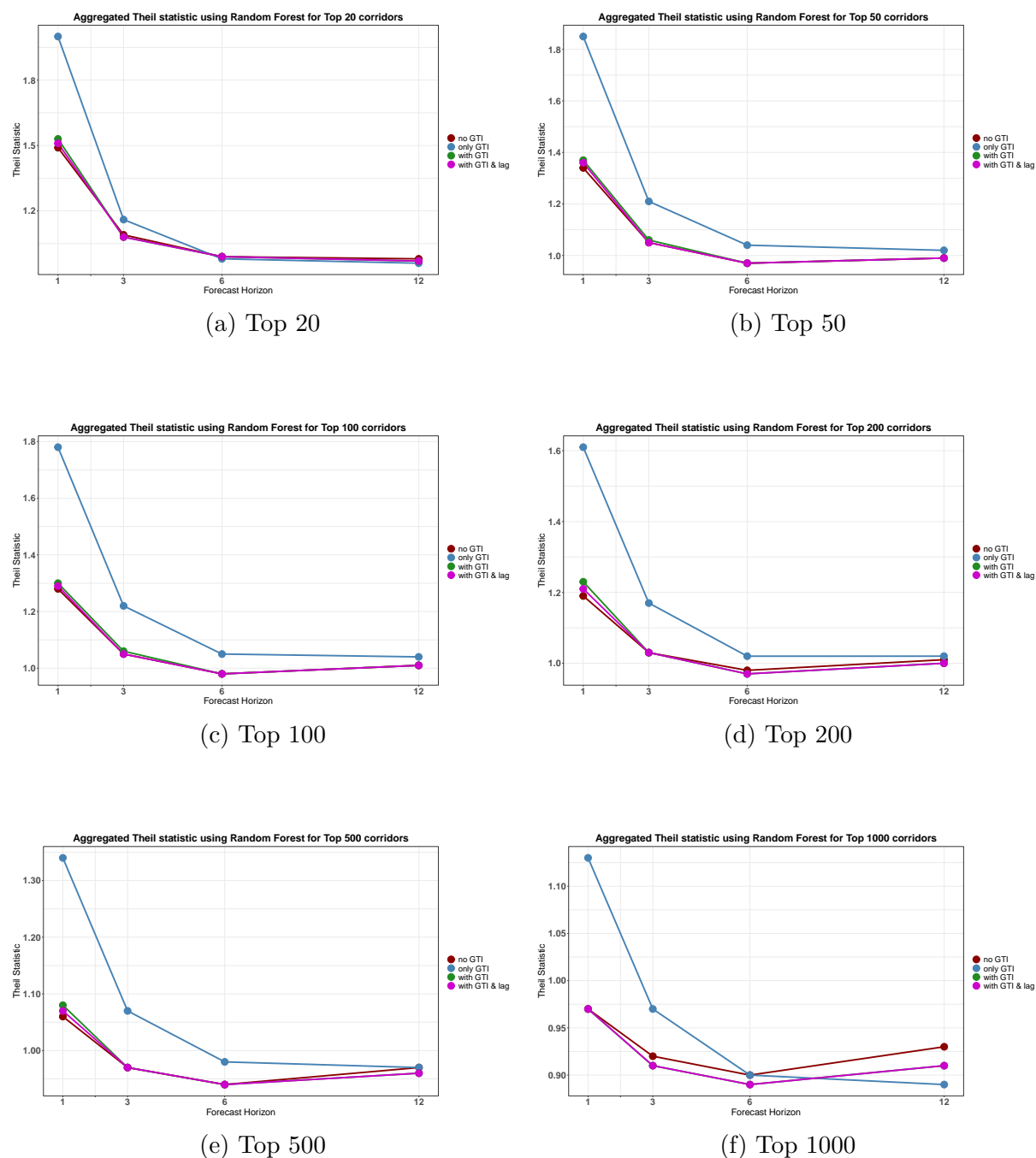
3.A.4 Additional Results

Table 3.A7: Theil statistics for specification with socioeconomic regressors, GTI and lagged asylum seekers

Horizon	EN	RF	XG	PC	EN_RF	EN_XG	XG_RF	RF_PC	EN_PC	XG_PC	EN_RF_XG	EN_RF_PC	RF_XG_PC	EN_XG_PC	EN_RF_XG_PC
A. Top 20 Corridors															
1	1.62939E+20	1.51	1.27	2.26	8.14694E+19	8.14694E+19	1.3	1.72	8.14694E+19	1.55	5.43129E+19	5.43129E+19	1.46	5.43129E+19	4.07347E+19
3	2.23858E+13	1.08	1.03	4.54	1.11929E+13	1.11929E+13	1.02	2.67	1.11929E+13	2.62	7.46195E+12	7.46195E+12	2.04	7.46195E+12	5.59646E+12
6	3.4422E+68	0.99	0.98	10059.3	1.7211E+68	1.7211E+68	0.96	5029.98	1.7211E+68	5029.96	1.1474E+68	1.1474E+68	3353.56	1.1474E+68	8.60551E+67
12	1.97939E+13	0.97	0.97	67.44	9.89697E+12	9.89697E+12	0.95	33.94	9.89697E+12	33.91	6.59798E+12	6.59798E+12	22.79	6.59798E+12	4.94848E+12
B. Top 50 Corridors															
1	1.45861E+35	1.36	1.25	10.97	7.29303E+34	7.29303E+34	1.22	6.04	7.29303E+34	5.92	4.86202E+34	4.86202E+34	4.34	4.86202E+34	3.64652E+34
3	5.44149E+42	1.05	1.07	24331.38	2.72074E+42	2.72074E+42	1.03	12166.11	2.72074E+42	12166.09	1.81383E+42	1.81383E+42	8111.03	1.81383E+42	1.36037E+42
6	1.43425E+68	0.97	1	4051.01	7.17126E+67	7.17126E+67	0.95	2025.83	7.17126E+67	2025.81	4.78084E+67	4.78084E+67	1350.79	4.78084E+67	3.58563E+67
12	1.34981E+61	0.99	1.02	8113.14	6.74906E+60	6.74906E+60	0.98	4056.81	6.74906E+60	4056.8	4.49937E+60	4.49937E+60	2704.73	4.49937E+60	3.37453E+60
C. Top 100 Corridors															
1	7.36822E+34	1.29	1.23	6.57	3.68411E+34	3.68411E+34	1.19	3.81	3.68411E+34	3.73	2.45607E+34	2.45607E+34	2.85	2.45607E+34	1.84206E+34
3	2.80996E+118	1.05	1.08	12273.24	1.40498E+118	1.40498E+118	1.03	6137.03	1.40498E+118	6137.01	9.36655E+117	9.36655E+117	4091.64	9.36655E+117	7.02491E+117
6	7.32384E+67	0.98	1	2248.08	3.66192E+67	3.66192E+67	0.96	1124.35	3.66192E+67	1124.32	2.44128E+67	2.44128E+67	749.78	2.44128E+67	1.83096E+67
12	1.41949E+110	1.01	1.02	4309.79	7.09743E+109	7.09743E+109	0.99	2155.12	7.09743E+109	2155.1	4.73162E+109	4.73162E+109	1436.92	4.73162E+109	3.54872E+109
D. Top 200 Corridors															
1	3.07435E+128	1.21	1.21	4522330347	1.53718E+128	1.53718E+128	1.14	2261165174	1.53718E+128	2261165174	1.02478E+128	1.02478E+128	1507443450	1.02478E+128	7.68588E+127
3	8.40792E+142	1.03	1.09	9.9928E+128	4.20396E+142	4.20396E+142	1.02	4.9964E+128	4.2253E+142	4.9964E+128	2.80264E+142	2.81686E+142	3.33093E+128	2.81686E+142	2.11265E+142
6	6.39105E+122	0.97	1.01	3.14692E+33	3.19552E+122	3.19552E+122	0.95	1.57346E+33	3.19552E+122	1.57346E+33	2.13035E+122	2.13035E+122	1.04897E+33	2.13035E+122	1.59776E+122
12	5.58381E+126	1	1.03	1.09711E+66	7.72258E+147	7.72258E+147	0.99	5.48556E+65	7.72258E+147	5.48556E+65	5.14839E+147	5.14839E+147	3.65704E+65	5.14839E+147	3.86129E+147
E. Top 500 Corridors															
1	5.95739E+138	1.07	1.17	3.9276E+110	2.9787E+138	2.9787E+138	1.05	1.9638E+110	2.98471E+138	1.9638E+110	1.9858E+138	1.98981E+138	1.3092E+110	1.98981E+138	1.49236E+138
3	9.68767E+145	0.97	1.08	4.00921E+128	4.84384E+145	4.84384E+145	0.97	2.0046E+128	4.8842E+145	2.0046E+128	3.22922E+145	3.25613E+145	1.3364E+128	3.25613E+145	2.4421E+145
6	8.41865E+146	0.94	1.02	1.44132E+103	4.20932E+146	4.20932E+146	0.93	7.20662E+102	4.25401E+146	7.20662E+102	2.80622E+146	2.83601E+146	4.80441E+102	2.83601E+146	1.79454E+149
12	8.18465E+147	0.96	1.03	6.22826E+135	7.2556E+147	7.2556E+147	0.95	3.11413E+135	7.33446E+147	3.11413E+135	4.83706E+147	4.88964E+147	2.07609E+135	4.88964E+147	3.66723E+147
F. Top 1000 Corridors															
1	2.98773E+138	0.97	1.08	7.19194E+142	1.49387E+138	1.49387E+138	0.96	3.59597E+142	3.62918E+142	3.59597E+142	9.9591E+137	2.41945E+142	2.39731E+142	2.41945E+142	1.81459E+142
3	4.80906E+145	0.91	1	1.8834E+149	2.40453E+145	2.40453E+145	0.9	9.41701E+148	9.63709E+148	9.41701E+148	1.60302E+145	3.3624E+149	3.28637E+149	3.3624E+149	2.5218E+149
6	4.15934E+146	0.89	0.96	1.6166E+146	2.07967E+146	2.07967E+146	0.87	8.08302E+145	2.99344E+146	8.08302E+145	1.38645E+146	1.99563E+146	5.38868E+145	1.99563E+146	9.09242E+148
12	6.2881E+147	0.91	0.96	4.74611E+146	4.70987E+147	4.70987E+147	0.88	2.37305E+146	5.35906E+147	2.37305E+146	3.13991E+147	4.02052E+149	3.74491E+149	4.02052E+149	3.01539E+149

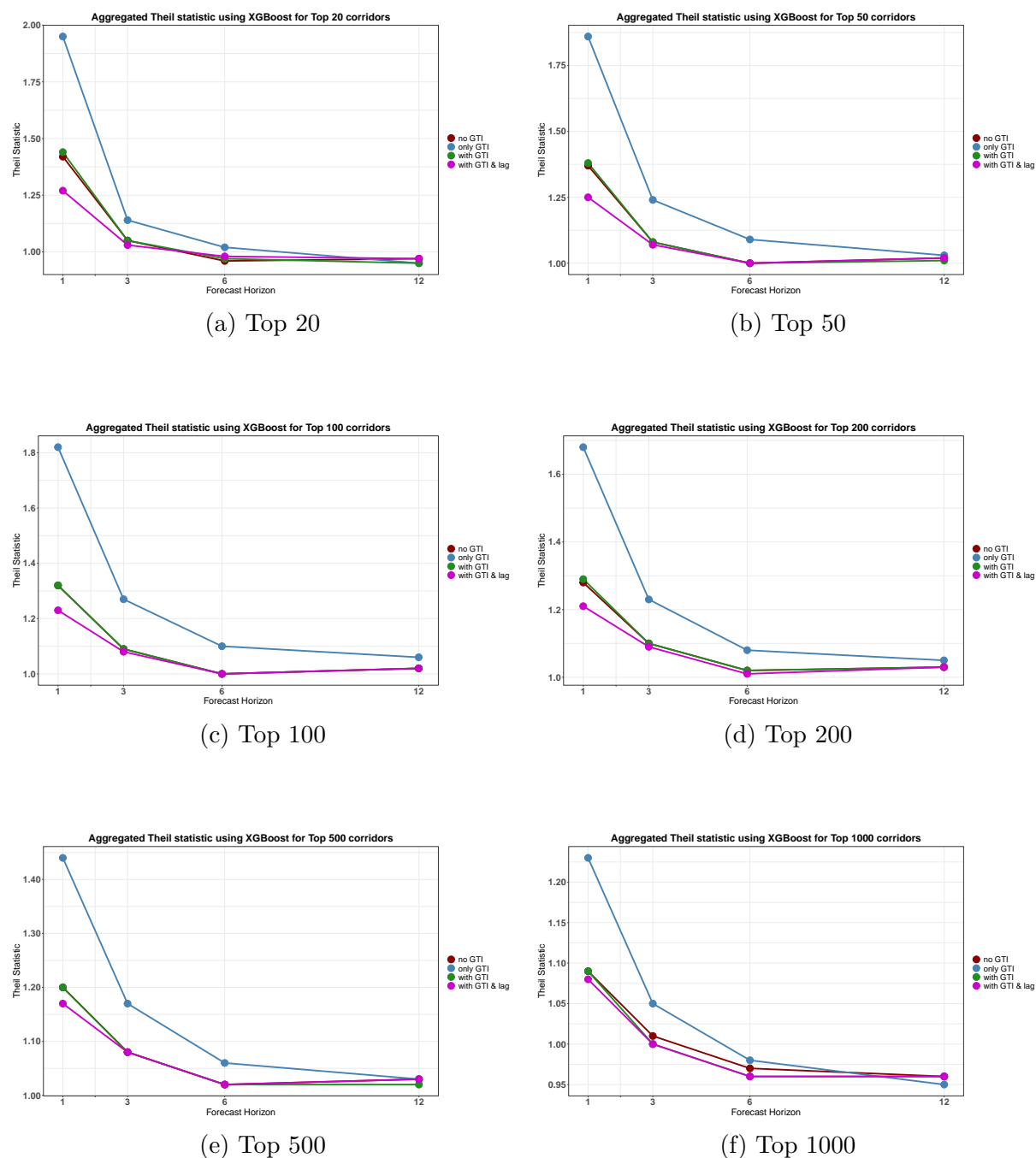
Note: Theil statistics are aggregated over all forecasts and all corridors in a given Top-X specification. For example, Top 20 refers to the twenty corridors in the sample with the largest number of asylum seekers during the sample period. A value below one indicates better performance than a random walk forecast for the same corridors over the same forecasting period. EN=Elastic Net; RF=Random Forest; XG=Extreme Gradient Boosting; PC=Principal Component. EN_RF=Ensemble model composed of Elastic Net and Random Fores.

Figure 3.A1: Main forecasting results from the Random Forest model for six subsamples of bilateral corridors of different importance for refugee flows



Note: Main forecasting results from the Random Forest model for six subsamples of bilateral corridors of different importance for refugee flows (top 20 - top 1000 corridors in terms of total asylum seeker numbers over the 2008 to 2020 period) over four horizons (1, 3, 6, and 12 months) and four specifications of the information set (no GTI, only GTI, with GTI, and with GTI and lagged dependent variable). The curves in each panel depict the Theil ratio of the RF model against the benchmark forecast based on the RW. Source: Author calculations.

Figure 3.A2: Main forecasting results from the XGBoost model for six subsamples of bilateral corridors of different importance for refugee flows



Note: Main forecasting results from the XGBoost model for six subsamples of bilateral corridors of different importance for refugee flows (top 20 - top 1000 corridors in terms of total asylum seeker numbers over the 2008 to 2020 period) over four horizons (1, 3, 6, and 12 months) and four specifications of the information set (no GTI, only GTI, with GTI, and with GTI and lagged dependent variable). The curves in each panel depict the Theil ratio of the XGBoost model against the benchmark forecast based on the RW. Source: Author calculations.

Chapter 4

Which Individual-Level Factors Explain Public Attitudes Toward Immigration? A Meta-Analysis

Lenka Dražanová, Jérôme Gonnot, Tobias Heidland & Finja Krüger

Abstract

Public attitudes toward immigration have attracted much scholarly interest and extensive empirical research in recent years. Despite a sizeable theoretical and empirical literature, no firm conclusions have been drawn regarding the factors affecting immigration opinion. We address this gap through a formal meta-analysis derived from the literature regarding immigration attitudes from the top journals of several social science disciplines in the years 2009-2019 and based on a population of 1185 estimates derived from 144 unique analyses on individual-level factors affecting attitudes to immigration. The meta-analytical findings show that two individual-level characteristics are most significantly associated with attitudes to immigration - education (positively) and age (negatively). Our results further reveal that the same individual characteristics do not necessarily explain immigration policy attitudes and attitudes toward immigrants' contribution. The findings challenge several conventional micro-level theories of attitudes to immigration. The meta-analysis can inform future research when planning the set of explanatory variables to avoid omitting key determinants.

Acknowledgment: Funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 882986 is acknowledged.

4.1 Introduction

Immigration is among voters' top concerns when asked about the main challenges for their country or other political entities such as the European Union (e.g. Commission, 2019a, 2019b). Public attitudes toward immigration are becoming part of a new political cleavage (Hobolt, 2016; Kriesi et al., 2012), particularly in the aftermath of the so-called "migration crisis". Consequently, explaining the reasons for individual differences in attitudes to immigration has attracted increased scholarly interest. Various hypotheses regarding factors affecting attitudes to immigration have been proposed, resulting in often highly correlated determinants, making it difficult to assess which of these are truly relevant. In this article, we conduct a meta-study asking which individual indicators are consistently found to influence attitudes to immigration within the broad social science literature. Meta-analyses are essential for formally structuring and summarizing the scholarly state-of-the-art on a topic. They also play a crucial role in explaining the origins of the heterogeneity of research results to academics who are non-experts in the field, policymakers, and practitioners. Our paper complements the influential review papers on attitudes toward migration (Ceobanu & Escandell, 2010; Hainmueller & Hopkins, 2014; Mayda, 2006) by providing a quantitative meta-analytical overview. Moreover, it is also worth highlighting that these reviews were done several years ago and a lot of new insights have emerged from the literature since then.

We systematize the knowledge regarding attitudes to immigration across various social science fields and cover the thirty top-ranked journals for each discipline across economics, political science, sociology, psychology, and migration/ethnic/demographic studies published between 2009 and 2019. From these, we select all 350 articles that quantitatively analyze the determinants of attitudes to immigration. After dropping all articles covering attitudes to immigration out of scope, we evaluate information from 140 academic articles and 1185 estimates in total. We thus provide an encompassing review of the research regarding attitudes to immigration published across different social science fields during the past decade.

Out of the 150 different types of attitudes to immigration that we have encountered in the literature, we focus on the two groups of dependent variables that are the most relevant and the two most commonly surveyed, which capture preferences toward migration policy (e.g., preferred levels of immigration) and views about immigrants' contribution to society. These two dependent variables complement each other as they measure two concepts - preferences regarding levels of immigration and opinions regarding the effect of immigration. We identify the relevant factors affecting attitudes to immigration based on statistically representative samples from all over the world.

Our approach focuses on eight individual-level indicators, namely age, gender, education, income, occupational and unemployment status, as well as respondents' minority

background and the type of area (urban versus rural) they live in. When selecting individual independent variables, we followed other reviews of public opinion on migration and only focused on the widely used determinants (see Ceobanu and Escandell (2010), Dražanová (2022), and Hainmueller and Hopkins (2014)) rather than understudied factors such as, for example, disgust sensitivity (Aarøe et al., 2017). After this initial pre-selection, we further reduced the number of independent variables, as some (for example subjective economic well-being, political affiliation, etc.) had too few observations for a meta-analysis. Our study's main objective is to assess recent empirical evidence on which of these individual-level factors are consistently (positively or negatively) linked with attitudes to immigration.

A key result of our meta-analysis is that two individual-level characteristics are most significantly associated with attitudes to immigration - education (positively) and age (negatively). More educated individuals are consistently found to hold significantly more positive attitudes toward immigration. Further, we find that education's effect is not only due to its connection to social class but also represents a value-based cleavage. On the other hand, age is negatively associated with negative attitudes toward immigration. Older respondents hold significantly more anti-immigration attitudes than younger respondents. Positive attitudes to immigration are also correlated with the economic standing of an individual - high-skill occupation as well as higher income lead to significantly more positive attitudes to immigration. Living in urban areas positively correlates with pro-immigration attitudes since urban residents have more contact and may also self-select into cities. Researchers should thus be careful in using and interpreting this variable in their statistical analyses due to its potential endogeneity. Furthermore, based on Bayesian model averaging (BMA), we show that if an analysis lacks certain variables, part of their effect might be wrongly attributed to other explanatory variables. For instance, we show that the effect of education, income, and age are prone to vary based on whether other individual characteristics such as minority background and gender are accounted for.

Finally, our results reveal that immigration policy attitudes and attitudes toward immigrants' contribution are not necessarily explained by the same individual characteristics. Thus, it appears that different factors should be tackled when addressing (or aiming at changing through informational campaigns) attitudes to immigration. These findings provide an original and insightful perspective on attitudes to immigration with implications for both researchers and policymakers.

The article's Theoretical Argument section presents the theoretical framework most of the literature uses to explain how and why different individual-level characteristics may affect attitudes to immigration. In the data and methodology section, we describe our research strategy and the sample of quantitative studies used in the meta-analysis. In the results section, we systematically analyze the individual factors most frequently used

in these quantitative analyses to explain attitudes toward immigration. Technical issues such as publication bias, study heterogeneity, and quality of the estimates are addressed as well. We conclude the paper in the Discussion section by summarizing the lessons learned and discussing some opportunities for further research.

4.2 Theoretical Arguments

Social scientists mostly agree that differences in attitudes to immigration are driven by two (sometimes overlapping) types of individual-level factors - economic and cultural. The theoretical framework most often used by social scientists in explaining opposition to immigration is the “competitive threat” theoretical model. According to the model, anti-immigrant sentiment should be understood as a reaction to the threat of competition (whether real or perceived) with immigrants either in the economic sphere (labor market, welfare system) or in the cultural sphere (cultural homogeneity of society, social values). Each discipline among those we investigate in this paper places more emphasis on studying one of the two groups. For instance, while economists usually focus on economic factors affecting attitudes to immigration, other disciplines, such as sociology or migration studies, predominantly study cultural factors. Within this framework, economic and cultural theories sometimes yield contradictory predictions. While economic theories suggest that natives should prefer immigrants who are different from themselves in terms of, for example, skills, cultural theories emphasize the importance of similarity.

Economic factors

Key theoretical arguments regarding economic concerns as drivers of anti-immigration attitudes extract two central motives: Firstly, the egocentric economic evaluation of consequences for individuals’ economic prospects due to migration. Secondly, the sociotropic economic evaluation of the effect of migration on the host country as a whole.

The labor market competition hypothesis suggests that the benefits and/or disadvantages of immigration are unevenly distributed within society. Natives who compete with immigrants in the labor market based on their income, employment status, and education or skill level¹ should perceive the newcomers more negatively than individuals who do not. Hence, according to the labor market competition theory, people who are in a weaker labor market situation (e.g., less education, lower wages, less job protection) are more likely to oppose low-skilled immigration because they may fear losing their jobs if immigrants can be hired at lower wages or worse working conditions (Gerber et al., 2017;

¹Empirically, skill is often approximated with formal education and the two terms are used interchangeably by many authors, although they are not the same. Trained crafts-persons may have scarcer and far better-remunerated skills than many college graduates.

Margalit, 2019; Scheve & Slaughter, 2001). At the same time, unemployed individuals may perceive that the presence of such immigrants makes their job search more difficult.

According to the labor market competition hypothesis, highly skilled and rich natives are expected to prefer low-skilled immigrants, but empirical literature documents the odds. Still, recent studies underline that the labor market threat plays a role in many contexts (Dancygier & Donnelly, 2013; Pardos-Prado & Xena, 2019).²

Another relevant economic channel for opposition to immigration is considerations about redistribution and the fiscal burden linked to the influx of immigrants. Immigration can be regarded as creating additional pressure on the national welfare system, with potential effects on the attitudes of both low- and high-skilled natives. On the one hand, native beneficiaries of a state's welfare system may fear an erosion of their welfare benefits if there is an increase in demand from immigrant recipients (Valentino et al., 2019). This concern should be more pronounced among low earners if additional benefits make adjustments necessary that result in lower transfers to natives. On the other hand, immigration may impose additional costs on tax-funded welfare systems, especially if most immigrants are low-skilled. If taxes are progressive, the brunt of these increased costs will be borne by higher-earning individuals, who could therefore develop opposing attitudes toward immigration (Gerber et al., 2017; Hainmueller & Hopkins, 2014).³

On aggregate, scholars find little evidence that personal economic circumstances influence immigration attitudes in a way that is consistent with predictions of the labor market theory (Hainmueller & Hopkins, 2014). Rather, multiple studies show that in most countries highly skilled migrants are preferred over low-skilled migrants irrespective of natives' skill levels (Gerber et al., 2017; Hainmueller & Hiscox, 2007, 2010; Hainmueller et al., 2015; Hainmueller & Hopkins, 2015; Helbling & Kriesi, 2014; Naumann et al., 2018). The authors reckon that natives prefer the arrival of high-skilled immigrants because they perceive them as contributing to the economy as a whole, i.e. as

²Dancygier and Donnelly (2013) find that individuals employed in growing sectors, where the creation of new jobs is likely, are more likely to support immigration than those employed in shrinking sectors. Pardos-Prado and Xena (2019) develop a new version of the labor market competition hypothesis that emphasizes the importance of skill transferability in theories about economic competition. The theory accounts for whether natives risk losing their jobs due to labor demand being limited in their industry and whether they would struggle to find a new job because their skills are highly industry-specific. Their results indicate that skill specificity is indeed a strong driver of negative attitudes toward immigration, regardless of natives' skill levels as measured by education.

³Most of these studies are based on the assumption that immigrants receive more welfare benefits than natives, despite a lack of evidence for this in many countries. Scholars studying the impact of immigration on the state budget often find immigrants' welfare utilization below natives' (Barrett & Maître, 2013; Huber & Oberdabernig, 2016). As a consequence, it is not entirely clear whether it is the perception of the effects of immigration on natives' welfare benefits or rather the actual situation regarding immigrants' receipt of welfare benefits driving the result (Card et al., 2012; Gang et al., 2013; Scheve & Slaughter, 2001). Other characteristics partly drive this perception. For example, more educated respondents tend to exhibit more optimism toward the economic impacts of immigration (Hainmueller & Hiscox, 2007), further strengthening the positive relationship between attitudes and education. Moreover, the extent to which immigrants depend on benefits is not only influenced by their income but is also the product of several institutional factors. These depend on entitlements and rules on eligibility for welfare benefits.

complements instead of substitutes to the existing labor force, as well as creating additional jobs. Citizens tend to exhibit a negative attitude toward immigrants who could potentially affect welfare services, either through a reduction in natives' welfare benefits or an increased tax burden if fiscal exposure to migration is high. Thus, economic self-interest does not seem to be the main driver of anti-immigration sentiments. In this regard, citizens' motivations are sociotropic rather than egocentric.

Cultural factors

Apart from the threat to economic interests, members of the majority population might feel threatened by the impact that immigrants exert on the national and cultural character of the society. Immigrants might therefore be perceived as a threat to cultural homogeneity, value system, national and cultural identity, purity of language, et cetera (Fetzer, 2000; Sniderman et al., 2004). Admitting immigrants familiar with national traditions and identity would be therefore preferred as they are perceived as more likely to culturally assimilate within the nation rather than taint the national culture and the homogeneous composition of the population (Gorodzeisky & Semyonov, 2019).

Intergroup social contact is another cultural driver of positive attitudes toward immigration, which does not originate from the "competitive threat" theoretical model and fear of competition (Pettigrew & Tropp, 2006). Intergroup contact under appropriate conditions (intimate personal experience, deeper knowledge, affective ties, and in-group reappraisal) can effectively reduce prejudice between majority and minority group members. In contrast, lack of contact is likely to preserve prejudice and negative attitudes toward out-groups (Allport, 1979). However, it is worth emphasizing that people are shown to have a general tendency to interact with others similar to them rather than with more socially distant individuals (Bogardus, 1959). Attitudes toward immigrants thus vary across groups, being more negative toward ethnic and religious minorities (Ben-Nun Bloom et al., 2015; Gorodzeisky & Semyonov, 2009; Hellwig & Sinno, 2017), such as Muslims and Roma in the European context (Gorodzeisky & Semyonov, 2019) as contacts are less frequent with this group. Nevertheless, although telling causes from consequences can be tricky, and the causal relations between contact and attitudes are not fully established, a large body of research lends firm support to the hypothesis that contact is likely to decrease negative attitudes and reduce hostility toward outsiders (Pettigrew & Tropp, 2006).

An important factor that is driven primarily by culture is the far more negative views that older people have against migrants on average. These differences are not because people become more anti-immigrant as they age. Several studies show that attitudes toward immigration are stable over adulthood (Hooghe & Wilkenfeld, 2008; Kustov et al., 2021) and remain remarkably persistent as the person grows older in a similar way to

other political predispositions acquired in youth (Neundorf et al., 2013; Vaisey & Kiley, 2021). These values and attitudes are then expected to persist through the individuals' lifetime and rarely be subjected to change (Visser & Krosnick, 1998). The correlation between age and anti-immigration views thus constitutes a cohort instead of a direct age effect. Such cohort effects are a prime example of why it can be complicated to analyze the causal mechanisms behind factors' effect on migration attitudes.

This becomes even trickier, when self-selection is involved, i.e. when people's characteristics are not exogenous (like their age) but can be adjusted based on their preferences or be shaped by circumstance. For instance, people living in urban settings are assumed to be more likely to have positive attitudes to immigration. There exist three major explanations for why residents of more urban areas often hold more positive immigration attitudes. First, this is explained by their higher exposure to immigrants, which makes them more likely to form favorable attitudes toward them through interactions with them as equals in work or personal environments (McLaren, 2003; Pettigrew & Tropp, 2006; Stolle et al., 2013). Second, urban residency can also be linked to compositional effects leading to more positive attitudes to immigration (Maxwell, 2019). This is, in part, based on self-selection. People more likely to hold positive attitudes to immigration are also more likely to self-select into large cities with their multicultural environments. Third, people in large cities are also more likely to be higher educated with higher-skill occupations and income, which, in turn, is associated with more positive attitudes to immigration. Education playing a role in cultural theories increases the expected positive correlation between education and immigration attitudes. For example, more-educated respondents often have higher self-esteem and confidence and attach higher values to cultural diversity (Hainmueller & Hiscox, 2007). The overall effect of education on attitudes is thus likely a combination of the labor market and cultural mechanisms.

Some factors cut across both theories, where the direction and predictions of their empirical relationship become ambiguous as mechanisms have different signs. Gender is a case in point. Studies generally assume men hold more anti-immigration attitudes due to their more authoritarian personalities (Adorno et al., 1950) and conservatism (Harteveld et al., 2015). However, with the recent politicization of gender in immigration debates (Farris, 2017), native women might view certain immigrants as a threat to gender equality (Ponce, 2017). Another example of theoretical associations going both ways is minority status. Despite potentially being prejudiced toward other immigrants, members of minority groups can identify more strongly with other immigrants due to their migration history and their shared outgroup status (Becker, 2019). On the other hand, members of minority groups already settled in a host country can perceive newcomers as a competitive threat. Our meta-study results partly agree with research showing that immigrants and ethnic and racial minorities are more favorable to immigration than native-born individuals (Just & Anderson, 2015).

In general, attitudes to immigration are thought to be shaped by the extent to which an individual feels threatened by different groups and ideas. These feelings of threat may be either economic or cultural or both. Many scholars (Card et al., 2012; Hainmueller & Hopkins, 2015; Margalit, 2019) argue that economic factors are consistently weaker predictors of attitudes to immigration than cultural concerns. Other studies (Mayda, 2006; Scheve & Slaughter, 2001), however, argue that economic concerns are of primary importance in explaining individual attitudes to immigration. Against this backdrop, the present meta-study may help to fill an important research gap by providing a systematic overview of the absolute and relative importance of different determinants of attitudes across studies and disciplines.

4.3 Data and Methods

4.3.1 Study selection and data generation

The selection of studies can influence the results and conclusions of a meta-analysis (Smets & van Ham, 2013), and statistically insignificant studies are less likely to be published - in both journals and preprints. Since the “file drawer” problem also exists for pre-prints, we follow researchers treating studies’ effect sizes and p-values as stemming from a censored distribution (Kasy, 2021). We then focus on the studies that can be considered “highest quality”, i.e., those published in top-ranked journals which have gone through rigorous peer review.

As the best available yet imperfect measure of journal quality, we use different rankings, namely the Journal Impact Factor (JIF) by Clarivate, the SCImago Journal Ranking (SJR), and the Google Scholar ranking, which are those most often provided by academic publishers hosting these journals (see Appendix for details). To compile a joint list of journals that can serve as the sampling frame for the selection of studies, we use the top 30 journals from the JIF ranking and the top 20 journals from the Google Scholar ranking⁴ for each discipline. We then exclude journals that are either assigned incorrectly to the field’s journal rankings (e.g., finance journals in the economics ranking) or that only publish review, methodological or theoretical papers. The resulting lists are then filled up with journals from the list of the top journals of the SJR index. We use a similar approach for the smaller disciplines (ethnic studies/migration studies/demography, see Appendix for details). The full list of journals included for each discipline can be found in Table 4.A1 in section 4.A.1 of the Appendix.

To select papers for the meta-analysis, we follow the Cochrane protocol (Higgins et al., 2019) and especially Dinesen et al. (2020) in identifying the population of studies. To identify potential articles of interest, we applied the following three criteria; The study

⁴Google Scholar only provides a top-20 ranking per discipline.

must be (1) published in one of the selected top 30 academic journals of the respective disciplines and has to be (2) in English. Since we want to provide a meta-analysis of the recent developments in the field, the study must (3) have been published between 2009 and 2019.

We then selected peer-reviewed articles using the keywords “immigrant” or “immigration” for the selected time frame (2009-2019). The identification process was carried out by two independent coders based on the additional criteria listed below. If both coders identified the same article, we included the article in our dataset. In case of a disagreement between the coders, a third coder made the decision.

When selecting analyses from articles to be included in the meta-analysis, one must bear in mind that an academic article may perform an analysis for different dependent variables. In this case, we include all these analyses separately. For further selection criteria regarding the inclusion of analyses from relevant papers, the full protocol and justification used to generate the universe of relevant studies and estimates are described in section 4.A.2 of the Appendix and the complete bibliography of the studies is provided in Table 4.A2. Figures 4.A1 - 4.A4 in the same section also present a series of visualizations of the information contained in articles regarding public opinion toward migration, which underlie our meta-analysis. The figures provide insights in terms of data, study design, and research methods used.

The dependent variables matter because they determine which analyses are included. We select analyses whose dependent variable refers to immigrants as a general category, illegal immigrants, refugees, asylum seekers, or migrants with specific ethnic, religious, and cultural backgrounds. The dependent variable must measure attitudes to immigration directly and express positive or negative opinions (as opposed to neutral statements toward immigrants or neutral policy preferences). To account for dependent variables that were created through dimension-reduction techniques, e.g., additive indices, the analysis is included only if the dependent variable includes all or a majority of attitudinal indicators directly related to immigration as defined previously.⁵ Finally, the dependent variable must be directly interpretable as a measure of attitudes, e.g., not a party vote, even if it is an anti-immigrant party.

Our initial inclusion criteria result in around 150 different measures of attitudes to immigration that are used in the literature. For simplicity, we have grouped analyses into ten higher-ordered groups of dependent variables based on the dependent variable they use. The list of those higher-ordered groups can be found in section 4.A.3 in the Appendix.

Many of the dependent variables covered above are not comparable to each other, and jointly considering them in a single analysis would not yield useful results. Instead,

⁵For example, a dependent variable that is an index composed of three questions (attitudes toward immigrants, religious minorities, and ethnic minorities would hence be excluded.

we focus on the two groups of dependent variables that are the most relevant and have the greatest number of analyses: “contribution and consequences of immigration (e.g., economic, cultural, social, political)” and “attitudes and policy preferences on immigration flows and level”. The former set of dependent variables includes mostly attitudes regarding the ex-post assessment of immigration’s impact on society and whether immigration is beneficial to the community, e.g., in terms of economy or culture. The latter set includes attitudes such as allowing more or fewer immigrants of different types into the country (labor, refugees, unskilled, certain religions). This concept engages with policy debates about levels of immigration and entry criteria, such as debates about the introduction of points systems that privilege potential migrants with higher skills. Furthermore, these two types of dependent variables complement each other, with the former covering ex-post and current assessment and the latter preferences for the future. A detailed list of attitudes included under these two categories is available in Table 4.A3 in the Appendix. Our focus on these two groups of dependent variables leaves us with a total of 144 analyses. When considering the different samples that can be included, we do not impose any demographic, geographic, ethnic, or other restrictions based on respondents’ characteristics, e.g., their religious background or minority status. Yet, we select samples based on their external validity, which we define as the extent to which attitudes of individuals in the sample are representative of a given population group and can serve as a meaningful basis for the analysis of attitudes toward immigration. Our meta-analysis thus covers both large and small-scale studies but excludes samples that lack information on representativeness. When an analysis consists of an experiment, we always use attitudes toward immigration from the pre-treatment period.

Based on previous studies (e.g. Ceobanu and Escandell (2010) and Dražanová (2022)) we initially pre-selected a sub-set of independent variables expected to affect attitudes to immigration and that appear in the literature frequently. Most independent variables offer little potential for our meta-analysis, as they occur in too few analyses to provide a large enough estimate sample. After the initial collection, we had to restrict the list of independent variables even further, as several originally included variables have not reached a critical number of estimates to be tested. The type of independent variables - continuous, categorical, binary – does not affect study inclusion as long as the corresponding coefficients can be exploited. Table 4.A4 in the Appendix lists all the independent variables included in our meta-analysis.

Based on the above, every study included in the analysis is a unique combination of the dependent variable (capturing the type of attitudes toward immigration), independent variable (capturing a specific individual factor), and sample (a given population group). For each of these studies, we choose to include a unique estimate, i.e. coefficients drawn from a single model. We do so to avoid within-study dependencies between estimates drawn from two or more models of the same study that differ only in the number of

controls included in the regression, estimation techniques, or other minor features whose inclusion would yield little additional information.⁶ As a result, each effect size in the dataset represents the estimated partial correlation between a type of attitude toward immigration and a specific individual factor for a given population group. We also exclude estimates for which neither the standard errors nor the p-value was reported and regressions with fewer than 30 degrees of freedom. Our final working dataset yields 1185 effect sizes across 110 studies. Table 4.A5 in the Appendix provides a list of all eighth individual factors covered frequently enough, along with the corresponding number of effect sizes.

4.3.2 The Meta-Analytical methodology

Because this meta-analysis aims to estimate the relationship between individual socio-economic characteristics - which cannot be subjected to experimental conditions - and attitudes toward immigration, the studies in our dataset are non-experimental. Incidentally, how individual characteristics are measured varies across studies, which use different empirical methods and therefore report different types of estimates (linear coefficients, odds ratio, probit, etc.). Moreover, because our meta-analysis contains data collected from different social sciences, each with specific guidelines for reporting estimation results, some studies present standard errors, others confidence intervals, and others simply p-values. Finally, how most individual characteristics are measured is not constant across studies. For instance, while some studies capture age through the actual age of individuals, others use a categorical variable or cohort classification. Likewise, education can be measured through years of schooling, categories of highest degree obtained, or through binary variables parting respondents between those with and without a post-secondary education.

As a result, we must transform the coefficients and the corresponding uncertainty estimates from each regression using a common metric. Following several existing meta-analyses (Cazachevici et al., 2020; Dinesen et al., 2020), we use the partial correlation coefficient (PCC) to standardize and compare these point estimates across studies. We first harmonize the direction of those estimates so that positive coefficients and t-statistics capture positive attitudes to immigration. We then derive the t-statistics and collect the sample size and the overall number of predictor variables contained in each regression. For each point estimate extracted from a regression, we then calculate the PCC as follows:

$$PCC = \frac{t}{t^2 + df}$$

⁶In the literature, this issue is sometimes addressed using Card's two-step methodology (2015) by meta-analyzing the coefficients of each study, thereby obtaining a study-specific meta-estimate called "study-pooled estimate". The overall meta-estimate is then obtained by meta-analyzing the study-pooled estimates.

where t denotes the corresponding t-statistic, and df corresponds to the number of degrees of freedom. We also calculate the standard error corresponding to the partial correlation coefficient using the following formula:

$$SE_{PCC} = \frac{PCC}{t}.$$

The partial correlation coefficient indicates the strength and direction of the relationship between a given individual characteristics and attitudes toward immigration. It captures the correlation between an individual factor and positive attitudes toward immigration, statistically adjusted for all other variables contained in the respective regression model. It can take values in the interval $[-1, 1]$, whose boundaries denote perfect negative and positive association, respectively.

To analyze the data, we use a meta-analytical multilevel random effects model. Like the fixed effects approach, the random effect model weighs the partial correlation coefficients by the inverse of their estimated variance and therefore considers the estimate's precision. Additionally, it accounts for between-study heterogeneity, as different studies will apply different methods to estimate the coefficient we use in our meta-analysis. Further details about the meta-analytical procedure can be found in the Appendix in section 4.A.5.

Moreover, in our application, the conventional assumption of meta-analytical studies that estimates are from independent samples is likely to be violated. As illustrated in Table 4.A6 in the Appendix, many estimates derive from the same or partly overlapping samples, such as the European Social Survey, the European Value Study, or the World Value Survey, creating dependencies between them. Following Dinesen et al. (2020), we address the issue of partly overlapping samples by adding random effects for the study dataset to our econometric model.

4.4 Results

4.4.1 Main results

Figure 4.1 shows the partial correlations between the individual characteristics of interest and positive attitudes to immigration. We first examine the results of the partial correlation between individual factors and both attitudes to immigration as an overarching attitude.

The strongest positive association among all individual-level factors covered in Figure 4.1 is the partial correlation between education and pro-immigration attitudes. Having a high-skill occupation⁷, as the second strongest positive association, has a substantially

⁷This captures the skill level of an individual's job.

smaller positive PCC and other economic factors such as income and unemployment have even smaller associations. This fits the expectation that the effect of education is not necessarily about its correlation with higher social class, income, or unemployment, but might also have a "cultural effect" on its own and may be understood as a value-based cleavage that adds to its role in the economic mechanisms determining attitudes.

In line with the expectation and as mentioned earlier, the PCC in Figure 4.1 shows that the higher an individual's income, the more pro-immigration their attitude. Unemployment is, on average, associated with more anti-immigration attitudes, but this effect is marginally insignificant at the 95% confidence level.

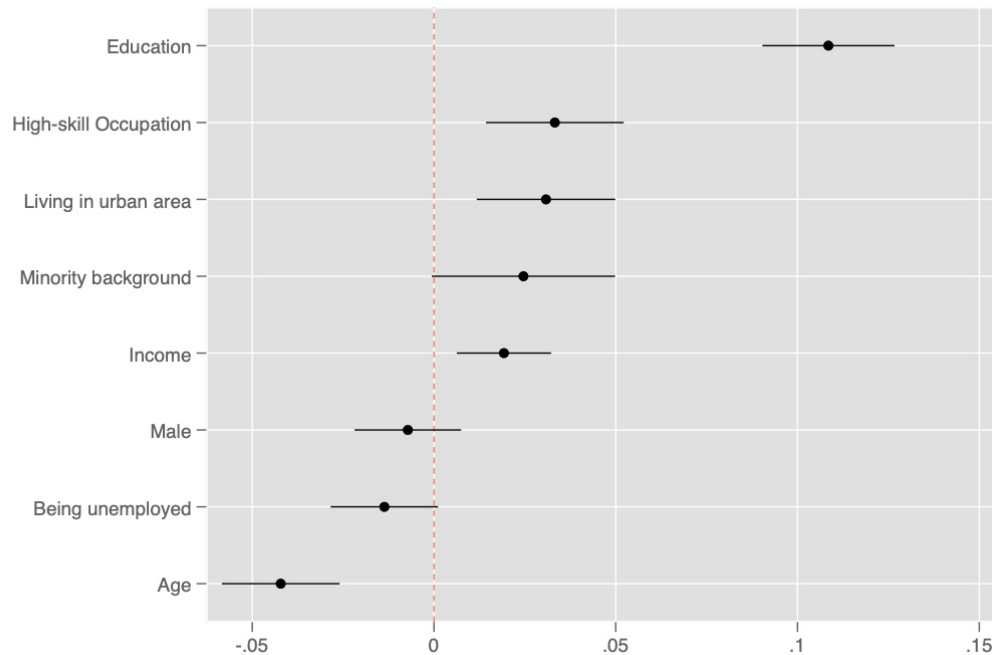
As expected, living in an urban area is positively associated with pro-immigration attitudes. Since immigrants are over-represented in urban areas, this coefficient highlights the cultural component of living in cities. Urban residents have more contact and may also self-select into cities - urban residence is thus potentially endogenous. In the extreme, there could be a direct link between anti-immigration attitudes and not wanting to move to cities - a fact that should make researchers tread carefully when using and interpreting this variable in their statistical analyses.

Having a minority background is positively associated with attitudes to immigration, but the coefficient does not significantly differ from zero at the 95 % confidence level suggesting that this is not always the case and may thus depend on the context in our PCC analysis. That fits the expected ambiguity of the relationship where cultural factors may lead to more openness and economic factors lead to less openness of minorities toward migrants. For instance, immigrants and ethnic and racial minorities may be more favorable to immigration than native-born individuals because they can identify more strongly with other immigrants due to their migration history or similar outgroup status (Becker, 2019). On the other hand, when immigrants perceive scarcity (of resources such as jobs or welfare) or competition from other immigrant groups, they may start to police national boundaries (Just & Anderson, 2015). Our data suggest that on average across studies, neither effect dominates.

Regarding gender differences, our data shows the expected empirical ambiguity. Men and women do not express significantly different attitudes to immigration in our PCC model, suggesting that gender differences are context-dependent and may furthermore depend on a study's specific framing.

Age plays a significant role across studies. Older respondents are significantly more likely to have anti-immigration views. As discussed in the theory section, the strong differences between older and younger people are not primarily because aging makes them less tolerant of migrants but due to the different contexts in which people have been socialized.

Figure 4.1: Partial correlation between individual characteristics and positive attitudes to immigration



Note: The figure visualizes the partial correlations between the individual characteristics of interest and positive attitudes to immigration.

For every individual factor, PCCs and their standard errors are plotted in Figures 4.A5 - 4.A12 in section 4.A.6 of the Appendix.

4.4.2 Heterogeneity analysis

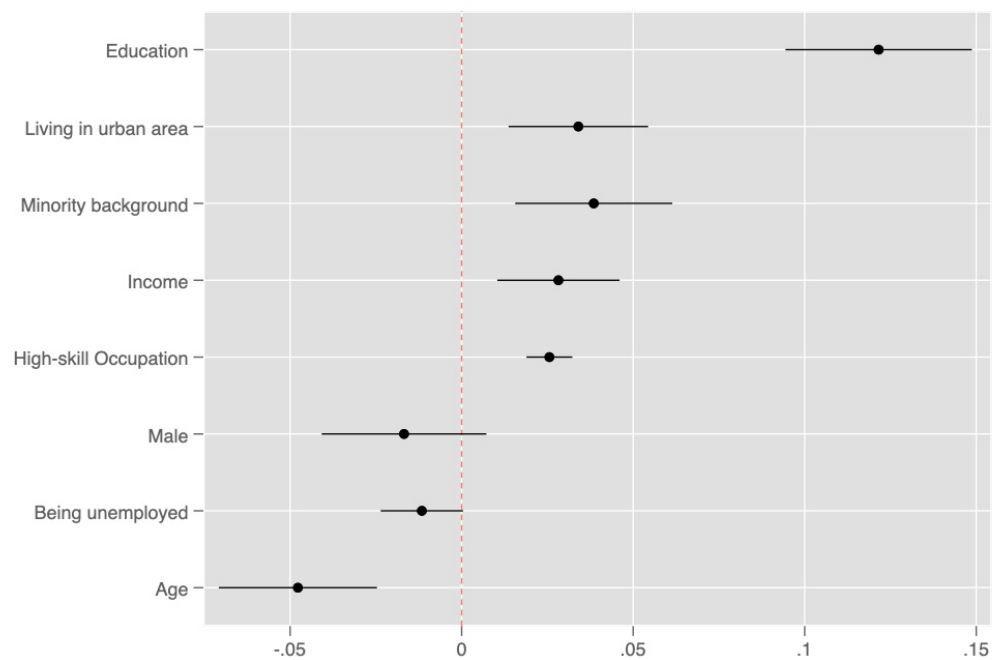
We next divide the dataset into two subsets, distinguishing between equations estimating the relationship between individual drivers and (1) public opinion toward migration policy and (2) individuals' assessment of migrants' contribution to their destination country. When looking at the results shown in Figures 4.2 and 4.3 it becomes evident why it is important to look also at the two sets of dependent variables separately. Figure 4.2 shows that several individual-level characteristics (education, living in an urban area, having a minority background, having a high-skill occupation and high-income level) are significantly positively associated with attitudes to immigrants' contribution, while age is significantly negatively associated with attitudes toward the effect of immigration. Being male and unemployed have no significant association. On the other hand, Figure 4.3 regarding support for more-/less-limiting immigration policies shows that only education (positively) and age (negatively) affect these attitudes. All other variables of interest appear not to be significantly associated with this type of attitude.

These results stress the importance of the analytical distinction between different types of attitudes to immigration in properly identifying the sources of public views

toward immigration. Therefore, a comprehensive and effective approach to address negative attitudes toward immigration necessitates careful attention to the type of attitudes in question to understand and disentangle the different factors that influence individual preferences and public opinions toward immigration.

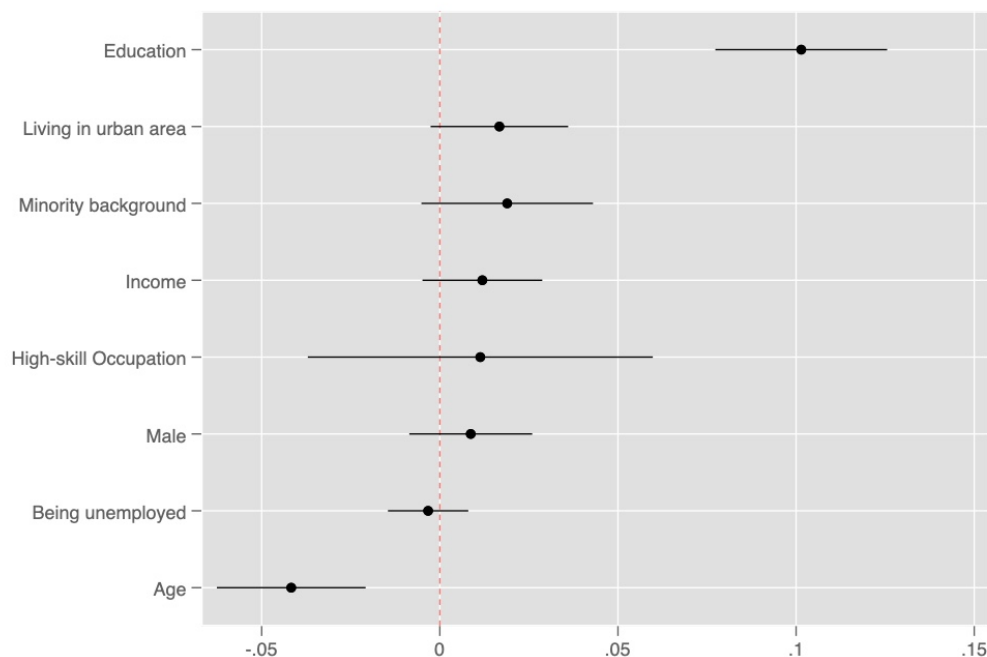
The perception that immigrants are a burden on society and challenge the status quo or, on the other hand, the unwillingness to allow an increase in arrivals of immigrants appears to be driven, apart from education and age, by different sources of opposition. Questions about "immigrants" or "immigration" are likely to be envisioned differently in the minds of different respondents.

Figure 4.2: Partial correlation between individual characteristics and positive attitudes to immigrants' contribution



Note: The figure visualizes the partial correlations between the individual characteristics of interest and positive attitudes to immigrants' contribution.

Figure 4.3: Partial correlation between individual characteristics and positive attitudes to open immigration policies



Note: The figure visualizes the partial correlations between the individual characteristics of interest and positive attitudes to open immigration policies

It is important to emphasize that the numbers reported above may be biased. First, they do not account for the fact that estimates with different signs and statistical significance may have a different probability of being reported, a problem usually referred to as publication bias.⁸ Second, although our random effects model accounts for heterogeneity across studies, it assumes it to be random, which may not be true. We, therefore, perform a publication bias test whose results can be found in Table 4.A7 in section 4.A.7 in the Appendix.

4.4.3 Study heterogeneity

As already outlined, the studies in our dataset differ in various respects, not least because they were collected from different social sciences, each with their own rules when it comes to quantitative analysis (methodology, the definition of the dependent variable, choice of empirical models, nature of population studied, etc.).

One such dimension of heterogeneity that is particularly relevant in our application is the correlation between our several variables of interest. Including specific independent variables like education and income might affect the magnitude of the coefficients of the

⁸In this respect, because our estimates are most often collected from coefficients of variables used as controls in regressions, selective reporting is less likely to be an issue considering that the publication of the results is much less dependent on the significance and direction of these coefficients.

other variables. For example, higher education often goes along with higher income. As a result, including or omitting some individual factors in a study is likely to influence the estimated effect of other individual characteristics. To assess the role of systematic heterogeneity in the inclusion of individual factors among studies on the estimated relationship between each individual factor and positive attitudes toward immigration, we fit the same multilevel weighted least square model with random effects for the study population as in Section 4.A.7 and extend the list of independent variables in the regression beyond the estimate's standard error to variables that capture features in which the studies in our dataset vary.

Besides the eight individual factors that we investigate in this paper, we also include as a precaution the choice of the dependent variable used to measure attitudes to immigration,⁹ the number of control variables and whether the data used in the study was original or not. Including these study characteristics allows extracting other potential sources of heterogeneity from the data at our disposal. It also helps to ensure that we correctly assess whether the inclusion of other correlated variables might significantly impact the estimated effect of some individual characteristics on attitudes to immigration. If we were to overlook these study characteristics, we could erroneously attribute the effect of the study's methodological features on our estimates to the inclusion of one or more correlated variables. Section 4.A.8 in the Appendix investigates the impact of study characteristics on our results and presents descriptive statistics of the explanatory variables included in the regression for each individual factor.

Only the results for the BMA of age appear relevant (Figure 4.A15 and Table 4.A10). The results for the BMA of the other variables provide little insight. Four variables have a strong or decisive effect on the estimation of age on attitudes toward immigrants.

First, without accounting for the minority status of the respondent, the effect of age on attitudes toward immigrants would be underestimated. Immigration attitudes of younger and older minority respondents differ less than majority respondents. For both younger and older minority members, socialization will ensure more positive attitudes toward immigrants. For the majority members, this is not the case. The true age difference is thus likely underestimated if a sample contains both majority and minority respondents. If the sample only contains majority respondents, this bias will not exist.

Second, including a measure for unemployment in the regression boosts the effect of age in the regression. The same holds for the urban residence variable. These effects emphasize that researchers should be cautious about self-selection patterns determining their sample composition. Both patterns could, for example, arise if older respondents are less mobile than younger respondents and therefore more likely to live in unemployment-

⁹As in section 4.4.2 we distinguish between studies investigating the relationship between individual drivers affecting public opinion toward migration policy and those regarding individuals' assessment of migrants' contribution to their destination country

prone or rural locations, respectively. Thus, these two variables would pick up part of the age differences and bias the age coefficient toward zero.

Third, without controlling for gender, the effect of age on attitudes toward immigrants becomes overestimated.¹⁰ Therefore, beyond checking for sample composition, we recommend researchers test whether adding non-linear interactions between variables affects their main estimation results. Such robustness exercises are particularly important if variables partly overlap in the concepts they measure. But as highlighted by our analysis, such interactions can potentially also bias results for explanatory variables that are less closely related.

4.5 Conclusion

Mass immigration is a global phenomenon affecting most countries. Public opinion on immigration has become a highly salient issue in many countries. Moreover, the debates surrounding migration are often highly polarized. Respondents' characteristics are found to play an important role in shaping their attitudes toward immigration. It is therefore crucial to understand what individual-level factors contribute to forming attitudes to migration and what (negative or positive) effects they have.

The scholarly literature on factors affecting attitudes to immigration often provides conflicting findings from individual studies. Therefore, to analyze the effect sizes of factors affecting attitudes to immigration, we synthesize recently published (between the years 2009-2019) individual studies from several social science disciplines (economics, political science, sociology, ethnic studies, demography, and migration studies). This quantitative summary of empirical results from studies addressing a common research question shall lead to a less biased interpretation of each factor's effect. By conducting a meta-analysis, we can provide more reliable answers to what individual-level factors are consistently associated with attitudes to immigration compared to single studies, especially if those studies have small sample sizes. Finally, it should be highlighted that a quality of a meta-analysis is determined by the quality of the studies it analyses. Based on our choice to include only studies published in top journals of each social science discipline, we tried to create only such quality samples.

Empirical results across multiple studies highlight that educational attainment is one of the key factors affecting individuals' attitudes toward immigration, with the higher-educated holding more liberal views. Studies analyzing individual-level factors affecting attitudes to immigration that cannot control the educational effect should be approached

¹⁰This could be due to several reasons, such as unbalanced samples by gender or non-linear age differences. For example, younger women might feel more threatened by immigrants than younger males, or women might be, on average, more tolerant than men. If the gender effect dissipates in older age groups, this could potentially bias the age effect.

with caution. From the results of our meta-analysis, which shows that its effect remains even after controlling for economic factors, it can be concluded that education may be understood not only as a proxy for skill level and social class but also as a source of value-based cleavage. Notably, education is also a stronger predictor of attitudes to immigration than economic variables such as income. Positive attitudes to immigration are also correlated with the economic standing of an individual – the higher his skill, occupation, and income, the more significantly positive attitudes to immigration the respondent holds. The respondent's age is also one of the most influential characteristics, with older respondents holding stronger views against immigration. Living in urban areas is positively associated with pro-immigration attitudes since urban residents have more contact and may self-select into cities. Researchers should thus be careful in using and interpreting this variable in their statistical analyses due to its potential endogeneity. In contrast, gender does not seem to be decisive in predicting attitudes to immigration, as does not being part of a minority and being unemployed.

We show that immigration policy attitudes and attitudes toward immigrants' contribution are not necessarily explained by the same individual characteristics. Even if the direction of the effect is the same, except for age and education, policy attitudes are less easily predicted by other individual characteristics used in the literature than attitudes toward immigrants' contribution.

The effect of education, income, and age are prone to vary based on whether other individual characteristics such as minority background and gender are accounted for. Our BMA analysis thus indicates that researchers should test 1) for overlap between variables and 2) whether adding non-linear interactions between variables affects their main estimation results.

The present meta-analysis can inform future research on which factors to consider in the empirical analysis of factors affecting attitudes to immigration and where to be cautious when planning the set of explanatory variables to avoid omitting key determinants. For policymakers, we provide results with greater external validity than single studies by aggregating results from various disciplines and samples across the world, therefore making our findings on the relationship between individual characteristics and attitudes to immigration more generalizable.

Each individual lives their life in a specific environment and period characterized by historical, economic, and political circumstances that may affect their attitudes toward and perception of minority groups. It is worth emphasizing that our meta-analysis does not specifically focus on the differential effects of individual-level factors affecting attitudes to immigration under different societal contexts of the country and region where individuals live. For instance, economic theories would predict that skilled individuals should favor immigration in countries where natives are more skilled than immigrants and oppose it otherwise. The literature has not extensively covered the role of macro-level

institutional and sociopolitical forces in shaping public attitudes. For instance, when testing for the effect of concerns regarding immigrants' benefits reception on individual attitudes to immigration, the extent of immigrants' benefit dependence is not only influenced by their income but also institutional factors regarding entitlements and rules on eligibility for welfare benefits. Thus, individual concerns might be stronger or weaker depending on the country where the individual resides.

