UNIVERSITÉ PARIS 1 PANTHÉON-SORBONNE
École Doctorale D'Économie (ED465)
Centre d'Économie de la Sorbonne (CES)

THÈSE
Pour l'obtention du titre de Docteur en Sciences Économiques

Présentée et soutenue publiquement le 22 avril 2024 par

Yasmine Elkhateeb

Trois Essais sur l'Impact Politique, Social, et Économique de l'Internet et de la Migration

---

Sous la co-direction de:

Mathilde Maurel
Directrice de Recherche CNRS
Université Paris 1 Panthéon-Sorbonne, CES

Jérôme Valette
Économiste au CEPII & Maître de conférence
Université Paris 1 Panthéon-Sorbonne, CES

Membres du jury:

Simone BERTOLI	Professeur des Universités Université Clermont Auvergne, CNRS, CERDI (Rapporteur)
Lisa CHAUVET	Professeur des Universités Université Paris 1 Panthéon-Sorbonne, CES (Présidente)
Flore GUBERT	Directrice de Recherche IRD Université Paris-Dauphine & PSL, LEDa-DIAL (Rapporteure)
Marion MERCIER	Chargée de recherche CNRS Université Paris-Dauphine & PSL, LEDa-DIAL (Examinatrice)

UNIVERSITÉ PARIS 1 PANTHÉON-SORBONNE
École Doctorale D'Économie (ED465)
Centre d'Économie de la Sorbonne (CES)

PhD THESIS
For the Degree of Doctor of Philosophy in Economics

Presented and publicly defended on April 22, 2024 by

Yasmine Elkhateeb

Three Essays on the Political, Social, and Economic Impact of the Internet and Migration

---

Under the co-supervision of:

Mathilde Maurel
Director of Research CNRS
University Paris 1 Panthéon-Sorbonne, CES

Jérôme Valette
Economist at CEPII & Associate Professor
University Paris 1 Panthéon-Sorbonne, CES

Members of the Committee:

Simone BERTOLI	Professor of Economics University Clermont Auvergne, CNRS, CERDI (Reviewer)
Lisa CHAUVET	Professor of Economics University Paris 1 Panthéon-Sorbonne, CES (President)
Flore GUBERT	Director of Research IRD University Paris-Dauphine & PSL, LEDa-DIAL (Reviewer)
Marion MERCIER	Researcher CNRS University Paris-Dauphine & PSL, LEDa-DIAL (Examiner)

To my family.

Acknoweldgments

Looking back five years ago, I recall being torn about returning to Egypt after completing my Master's degree in Development Economics. I doubted whether I had what it takes to apply for this Ph.D. program, let alone survive the thesis years. Yet here I am, just a day away from submitting the final version of my thesis, filled with immense gratitude for all those who believed in me and encouraged me to reach this day. This thesis would not have been achieved without your unwavering support, and know that the next words will never be enough to thank you for your presence throughout this journey.

First and foremost, I wish to express my sincere gratitude to my incredible supervisors, Mathilde Maurel and Jérôme Valette. Your invaluable guidance, knowledge, and support have been the pillars in achieving this milestone. Your positivity, flexibility in research, and thoughtful advice kept me motivated. Mathilde Maurel, you believed in me since day one when I presented my Master's thesis research idea to you, and you steered me away from the mistake of not applying for this Ph.D. Your trust in me and consistent reassurance throughout those years have been invaluable. Jérôme Valette, I vividly recall receiving an email from you offering thorough feedback on my presentation during the trial auditions for the doctoral contract. Since then, you have remained consistently available, reviewing my work and encouraging me, long before you officially became my co-supervisor. I consider myself fortunate to have had both of you as my mentors and to have worked with you as co-authors on two chapters of this thesis. You taught me how to be a rigorous researcher and I am deeply grateful to both of you. Thank you!

I am also honored to have Simone Bertoli, Lisa Chauvet, Flore Gubert, and Marion Mercier as members of my jury committee. I would like to express my gratitude to Lisa Chauvet, who was also part of my thesis committee, for all the discussions we had and for her insightful comments throughout the development of the chapters of this thesis. Additionally, I extend my thanks to Simone Bertoli and Flore Gubert for graciously agreeing to be my two referees. Your invaluable insights and suggestions have significantly enhanced the quality of my work. I am also thankful to Marion Mercier for kindly agreeing to read and review the chapters of my thesis.