Future research is needed to convert these caveats on country-level factors affecting individual factors' relationship with attitudes to immigration into firmly based explanations. Including additional evidence to better understand the sources of country-level variation in attitudes to immigration and how macro-level factors mediate individual-level relationships remains in the purview of future research.

Bibliography

- Aarøe, L., Bang Petersen, M., & Arceneaux, K. (2017). The behavioral immune system shapes political intuitions: Why and how individual differences in disgust sensitivity underlie opposition to immigration. *American Political Science Review*, *111*(2), 277–294.
- Adorno, T., Frenkel-Brenswik, E., Levinson, D., & Sanford, R. (1950). *The authoritarian personality*. Harper & Brothers.
- Allport, G. (1979). *The nature of prejudice*. Basic Books.
- Bagos, P. G. (2015). Meta-analysis in Stata using gllamm. *Research Synthesis Methods*, *6*(4), 310–332.
- Barrett, A., & Maître, B. (2013). Immigrant welfare receipt across Europe. *International Journal of Manpower*, *34*(1), 8–23.
- Becker, C. C. (2019). The influence of a migration background on attitudes towards immigration. *Social Inclusion*, *7*(4), 279–292.
- Ben-Nun Bloom, P., Arikan, G., & Lahav, G. (2015). The effect of perceived cultural and material threats on ethnic preferences in immigration attitudes. *Ethnic and Racial Studies*, *38*(10), 1760–1778.
- Bogardus, E. S. (1959). *Social distance*. University of Southern California Press.
- Card, D., Dustmann, C., & Preston, I. (2012). Immigration, wages, and compositional amenities. *Journal of the European Economic Association*, *10*(1), 78–119.
- Cazachevici, A., Havranek, T., & Horvath, R. (2020). Remittances and economic growth: A meta-analysis. *World Development*, *134*, 105021.
- Ceobanu, A. M., & Escandell, X. (2010). Comparative analyses of public attitudes toward immigrants and immigration using multinational survey data: A review of theories and research. *Annual Review of Sociology*, *36*(1), 309–328.
- Commission, E. (2019a). *Standard Eurobarometer 92 - Autumn 2019*.
- Commission, E. (2019b). *Standard Eurobarometer 95 - September 2021*.
- Dancygier, R. M., & Donnelly, M. J. (2013). Sectoral economies, economic contexts, and attitudes toward immigration. *The Journal of Politics*, *75*(1), 17–35.
- Dinesen, P. T., Schaeffer, M., & Sønderskov, K. M. (2020). Ethnic diversity and social trust: A narrative and meta-analytical review. *Annual Review of Political Science*, *23*, 441–465.

- Dražanová, L. (2022). Sometimes it is the little things: A meta-analysis of individual and contextual determinants of attitudes toward immigration (2009–2019). *International Journal of Intercultural Relations*, *87*, 85–97.
- Farris, S. R. (2017). *In the name of women's rights*. Duke University Press.
- Fetzer, J. S. (2000). *Public attitudes toward immigration in the United States, France, and Germany*. Cambridge University Press.
- Gang, I. N., Rivera-Batiz, F. L., & Yun, M.-S. (2013). Economic strain, education and attitudes towards foreigners in the European Union. *Review of International Economics*, *21*(2), 177–190.
- Gerber, A. S., Huber, G. A., Biggers, D. R., & Hendry, D. J. (2017). Self-interest, beliefs, and policy opinions: Understanding how economic beliefs affect immigration policy preferences. *Political Research Quarterly*, *70*(1), 155–171.
- Gorodzeisky, A., & Semyonov, M. (2009). Terms of exclusion: Public views towards admission and allocation of rights to immigrants in European countries. *Ethnic and Racial Studies*, *32*(3), 401–423.
- Gorodzeisky, A., & Semyonov, M. (2019). Unwelcome immigrants: Sources of opposition to different immigrant groups among Europeans. *Frontiers in Sociology*, *4*.
- Hainmueller, J., & Hiscox, M. J. (2007). Educated preferences: Explaining attitudes toward immigration in Europe. *International Organization*, *61*(2), 399–442.
- Hainmueller, J., & Hiscox, M. J. (2010). Attitudes toward highly skilled and low-skilled immigration: Evidence from a survey experiment. *American Political Science Review*, *104*(1), 61–84.
- Hainmueller, J., Hiscox, M. J., & Margalit, Y. (2015). Do concerns about labor market competition shape attitudes toward immigration? New evidence. *Journal of International Economics*, *97*(1), 193–207.
- Hainmueller, J., & Hopkins, D. J. (2014). Public attitudes toward immigration. *Annual Review of Political Science*, *17*, 225–249.
- Hainmueller, J., & Hopkins, D. J. (2015). The hidden American immigration consensus: A conjoint analysis of attitudes toward immigrants. *American Journal of Political Science*, *59*(3), 529–548.
- Harteveld, E., Van Der Brug, W., Dahlberg, S., & Kokkonen, A. (2015). The gender gap in populist radical-right voting: Examining the demand side in Western and Eastern Europe. *Patterns of Prejudice*, *49*(1-2), 103–134.
- Havranek, T., & Sokolova, A. (2020). Do consumers really follow a rule of thumb? Three thousand estimates from 144 studies say “probably not”. *Review of Economic Dynamics*, *35*, 97–122.
- Helbling, M., & Kriesi, H. (2014). Why citizens prefer high- over low-skilled immigrants. Labor market competition, welfare state, and deservingness. *European Sociological Review*, *30*(5), 595–614.

- Hellwig, T., & Sinno, A. (2017). Different groups, different threats: Public attitudes towards immigrants. *Journal of Ethnic and Migration Studies*, *43*(3), 339–358.
- Higgins, J. P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M., & Welch, V. (2019). Introduction. In *Cochrane handbook for systematic reviews of interventions*. John Wiley & Sons, Ltd.
- Hobolt, S. B. (2016). The Brexit vote: A divided nation, a divided continent. *Journal of European Public Policy*, *23*(9), 1259–1277.
- Hoeting, J. A., Madigan, D., Raftery, A. E., & Volinsky, C. T. (1999). Bayesian model averaging: A tutorial. *Statistical science*, *14*(4), 382–417.
- Hooghe, M., & Wilkenfeld, B. (2008). The stability of political attitudes and behaviors across adolescence and early adulthood: A comparison of survey data on adolescents and young adults in eight countries. *Journal of Youth and Adolescence*, *37*(2), 155–167.
- Huber, P., & Oberdabernig, D. A. (2016). Decomposing welfare wedges: An analysis of welfare dependence of immigrants and natives in Europe. *Kyklos*, *69*(1), 82–107.
- Just, A., & Anderson, C. J. (2015). Dual allegiances? Immigrants' attitudes toward immigration. *The Journal of Politics*, *77*(1), 188–201.
- Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430), 773–795.
- Kasy, M. (2021). Of forking paths and tied hands: Selective publication of findings, and what economists should do about it. *Journal of Economic Perspectives*, *35*(3), 175–92.
- Kriesi, H., Grande, E., Dolezal, M., Helbling, M., Hoeglinger, D., Hutter, S., & Wüest, B. (2012). *Political conflict in Western Europe*. Cambridge University Press.
- Kustov, A., Laaker, D., & Reller, C. (2021). The stability of immigration attitudes: Evidence and implications. *The Journal of Politics*, *83*(4), 1478–1494.
- Margalit, Y. (2019). Economic insecurity and the causes of populism, reconsidered. *Journal of Economic Perspectives*, *33*(4), 152–70.
- Maxwell, R. (2019). Cosmopolitan immigration attitudes in large European cities: Contextual or compositional effects? *American Political Science Review*, *113*(2), 456–474.
- Mayda, A. M. (2006). Who is against immigration? A cross-country investigation of individual attitudes toward immigrants. *The Review of Economics and Statistics*, *88*(3), 510–530.
- McLaren, L. M. (2003). Anti-immigrant prejudice in Europe: Contact, threat perception, and preferences for the exclusion of migrants. *Social Forces*, *81*(3), 909–936.
- Naumann, E., Stoetzer, L. F., & Pietrantuono, G. (2018). Attitudes towards highly skilled and low-skilled immigration in Europe: A survey experiment in 15 European countries. *European Journal of Political Research*, *57*(4), 1009–1030.

- Neundorff, A., Smets, K., & García-Albacete, G. M. (2013). Homemade citizens: The development of political interest during adolescence and young adulthood. *Acta politica*, *48*(1), 92–116.
- Pardos-Prado, S., & Xena, C. (2019). Skill specificity and attitudes toward immigration. *American Journal of Political Science*, *63*(2), 286–304.
- Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, *90*(5), 751.
- Ponce, A. (2017). Gender and anti-immigrant attitudes in Europe. *Socius*, *3*.
- Raftery, A. E., Madigan, D., & Hoeting, J. A. (1997). Bayesian model averaging for linear regression models. *Journal of the American Statistical Association*, *92*(437), 179–191.
- Scheve, K. F., & Slaughter, M. J. (2001). Labor market competition and individual preferences over immigration policy. *Review of Economics and Statistics*, *83*(1), 133–145.
- Smets, K., & van Ham, C. (2013). The embarrassment of riches? A meta-analysis of individual-level research on voter turnout. *Electoral Studies*, *32*(2), 344–359.
- Sniderman, P. M., Hagendoorn, L., & Prior, M. (2004). Predisposing factors and situational triggers: Exclusionary reactions to immigrant minorities. *American Political Science Review*, *98*(1), 35–49.
- Stanley, T. D., & Doucouliagos, H. (2015). Neither fixed nor random: Weighted least squares meta-analysis. *Statistics in Medicine*, *34*(13), 2116–2127.
- Stolle, D., Petermann, S., Schmid, K., Schönwälder, K., Hewstone, M., Vertovec, S., Schmitt, T., & Heywood, J. (2013). Immigration-related diversity and trust in German cities: The role of intergroup contact. *Journal of Elections, Public Opinion & Parties*, *23*(3), 279–298.
- Vaisey, S., & Kiley, K. (2021). A model-based method for detecting persistent cultural change using panel data. *Sociological Science*, *8*, 83–95.
- Valentino, N. A., Soroka, S. N., Iyengar, S., Aalberg, T., Duch, R., Fraile, M., Hahn, K. S., Hansen, K. M., Harell, A., Helbling, M., et al. (2019). Economic and cultural drivers of immigrant support worldwide. *British Journal of Political Science*, *49*(4), 1201–1226.
- Visser, P. S., & Krosnick, J. A. (1998). Development of attitude strength over the life cycle: Surge and decline. *Journal of Personality and Social Psychology*, *75*(6), 1389.
- Zeugner, S., & Feldkircher, M. (2015). Bayesian model averaging employing fixed and flexible priors: The BMS package for R. *Journal of Statistical Software*, *68*, 1–37.

Appendix

4.A.1 List of journals

Table 4.A1: List of top-ranked journals included in the meta-analysis for each discipline

Political Science
African Affairs
American Journal of Political Science
American Political Science Review
Annual Review of Political Science
British Journal of Political Science
Comparative Political Studies
Democratization
Electoral Studies
European Journal of Political Research
Governance
International Organization
JCMS: Journal of Common Market Studies
Journal of Conflict Resolution
Journal of Democracy
Journal of European Public Policy
Party Politics
Perspectives on Politics
Political Analysis
Political Behavior
Political Psychology
Political Research Quarterly
Political Studies
Public Administration
Public Opinion Quarterly
Regulation and Governance
Review of International Political Economy
Socio-economic Review
The Journal of Politics
West European Politics
World Politics

Continued on next page

Table 4.A1 – *Continued from previous page*

Sociology
American Sociological Review
Sociology of Education
Annual Review of Sociology
American Journal of Sociology
New Media and Society
Socio-Economic Review
European Sociological Review
Work and Occupations
Gender and Society
Theory and Research in Social Education
Sociological Theory
Work, Employment and Society
Social Forces
Sociological Methods and Research
Sociology
Theory, Culture and Society
International Political Sociology
Sociological Review
Social Problems
Sociologia Ruralis
British Journal of Sociology
American Journal of Cultural Sociology
British Journal of Sociology of Education
Social networks
Journal of Consumer Culture
European Journal of Social Theory
Social Science Research
Chinese Sociological Review
Journal of Marriage and Family
Sociological Forum
Psychology
Cyberpsychology Behavior and Social Networking
Personality and Social Psychology Bulletin
Journal of research in personality
Journal of experimental social psychology
Personality and Individual differences

Continued on next page

Table 4.A1 – *Continued from previous page*

Social Psychological and Personality Science
 Social and Personality Psychology Compass
 European Journal of Social Psychology
 British Journal of Social Psychology
 Group Processes & Intergroup Relations
 Psychology of Popular Media Culture
 Personality and Social psychology review
 Social Issues and Policy Review
 Journal of Personality and social psychology
 European Journal of personality
 Journal of Personality
 Journal of Social and Personal Relationships
 Social Behavior and Personality: An International Journal
 Self and Identity
 Annual Review of Organizational Psychology and Organizational Behavior
 Nature Human Behaviour
 Organizational Psychology Review
 Research in Organizational Behavior
 Journal of Research in Crime and Delinquency
 European Review of Social Psychology
 The Journal of Social Psychology
 Journal of Social and Political Psychology
 International Review of Social Psychology
 Media Psychology
 Journal of Counseling Psychology

Economy

American Economic Review
 Econometrica
 Journal of Political Economy
 Quarterly Journal of Economics
 Review of Economic Studies
 American Economic Journal: Macroeconomics
 American Economic Journal: Economic Policy
 Journal of labour Economics
 American Economic Journal: Applied Economics
 Journal of Human Resources
 Journal of Monetary Economics

Continued on next page

Table 4.A1 – *Continued from previous page*

Review of Economics and Statistics
Journal of the European Economic Association
Theoretical Economics
Journal of Economic Growth
Journal of Econometrics
Economic Journal
American Economic Journal: Microeconomics
Quantitative Economics
Journal of International Economics
Journal of Applied Econometrics
Review of Economic Dynamics
Journal of Economic Theory
Journal of Business and Economic Statistics
RAND Journal of Economics (formerly: Bell Journal of Economics)
Economic Policy (formerly: Economic Policy: A European Forum)
Journal of Public Economics
IMF Economic Review (formerly: IMF Staff Papers International Monetary Fund Staff Papers)
International Economic Review
Journal of Development Economics
Human Migration/Ethnic Studies
Journal of Ethnic and Migration Studies
Ethnic and Racial Studies
Journal of Refugee Studies
Journal of Immigrant & Refugee Studies
Identities
Ethnicities
Ethnicity & Health
Mobilities
International Migration Review
Global Networks
International Migration
Citizenship Studies
Comparative Migration Studies
Journal of International Migration and Integration
Migration Studies
IZA Journal of Migration
Journal of Intercultural Studies

Continued on next page

Table 4.A1 – *Continued from previous page*

Refugee Survey Quarterly
Migration Letters
International Journal of Refugee Law
Demography
Demography
Journal of Population Economics
Population and Development Review
Studies in Family Planning
Population, Space and Place
Perspectives on Sexual and Reproductive Health
European Journal of Population
Population and Environment
Demographic Research
Population Studies

4.A.2 Study selection approach

We rely on three further selection criteria to include analyses from relevant papers apart from those mentioned in the main text. We include (4) quantitative analyses only. Being interested in individual attitudes toward immigration, we (5) only include studies using individuals as the unit of observation. Finally, to be included, analyses (6) must measure how respondents' characteristics and circumstances affect their attitudes to immigration, i.e., not how immigrants' characteristics affect respondents' attitudes. The analyses included in the meta-analysis (7) must contain information about factors affecting the variation in individual attitudes to immigration, but their principal focus does not have to be attitudes to immigration. Table 4.A2 provides the bibliography of the selected articles according to the criteria listed.

Table 4.A2: Bibliographical list of articles included in the meta-analysis

Abrajano, Marisa, and Lydia Lundgren. 2015. "How Watershed Immigration Policies Affect American Public Opinion over a Lifetime." <i>International Migration Review</i> , 49(1): 70–105.
Aksoy, Deniz. 2012. "The Flag or the Pocketbook: To What are Immigrants a Threat: The flag or the pocketbook." <i>International Migration</i> , 50(6): 28–41.
Alarian, Hannah M., and Michael Neureiter. 2021. "Values or origin? Mandatory immigrant integration and immigration attitudes in Europe." <i>Journal of Ethnic and Migration Studies</i> , 47(5): 1006–1027.
Ariely, Gal. 2012. "Globalization, immigration and national identity: How the level of globalization affects the relations between nationalism, constructive patriotism and attitudes toward immigrants?" <i>Group Processes & Intergroup Relations</i> , 15(4): 539–557.
Asbrock, Frank, Oliver Christ, John Duckitt, and Chris G. Sibley. 2012. "Differential Effects of Intergroup Contact for Authoritarians and Social Dominators: A Dual Process Model Perspective." <i>Personality and Social Psychology Bulletin</i> , 38(4): 477–490.
Bearce, David H., and Megan Roosevelt. 2019. "A Sometimes Hidden Economic Dimension to Individual Immigration Preferences: Cross-National Evidence in Support of the Labor Competition Hypothesis." <i>Political Research Quarterly</i> , 72(4): 894–909.
Bilodeau, Antoine, and Nada Fadol. 2011. "The roots of contemporary attitudes toward immigration in Australia: contextual and individual-level influences." <i>Ethnic and Racial Studies</i> , 34(6): 1088–1109.
Blinder, Scott. 2015. "Imagined Immigration: The Impact of Different Meanings of 'Immigrants' in Public Opinion and Policy Debates in Britain." <i>Political Studies</i> , 63(1): 80–100.
Blum, Rachel Marie, and Christopher Sebastian Parker. 2019. "Trumping Foreign Affairs: Status Threat and Foreign Policy Preferences on the Right." <i>Perspectives on Politics</i> , 17(3): 737–755.

Continued on next page

Table 4.A2 – *Continued from previous page*

Bohman, Andrea, and Mikael Hjerm. 2014. “How the religious context affects the relationship between religiosity and attitudes toward immigration.” <i>Ethnic and Racial Studies</i> , 37(6): 937–957.
Böhmelt, Tobias, and Vincenzo Bove. 2020. “Regional Integration Support: A Positive Externality toward Migration Attitudes.” <i>JCMS: Journal of Common Market Studies</i> , 58(2): 309–327.
Borwein, Sophie, and Michael J. Donnelly. 2021. “Fiscal burdens and knowledge of immigrant selection criteria.” <i>Journal of Ethnic and Migration Studies</i> , 47(7): 1704–1723.
Bramlett, Brittany H., James G. Gimpel, and Frances E. Lee. 2011. “The Political Ecology of Opinion in Big-Donor Neighborhoods.” <i>Political Behavior</i> , 33(4): 565–600.
Branton, Regina, Erin C. Cassese, Bradford S. Jones, and Chad Westerland. 2011. “All Along the Watchtower: Acculturation Fear, Anti-Latino Affect, and Immigration.” <i>The Journal of Politics</i> , 73(3): 664–679.
Bruneau, Emile, Nour Kteily, and Lasse Laustsen. 2018. “The unique effects of blatant dehumanization on attitudes and behavior towards Muslim refugees during the European ‘refugee crisis’ across four countries: Blatant dehumanization and the European ‘Refugee Crisis.’” <i>European Journal of Social Psychology</i> , 48(5): 645–662.
Cameron, G. 2013. “Exodus: How migration is changing our world. By Paul Collier.” <i>Migration Studies</i> , 1(3): 375–377.
Careja, Romana, and Hans-Jürgen Andreß. 2013. “Needed but Not Liked – The Impact of Labor Market Policies on Natives’ Opinions about Immigrants.” <i>International Migration Review</i> , 47(2): 374–413.
Castanho Silva, Bruno. 2018. “The (Non)Impact of the 2015 Paris Terrorist Attacks on Political Attitudes.” <i>Personality and Social Psychology Bulletin</i> , 44(6): 838–850.
Cizmar, Anne M., Geoffrey C. Layman, John McTague, Shanna Pearson-Merkowitz, and Michael Spivey. 2014. “Authoritarianism and American Political Behavior from 1952 to 2008.” <i>Political Research Quarterly</i> , 67(1): 71–83.
Colantone, Italo, and Piero Stanig. 2018. “Global Competition and Brexit.” <i>American Political Science Review</i> , 112(2): 201–218.
Creighton, Mathew J. 2016. “Prominence and exclusivity: Identity and opposition to immigration in a multinational context.” <i>Journal of Ethnic and Migration Studies</i> , 42(3): 458–479.
Cruz, Erik, Stewart J. D’Alessio, and Lisa Stolzenberg. 2020. “Decisions made in terror: Testing the relationship between terrorism and immigration.” <i>Migration Studies</i> , 8(4): 573–588.

Continued on next page

Table 4.A2 – *Continued from previous page*

Danckert, Bolette, Peter Thisted Dinesen, Robert Klemmensen, Asbjørn Sonne Nørgaard, Dietlind Stolle, and Kim Mannemar Sønderskov. 2017. “With an Open Mind: Openness to Experience Moderates the Effect of Interethnic Encounters on Support for Immigration.” <i>European Sociological Review</i> , 33(5): 721–733.
Dancygier, Rafaela M., and Michael J. Donnelly. 2013. “Sectoral Economies, Economic Contexts, and Attitudes toward Immigration.” <i>The Journal of Politics</i> , 75(1): 17–35.
Davidov, Eldad, and Bart Meuleman. 2012. “Explaining attitudes toward immigration Policies in European Countries: The Role of Human Values.” <i>Journal of Ethnic and Migration Studies</i> , 38(5): 757–775.
DeWaard, Jack. 2015. “Beyond Group-threat: Temporal Dynamics of International Migration and Linkages to Anti-foreigner Sentiment.” <i>Journal of Ethnic and Migration Studies</i> , 41(7): 1041–1067.
Doebler, Stefanie, Ruth McAreavey, and Sally Shortall. 2018. “Is racism the new sectarianism? Negativity towards immigrants and ethnic minorities in Northern Ireland from 2004 to 2015.” <i>Ethnic and Racial Studies</i> , 41(14): 2426–2444.
Donnelly, Michael J. 2016. “Competition and solidarity: union members and immigration in Europe.” <i>West European Politics</i> , 39(4): 688–709.
Dustmann, Christian, Kristine Vasiljeva, and Anna Piil Damm. 2019. “Refugee Migration and Electoral Outcomes.” <i>The Review of Economic Studies</i> , 86(5): 2035–2091.
Feinstein, Yuval, and Bart Bonikowski. 2021. “Nationalist narratives and anti-Immigrant attitudes: exceptionalism and collective victimhood in contemporary Israel.” <i>Journal of Ethnic and Migration Studies</i> , 47(3): 741–761.
Fielding, David. 2018. “The Co-Evolution of Education and Tolerance: Evidence from England.” <i>Social Forces</i> , 96(4): 1825–1850.
Frølund Thomsen, Jens Peter, and Mark Olsen. 2017. “Re-examining Socialization Theory: How Does Democracy Influence the Impact of Education on Anti-Foreigner Sentiment?” <i>British Journal of Political Science</i> , 47(4): 915–938.
Frølund Thomsen, Jens Peter, and Arzoo Rafiqi. 2019. “Intergroup contact and its right-wing ideological constraint.” <i>Journal of Ethnic and Migration Studies</i> , 45(15): 2739–2757.
Frølund Thomsen, Jens Peter, and Arzoo Rafiqi. 2020. “Ideological Biases Weaken the Impact of Social Trust on Ethnic Outgroup Threat.” <i>Political Studies</i> , 68(2): 523–540.
Gallego, Aina, and Sergi Pardos-Prado. 2014. “The Big Five Personality Traits and Attitudes towards Immigrants.” <i>Journal of Ethnic and Migration Studies</i> , 40(1): 79–99.
Gerber, Alan S., Gregory A. Huber, Daniel R. Biggers, and David J. Hendry. 2017. “Self-Interest, Beliefs, and Policy Opinions: Understanding How Economic Beliefs Affect Immigration Policy Preferences.” <i>Political Research Quarterly</i> , 70(1): 155–171.

Continued on next page

Table 4.A2 – *Continued from previous page*

Gidron, Noam, and Peter A. Hall. 2017. “The politics of social status: Economic and cultural roots of the populist right.” <i>The British Journal of Sociology</i> , 68: S57–S84.
Gordon, Steven Lawrence. 2017. “Waiting for the Barbarians: a public opinion analysis of South African attitudes towards international migrants.” <i>Ethnic and Racial Studies</i> , 40(10): 1700–1719.
Gorodzeisky, Anastasia, and Anya Glikman. 2017. “Two Peoples – Two Stories: Anti-Immigrant Attitudes in Post-Socialist Russia.” <i>Social Problems</i> , 65(4):543–563.
Gorodzeisky, Anastasia, and Moshe Semyonov. 2018. “Competitive threat and temporal change in anti-immigrant sentiment: Insights from a hierarchical age-period-cohort model.” <i>Social Science Research</i> , 73: 31–44.
Gravelle, Timothy B. 2019. “Party identification, local context, and Australian attitudes toward immigration and asylum policy.” <i>Social Science Research</i> , 81: 77–90.
Hainmueller, Jens, and Michael J. Hiscox. 2010. “Attitudes toward Highly Skilled and Low-skilled Immigration: Evidence from a Survey Experiment.” <i>American Political Science Review</i> , 104(1): 61–84.
Hainmueller, Jens, Michael J. Hiscox, and Yotam Margalit. 2015. “Do concerns about labor market competition shape attitudes toward immigration? New evidence.” <i>Journal of International Economics</i> , 97(1): 193–207.
Hall, Jonathan, Carolin Rapp, and Terje A. Eikemo. 2019. “Does Individual Health Have Implications for Individuals’ Attitudes towards Minority Groups? A Case Study from the Greek Population.” <i>Journal of Refugee Studies</i> , 32(Special Issue 1): i238–i252.
Harell, Allison, Stuart Soroka, and Shanto Iyengar. 2017. “Locus of Control and Anti-Immigrant Sentiment in Canada, the United States, and the United Kingdom: Locus of Control and Anti-Immigrant Sentiment.” <i>Political Psychology</i> , 38(2): 245–260.
Ha, Shang E. 2010. “The Consequences of Multiracial Contexts on Public Attitudes toward Immigration.” <i>Political Research Quarterly</i> , 63(1): 29–42.
Hatemi, Peter K., and Zoltan Fazekas. 2018. “Narcissism and Political Orientations.” <i>American Journal of Political Science</i> , 62(4): 873–888.
Hatton, Timothy J. 2016. “Immigration, public opinion and the recession in Europe.” <i>Economic Policy</i> , 31(86): 205–246.
Hatton, Timothy J. 2017. “Refugees and asylum seekers, the crisis in Europe and the future of policy.” <i>Economic Policy</i> , 32(91): 447–496.
Helbling, Marc, and Richard Traummüller. 2016. “How State Support of Religion Shapes Attitudes Toward Muslim Immigrants: New Evidence From a Sub-National Comparison.” <i>Comparative Political Studies</i> , 49(3): 391–424.

Continued on next page

Table 4.A2 – *Continued from previous page*

Hellwig, Timothy, and Abdulkader Sinno. 2017. "Different groups, different threats: public attitudes towards immigrants." <i>Journal of Ethnic and Migration Studies</i> , 43(3): 339–358.
Hellwig, Timothy, and Yesola Kweon. 2016. "Taking cues on multidimensional issues: the case of attitudes toward immigration." <i>West European Politics</i> , 39(4): 710–730.
Hochman, Oshrat. 2015. "Infiltrators or Asylum Seekers? Framing and Attitudes Toward Asylum Seekers in Israel." <i>Journal of Immigrant & Refugee Studies</i> , 13(4): 358–378.
Homola, Jonathan, and Margit Tavits. 2018. "Contact Reduces Immigration-Related Fears for Leftist but Not for Rightist Voters." <i>Comparative Political Studies</i> , 51(13): 1789–1820.
Hooghe, Marc, and Ellen Quintelier. 2013. "Do All Associations Lead to Lower Levels of Ethnocentrism? A Two-Year Longitudinal Test of the Selection and Adaptation Model." <i>Political Behavior</i> , 35(2): 289–309.
Hooghe, Marc, and Thomas de Vroome. 2015. "The perception of ethnic diversity and anti-immigrant sentiments: a multilevel analysis of local communities in Belgium." <i>Ethnic and Racial Studies</i> , 38(1): 38–56.
Hopkins, Daniel J. 2011. "National Debates, Local Responses: The Origins of Local Concern about Immigration in Britain and the United States." <i>British Journal of Political Science</i> , 41(3): 499–524.
Jacobson, Robin D., and T. Elizabeth Durden. 2014. "Old poison in new security bottles: Contemporary immigration restriction and the detention regime." <i>Migration Studies</i> , 2(2): 235–254.
Janmaat, Jan G., and Avril Keating. 2019. "Are today's youth more tolerant? Trends in tolerance among young people in Britain." <i>Ethnicities</i> , 19(1): 44–65.
Jeong, Hoi Ok. 2013. "Do National Feelings Influence Public attitudes toward immigration?" <i>Journal of Ethnic and Migration Studies</i> , 39(9): 1461–1477.
Johnston, Christopher D., Benjamin J. Newman, and Yamil Velez. 2015. "Ethnic Change, Personality, and Polarization Over Immigration in the American Public." <i>Public Opinion Quarterly</i> , 79(3): 662–686.
Just, Aida, and Christopher J. Anderson. 2015. "Dual Allegiances? Immigrants' Attitudes toward Immigration." <i>The Journal of Politics</i> , 77(1): 188–201.
Karreth, Johannes, Shane P. Singh, and Szymon M. Stojek. 2015. "Explaining Attitudes toward Immigration: The Role of Regional Context and Individual Predispositions." <i>West European Politics</i> , 38(6): 1174–1202.
Konitzer, Tobias B., Shanto Iyengar, Nicholas A. Valentino, Stuart Soroka, and Raymond M. Duch. 2019. "Ethnocentrism versus groupspecific stereotyping in immigration opinion: cross-national evidence on the distinctiveness of immigrant groups." <i>Journal of Ethnic and Migration Studies</i> , 45(7): 1051–1074.

Continued on next page

Table 4.A2 – *Continued from previous page*

Kunovich, Robert M. 2013. “Labor Market Competition and Anti-Immigrant Sentiment: Occupations as Contexts.” <i>International Migration Review</i> , 47(3): 643–685.
Kunovich, Robert M. 2017. “Labour market competition and immigration attitudes in an established gateway.” <i>Ethnic and Racial Studies</i> , 40(11): 1961–1980.
Kuppens, Toon, Matthew J. Easterbrook, Russell Spears, and Antony S. R. Manstead. 2015. “Life at Both Ends of the Ladder: Education-Based Identification and Its Association With Well-Being and Social Attitudes.” <i>Personality and Social Psychology Bulletin</i> , 41(9): 1260–1275.
Kustov, Alexander. 2019. “Is There a Backlash Against Immigration From Richer Countries? International Hierarchy and the Limits of Group Threat.” <i>Political Psychology</i> , 40(5): 973–1000.
Langsæther, Peter Egge, and Rune Stubager. 2019. “Old wine in new bottles? Reassessing the effects of globalisation on political preferences in Western Europe.” <i>European Journal of Political Research</i> , 58(4): 1213–1233.
Laurence, James, and Lee Bentley. 2018. “Countervailing contact: Community ethnic diversity, anti-immigrant attitudes and mediating pathways of positive and negative inter-ethnic contact in European societies.” <i>Social Science Research</i> , 69: 83–110.
Lee, Naeyun, and Cheol-Sung Lee. 2015. “Relational skill assets and antiimmigrant sentiments.” <i>Social Science Research</i> , 52: 270–289.
Malhotra, Neil, Yotam Margalit, and Cecilia Hyunjung Mo. 2013. “Economic Explanations for Opposition to Immigration: Distinguishing between Prevalence and Conditional Impact.” <i>American Journal of Political Science</i> , 57(2): 391–410.
Marfouk, Abdeslam. 2019. “I’m neither racist nor xenophobic, but: dissecting European attitudes towards a ban on Muslims’ immigration.” <i>Ethnic and Racial Studies</i> , 42(10): 1747–1765.
Markaki, Yvonne, and Simonetta Longhi. 2013. “What determines attitudes to immigration in European countries? An analysis at the regional level.” <i>Migration Studies</i> , 1(3): 311–337.
Masso, Anu. 2009. “A Readiness to Accept Immigrants in Europe? Individual and Country-Level Characteristics.” <i>Journal of Ethnic and Migration Studies</i> , 35(2): 251–270.
Meeusen, Cecil, and Anna Kern. 2016. “The relation between societal factors and different forms of prejudice: A cross-national approach on targetspecific and generalized prejudice.” <i>Social Science Research</i> , 55: 1–15.
Meidert, Nadine, and Carolin Rapp. 2019. “Public Attitudes towards Refugees in Germany: What Drives Attitudes towards Refugees in Comparison with Immigrant Workers from European Union Countries?” <i>Journal of Refugee Studies</i> , 32(Special Issue 1): i209–i218.

Continued on next page

Table 4.A2 – *Continued from previous page*

Meseguer, Covadonga, and Achim Kemmerling. 2018. “What Do You Fear? Anti-immigrant Sentiment in Latin America.” <i>International Migration Review</i> , 52(1): 236–272.
Mitchell, Jeffrey. 2019. “Prejudice in the classroom: a longitudinal analysis of anti-immigrant attitudes.” <i>Ethnic and Racial Studies</i> , 42(9): 1514–1533.
Mora, G. Cristina, and Tianna S. Paschel. 2020. “Antiblackness as a Logic for Anti-Immigrant Resentment: Evidence From California.” <i>Sociological Forum</i> , 35(S1): 918–940.
Mustafa, Asma, and Lindsay Richards. 2019. “Immigration attitudes amongst European Muslims: social identity, economic threat and familiar experiences.” <i>Ethnic and Racial Studies</i> , 42(7): 1050–1069.
Newman, Benjamin J. 2013. “Acculturating Contexts and Anglo Opposition to Immigration in the United States.” <i>American Journal of Political Science</i> , 57(2): 374–390.
Newman, Benjamin J., Todd K. Hartman, and Charles S. Taber. 2012. “Foreign Language Exposure, Cultural Threat, and Opposition to Immigration: Language, Threat, and Opposition to Immigration.” <i>Political Psychology</i> , 33(5): 635–657.
Newman, Benjamin J., Todd K. Hartman, Patrick L. Lown, and Stanley Feldman. 2015. “Easing the Heavy Hand: Humanitarian Concern, Empathy, and Opinion on Immigration.” <i>British Journal of Political Science</i> , 45(3): 583–607.
Panichella, Nazareno, and Maurizio Ambrosini. 2018. “Between Fears, Contacts and Family Dynamics: the Anti-Immigrant Attitudes in Italy.” <i>Journal of International Migration and Integration</i> , 19(2): 391–411.
Pardos-Prado, Sergi. 2011. “Framing Attitudes Towards Immigrants in Europe: When Competition Does Not Matter.” <i>Journal of Ethnic and Migration Studies</i> , 37(7): 999–1015.
Pardos-Prado, Sergi, and Carla Xena. 2019. “Skill Specificity and Attitudes toward Immigration.” <i>American Journal of Political Science</i> , 63(2): 286–304.
Percival, Garrick L., and Mary Currin-Percival. 2013. “Exploring the Contextual Determinants of Individual Attitudes toward Immigrants and Criminal Activity and their Spillover Policy Implications: Attitudes toward immigrants.” <i>International Migration</i> , 51(6): 1–23.
Rad, Mostafa Salari, and Jeremy Ginges. 2018. “Folk theories of nationality and anti-immigrant attitudes.” <i>Nature Human Behaviour</i> , 2(5): 343–347.
Raijman, Rebeca. 2010. “Citizenship Status, Ethno-National Origin and Entitlement to Rights: Majority Attitudes towards Minorities and Immigrants in Israel.” <i>Journal of Ethnic and Migration Studies</i> , 36(1): 87–106.
Ruedin, Didier. 2019. “Attitudes to immigrants in South Africa: personality and vulnerability.” <i>Journal of Ethnic and Migration Studies</i> , 45(7): 1108–1126.
Ruist, Joakim. 2016. “How the macroeconomic context impacts on attitudes to immigration: Evidence from within-country variation.” <i>Social Science Research</i> , 60: 125–134.

Continued on next page

Table 4.A2 – *Continued from previous page*

Sarrasin, Oriane, Eva G. T. Green, Claudio Bolzman, Emilio Paolo Visintin, and Emanuele Politi. 2018. “Competition- and Identity-based Roots of Anti-immigration Prejudice among Individuals with and without an Immigrant Background.” <i>International Review of Social Psychology</i> , 31(1): 12.
Schlueter, Elmar, and Eldad Davidov. 2013. “Contextual Sources of Perceived Group Threat: Negative Immigration-Related News Reports, Immigrant Group Size and their Interaction, Spain 1996–2007.” <i>European Sociological Review</i> , 29(2): 179–191.
Schlueter, Elmar, Bart Meuleman, and Eldad Davidov. 2013. “Immigrant Integration policies and perceived Group Threat: A Multilevel Study of 27 Western and Eastern European Countries.” <i>Social Science Research</i> , 42(3): 670–682.
Schmuck, Desirée, and Jörg Matthes. 2015. “How Anti-immigrant Right-wing Populist Advertisements Affect Young Voters: Symbolic Threats, Economic Threats and the Moderating Role of Education.” <i>Journal of Ethnic and Migration Studies</i> , 41(10): 1577–1599.
Schotte, Simone, and Hernan Winkler. 2018. “Why Are the Elderly More Averse to Immigration When They Are More Likely to Benefit? Evidence across Countries.” <i>International Migration Review</i> , 52(4): 1250–1282.
Shrira, Ilan. 2020. “Population Diversity and Ancestral Diversity As Distinct Contributors to Outgroup Prejudice.” <i>Personality and Social Psychology Bulletin</i> , 46(6): 885–895.
Sibley, Chris G., John Duckitt, Robin Bergh, Danny Osborne, Ryan Perry, Frank Asbrock, Andrew Robertson, Gavin Armstrong, Marc Stewart Wilson, and Fiona Kate Barlow. 2013. “A Dual Process Model of attitudes toward immigration: Person × Residential Area Effects in a National Sample: A Dual Process Model of attitudes toward immigration.” <i>Political Psychology</i> , 34(4): 553–572.
Smiley, Kevin T., Michael Oluf Emerson, and Julie Werner Markussen. 2017. “Immigration Attitudes Before and After Tragedy in Copenhagen: The Importance of Political Affiliation and Safety Concerns.” <i>Sociological Forum</i> , 32(2): 321–338.
Solheim, Øyvind Bugge. 2021. “Are we all Charlie? How media priming and framing affect immigration policy preferences after terrorist attacks.” <i>West European Politics</i> , 44(2): 204–228.
Strabac, Zan, Toril Aalberg, and Marko Valenta. 2014. “Attitudes towards Muslim Immigrants: Evidence from Survey Experiments across Four Countries.” <i>Journal of Ethnic and Migration Studies</i> , 40(1): 100–118.
Valentova, Marie, and Marie-Sophie Callens. 2018. “Did the escalation of the financial crisis of 2008 affect the perception of immigration-related threats? A natural experiment.” <i>Journal of Ethnic and Migration Studies</i> , 44(3): 439–461.

Continued on next page

Table 4.A2 – *Continued from previous page*

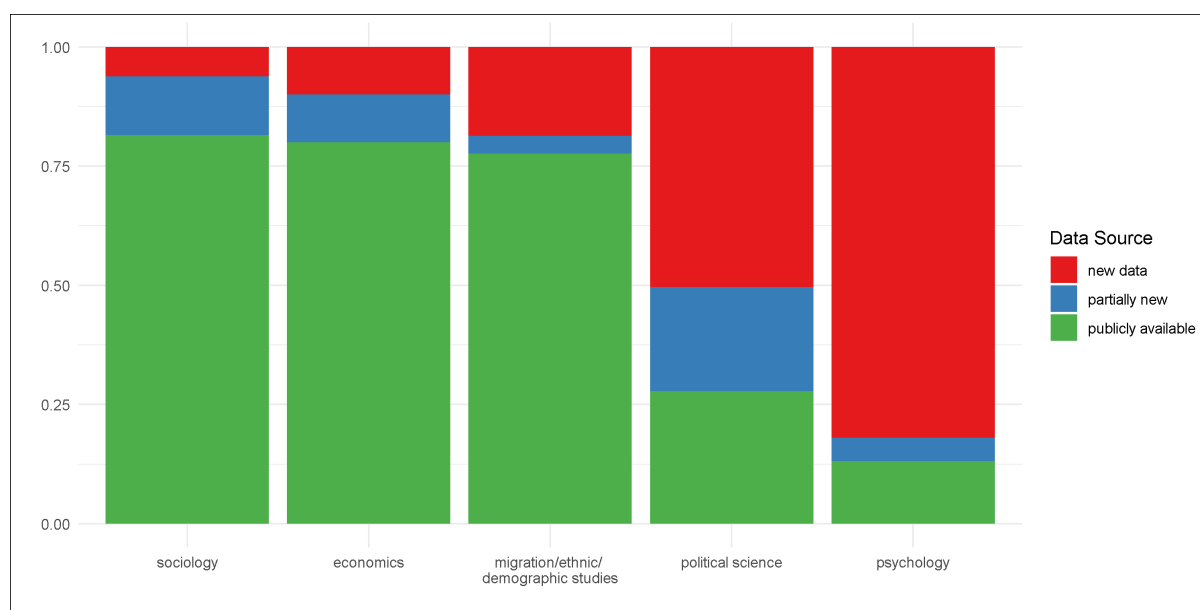
van der Zwan, Roos, Per Bles, and Marcel Lubbers. 2017. “Perceived Migrant Threat among Migrants in Europe.” <i>European Sociological Review</i> , 33(4): 518–533.
van Prooijen, Jan-Willem, André P. M. Krouwel, and Julia Emmer. 2018. “Ideological Responses to the EU Refugee Crisis: The Left, the Right, and the Extremes.” <i>Social Psychological and Personality Science</i> , 9(2): 143–150.
van Prooijen, Jan-Willem, André P. M. Krouwel, Max Boiten, and Lennart Eendebak. 2015. “Fear Among the Extremes: How Political Ideology Predicts Negative Emotions and Outgroup Derogation.” <i>Personality and Social Psychology Bulletin</i> , 41(4): 485–497.
Wallrich, Lukas, Keon West, and Adam Rutland. 2020. “Painting all foreigners with one brush? How the salience of Muslims and refugees shapes judgements.” <i>Journal of Social and Political Psychology</i> , 8(1): 246–265.
Whitaker, Beth Elise, and Jason Giersch. 2015. “Political Competition and attitudes toward immigration in Africa.” <i>Journal of Ethnic and Migration Studies</i> , 41(10): 1536–1557.
Wohl, Michael J. A., Anna Stefaniak, and Anouk Smeekes. 2020. “Longing is in the memory of the beholder: Collective nostalgia content determines the method members will support to make their group great again.” <i>Journal of Experimental Social Psychology</i> , 91: 104044.
Zahl-Thanem, Alexander, and Marit S. Haugen. 2019. “Attitudes Toward Immigrants in Rural Norway. A Rural-Urban Comparison.” <i>Sociologia Ruralis</i> , 59(4): 685–700.

The quantitative study of attitudes to immigration relies on public opinion surveys. Figure 4.A1 shows each discipline’s data source usage in attitudes to immigration research. We classified data sources as either commonly and/or publicly available datasets (such as the European Social Survey, World Value Survey, and the like), data with partial data development, and as original data developed by the respective author(s). We counted the author(s)’ efforts as partial data development if they modified or extended, in either substantive, temporal, or geographic terms, at least one variable contained in a publicly available dataset (but not if they only recoded or aggregated existing measures). New data means introducing entirely original measures, such as a completely new dataset or new and original data generated for at least one variable. As is apparent from Figure 4.A1 while sociology, economics, and migration/ethnic/demographic studies rely mostly on publicly and/or commonly available datasets in their study of public attitudes to immigration, political scientists and especially psychologists mostly rely on original data developed by the respective authors.

While political scientists and psychologists invest in developing original data, sociologists, economists, and migration/ethnic/demography scholars appear to accept the constraints imposed by data availability. On the other hand, original data development leads to fewer opportunities for the reproducibility of empirical results. It is worth mentioning that the replication crisis is highly salient, particularly in psychology. Nevertheless, our

intent here is diagnostic rather than normative. We wish to empirically establish some structural features of research conducted on attitudes to immigration across disciplines rather than normatively evaluating data usage. Evaluating whether original or common/publicly available (and thus, somehow limited in country coverage based on their availability) data usage across disciplines leads to more or fewer novel findings would require a prior examination of their quality. Such an evaluative assessment lies beyond the scope of this article.

Figure 4.A1: Data source according to discipline

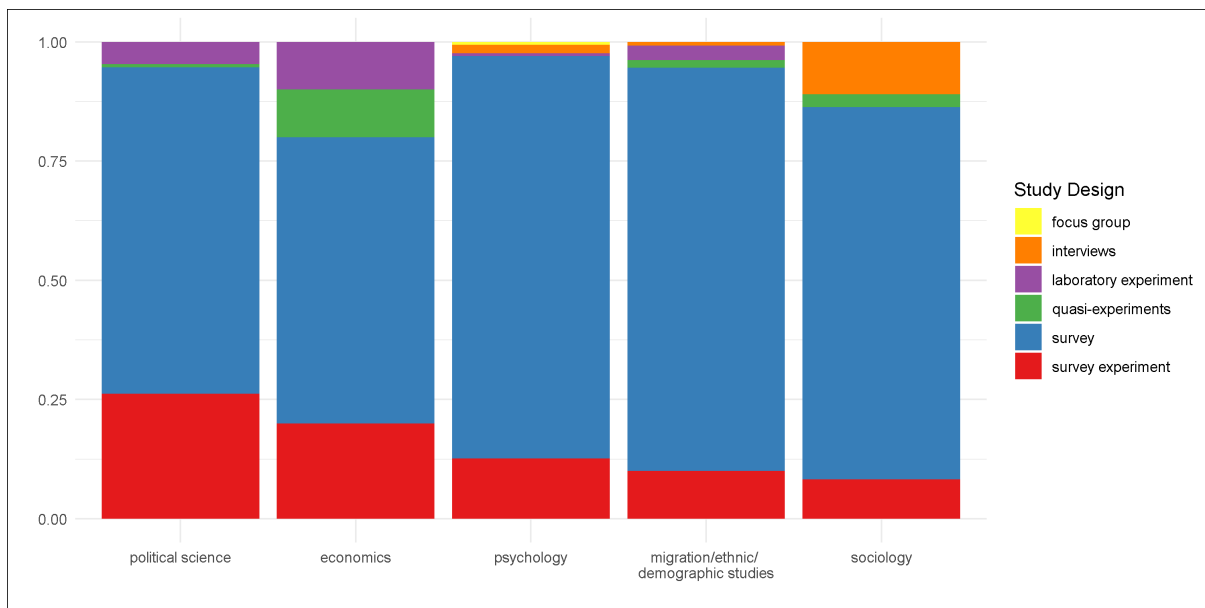


We also looked at what method was used to generate the new data.¹¹ Figure 4.A2 shows the distribution of research designs across disciplines. It shows that in all disciplines, data generation mostly relies on surveys. While survey experiments also have a large share as data sources in political science (above one-fourth), other fields use them less. Economics, for example, appears to rely frequently on laboratory experiments as well as quasi-experiments when studying attitudes to immigration. Interviews are a widespread data collection method in sociology. All research designs seem to require similar levels of measurement innovation.¹²

¹¹If more than one method is used, we selected multiple values.

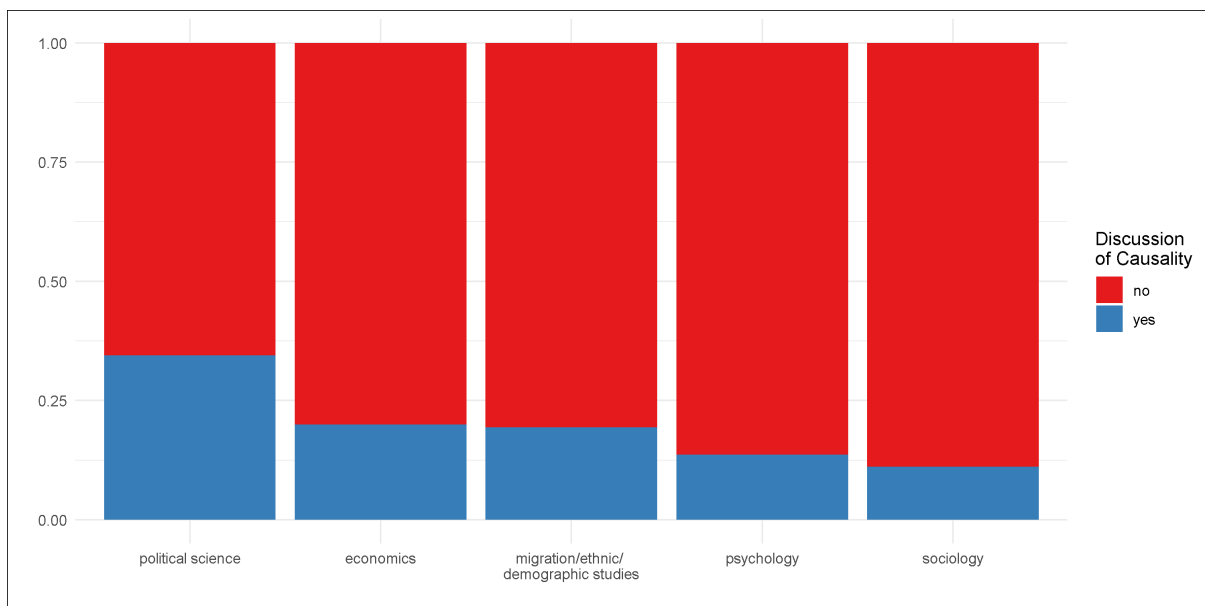
¹²Figure 4.A3 in the Appendix shows when causality in relation to attitudes to immigration is explicitly discussed in our sample studies across disciplines. Generally, the explicit discussion of causality is rather seldom across all disciplines. Future research on factors affecting attitudes to immigration shall discuss causality explicitly.

Figure 4.A2: Study design according to discipline



Note: The study design refers to how original data was generated.

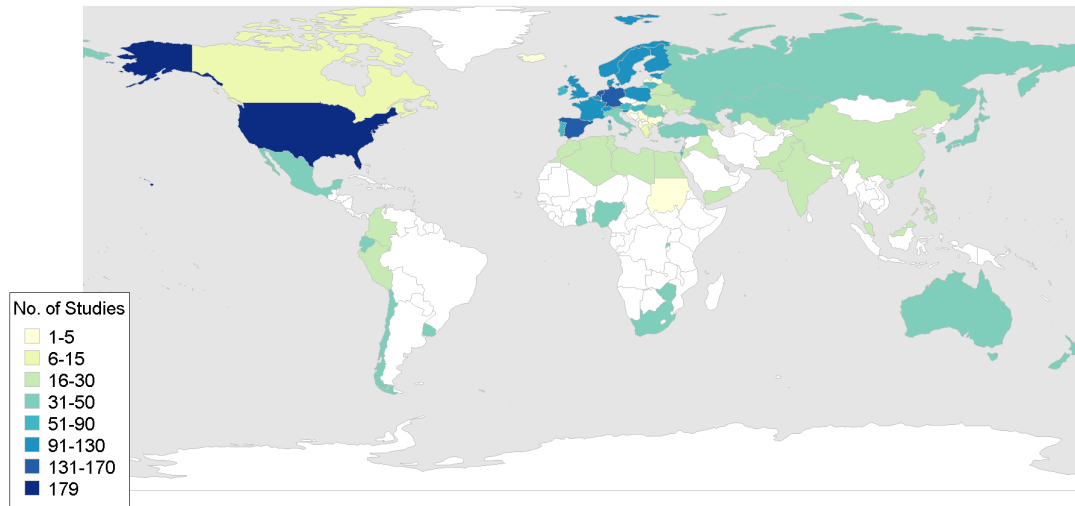
Figure 4.A3: Discussion of causality according to discipline



The map in Figure 4.A4 displays the number of studies that study specific countries as a heatmap overlay. We can read this as a rough measurement that answers the question about which countries the scholars in our dataset write about the most. It is immediately noticeable that the United States is the country which is mentioned by far the most, followed by European countries such as Spain, Germany, and the Netherlands. Figure 4.A4 shows that our dataset's sample of countries studied is skewed toward the countries in the Global North, particularly the Western hemisphere. This might lead to gaps in the knowledge of whether theories confirmed in the West work the same way also in countries

with other political, cultural, and socio-demographic backgrounds.

Figure 4.A4: Map of countries studied



Note: If countries worldwide were studied, we included the list of countries from the World Value Survey 2010-2014 wave. If countries Europe-wide were studied, we included the list of countries from the European Social Survey 2012-2014 wave.

4.A.3 List of measures of attitudes to immigration

Our initial inclusion criteria result in around 150 different measures of attitudes to immigration that are used in the literature. For simplicity, we have grouped analyses into ten higher-ordered groups of dependent variables based on the dependent variable they use. Ten higher-ordered groups of dependent variables based on the dependent variable they use are;

- Attitudes and policy preferences on integration issues (rights and opportunities)
- Attitudes and policy preferences on cultural issues
- Concerns and feelings toward immigrants
- Contribution and consequences of immigration (e.g., economic, cultural, social, political)
- Attitudes and policy preferences on refugees/asylum seekers management issues (e.g., border management, support, management of flows)
- Attitudes and policy preferences on immigration flows and level
- Individual behavior toward immigrants (e.g., financial support, social distance, assisting in arresting immigrants)
- A mix of attitudes to immigrants (indistinct)
- Prejudice and trust toward immigrants

Table 4.A3: List of attitudes included in the meta-analysis

Contribution and consequences of immigration
Benefits of immigrants for the country
Benefits of immigrants from Ethiopia for country
Benefits of immigrants from FSU (former Soviet Union) for country
Benefits of immigrants from Western countries for country
Economic competition with immigrants
Benefits of immigrants for the country
Benefits of refugees for the country
Immigrants steal jobs
Immigrants are a strain on the welfare system
Immigrants are a problem for the community

Continued on next page

Table 4.A3 – *Continued from previous page*

Immigrants enrich culture
Immigrants are good for the economy
Immigrants make the country a better place
Immigrants are a problem for security
Immigrants are a problem for security in the community
Immigrants are a problem for security/Afraid of immigrants
Mix of attitudes to benefits of Muslim immigrants for the country
Mix of attitudes to benefits of high-skill immigration for the country
Mix of attitudes to benefits of immigrants for the country
Mix of attitudes to benefits of low-skill immigration
Mix of attitudes to benefits of refugees/asylum seekers
Threat from asylum seekers
Threat from immigrants from Ethiopia
Threat from immigrants from FSU (former Soviet Union)
Threat from immigrants from Western countries
Threat from immigrants to the economic welfare of the country
Threat from immigrants to the economic welfare of the household
Threat from immigrants to the economic welfare system of the country
Threat from immigrants to national identity
Refugees' contribution to the country
Refugees are a problem for security
Attitudes and policy preferences on immigration flows and level
Allow refugees to bring family
Ban on Muslim immigrants
Government judgement of refugee applications
Government should accept refugees
Mix of attitudes to immigration policy
Mix of attitudes to refugees/asylum seekers policy
Mix of attitudes to selective admission of immigrants
More/less Arab immigrants
More/less asylum seekers
More/less immigrants from a different ethnic group
More/less immigrants from poor countries
More/less immigrants from poor European countries
More/less immigrants from the same ethnic group
More/less immigrants
More/less Jewish immigrants from poor countries

Continued on next page

Table 4.A3 – *Continued from previous page*

More/less labor immigrants
More/less Muslim immigrants
More/less Muslim immigrants from poor countries
More/less refugees
More/less refugees/Mix of attitudes to refugees
More/less refugees from countries with terrorists
More/less Roma immigrants from poor countries
More/less same ethnic group
More/less skilled immigrants
More/less unskilled immigrants

4.A.4 List of independent variables

Table 4.A4: Independent variables used in the meta-analysis and effect sizes

Independent variables	Definition
Age	Age measured in years or in decades; cohorts
Education	Years spent in education; highest degree obtained
Gender	Gender of the respondent (being male is the reference)
High-skill occupation	Based on the skill intensity of an individual's job
High Income	High income measured as annual/monthly personal/household income
Minority status	Measured as a self-reported ethnic minority, or migration background
Unemployed	Being unemployed
Urban residence	Rural or urban area of residence

Table 4.A5: Number of effect sizes

Individual factor	Number of estimates
Age	220
Education	248
Male	177
High-skill occupation	83
High Income	112
Minority status	147
Unemployed	98
Urban residence	100
Total	1185

4.A.5 Meta-analytical approach

This section gives a more detailed description of the meta-analytical technique used in the paper.

To analyze our metadata, we use meta-analytical multilevel random effects models. Stata contains several programs performing meta-analyses, such as the *metan* package, which can fit the random effects model. However, there is no dedicated module for performing meta-analysis with correlated estimates or, in our application, multilevel meta-analysis, where multiple estimates are "nested" within samples. A specific issue with our dataset is that dependencies between estimates may arise from the fact that most studies analyze widely-used survey data such as the European Social Survey, the European Value Study or the World Value Survey, or the World Values. To analyze our metadata, we use the method developed by Bagos (2015) using the *gllam* package in Stata, where estimates from each study are grouped - or nested - based on the data used.

It is worth stressing that despite the inclusion of random effects for the sample used, it remains impossible to identify the dependencies between estimates perfectly. For instance, some regressions are performed on a single wave and a single country of the ESS, while others pool together several waves and/or countries. The corresponding samples are therefore partly overlapping in a way that is impossible to identify without access to the actual data. We, therefore, follow Dinesen et al. (2020) and add random effects for the data used but ignore specific years of survey waves.

Table 4.A6 below shows how many estimates are drawn from partly or completely overlapping samples.

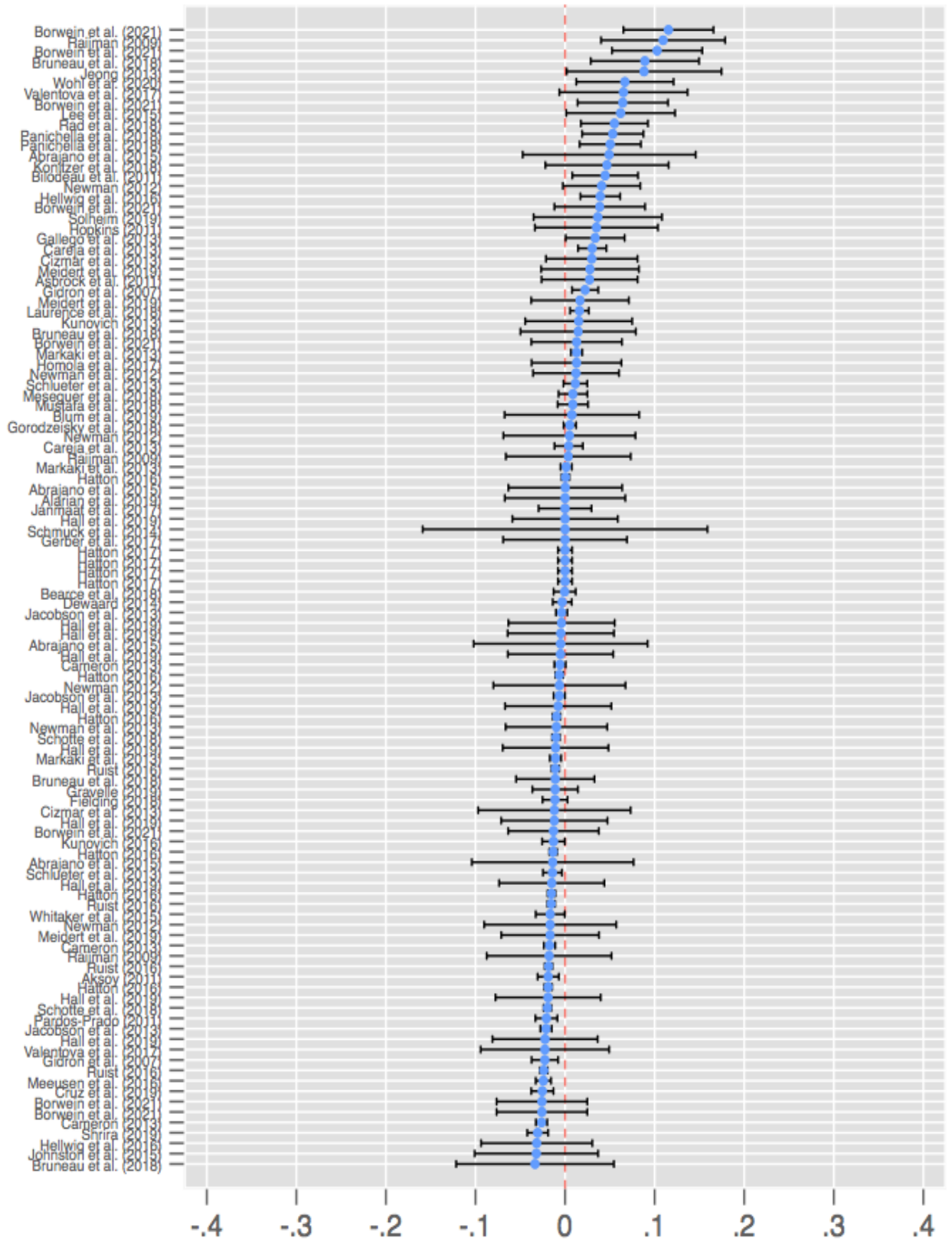
Table 4.A6: Datasets

Dataset	Number of estimates
European Social Survey	352
Own dataset	315
Citizenship, involvement, democracy (CID)	54
American National Election Studies	46
Pew Research Center surveys	36
International Social Survey Programme	35
European Value Study	30
British Election Study (BES)	28
Attitudes toward minority workers survey (Eurobarometer)	24
The American Panel Survey	24
Social Capital Community Benchmark Survey	20
General Social Survey (GSS)	26
CEO	18
German General Social Survey (ALLBUS)	17
South African Social Attitudes Survey (SASAS)	17
HAS	15
Copenhagen Area Surveys	14
Canadian Election Studies	11
AES	10
Longitudinal Internet Studies for the Social sciences	10
The Latin American Public Opinion Project (LAPOP)	8
Northern Ireland Life and Times Survey	8
BSA	7
CIS	7
World Value Survey	7
Danish National Election Studies (DNES)	6
Eurobarometer	6
SCIF	6
MOSAiCH Survey (Switzerland)	5
ASEP	4
SCDS (State Convention Delegate Study)	4
Swiss Electoral Studies (SELECTS)	4
Belgian Political Panel Survey	3
CMS	3
NAES	3
USS	1
Youth and society dataset	1

4.A.6 Forest plots

The following are forest plots of estimates of the relationship between each individual factor and positive attitudes toward immigration. They represent partial correlation coefficients with associated 95% confidence intervals.

Figure 4.A5: Age



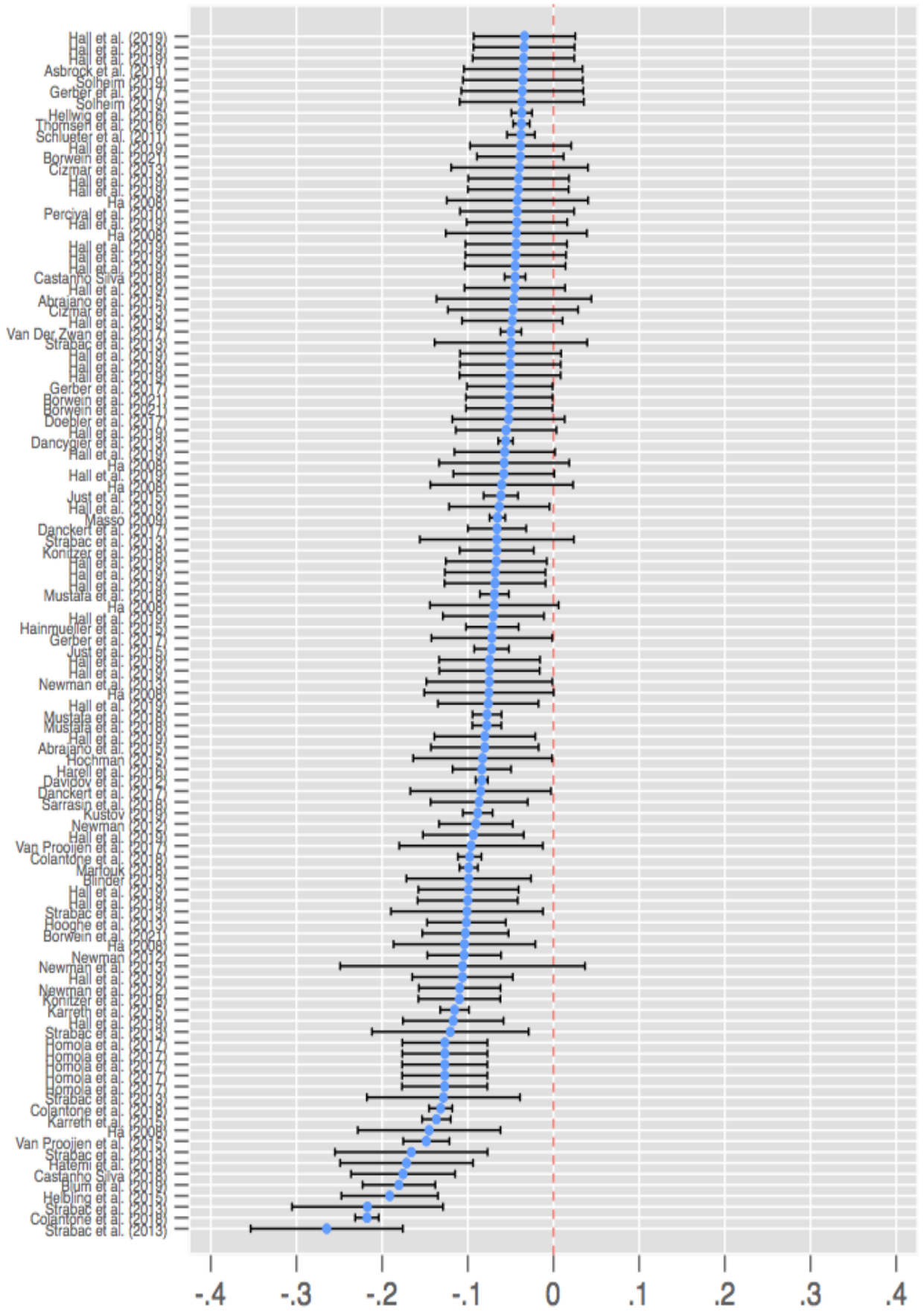
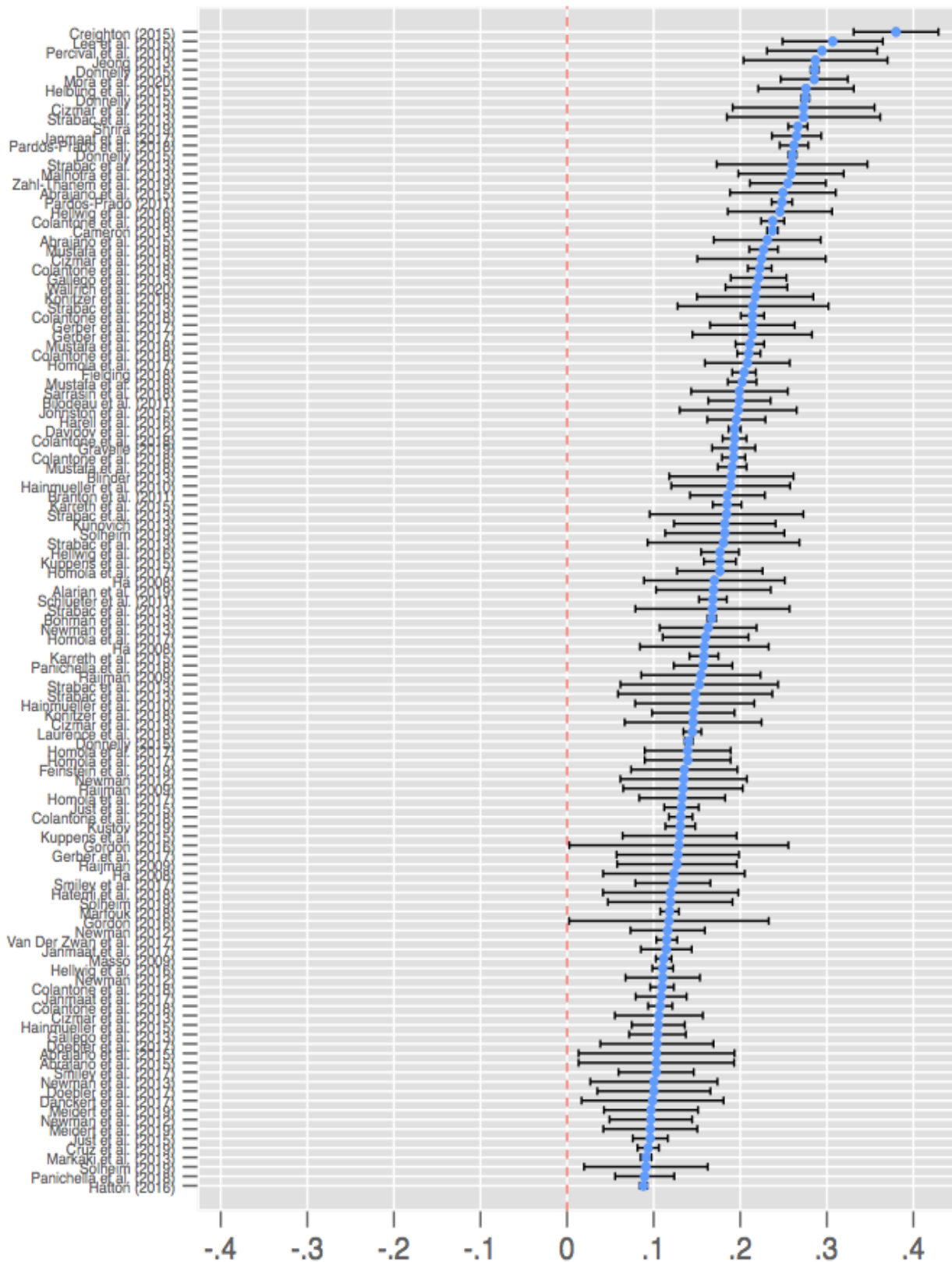


Figure 4.A6: Education



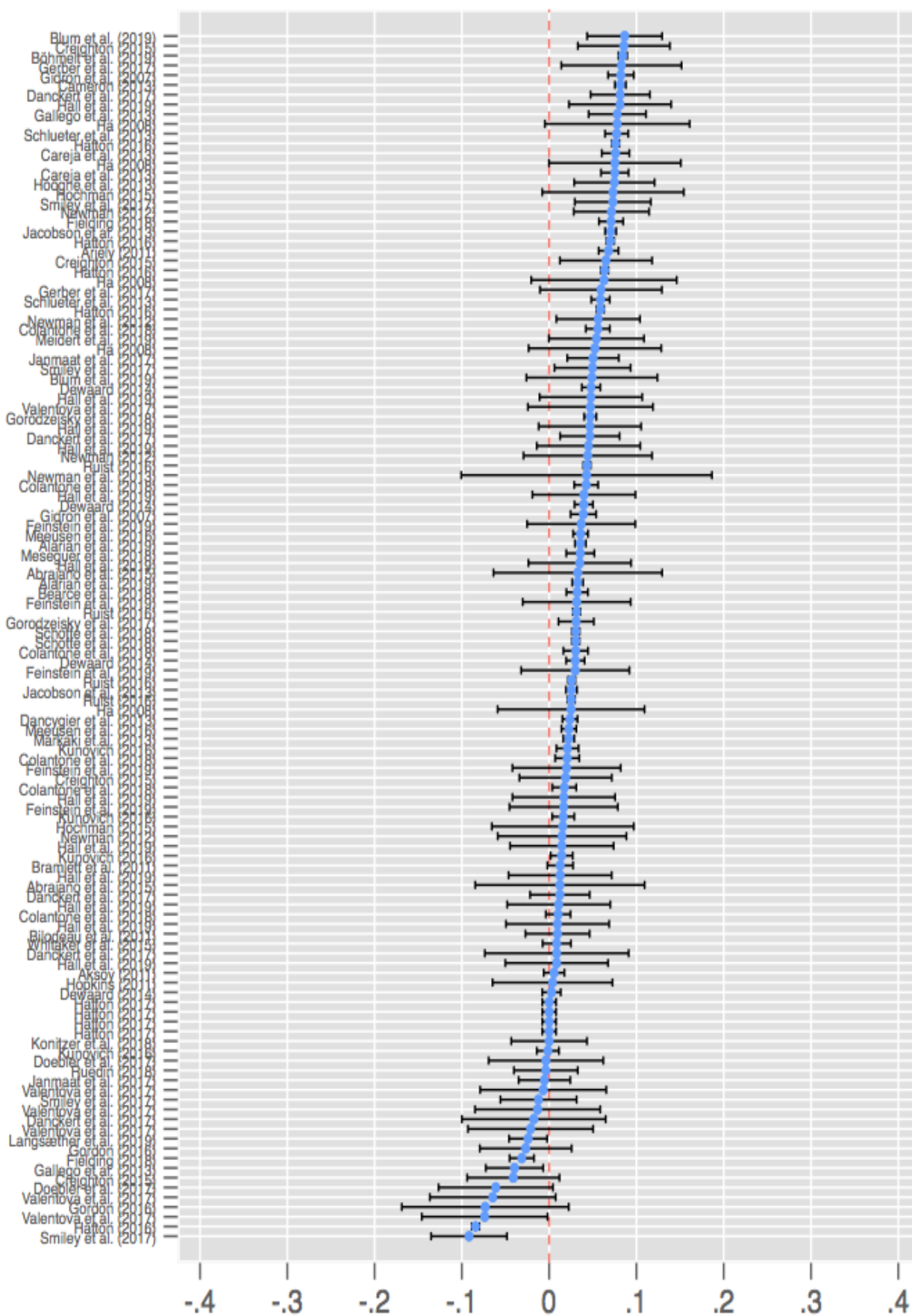
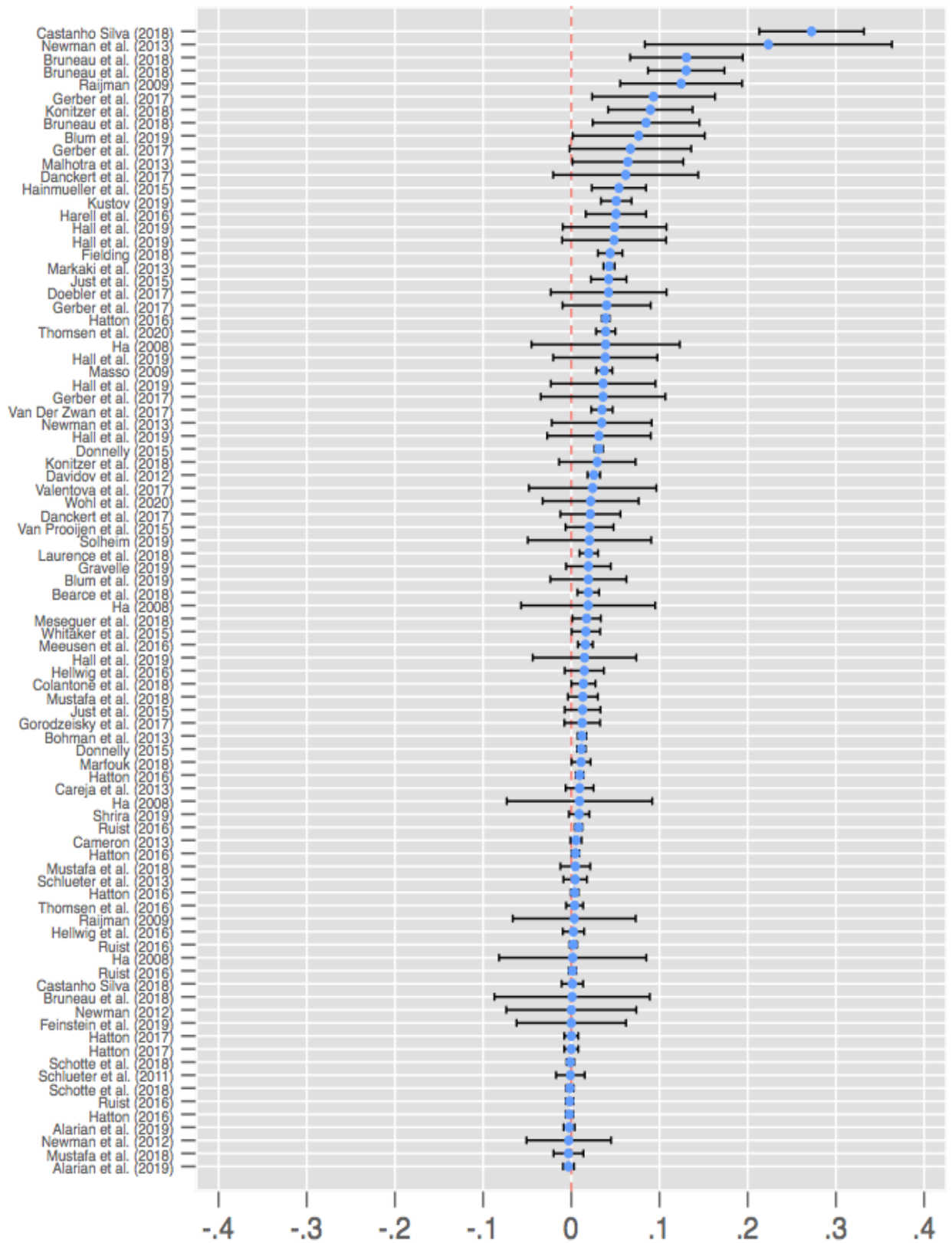


Figure 4.A7: Male



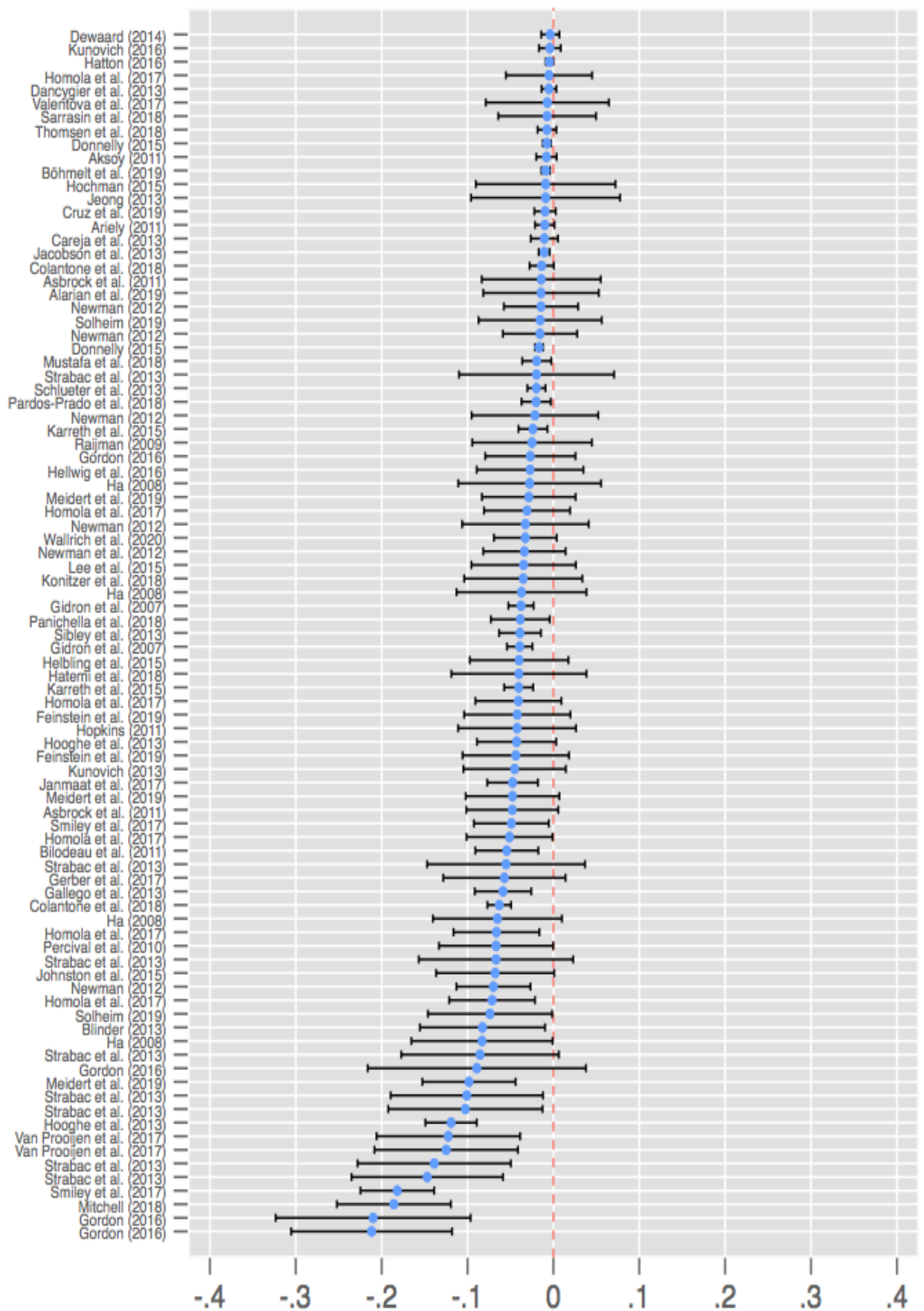
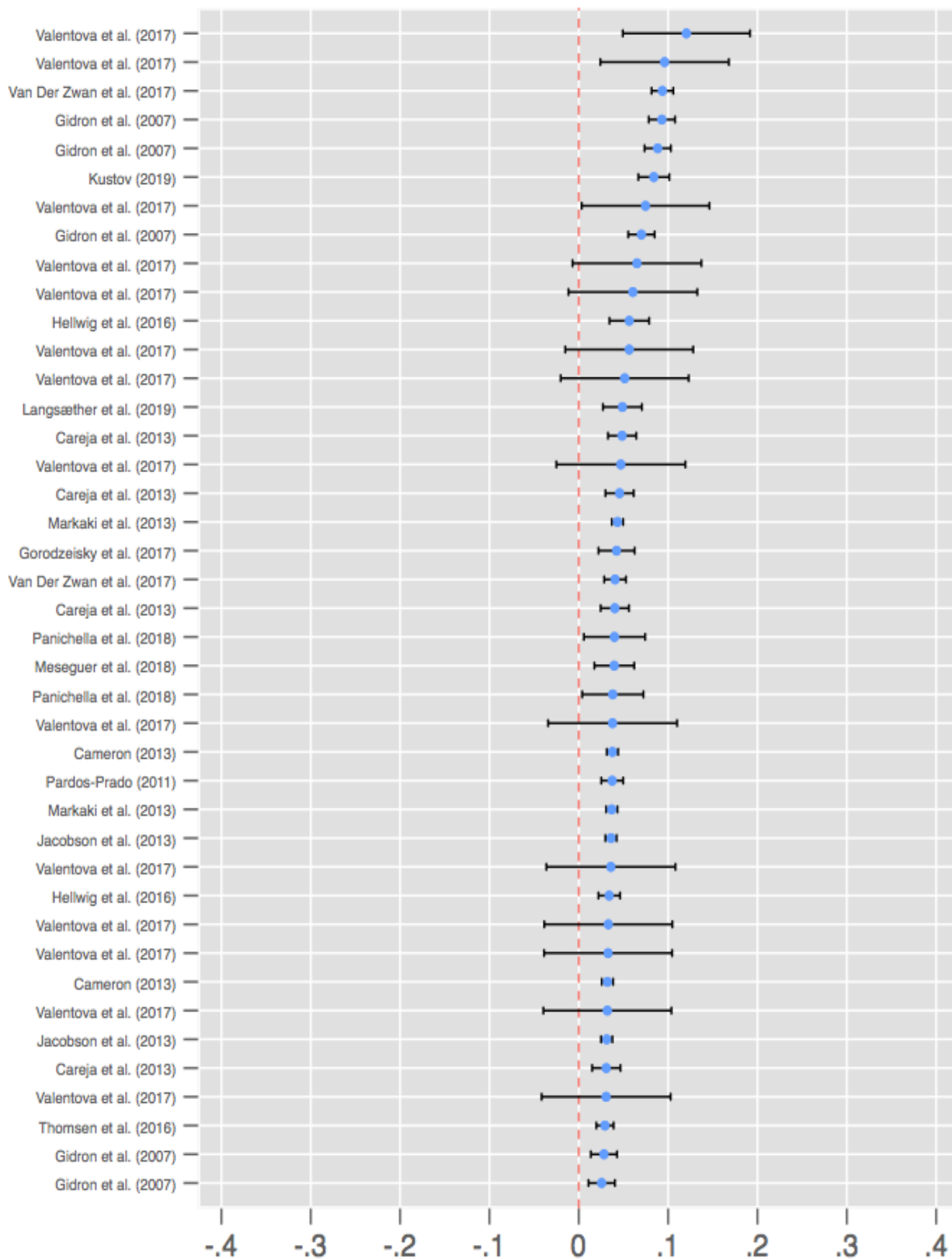


Figure 4.A8: High-skilled occupation



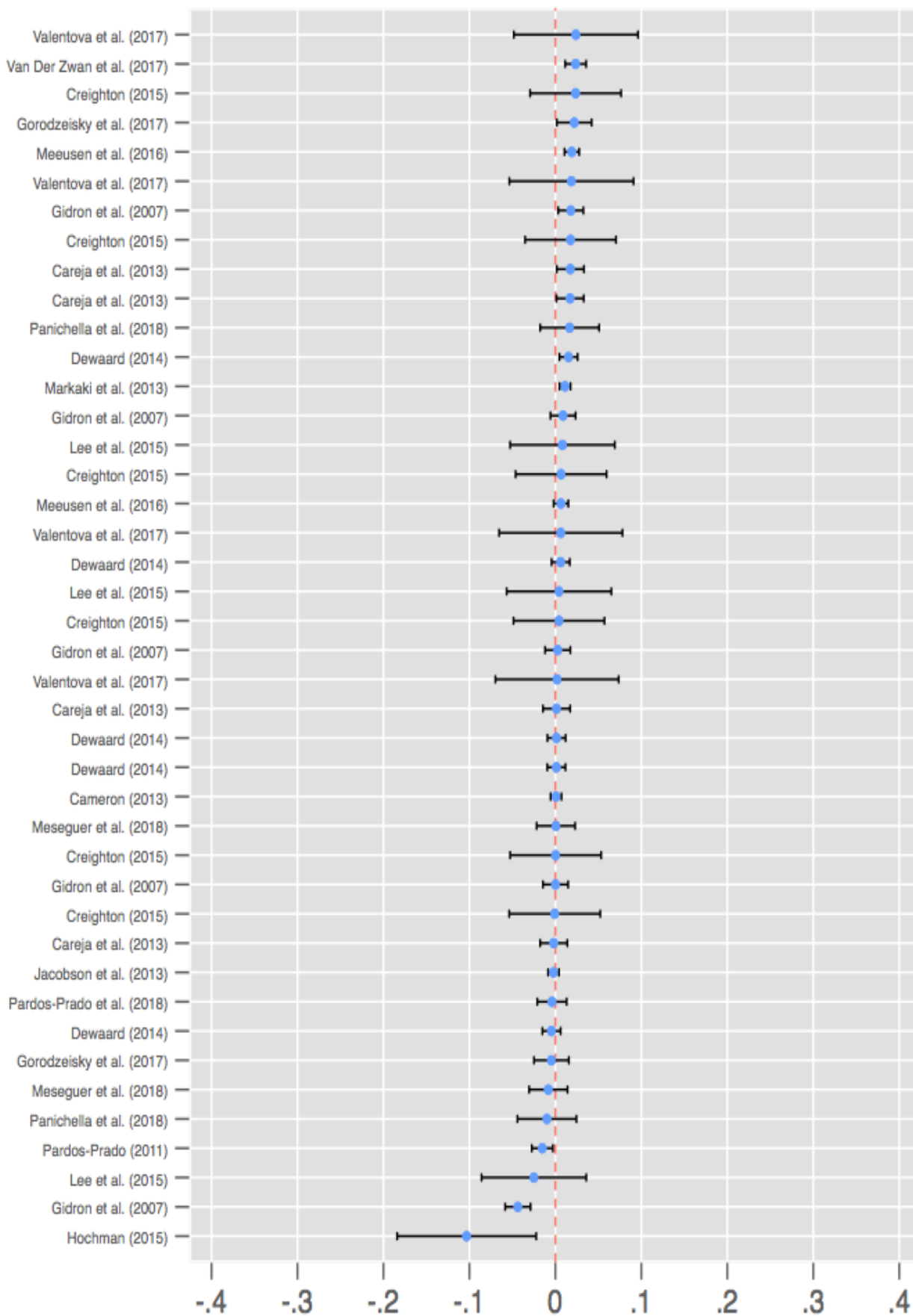
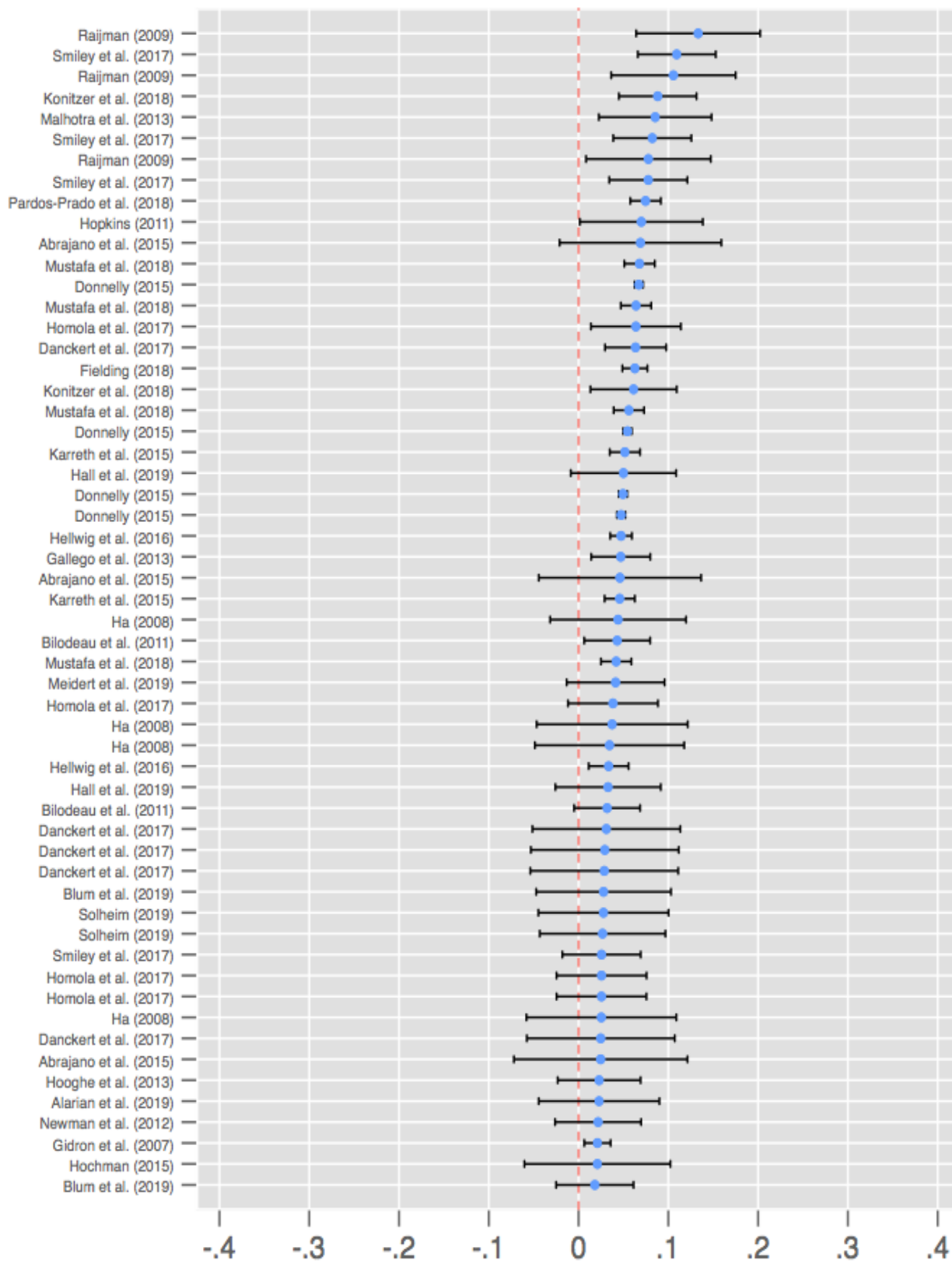


Figure 4.A9: High income



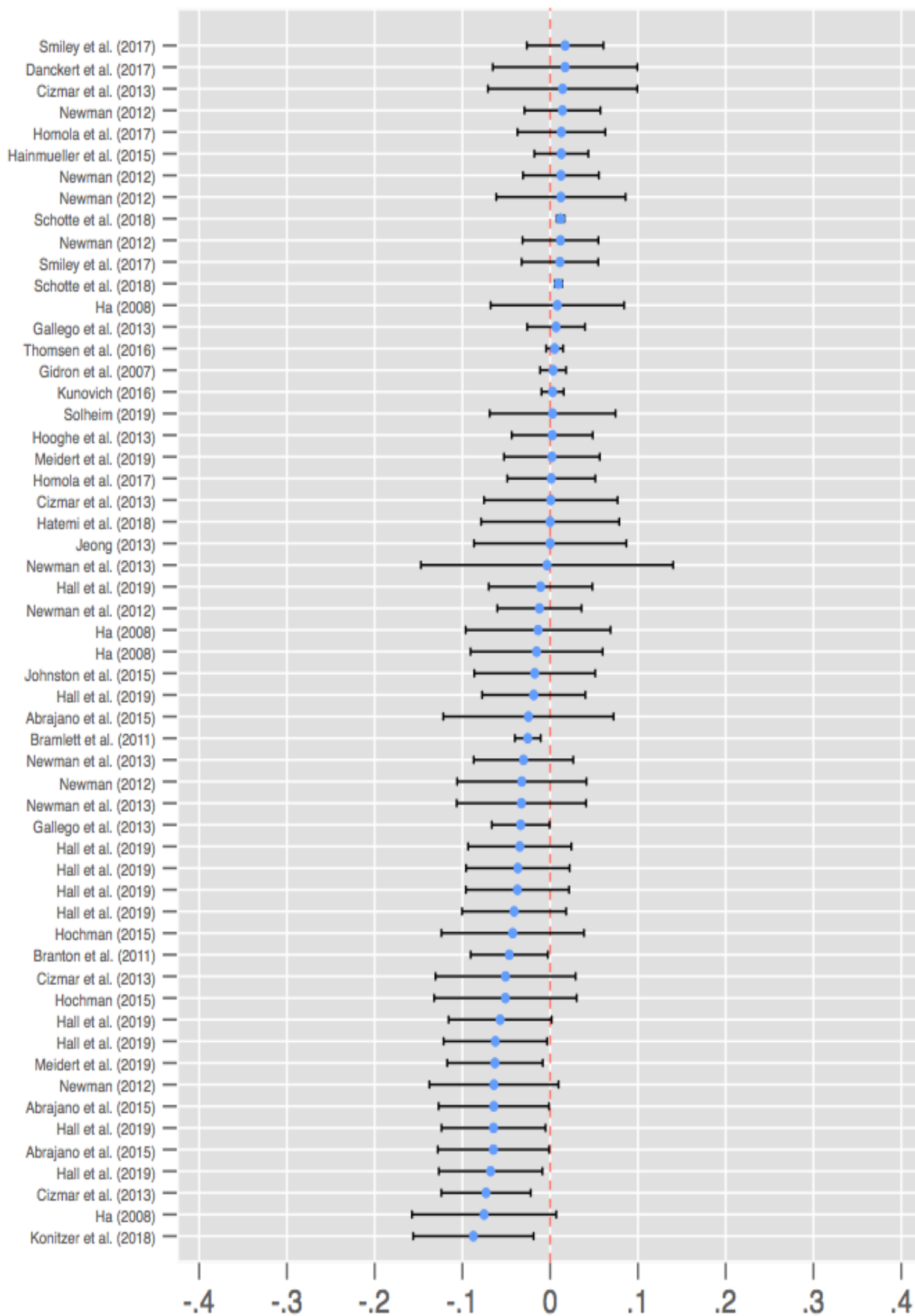
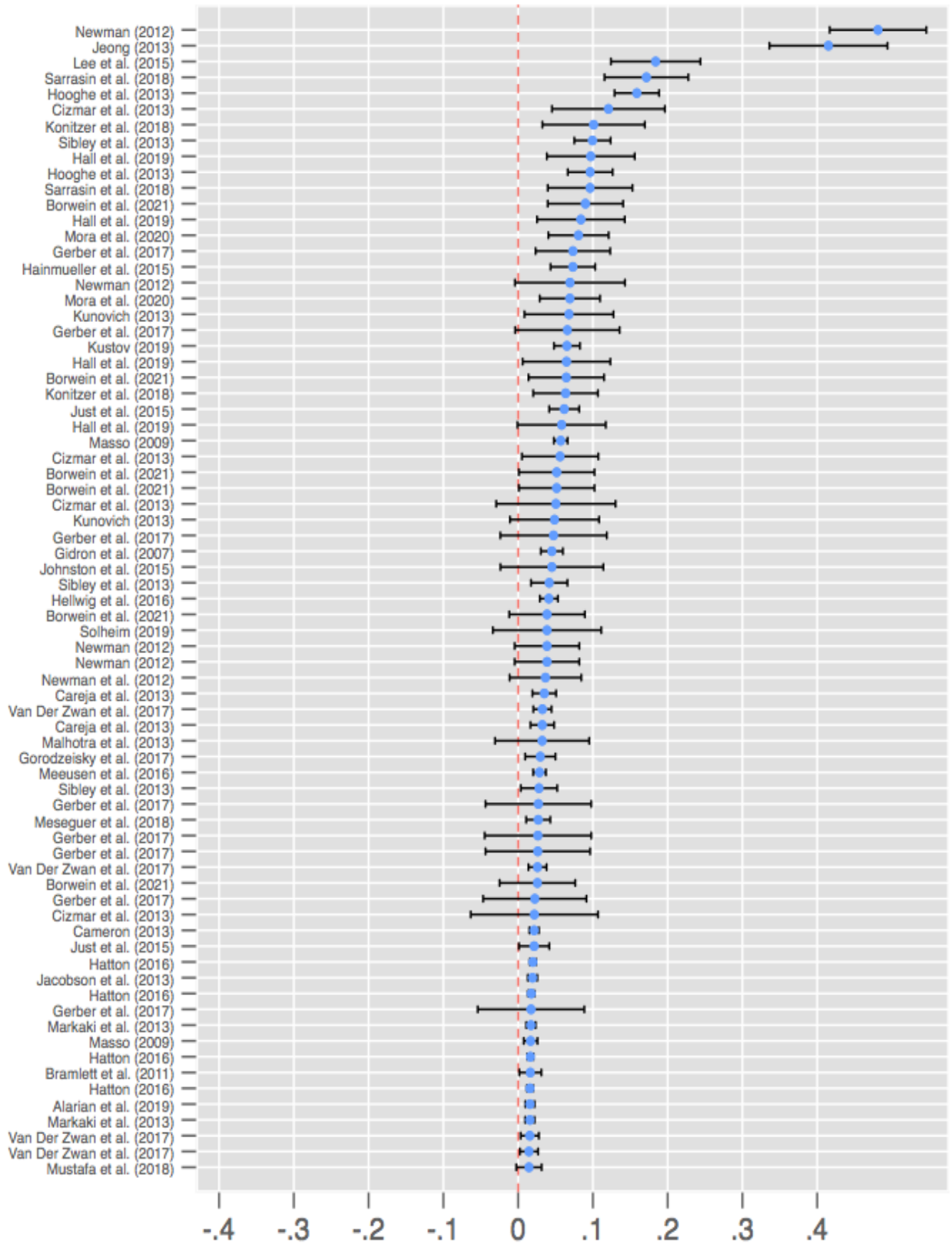


Figure 4.A10: Minority status



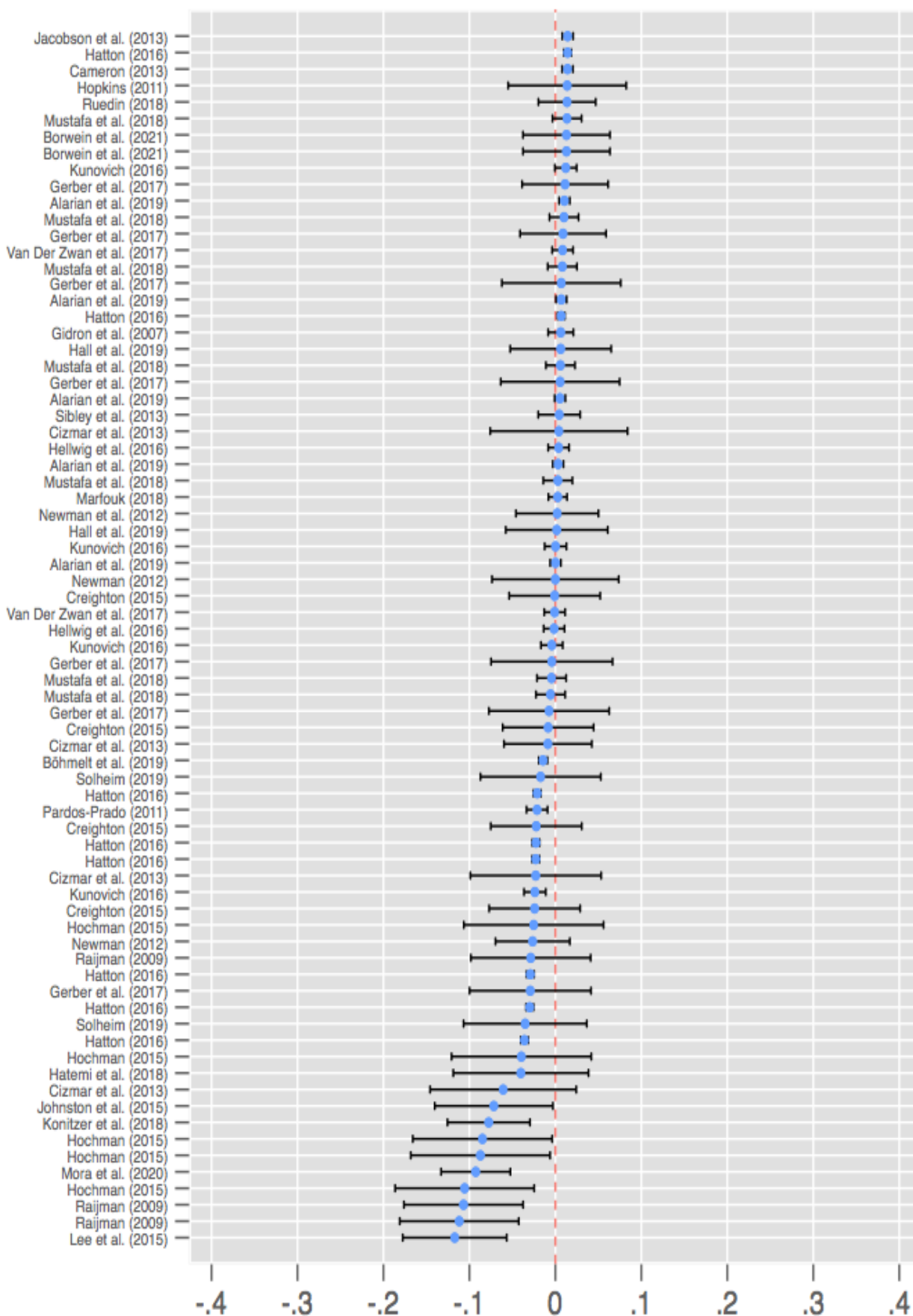
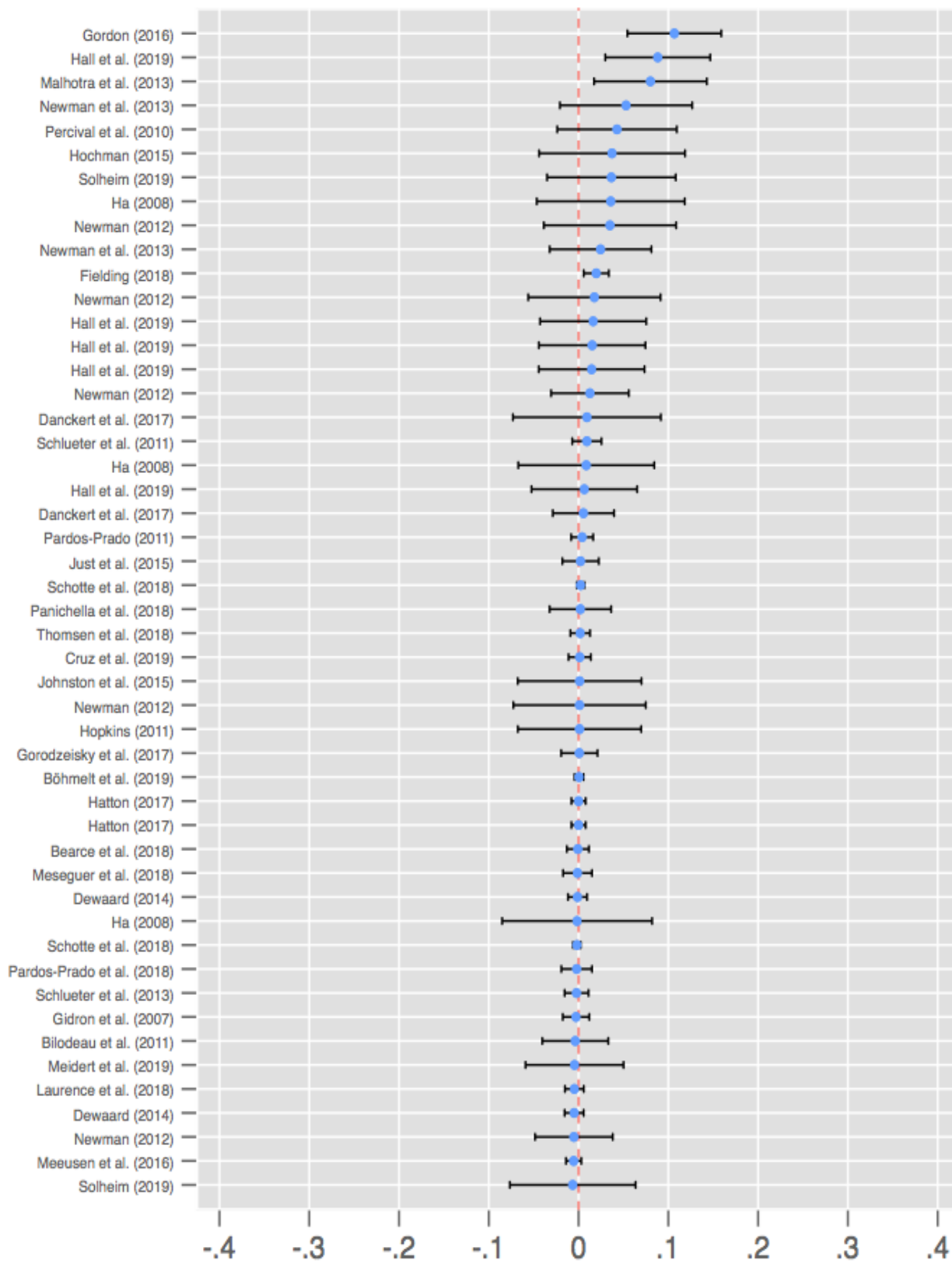


Figure 4.A11: Unemployed



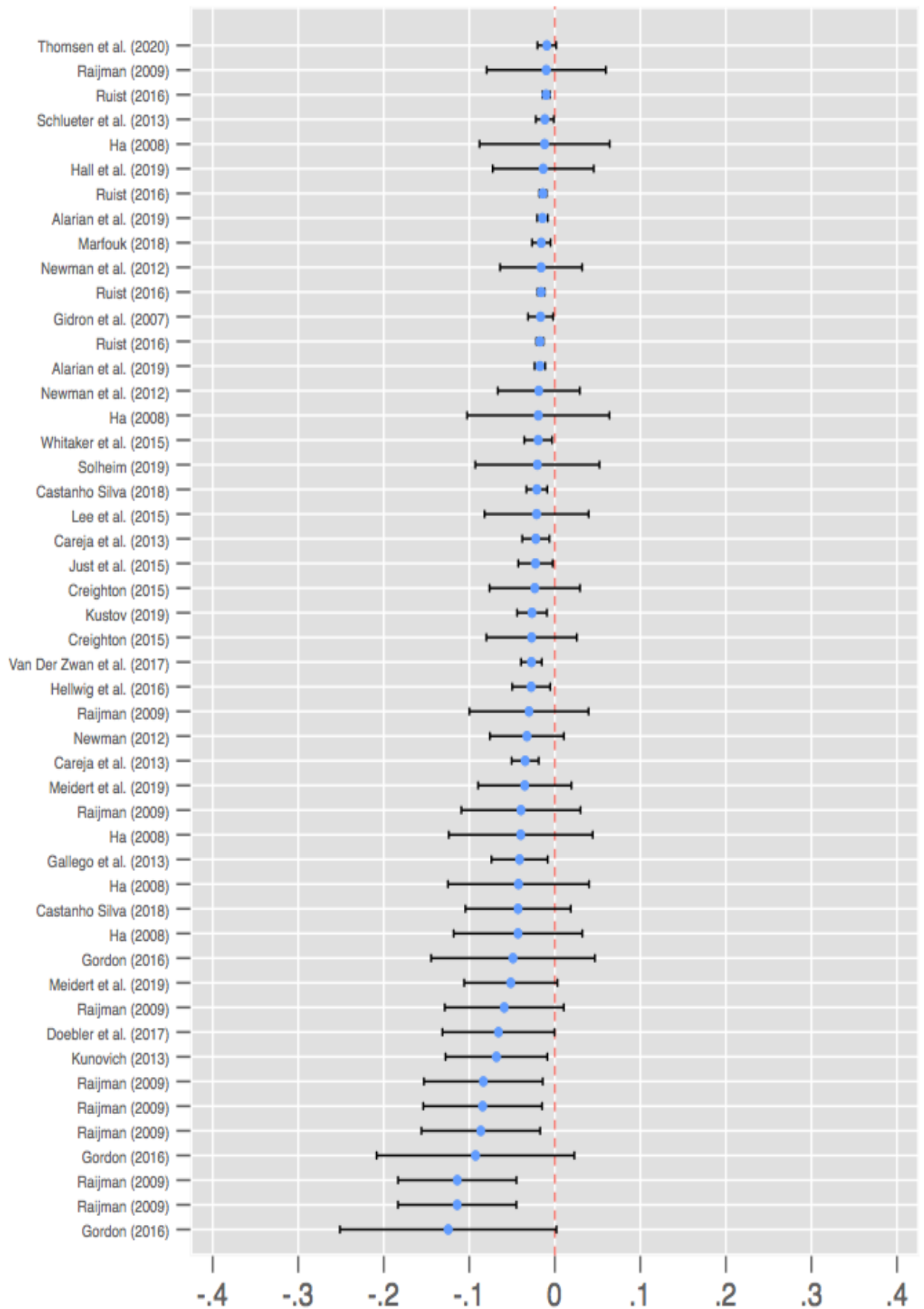
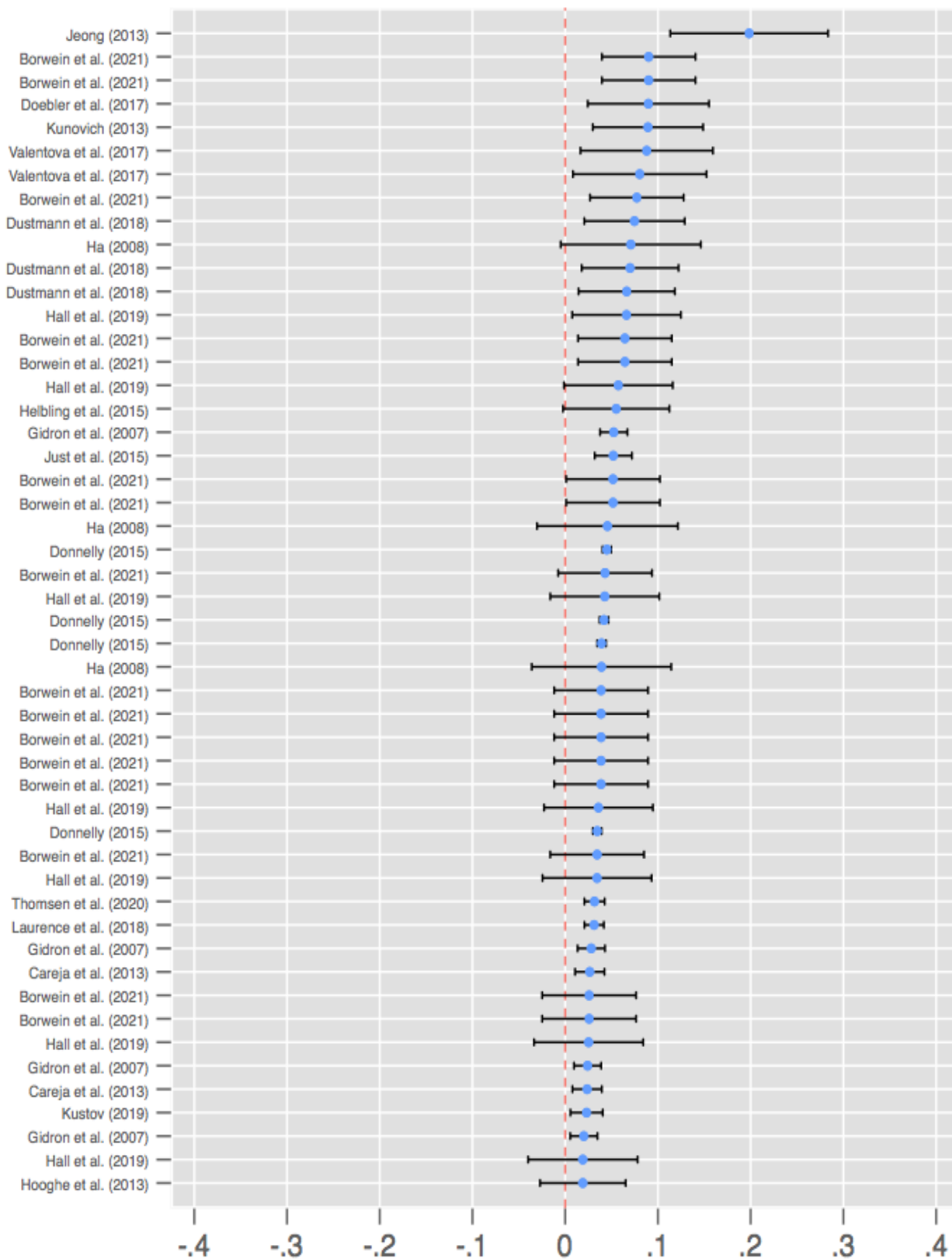
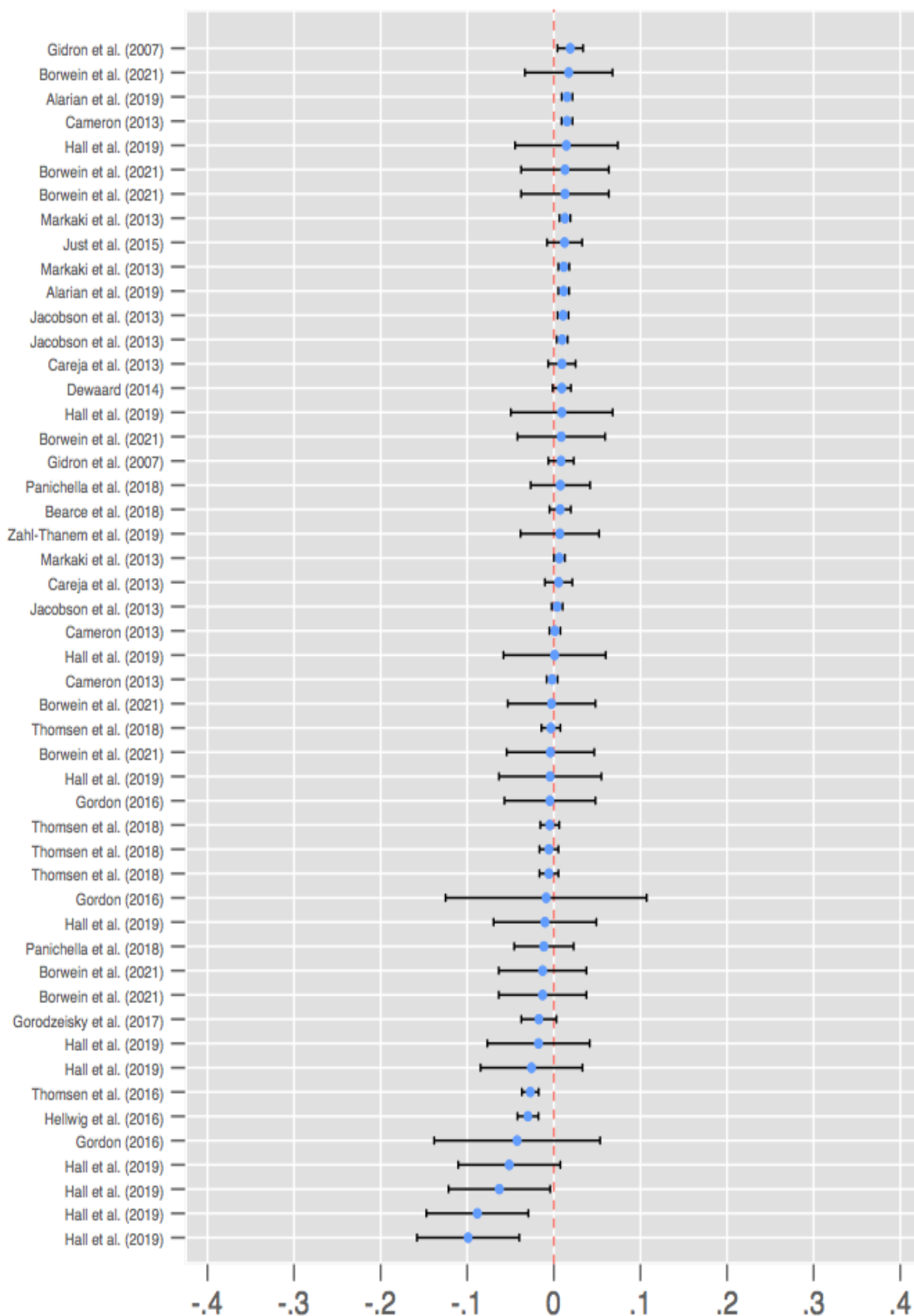


Figure 4.A12: Urban residence





4.A.7 Publication bias

In what follows, we first perform a publication bias test and explore how study characteristics influence the results.

Following Stanley and Doucouliagos (2015) and Cazachevici et al. (2020), we analyze publication bias using an estimate's standard error as a predictor in a regression weighted by the inverse of the variance of that estimate. The idea behind this test is that in the absence of publication bias, the estimated effect should be randomly distributed across studies, and the estimated effect size should not be correlated with its standard error. If the opposite is true and the standard error significantly predicts the size of the partial correlation, then estimates that suggest a particularly strong relationship between individual factors and attitudes to immigration would be systematically less precise. In that latter case, publication bias will have made results less robust and less likely to be replicable. Their internal and external validity would be lower and their results should thus be trusted less than in the absence of publication bias.