A heartfelt thanks goes to my exceptional co-authors, Joël Cariolle and Riccardo Turati. Collaborating with you on two chapters of this thesis has made the process thoroughly enjoyable. Thank you for your invaluable assistance and for the fruitful discussions we shared. I have learned a great deal from both of you. Additionally, I want to extend my thanks to Riccardo for inviting me to spend one month at Universitat Autònoma de

Barcelona as a visiting Ph.D. student in February 2023, where I had a wonderful experience and had the opportunity to meet remarkable researchers in the Applied Economics Department.

I would also like to extend my gratitude to all the professors at Paris 1 Panthéon-Sorbonne University, particularly those in the Development group, and at the Centre d'Économie de la Sorbonne, who taught me courses during my Master's degree, assisted me in various ways, or simply inspired me throughout this journey. Specifically, I want to thank Rémi Bazillier, Marie Boltz, Josselin Thuillez, and Marc-Arthur Diaye.

Teaching Economics has been and will always be one of my greatest passions, and my time at Paris 1 Panthéon-Sorbonne University has been profoundly enriching both professionally and personally. Therefore, I am deeply grateful to all the professors with whom I have had the privilege of teaching courses in Macroeconomics of Growth, Statistics, and Mathematics including Luigi Moretti, Goulven Rubin, Pierre-Charles Pradier, Nouhoum Touré, Biagio Speciale, Bastien Blain, Muriel Pucci, and Sophie Jallais.

I would also like to thank the entire administrative team at the Sorbonne Doctoral School of Economics and the Centre d'Économie de la Sorbonne, who were consistently available to assist and facilitate all administrative processes. Special thanks to Christelle Gauvrit, Lucie Label, Nathalie Lounie, Mathieu Collin, and Ngoc-Diep Ton for their exceptional assistance.

I am also grateful to my professors at Cairo University, who ignited my passion for economics, assisted me in applying for the Eiffel Scholarship, and inspired me to pursue this Ph.D. Specifically, I am thankful to Chahir Zaki, Nora Aboushady, Dina Kassab, Riham Ezzat, Myriam Ramzi, and Racha Ramadan.

Throughout this journey, I have been fortunate to be surrounded by many wonderful friends and colleagues at the Maison des Sciences Économiques (MSE). To Farida, Jala, and Nada, thank you for your unwavering support and encouragement; you have been true friends to me. I will always cherish our coffee breaks, corridor chitchats, and venting discussions and outings. To my Ph.D. fellows in the 316 office—Andrea, Thibault, Fatou, Clémence, Malak, Natalia, Stephen, Coralie, Étienne, José, Musa, Hugo, and Su—I am grateful for sharing the office with you throughout these years and for the enriching discussions and great moments we've had. I would also like to express my gratitude to Antonio, Robert, Johann, Beatrice, and Maria for their friendliness and the memorable moments we shared inside and outside the MSE. I am also thankful for the amazing individuals I met during my first two Ph.D. years, who made the COVID era more bearable

at the MSE and more pleasant— Adham, Mona, Mussa, Aref, Zeinab, Davide, Gaspard, Enxhi, Camille, and Mehdi. I would also like to thank Adham for all his guidance and support during my early Ph.D. years. Additional thanks go to my colleagues in the neighboring office for the enjoyable second-floor lunch breaks and after-work outings, including Pol, Maxence, Afifi, and Elsa.

I am also grateful to have met Hadir, Mariam, Alia, Hagar, Mayar, Bassem, and Karim in Paris. You have made my Ph.D. years a little easier, my time in Paris more enjoyable, and helped me disconnect during stressful moments. Thank you for the wonderful trips, fun nights out, cozy gatherings, and all the great memories. Special thanks to Gehad, Rowayda, and Hala for all the moments we shared in Paris at different times. I also want to extend my gratitude to Zeina, Salma, Marwa, Mirna, Mariam, Alia, Dana, and Dareen, my constant friends, for their love and encouragement throughout the years, despite the distance.

Last but not least, I am profoundly indebted to my wonderful family and extended family. Words will never be enough to express my gratitude and love to you. To my father, thank you for teaching me how to be independent, for encouraging me to pursue my dreams, and for never letting me give up. To my mother, thank you for all your kindness, understanding, and for being my best friend before being my mother. Thank you both for your unconditional love and support and for always trusting all my choices. To my two sisters, Délara and Hana, thank you for all your love and support, and for always having my back. Lastly, to my partner, Mido, thank you for all your love, kind words, and encouragement through those last few months of my Ph.D. journey.