Table 4.A7 presents two models for each individual factor. The first one fits a weighted least square model as described above, while the second one is a weighted multilevel model with random effects clustering that accounts for dependencies between studies relying on the same samples. The only difference between the two models is that the second one takes the clustering of study populations into account. The interpretation of the results presented in Table 4.A7 is relatively straightforward. Following Dinesen et al. (2020), we interpret the intercept as the expected value of the partial correlation between a given individual factor and positive attitudes toward immigration if the standard error were equal to zero. Therefore, an intercept significantly different from zero suggests an overall systematic association between individual factors and attitudes toward immigration despite publication bias. The results in Table 4.A7 suggest that economic factors are robust predictors of attitudes to immigration. On the other hand, Table 4.A7 also reveals that the effect of age should be interpreted with caution since once we account for sample dependency, the intercept becomes small and insignificant.

Table 4.A7: Publication bias

<i>Indiv. factor</i>	Age	Age	Education	Education	Male	Male	HS Occ.	HS Occ.
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Intercept	-0.0129*** (-5.21)	-0.00647 (-1.83)	0.0795*** (10.54)	0.0795*** (10.58)	0.00689*** (3.61)	0.00689*** (3.63)	0.0231*** (5.32)	0.0231*** (5.39)
Standard Error	-1.330*** (-3.62)	-1.686** (-3.27)	1.365 (1.13)	1.365 (1.13)	-0.715* (-2.17)	-0.715* (-2.18)	0.168 (0.27)	0.168 (0.27)
Observations	220	220	248	248	177	177	83	83

<i>Indiv. factor</i>	Income	Income	Minority	Minority	Unemployed	Unemployed	Urban res.	Urban res.
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Intercept	0.0394*** (12.83)	0.0374*** (11.25)	-0.00350 (-1.31)	-0.00350 (-1.32)	-0.00804*** (-5.36)	-0.00819*** (-4.70)	0.0188*** (6.87)	0.0206*** (6.65)
Standard Error	-0.885* (-2.50)	-0.775* (-2.06)	1.779*** (3.73)	1.779*** (3.75)	-0.106 (-0.45)	-0.0799 (-0.31)	-0.112 (-0.27)	-0.174 (-0.37)
Observations	112	112	147	147	98	98	100	100

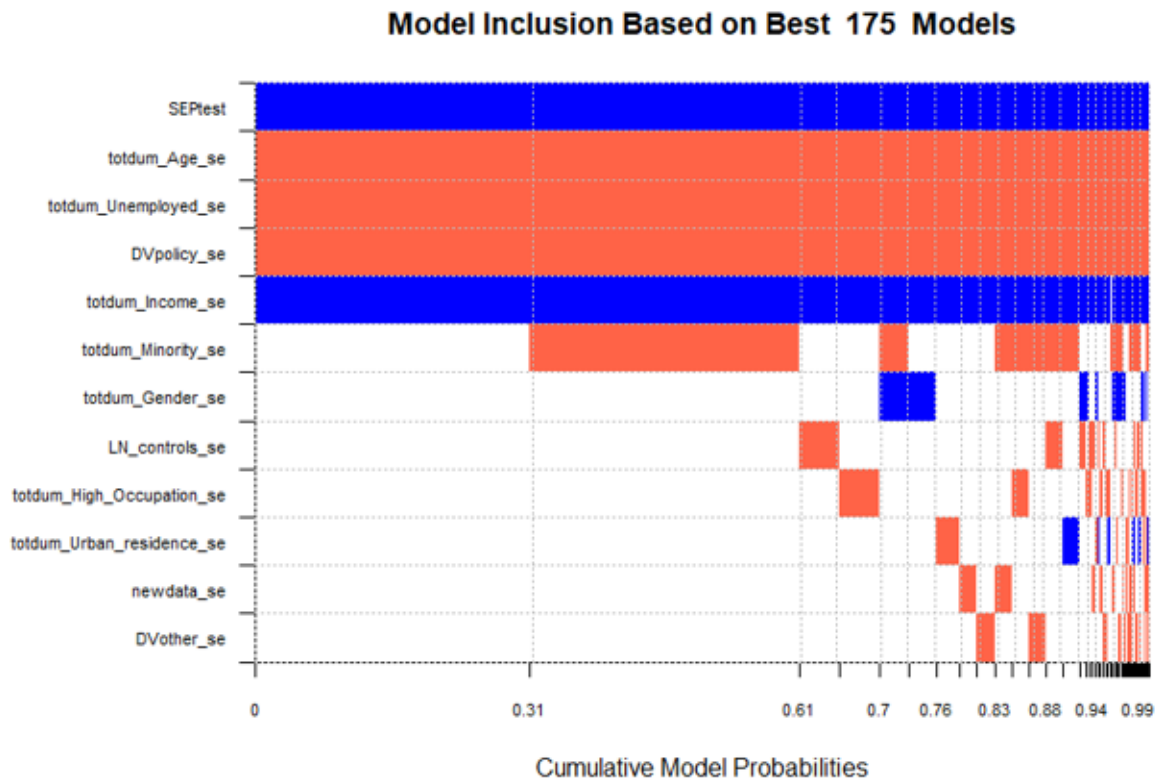
Note: Standard errors in parentheses. Notes: * $p < .10$, ** $p < .05$, *** $p < .01$

The table presents the estimates of the publication bias for the different independent variables. Model 1 fits a weighted least square model. Model 2 fits a weighted multilevel model with random effects clustering that accounts for dependencies between studies relying on the same samples.

4.A.8 Bayesian model averaging

With twelve potential explanatory variables, many correlated with each other. The results of a simple linear regression may suffer from over-specification bias due to model uncertainty. While the effect of each individual factor taken separately has been extensively discussed in the literature, there exist few theoretical frameworks that could help us decide which variables among them influence their partial correlation with attitudes to immigration. To address the resulting regression model uncertainty, we apply Bayesian model averaging (BMA) (Hoeting et al., 1999; Raftery et al., 1997)). This method is frequently used in meta-analyses, for example by Cazachevici et al. (2020) and Havranek and Sokolova (2020). BMA runs numerous different regressions with different subsets of the explanatory variables. Each of these models is then assigned the posterior model probability (PMP), measuring how well the model fits the data conditional on the model's size (this measure is equivalent to R^2). The final result is computed by the average model coefficients weighted by their PMP. Analogous to statistical significance in regression models, the posterior inclusion probability (PIP) is composed of the sum of the PMPs for all the models, including the variable. We use the `bms` R package developed by Zeugner and Feldkircher (2015) to estimate the BMA for the relationship between the different explanatory variables and attitudes toward immigrants. We visualize the graphical results of the BMA for education in Figure 4.A13.

Figure 4.A13: Model inclusion in Bayesian model averaging.



Note: The response variable is the effect of education on attitudes toward immigrants (partial correlation coefficient). A blue cell indicates a positive effect and a red cell indicates that the estimated coefficient of a variable is negative

The different explanatory variables are displayed on the vertical axis, sorted in descending order according to their PIPs. The horizontal axis represents the individual regression models sorted in descending order according to their inclusion probability. The color of the individual cell visualizes the sign of the regression coefficient. A blue cell indicates a positive effect, i.e., the variable causes the estimated effect of education on attitudes toward immigrants to be larger. By contrast, a red cell indicates that the estimated coefficient of a variable is negative. An empty cell indicates that the variable is not included in the model. Table 4.A8 presents the posterior mean, the standard deviation, and the PIP for each explanatory variable.

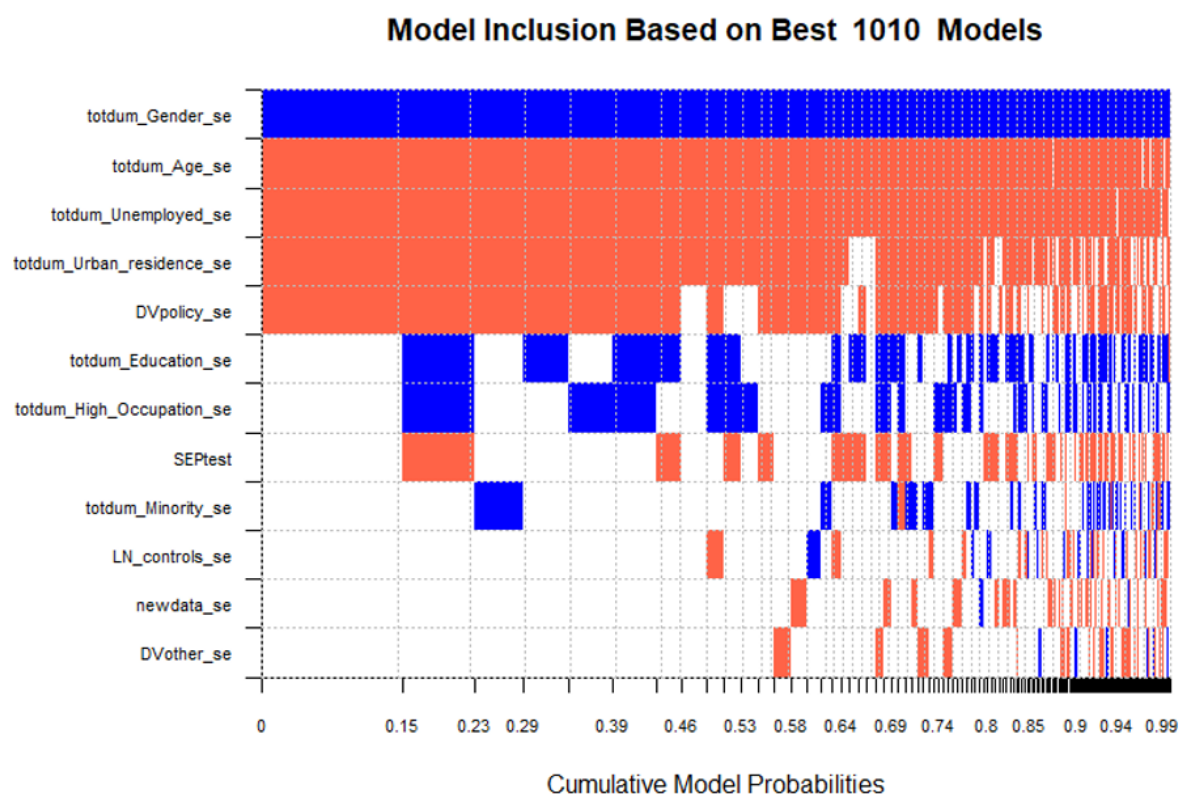
Table 4.A8: Numerical results of BMA for the education variable

	Posterior Mean	Posterior SD	PIP
SEPtest	0.190	0.018	1.000
LN_control	-0.001	0.005	0.093
newdata	-0.001	0.012	0.061
DVpolicy	-0.053	0.009	1.000
DVother	-0.001	0.018	0.060
Age	-0.078	0.011	1.000
Gender	0.003	0.011	0.098
High-skill_Occupation	-0.001	0.006	0.091
Income	0.044	0.011	0.996
Minority	-0.010	0.012	0.461
Unemployed	-0.061	0.009	1.000
Urban_residence	0.000	0.003	0.067
(Intercept)	0.873	NA	1.000

Note: PIP = posterior inclusion probability. SD = standard deviation. The table shows unconditional moments from Bayesian model averaging. PIPs with values below 0.5 denote a negligible effect, and PIPs above 0.99 indicate a decisive effect.

Kass and Raftery (1995) provide a rule of thumb for the interpretation of the PIPs. PIPs with values below 0.5 denote a negligible effect, and PIPs above 0.99 indicate a decisive effect. For the variables in question – namely income, occupation, and unemployment – the PIPs suggest a decisive impact of income and employment status, while the occupation is negligible. The results from Figure 4.A13 indicate that controlling for income strengthens the effect of education on attitudes toward immigrants. These results are expected because income and education capture partly overlapping concepts in the analysis. Without controlling for unemployment, the effect of education would be overestimated. Higher education commonly goes along with a lower possibility of being unemployed. Therefore, if an analysis lacks the unemployment variable, part of this effect is wrongly attributed to education. The same holds for the age variable.

Figure 4.A14: Model inclusion in Bayesian model averaging.



Note: The response variable is the effect of income on attitudes toward immigrants (partial correlation coefficient). A blue cell indicates a positive effect and a red cell indicates that the estimated coefficient of a variable is negative

Table 4.A9: Numerical results of BMA for the income variable

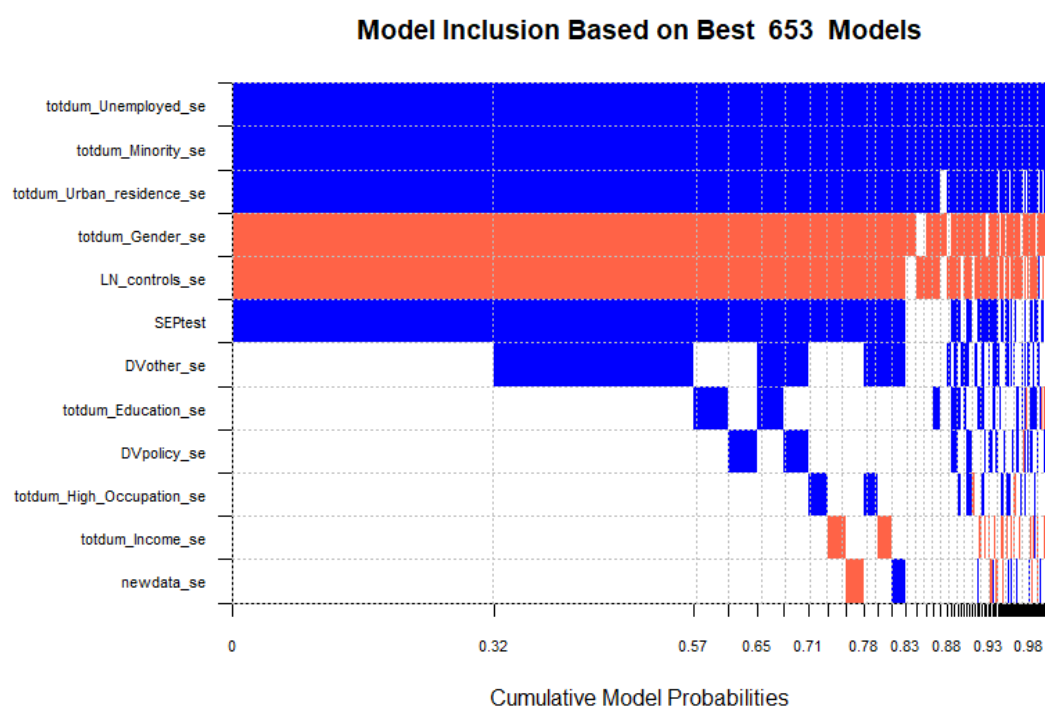
	Posterior Mean	Posterior SD	PIP
SEptest	-0.010	0.018	0.340
LN_control	0.000	0.004	0.126
newdata	-0.001	0.006	0.116
DVpolicy	-0.009	0.006	0.792
DVother	-0.001	0.005	0.097
Age	-0.032	0.009	0.993
Education	0.011	0.015	0.441
Gender	0.076	0.013	1.000
High-skill_Occupation	0.007	0.010	0.394
Minority	0.001	0.004	0.175
Unemployed	-0.022	0.006	0.991
Urban_residence	-0.020	0.010	0.906
(Intercept)	0.090	NA	1.000

Note: PIP = posterior inclusion probability. SD = standard deviation. The table shows unconditional moments from Bayesian model averaging. PIPs with values below 0.5 denote a negligible effect, and PIPs above 0.99 indicate a decisive effect.

Equivalently to the education variable, are the results regarding the income variable. Age and unemployment both have a strong and negative effect on the estimation of income on attitudes toward immigration (Figure 4.A14 and Table 4.A9). Similarly, we find a strong negative effect of (urban) residence on the estimation. Without controlling for these variables, the effect of income becomes overestimated. Without controlling for gender, the effect becomes decisively underestimated due to males earning more than females on average.

The results of the BMA for age are shown in Figure 4.A15 and Table 4.A10.

Figure 4.A15: Model inclusion in Bayesian model averaging.



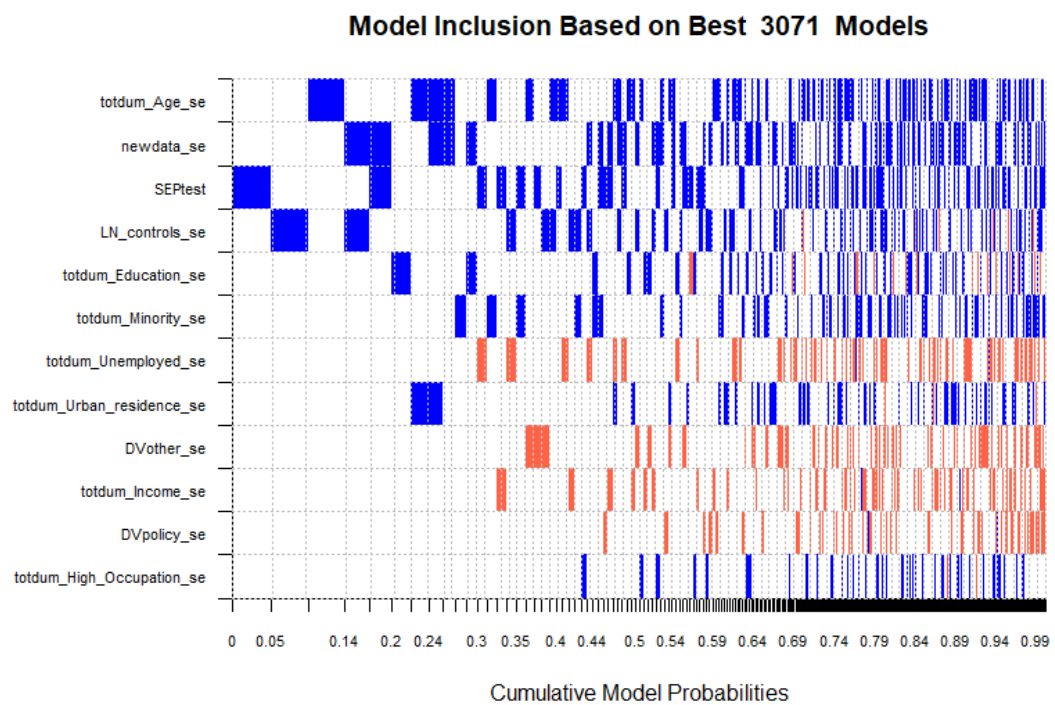
Note: The response variable is the effect of age on attitudes toward immigrants (partial correlation coefficient). A blue cell indicates a positive effect and a red cell indicates that the estimated coefficient of a variable is negative

Table 4.A10: Numerical results of BMA for the age variable

	Posterior Mean	Posterior SD	PIP
SEptest	0.056	0.026	0.892
LN_control	-0.032	0.011	0.951
newdata	0.000	0.004	0.065
DVpolicy	0.000	0.002	0.105
DVother	0.042	0.059	0.416
Education	0.002	0.007	0.128
Gender	-0.040	0.013	0.968
High-skill_Occupation	0.000	0.003	0.077
Income	0.000	0.001	0.067
Minority	0.026	0.005	1.000
Unemployed	0.025	0.005	1.000
Urban_residence	0.038	0.011	0.981
(Intercept)	-1.593	NA	1.000

Note: PIP = posterior inclusion probability. SD = standard deviation. The table shows unconditional moments from Bayesian model averaging. PIPs with values below 0.5 denote a negligible effect, and PIPs above 0.99 indicate a decisive effect.

Figure 4.A16: Model inclusion in Bayesian model averaging.



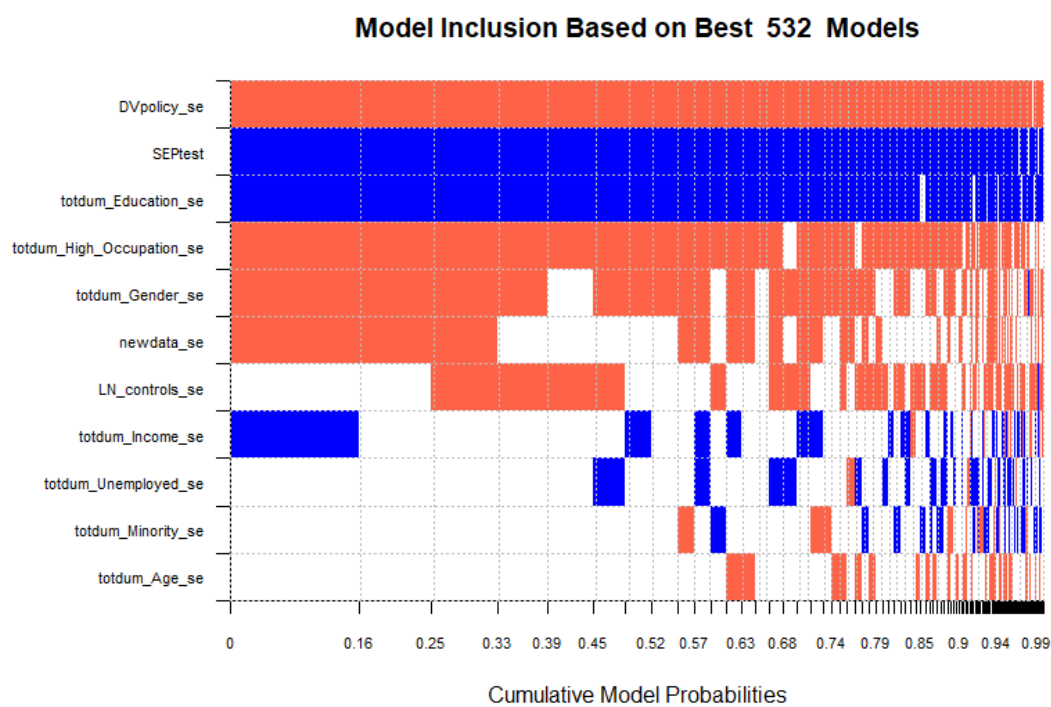
Note: The response variable is the effect of gender on attitudes toward immigrants (partial correlation coefficient). A blue cell indicates a positive effect and a red cell indicates that the estimated coefficient of a variable is negative

Table 4.A11: Numerical results of BMA for gender

	Posterior Mean	Posterior SD	PIP
SEPtest	0.003	0.004	0.333
LN_control	0.001	0.002	0.283
newdata	0.011	0.018	0.345
DVpolicy	0.000	0.001	0.111
DVother	-0.009	0.026	0.158
Age	0.002	0.004	0.356
Education	0.001	0.003	0.187
High-skill_Occupation	0.000	0.002	0.102
Income	-0.001	0.002	0.148
Minority	0.001	0.002	0.185
Unemployed	-0.001	0.002	0.176
Urban_residence	0.001	0.003	0.174
(Intercept)	-0.791	NA	1.000

Note: PIP = posterior inclusion probability. SD = standard deviation. The table shows unconditional moments from Bayesian model averaging. PIPs with values below 0.5 denote a negligible effect, and PIPs above 0.99 indicate a decisive effect.

Figure 4.A17: Model inclusion in Bayesian model averaging.



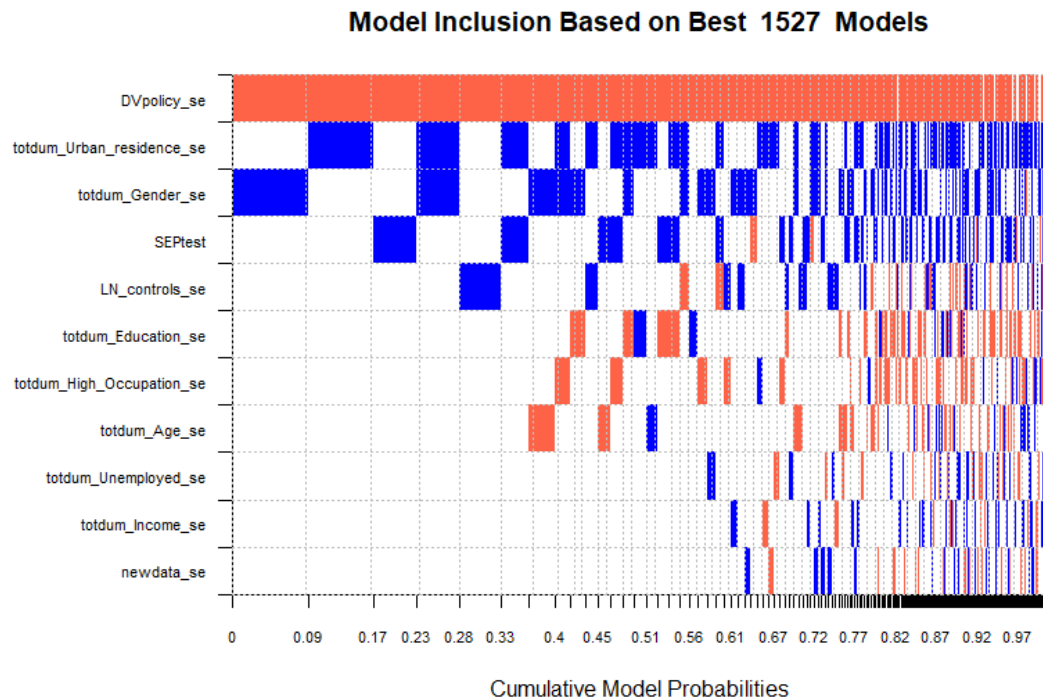
Note: The response variable is the effect of urban residence on attitudes toward immigrants (partial correlation coefficient). A blue cell indicates a positive effect and a red cell indicates that the estimated coefficient of a variable is negative

Table 4.A12: Numerical results of BMA for (urban) residence

	Posterior Mean	Posterior SD	PIP
SEPtest	0.062	0.022	0.995
LN_control	-0.007	0.009	0.478
newdata	-0.017	0.019	0.544
DVpolicy	-0.019	0.004	0.999
Age	-0.001	0.003	0.139
Education	0.023	0.006	0.983
Gender	-0.034	0.022	0.807
High-skill_Occupation	-0.025	0.009	0.945
Income	0.002	0.004	0.344
Minority	0.000	0.004	0.159
Unemployed	0.001	0.003	0.200
(Intercept)	0.438	NA	1.000

Note: PIP = posterior inclusion probability. SD = standard deviation. The table shows unconditional moments from Bayesian model averaging. PIPs with values below 0.5 denote a negligible effect, and PIPs above 0.99 indicate a decisive effect.

Figure 4.A18: Model inclusion in Bayesian model averaging.



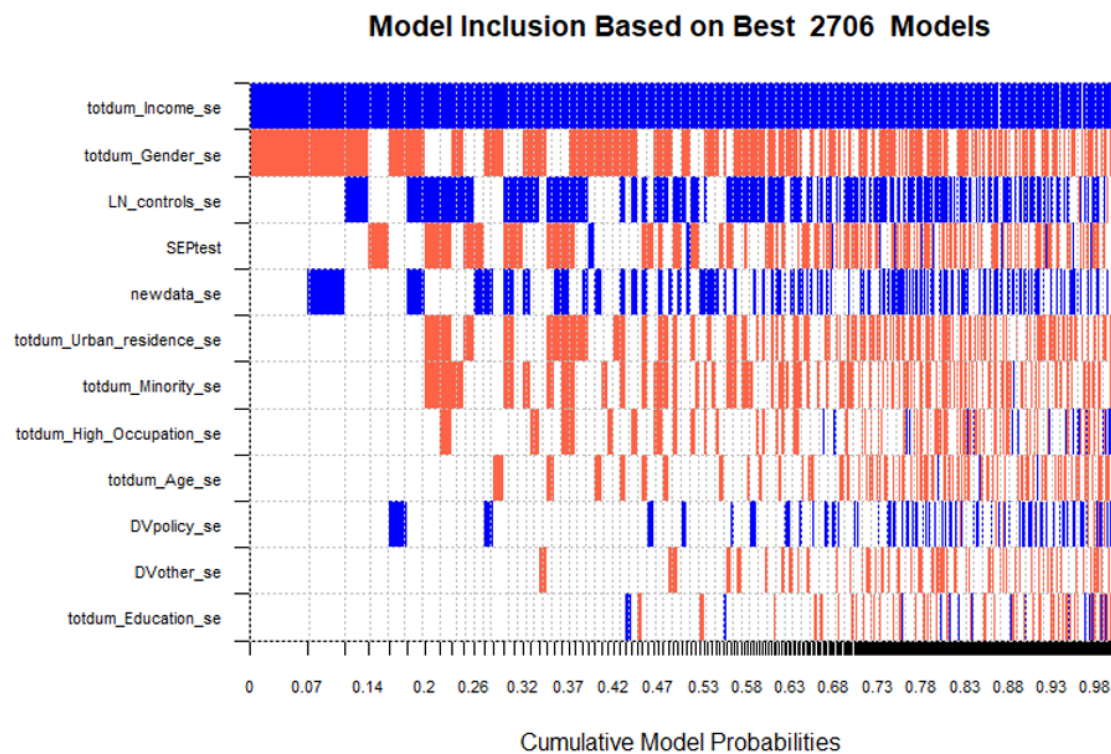
Note: The response variable is the effect of minority background on attitudes toward immigrants (partial correlation coefficient). A blue cell indicates a positive effect and a red cell indicates that the estimated coefficient of a variable is negative

Table 4.A13: Numerical results of BMA for minority status

	Posterior Mean	Posterior SD	PIP
SEPtest	0.007	0.015	0.299
LN_control	0.000	0.004	0.213
newdata	0.000	0.006	0.083
DVpolicy	-0.019	0.007	0.982
Age	-0.001	0.003	0.155
Education	-0.003	0.012	0.198
Gender	0.008	0.012	0.428
High-skill_Occupation	-0.001	0.005	0.156
Income	0.000	0.002	0.084
Unemployed	0.000	0.002	0.088
Urban_residence	0.005	0.007	0.474
(Intercept)	1.289	NA	1.000

Note: PIP = posterior inclusion probability. SD = standard deviation. The table shows unconditional moments from Bayesian model averaging. PIPs with values below 0.5 denote a negligible effect, and PIPs above 0.99 indicate a decisive effect.

Figure 4.A19: Model inclusion in Bayesian model averaging.



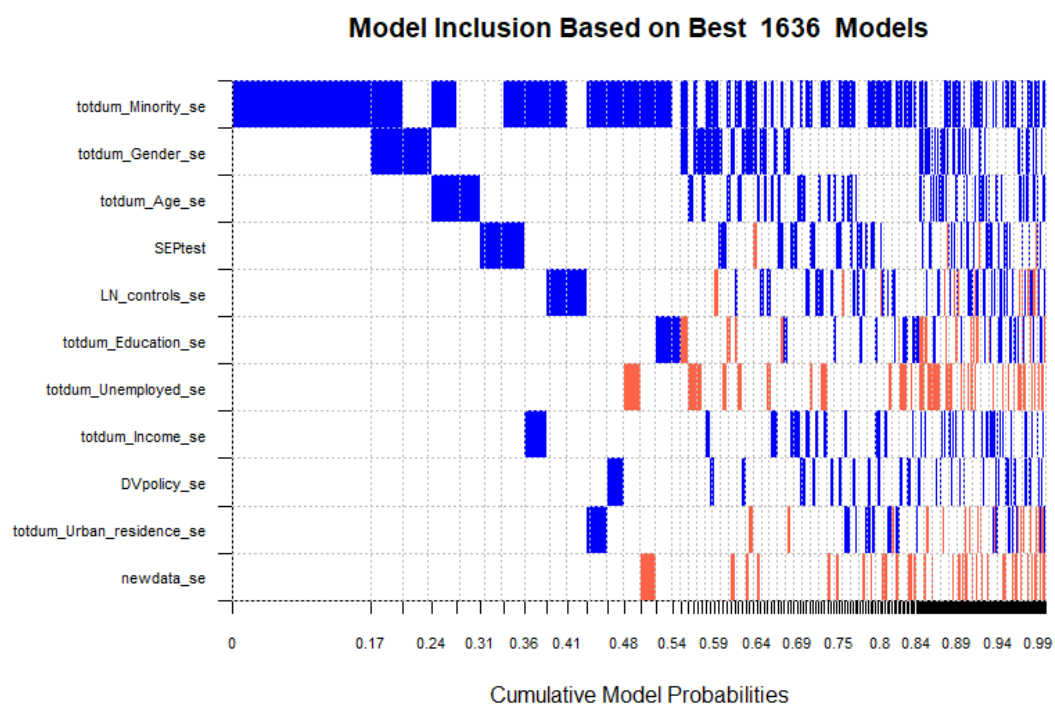
Note: The response variable is the effect of unemployment on attitudes toward immigrants (partial correlation coefficient). A blue cell indicates a positive effect and a red cell indicates that the estimated coefficient of a variable is negative.

Table 4.A14: Numerical results of BMA for unemployment

	Posterior Mean	Posterior SD	PIP
SEPtest	-0.009	0.015	0.416
LN_control	0.006	0.007	0.554
newdata	0.010	0.015	0.397
DVpolicy	0.000	0.001	0.175
DVother	-0.003	0.010	0.141
Age	-0.001	0.002	0.185
Education	0.000	0.002	0.126
Gender	-0.009	0.009	0.652
High-skill_Occupation	-0.001	0.003	0.210
Income	0.009	0.003	0.987
Minority	-0.002	0.004	0.351
Urban_residence	-0.003	0.005	0.357
(Intercept)	-0.349	NA	1.000

Note: PIP = posterior inclusion probability. SD = standard deviation. The table shows unconditional moments from Bayesian model averaging. PIPs with values below 0.5 denote a negligible effect, and PIPs above 0.99 indicate a decisive effect.

Figure 4.A20: Model inclusion in Bayesian model averaging.



Note: The response variable is the effect of (high-skill) occupation on attitudes toward immigrants (partial correlation coefficient). A blue cell indicates a positive effect and a red cell indicates that the estimated coefficient of a variable is negative

Table 4.A15: Numerical results of BMA for (high-skill) occupation

	Posterior Mean	Posterior SD	PIP
SEPttest	0.002	0.008	0.171
LN_control	0.001	0.003	0.163
newdata	-0.001	0.011	0.106
DVpolicy	0.001	0.007	0.111
Age	0.002	0.007	0.189
Education	0.000	0.006	0.137
Gender	0.003	0.008	0.227
High_Occupation	0.001	0.003	0.125
Income	0.012	0.010	0.665
Minority	-0.001	0.003	0.131
Unemployed	0.000	0.003	0.111
(Intercept)	0.802	NA	1.000

Note: PIP = posterior inclusion probability. SD = standard deviation. The table shows unconditional moments from Bayesian model averaging. PIPs with values below 0.5 denote a negligible effect, and PIPs above 0.99 indicate a decisive effect.

4.A.9 Description and Summary Statistics

Table 4.A16: AGE - Description and summary statistics of explanatory variables

Variable	Definition	Mean	St. Dev.
TSTAT	Estimated t-statistic of the effect size	-2,55	4,46
PCC	Partial correlation coefficient	-0,04	0,06
Precision	Precision of the estimated partial correlation coefficient (the inverse of the standard error)	94,35	114,84
Number of variables	Logarithm of number of explanatory variables	2,24	0,54
Original data	Dummy, 1 if original data used, 0 otherwise	0,41	0,49
<i>Measure of attitudes</i>			
Migration policy	Dummy, 1 if attitudes toward the number of immigrants are used, 0 otherwise	0,55	0,50
Mixed attitudes	Dummy, 1 if attitudes toward both the number and contribution of immigrants are used, 0 otherwise	0,01	0,10
Contribution of immigrants (base cat.)	Dummy, 1 if attitudes toward the contribution of immigrants are used, 0 otherwise	0,44	0,50
<i>Control variables - factors</i>			
Age	Dummy, 1 if age is included, 0 otherwise	1,00	0,00
Education	Dummy, 1 if education is included, 0 otherwise	0,86	0,34
Gender	Dummy, 1 if gender is included, 0 otherwise	0,85	0,35
High-skill occupation	Dummy, 1 if occupation is included, 0 otherwise	0,12	0,33
Income	Dummy, 1 if income is included, 0 otherwise	0,52	0,50
Minority status	Dummy, 1 if minority status is included, 0 otherwise	0,55	0,50
Unemployed	Dummy, 1 if employment status is included, 0 otherwise	0,51	0,50
Urban residence	Dummy, 1 if a measure of whether individuals lives in an urban or rural area is included, 0 otherwise	0,40	0,49

Table 4.A17: EDUCATION - Description and summary statistics of explanatory variables

Variable	Definition	Mean	St. Dev.
TSTAT	Estimated t-statistic of the effect size	9,96	16,92
PCC	Partial correlation coefficient	0,10	0,09
Precision	Precision of the estimated partial correlation coefficient (the inverse of the standard error)	108,14	118,76
Number of variables	Logarithm of number of explanatory variables	2,18	0,49
Original data	Dummy, 1 if original data used, 0 otherwise	0,21	0,40
<i>Measure of attitudes</i>			
Migration policy	Dummy, 1 if attitudes toward the number of immigrants are used, 0 otherwise	0,46	0,50
Mixed attitudes	Dummy, 1 if attitudes toward both the number and contribution of immigrants are used, 0 otherwise	0,02	0,15
Contribution of immigrants (base cat.)	Dummy, 1 if attitudes toward the contribution of immigrants are used, 0 otherwise	0,52	0,50
<i>Control variables - factors</i>			
Age	Dummy, 1 if age is included, 0 otherwise	0,79	0,41
Education	Dummy, 1 if education is included, 0 otherwise	1,00	0,00
Gender	Dummy, 1 if gender is included, 0 otherwise	0,87	0,34
High-skill occupation	Dummy, 1 if occupation is included, 0 otherwise	0,17	0,37
Income	Dummy, 1 if income is included, 0 otherwise	0,45	0,50
Minority status	Dummy, 1 if minority status is included, 0 otherwise	0,38	0,49
Unemployed	Dummy, 1 if employment status is included, 0 otherwise	0,46	0,50
Urban residence	Dummy, 1 if a measure of whether individuals lives in an urban or rural area is included, 0 otherwise	0,25	0,43

Table 4.A18: GENDER - Description and summary statistics of explanatory variables

Variable	Definition	Mean	St. Dev.
TSTAT	Estimated t-statistic of the effect size	0,07	3,40
PCC	Partial correlation coefficient	-0,01	0,06
Precision	Precision of the estimated partial correlation coefficient (the inverse of the standard error)	114,49	129,98
Number of variables	Logarithm of number of explanatory variables	2,07	0,48
Original data	Dummy, 1 if original data used, 0 otherwise	0,26	0,44
<i>Measure of attitudes</i>			
Migration policy	Dummy, 1 if attitudes toward the number of immigrants are used, 0 otherwise	0,47	0,50
Mixed attitudes	Dummy, 1 if attitudes toward both the number and contribution of immigrants are used, 0 otherwise	0,01	0,11
Contribution of immigrants (base cat.)	Dummy, 1 if attitudes toward the contribution of immigrants are used, 0 otherwise	0,52	0,50
<i>Control variables - factors</i>			
Age	Dummy, 1 if age is included, 0 otherwise	0,80	0,40
Education	Dummy, 1 if education is included, 0 otherwise	0,88	0,32
Gender	Dummy, 1 if gender is included, 0 otherwise	1,00	0,00
High-skill occupation	Dummy, 1 if the occupation is included, 0 otherwise	0,12	0,32
Income	Dummy, 1 if income is included, 0 otherwise	0,42	0,50
Minority status	Dummy, 1 if minority status is included, 0 otherwise	0,40	0,49
Unemployed	Dummy, 1 if employment status is included, 0 otherwise	0,47	0,50
Urban residence	Dummy, 1 if a measure of whether individuals live in an urban or rural area is included, 0 otherwise	0,25	0,43

Table 4.A19: HIGH-SKILL OCCUPATION - Description and summary statistics of explanatory variables

Variable	Definition	Mean	St. Dev.
TSTAT	Estimated t-statistic of the effect size	2,87	4,02
PCC	Partial correlation coefficient	0,03	0,03
Precision	Precision of the estimated partial correlation coefficient (the inverse of the standard error)	116,90	88,58
Number of variables	Logarithm of number of explanatory variables	2,64	0,36
Original data	Dummy, 1 if original data used, 0 otherwise	0,06	0,24
<i>Measure of attitudes</i>			
Migration policy	Dummy, 1 if attitudes toward the number of immigrants are used, 0 otherwise	0,07	0,26
Mixed attitudes	Dummy, 1 if attitudes toward both the number and contribution of immigrants are used, 0 otherwise	0,00	0,00
Contribution of immigrants (base cat.)	Dummy, 1 if attitudes toward the contribution of immigrants are used, 0 otherwise	0,93	0,26
<i>Control variables - factors</i>			
Age	Dummy, 1 if age is included, 0 otherwise	0,72	0,45
Education	Dummy, 1 if education is included, 0 otherwise	0,95	0,22
Gender	Dummy, 1 if gender is included, 0 otherwise	0,86	0,35
High-skill occupation	Dummy, 1 if the occupation is included, 0 otherwise	1,00	0,00
Income	Dummy, 1 if income is included, 0 otherwise	0,18	0,39
Minority status	Dummy, 1 if minority status is included, 0 otherwise	0,59	0,49
Unemployed	Dummy, 1 if employment status is included, 0 otherwise	0,60	0,49
Urban residence	Dummy, 1 if a measure of whether individuals live in an urban or rural area is included, 0 otherwise	0,72	0,45

Table 4.A20: INCOME - Description and summary statistics of explanatory variables

Variable	Definition	Mean	St. Dev.
TSTAT	Estimated t-statistic of the effect size	1,79	4,74
PCC	Partial correlation coefficient	0,01	0,04
Precision	Precision of the estimated partial correlation coefficient (the inverse of the standard error)	67,93	93,15
Number of variables	Logarithm of number of explanatory variables	2,28	0,48
Original data	Dummy, 1 if original data used, 0 otherwise	0,21	0,41
<i>Measure of attitudes</i>			
Migration policy	Dummy, 1 if attitudes toward the number of immigrants are used, 0 otherwise	0,62	0,49
Mixed attitudes	Dummy, 1 if attitudes toward both the number and contribution of immigrants are used, 0 otherwise	0,04	0,21
Contribution of immigrants (base cat.)	Dummy, 1 if attitudes toward the contribution of immigrants are used, 0 otherwise	0,34	0,48
<i>Control variables - factors</i>			
Age	Dummy, 1 if age is included, 0 otherwise	0,84	0,37
Education	Dummy, 1 if education is included, 0 otherwise	0,96	0,21
Gender	Dummy, 1 if gender is included, 0 otherwise	0,85	0,36
High-skill occupation	Dummy, 1 if the occupation is included, 0 otherwise	0,08	0,27
Income	Dummy, 1 if income is included, 0 otherwise	1,00	0,00
Minority status	Dummy, 1 if minority status is included, 0 otherwise	0,45	0,50
Unemployed	Dummy, 1 if employment status is included, 0 otherwise	0,54	0,50
Urban residence	Dummy, 1 if a measure of whether individuals live in an urban or rural area is included, 0 otherwise	0,23	0,42

Table 4.A21: MINORITY STATUS - Description and summary statistics of explanatory variables

Variable	Definition	Mean	St. Dev.
TSTAT	Estimated t-statistic of the effect size	1,35	4,25
PCC	Partial correlation coefficient	0,02	0,07
Precision	Precision of the estimated partial correlation coefficient (the inverse of the standard error)	121,55	131,26
Number of variables	Logarithm of number of explanatory variables	2,30	0,40
Original data	Dummy, 1 if original data used, 0 otherwise	0,30	0,46
<i>Measure of attitudes</i>			
Migration policy	Dummy, 1 if attitudes toward the number of immigrants are used, 0 otherwise	0,55	0,50
Mixed attitudes	Dummy, 1 if attitudes toward both the number and contribution of immigrants are used, 0 otherwise	0,00	0,00
Contribution of immigrants (base cat.)	Dummy, 1 if attitudes toward the contribution of immigrants are used, 0 otherwise	0,45	0,50
<i>Control variables - factors</i>			
Age	Dummy, 1 if age is included, 0 otherwise	0,81	0,39
Education	Dummy, 1 if education is included, 0 otherwise	0,89	0,31
Gender	Dummy, 1 if gender is included, 0 otherwise	0,82	0,39
High-skill occupation	Dummy, 1 if the occupation is included, 0 otherwise	0,23	0,42
Income	Dummy, 1 if income is included, 0 otherwise	0,41	0,49
Minority status	Dummy, 1 if minority status is included, 0 otherwise	1,00	0,00
Unemployed	Dummy, 1 if employment status is included, 0 otherwise	0,42	0,50
Urban residence	Dummy, 1 if a measure of whether individuals live in an urban or rural area is included, 0 otherwise	0,27	0,45

Table 4.A22: EMPLOYMENT STATUS - Description and summary statistics of explanatory variables

Variable	Definition	Mean	St. Dev.
TSTAT	Estimated t-statistic of the effect size	-0,95	1,99
PCC	Partial correlation coefficient	-0,01	0,04
Precision	Precision of the estimated partial correlation coefficient (the inverse of the standard error)	104,46	118,58
Number of variables	Logarithm of number of explanatory variables	2,26	0,40
Original data	Dummy, 1 if original data used, 0 otherwise	0,10	0,30
<i>Measure of attitudes</i>			
Migration policy	Dummy, 1 if attitudes toward the number of immigrants are used, 0 otherwise	0,38	0,49
Mixed attitudes	Dummy, 1 if attitudes toward both the number and contribution of immigrants are used, 0 otherwise	0,02	0,14
Contribution of immigrants (base cat.)	Dummy, 1 if attitudes toward the contribution of immigrants are used, 0 otherwise	0,60	0,49
<i>Control variables - factors</i>			
Age	Dummy, 1 if age is included, 0 otherwise	0,82	0,39
Education	Dummy, 1 if education is included, 0 otherwise	0,92	0,28
Gender	Dummy, 1 if gender is included, 0 otherwise	0,92	0,28
High-skill occupation	Dummy, 1 if the occupation is included, 0 otherwise	0,18	0,39
Income	Dummy, 1 if income is included, 0 otherwise	0,52	0,50
Minority status	Dummy, 1 if minority status is included, 0 otherwise	0,51	0,50
Unemployed	Dummy, 1 if employment status is included, 0 otherwise	1,00	0,00
Urban residence	Dummy, 1 if a measure of whether individuals live in an urban or rural area is included, 0 otherwise	0,32	0,47

Table 4.A23: URBAN / RURAL RESIDENCE - Description and summary statistics of explanatory variables

Variable	Definition	Mean	St. Dev.
TSTAT	Estimated t-statistic of the effect size	1,91	3,65
PCC	Partial correlation coefficient	0,02	0,04
Precision	Precision of the estimated partial correlation coefficient (the inverse of the standard error)	107,40	110,49
Number of variables	Logarithm of number of explanatory variables	2,52	0,44
Original data	Dummy, 1 if original data used, 0 otherwise	0,44	0,50
<i>Measure of attitudes</i>			
Migration policy	Dummy, 1 if attitudes toward the number of immigrants are used, 0 otherwise	0,38	0,49
Mixed attitudes	Dummy, 1 if attitudes toward both the number and contribution of immigrants are used, 0 otherwise	0,00	0,00
Contribution of immigrants (base cat.)	Dummy, 1 if attitudes toward the contribution of immigrants are used, 0 otherwise	0,62	0,49
<i>Control variables - factors</i>			
Age	Dummy, 1 if age is included, 0 otherwise	0,80	0,40
Education	Dummy, 1 if education is included, 0 otherwise	0,67	0,47
Gender	Dummy, 1 if gender is included, 0 otherwise	0,72	0,45
High-skill occupation	Dummy, 1 if the occupation is included, 0 otherwise	0,28	0,45
Income	Dummy, 1 if income is included, 0 otherwise	0,35	0,48
Minority status	Dummy, 1 if minority status is included, 0 otherwise	0,70	0,46
Unemployed	Dummy, 1 if employment status is included, 0 otherwise	0,52	0,50
Urban residence	Dummy, 1 if a measure of whether individuals live in an urban or rural area is included, 0 otherwise	1,00	0,00

Chapter 5

Country, Culture or Competition - What Drives Attitudes toward Immigrants in Sub-Saharan Africa?

Malte Becker, Finja Krüger & Tobias Heidland

Abstract

Sub-Saharan Africa is becoming an increasingly important destination for international migration. The region hosts immigrants from other African countries and other parts of the world, such as China. Given high poverty levels and weak social security systems in Sub-Saharan Africa, host populations might fear increasing competition for resources and labor, potentially resulting in negative attitudes toward immigrants. We provide the first systematic causal study of attitudes toward immigrants in Sub-Saharan African countries. Using survey experiments in Uganda and Senegal, we study both attitudes toward immigrants in general and toward specific immigrant groups. In particular, we focus on Chinese immigrants, whose increasing presence in Africa is highly controversial and politicized. We find that attitudes toward immigrants are mainly driven by sociotropic cultural and sociotropic economic concerns. Immigrants from China are seen less positively than the average immigrant despite few differences in perceived economic and cultural contribution.

Acknowledgment: Funding from the Mercator Foundation, project "MEDAM II" (19002900), and the Leibniz Association, project "China in Africa", is acknowledged.

5.1 Introduction

International migration shapes societies, politics and economies (Alesina & Tabellini, 2022) - and Sub-Saharan Africa is no exception. While typically associated with out-migration toward Europe and the MENA region, Sub-Saharan Africa is also an important destination region for international migrants. Africa hosts roughly 10% of international migrants globally, out of which over 20% come from outside the continent. Within the continent, Sub-Saharan Africa attracts the majority of immigrants, with Western and Eastern Africa hosting almost 60% of them (Achieng et al., 2020). Given strong population growth and increasing regional integration in Africa, immigrant populations are projected to grow substantially in the future (OECD, 2021), making Africa the most dynamic region in terms of migration in the next decades.

Understanding public attitudes is important for the successful integration of immigrants, for shaping future migration flows, and for informing migration policy (Facchini et al., 2008; Friebel et al., 2013; Ruhs, 2022). But to date, the scientific literature on attitudes toward immigrants has focused almost exclusively on high-income countries, especially Europe and the USA, see for example Ajzenman et al. (2022), Alesina et al. (2023), Card et al. (2012), Hainmueller and Hopkins (2015), and Valentino et al. (2019). We know very little about how attitudes toward immigrants are determined in Sub-Saharan Africa and other developing countries. As a result, policy-makers from these regions lack quantitative evidence on the impact of migration on the attitudes of their citizens - a gap this paper seeks to close.

For developed countries, research shows that public attitudes are predominantly shaped by sociotropic concerns about immigrants' cultural and economic impacts. Contrary to common belief, egocentric economic concerns such as fear of labor market competition play a less important role (Alesina & Tabellini, 2022). However, the extent to which these results depend on the Western context, with its high average income, strong welfare states, and cultural peculiarities, remains unclear.

In Sub-Saharan Africa, the drivers of public opinion toward immigrants might differ from those in developed countries. Egocentric economic concerns may play a larger role because labor market competition is fiercer and social protection systems are less developed (Berg et al., 2021). By contrast, sociotropic economic and cultural concerns may be less important due to the absence of competition for social services and higher cultural heterogeneity in the respective countries (Alrababa'h et al., 2021).

In this paper, we study attitudes toward immigrants with new experimental and observational survey data from Uganda and Senegal. To our knowledge, we provide the first large-scale causal study of the determinants of attitudes toward immigrants in Sub-Saharan Africa. We focus explicitly on attitudes toward immigrants and exclude attitudes toward refugees from our analysis.

We estimate relative preferences for the characteristics of incoming immigrants using a fully randomized conjoint experiment. Here, respondents select their preferred immigrant profile among different profiles of hypothetical immigrants. Randomly varying the profiles' characteristics allows us to estimate the effect of every specific characteristic on the probability of a profile being selected. Conjoint experiments are routinely employed in the literature on attitudes toward immigrants because they allow causally estimating treatment effects in multidimensional choice settings and are robust to social desirability bias. They have also been validated with real-world behavior (Hainmueller et al., 2014).

Additionally, we take inspiration from Card et al. (2012) and use observational survey data to strengthen the results from the conjoint experiment. We use different indicator questions to measure absolute attitudes toward immigrants and immigration, differentiating between egocentric economic concerns, sociotropic economic concerns, sociotropic cultural concerns, and concerns about foreign influence - which we call power concerns. We then estimate their predictive power for explaining respondents' attitudes toward immigrants.

Furthermore, we differentiate between immigrants in general and Chinese immigrants in particular. Along with China's increasing engagement in Sub-Saharan Africa, Chinese immigrants have recently gained importance in the region (Cervellati et al., 2022). Conservative estimates talk about 1 to 2 Million Chinese immigrants in Sub-Saharan Africa (Hairong, 2020), but actual numbers are likely to be much higher (French, 2015). Official data from the Chinese Ministry of Commerce, likely substantially underestimating the true figures, put the numbers of Chinese immigrants at around 20,000 in Uganda and 8,000 in Senegal (MOFCOM, 2021). The increasing presence of Chinese in Africa is controversial and Chinese immigration is highly salient, provoking policy debate and sometimes even civil unrest (Park, 2022).

While Uganda and Senegal are affected by immigration differently, their labor markets show strong similarities to other countries in Sub-Saharan Africa. Underemployment and informality are widespread. Especially youths suffer from a lack of decent working conditions, and growing working-age populations are projected to further increase competition in the labor market (Berg et al., 2021). Culturally, both countries are highly diverse, yet citizens from different ethnic and cultural groups in these countries have very similar views. And while both countries are 5000 kilometers apart, the determinants of attitudes turn out to be remarkably similar. We, therefore, think that our carefully selected case studies hold substantial external validity for other Sub-Saharan African countries.

Our paper adds to the literature in three ways. First, we investigate whether drivers found to be influential in developed countries also apply to Sub-Saharan Africa. Our experiment is designed such that the measured attitudes regard attitudes toward additional migrants. Thus, our results are not only informative about drivers of attitudes *per se*, but can also inform about acceptance of rising migration that is projected due

to increasing regional integration in Africa and, potentially, European migration policies that affect the African continent.

Second, we extend the three well-established theoretical drivers of public attitudes – sociotropic economic concerns, egocentric economic concerns, and sociotropic cultural concerns – by a fourth one: concerns about foreign influence, which we call power concerns. Power concerns may be particularly relevant in Sub-Saharan Africa due to the region’s history of colonialism and dependence, as contrasted with China’s official non-interference policy (Dreher et al., 2018).

Third, we differentiate between different immigrant groups, focusing especially on attitudes toward Chinese immigrants, a group that has strongly grown in recent years and is perceived controversially by local populations and policymakers alike (Park, 2022).

Our results show that sociotropic concerns, both economic and cultural, are decisive in driving preferences for immigrants. Egocentric economic concerns and power concerns do not seem to be major drivers. In the conjoint experiment, respondents strongly prefer immigrants who are willing to integrate and who have high-paying jobs, irrespective of the respondents’ skill levels. Moreover, respondents prefer African immigrants over non-African immigrants. Our survey measures strengthen the results from the conjoint experiment, underpinning the importance of sociotropic concerns. While Chinese immigrants are viewed somewhat more negatively than the average immigrant, attitudes toward both groups are shaped by the same mechanisms. Power concerns exist but have no explanatory power for attitudes toward immigrants.

Our results are surprisingly homogeneous across multiple subgroups and geographic locations. Similar to findings from the USA and Europe, different respondent subgroups share the same concerns and give importance to the same immigrant characteristics – a hidden consensus (cf. Hainmueller & Hopkins, 2015) that holds within and across the different contexts we study.

The remainder of the paper is structured as follows. In section 5.2, we discuss the relevant theory and existing evidence on which our analysis builds. We then introduce our hypotheses and study design in section 5.3. Section 5.4 covers the estimation strategy before presenting the results in section 5.5. Section 5.6 concludes.

5.2 Drivers of attitudes toward immigrants: theory and evidence

The economic and political science literature distinguishes three main theoretical drivers of attitudes toward immigrants: egocentric economic concerns, sociotropic economic concerns, and sociotropic cultural concerns (see for example Hainmueller and Hopkins (2015))

and Valentino et al. (2019))¹.

Egocentric economic concerns describe the fear of negative consequences for individuals' economic prospects resulting from immigration, for example by competing with immigrants in the labor market (also often referred to as the labor market competition hypothesis). If natives fear potential job losses or a decrease in wages, they should particularly oppose immigrants with a similar set of skills as themselves. Consequently, attitudes should depend on both the natives' and the immigrants' skill-levels and the local labor market conditions.

Sociotropic economic concerns describe concerns about the host economy as a whole, its welfare system, and social services (Alrababa'h et al., 2021). In general, high-skilled immigrants are expected to contribute more to the host country's economy by making higher net contributions to the welfare state than low-skilled immigrants. Accordingly, natives should strictly prefer high-skilled immigrants, i.e. immigrants with a university education or other highly valued degrees, over low-skilled immigrants, irrespective of the natives' skill-level.

Sociotropic cultural concerns capture the perceived threats of immigrants to the host country's culture, its norms, and values. Based on cultural concerns, one expects natives to prefer immigrants who share a similar cultural background over immigrants who are perceived to be culturally distant. Cultural concerns are often proxied using either nationalities (Cogley et al., 2019; Hainmueller & Hopkins, 2015), religion (Alrababa'h et al., 2021) or a combination of both (Bansak et al., 2016).

We introduce a fourth concern which we call power concerns. Power concerns capture citizens' fear that immigrants and their country of origin could gain an out-sized influence in the host country, potentially threatening the host countries' sovereignty. Power concerns are a nuanced form of sociotropic concerns. But instead of looking at attitudes shaped by immigrants' direct impact on the host countries' economy and culture, we intend to capture potential negative attitudes based on foreign countries' government and business activities that are then projected on individual immigrants. Foreign governments and international relations have been shown to shape public opinion and attitudes toward migrants in Western contexts (Balmas, 2018; Goldsmith et al., 2021; Matush, 2023). Anecdotal evidence from Sub-Saharan Africa goes in a similar direction. For example, in early 2021 French businesses in Dakar were attacked during protests about the economic situation and the French influence in Senegal (Thomas-Johnson, 2021). Several studies, for example by Iacoella et al. (2021) and Sardoschau and Jarotschkin (2019), linking Chinese projects to political participation and conflict, point toward similar concerns about foreign influence in Sub-Saharan Africa.

¹Studies focusing on attitudes toward refugees in particular additionally investigate humanitarian concerns, see for example Bansak et al. (2016) and Alrababa'h et al. (2021).

Recent studies from Europe and the USA broadly agree that sociotropic concerns, both economic and cultural, are the key drivers of public attitudes toward immigrants (Adida et al., 2019; Bansak et al., 2016; Card et al., 2012; Denney & Green, 2021; Haaland & Roth, 2020; Hainmueller & Hopkins, 2014, 2015; Valentino et al., 2019). In these studies, respondents consistently favor skilled and educated immigrants, irrespective of the respondents' qualification or education. Respondents also prefer immigrants who are perceived to be culturally similar over immigrants that are perceived to be culturally distant. Egocentric economic concerns are of little relevance in Western societies: While earlier studies stress the importance of potential labor market competition (Mayda, 2006; Scheve & Slaughter, 2001), more recent experimental studies find little to no support for a bias against immigrants with a similar set of skills (Hainmueller et al., 2015; Hainmueller & Hopkins, 2014; Hainmueller et al., 2014; Valentino et al., 2019). These findings are in line with literature that estimates the economic and wage effects of immigrants on host countries to be small in general (Dustmann et al., 2016).

Although developing countries receive the majority of immigrants worldwide, quantitative studies on attitudes toward immigrants in developing countries are scarce. Most studies are purely observational (e.g., Adida, 2011; Buehler & Han, 2018; Facchini et al., 2013; Ghosn et al., 2019; Gordon, 2018; Harris et al., 2018; Hartman & Morse, 2020; Ruedin, 2019; Whitaker & Giersch, 2015) and lack causal identification.

An emerging literature investigates attitudes toward *refugees* in low and middle income countries, see for example Aksoy and Ginn (2022), Alrababa'h et al. (2021), Betts et al. (2023), Lehmann and Masterson (2020), and Zhou et al. (2023). Overall, the effects of refugees on public attitudes are limited, with little support for a public backlash which is often observed in high-income countries (Dustmann et al., 2019). But attitudes toward forced migrants typically differ from attitudes toward other migrants, who typically migrate for economic reasons. The extant literature on migrants in developing countries who are not forced is very limited.

The study perhaps closest to ours is by Cogley et al. (2019) and investigates attitudes toward the naturalization of immigrants in the Ivory Coast using a vignette-experiment. Respondents were presented with single randomly composed profiles of hypothetical immigrants that were assumed to already reside in the Ivory Coast. Respondents were then asked to decide whether a given immigrant profile should be naturalized, tolerated, or deported. The authors find that immigrants' legal status, education, and employment play an important role in respondents' decisions. Cultural concerns, measured by religion, language skills, and origin country have mixed effects on respondents' attitudes and do not provide clear results. Despite some overlap, our study differs from Cogley et al. (2019) in several dimensions. Their main focus lies on political and legal determinants of attitudes, they focus on the naturalization of immigrants, and their method differs.

There is hardly any quantitative evidence that causally investigates the effects of Chinese migration on public attitudes in Africa. Even the actual number of Chinese immigrants on the continent is unclear as no official figures exist. Guesstimates talk about 1 to 2 million Chinese immigrants in Africa (Hairong, 2020)², but actual numbers are likely a multiple of that (Botchwey et al., 2019; French, 2015). Park (2022) gives a brief overview of the history of Chinese migration to Africa and provides anecdotal evidence for increasing anti-China sentiment fueled by a mixture of geopolitics, local rumors, negative (Western) media coverage, and economic factors.

An emerging literature investigates the effect of Chinese aid on public opinion toward China in Africa, with mixed and partially conflicting results (Blair et al., 2022; Mandon & Woldemichael, 2023; McCauley et al., 2022; Wegenast et al., 2019; Xu & Zhang, 2020). Another branch of the literature investigates the effects of Chinese aid on conflict, which is an important manifestation of public opinion. Again, the results are mixed. While Gehring et al. (2022) do not find any relationship between Chinese aid projects and conflict in Africa, Iacoella et al. (2021) find that proximity to Chinese projects is associated with an increase in civil unrest and political participation, potentially triggered by negative perceptions about China's influence in the region and loss of trust in the local government. Similarly, Sardoschau and Jarotschkin (2019) link conflict in Sub-Saharan Africa to Chinese development projects. Interestingly, they find that the relationship is driven by projects that do not involve the immigration of Chinese workers. Rather, it is projects involving rent-seeking that drive their results.

To our knowledge, attitudes toward Chinese immigrants have been quantitatively investigated only once before in Africa. When studying attitudes toward naturalization in Ivory Coast, Cogley et al. (2019) include immigrants' origin as a dimension in their vignette experiment, one of the origin countries being China. Contrary to anecdotal evidence and common belief, they find that Chinese immigrants were significantly more likely to be recommended for naturalization than immigrants from other countries. However, the positive effect for Chinese immigrants does not hold when looking at deportations.

5.3 Hypotheses, data & study design

5.3.1 Hypotheses

We test whether the three main theoretical drivers of attitudes toward immigrants - egocentric economic concerns, sociotropic economic concerns, and sociotropic cultural concerns - are relevant in the Sub-Saharan African context, and add power concerns as a potential fourth driver.

²Proxies for the presence of Chinese workers in Africa are, for example, the number of Weibo users, a Chinese Twitter equivalent (Cervellati et al., 2022).

Building on the aforementioned literature, we test the following working hypotheses: (1) egocentric economic concerns are important drivers of attitudes toward immigrants in Sub-Saharan Africa, due to high poverty rates, weak labor markets, and large working-age populations; (2) sociotropic economic concerns are less important drivers, e.g., because people are more self-reliant due to the absence of public services and income support in many areas, decreasing competition for public services; (3) sociotropic cultural concerns are less important drivers, e.g., due to host countries' high degree of cultural diversity or relatively higher importance of economic concerns; (4) power concerns are important drivers of immigration attitudes in Sub-Saharan Africa, where development cooperation is often highly politicized; (5) attitudes toward Chinese immigrants are more critical than attitudes toward immigrants in general.

All hypotheses have been pre-registered before the data collection.³

5.3.2 Sampling and data collection

To investigate our hypotheses, we collected original individual-level survey data in Uganda and Senegal. While we are aware that these two countries cannot speak for the whole of Sub-Saharan Africa, they provide two very interesting case studies from two important regions and show many similarities to other countries in Sub-Saharan Africa. Table 5.1 summarizes different development indicators for Uganda, Senegal, and Sub-Saharan Africa.

Sub-Saharan Africa is characterized by relatively high poverty levels and a lack of social protection. As a consequence, people cannot afford to be unemployed. Underemployment and informality, however, are the norm. Especially youths suffer from a lack of decent working conditions, and growing working-age populations are projected to further increase competition in the labor market (Berg et al., 2021). Similarly, unemployment in Uganda and Senegal is negligible, but self-employment and informality are high: According to World Bank (2022), 73% of employment in Uganda and 63% of employment in Senegal are vulnerable - and might thus be particularly susceptible to competition caused by immigration.

Detailed up-to-date data on immigration in Uganda and Senegal does not exist. Most recent numbers from 2015 estimate the migrant stock at 1.9% in Uganda and 1.7% in Senegal (World Bank, 2022). There is also no systematic information on immigrants' occupations in Uganda and Senegal. Anecdotal evidence and information provided by our local partners suggest that immigrants from neighboring countries rather work in informal low-skilled jobs, such as agriculture, construction, or retail (Merotto, 2020). Non-African immigrants tend to work in high-skilled jobs (Leichtman, 2005). Especially in the case of Chinese immigrants, lower-skilled occupations are also common. Park

³The pre-analysis plan has been submitted to OSF on 08 October 2021.

(2022) states construction, small-scale entrepreneurship, and agriculture as common jobs among Chinese immigrants.

China’s increasing influence in Sub-Saharan Africa is evident in both Uganda and Senegal. The two countries have received large Chinese infrastructure projects, loans, and development aid. Custer et al. (2021) recorded 144 Chinese projects in Uganda and 83 Chinese projects in Senegal between 2000 and 2017. Chinese projects in Sub-Saharan Africa typically cause sizeable immigration of Chinese workers (Cervellati et al., 2022). According to official figures from the Chinese Ministry of Commerce, Uganda, and Senegal are home to around 20,000 and 8,000 Chinese immigrants, respectively (MOFCOM, 2021). The actual numbers, however, are likely to be much higher, as Chinese immigration in Sub-Saharan Africa is often irregular (Botchwey et al., 2019; French, 2015). For more details, a brief history, and anecdotal evidence on immigration in Uganda and Senegal, see Appendix 5.A.1.

Table 5.1: Country characteristics

	Uganda	Senegal	SSA
GDP per capita, PPP (constant 2017 international \$) ¹	2183	3430	3764
Population 18-40 (%) ¹	29	30	30
Immigrant population % (2015) ¹	1.9	1.7	1.9
Unemployment % (2020) ¹	2.8	3.6	7.3
Unemployment youth % (2020) ¹	4.0	4.9	14.5
Vulnerable employment % (2019) ¹	73.1	62.8	73.1
Number of Chinese projects ²	144	83	5995

¹ Source: World Bank (2022)

² Source: Custer et al. (2021)

To investigate public attitudes toward immigrants in Uganda and Senegal, we collected detailed primary data from 2,704 individuals. In Uganda, we surveyed 1,204 individuals between October and November 2021. In Senegal, we surveyed 1,500 individuals between February and April 2022. Together with our local partners, we identified suitable enumeration areas from different regions to conduct the surveys. In Uganda, we interviewed individuals from 36 enumeration areas in the districts Kampala, Mbale, Gulu, Wakiso, Masaka, and Mbarara (Figure 5.A1). In Senegal, we interviewed 1,500 individuals in 60 different enumeration areas in the districts Dakar, Diourbel, Matam, Saint Louis, Tambacounda, and Ziguinchor (Figure 5.A2).

Each sample aims at being representative of the respective country’s young and mobile population. After a complete household listing in each enumeration area, we randomly sampled individuals aged between 18 and 40 years for the main interview ⁴. In both

⁴Our study is part of a larger survey mainly investigating migration aspirations and decisions. For

Uganda and Senegal, this age group accounts for approximately 30% of the countries' population (World Bank, 2022). The response rate for the primary interview in Uganda was a remarkable 99%. In Senegal, we had a response rate of 75%. Including our experiment, each survey took roughly between 90 to 120 minutes. In addition to the conjoint tasks, we collected information on household and individual-level socioeconomic characteristics, individual migration intentions, and attitudes toward immigrants⁵. Note that respondents completed the conjoint experiment before answering the survey questions. This order minimizes priming in two ways: First, by starting with the experiment, we can rule out that the survey questions have primed responses in the conjoint experiment, which is the central part of the survey. Second, respondents' answers to the subsequent survey questions can be systematically primed by previous treatments. However, such systematic priming can be ruled out in our experiment because profiles in the conjoint experiment were randomly composed.

5.3.3 Sample characteristics

Table 5.2 summarizes the socio-demographic characteristics of our full sample. The average respondent is 28 years old, female, and has no secondary education. 49% of respondents engaged in income-generating activities during the last 7 days before the interview. 26% were not working but wanted to, and 11% were studying. Under ideal circumstances, 67% of our sample would like to move within their country, and 71% would like to migrate internationally if given the chance. 49% of the respondents reported having contact with any foreigner at least once a month, and 34% have contact with a person from China at least once a month - for example on public transport, in the street, in shops, or in their neighborhood.

5.4 Empirical approach

5.4.1 Conjoint experiment

The central part of our survey is a conjoint experiment, which allows us to causally assess the relative importance of immigrants' characteristics for respondents' preferences toward potential incoming immigrants. The conjoint experiment is fully randomized, consisting of three choice tasks per respondent. Each time, respondents were asked to compare two

that purpose, the data collection was targeted to a young and mobile population.

⁵A conjoint experiment relies on randomization of the different attribute levels. Due to errors in the survey program, this randomization did not work correctly for the first 500 experiments we conducted in Uganda, making the data invalid. We tried to re-visit all of these 500 respondents to repeat the conjoint experiment, but 33 respondents (6.6% of the 500 respondents to be re-visited) had to be replaced, either because they refused to re-do parts of the questionnaire ($N = 4$) or because they could not be re-contacted ($N = 29$).

Table 5.2: Respondents' characteristics

Variable	Senegal		Uganda		Total	
	N	Mean/SE	N	Mean/SE	N	Mean/SE
Age	1500	28.480 (0.178)	1204	26.743 (0.180)	2704	27.707 (0.128)
Female	1500	0.538 (0.013)	1204	0.627 (0.014)	2704	0.578 (0.010)
Low skilled (no sec. education)	1500	0.621 (0.013)	1204	0.264 (0.013)	2704	0.462 (0.010)
Working	1500	0.488 (0.013)	1204	0.491 (0.014)	2704	0.489 (0.010)
Not working, not wanting to	1500	0.029 (0.004)	1204	0.027 (0.005)	2704	0.028 (0.003)
Not working, wanting to	1500	0.262 (0.011)	1204	0.247 (0.012)	2704	0.255 (0.008)
Studying (university or school)	1500	0.137 (0.009)	1204	0.067 (0.007)	2704	0.106 (0.006)
Would like to move internally	1487	0.516 (0.013)	1201	0.867 (0.010)	2688	0.673 (0.009)
Would like to move internationally	1482	0.626 (0.013)	1204	0.820 (0.011)	2686	0.713 (0.009)
Contact foreigner at least once a month	1468	0.441 (0.013)	1193	0.559 (0.014)	2661	0.494 (0.010)
Contact Chinese at least once a month	1490	0.225 (0.011)	1195	0.481 (0.014)	2685	0.339 (0.009)
Rural	1500	0.449 (0.013)	1204	0.290 (0.013)	2704	0.378 (0.009)

Note: Don't know and Refused to answer are coded as missing.

profiles of hypothetical immigrants and choose the one they would prefer to come to their country⁶. That implies that we are measuring relative preferences for immigrants at the margin, i.e., for one additional immigrant. Each immigrant profile consists of four dimensions: *immigrant's job*, *immigrant's origin country/region*, *immigrant's willingness to integrate into the host country's society* and *immigrant's location within the host country*. Each dimension takes on a specific attribute level that is randomly drawn from a pool of pre-defined attribute levels. The two profiles to be compared in each task differ in at least one dimension. We do not restrict combinations of specific levels. We randomized the order of dimensions across respondents to avoid listing effects but held it constant across tasks per respondent. A total of 2,704 respondents, three tasks per respondent, and two profiles per task give us an effective sample size of $N = 16,224$ for the conjoint experiment.

The different immigrant profiles are presented to the respondent in a visualized form while being read out and explained by the enumerator. We co-developed the attribute levels and visualizations in close collaboration with local experts to ensure they suit the local context. Table 5.3 shows the conjoint experiment's different dimensions and attribute levels. Attribute levels for Uganda and Senegal are almost identical and differ only in one level of the origin dimension. In Uganda, we included Indian immigrants, while in Senegal, we used Lebanese immigrants. Indians have a long-standing and controversial history of immigration in Uganda, as do Lebanese people in Senegal (Park, 2022).

We explicitly designed the conjoint experiment to investigate the different potential drivers of attitudes toward immigrants. First, suppose sociotropic cultural concerns are important in Sub-Saharan Africa. In that case, one should expect respondents to prefer African immigrants from a similar cultural background and immigrants willing to integrate into the local society. To reduce the complexity of the conjoint experiment and conferring with our local partners, we do not differentiate between specific origin countries within Africa. Notably, the conjoint explicitly refers to African immigrants, not African refugees. We exclude refugees because certain combinations of attribute levels, such as refugees that work as investors or supermarket owners, would be unrealistic in the local setting.

Second, if sociotropic economic concerns are salient, we can expect respondents to strictly prefer high-skilled immigrants, namely investors and supermarket owners, over low-skilled immigrants, namely small shop owners and construction workers - irrespective of the respondents' characteristics. If, however, egocentric economic concerns are important, we expect respondents to prefer immigrants with whom they would not compete in the labor market. For example, we would expect higher-skilled respondents to prefer lower-skilled immigrants (small shop owners or construction workers) and lower-skilled

⁶Respondents also had the option to select "don't know" or "refuse to answer". In Uganda, this was never the case. In Senegal, it happened 98 times (about 1% of all conjoint decisions in Senegal).

Table 5.3: Conjoint experiment: dimensions and attribute levels

Dimension	Attribute Levels	Visualization
Job	• Construction worker	
	• Investor	
	• Small shop owner	
	• Supermarket owner	
Location in Uganda	• Within community	
	• Outside of community	
Willingness to integrate	• Willing	
	• Not willing	
Origin	• Africa	
	• China	
	• Europe	
	• India	
	• Lebanon	

Note: Dimensions, attribute levels, and visualizations for the conjoint experiment in Uganda. We replaced the origin attribute level "India" with "Lebanon" in the Senegalese localization of the experiment. Attribute level "Africa" explicitly excludes refugees.

respondents to prefer higher-skilled immigrants (investors or supermarket owners).

To proxy potential labor market competition, we follow the literature and focus on respondents' educational attainment (Hainmueller & Hopkins, 2015), employment status (Alrababa'h et al., 2021) and income levels (Bansak et al., 2016). Given the low levels of education in the local context, we use "some secondary education" as the cutoff to differentiate between higher and lower skilled individuals. To capture respondents' employment status, we rely on respondents' main job during the last seven days and group respondents into working; unemployed and wanting to work; unemployed and not wanting to work; and studying. To distinguish between different income levels, we create country-specific household income per capita quintiles for our sample, relying on the reported household income for the last 12 months.

Next, the immigrants' location allows us to capture respondents' overall evaluation of immigrants. Kadigo et al. (2022) and Zhou et al. (2023) find that proximity to refugee settlements in Uganda is associated with benefits for the host population, such as an increase in local households' welfare or improved access to public goods. If respondents indeed evaluate immigration as beneficial, we would expect them to prefer immigrants that live close by. Conversely, opposing immigration should lead to a "not in my backyard" mentality, driven by concerns about negative externalities in respondents' regions (Cogley et al., 2019).

Finally, we combine the conjoint experiment with observational survey data to estimate the importance of power concerns. Using the survey data, we create a dummy for respondents who think foreign governments and businesses have too much influence in their country. Based on this dummy, we conduct subgroup analysis in the conjoint experiment to investigate whether power concerns influence respondents' preferences for specific immigrant characteristics.

Among the experimental dimensions shown in Table 5.3, attribute levels for origin, job, and location are rather self-explanatory. Willingness to integrate, however, can leave some room for interpretation. It could also capture migrants' willingness to integrate into society economically or willingness to obey rules and laws. To avoid misunderstandings, we explicitly trained examples for integration, such as learning the local languages, participating in local celebrations, and making local friends. This way, we ensured that explanations and paraphrasing aligned with what we intended to capture.

Conjoint experiments have become a popular tool for studying attitudes and preferences in political and economic contexts. However, they come with specific disadvantages that must be addressed in the research design and the interpretation of results. First, by construction, conjoint experiments allow studying the causal effect of specific immigrant characteristics *relative* to other characteristics. The identification of relative preferences, for example, whether respondents prefer immigrants from a certain origin over immigrants from another origin, needs to be accompanied by survey questions that inform about gen-

eral attitudes toward immigrants, for example, whether respondents prefer more or less immigration of certain nationalities in absolute terms. Second, conjoint experiments make the comparison of estimated effect sizes across different dimensions complex if each dimension follows a different scale. As we include categorical values such as specific occupations or origins, we cannot directly compare the share of variance explained by each dimension. Third, a common concern is that the attribute levels included in a conjoint experiment capture only some of the population’s concerns and potentially miss other important profile characteristics. Ultimately, a conjoint experiment has to choose the most relevant dimensions to remain tractable for respondents. We sought to address this challenge by closely working with local experts when designing and picking the most relevant conjoint dimensions - both from an academic and a local-societal perspective. Fourth, the concrete wording of dimensions needs to be well-designed to capture the intended underlying aspect. In our study, this difficulty is most apparent in the willingness to integrate, which measures a large share of sociotropic cultural concerns but likely cannot capture them entirely. The hypothetical immigrants’ nationality likely captures a mix of cultural concerns and other aspects, such as group-based xenophobia or racism. This speaks to the broader point that readers should always critically reflect on what conjoint dimensions actually capture.

To evaluate the conjoint experiment, we estimate marginal means (MMs), following Leeper et al. (2020). In a forced conjoint experiment, the marginal mean of a given attribute level represents the average effect of that specific level on the overall probability of the profile being preferred, relative to a zero effect of 0.5 (randomly choosing one of two profiles), averaging across all respondents and all other levels (Leeper et al., 2020). Marginal means above 0.5 indicate a positive effect of an attribute level on the probability of a profile being selected, while marginal means below 0.5 indicate a negative effect. For example, a marginal mean of 0.6 for a given attribute level indicates that profiles with that attribute level are selected with a probability of 60%, on average.

We estimate marginal means corresponding to the following regression equation

$$Y_{ijk} = \beta_0 + \beta_1 Job_{ijk} + \beta_2 Location_{ijk} + \beta_3 Integration_{ijk} + \beta_4 Origin_{ijk} + \epsilon_{ijk} \quad (5.1)$$

where Y_{ijk} denotes the choice of respondent i in task j for profile k , β_1 captures marginal means of the immigrant’s jobs, β_2 captures marginal means of the immigrant’s locations in the host country, β_3 captures marginal means of the immigrant’s willingness to integrate, and β_4 captures marginal means of the immigrant’s origin. For subgroup analysis, we add interactions between each attribute level and categorical variables defin-

ing the specific subgroups ⁷.

One might be concerned that respondents associate immigrants' origin with other specific stereotypes. For example, respondents might associate Indian immigrants with high-paying jobs while associating African immigrants with low-paying jobs. Though, we estimate the marginal means (or AMCE) of each specific profile characteristic automatically controlling for all other dimensions. This implies that we can account for potential stereotypes that are captured by attribute levels included in the conjoint experiment. However, stereotypes that are not included in the conjoint experiment can still affect the coefficient for a given nationality. When interpreting the results, it is important to note that we sought to include major local stereotypes to avoid that they drive the experimental estimates. For example, while stereotypes about criminality might be salient in Europe, they do not play a role in our study areas for the chosen immigrant groups.

For all estimations, we cluster standard errors on the respondent level. Following Hainmueller et al. (2014), we test for carryover effects, profile order effects, attribute order effects, interactions between different attribute levels, and balance.

5.4.2 Observational data

Observational survey data can help us strengthen the conjoint experiment's results. To measure attitudes toward immigrants using observational data, we build on survey questions from the European Social Survey, adapting them to the local context.⁸ The outcome of interest, namely absolute attitudes toward immigrants in general and toward Chinese immigrants in particular, is measured by the following two questions:

- To what extent do you think should Uganda (Senegal) allow people from other countries to come and live here?
- To what extent do you think should Uganda (Senegal) allow people from China to come and live here?

⁷We also estimate Average Marginal Component Effects (AMCEs) for equation 5.1, following Hainmueller et al. (2014). AMCEs estimate causal differences in preferences for a given characteristic relative to a baseline attribute level rather than relative to a probability of 0.5.

⁸The ESS differentiates between different types of immigrants: same/different ethnicity to that of the majority community; from richer/poorer countries; from inside/outside of Europe. Our questionnaire differentiates between immigrants in general and Chinese immigrants in particular. The following two ESS questions measure sociotropic economic concerns: *Would you say it is generally bad or good for [country]'s economy that people come to live here from other countries?* and *Would you say that people who come to live here generally take jobs away from workers in [country], or generally help to create new jobs?* Egocentric economic concerns are not captured in the ESS. Sociotropic cultural concerns are measured by the ESS question *Would you say that [country]'s cultural life is generally undermined or enriched by people coming to live here from other countries?* We focused on the potential threat of immigrants following Sniderman et al. (2004): *"These days, I am afraid that the Dutch culture is threatened by ethnic minorities."* The ESS does not assess power concerns.

We then follow Card et al. (2012) and ask respondents different indicator questions to distinguish between egocentric economic, sociotropic economic, and sociotropic cultural concerns. We furthermore assess power concerns as potential drivers of attitudes toward (Chinese) immigrants. Answers to the outcome and indicator questions are measured using different Likert scales. Table 5.4 gives an overview of the outcome and indicator variables, including the corresponding questions. The outcome variable asks about the preferred level of immigration and can thus be interpreted as a policy preference. With the analysis, we can study how beliefs about different effects of immigration influence the overall attitude toward immigration, complementing the relative preferences toward individuals from the conjoint experiment.

Table 5.4: Survey questions for observational analysis

	Question	Answer Options
Outcome questions	To what extent do you think should Uganda (Senegal) allow people from other countries to come and live here?	Allow none Allow a small number Allow some Allow many
	To what extent do you think should Uganda (Senegal) allow people from China to come and live here?	Allow none Allow a small number Allow some Allow many
Sociotropic economic concerns	Would you say it is generally good or bad for Uganda's (Senegal's) economy that people from other countries come to live here?	Very bad Somewhat bad Neither bad nor good Somewhat good Very good
	Would you say it is generally good or bad for Uganda's (Senegal's) economy that people from China come to live here?	Very bad Somewhat bad Neither bad nor good Somewhat good Very good
	Do foreigners who come to live here take jobs or help to create new ones?	0 (take away jobs) to 10 (help to create new ones)
	Do Chinese who come to live here take jobs or help to create new ones?	0 (take away jobs) to 10 (help to create new ones)
Egocentric economic	Would you say it is generally good or bad for	Very bad Somewhat bad

Continued on next page

Table 5.4 continued from previous page

	Question	Answer Options
concerns	you and your household's economic prospects that people from other countries come to live here?	Neither bad nor good Somewhat good Very good
	Would you say it is generally good or bad for you and your household's economic prospects that people from China come to live here?	Very bad Somewhat bad Neither bad nor good Somewhat good Very good
Sociotropic cultural concerns	How much do you agree with the following statement: Countries' norms and values are being threatened by people from other countries who come to live here?	0 (not at all) to 10 (a lot)
	How much do you agree with the following statement: Countries' norms and values are being threatened by people from China who come to live here?	0 (not at all) to 10 (a lot)
Power concerns	Do you think that foreign governments have too much or too little little influence in Uganda (Senegal)	0 (too little) to 10 (too much)
	Do you think that foreign businesses have too much or too little little influence in Uganda (Senegal)	0 (too little) to 10 (too much)
	Do you think that Chinese governments have too much or too little little influence in Uganda (Senegal)	0 (too little) to 10 (too much)
	Do you think that Chinese businesses have too much or too little little influence in Uganda (Senegal)	0 (too little) to 10 (too much)

Following the intuition provided in Card et al. (2012), we estimate the predictive power of different concerns for explaining attitudes toward immigrants. We assume that

respondent i 's answers to our outcome questions depend on respondent i 's concerns about (Chinese) immigrants' effects on i 's individual economic situation, on the economic situation of i 's home country, on the norms and values of i 's home country, as well as i 's concerns about the influence of foreign governments and businesses. While the true concerns are unobserved, we use respondents' answers to the indicator questions as proxies.

We then run the following OLS regression using the variables from Table 5.4:

$$Y_i = \beta_0 + \beta_1 \text{EgocentricEconomic}_i + \beta_2 \text{SociotropicEconomic}_i + \beta_3 \text{SociotropicJob}_i + \beta_4 \text{SociotropicCultural}_i + \beta_5 \text{PowerGovernment}_i + \beta_6 \text{PowerBusiness}_i + \beta_7 \text{Controls}_i + \epsilon_{ijk} \quad (5.2)$$

where Y_i represents respondent i 's opinion on the number of people from other countries that should be allowed to come and live in the host country; β_1 captures egocentric economic concerns; β_2 captures concerns about the expected effect of immigrants on the economy as a whole; β_3 estimates concerns about the expected effect of immigration on the labor market; β_4 captures sociotropic cultural concerns; and β_5 and β_6 capture power concerns related to foreign business and foreign governments respectively. β_7 captures the effect of a wide range of respondent-level control variables, namely age, gender, education, self-reported employment during the last 7 days (4 levels), location (rural-urban), self-reported contact with immigrants, country-specific quintiles for self-reported household income per capita during the last 12 months, life satisfaction (0-10), satisfaction with city/area of residence (1-5), and country fixed effects. Outcome and indicator variables are standardized to make estimates comparable across different answer scales, standard errors are clustered at the district level. Running the analysis both for attitudes toward immigrants in general and toward Chinese immigrants in particular, we investigate whether the relative importance of the concerns differs between the two immigrant groups.

Some of the differences between absolute attitudes in the survey section and relative attitudes in the conjoint experiment may be driven by respondents' individual beliefs about certain migrant origins. While we are mainly able to control for that in the conjoint experiment, we cannot do so in the survey section. For example, if Chinese immigrants are not expected to integrate well, that would result in more negative attitudes in the survey question compared to the experimental estimate.

5.5 Results

What drives relative attitudes toward immigrants in Sub-Saharan African countries?

Figure 5.1 plots the marginal means corresponding to equation 5.1. The dots indicate point estimates for the marginal mean of each attribute level. Whiskers around the dots show the 95% confidence intervals. The solid vertical line at the x-value of 0.5 represents the reference effect relative to which the marginal means are estimated. Intuitively, if a given attribute level does not affect the probability of a profile being selected, a profile with this attribute level should, on average, be selected with a probability of 0.5, holding all other attribute levels constant - like flipping a coin. Thus, the line helps to assess whether an effect is statistically different from zero at the 95% confidence level. The AMCE equivalent can be found in the appendix in Table 5.A2 and Figure 5.A4.

Economic versus cultural preferences

We start our analysis by focusing on immigrants' economic role in society and assess the effect of a profile's occupation on relative preferences. The top panel in Figure 5.1 shows that, on average, respondents prefer immigrants with occupations that are high-paying and have the potential to create additional jobs: Profiles including investors or supermarket owners are selected with a probability of 0.65 and 0.56, respectively. Low-skilled jobs have a negative effect on the probability of a profile being selected: marginal means for small shop owners and construction workers are 0.38 and 0.41, respectively. Occupational differences matter, indicating that economic concerns play an important role.

The second panel shows that contrary to the dominant views in many developed countries, our respondents prefer immigrants to move nearby. The marginal mean for immigrants who will stay within the respondents' community is 0.54, as compared to 0.46 for immigrants who will stay outside the respondents' community. We interpret this as evidence for an overall positive evaluation of immigrants, at least for those immigrants that are well-proxied by the characteristics included in our experiment: Respondents rather want the average immigrants to be their neighbors instead of them living elsewhere. That is in line with results from Cogley et al. (2019), who do not find any evidence for a not-in-my-backyard mentality, and with Kadigo et al. (2022), who find positive effects of proximity to refugee settlements on households' welfare.

As the third panel shows, respondents want these new neighbors to integrate socially. Immigrants' willingness to integrate - for example, to learn the local language, make local friends, and participate in cultural events - strongly affects a profile's probability of being preferred. Profiles with immigrants that are willing to integrate are selected with

a probability of 0.67, as compared to 0.32 for their non-willing counterparts. This large effect is in line with anecdotal evidence from conversations with people in the field who emphasized the importance of immigrants' integration. The particularly sizable effect underlines the importance of cultural concerns in the host population.

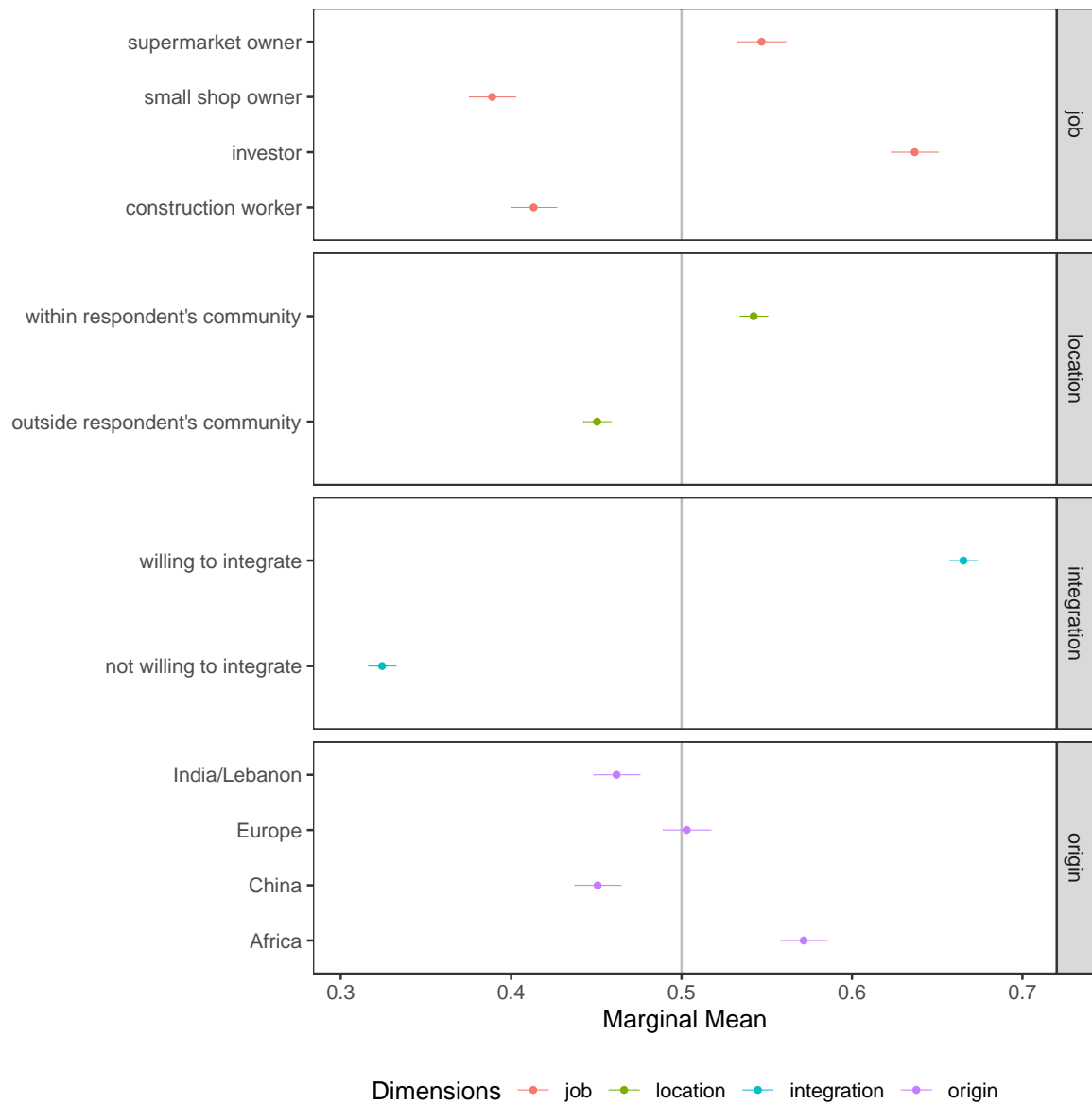
The fourth panel assesses another important part of the cultural dimension by contrasting attitudes toward different nationalities. Respondents significantly prefer fellow Africans (marginal mean of 0.56), whereas immigrants from India/Lebanon and China decrease the probability of a profile being selected (marginal means of 0.46 for both groups). The effect for immigrants from Europe is not significantly different from the null effect of 0.5. While respondents do penalize immigrants of Chinese origin, they do *not* dislike them more than immigrants from India/Lebanon, who have a much longer history of immigration in the respective countries. European immigrants are seen more negatively than African immigrants but significantly more positively than immigrants from China or India/Lebanon. Given that we do not differentiate between African countries or ethnicity, we likely underestimate the importance of the origin dimension. Plausibly, there will be differences in attitudes toward culturally more similar and culturally more distant African countries, which we cannot observe due to the limited number of attribute levels per dimension we could include in our experiment.

Egocentric versus sociotropic economic preferences

In the previous section, we established that economic concerns are an important driver of relative attitudes toward immigrants. But are these economic concerns sociotropic or egocentric? To answer this question, we follow Hainmueller and Hopkins (2015), Alrababa'h et al. (2021) and Bansak et al. (2016) and estimate marginal means for different respondent sub-groups, some of whom are more likely to compete with the immigrants in our experiment. We differentiate between (i) higher- and lower-skilled respondents (Figure 5.A5), (ii) different employment statuses (Figure 5.A6), and (iii) different household income per capita levels (Figure 5.A7).⁹ Furthermore, we interact respondents' skill levels with the immigrant's job and the immigrant's location in the host country since competition will occur where workers are substitutable. Respondents might only fear competition with similarly skilled immigrants who also live inside their community and not with those

⁹Given that education levels, employment status, and household income are very broad measures of whether respondents compete with the potential immigrants, one should generally be critical of this standard way of measuring economic competition. A better measure for competition may be the profession because it allows a more precise measurement of competition based on substitutability (Hainmueller & Hopkins, 2015). We did not differentiate between specific professions because, in the Sub-Saharan African context, professions are less regulated than in the United States or parts of Europe. There is typically no occupational licensing and workers change their occupations often and also work in different jobs simultaneously. Assuming that competition occurs within cells defined by occupation did not strike our local partners or us as a particularly realistic way of capturing competition. We rather included additional survey questions to assess the concerns about direct competition in more detail.

Figure 5.1: Marginal means estimation for equation 5.1



Note: The graph plots marginal means for equation 5.1. Dots represent point estimates, whiskers around the dots represent 95% confidence intervals. The vertical line at 0.5 indicates the null effect. Standard errors are clustered on the respondent level. *Don't know* or *Refuse to answer* are dropped from the estimation. The corresponding numbers can be found in Table 5.A1. The AMCE equivalent can be found in Figure 5.A4 and Table 5.A2.

who are locally distant (Figure 5.A8).

The results are clear. No matter how we proxy potential labor market competition, marginal means do not differ substantially between respondent sub-groups. Despite some variation, overall preferences remain unchanged, with respondents strictly preferring high-skilled over low-skilled immigrants, irrespective of their own skill level, employment status, income level, and the immigrant's location. Thus, we find little evidence that concerns about direct competition with immigrants play a significant role in shaping attitudes toward them. The lack of egocentric economic concerns equally holds when we consider potential benefits that respondents might expect. Job-seeking respondents and respondents in work do not hold different attitudes toward immigrant occupations that could create jobs (supermarket owner and investor)(Figure 5.A6), suggesting that the hope of personally benefiting from new jobs created by immigrants is neither important. Egocentric concerns are minor drivers of attitudes at best. Hence, the economic concerns driving the attitudes toward different occupations in Figure 5.1 are sociotropic concerns.

Relative versus absolute drivers of attitudes

While the conjoint experiment can inform about *relative* attitudes within attribute dimensions, our survey data additionally allows us to investigate respondents' *absolute* attitudes toward immigrants. These are important to understand whether the host population is generally open to immigration or not, and what its beliefs about the net effects of immigration are.

Column 1 of Table 5.5 summarizes answers to our survey questions regarding immigrants in general, i.e. without explicitly inquiring about specific nationalities. On average, attitudes toward immigrants are positive, as most respondents would allow either some or many immigrants to come to their home country (2.85 on a scale from 1 (none) to 4 (many)). Respondents also perceive immigration as being economically neutral to beneficial: They rate immigrants' impact on host countries' economies and job creation as slightly positive (2.66 on a scale from 1 (good) to 5 (bad), and 4.58 on a scale from 0 (create new jobs) to 10 (take away jobs), respectively). In line with the absence of egocentric economic concerns in the conjoint experiment, respondents perceive immigration as relatively irrelevant to their own and their households' economic situation (2.78 on a scale from 1 (good) to 5 (bad)). Finally, respondents do not think that immigration undermines host countries' norms and values (3.7 on a scale from 0 (not at all) to 10 (a lot)).

These absolute attitudes are similar to those in Europe: For comparison, in the ESS, European respondents rate immigration as neutral to slightly beneficial for their countries' economy (5.2 on a scale from 0 (bad for the economy) to 10 (good for the economy)) and enriching for their countries' culture (5.4 on a scale from 0 (cultural life undermined) to

10 (cultural life enriched)).¹⁰

In absolute terms, respondents are not concerned about potential adverse impacts that might arise from immigration. While the conjoint experiment shows that respondents want immigrants to contribute to the host society economically and culturally, the survey questions show that respondents perceive immigrants as doing exactly that: benefiting the host countries' economy and creating jobs, without undermining respondents' livelihoods or the local culture. The alignment of relative and absolute concerns explains the overall positive evaluation of immigrants which we can also observe in the conjoint experiment where respondents want the average immigrant to live close by.

Chinese immigrants versus immigrants in general

In the conjoint experiment, respondents discriminated against Chinese and Indian/Lebanese immigrants relative to immigrants from Europe or Africa. Is there also a difference in absolute attitudes toward Chinese immigrants compared to immigrants in general? In Table 5.5, column 2 summarizes absolute attitudes toward Chinese immigrants, while column 3 shows t-tests for differences in attitudes toward immigrants in general. In line with the results from our conjoint experiment, respondents are significantly less open to immigration from China than to immigration in general. Still, respondents prefer, on average, "some" immigration of either group, indicating slightly positive absolute attitudes toward both groups.

The perceived economic effects of Chinese immigrants are rather positive. On average, respondents rate Chinese immigrants as slightly beneficial for their economic situation, the economy at large, and the creation of jobs. But when comparing economic concerns toward Chinese immigrants to those toward the average immigrant, the picture becomes more nuanced. Respondents perceive Chinese immigrants as having a slightly worse effect on the economy as a whole than immigrants in general. While there is no significant difference in perceived economic threat to respondents' personal situation, Chinese immigrants are perceived as significantly more likely to create jobs than immigrants in general. That suggests that the rather negative sociotropic economic assessment of immigrants is not only based on the creation of jobs. There must be another reason for the overall worse economic assessment of Chinese immigrants.

A similarly nuanced picture emerges when looking at absolute cultural concerns. Respondents do not think that Chinese immigrants undermine norms and values in absolute terms. Interestingly, cultural concerns are significantly lower for Chinese immigrants as compared to immigrants in general. This difference goes in the opposite direction of the generally more negative attitudes toward the Chinese in the conjoint and the survey

¹⁰A key feature of the ESS data is that the modal respondents select the middle category and only between half and two-thirds of respondents select other values (Ademmer & Stöhr, 2018). In our data, respondents' opinions are much less centered around the middle of the range, see Figure 5.A3.

section¹¹.

To sum up, the more negative view of the Chinese in both the conjoint and in the survey section cannot be explained by threats to norms and values, nor by job-related concerns. One alternative explanation for the more negative view of Chinese immigrants could be power concerns, i.e. respondents' fear that the Chinese government or Chinese businesses gain too much influence in Chinese immigrants' host countries. We study potential power concerns in the next section.

Do power concerns determine attitudes?

To capture power concerns, we ask about the perceived influence of foreign countries and foreign businesses. These two variables, however, should be taken with a grain of salt. In Uganda, we did not differentiate between foreign governments/businesses in general and those from China in particular. In Senegal, we did, but many data points contain "don't know" or "refuse to answer".¹²

Interestingly, while respondents' views on immigrants are generally positive, their opinions on foreign governments and businesses are, on average, negative. And while respondents seem to view Chinese immigrants more negatively than immigrants overall, when it comes to power concerns, the picture is reversed (Table 5.5). In line with many other surveys, such as Afrobarometer (Sanny & Selormey, 2020), power concerns in Senegal are less negative for Chinese governments and businesses compared to those from Europe or America (keeping in mind that we inquire about the perceived influence of the Chinese government and businesses only in Senegal). The influence of the Chinese government is seen as too small by the average respondent. That may surprise Westerners, who often perceive or frame Chinese involvement as a risk for Africans. However, Westerners tend to overlook that their economic involvement is also regularly seen as self-interested and of little value to locals. Even some development cooperation is critically perceived by Africans, especially when it props up unpopular regimes. By contrast, a substantial share of the population sees Chinese activity on the continent as an exchange that is not rooted in colonial patterns and (often) involves less finger-wagging (Carbone, 2011).

We also test whether power concerns explain heterogeneity in attitudes within the conjoint experiment by estimating subgroup differences. Respondents with strong power concerns may be more critical of immigrant groups whose origin country the respon-

¹¹An additional test by separately estimating the marginal means by respondents' assessment of the risks of immigration in general or the Chinese in particular pose to norms and values shows that the differences in attitudes between groups in the conjoint experiment are not significantly affected by the differential assessment of cultural threat (Figure 5.A26).

¹²In total, we have 243 unusable data points for the variable referring to the influence of the Chinese government and 235 for Chinese businesses. Especially in rural locations, respondents were often unaware of the influence of the Chinese government or Chinese businesses.

Table 5.5: Difference in means for main observational variables of interest, overall immigrants vs. Chinese immigrants

Variable	(1)		(2)		T-test Difference (1)-(2)
	Immigrants overall N	Mean/SE	Immigrants from China N	Mean/SE	
Allow immigrants to come to home country: none (=1) to many (=4)	2668	2.851 (0.015)	2652	2.580 (0.015)	0.272***
Are immigrants good or bad for respondent's/household's economy: good (=1) to bad (=5)	2633	2.782 (0.020)	2634	2.802 (0.019)	-0.020
Are immigrants good or bad for country's economy: good (=1) to bad (=5)	2632	2.656 (0.023)	2624	2.796 (0.023)	-0.140***
Do immigrants take or create jobs: create (=0) to take (=10)	2629	4.580 (0.061)	2622	3.990 (0.062)	0.590***
Do immigrants undermine country's norms and values: not at all (=0) to a lot (=10)	2638	3.697 (0.065)	2618	2.494 (0.056)	1.202***
Do foreign governments have too little or too much influence in country: too little (=0) to too much (=10)	2422	5.966 (0.070)	1257	3.437 (0.096)	2.529***
Do foreign businesses have too little or too much influence in country: too little (=0) to too much (=10)	2420	5.907 (0.066)	1265	4.570 (0.100)	1.337***

Notes: The value displayed for t-tests are the differences in the means across the groups. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level. Concerns regarding the influence of the Chinese government and Chinese businesses have only been asked separately in Senegal, and are missing for observations from Uganda. *Don't know* and *Refuse to answer* are coded as missing values.

dents associate with having too much influence. When estimating subgroup preferences for people who think that foreign governments and/or businesses, in general, have too much influence in their countries ¹³, results do not change substantially (Figures 5.A9 and 5.A10). The results thus suggest that relative preferences for certain immigrant characteristics do not depend on respondents' power concerns.

It seems clear that the more negative view of Chinese immigrants does not stem from power concerns. If anything, these might reduce the gap between the Chinese and other immigrant groups. If the difference in absolute attitudes is not driven by power concerns and differences in economic and cultural valuations between both groups are very small (and often more positive for Chinese immigrants), it may be taste-based. Indeed, during data collection, we noticed that Chinese immigrants were viewed more critically than other immigrants, even if they behaved identically.

What type of concern is the most important driver of attitudes?

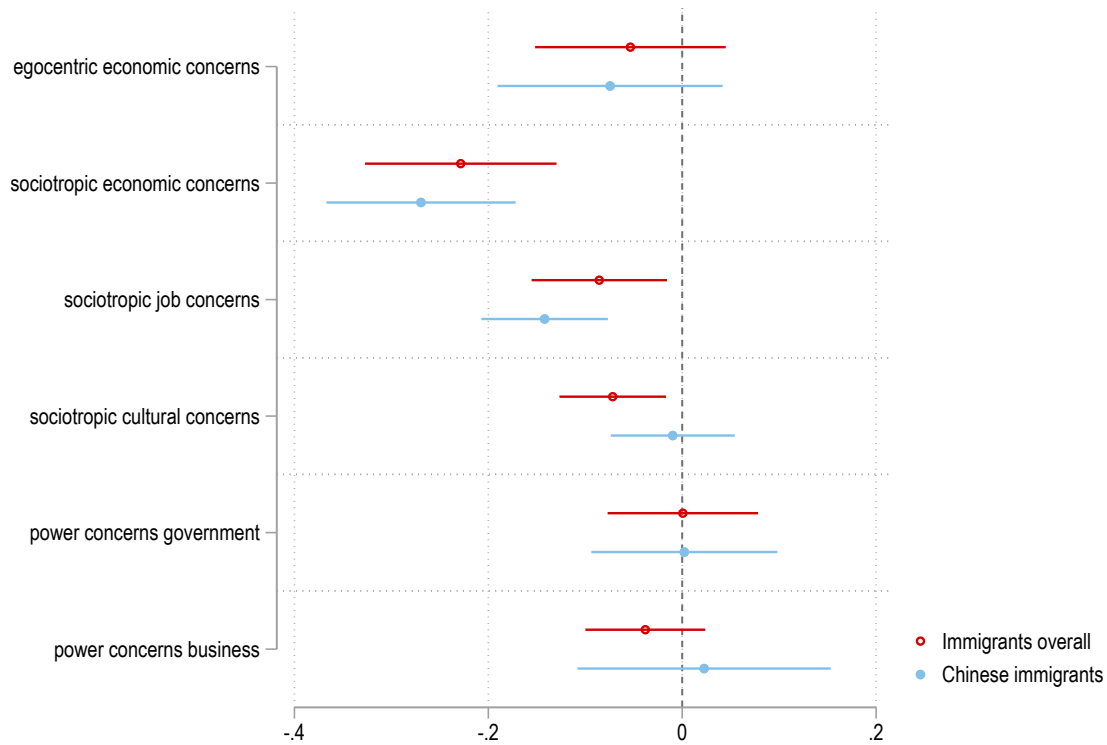
As mentioned before, the conjoint experiment does not allow to directly compare effect sizes across different dimensions. To identify the contribution of different determinants of attitudes toward immigrants in general and Chinese immigrants in particular, we estimate correlates of absolute attitudes toward immigrants using equation 5.2. We standardize outcome and indicator variables and cluster standard errors at the district level. We run different specifications: without and with control variables ¹⁴, and without and with including power concerns. When including power concerns, for the Ugandan sample we extrapolate answers to indicator questions on the perceived influence of foreign governments and businesses to the missing variables for the perceived influence of Chinese governments and businesses. Thereby, we assume that power concerns in Uganda are the same for foreign governments and businesses in general and Chinese ones in particular. Results for the different specifications are reported in Table 5.A6.

Figure 5.2 illustrates the main results from our preferred specification (standardized outcome and indicator variables, including control variables, extrapolating power concerns for the Ugandan sample), overlaying findings for attitudes toward immigrants in general (red, column (7) in Table 5.A6) and Chinese immigrants in particular (blue, column (8) in Table 5.A6). Dots plot point estimates, whiskers around dots plot 95% confidence intervals. Indicator variables proxying respondents' concerns are coded such that more negative values represent stronger concerns. Outcome variables proxying respondents' attitudes toward immigrants are coded such that more positive values represent

¹³Bigger than 5 on a scale from 0 (=too little) to 10 (=too much). We have this information for both Uganda and Senegal.

¹⁴Respondents' age, gender, education, self-reported employment during the last 7 days (4 levels), location (rural-urban), self-reported contact with immigrants, country-specific quintiles for self-reported household income per capita during the last 12 months, life satisfaction (0-10), and satisfaction with city/area of residence (1-5), and country fixed effects

Figure 5.2: Contributions of different types of concerns to overall attitudes



Note: The graph plots OLS estimates for equation 5.2. Table version in Table 5.A6. Outcome and indicator variables are standardized. The regression includes the full set of controls and country fixed effects. In Senegal, we collected power concerns separately for foreign governments/businesses in general and Chinese ones in particular. In Uganda, we assume that power concerns are the same for foreign governments and businesses in general and Chinese ones in particular. Standard errors are clustered at the district level. *Don't know* and *Refuse to answer* are coded as missing. For tables see 5.A6, Columns (7) and (8).

more positive attitudes. Thus, if a specific concern is a significant predictor of attitudes toward immigrants, their coefficient should be negative and significant.

The results complement the experimental findings from the conjoint experiment, adding nuance to them. Again, egocentric economic concerns do not seem to play an important role. Despite explicitly asking respondents about immigrants' threat to their own and their households' economic situation, egocentric economic concerns are not a significant predictor of attitudes toward immigrants in general and are also not significantly related to attitudes toward Chinese immigrants.

What could explain the apparent lack of egocentric economic concerns in the context of strong competition for limited resources? In our sample, as in many other countries in Sub-Saharan Africa, households and even individuals often work multiple jobs at the time (Barrett et al., 2001; Berg et al., 2021). The high degree of income diversification and informality among people in Sub-Saharan Africa implies that locals are less affected by competition with incoming immigrants. Moreover, many African societies are less individualistic than Western societies (Gorodnichenko & Roland, 2017). This might explain why individuals in Sub-Saharan Africa give more importance to sociotropic factors,

despite the economic challenges that respondents often face.

Broader sociotropic economic concerns, measured by asking respondents about immigrants' effects on the host economy as a whole and on job creation, are associated with significantly more negative attitudes toward immigrants, underpinning the importance of immigrants' ability to contribute to the host country's economy. A one standard deviation increase in sociotropic economic concerns is associated with a 0.23 standard deviation decrease in attitudes toward immigrants in general and a 0.27 standard deviation decrease in attitudes toward Chinese immigrants in particular. The narrower job-creating concern can add to this. If job concerns are one standard deviation higher, the overall attitude toward immigration is 0.09 standard deviation lower for immigrants and 0.14 standard deviations for the Chinese.

While results from the conjoint experiment highlight the importance of sociotropic cultural concerns (measured by immigrants' origin and willingness to integrate), fear of immigrants undermining the country's norms and values seems to play a minor role in determining overall attitudes. The effect of cultural concerns is very small and only statistically different from zero for immigrants in general. These concerns thus hardly influence overall attitudes.

Although we find considerable variation in power concerns in absolute terms, their predictive power for attitudes toward immigration is statistically insignificant in both specifications. Respondents who have relatively greater power concerns are not more critical of immigrants. While this finding contradicts our hypothesis about power concerns, it is in line with several conversations we had with locals in the field. In Uganda and Senegal, people are worried about the influence of foreign governments and businesses, but they differentiate between governments and businesses and immigrants themselves.

Overall, sociotropic economic concerns are the most important determinant of attitudes toward both groups of immigrants. Point estimates for them are significantly larger than those for job concerns or any other type of concerns (Table 5.A7). Altogether, while Chinese immigrants are perceived slightly more negatively and economically worse for the economy despite creating more jobs, the analysis shows that attitude formation toward this group is not systematically different from that toward immigration overall.

A hidden consensus within and across countries

The analyses seen so far provide average results for our entire study population. To better understand the possible implications of the drivers of attitudes, it is also important to investigate whether there is broad agreement in society or whether the average results mask a highly polarized view that would require different political answers. We study differences in attitudes and attitude formation across respondent subgroups in two steps. First, we analyze whether there is heterogeneity in the determinants of relative attitudes,

i.e. if certain individuals have stronger preferences for certain characteristics of migrants in the conjoint experiment. Second, we study the determinants of absolute attitudes by running regressions to see which respondent-level variables determine absolute attitudes in the survey section.

We have already established that relative preferences do not differ by respondents' education or economic situation. Testing for gender differences in the conjoint (Figure 5.A17) shows that, while estimates vary significantly at the 5% level, differences are small in absolute terms and coefficients do not change their sign¹⁵.

Differences between rural-urban subgroups are insignificant, indicating that people have broadly similar views about the importance of our conjoint dimensions regardless of the place they live (Figure 5.A19).

We also test for between and within-country heterogeneity. Conducting our main estimations separately for Uganda and Senegal leaves the overall picture unchanged, with only minor differences between the two countries (Figure 5.A18). In Uganda, the preference for immigrants that live in the same community is stronger. Respondents in Senegal view European immigrants significantly more negatively than respondents in Uganda, which fits the impression from the field. People in Uganda and Senegal have different opinions toward immigrants from former colonial powers. Based on our experience, there are no signs of particularly negative sentiment toward British immigrants in Uganda. However, in Senegal, French immigrants have been seen much more critically - at least during personal conversations in the field.

Looking at within-country heterogeneity, attitudes might vary on the geographical level. Within Uganda and Senegal, there is substantial cultural, political, and economic heterogeneity across regions, and regional peculiarities such as hosting refugees or hosting special economic zones might influence respondents' preferences. When estimating marginal means for the different districts in Uganda and Senegal, estimates remain largely unchanged (Figure 5.A20).

We include several additional subgroup analyses to investigate potential heterogeneous effects among respondent groups, which were not pre-registered. First, we test whether attitudes toward immigrants vary with self-reported contact with (Chinese) immigrants. Based on the contact hypothesis, increased contact with (Chinese) immigrants should lead to more positive attitudes toward these groups (Pettigrew & Tropp, 2006). In our sample, however, we do not observe such a pattern. The overall picture remains unchanged, irrespective of self-reported contact with immigrants in general (Figure 5.A21)

¹⁵When assessing differences by gender on absolute attitudes in the survey questions, we also find some differences (Tables 5.A8 to 5.A14). Conditional on their age, education, labor market status, and location, women prefer slightly less immigration, feel more economically threatened by Chinese immigrants, and have lower power concerns. But with coefficients typically around 0.1-0.15 on a five or ten-point scale, none of these differences is large. Thus, males and females have very similar attitudes toward immigrants, the optimal level of immigration, and assess its effects similarly.

or Chinese immigrants in particular (Figure 5.A22). Alternatively to contact, respondents' beliefs about how many (Chinese) immigrants already live in their country could matter (Bursztyn & Yang, 2022). To check this, we create country-specific quintiles for respondents' estimates of the number of (Chinese) immigrants currently living in their home country. Again, overall results remain unchanged (Figures 5.A23 and 5.A24). Contact or beliefs about numbers do not seem to determine the attitudes toward migrant groups, similar to recent studies from Europe or the US (Alesina et al., 2023; Grigorieff et al., 2020)

We also test whether the respondents' preferred destination region for migration influences their preferences for the origins of immigrants. We group respondents' preferred destination countries into the regions Africa, Europe, South and Western Asia (including India and Lebanon), and Eastern Asia (including China). Marginal means do not differ much between groups. The only point estimate that deviates substantially is the one for immigrants from China among respondents whose preferred destination region is East Asia (Figure 5.A25). The coefficient is positive and close to zero, but not significant. Note that some of the respondent subgroups are small, leading to large standard errors and loss of power.

Altogether, the drivers of relative attitudes are very similar across different subgroups of the population and across countries. What could nonetheless lead to differences in attitudes are differences in *absolute* beliefs about the economic, cultural, or power-related influences that immigrants have. For example, more educated people might think very similarly about the relative importance of immigrants' economic contribution to society. Still, compared to their less educated compatriots, they might perceive the absolute economic contribution differently, resulting in different overall attitudes. These differences could exist despite not finding differences in the conjoint experiment or the distinction of mechanisms in Figure 5.2. To study this, we use the outcome and indicator variables from the survey section and predict them with our usual set of respondent-level controls.¹⁶ The results emphasize the consensus in views (Tables 5.A8 to 5.A14). There are only minor differences in perceptions. Given the risk of false positives in hypothesis testing, we refrain from discussing weaker or smaller differences. One of them is large and worth noting, though: Compared to individuals without education, the more educated individuals are more critical of the role of foreign businesses and governments in general.

In summary, our results are surprisingly homogeneous and vary hardly with respondents' characteristics or geographic location. Similar to findings from the US (Hainmueller & Hopkins, 2015), there seems to be a "consensus" about attitudes toward immigrants in Uganda and Senegal. While one might argue that the absence of heterogeneity stems from respondents' lack of interest or awareness, our survey questions show the opposite: respondents have strong opinions on immigration to their country, leading to answers

¹⁶Note that differences will be captured at least in part by these controls in Figure 5.2.

which are more skewed than answers to similar questions in the ESS (Figure 5.A3). The homogeneity of our results across a multitude of subgroups from two countries that share similar labor market characteristics with many other countries in the region strengthens our confidence that our results can also speak for many other parts of Sub-Saharan Africa.

Robustness

Our experimental results are robust to several additional specifications. We test for profile order effects (Figure 5.A28), attribute order effects (Figure 5.A29), and carryover effects (Figure 5.A30) - none of which alter our results substantially. Preferences for certain attribute levels do not depend on the profile, the order, or the task they appear in. Attribute levels are balanced within dimensions and have the same probability of being sampled (Figure 5.A31). To control for potential experimental fatigue among respondents that had to be revisited, we estimate marginal means by whether respondents were revisited or not. Estimates remain largely unchanged, reassuring us that revisiting households did not impact our estimates (Figure 5.A32). Finally, we re-estimate 5.1, (i) using the Bonferroni-method to control for multiple hypothesis testing (29 different conjoint specifications) and (ii) applying sampling and non-response weights. Again, our results remain virtually unchanged (Table 5.A15).

As an additional exercise, we control for two-way interaction effects between different attribute levels by interacting all attribute levels with each other (Egami & Imai, 2018). Results are invariant: Respondents strictly prefer high-skilled over low-skilled immigrants, African over non-African immigrants, immigrants who are willing to integrate over those who are not willing to integrate, and immigrants within their community over those outside - irrespective of the other profiles' attribute levels (Figures 5.A11 to 5.A16).

The results from our survey data analysis are robust to different specifications. Given the shortcomings of the variables measuring power concerns, we re-estimate our model without power concerns. We also estimate the main specification without control variables to maximize our sample size. Estimated coefficients remain robust in both additional specifications (Table 5.A6).

5.6 Conclusion

While there is vast literature on attitudes toward immigrants in Western countries, little systematic knowledge of the perception of immigration in Sub-Saharan Africa exists. Understanding what Africans think about migration and what influences their opinions is important for key policy developments, such as the planned African-Union-wide visa-free travel and similar plans made in tandem with the African Continental Free Trade Agreement. Another critical trend affecting countries in Sub-Saharan Africa is the activity of

China. There has been an increasing Chinese immigration to Sub-Saharan Africa, yet the existing knowledge about the perception of this immigration is mostly anecdotal.

To fill this gap in the literature, we provide the first large-scale study of attitudes toward immigrants in Sub-Saharan Africa that uses a causal framework. We study attitudes toward immigrants in general and Chinese immigrants in particular in Uganda and Senegal. While Uganda and Senegal differ in immigration patterns, the labor markets in both countries are characterized by high levels of underemployment and informality - similar to many other countries in Sub-Saharan Africa.

Using experimental and survey-based estimates, we analyze whether egocentric economic, sociotropic economic, cultural, or power concerns determine attitudes toward immigrants. We provide quantitative evidence that sociotropic economic concerns and sociotropic cultural concerns are strong drivers of attitudes toward immigrants. As a consequence, citizens with a negative view of the economic or cultural effects of migrants are more opposed to further immigration. In particular, immigrants' perceived impact on the economy and the creation of jobs is essential. If immigrants have the potential to create jobs, they are particularly welcome. Furthermore, our results highlight that respondents' attitudes are not significantly predicted by their self-interest, such as the potential employment benefits they could receive. Immigrants' cultural effect is the other major driver of attitudes. Respondents react strongly to immigrants' willingness to integrate into local society and the perceived effect on norms and values matters for the attitude toward further immigration.

Despite the fierce competition for resources and jobs in Sub-Saharan Africa, egocentric economic concerns do not play a significant role in determining attitudes. Similar to the US and Europe, sociotropic concerns are what matter in Sub-Saharan Africa. This finding, which differs from our pre-registered hypotheses, is one of the most important take-aways from our study. Our paper shows, for the first time, that the determinants of attitudes toward immigrants are not that different between the rich industrialized economies that have mostly been studied and the low and lower-middle income countries that we focus on.

Immigrants from China, in particular, are perceived as economically and culturally less threatening than immigrants in general. Yet, Chinese immigrants are perceived less positively than other immigrants in absolute terms. We interpret this as evidence for taste-based differences.

Our results on power concerns indicate that respondents are generally somewhat concerned about the influence of foreign governments and businesses. In Senegal, where we distinguished power concerns in general and regarding the Chinese, respondents perceive the influence of Chinese businesses and especially the Chinese government as too small. That stands very much in contrast to Western rhetoric. However, power concerns do not influence the overall attitudes toward immigration or the relative preference of Chinese

migrants vis-à-vis other immigrants. We conclude that respondents distinguish between the more abstract power concerns, their attitudes toward immigrants, and their resulting preferences for immigration policy. Our results are surprisingly homogeneous and vary hardly with respondents' characteristics or geographic location.

Overall, our study shows that despite the vastly different economic and contextual factors, the determinants of attitudes toward migrants are remarkably similar to those already established by research conducted in the US and Europe: sociotropic concerns outweigh egocentric concerns. Attitude formation processes thus seem to carry over from context to context, even if the level of competition required to make a living is vastly different. Furthermore, similar to the US, there seems to be a "hidden consensus" about attitudes toward immigrants in Uganda and Senegal (Hainmueller & Hopkins, 2015). Respondents agree on what kind of immigration they prefer: Immigrants who benefit the overall economy, are willing to integrate and do not create a threat to norms and values. In contrast to many of the rich industrialized economies, overall attitudes toward immigrants are positive: Respondents perceive immigrants as beneficial for the economy and, in stark contrast to strong not in my backyard concerns in the West, prefer them to live close by.

Our results provide support for policymakers who are planning to allow more immigration, for example, by easing migration restrictions within the African Union. They also suggest that when aiming for a reduction in tensions and xenophobia, communicating the positive impacts of immigration on society at large and promoting immigrants' willingness to integrate into host societies are worthwhile mechanisms to target.

Bibliography

- Achieng, M., El Fadil, A., Adepoju, A., Fumagalli, C., & Nyabola, N. (2020). *Africa migration report. Challenging the narrative*. International Organization for Migration, Addis Ababa.
- Ademmer, E., & Stöhr, T. (2018). Europeans are more accepting of immigrants today than 15 years ago. *MEDAM Policy Brief, 2018*(1).
- Adida, C. L. (2011). Too close for comfort? Immigrant exclusion in Africa. *Comparative Political Studies, 44*(10), 1370–1396.
- Adida, C. L., Lo, A., & Platas, M. R. (2019). Americans preferred Syrian refugees who are female, English-speaking, and christian on the eve of Donald Trump’s election. *PloS one, 14*(10), e0222504.
- Ajzenman, N., Aksoy, C. G., & Guriev, S. (2022). Exposure to transit migration: Public attitudes and entrepreneurship. *Journal of Development Economics, 102*899.
- Aksoy, C. G., & Ginn, T. (2022). Attitudes and policies toward refugees. *Policy Research Working Paper*.
- Alesina, A., Miano, A., & Stantcheva, S. (2023). Immigration and redistribution. *The Review of Economic Studies, 90*(1), 1–39.
- Alesina, A., & Tabellini, M. (2022). *The political effects of immigration: Culture or economics?* (Tech. rep.). National Bureau of Economic Research.
- Alrababa’h, A., Dillon, A., Williamson, S., Hainmueller, J., Hangartner, D., & Weinstein, J. (2021). Attitudes toward migrants in a highly impacted economy: Evidence from the Syrian refugee crisis in Jordan. *Comparative Political Studies, 54*(1), 33–76.
- Balmas, M. (2018). Tell me who is your leader, and i will tell you who you are: Foreign leaders’ perceived personality and public attitudes toward their countries and citizenry. *American Journal of Political Science, 62*(2), 499–514.
- Bansak, K., Hainmueller, J., & Hangartner, D. (2016). How economic, humanitarian, and religious concerns shape European attitudes toward asylum seekers. *Science, 354*(6309), 217–222.
- Barrett, C. B., Reardon, T., & Webb, P. (2001). Nonfarm income diversification and household livelihood strategies in rural Africa: Concepts, dynamics, and policy implications. *Food Policy, 26*(4), 315–331.

- Berg, J., Hilal, A., El, S., Horne, R., et al. (2021). *World employment and social outlook: Trends 2021*. International Labour Organization.
- Betts, A., Stierna, M. F., Omata, N., & Sterck, O. (2023). Refugees welcome? Inter-group interaction and host community attitude formation. *World Development*, *161*, 106088.
- Blair, R. A., Marty, R., & Roessler, P. (2022). Foreign aid and soft power: Great power competition in Africa in the early twenty-first century. *British Journal of Political Science*, *52*(3), 1355–1376.
- Botchwey, G., Crawford, G., Loubere, N., & Lu, J. (2019). South-south irregular migration: The impacts of China's informal gold rush in Ghana. *International Migration*, *57*(4), 310–328.
- Boumedouha, S. (1990). Adjustment to West African realities: The Lebanese in Senegal. *Africa*, *60*(4), 538–549.
- Buehler, M., & Han, K. J. (2018). Who endorses amnesty? An original survey from Morocco's Casablanca region assessing citizen support for regularizing clandestine migrants. *Review of Middle East Studies*, *52*(2), 283–298.
- Bursztyn, L., & Yang, D. Y. (2022). Misperceptions about others. *Annual Review of Economics*, *14*, 425–452.
- Carbone, M. (2011). The European Union and China's rise in Africa: Competing visions, external coherence and trilateral cooperation. *Journal of Contemporary African Studies*, *29*(2), 203–221.
- Card, D., Dustmann, C., & Preston, I. (2012). Immigration, wages, and compositional amenities. *Journal of the European Economic Association*, *10*(1), 78–119.
- Cervellati, M., Esposito, E., Sunde, U., & Yuan, S. (2022). Malaria and Chinese economic activities in Africa. *Journal of Development Economics*, *154*, 102739.
- Cogley, N. T., Doces, J. A., & Whitaker, B. E. (2019). Which immigrants should be naturalized? Which should be deported? Evidence from a survey experiment in Cote d'Ivoire. *Political Research Quarterly*, *72*(3), 653–668.
- Custer, S., Dreher, A., Elston, T.-B., Fuchs, A., Ghose, S., Lin, J., Malik, A., Parks, B. C., Russell, B., Solomon, K., et al. (2021). *Tracking Chinese development finance: An application of AidData's TUFF 2.0 methodology*. Williamsburg, VA: AidData at William & Mary.
- Denney, S., & Green, C. (2021). Who should be admitted? Conjoint analysis of South Korean attitudes toward immigrants. *Ethnicities*, *21*(1), 120–145.
- Devillard, A., Bacchi, A., & Marion, N. (2015). *A survey on migration policies in West Africa*. International Centre for Migration Policy Development.
- Dreher, A., Fuchs, A., Parks, B., Strange, A. M., & Tierney, M. J. (2018). Apples and dragon fruits: The determinants of aid and other forms of state financing from China to Africa. *International Studies Quarterly*, *62*(1), 182–194.

- Dustmann, C., Schönberg, U., & Stuhler, J. (2016). The impact of immigration: Why do studies reach such different results? *Journal of Economic Perspectives*, 30(4), 31–56.
- Dustmann, C., Vasiljeva, K., & Piil Damm, A. (2019). Refugee migration and electoral outcomes. *The Review of Economic Studies*, 86(5), 2035–2091.
- Egami, N., & Imai, K. (2018). Causal interaction in factorial experiments: Application to conjoint analysis. *Journal of the American Statistical Association*.
- Ember, M., Ember, C. R., & Skoggard, I. (2004). *Encyclopedia of diasporas: Immigrant and refugee cultures around the world*. Springer Science & Business Media.
- Facchini, G., Mayda, A. M., Guiso, L., & Schultz, C. (2008). From individual attitudes towards migrants to migration policy outcomes: Theory and evidence. *Economic Policy*, 23(56), 651–713.
- Facchini, G., Mayda, A. M., & Mendola, M. (2013). What drives individual attitudes towards immigration in South Africa? *Review of International Economics*, 21(2), 326–341.
- French, H. W. (2015). *China's second continent: How a million migrants are building a new empire in Africa*. Vintage Books.
- Friebel, G., Gallego, J., & Mendola, M. (2013). Xenophobic attacks, migration intentions, and networks: Evidence from the South of Africa. *Journal of Population Economics*, 26(2), 555–591.
- Gaye, A. (2008). China in Africa: After the gun and the bible: A West African perspective. *China Returns to Africa—A Rising Power and a Continent Embrace*. London: Hurst & Company, 129–142.
- Gehring, K., Kaplan, L. C., & Wong, M. H. (2022). China and the World Bank—how contrasting development approaches affect the stability of African states. *Journal of Development Economics*, 158, 102902.
- Ghosn, F., Braithwaite, A., & Chu, T. S. (2019). Violence, displacement, contact, and attitudes toward hosting refugees. *Journal of Peace Research*, 56(1), 118–133.
- Goldsmith, B. E., Horiuchi, Y., & Matush, K. (2021). Does public diplomacy sway foreign public opinion? Identifying the effect of high-level visits. *American Political Science Review*, 115(4), 1342–1357.
- Gordon, S. (2018). Who is welcoming and who is not? An attitudinal analysis of anti-immigrant sentiment in South Africa. *South African Review of Sociology*, 49(1), 72–90.
- Gorodnichenko, Y., & Roland, G. (2017). Culture, institutions, and the wealth of nations. *Review of Economics and Statistics*, 99(3), 402–416.
- Grigorieff, A., Roth, C., & Ubfal, D. (2020). Does information change attitudes toward immigrants? *Demography*, 57(3), 1117–1143.

- Haaland, I., & Roth, C. (2020). Labor market concerns and support for immigration. *Journal of Public Economics*, 191, 104256.
- Hainmueller, J., Hiscox, M. J., & Margalit, Y. (2015). Do concerns about labor market competition shape attitudes toward immigration? New evidence. *Journal of International Economics*, 97(1), 193–207.
- Hainmueller, J., & Hopkins, D. J. (2014). Public attitudes toward immigration. *Annual Review of Political Science*, 17(1), 225–249.
- Hainmueller, J., & Hopkins, D. J. (2015). The hidden American immigration consensus: A conjoint analysis of attitudes toward immigrants. *American Journal of Political Science*, 59(3), 529–548.
- Hainmueller, J., Hopkins, D. J., & Yamamoto, T. (2014). Causal inference in conjoint analysis: Understanding multidimensional choices via stated preference experiments. *Political Analysis*, 22(1), 1–30.
- Hairong, Y. (2020). *The truth about Chinese migrants in Africa and their self-segregation* [Downloaded from [<https://apmigration.ilo.org/news/the-truth-about-chinese-migrants-in-africa-and-their-self-segregation>] on 28 March 2022]. Quartz.
- Harris, A. S., Findley, M. G., Nielson, D. L., & Noyes, K. L. (2018). The economic roots of anti-immigrant prejudice in the global south: Evidence from South Africa. *Political Research Quarterly*, 71(1), 228–241.
- Hartman, A. C., & Morse, B. S. (2020). Violence, empathy and altruism: Evidence from the Ivorian refugee crisis in Liberia. *British Journal of Political Science*, 50(2), 731–755.
- Iacoella, F., Martorano, B., Metzger, L., & Sanfilippo, M. (2021). Chinese official finance and political participation in Africa. *European Economic Review*, 136, 103741.
- Kadigo, M. M., Diallo, N. O., & Maystadt, J. F. P. C. (2022). *How to cope with a refugee shock? Evidence from Uganda*. Washington, DC: World Bank.
- Leeper, T. J., Hobolt, S. B., & Tilley, J. (2020). Measuring subgroup preferences in conjoint experiments. *Political Analysis*, 28(2), 207–221.
- Lehmann, M. C., & Masterson, D. T. R. (2020). Does aid reduce anti-refugee violence? Evidence from Syrian refugees in Lebanon. *American Political Science Review*, 114(4), 1335–1342.
- Leichtman, M. A. (2005). The legacy of transnational lives: Beyond the first generation of Lebanese in Senegal. *Ethnic and Racial Studies*, 28(4), 663–686.
- Mandon, P., & Woldemichael, M. T. (2023). Has Chinese aid benefited recipient countries? Evidence from a meta-regression analysis. *World Development*, 166, 106211.
- Matush, K. (2023). Harnessing backlash: How leaders can benefit from antagonizing foreign actors. *British Journal of Political Science*, 1–17.

- Mayda, A. M. (2006). Who is against immigration? A cross-country investigation of individual attitudes toward immigrants. *The Review of Economics and Statistics*, 88(3), 510–530.
- McCauley, J. F., Pearson, M. M., & Wang, X. (2022). Does Chinese FDI in Africa inspire support for a China model of development? *World Development*, 150, 105738.
- Merotto, D. (2020). *Uganda: Jobs strategy for inclusive growth*. Washington, DC: World Bank.
- MOFCOM. (2021). Foreign investment cooperation country (region) guide. *Ministry of Commerce of the People's Republic of China*.
- Mosel, I., Leach, A., & Hargrave, K. (2020). *Public narratives and attitudes towards refugees and other migrants: Uganda country profile*. ODI.
- OECD. (2021). *International migration outlook 2021*. Organisation for Economic Co-operation and Development Publishing.
- Park, Y. J. (2022). Forever foreign? Is there a future for Chinese people in Africa? *Journal of Ethnic and Migration Studies*, 48(4), 894–912.
- Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, 90(5), 751.
- Ruedin, D. (2019). Attitudes to immigrants in South Africa: Personality and vulnerability. *Journal of Ethnic and Migration Studies*, 45(7), 1108–1126.
- Ruhs, M. (2022). Who cares what the people think? Public attitudes and refugee protection in Europe. *Politics, Philosophy & Economics*, 21(3), 313–344.
- Sanny, J. A.-N., & Selormey, E. (2020). Africans regard China's influence as significant and positive, but slipping. *Afrobarometer*.
- Sardoschau, S., & Jarotschkin, A. (2019). Chinese aid in Africa: Attitudes and conflict. *SSRN*, 3383978.
- Scheve, K. F., & Slaughter, M. J. (2001). Labor market competition and individual preferences over immigration policy. *Review of Economics and Statistics*, 83(1), 133–145.
- Sniderman, P. M., Hagendoorn, L., & Prior, M. (2004). Predisposing factors and situational triggers: Exclusionary reactions to immigrant minorities. *American Political Science Review*, 98(1), 35–49.
- Thomas-Johnson, A. (2021). *Senegal: Anti-French sentiment on the rise as protests continue* [Downloaded from [<https://www.aljazeera.com/news/2021/3/12/senegal-anti-french-sentiments-on-the-rise-amid-ongoing-protests>] on 08 September 2022]. AlJazeera.
- Valentino, N. A., Soroka, S. N., Iyengar, S., Aalberg, T., Duch, R., Fraile, M., Hahn, K. S., Hansen, K. M., Harell, A., Helbling, M., Jackman, S. D., & Kobayashi, T. (2019). Economic and cultural drivers of immigrant support worldwide. *British Journal of Political Science*, 49(4), 1201–1226.

- Wegenast, T., Krauser, M., Strüver, G., & Giesen, J. (2019). At Africa's expense? Disaggregating the employment effects of Chinese mining operations in Sub-Saharan Africa. *World Development*, *118*, 39–51.
- Whitaker, B. E., & Giersch, J. (2015). Political competition and attitudes towards immigration in Africa. *Journal of Ethnic and Migration Studies*, *41*(10), 1536–1557.
- World Bank. (2022). World development indicators [Downloaded from <https://databank.worldbank.org/source/world-development-indicators>] on 02 December 2022].
- Xu, Z. P., & Zhang, Y. (2020). Can Chinese aid win the hearts and minds of Africa's local population? *Economic Modelling*, *90*, 322–330.
- Zhou, Y.-Y., Grossman, G., & Ge, S. (2023). Inclusive refugee-hosting can improve local development and prevent public backlash. *World Development*, *166*, 106203.

Appendix

5.A.1 Background Uganda and Senegal

Uganda was home to over 1.7 million immigrants in 2019. It is among the top five refugee-hosting countries in the world and the leading country in Africa. Uganda also hosts significant numbers of regular labor or transit migrants from neighboring countries such as Rwanda and Tanzania (Mosel et al., 2020). While non-African immigrants account for less than 1% of Uganda's immigrant stock (Mosel et al., 2020), they are very visible in daily life and the media. Indians are historically the most relevant group of non-African immigrants to Uganda, settling there mostly when Uganda and India were both British colonies. They constitute a group of immigrants who have been members of society for many decades. However, this relationship has not always been a harmonious one. Notably, several tens of thousands of Indians were expelled in 1972 by the regime of Idi Amin. By the end of his regime's downfall in 1979, almost all Indian citizens had left the country and thus lost their businesses and most other possessions. Under the still ruling Museveni government, Indians were invited back to Uganda in the mid-1980s (Ember et al., 2004). Official data from 2012/13 shows that the highest number of work permits in Uganda was issued to Indian citizens (39%).

Senegal is an important destination for migrants within the Economic Community Of West African States. The majority of Senegal's immigrant population are regular migrants from neighboring countries such as Mauritania, Mali, Gambia, and Guinea Bissau (Devillard et al., 2015). What the Indians are to Uganda are the Lebanese to Senegal. The Lebanese formed an important part of the Senegalese society and economy already during colonial times, engaging in trade and politics alike. They competed with French traders in the peanut trade, making them the target of anti-immigration propaganda and lobbying. After Senegal gained independence in 1960 (supported by the Lebanese community) the Lebanese stayed in the country, increasingly competing with Senegalese traders (Boumedouha, 1990). Today, the Lebanese form a well-established yet distinct population group in Senegal (Leichtman, 2005). They are an economically strong minority but face increasing competition from Chinese immigrants in recent years (Gaye, 2008).

5.A.1.1 Study area

Figure 5.A1: Interview locations in Uganda

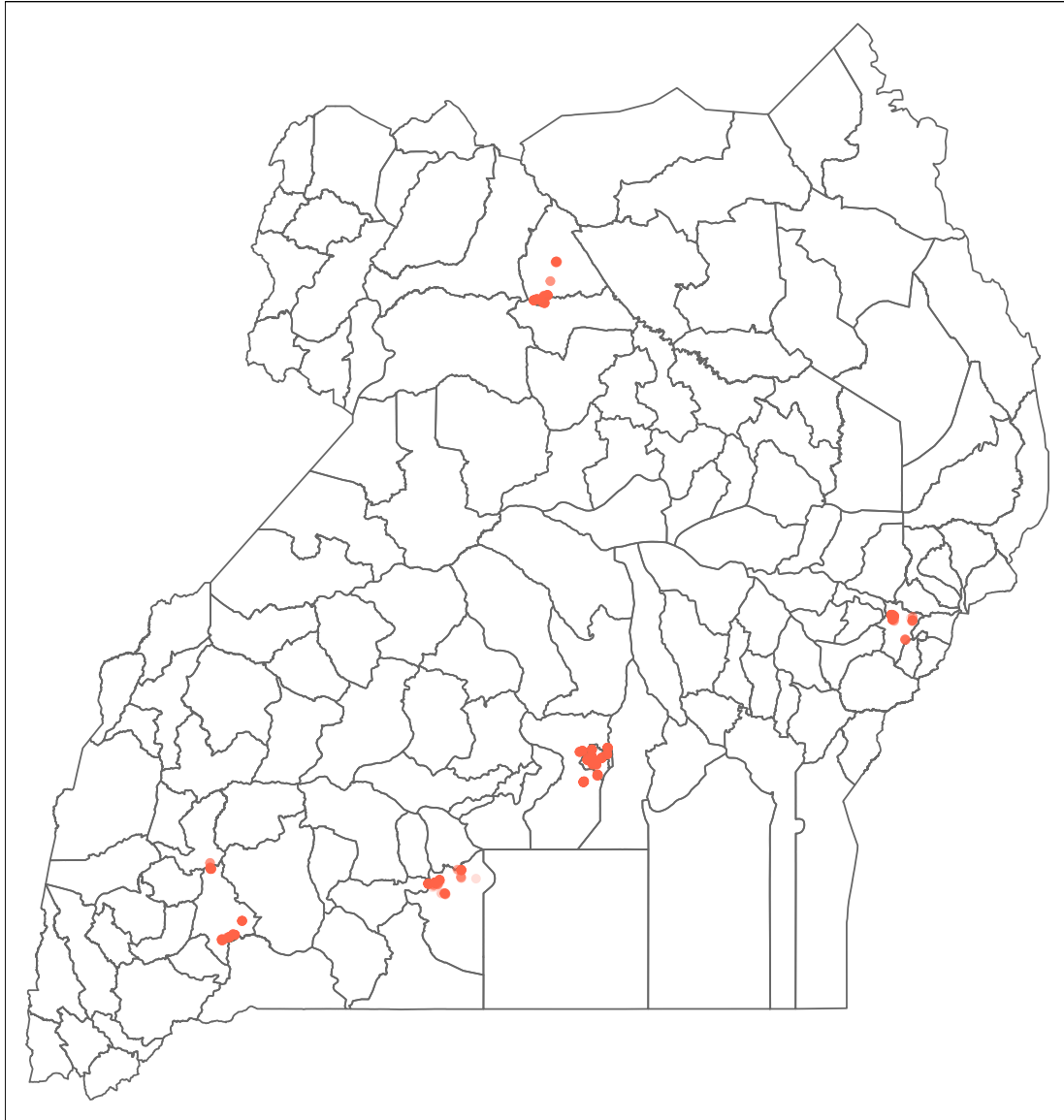
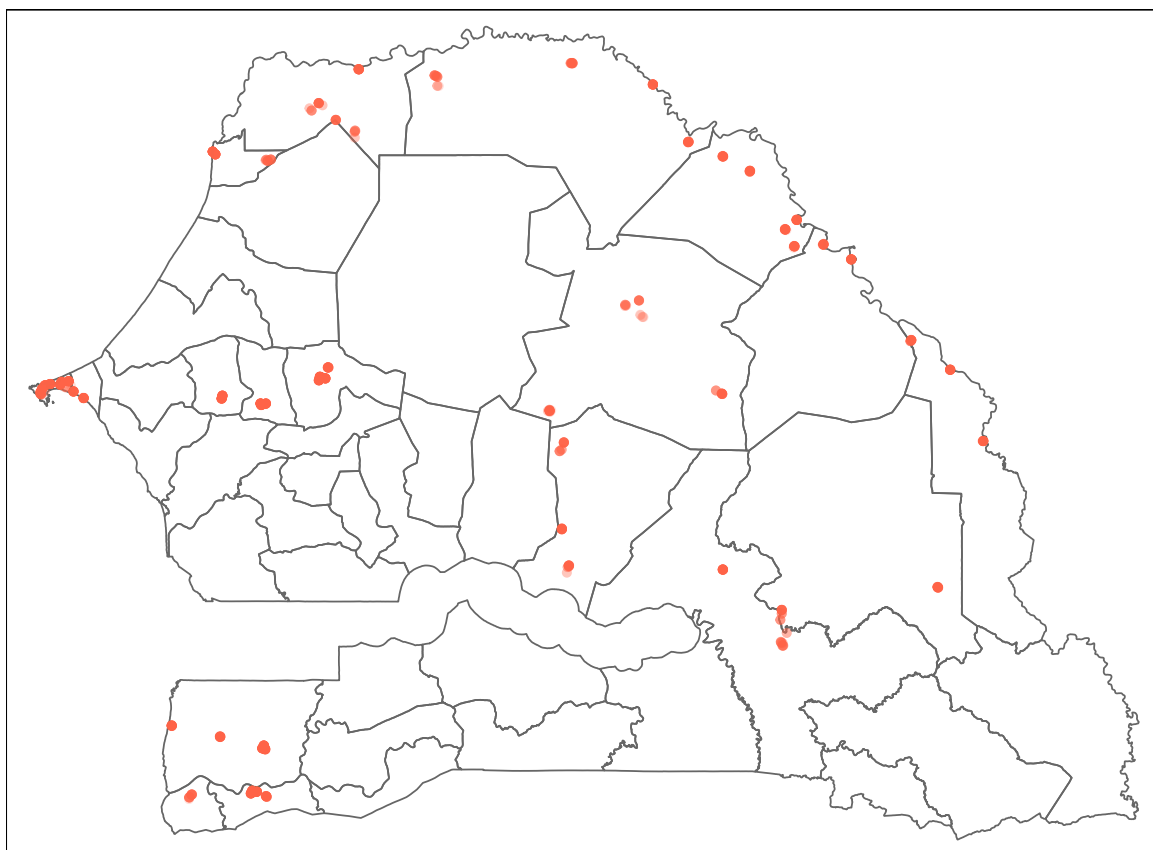


Figure 5.A2: Interview locations in Senegal



5.A.2 Figures

Figure 5.A3: Histograms for main dependent and explanatory variables

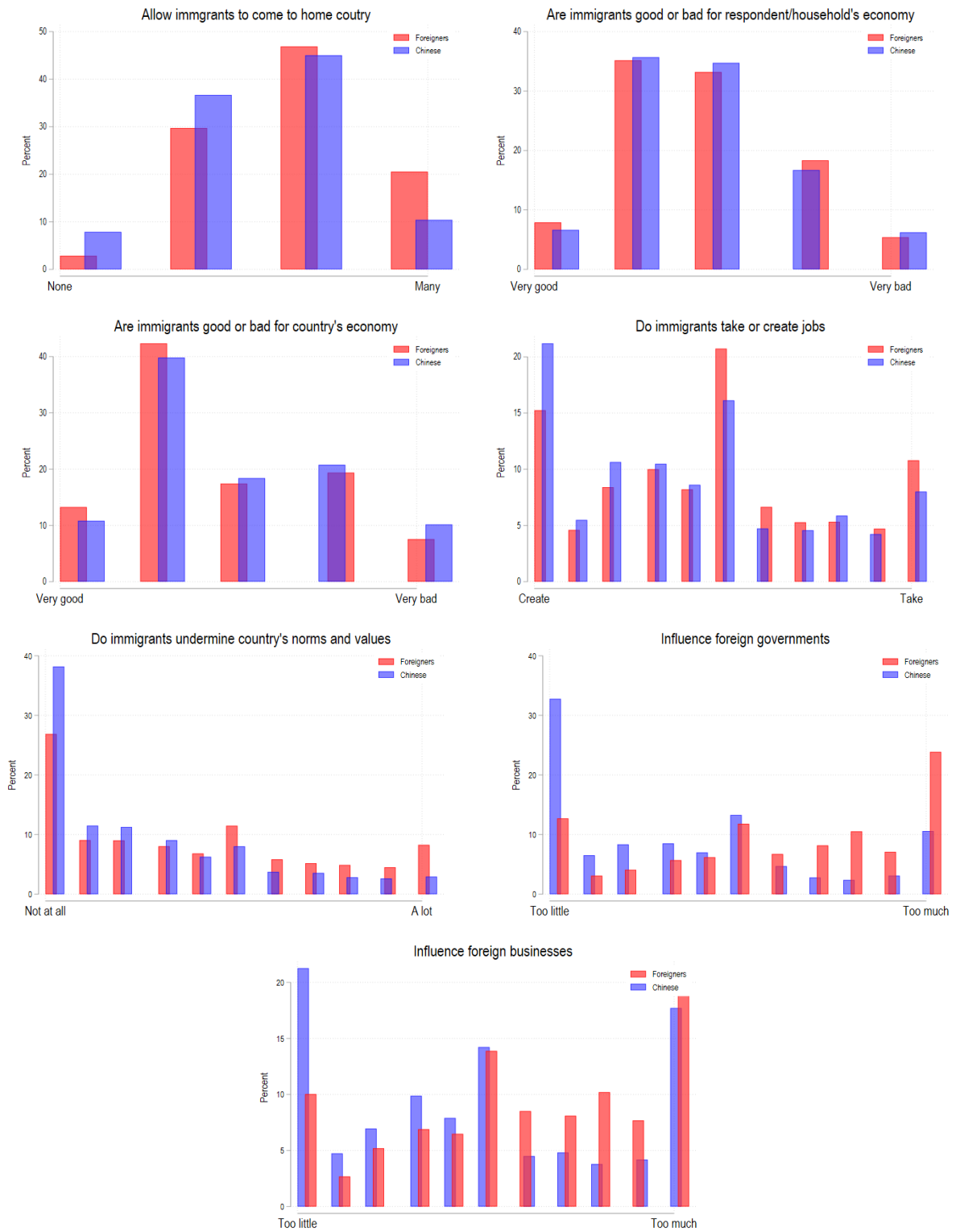
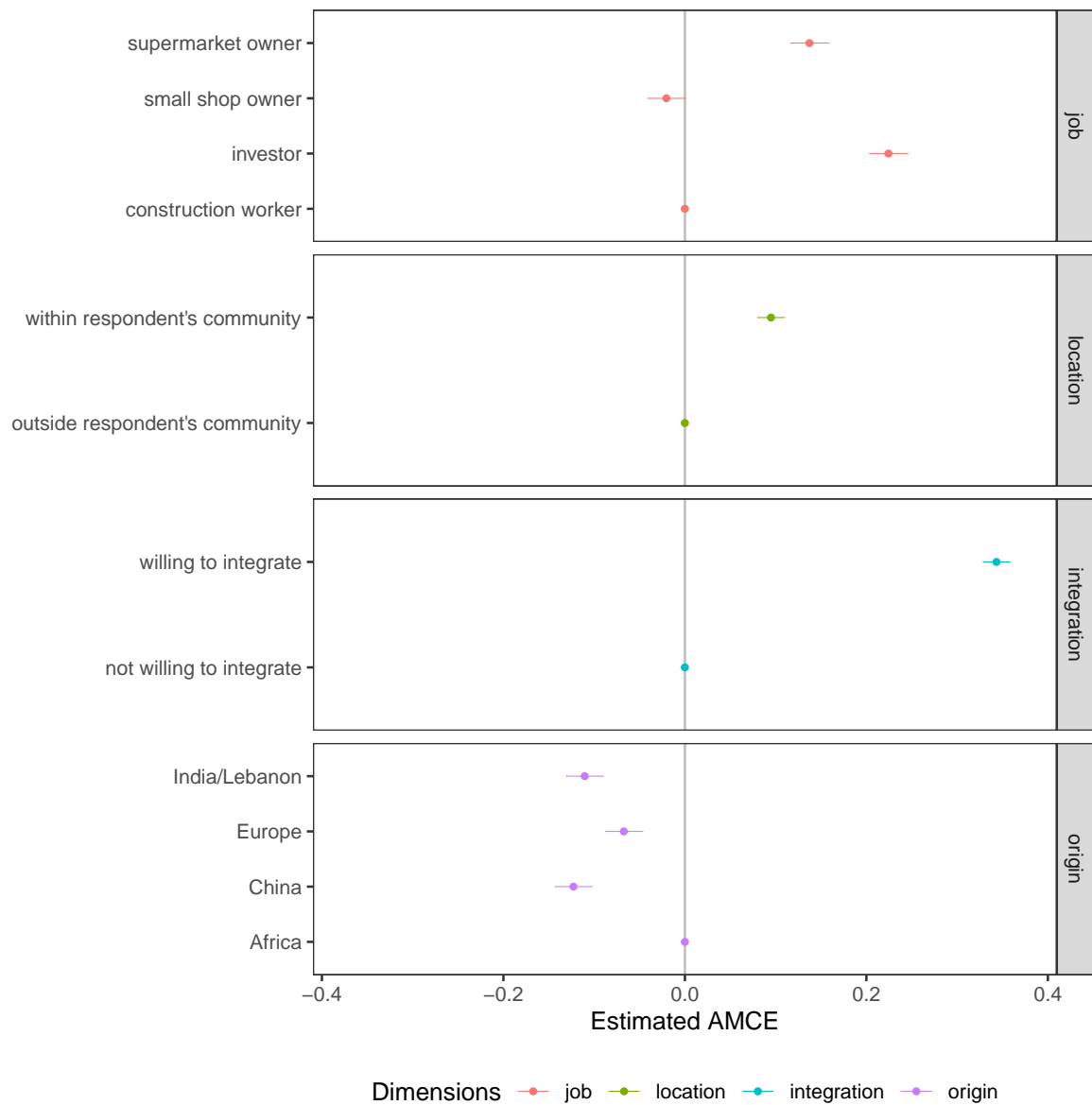
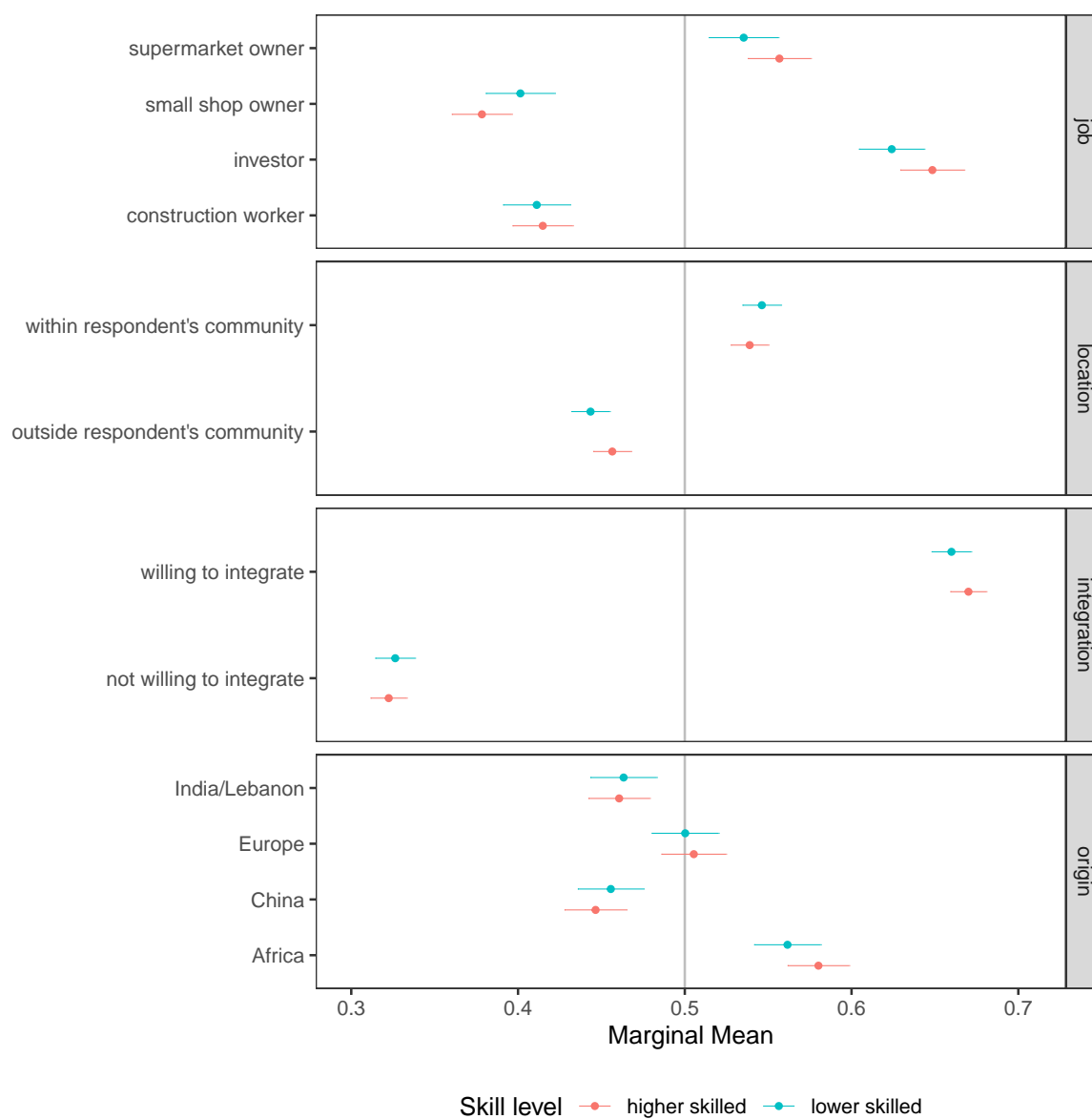


Figure 5.A4: Average marginal component effects for equation 5.1



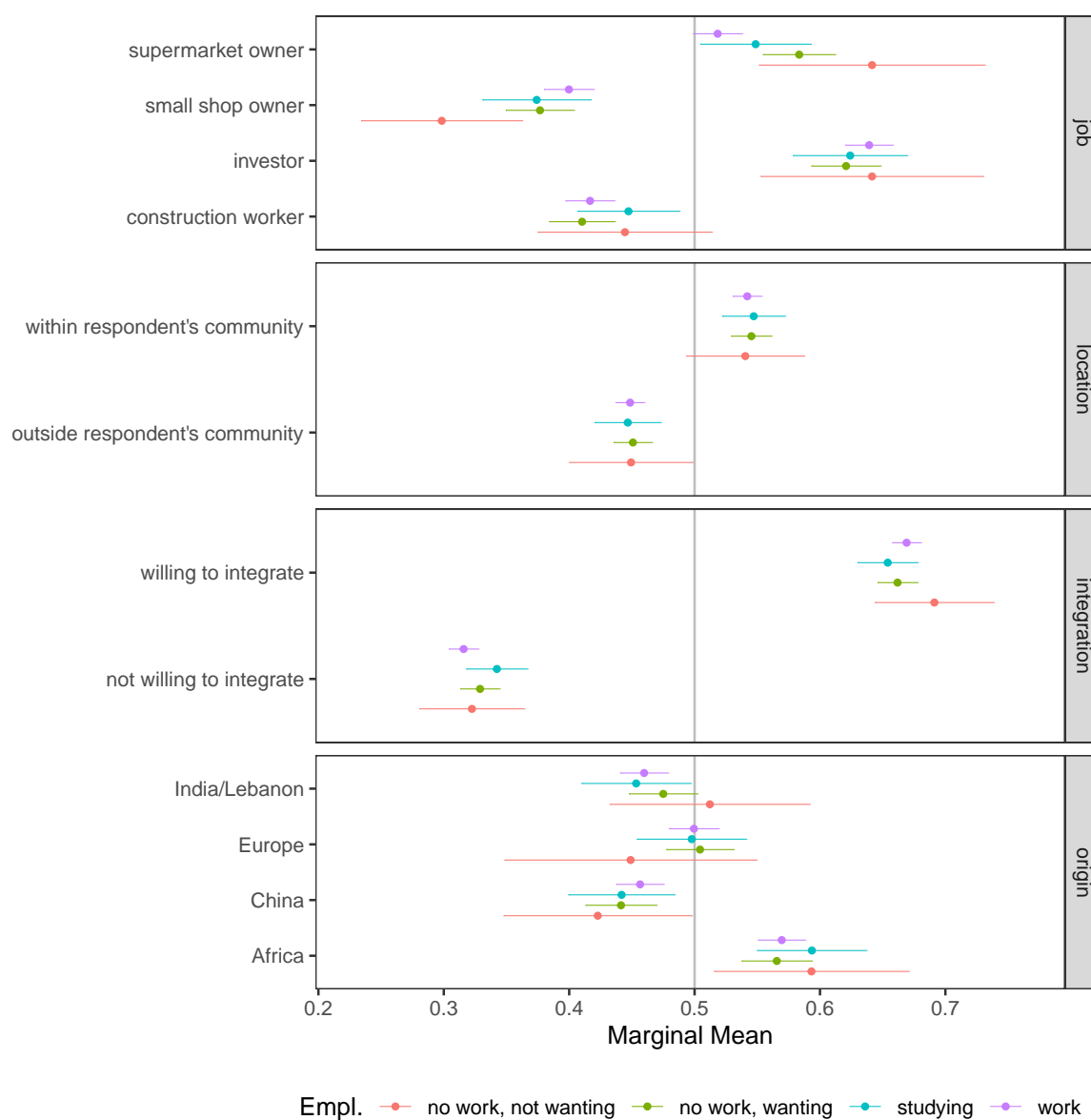
Note: The graph plots AMCEs for equation 5.1. Dots represent point estimates, whiskers around the dots represent 95% confidence intervals. Dots without whiskers represent the reference attribute level. Standard errors are clustered on the respondent level. *Don't know* or *Refuse to answer* are dropped from the estimation. The corresponding numbers can be found in Table 5.A2.

Figure 5.A5: Marginal means for subgroup analysis by skill level



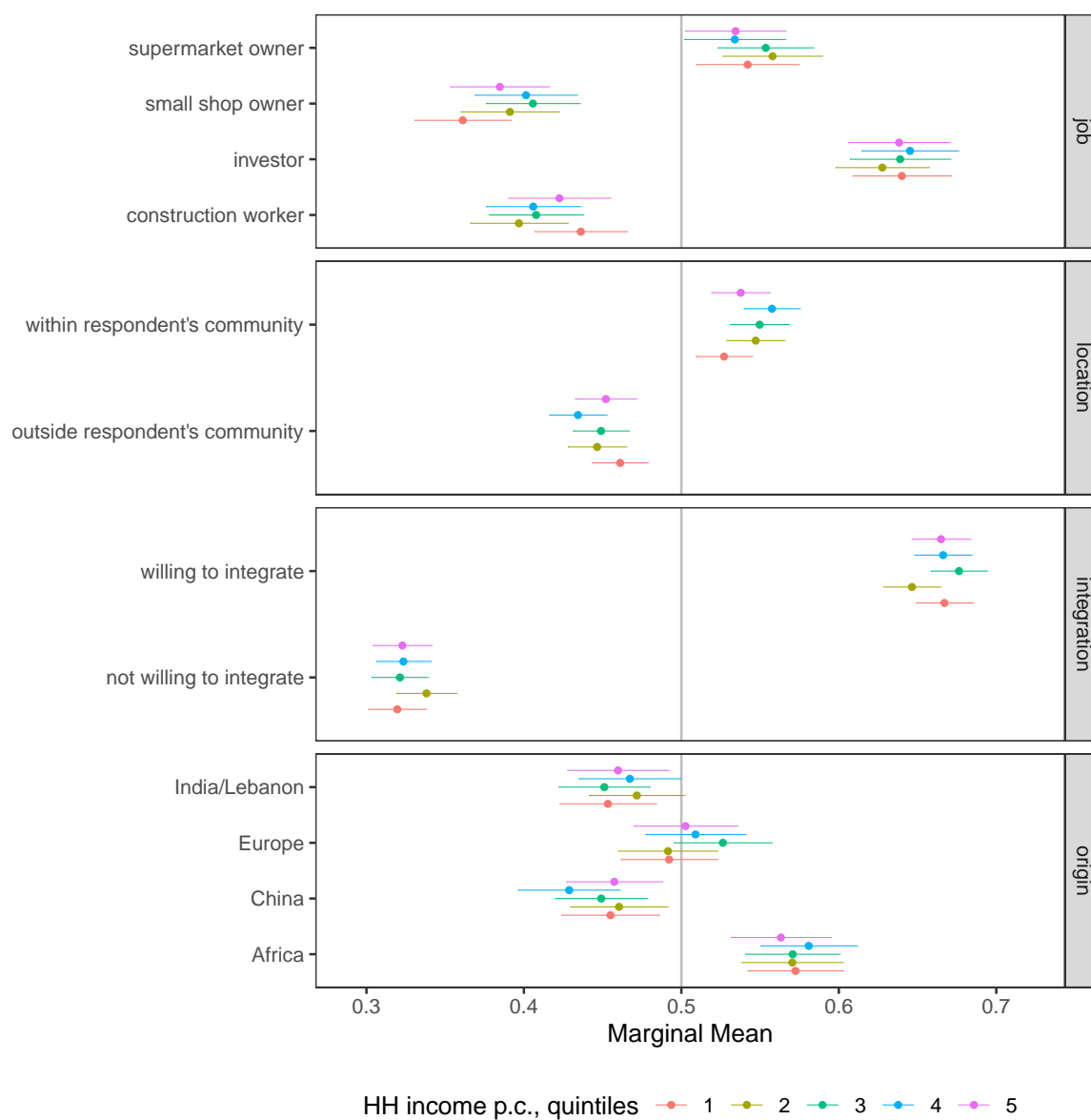
Note: The graph plots marginal means for subgroup analysis by skill level, measured by educational attainment. *Don't know* or *Refuse to answer* are dropped from the estimation. The corresponding numbers can be found in Table 5.A3.

Figure 5.A6: Marginal means for subgroup analysis by employment group



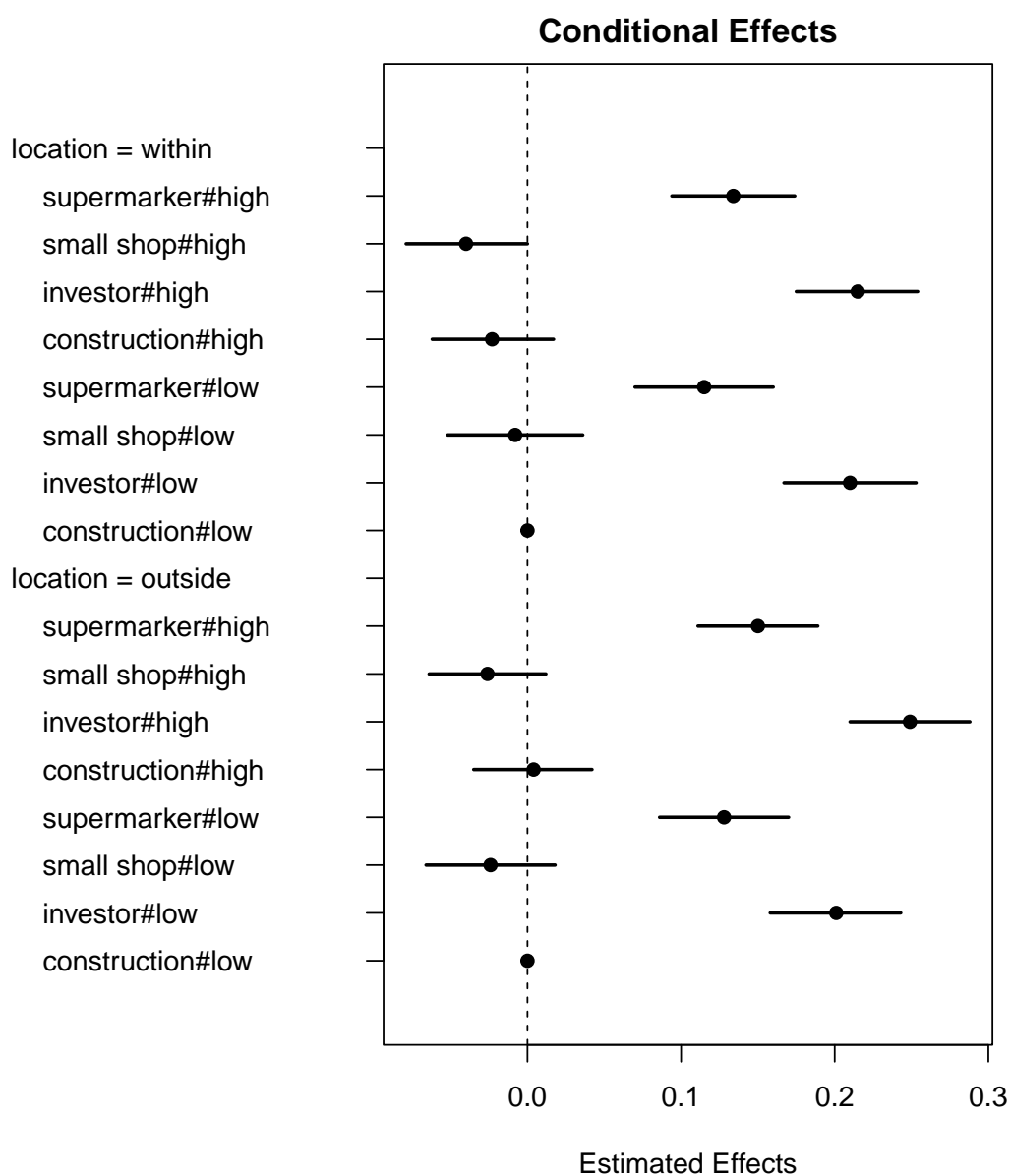
Note: The graph plots marginal means for subgroup analysis by employment group, based on the last 7 days before the interview. *Don't know* or *Refuse to answer* are dropped from the estimation. The corresponding numbers can be found in Table 5.A4.

Figure 5.A7: Marginal means for subgroup analysis by household income quintiles



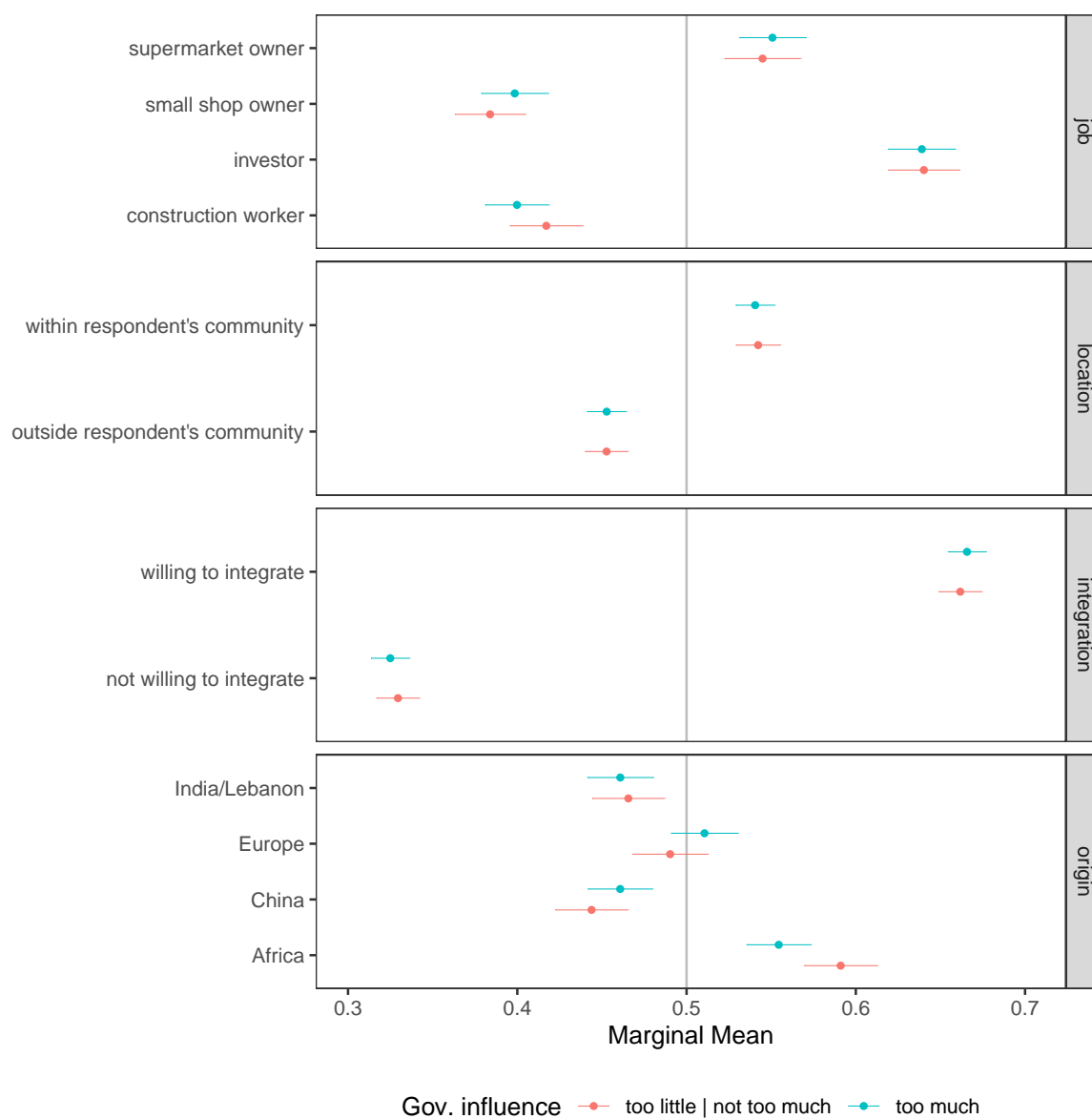
Note: The graph plots marginal means for subgroup analysis by household income per capita quintiles, based on the households' reported income during the last 12 months. The first quintile represents the bottom 20% of the income distribution. *Don't know* or *Refuse to answer* are dropped from the estimation. The corresponding numbers can be found in Table 5.A5.

Figure 5.A8: AMIE for skill level and location



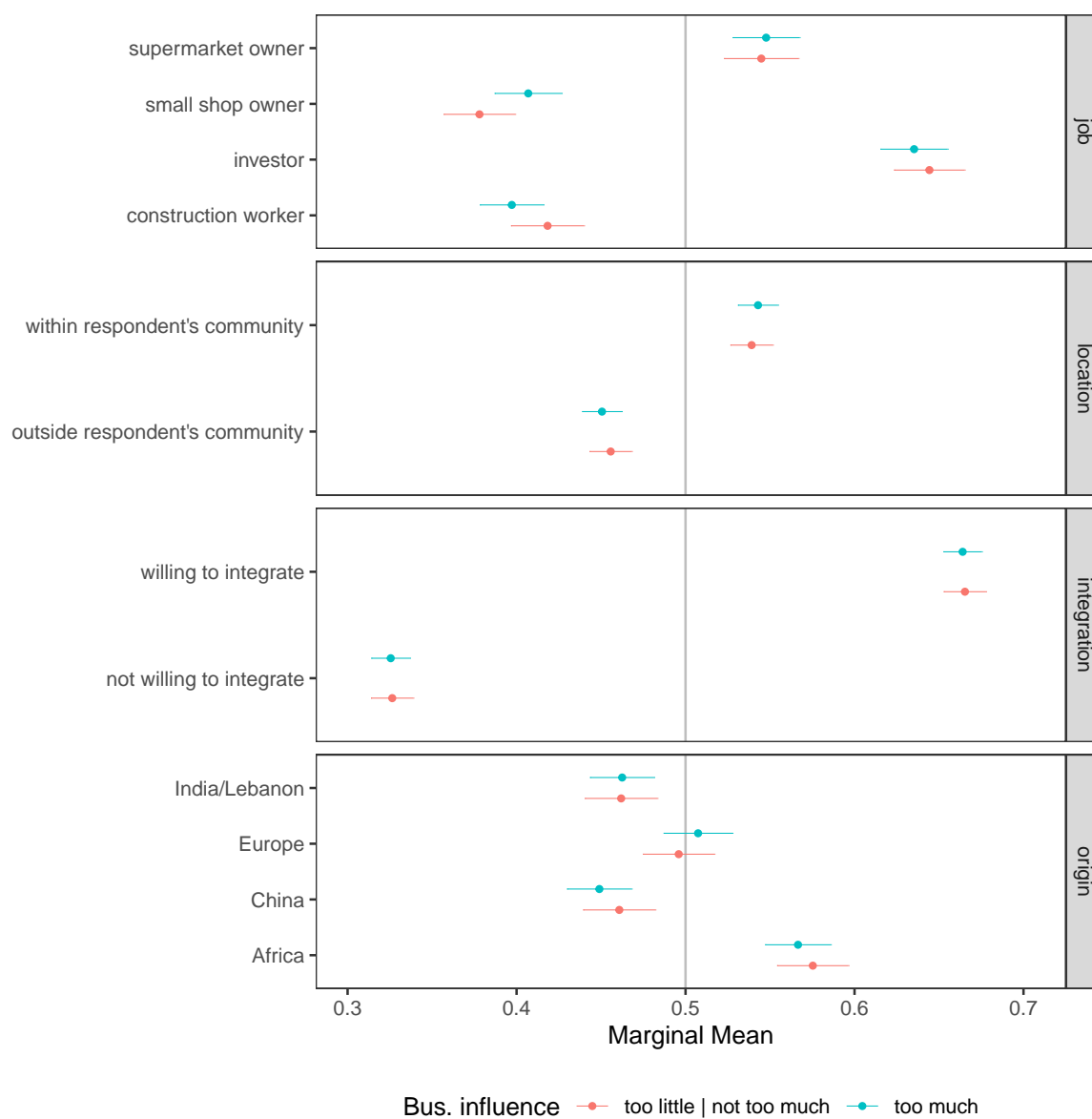
Note: The graph plots Average Marginal Interaction Effects for immigrants' job, immigrants' location, and respondents' skill level, measured by educational attainment. *Don't know* or *Refuse to answer* are dropped from the estimation.

Figure 5.A9: MMs for equation 5.1 by concerns about foreign governments' influence



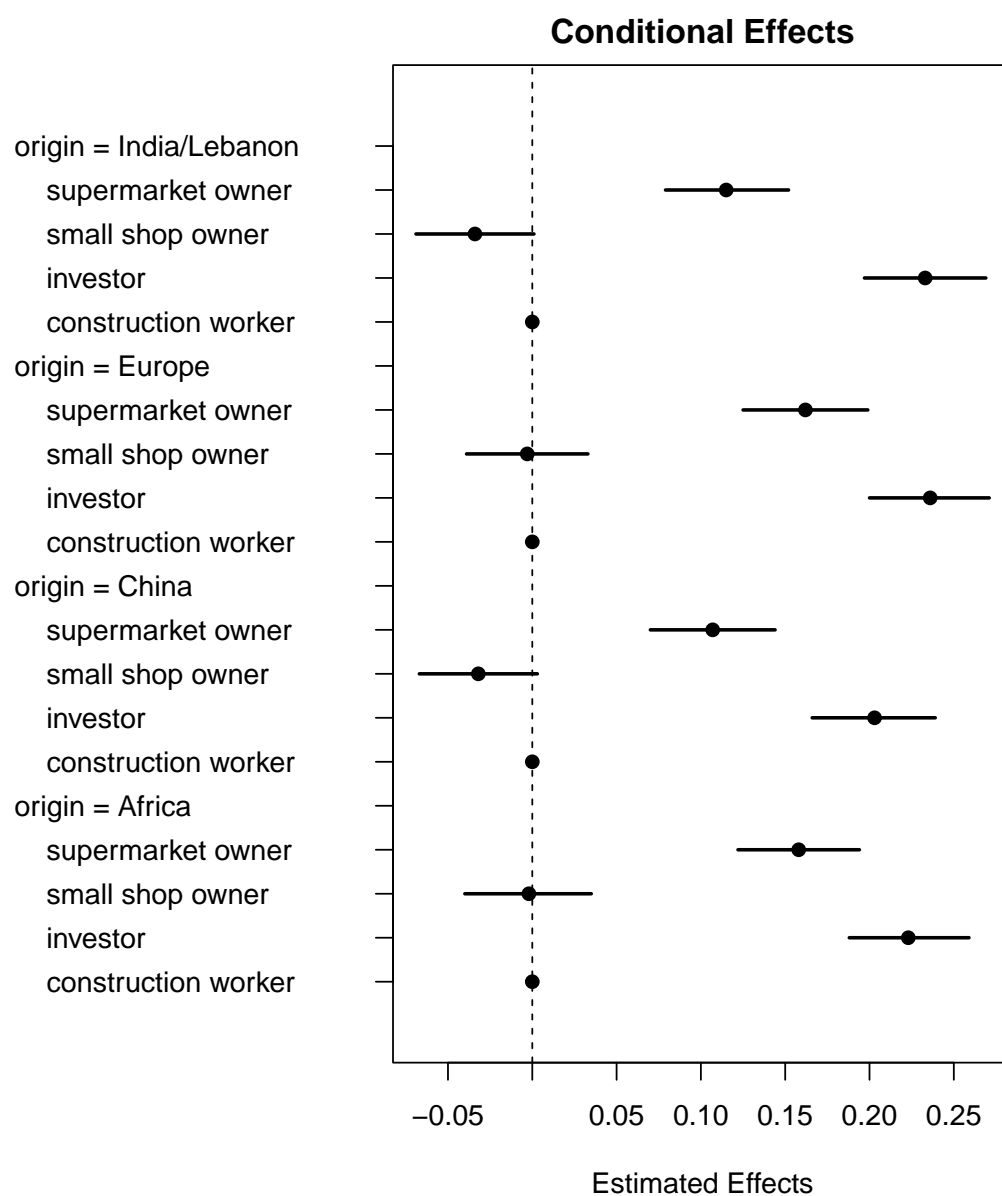
Note: The graph shows marginal means for equation 5.1 by respondents' concerns about foreign governments' influence. *Don't know* or *Refuse to answer* are dropped from the estimation.

Figure 5.A10: MMs for equation 5.1 by concerns about foreign businesses' influence



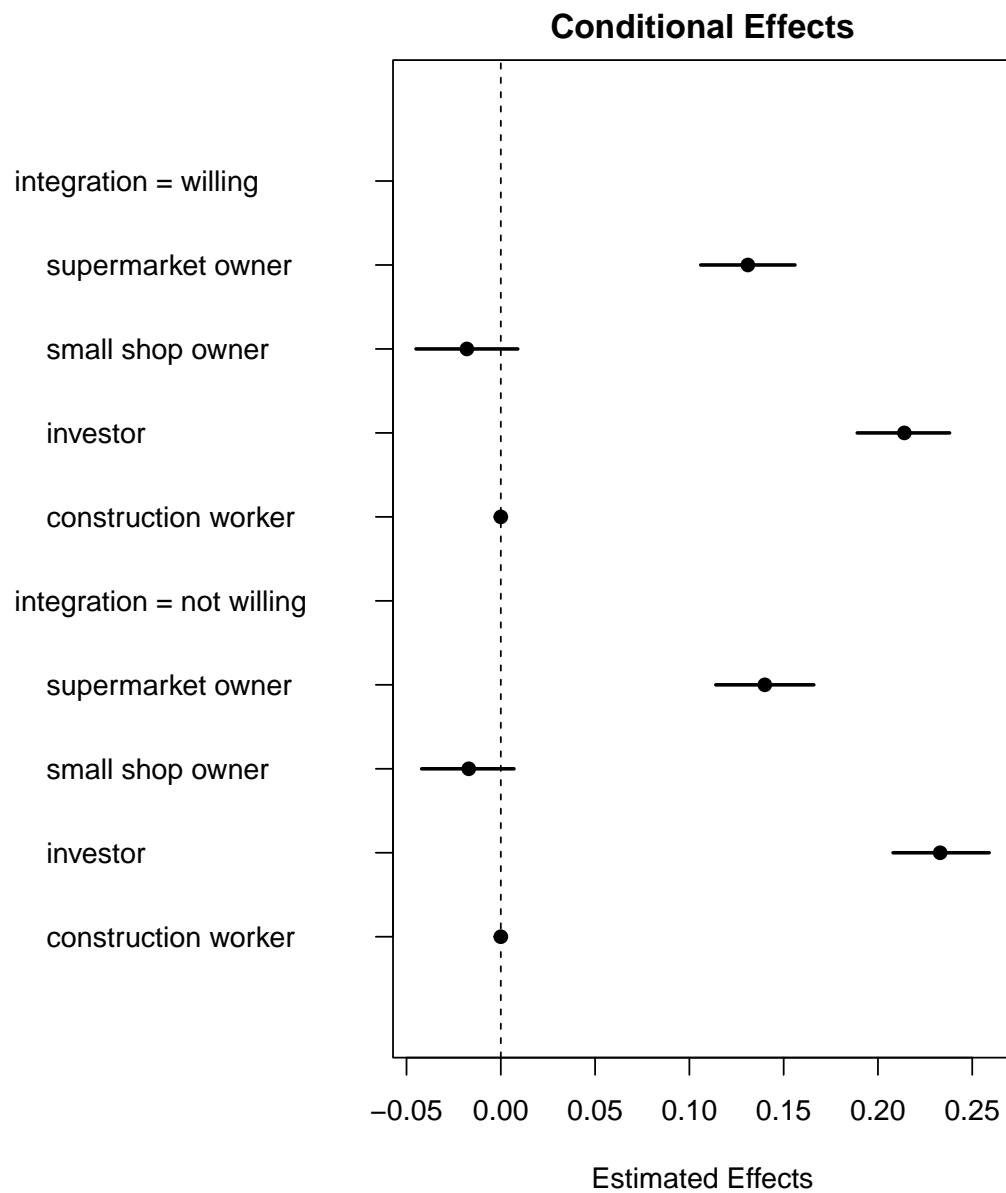
Note: The graph shows marginal means for equation 5.1 by respondents' concerns about foreign businesses' influence. *Don't know* or *Refuse to answer* are dropped from the estimation.

Figure 5.A11: AMIE for job and location



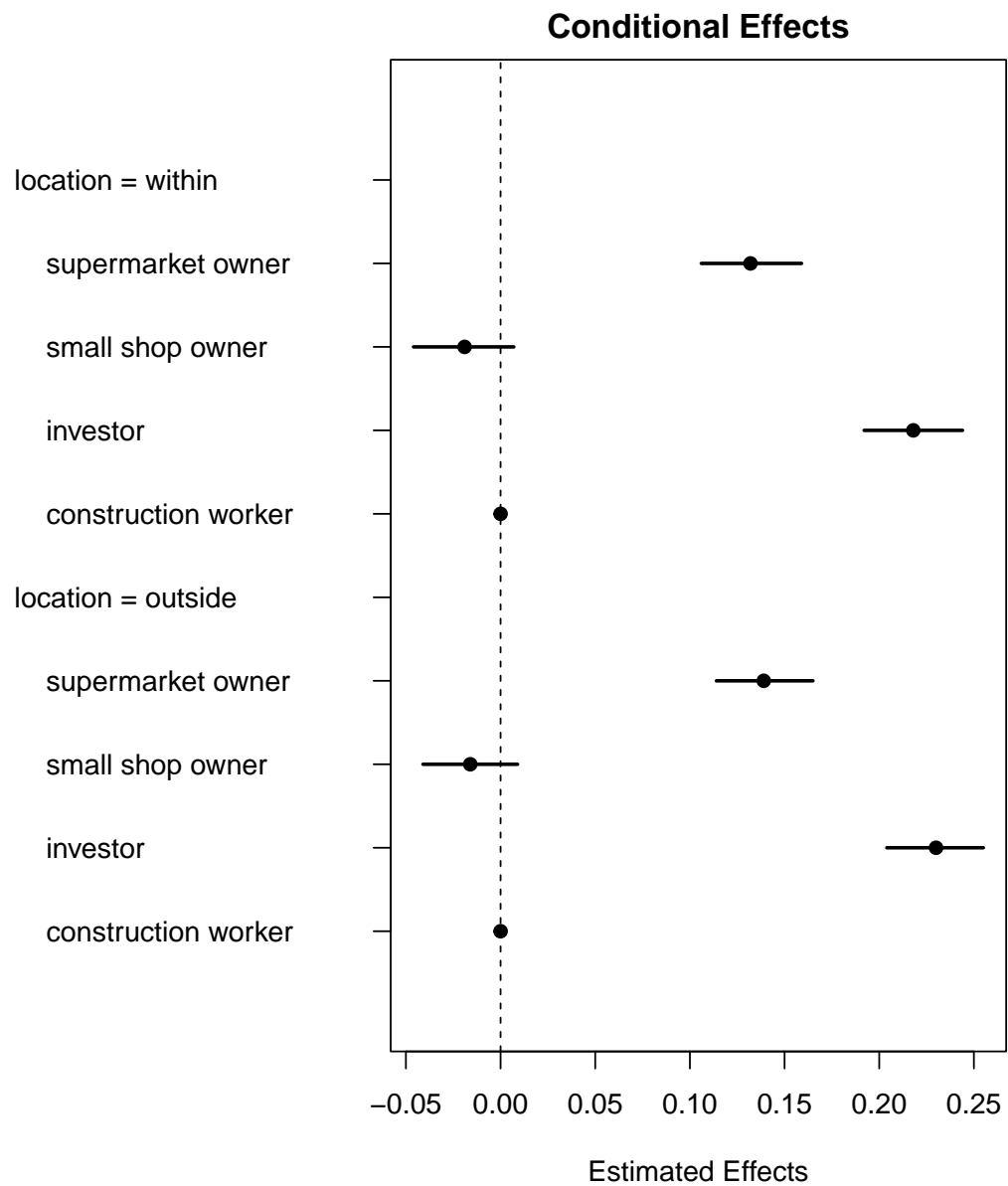
Note: The graph plots Average Marginal Interaction Effects for immigrants' job and immigrants' location. *Don't know* or *Refuse to answer* are dropped from the estimation.

Figure 5.A12: AMIE for job and willingness to integrate



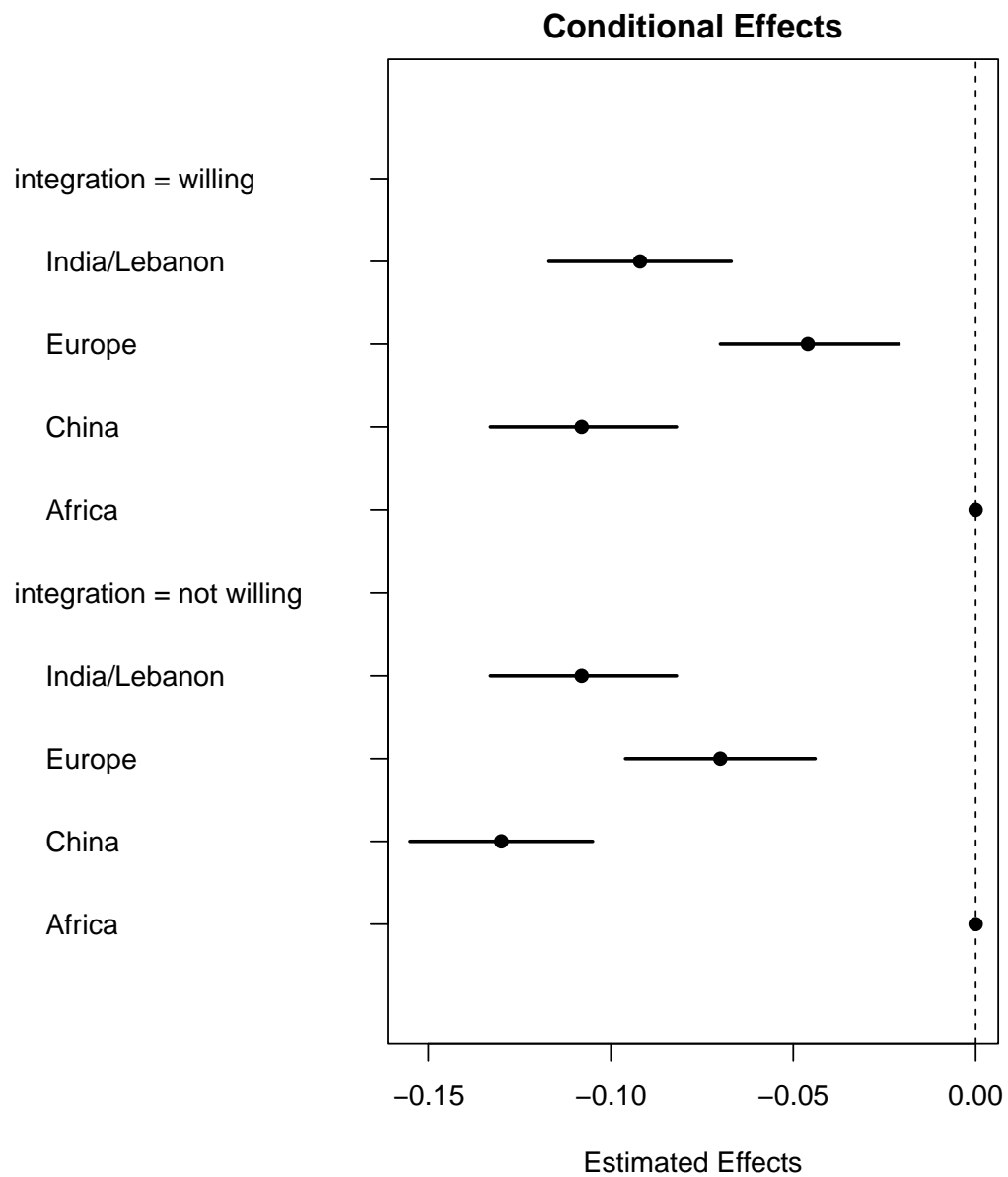
Note: The graph plots Average Marginal Interaction Effects for immigrants' job and immigrants' willingness to integrate. *Don't know* or *Refuse to answer* are dropped from the estimation.

Figure 5.A13: AMIE for job and location



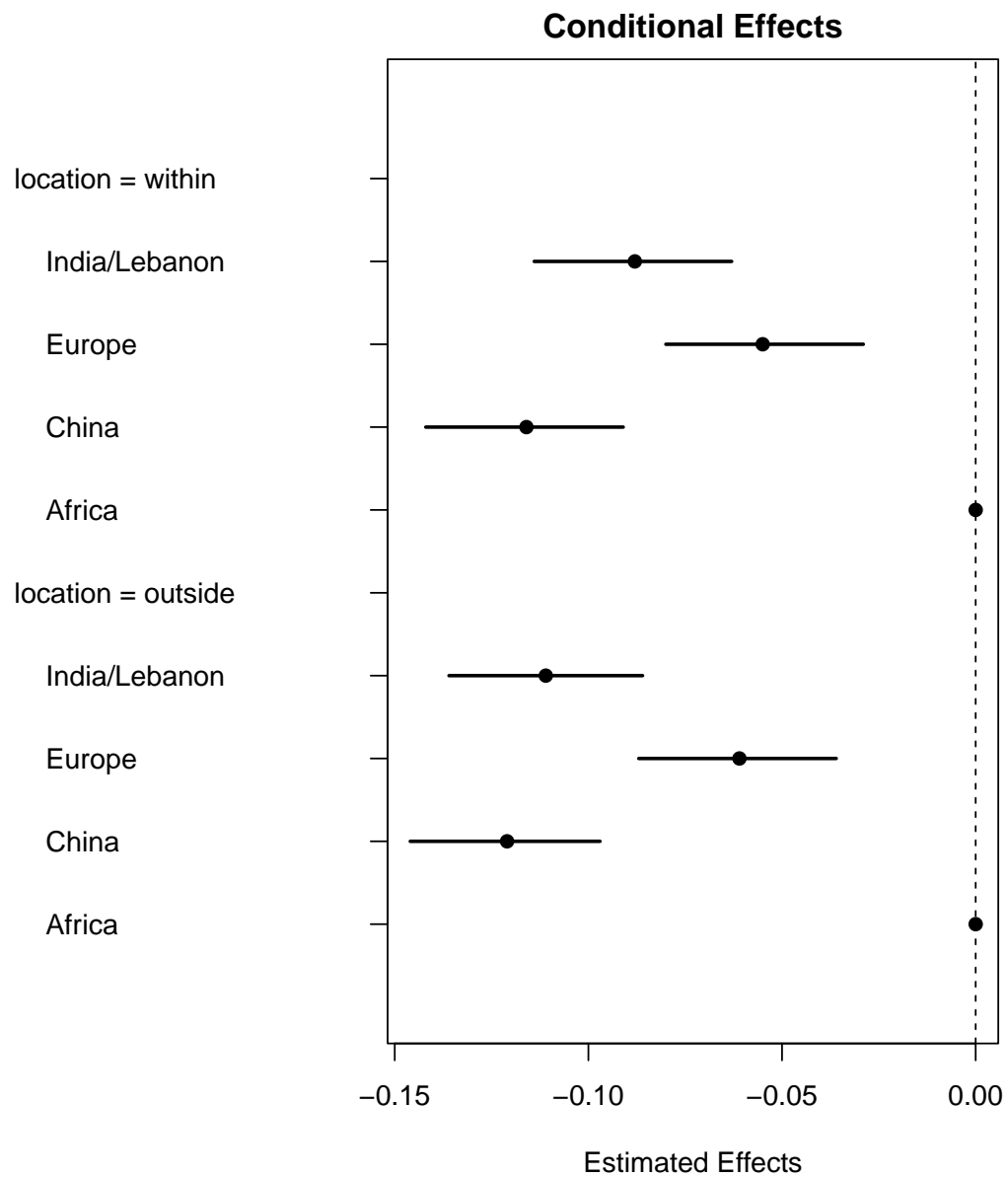
Note: The graph plots Average Marginal Interaction Effects for immigrants' job and immigrants' location. *Don't know* or *Refuse to answer* are dropped from the estimation.

Figure 5.A14: AMIE for origin and willingness to integrate



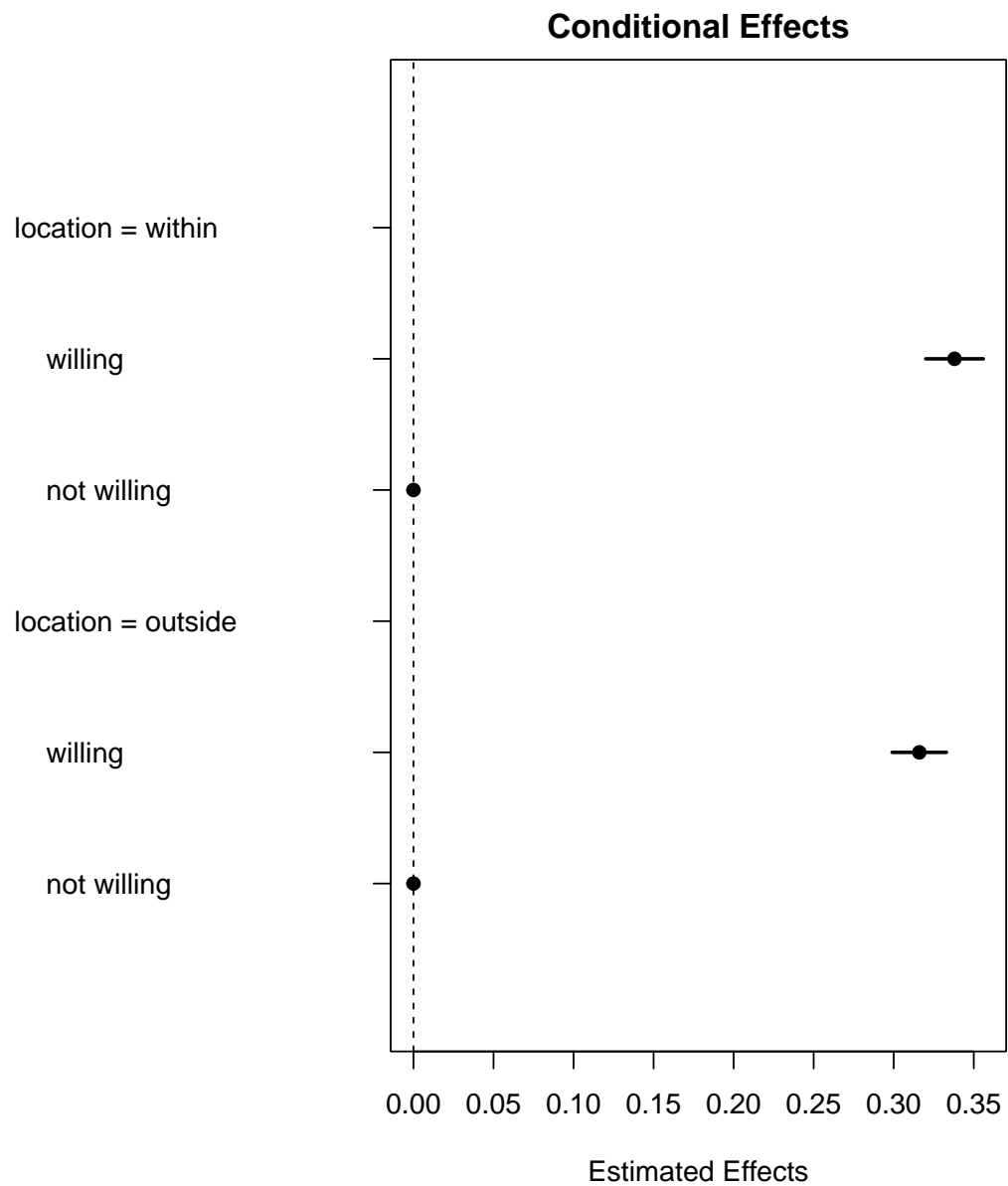
Note: The graph plots Average Marginal Interaction Effects for immigrants' origin and immigrants' willingness to integrate. *Don't know* or *Refuse to answer* are dropped from the estimation.

Figure 5.A15: AMIE for origin and location



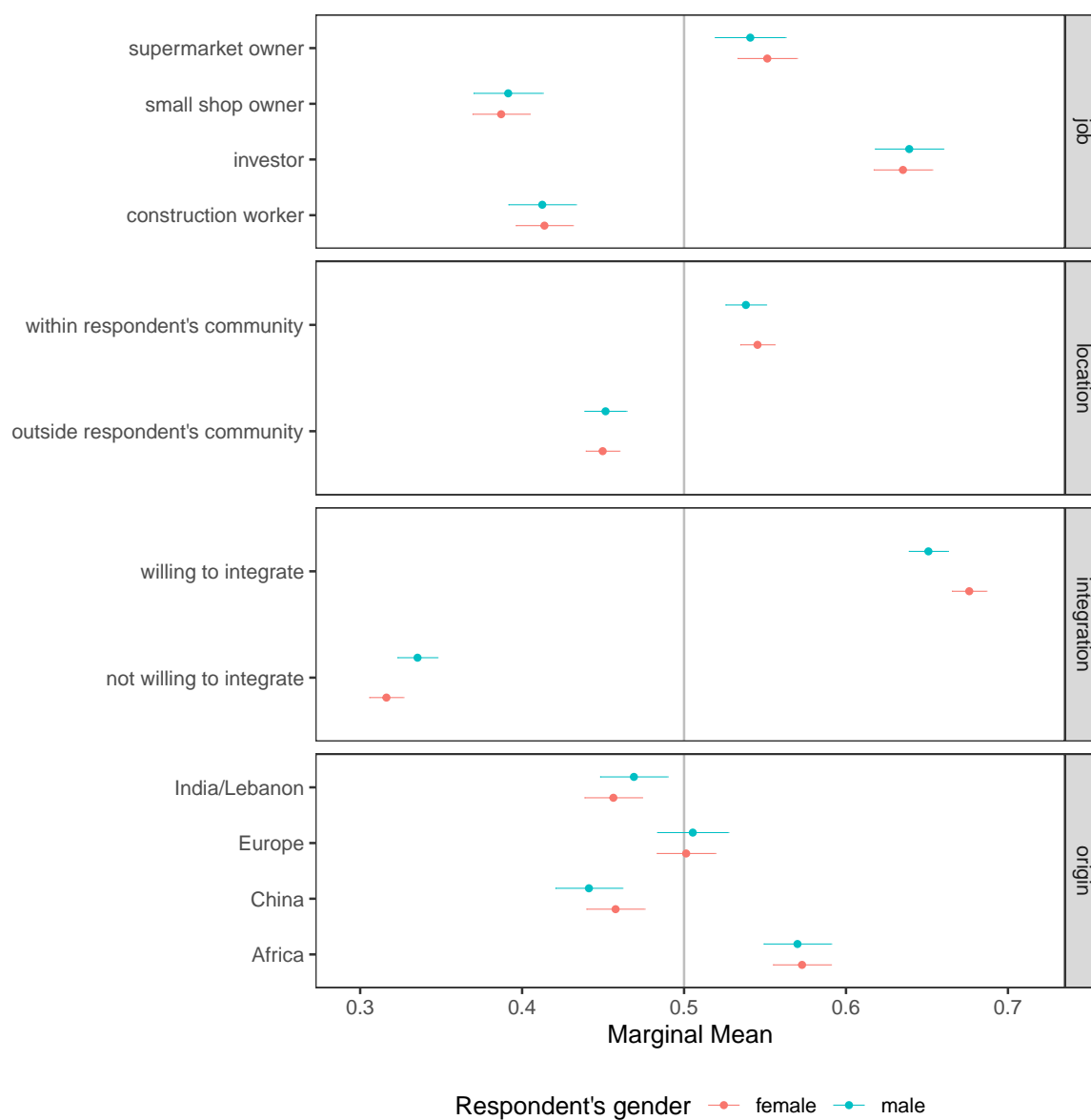
Note: The graph plots Average Marginal Interaction Effects for immigrants' origin and immigrants' location. *Don't know* or *Refuse to answer* are dropped from the estimation.

Figure 5.A16: AMIE for willingness to integrate and location



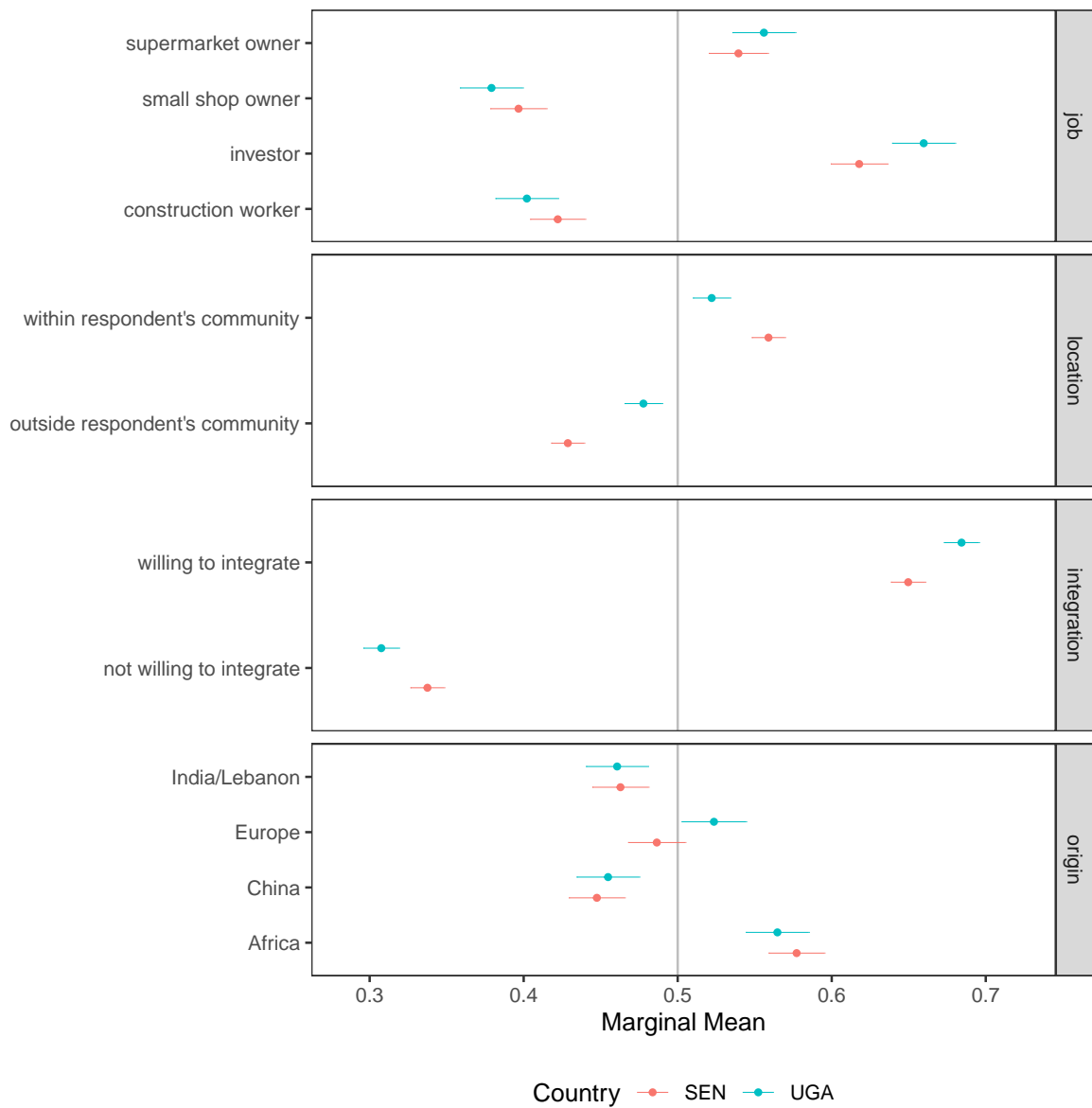
Note: The graph plots Average Marginal Interaction Effects for immigrants' willingness to integrate and immigrants' location. *Don't know* or *Refuse to answer* are dropped from the estimation.

Figure 5.A17: MMs for equation 5.1 by respondent's gender



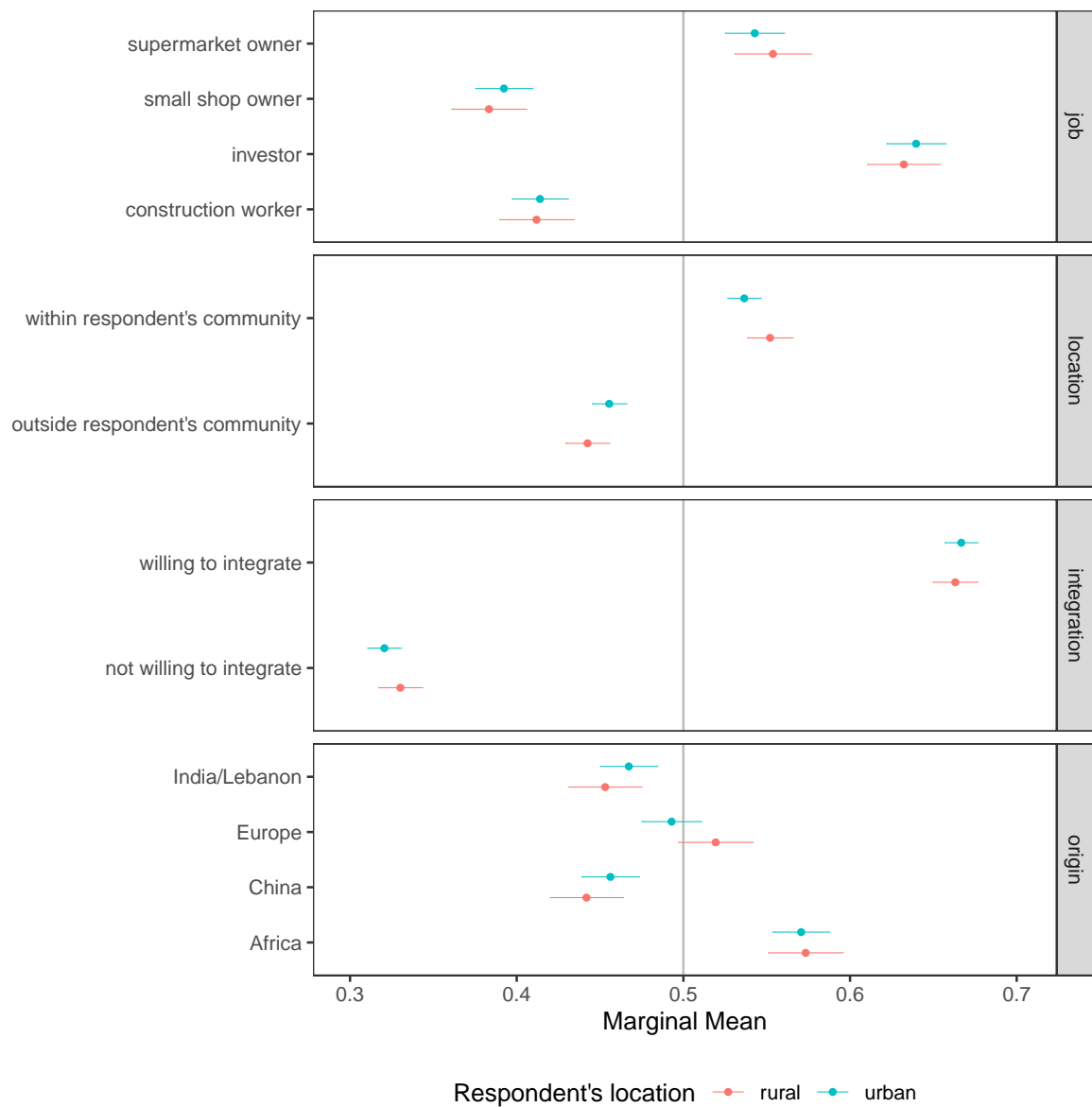
Note: The graph shows marginal means for equation 5.1 by respondent's gender. *Don't know* or *Refuse to answer* are dropped from the estimation.

Figure 5.A18: MMs for equation 5.1 by country



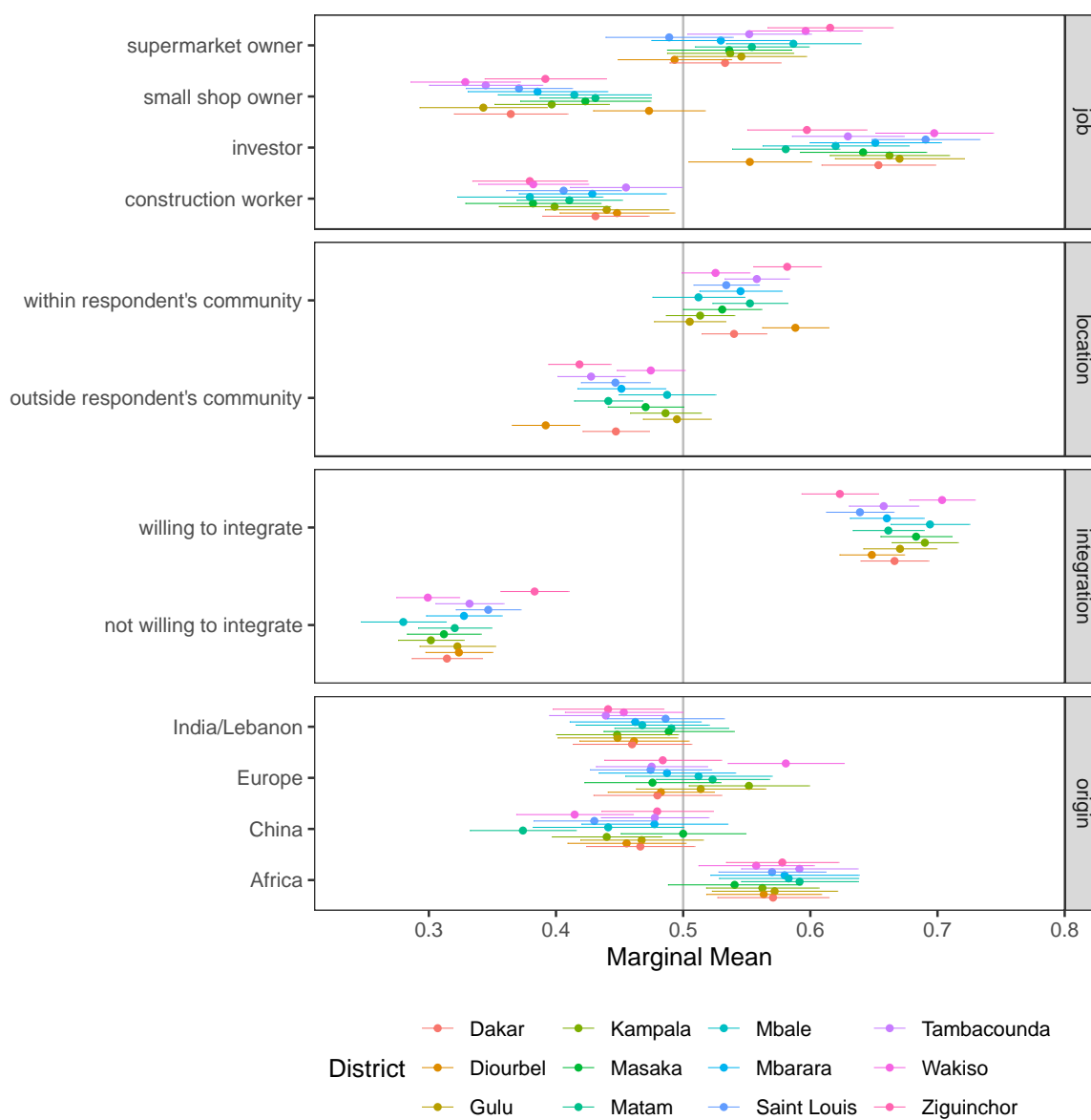
Note: The graph shows marginal means for equation 5.1 by country. Don't know or Refuse to answer are dropped from the estimation.

Figure 5.A19: MMs for equation 5.1 by respondent's location



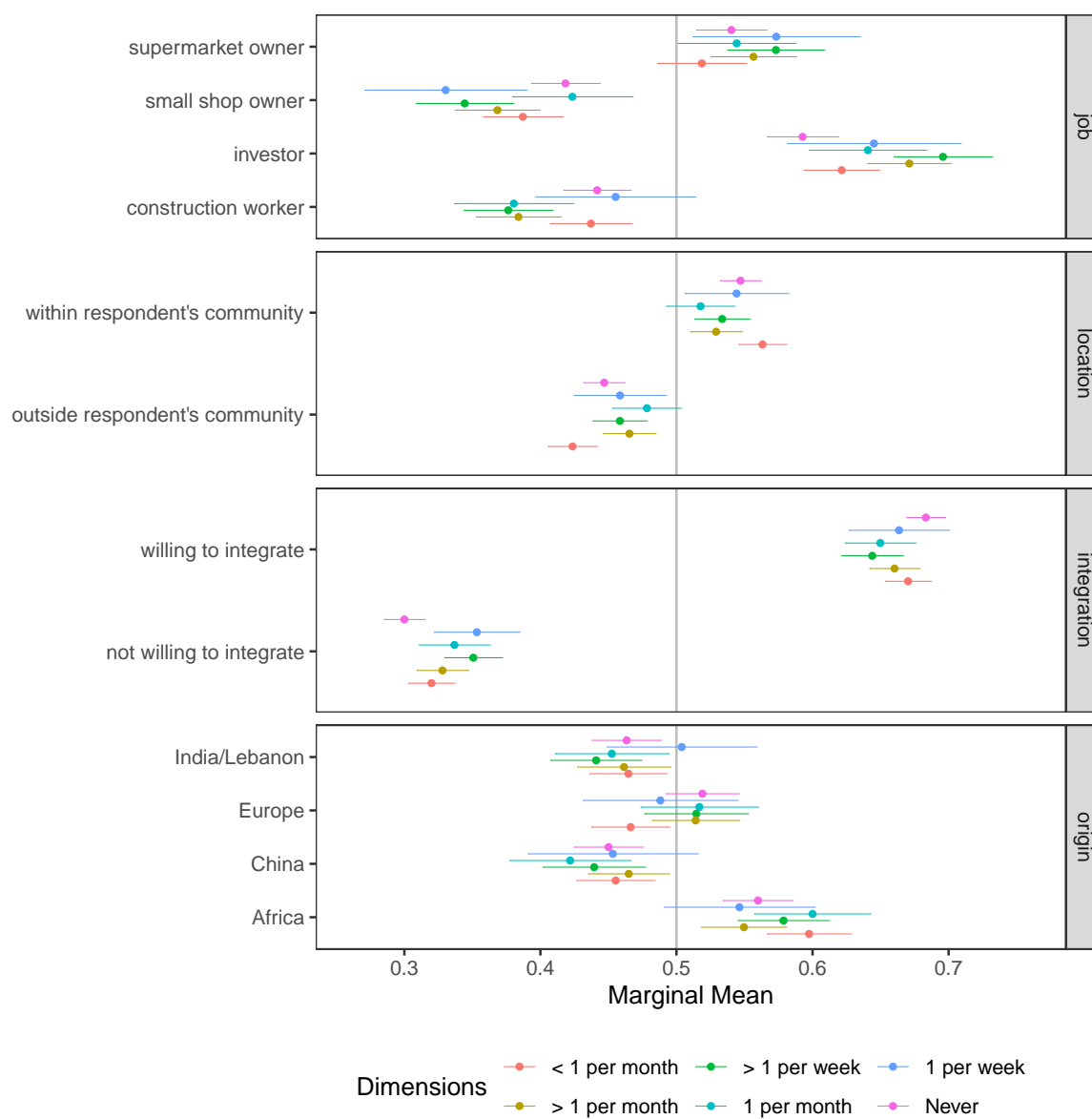
Note: The graph shows marginal means for equation 5.1 by respondent's location (rural-urban). *Don't know* or *Refuse to answer* are dropped from the estimation.

Figure 5.A20: MMs for equation 5.1 by district



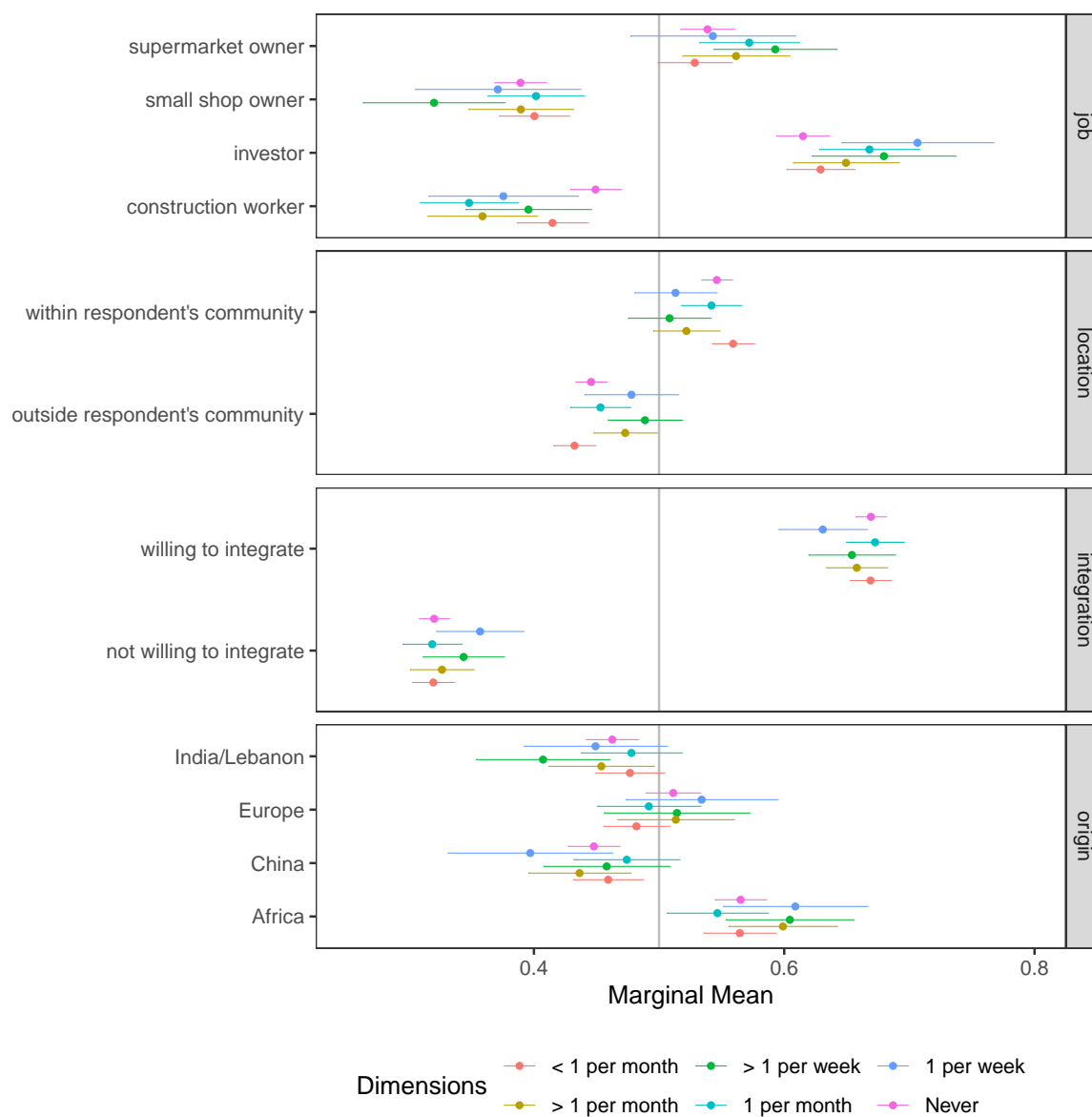
Note: The graph shows marginal means for equation 5.1 by respondents' district of residence. *Don't know* or *Refuse to answer* are dropped from the estimation.

Figure 5.A21: MMs for equation 5.1 by self-reported contact with immigrants in general



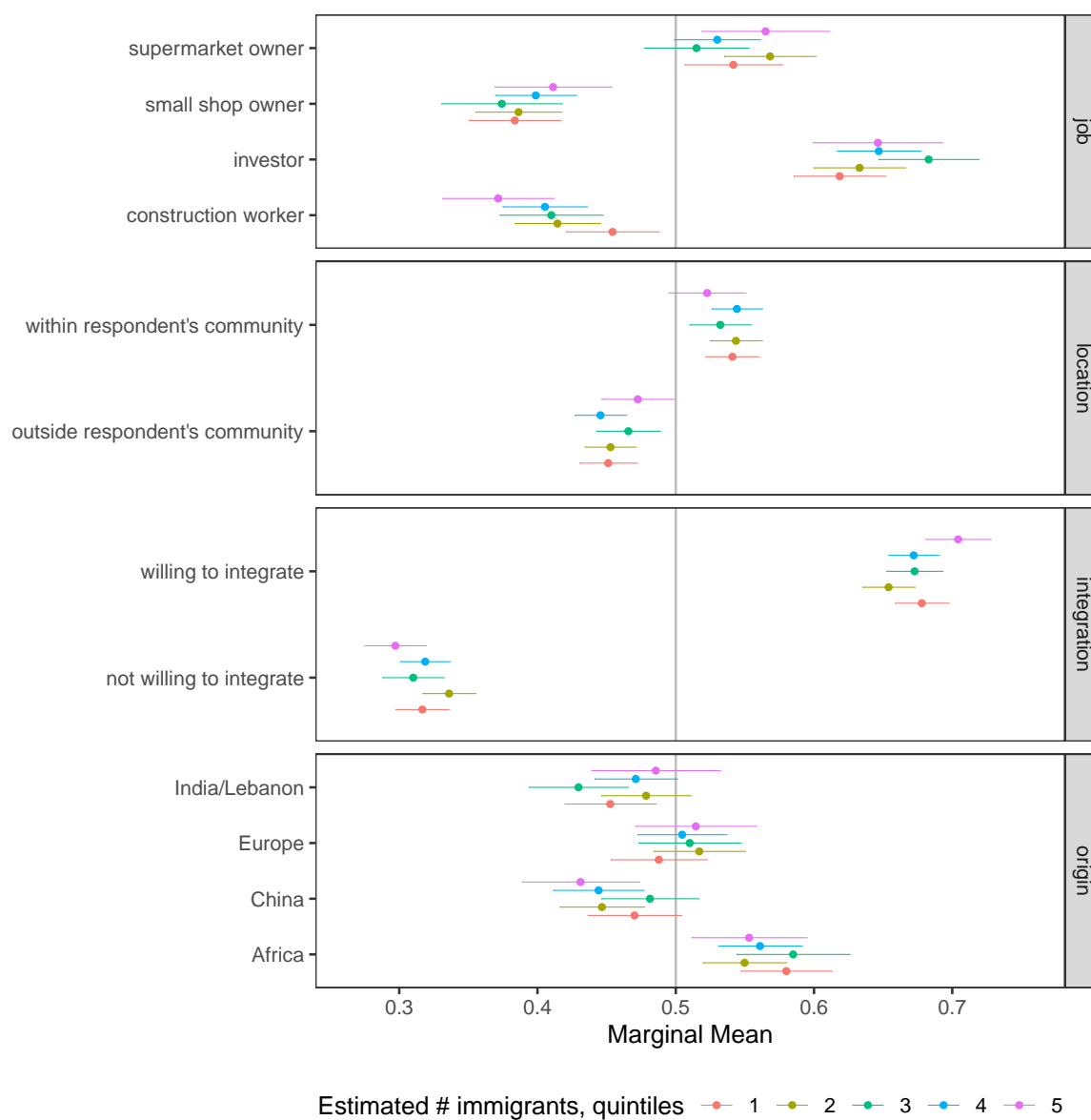
Note: The graph shows marginal means for equation 5.1 by respondents' self-reported contact with immigrants in general. Don't know or Refuse to answer are dropped from the estimation.

Figure 5.A22: MMs for equation 5.1 by self-reported contact with with Chinese immigrants



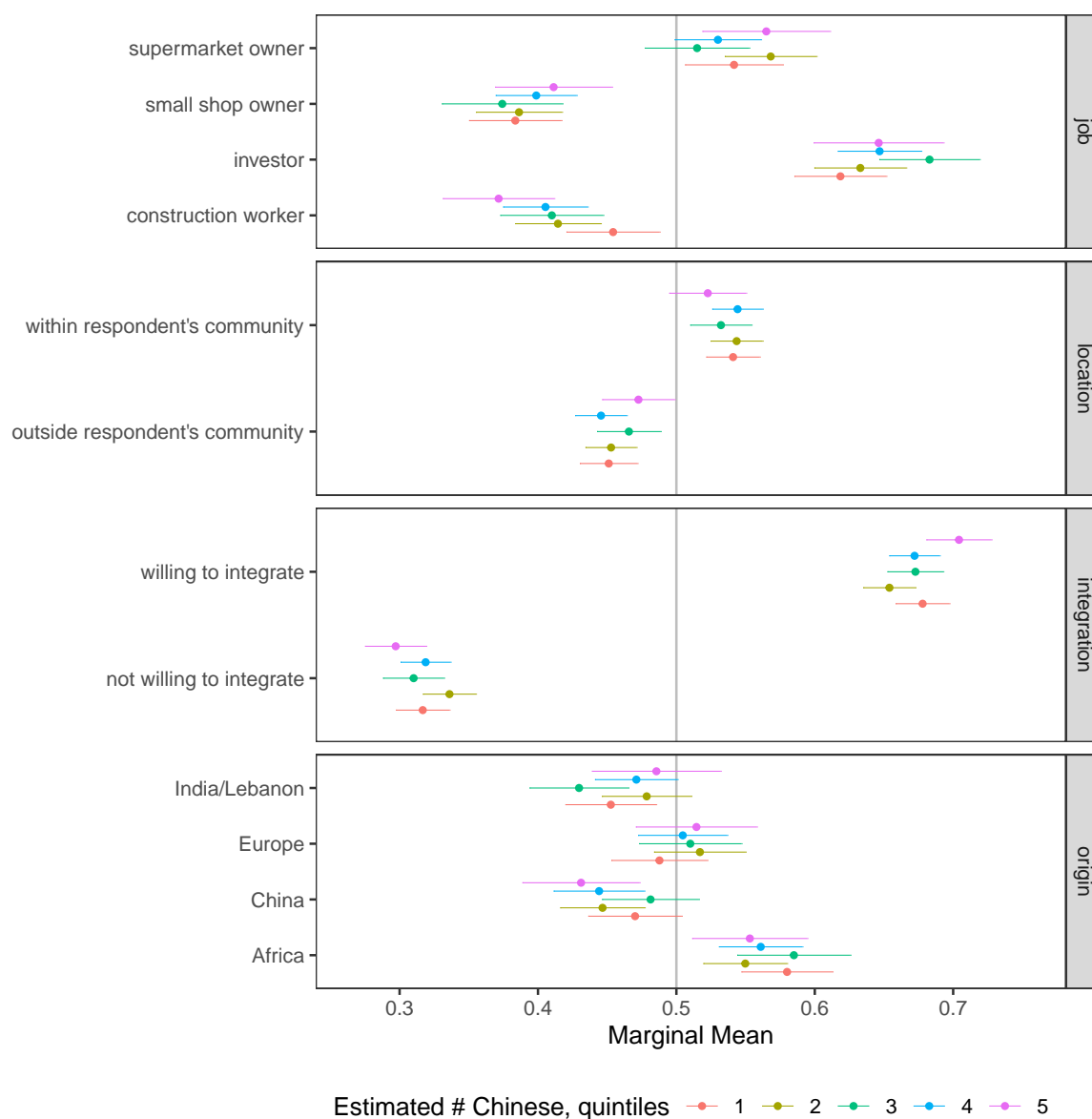
Note: The graph shows marginal means for equation 5.1 by respondents' self-reported contact with Chinese immigrants. Don't know or Refuse to answer are dropped from the estimation.

Figure 5.A23: MMs for equation 5.1 by estimated number of immigrants in general



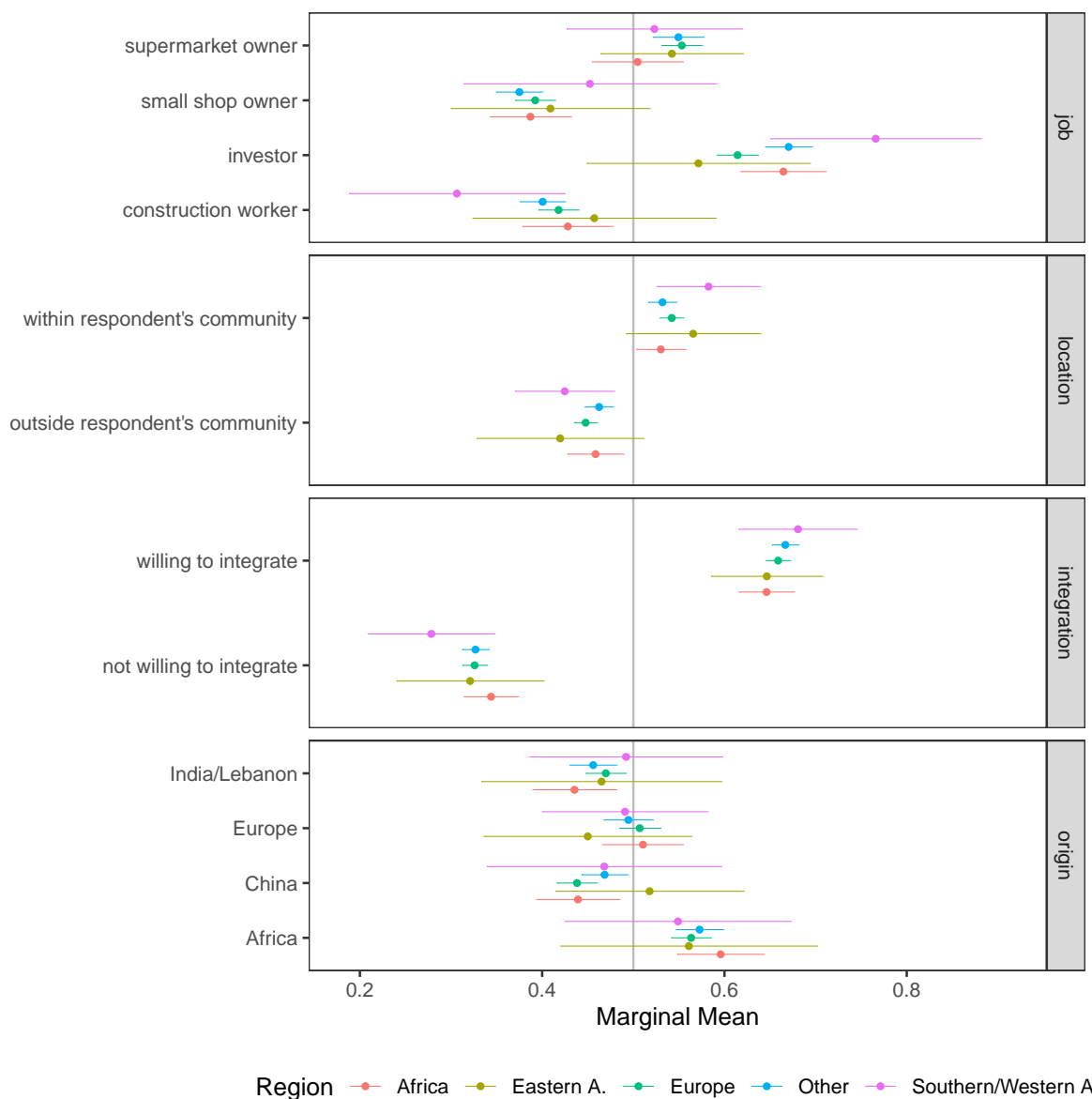
Note: The graph shows marginal means for equation 5.1 by country-specific quintiles of the estimated number of immigrants in general in respondents' home region. *Don't know* or *Refuse to answer* are dropped from the estimation.

Figure 5.A24: MMs for equation 5.1 by estimated number of Chinese immigrants



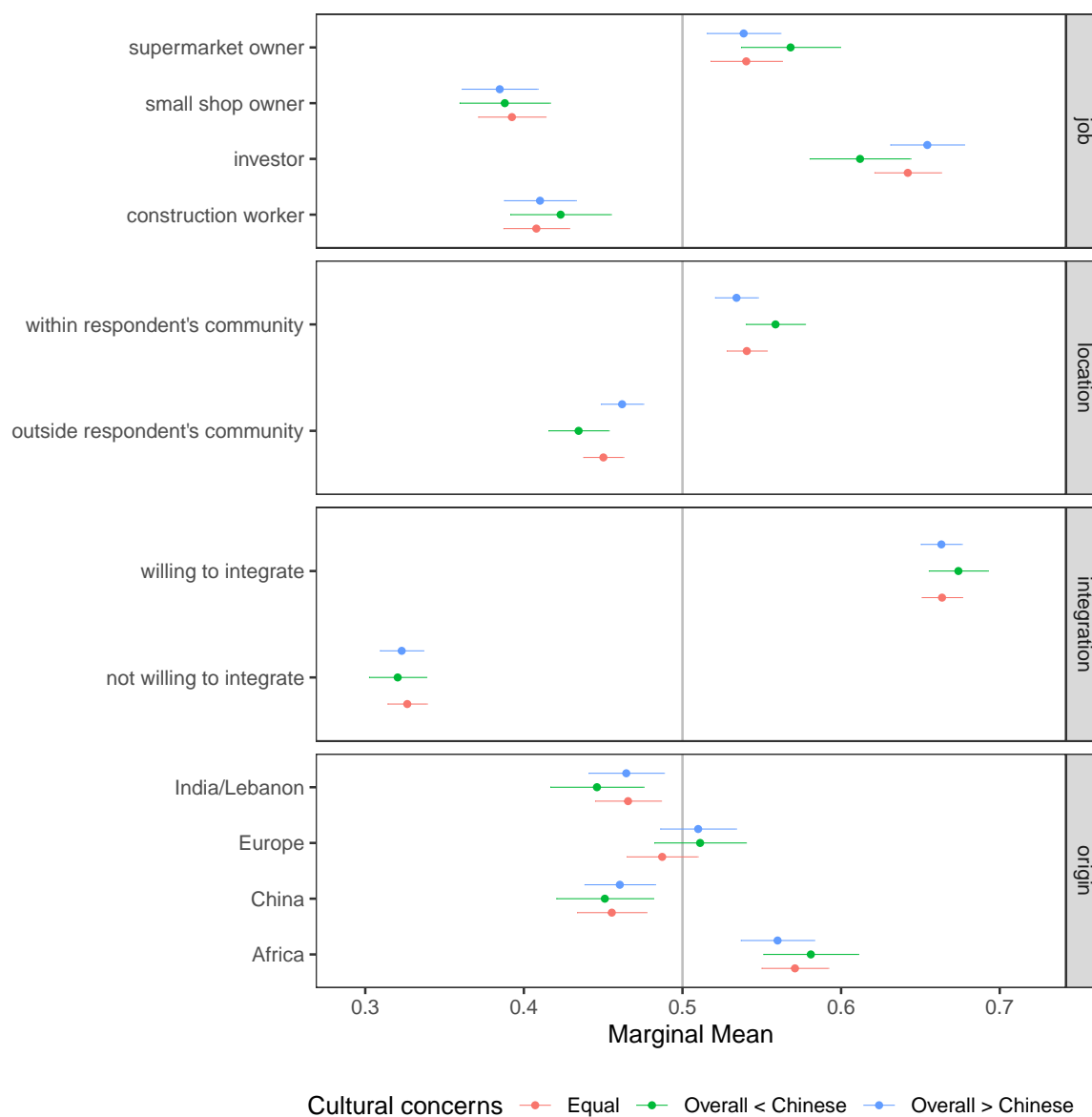
Note: The graph shows marginal means for equation 5.1 by country-specific quintiles of the estimated number of Chinese immigrants in respondents' home country. *Don't know* or *Refuse to answer* are dropped from the estimation.

Figure 5.A25: MMs for equation 5.1 by favorite destination region



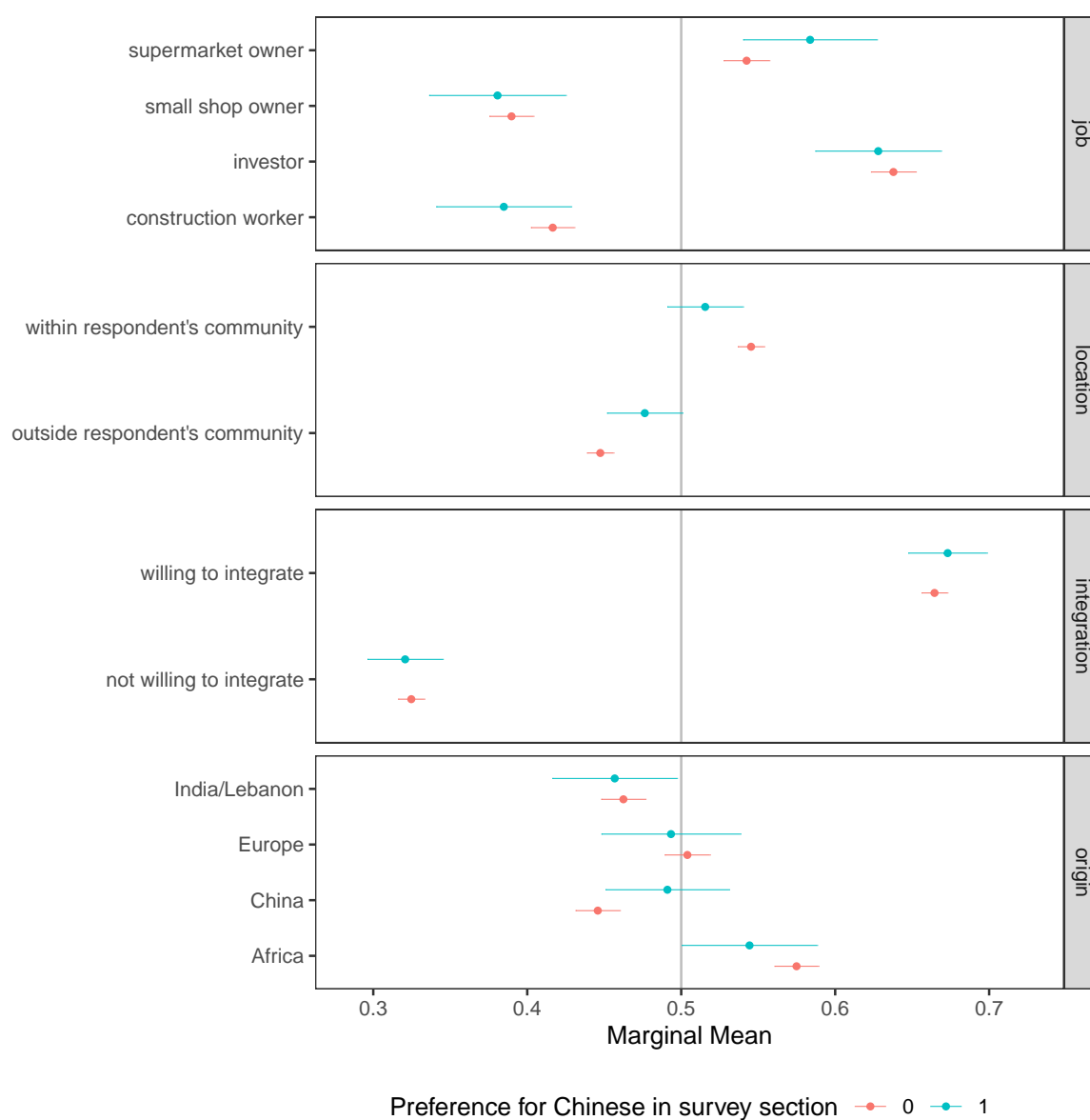
Note: The graph shows marginal means for equation 5.1 by respondents' favorite destination region. *Don't know* or *Refuse to answer* are dropped from the estimation.

Figure 5.A26: MMs for equation 5.1 by differences in sociotropic cultural concerns between immigrants in general and Chinese immigrants



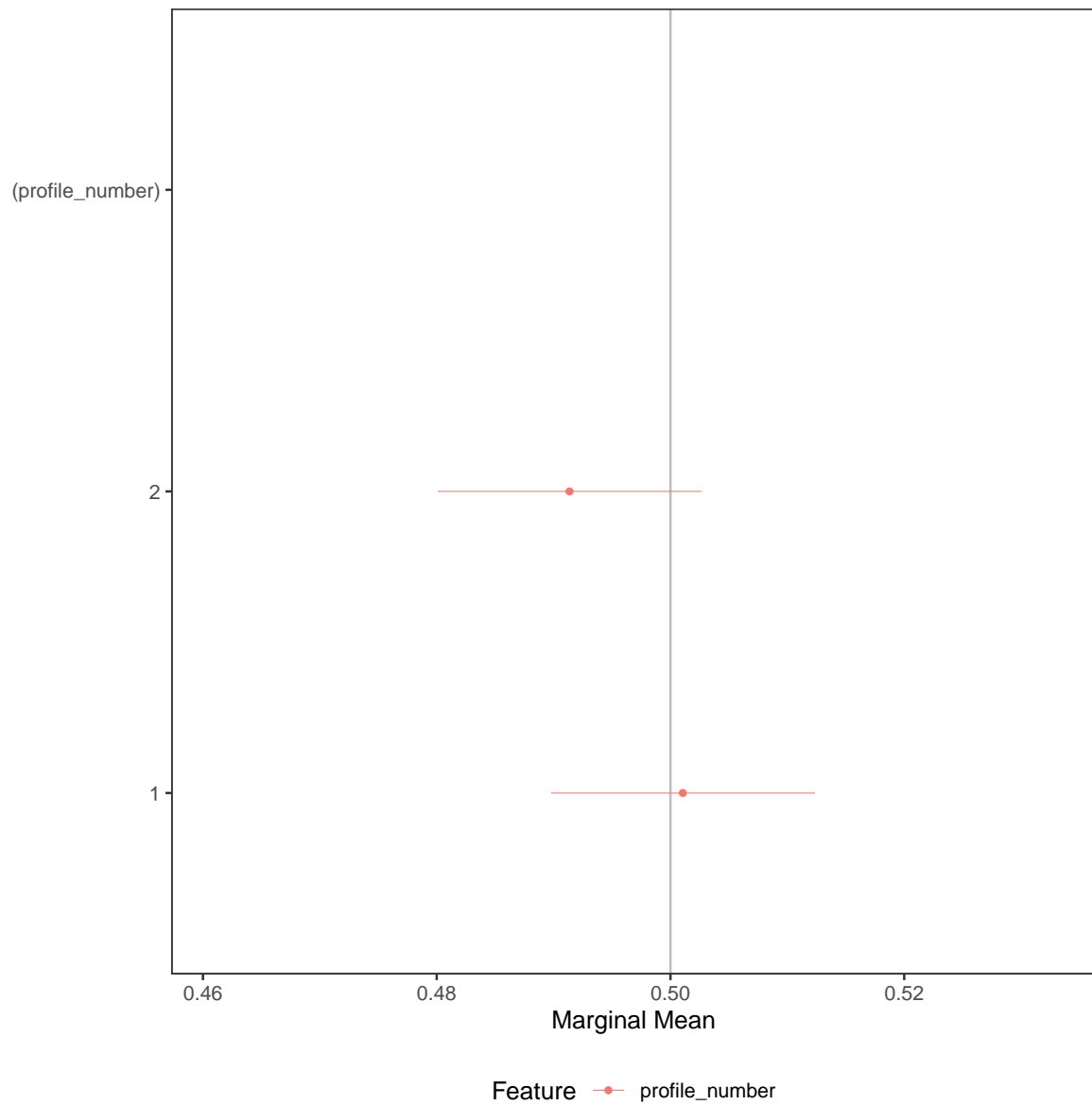
Note: The graph shows marginal means for equation 5.1 by differences in sociotropic cultural concerns between immigrants in general and Chinese immigrants. *Don't know* or *Refuse to answer* are dropped from the estimation.

Figure 5.A27: MMs for equation 5.1 by relative preference for Chinese in survey section



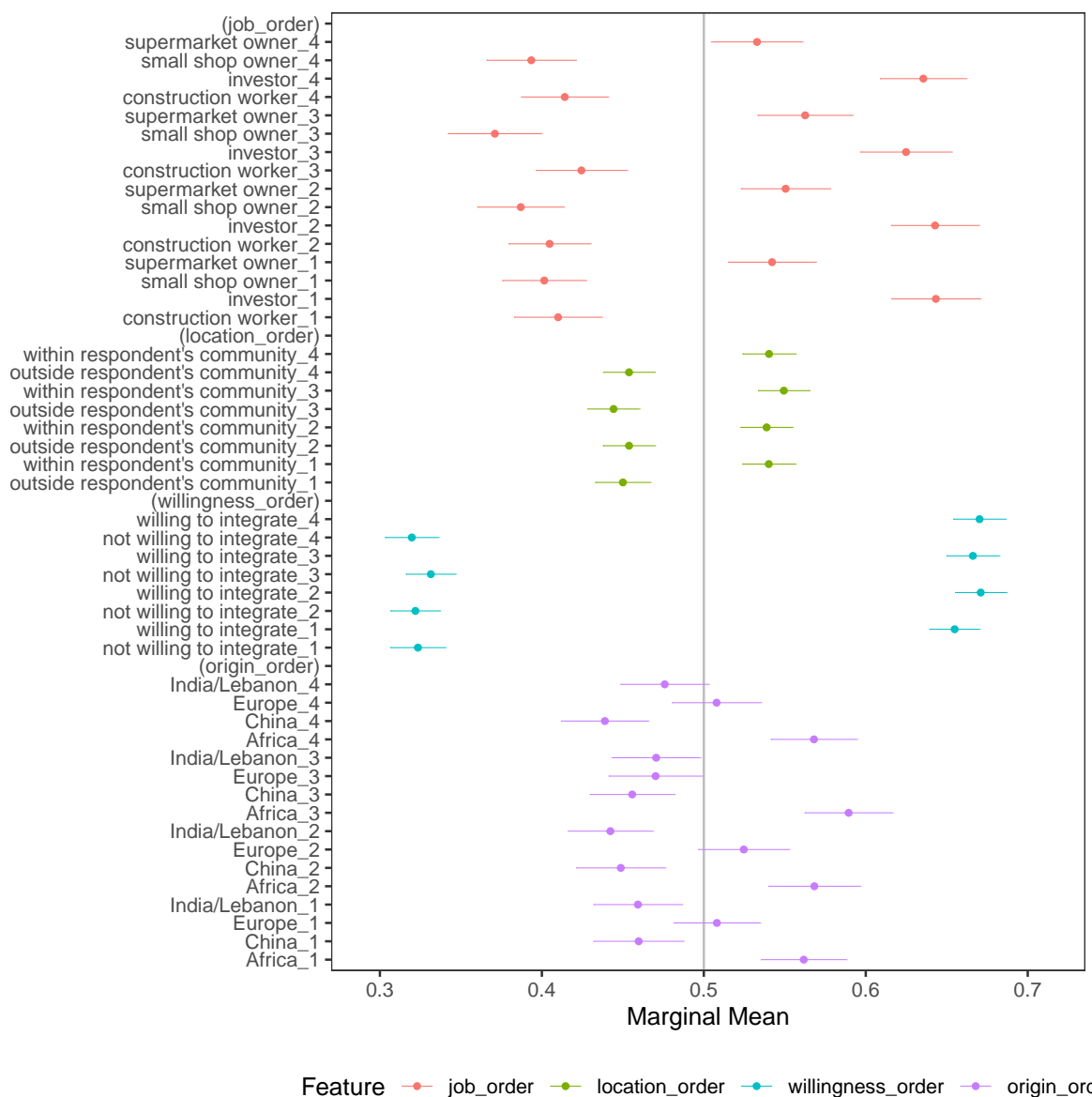
Note: The graph shows marginal means for equation 5.1 by relative preference for Chinese in the survey section. Respondents that preferred more Chinese immigrants than immigrants in general to come to their country are coded as 1. *Don't know* or *Refuse to answer* are dropped from the estimation.

Figure 5.A28: MMs for equation 5.1 by profile order



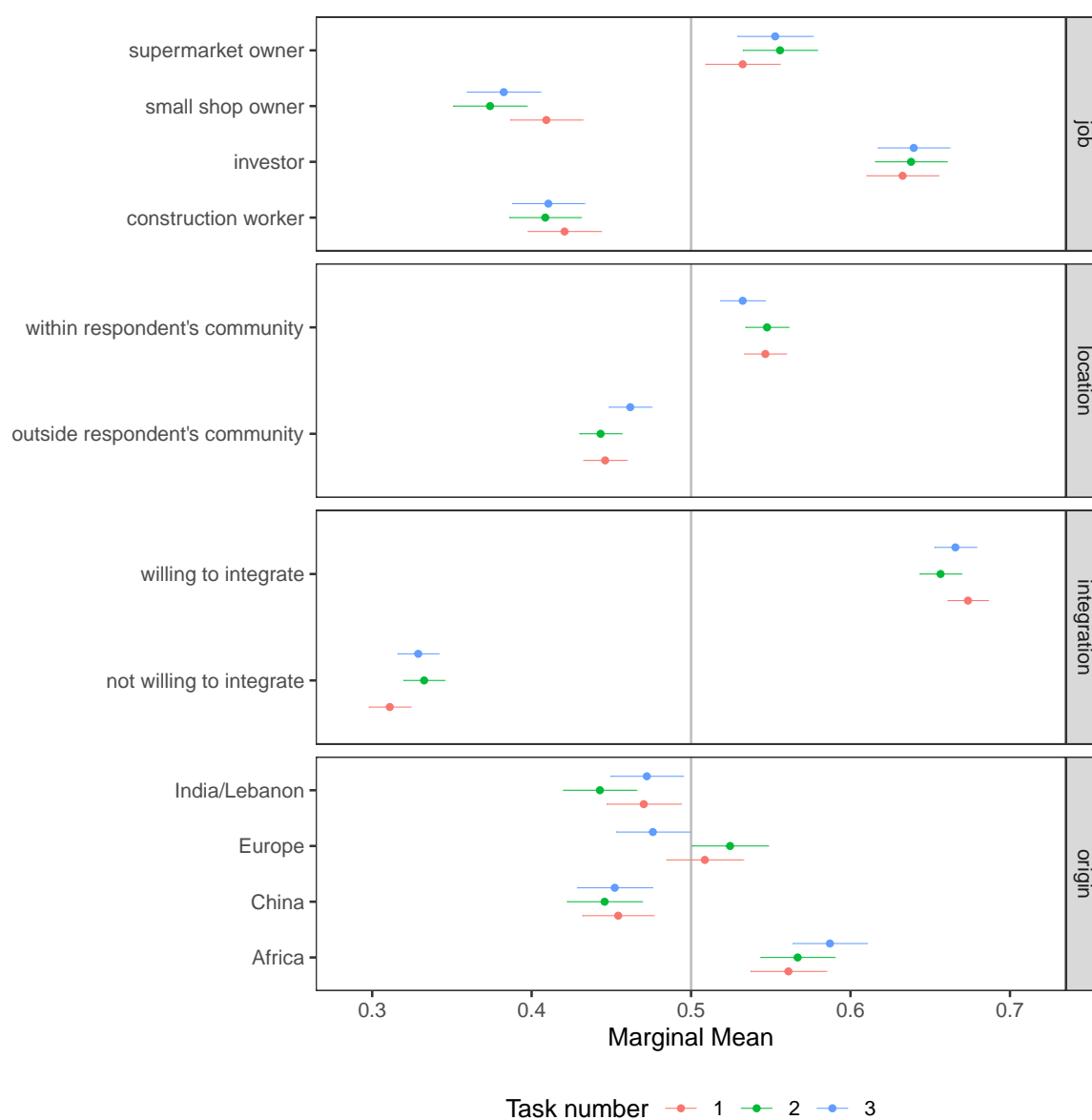
Note: The graph shows marginal means for equation 5.1 by profile order to test for profile order effects. *Don't know* or *Refuse to answer* are dropped from the estimation.

Figure 5.A29: MMs for equation 5.1 by attribute order



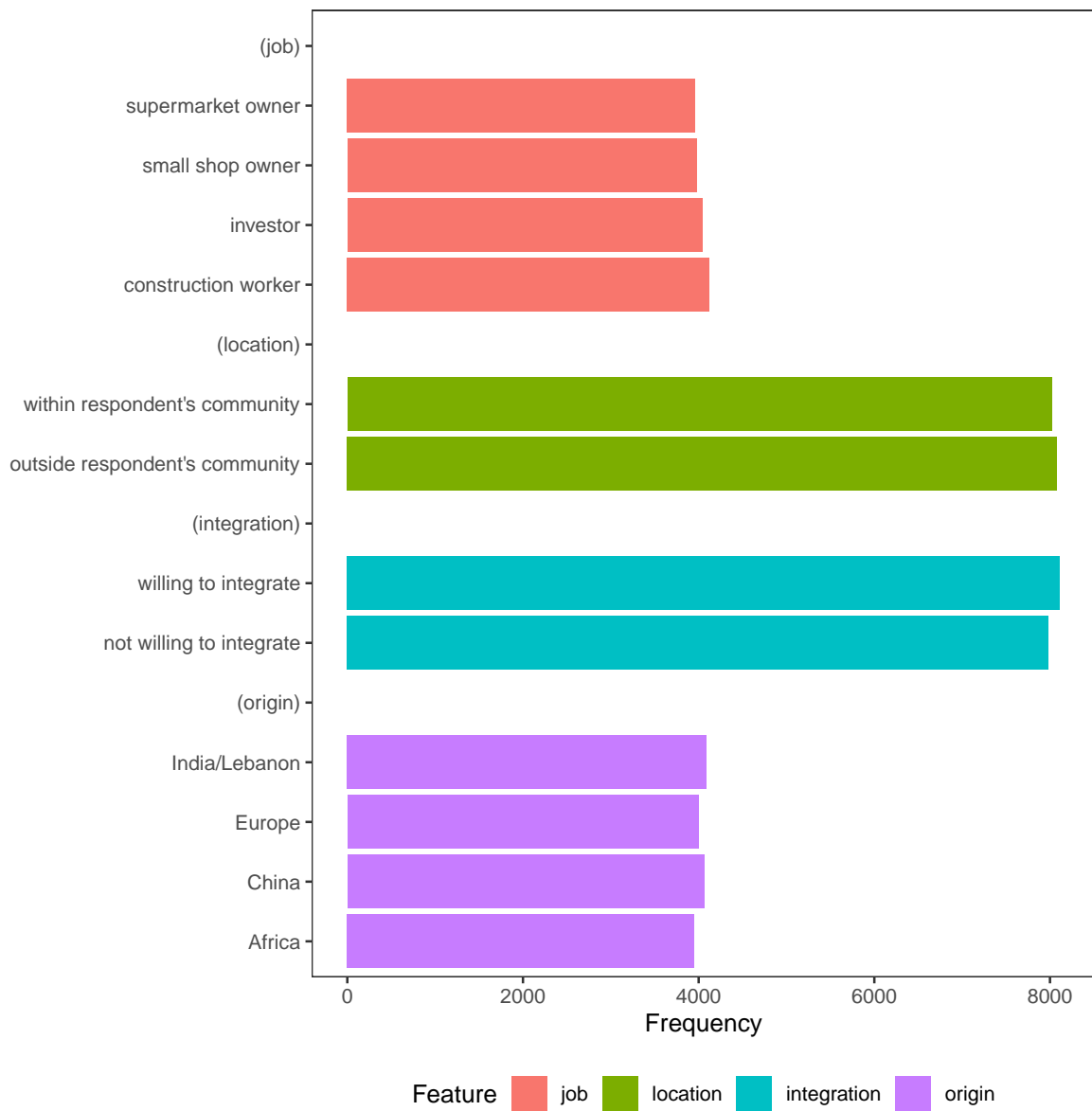
Note: The graph shows marginal means for equation 5.1 by attribute order to test for attribute order effects. *Don't know* or *Refuse to answer* are dropped from the estimation.

Figure 5.A30: MMs for equation 5.1 by task



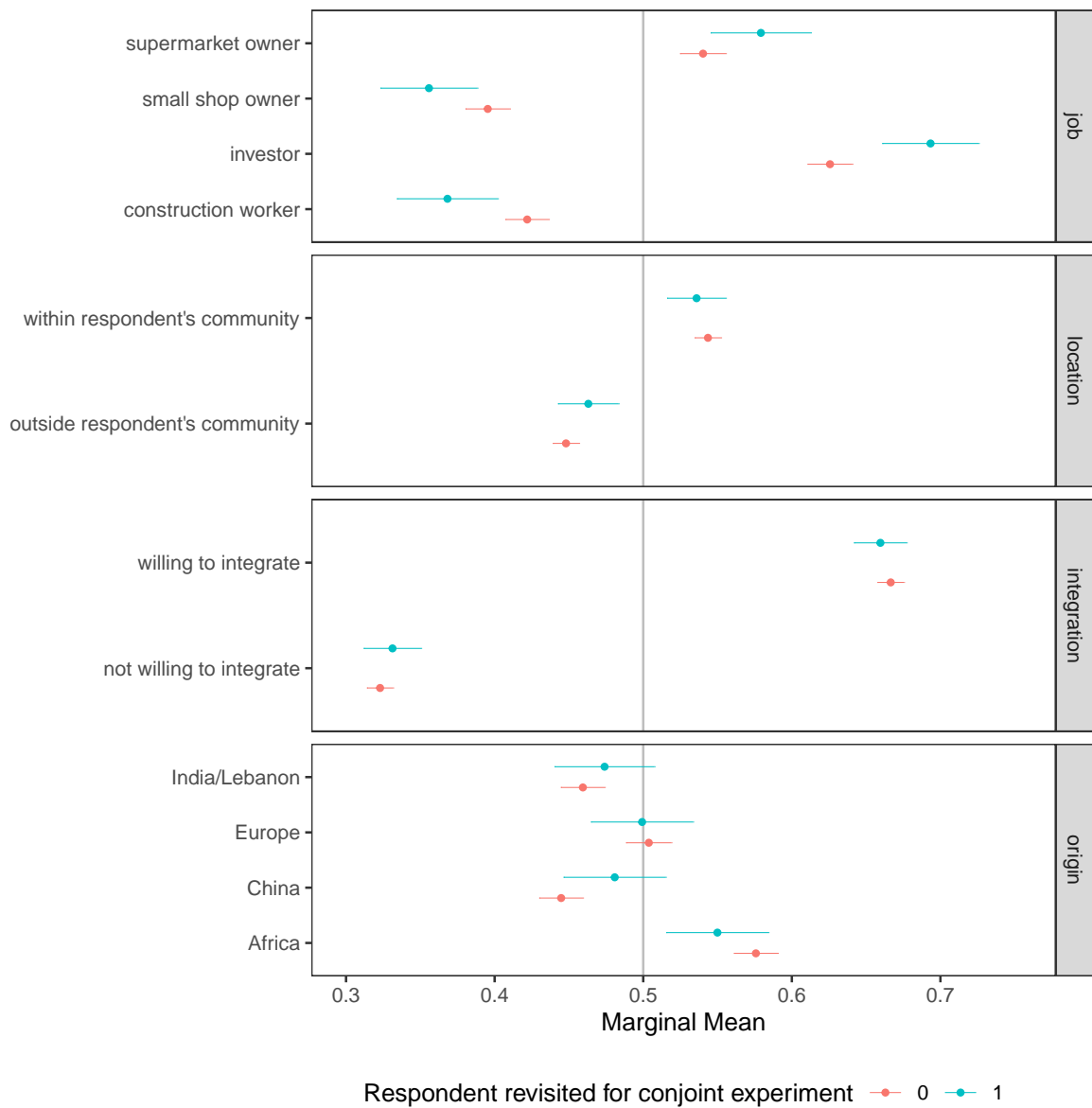
Note: The graph shows marginal means for equation 5.1 by task to test for carryover effects. Task number four has only been executed 6 times (due to programming mistakes), therefore standard errors are huge. *Don't know* or *Refuse to answer* are dropped from the estimation.

Figure 5.A31: Frequency of attribute levels



Note: The graph shows the frequency of each attribute level in the sample.

Figure 5.A32: MMs for equation 5.1 by whether or not respondents that had to be revisited



Note: The graph shows MMs for equation 5.1 by whether or not respondents had to be revisited. Respondents which had to be re-visited are coded as 1. *Don't know* or *Refuse to answer* are dropped from the estimation.

5.A.3 Tables

Table 5.A1: Attitudes toward immigrants - marginal means

feature	level	estimate	std.error	p
job	construction worker	0.41	0.01	0.00
job	investor	0.64	0.01	0.00
job	small shop owner	0.39	0.01	0.00
job	supermarket owner	0.55	0.01	0.00
location	outside respondent's community	0.45	0.00	0.00
location	within respondent's community	0.54	0.00	0.00
integration	not willing to integrate	0.32	0.00	0.00
integration	willing to integrate	0.67	0.00	0.00
origin	Africa	0.57	0.01	0.00
origin	China	0.45	0.01	0.00
origin	Europe	0.50	0.01	0.67
origin	India/Lebanon	0.46	0.01	0.00

Note: Marginal means for equation 5.1. "Don't know" or "Refuse to answer" are dropped from the estimation. Standard errors are clustered on the respondent level. The corresponding figure is 5.1. Results from the conjoint experiment align with the answers from the survey questions. E.g., respondents who would allow more Chinese immigrants than immigrants in general to come to their country in the survey section also have a preference for Chinese immigrants in the conjoint experiment, see figure 5.A27.

Table 5.A2: Attitudes toward immigrants - AMCE

feature	level	estimate	std.error	p
job	construction worker	0.00		
job	investor	0.22	0.01	0.00
job	small shop owner	-0.02	0.01	0.05
job	supermarket owner	0.14	0.01	0.00
location	outside respondent's community	0.00		
location	within respondent's community	0.09	0.01	0.00
integration	not willing to integrate	0.00		
integration	willing to integrate	0.34	0.01	0.00
origin	Africa	0.00		
origin	China	-0.12	0.01	0.00
origin	Europe	-0.07	0.01	0.00
origin	India/Lebanon	-0.11	0.01	0.00

Note: AMCEs for equation 5.1. Null estimates indicate the reference attribute level. "Don't know" or "Refuse to answer" are dropped from the estimation. Standard errors are clustered on the respondent level. The corresponding figure is 5.A4.

Table 5.A3: Attitudes toward immigrants - marginal means by skill level

BY	feature	level	estimate	std.error	p
higher skilled	job	construction worker	0.41	0.01	0.00
higher skilled	job	investor	0.65	0.01	0.00
higher skilled	job	small shop owner	0.38	0.01	0.00
higher skilled	job	supermarket owner	0.56	0.01	0.00
higher skilled	location	outside respondent's community	0.46	0.01	0.00
higher skilled	location	within respondent's community	0.54	0.01	0.00
higher skilled	integration	not willing to integrate	0.32	0.01	0.00
higher skilled	integration	willing to integrate	0.67	0.01	0.00
higher skilled	origin	Africa	0.58	0.01	0.00
higher skilled	origin	China	0.45	0.01	0.00
higher skilled	origin	Europe	0.51	0.01	0.58
higher skilled	origin	India/Lebanon	0.46	0.01	0.00
lower skilled	job	construction worker	0.41	0.01	0.00
lower skilled	job	investor	0.62	0.01	0.00
lower skilled	job	small shop owner	0.40	0.01	0.00
lower skilled	job	supermarket owner	0.54	0.01	0.00
lower skilled	location	outside respondent's community	0.44	0.01	0.00
lower skilled	location	within respondent's community	0.55	0.01	0.00
lower skilled	integration	not willing to integrate	0.33	0.01	0.00
lower skilled	integration	willing to integrate	0.66	0.01	0.00
lower skilled	origin	Africa	0.56	0.01	0.00
lower skilled	origin	China	0.46	0.01	0.00
lower skilled	origin	Europe	0.50	0.01	0.98
lower skilled	origin	India/Lebanon	0.46	0.01	0.00

Note: Marginal means for equation 5.1 by respondents' skill level proxied by educational attainment, whereas at least some secondary education counts as higher-skilled. "Don't know" or "Refuse to answer" are dropped from the estimation. Standard errors are clustered on the respondent level. The corresponding figure is 5.A5.

Table 5.A4: Attitudes toward immigrants - marginal means by employment status

feature	level	estimate	std.error	p	
not working, doesn't want	job	construction worker	0.44	0.04	0.12
not working, doesn't want	job	investor	0.64	0.05	0.00
not working, doesn't want	job	small shop owner	0.30	0.03	0.00
not working, doesn't want	job	supermarket owner	0.64	0.05	0.00
not working, doesn't want	location	outside respondent's community	0.45	0.03	0.04
not working, doesn't want	location	within respondent's community	0.54	0.02	0.09
not working, doesn't want	integration	not willing to integrate	0.32	0.02	0.00
not working, doesn't want	integration	willing to integrate	0.69	0.02	0.00
not working, doesn't want	origin	Africa	0.59	0.04	0.02
not working, doesn't want	origin	China	0.42	0.04	0.04
not working, doesn't want	origin	Europe	0.45	0.05	0.32
not working, doesn't want	origin	India/Lebanon	0.51	0.04	0.76
not working, wanting	job	construction worker	0.41	0.01	0.00
not working, wanting	job	investor	0.62	0.01	0.00
not working, wanting	job	small shop owner	0.38	0.01	0.00
not working, wanting	job	supermarket owner	0.58	0.01	0.00
not working, wanting	location	outside respondent's community	0.45	0.01	0.00
not working, wanting	location	within respondent's community	0.55	0.01	0.00
not working, wanting	integration	not willing to integrate	0.33	0.01	0.00
not working, wanting	integration	willing to integrate	0.66	0.01	0.00
not working, wanting	origin	Africa	0.57	0.01	0.00
not working, wanting	origin	China	0.44	0.01	0.00
not working, wanting	origin	Europe	0.50	0.01	0.75
not working, wanting	origin	India/Lebanon	0.48	0.01	0.07
studying	job	construction worker	0.45	0.02	0.01
studying	job	investor	0.62	0.02	0.00
studying	job	small shop owner	0.37	0.02	0.00
studying	job	supermarket owner	0.55	0.02	0.03
studying	location	outside respondent's community	0.45	0.01	0.00
studying	location	within respondent's community	0.55	0.01	0.00
studying	integration	not willing to integrate	0.34	0.01	0.00
studying	integration	willing to integrate	0.65	0.01	0.00
studying	origin	Africa	0.59	0.02	0.00
studying	origin	China	0.44	0.02	0.01
studying	origin	Europe	0.50	0.02	0.92
studying	origin	India/Lebanon	0.45	0.02	0.04
working	job	construction worker	0.42	0.01	0.00
working	job	investor	0.64	0.01	0.00
working	job	small shop owner	0.40	0.01	0.00
working	job	supermarket owner	0.52	0.01	0.06
working	location	outside respondent's community	0.45	0.01	0.00
working	location	within respondent's community	0.54	0.01	0.00
working	integration	not willing to integrate	0.32	0.01	0.00
working	integration	willing to integrate	0.67	0.01	0.00
working	origin	Africa	0.57	0.01	0.00
working	origin	China	0.46	0.01	0.00
working	origin	Europe	0.50	0.01	0.96
working	origin	India/Lebanon	0.46	0.01	0.00

Note: Marginal means for equation 5.1 by respondents' employment status during the past 7 days. "Don't know" or "Refuse to answer" are dropped from the estimation. Standard errors are clustered on the respondent level. The corresponding figure is 5.A6.

Table 5.A5: Attitudes toward immigrants - marginal means by household income per capita quintiles

BY	feature	level	estimate	std.error	p
1	job	construction worker	0.44	0.01	0.00
1	job	investor	0.63	0.02	0.00
1	job	small shop owner	0.37	0.01	0.00
1	job	supermarket owner	0.54	0.02	0.01
1	location	outside respondent's community	0.45	0.01	0.00
1	location	within respondent's community	0.54	0.01	0.00
1	integration	not willing to integrate	0.32	0.01	0.00
1	integration	willing to integrate	0.67	0.01	0.00
1	origin	Africa	0.57	0.01	0.00
1	origin	China	0.46	0.02	0.01
1	origin	Europe	0.50	0.02	0.75
1	origin	India/Lebanon	0.45	0.01	0.00
2	job	construction worker	0.40	0.01	0.00
2	job	investor	0.65	0.01	0.00
2	job	small shop owner	0.39	0.01	0.00
2	job	supermarket owner	0.55	0.01	0.00
2	location	outside respondent's community	0.47	0.01	0.00
2	location	within respondent's community	0.53	0.01	0.00
2	integration	not willing to integrate	0.33	0.01	0.00
2	integration	willing to integrate	0.65	0.01	0.00
2	origin	Africa	0.57	0.02	0.00
2	origin	China	0.45	0.01	0.00
2	origin	Europe	0.52	0.01	0.19
2	origin	India/Lebanon	0.45	0.01	0.00
3	job	construction worker	0.43	0.02	0.00
3	job	investor	0.61	0.02	0.00
3	job	small shop owner	0.39	0.02	0.00
3	job	supermarket owner	0.56	0.02	0.00
3	location	outside respondent's community	0.43	0.01	0.00
3	location	within respondent's community	0.57	0.01	0.00
3	integration	not willing to integrate	0.32	0.01	0.00
3	integration	willing to integrate	0.67	0.01	0.00
3	origin	Africa	0.60	0.02	0.00
3	origin	China	0.43	0.02	0.00
3	origin	Europe	0.48	0.02	0.26
3	origin	India/Lebanon	0.47	0.02	0.06
4	job	construction worker	0.39	0.02	0.00
4	job	investor	0.67	0.02	0.00
4	job	small shop owner	0.39	0.02	0.00
4	job	supermarket owner	0.53	0.02	0.08
4	location	outside respondent's community	0.45	0.01	0.00
4	location	within respondent's community	0.54	0.01	0.00
4	integration	not willing to integrate	0.33	0.01	0.00
4	integration	willing to integrate	0.66	0.01	0.00
4	origin	Africa	0.55	0.02	0.00

Continued on next page

Table 5.A5 continued from previous page

BY	feature	level	estimate	std.error	p
4	origin	China	0.46	0.02	0.01
4	origin	Europe	0.49	0.02	0.77
4	origin	India/Lebanon	0.48	0.02	0.37
5	job	construction worker	0.42	0.02	0.00
5	job	investor	0.63	0.02	0.00
5	job	small shop owner	0.41	0.02	0.00
5	job	supermarket owner	0.54	0.02	0.02
5	location	outside respondent's community	0.44	0.01	0.00
5	location	within respondent's community	0.55	0.01	0.00
5	integration	not willing to integrate	0.32	0.01	0.00
5	integration	willing to integrate	0.67	0.01	0.00
5	origin	Africa	0.56	0.02	0.00
5	origin	China	0.45	0.02	0.00
5	origin	Europe	0.53	0.02	0.08
5	origin	India/Lebanon	0.45	0.02	0.00

Note: Marginal means for equation 5.1 by country-specific quintiles of respondents' self-reported per household income per capita during the past 12 months. "Don't know" or "Refuse to answer" are dropped from the estimation. Standard errors are clustered on the respondent level. The corresponding figure is 5.A7.

Table 5.A6: Absolute attitudes toward immigrants: Overall vs. Chinese

	No controls, no power		No controls, no extrapol.		No controls, extrapol.		Full model		Full, non-standardized	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Overall	Chinese	Overall	Chinese	Overall	Chinese	Overall	Chinese	Overall	Chinese
Egocentric economic concerns	-0.0684* (0.0352)	-0.0575 (0.0389)	-0.0573 (0.0391)	-0.0463 (0.0746)	-0.0573 (0.0391)	-0.0605 (0.0465)	-0.0416 (0.0427)	-0.0676 (0.0543)	-0.0318 (0.0326)	-0.0528 (0.0423)
Sociotropic economic concerns	-0.240*** (0.0395)	-0.301*** (0.0358)	-0.235*** (0.0400)	-0.257** (0.0712)	-0.235*** (0.0400)	-0.280*** (0.0434)	-0.230*** (0.0457)	-0.282*** (0.0486)	-0.154*** (0.0306)	-0.186*** (0.0320)
Sociotropic job concerns	-0.0810** (0.0273)	-0.138*** (0.0321)	-0.0812** (0.0280)	-0.133** (0.0414)	-0.0812** (0.0280)	-0.147*** (0.0300)	-0.0909** (0.0311)	-0.146*** (0.0318)	-0.0224** (0.00765)	-0.0358*** (0.00784)
Sociotropic cultural concerns	-0.0950*** (0.0284)	-0.0220 (0.0283)	-0.0752** (0.0287)	-0.0584 (0.0370)	-0.0752** (0.0287)	-0.0105 (0.0260)	-0.0738** (0.0244)	-0.00133 (0.0302)	-0.0170** (0.00563)	-0.000360 (0.00827)
Power concerns government			0.00560 (0.0341)	-0.00323 (0.0627)	0.00560 (0.0341)	0.0307 (0.0404)	-0.00167 (0.0352)	0.0153 (0.0437)	-0.000375 (0.00789)	0.00286 (0.00989)
Power concerns business			-0.0111 (0.0302)	0.0731 (0.0842)	-0.0111 (0.0302)	-0.00697 (0.0548)	-0.0185 (0.0308)	0.0119 (0.0557)	-0.00439 (0.00731)	0.00371 (0.0128)
Control variables	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Country FEs	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
Standardized variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Observations	2538	2511	2279	1166	2279	2308	1943	1969	1943	1969
R ²	0.170	0.240	0.152	0.150	0.152	0.226	0.164	0.239	0.164	0.239

Notes: Estimated coefficients for equation 5.2. Negative coefficients indicate a negative relationship between specific concerns and attitudes toward immigrants. Outcome and indicator variables are standardized in columns (1) to (8) and non-standardized in columns (9) and (10). Power concerns were not asked separately for the Chinese government and businesses in Uganda. Column (4) only contains information for Senegal due to missing information regarding the Chinese government and businesses in Uganda. Columns (5) to (10) assume attitudes toward the Chinese government and businesses to be equal to attitudes toward foreign governments and businesses in general in Uganda. Control variables include respondents' age, gender, education, self-reported employment during the last 7 days (4 levels), location (rural-urban), self-reported contact with immigrants, country-specific quintiles for self-reported household income during the last 12 months, life satisfaction (0-10), and satisfaction with city/area of residence (1-5), and country fixed effects. Standard errors are in parentheses and clustered on the district level. *Don't know* and *Refuse to answer* are coded as missing. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5.A7: Linear hypothesis tests for differences in estimated coefficients from equation 5.2

(a) Immigrants in general						(b) Chinese immigrants					
	β_2	β_3	β_4	β_5	β_6		β_2	β_3	β_4	β_5	β_6
β_1	0.035	0.34	0.502	0.43	0.742	β_1	0.048	0.26	0.483	0.31	0.274
β_2		0.03	0.016	0.01	0.002	β_2		0.02	0.000	0.00	0.003
β_3			0.920	0.02	0.202	β_3			0.009	0.01	0.052
β_4				0.07	0.214	β_4				0.54	0.690
β_5					0.695	β_5					0.983

Wald test for differences between estimated coefficients, estimated using the *test* command in Stata.

Table 5.A8: Policy attitudes: predictors by immigrant group

	Overall		Chinese	
	Restricted	Full	Restricted	Full
Age	-0.00606 (0.00352)	-0.00681* (0.00341)	0.00108 (0.00408)	0.00129 (0.00408)
Female dummy	-0.130** (0.0517)	-0.0988** (0.0435)	-0.168*** (0.0433)	-0.161*** (0.0402)
1 = No education	Baseline category			
2 = Some primary education	0.130 (0.0756)	0.103 (0.0778)	0.0604 (0.0612)	0.0478 (0.0677)
3 = Completed primary education	-0.108 (0.123)	-0.109 (0.128)	0.0697 (0.173)	0.0582 (0.174)
4 = Some secondary education	-0.0139 (0.101)	-0.0468 (0.105)	0.0498 (0.113)	0.0400 (0.118)
5 = Completed secondary education	-0.000276 (0.137)	-0.0249 (0.134)	0.103 (0.145)	0.0990 (0.143)
6 = Vocational and technical training	-0.0651 (0.123)	-0.128 (0.126)	0.0907 (0.149)	0.108 (0.156)
7 = Tertiary education (Diploma/Degree)	-0.0219 (0.101)	-0.0681 (0.0961)	0.106 (0.115)	0.0922 (0.120)
Working	Baseline category			
Not working, does not want to	0.0976 (0.119)	0.0740 (0.117)	0.0501 (0.0625)	0.0488 (0.0634)
Not working but wants to	-0.0365 (0.0376)	-0.0303 (0.0363)	-0.0989* (0.0470)	-0.0934* (0.0504)
Studying	-0.0124 (0.0778)	-0.00707 (0.0747)	-0.0282 (0.0472)	-0.0219 (0.0489)
Senegal	Baseline category			
Uganda	-0.408*** (0.131)	-0.414*** (0.121)	-0.497*** (0.106)	-0.511*** (0.109)
Rural dummy	0.0513 (0.0861)	0.0632 (0.0847)	-0.0305 (0.0812)	-0.0239 (0.0837)
Contact with immigrants		0.0596** (0.0268)		0.0277 (0.0341)
Constant	0.391* (0.206)	0.236 (0.246)	0.269 (0.241)	0.204 (0.236)
Observations	2342	2316	2333	2319
R^2	0.056	0.065	0.069	0.070

The outcome variable is standardized. Higher values in the outcome variable imply more positive attitudes toward immigration. Standard errors are in parentheses. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5.A9: Egocentric economic concerns: predictors by immigrant group

	Overall		Chinese	
	Restricted	Full	Restricted	Full
Age	0.00489 (0.00343)	0.00516 (0.00342)	0.00806** (0.00355)	0.00771** (0.00337)
Female dummy	0.0348 (0.0468)	0.0118 (0.0452)	0.0974* (0.0541)	0.0834 (0.0541)
1 = No education	Baseline category			
2 = Some primary education	-0.0768 (0.101)	-0.0756 (0.105)	0.0837 (0.101)	0.0986 (0.100)
3 = Completed primary education	-0.0936 (0.0841)	-0.0932 (0.0858)	-0.00694 (0.134)	-0.00126 (0.132)
4 = Some secondary education	-0.0135 (0.0759)	0.00340 (0.0803)	0.0704 (0.0553)	0.0874 (0.0576)
5 = Completed secondary education	-0.142 (0.113)	-0.128 (0.116)	-0.0412 (0.131)	-0.0257 (0.131)
6 = Vocational and technical training	-0.184 (0.108)	-0.146 (0.103)	-0.0916 (0.125)	-0.0784 (0.121)
7 = Tertiary education (Diploma/Degree)	-0.160 (0.0926)	-0.139 (0.0989)	0.0856 (0.104)	0.108 (0.105)
Working	Baseline category			
Not working, does not want to	-0.208 (0.125)	-0.211 (0.125)	-0.0463 (0.0915)	-0.0523 (0.0931)
Not working but wants to	0.0861 (0.0583)	0.0821 (0.0584)	0.131* (0.0677)	0.118 (0.0730)
Studying	-0.0797 (0.0839)	-0.0848 (0.0813)	-0.0789 (0.0563)	-0.0788 (0.0617)
Senegal	Baseline category			
Uganda	-0.275** (0.0902)	-0.268** (0.0869)	-0.182 (0.109)	-0.144 (0.103)
Rural dummy	-0.00950 (0.0701)	-0.0131 (0.0731)	0.0431 (0.0760)	0.0289 (0.0769)
Contact with immigrants		-0.0367* (0.0167)		-0.0552* (0.0284)
Constant	0.0454 (0.148)	0.154 (0.153)	-0.274* (0.152)	-0.144 (0.165)
Observations	2316	2290	2318	2307
R^2	0.033	0.037	0.022	0.027

The outcome variable is standardized. High values in the outcome variable imply stronger egocentric economic concerns. Standard errors are in parentheses. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5.A10: Sociotropic economic concerns: predictors by immigrant group

	Overall		Chinese	
	Restricted	Full	Restricted	Full
Age	0.0000275 (0.00374)	0.000621 (0.00368)	0.00732* (0.00363)	0.00699* (0.00364)
Female dummy	0.0531 (0.0338)	0.0318 (0.0373)	0.138*** (0.0363)	0.132*** (0.0322)
1 = No education	Baseline category			
2 = Some primary education	-0.0580 (0.118)	-0.0525 (0.119)	-0.0286 (0.109)	-0.0204 (0.108)
3 = Completed primary education	-0.0500 (0.170)	-0.0483 (0.173)	-0.122 (0.138)	-0.119 (0.137)
4 = Some secondary education	0.0110 (0.121)	0.0252 (0.125)	-0.00874 (0.0858)	-0.00311 (0.0877)
5 = Completed secondary education	-0.118 (0.167)	-0.107 (0.172)	-0.0613 (0.140)	-0.0565 (0.141)
6 = Vocational and technical training	-0.222* (0.106)	-0.190 (0.110)	-0.241* (0.125)	-0.259* (0.129)
7 = Tertiary education (Diploma/Degree)	-0.142 (0.138)	-0.115 (0.140)	-0.0597 (0.0929)	-0.0492 (0.0919)
Working	Baseline category			
Not working, does not want to	-0.100 (0.133)	-0.0774 (0.127)	-0.227** (0.103)	-0.229* (0.106)
Not working but wants to	0.0104 (0.0530)	0.00644 (0.0546)	0.0991 (0.0650)	0.0938 (0.0692)
Studying	-0.115 (0.0696)	-0.113 (0.0693)	-0.0606 (0.0584)	-0.0581 (0.0597)
Senegal	Baseline category			
Uganda	-0.346** (0.132)	-0.338** (0.130)	-0.261** (0.106)	-0.245** (0.0998)
Rural dummy	-0.0109 (0.104)	-0.0188 (0.108)	0.0583 (0.0872)	0.0513 (0.0908)
Contact with immigrants		-0.0352 (0.0233)		-0.0263 (0.0347)
Constant	0.195 (0.199)	0.287 (0.212)	-0.164 (0.163)	-0.0963 (0.191)
Observations	2315	2292	2304	2290
R^2	0.037	0.040	0.036	0.038

The outcome variable is standardized. Higher values in the outcome variable imply stronger sociotropic economic concerns. Standard errors are in parentheses. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5.A11: Sociotropic job concerns: predictors by immigrant group

	Overall		Chinese	
	Restricted	Full	Restricted	Full
Age	-0.00268 (0.00340)	-0.00264 (0.00325)	-0.00230 (0.00341)	-0.00196 (0.00329)
Female dummy	0.137*** (0.0385)	0.128*** (0.0343)	0.0809 (0.0544)	0.0727 (0.0500)
1 = No education	Baseline category			
2 = Some primary education	-0.124 (0.152)	-0.138 (0.155)	-0.121 (0.101)	-0.114 (0.100)
3 = Completed primary education	-0.183 (0.260)	-0.183 (0.261)	-0.200 (0.203)	-0.195 (0.201)
4 = Some secondary education	-0.0322 (0.138)	-0.0373 (0.141)	-0.183** (0.0799)	-0.176* (0.0821)
5 = Completed secondary education	-0.221 (0.201)	-0.217 (0.204)	-0.277** (0.121)	-0.262** (0.118)
6 = Vocational and technical training	-0.100 (0.149)	-0.0850 (0.144)	-0.230* (0.119)	-0.210* (0.114)
7 = Tertiary education (Diploma/Degree)	-0.0594 (0.149)	-0.0568 (0.153)	-0.106 (0.113)	-0.0912 (0.112)
Working	Baseline category			
Not working, does not want to	-0.160 (0.120)	-0.135 (0.116)	-0.0185 (0.167)	-0.0214 (0.169)
Not working but wants to	0.0209 (0.0337)	0.0325 (0.0321)	0.0624 (0.0700)	0.0592 (0.0724)
Studying	-0.205** (0.0727)	-0.202** (0.0709)	-0.137 (0.111)	-0.141 (0.110)
Senegal	Baseline category			
Uganda	-0.296 (0.171)	-0.292 (0.173)	0.171 (0.117)	0.191 (0.124)
Rural dummy	-0.0892 (0.107)	-0.0879 (0.107)	0.0157 (0.0781)	0.00926 (0.0839)
Contact with immigrants		-0.00359 (0.0252)		-0.0299 (0.0310)
Constant	0.268 (0.154)	0.277 (0.164)	0.0705 (0.121)	0.126 (0.150)
Observations	2315	2291	2306	2293
R^2	0.035	0.034	0.017	0.019

The outcome variable is standardized. Higher values in the outcome variable imply stronger sociotropic concerns about the effect of immigration on jobs. Standard errors are in parentheses. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5.A12: Sociotropic cultural concerns: predictors by immigrant group

	Overall		Chinese	
	Restricted	Full	Restricted	Full
Age	0.00500 (0.00310)	0.00455 (0.00339)	-0.00000650 (0.00393)	-0.0000485 (0.00399)
Female dummy	-0.00254 (0.0541)	0.00915 (0.0557)	0.0119 (0.0481)	0.0279 (0.0510)
1 = No education	Baseline category			
2 = Some primary education	0.158* (0.0878)	0.153* (0.0791)	0.00377 (0.0871)	-0.0127 (0.0894)
3 = Completed primary education	0.0912 (0.110)	0.0995 (0.106)	0.0351 (0.0752)	0.0316 (0.0775)
4 = Some secondary education	0.196* (0.0920)	0.183** (0.0808)	0.0961 (0.0638)	0.0860 (0.0667)
5 = Completed secondary education	0.0714 (0.133)	0.0637 (0.133)	0.112 (0.101)	0.0925 (0.107)
6 = Vocational and technical training	0.173 (0.114)	0.141 (0.100)	0.00355 (0.219)	-0.0377 (0.219)
7 = Tertiary education (Diploma/Degree)	0.193 (0.111)	0.188* (0.0986)	0.0767 (0.0773)	0.0676 (0.0759)
Working	Baseline category			
Not working, does not want to	-0.149 (0.123)	-0.135 (0.115)	-0.0312 (0.115)	-0.0315 (0.112)
Not working but wants to	-0.0131 (0.0528)	-0.0131 (0.0524)	0.00450 (0.0388)	0.00865 (0.0382)
Studying	0.0435 (0.0702)	0.0520 (0.0679)	-0.0851 (0.0583)	-0.0788 (0.0554)
Senegal	Baseline category			
Uganda	0.315** (0.136)	0.311** (0.140)	0.455*** (0.0958)	0.433*** (0.104)
Rural dummy	-0.143 (0.0937)	-0.138 (0.0977)	-0.175** (0.0638)	-0.171** (0.0670)
Contact with immigrants		0.0278 (0.0280)		0.0309 (0.0177)
Constant	-0.352 (0.203)	-0.428 (0.253)	-0.191 (0.153)	-0.256 (0.149)
Observations	2319	2292	2302	2289
R^2	0.049	0.052	0.076	0.078

The outcome variable is standardized. Higher values in the outcome variable imply stronger sociotropic cultural concerns. Standard errors are in parentheses. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5.A13: Power concerns foreign governments: predictors by immigrant group

	Overall		Chinese	
	Restricted	Full	Restricted	Full
Age	-0.00405 (0.00316)	-0.00479 (0.00312)	0.000537 (0.00791)	0.000665 (0.00722)
Female dummy	-0.168*** (0.0423)	-0.161*** (0.0406)	0.00698 (0.0282)	0.0279 (0.0302)
1 = No education	Baseline category			
2 = Some primary education	0.270*** (0.0633)	0.251*** (0.0678)	0.166** (0.0602)	0.146** (0.0544)
3 = Completed primary education	0.0529 (0.0965)	0.0472 (0.0955)	0.205 (0.136)	0.202 (0.131)
4 = Some secondary education	0.316*** (0.0919)	0.297*** (0.0924)	0.170* (0.0707)	0.154* (0.0637)
5 = Completed secondary education	0.421*** (0.0767)	0.401*** (0.0759)	0.199 (0.139)	0.178 (0.136)
6 = Vocational and technical training	0.209* (0.111)	0.181 (0.105)	0.287 (0.292)	0.263 (0.284)
7 = Tertiary education (Diploma/Degree)	0.368*** (0.0731)	0.346*** (0.0756)	0.175 (0.141)	0.150 (0.136)
Working	Baseline category			
Not working, does not want to	0.148 (0.136)	0.139 (0.127)	-0.153 (0.231)	-0.147 (0.220)
Not working but wants to	-0.0213 (0.0495)	-0.0131 (0.0488)	0.0751* (0.0324)	0.0869** (0.0310)
Studying	-0.0308 (0.0741)	-0.0240 (0.0744)	-0.0740 (0.0814)	-0.0468 (0.0750)
Senegal	Baseline category			
Uganda	0.00282 (0.116)	0.0105 (0.115)		
Rural dummy	-0.112 (0.0957)	-0.108 (0.0932)	-0.0468 (0.102)	-0.0426 (0.109)
Contact with immigrants		0.0199 (0.0181)		0.0545 (0.0391)
Constant	0.00862 (0.205)	-0.0306 (0.237)	-0.119 (0.271)	-0.242 (0.209)
Observations	2145	2124	1170	1165
R^2	0.036	0.036	0.011	0.015

The outcome variable is standardized. Higher values in the outcome variable imply stronger concerns about the influence of foreign governments. Power concerns toward the Chinese government in particular are only available for Senegal. Standard errors are in parentheses. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5.A14: Power concerns foreign businesses: predictors by immigrant group

	Overall		Chinese	
	Restricted	Full	Restricted	Full
Age	0.00194 (0.00401)	0.00112 (0.00431)	-0.00420 (0.00429)	-0.00437 (0.00353)
Female dummy	-0.133** (0.0558)	-0.108* (0.0543)	0.0615 (0.0747)	0.0794 (0.0701)
1 = No education	Baseline category			
2 = Some primary education	0.208 (0.123)	0.180 (0.124)	0.163* (0.0721)	0.143 (0.0786)
3 = Completed primary education	0.119 (0.159)	0.116 (0.157)	0.0592 (0.107)	0.0582 (0.105)
4 = Some secondary education	0.420*** (0.117)	0.382** (0.124)	0.278* (0.108)	0.267* (0.115)
5 = Completed secondary education	0.411** (0.168)	0.374** (0.169)	0.364 (0.231)	0.346 (0.232)
6 = Vocational and technical training	0.453** (0.167)	0.393** (0.170)	0.225 (0.163)	0.205 (0.152)
7 = Tertiary education (Diploma/Degree)	0.565*** (0.133)	0.512*** (0.139)	0.273* (0.132)	0.259* (0.127)
Working	Baseline category			
Not working, does not want to	0.0440 (0.183)	0.0109 (0.176)	-0.308* (0.141)	-0.301* (0.126)
Not working but wants to	-0.0188 (0.0566)	-0.0217 (0.0549)	-0.0342 (0.0664)	-0.0156 (0.0603)
Studying	-0.0136 (0.0802)	-0.0123 (0.0764)	-0.147* (0.0709)	-0.124 (0.0694)
Senegal	Baseline category			
Uganda	0.00441 (0.128)	-0.000940 (0.141)		
Rural dummy	-0.0219 (0.0770)	-0.0194 (0.0884)	-0.0182 (0.124)	-0.00913 (0.129)
Contact with immigrants		0.0669*** (0.0214)		0.0614 (0.0306)
Constant	-0.273 (0.183)	-0.434** (0.193)	-0.00539 (0.216)	-0.140 (0.194)
Observations	2135	2116	1180	1176
R^2	0.041	0.054	0.018	0.025

The outcome variable is standardized. Higher values in the outcome variable imply stronger concerns about the influence of foreign businesses. Power concerns toward Chinese businesses in particular are only available for Senegal. Standard errors are in parentheses. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5.A15: Marginal Means: Bonferroni and Weights

feature	level	Baseline		Bonferroni		Weighted	
		estimate	p	estimate	p	estimate	p
job	construction worker	0.41	1.08E-36	0.41	3.14E-35	0.44	1.32E-07
job	investor	0.64	5.63E-85	0.64	1.63E-83	0.61	9.52E-20
job	small shop owner	0.39	1.48E-57	0.39	4.29E-56	0.40	1.72E-15
job	supermarket owner	0.55	5.37E-11	0.55	1.56E-09	0.53	7.37E-03
location	outside respondent's community	0.45	5.81E-33	0.45	1.69E-31	0.43	5.95E-26
location	within respondent's community	0.54	1.33E-24	0.54	3.84E-23	0.56	2.33E-18
integration	not willing to integrate	0.32	0.00E+00	0.32	0.00E+00	0.32	1.69E-152
integration	willing to integrate	0.67	0.00E+00	0.67	0.00E+00	0.66	1.43E-126
origin	Africa	0.57	3.76E-25	0.57	1.09E-23	0.56	2.33E-08
origin	China	0.45	1.09E-12	0.45	3.16E-11	0.46	2.66E-03
origin	Europe	0.50	6.73E-01	0.50	1.00E+00	0.50	7.21E-01
origin	India/Lebanon	0.46	3.23E-08	0.46	9.35E-07	0.46	4.07E-04

Columns 1 and 2 display estimates and p-values for our baseline Marginal Means estimation. Columns 3 and 4 adjust for 29 hypotheses that we tested in the conjoint experiment using the Bonferroni method. Columns 5 and 6 use weights to adjust for sampling design and non-response. P-values are shown in scientific notation to display changes for very small p-values.

Erklärungen

Eidesstattliche Erklärung

Erklärung zum selbständigen Verfassen der Arbeit:

Ich erkläre hiermit, dass ich meine Doktorarbeit "Understanding migration decisions and their reverberations" selbstständig und ohne fremde Hilfe angefertigt habe und dass ich als Koautor maßgeblich zu den weiteren Fachartikeln beigetragen habe. Alle von anderen Autoren wörtlich übernommenen Stellen, wie auch die sich an die Gedanken anderer Autoren eng anlehenden Ausführungen der aufgeführten Beiträge wurden besonders gekennzeichnet und die Quellen nach den mir angegebenen Richtlinien zitiert.

Kiel, den 21. April 2023

Finja K. Krüger

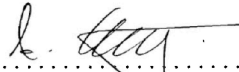
Erklärung der Ko-autoren

I hereby declare that I am co-author of the article listed below. In the article written under my co-authorship, all co-authors have contributed significantly to the development of the research question, the analysis of the research data, and the preparation of the manuscript, in accordance with the DFG guidelines for ensuring good scientific practice.

Boss, K., Gröger, A., Heidland, T., Krüger, F., & Zheng, C. (2023). Forecasting Bilateral Refugee Flows with High-dimensional Data and Machine Learning Techniques

KONSTANTIN
BOSS - DNI
Y7499378N

Digital unterschrieben
von KONSTANTIN BOSS -
DNI Y7499378N
Datum: 2023.04.12
13:17:33 +02'00'



Konstantin Boss

Andre Gröger



Tobias Heidland

CONGHAN
ZHENG - DNI
Y6499537G

Digitally signed by
CONGHAN ZHENG
DNI Y6499537G
Date: 2023.04.13
11:30:06 +02'00'

Conghan Zheng

Dražanová, L., Gonnot, J., Heidland, T., & Krüger, F. (2023). Which Individual-Level Factors Explain Public Attitudes Toward Immigration? A Meta-Analysis



Lenka Dražanová

Jérôme Gonnot

Tobias Heidland

Becker, M., Krüger, F., & Heidland, T. (2023). Country, Culture or Competition - What Drives Attitudes toward Immigrants in Sub-Saharan Africa?



Malte Becker



Tobias Heidland