Abstract

This thesis presents three essays investigating the impact of Internet use and international migration, as two globalization forces, on distinct political, social, and economic outcomes. The first chapter assesses the impact of using the Internet as a source of news on African citizens' perceptions of democracy. We use repeated cross-sectional data from the Afrobarometer survey across 35 African countries over the period 2011-2018, along with an instrumental variable approach, to address potential endogeneity bias. We find a negative significant effect of Internet use on both preference for and perception of the extent of democracy. This effect is channeled through a decreased trust in political institutions and the Internet acting as a potential misinformation channel. In the second chapter, we investigate the impact of immigration on cultural heterogeneity in Europe from 2004 to 2018 at the regional level. It combines European Social Survey data, to measure cultural heterogeneity across several cultural traits, with immigrant data from the European Labor Force Survey. Using shift-share IV strategy, our findings show that the cultural heterogeneity is negatively influenced by inflows of immigrants. The results indicate that low-skilled and non-European immigrants increase cultural heterogeneity by introducing new values in destination countries. However, this effect vanishes with assimilation and it is outweighed by the cultural reaction of natives. Finally, the third chapter examines the impact of male household members' migration on the labor supply of females left behind. Leveraging data from the last three waves of the Egypt Labor Market Panel Survey and employing a fixed effect estimation strategy, this study assesses the impact of both current and return male migration on female employment status in urban and rural regions. The results reveal that the presence of a current migrant increases the probability of engaging in wage work in urban areas and unpaid family work in rural areas. Additionally, the findings indicate that return migrants have a negative impact on female wage work, predominantly in rural areas. The underlying mechanisms shed light on the intricate interplay of economic factors and cultural norms that shape female labor supply in distinct geographic settings in response to male migration.

Keywords: Internet news, Democracy, Misinformation, Immigration, Cultural heterogeneity, Migration, Labor, Gender, Africa, Europe, Egypt.

Résumé

Cette thèse présente trois essais qui étudient l'impact de l'utilisation d'Internet et de la migration internationale, comme deux forces de globalization, sur différentes variables politiques, sociales et économique. Le premier chapitre évalue l'impact de l'utilisation d'Internet comme source d'information sur les perceptions de la démocratie par les citoyens africains. Nous utilisons des données transversales répétées provenant de l'enquête Afrobaromètre dans 35 pays africains sur la période 2011-2018, ainsi qu'une approche de variable instrumentale, pour traiter le biais d'endogénéité potentiel. Nous constatons un effet négatif significatif de l'utilisation d'Internet sur la préférence et la perception de l'étendue de la démocratie. Cet effet est canalisé par une diminution de la confiance dans les institutions politiques et par le fait qu'Internet agit comme un canal potentiel de mésinformation. Dans le second chapitre, nous étudions l'impact de l'immigration sur l'hétérogénéité culturelle en Europe de 2004 à 2018 au niveau régional. Il combine les données de l'European Social Survey, pour mesurer l'hétérogénéité culturelle à travers plusieurs traits culturels, avec les données sur les immigrants de l'European Labor Force Survey. En utilisant une variable instrumentale en shift-share, nos résultats montrent que l'hétérogénéité culturelle est influencée négativement par les flux d'immigrants. Les résultats indiquent que les immigrants peu qualifiés et non-européens augmentent l'hétérogénéité culturelle en introduisant de nouvelles valeurs dans les pays de destination. Toutefois, cet effet disparaît avec l'assimilation et est compensé par la réaction culturelle des natifs. Enfin, le troisième chapitre examine l'impact de la migration des membres masculins du ménage sur l'offre de travail des femmes restées au pays. En s'appuyant sur les données des trois dernières vagues de l'Egypt Labor Market Panel Survey et en utilisant une stratégie d'estimation à effet fixe, cette étude évalue l'impact de la migration actuelle et du retour des hommes sur le statut de l'emploi des femmes dans les régions urbaines et rurales. Les résultats révèlent que la présence d'un migrant actuel augmente la probabilité d'exercer un travail salarié dans les zones urbaines et un travail familial non rémunéré dans les zones rurales. En outre, les résultats indiquent que les migrants de retour ont un impact négatif sur le travail salarié des femmes, principalement dans les zones rurales. Les mécanismes sous-jacents mettent en lumière l'interaction complexe des facteurs économiques et des normes culturelles qui façonnent la participation des femmes au travail dans des contextes géographiques distincts en réponse à la migration masculine.

Mots Clés: Actualités d'Internet, Démocratie, Mésinformation, Immigration, Hétérogénéité culturelle, Migration, Travail, Genre, Afrique, Europe, Égypte.

Contents

List of Figures

List of Tables

List of Acronyms

ACLED	Armed Conflict Location and Event Data
CAPMAS	Egyptian Central Agency for Public Mobilization and Statistics
CF	Cultural Fractionalization
CLR	Conditional Likelihood Ratio
ELMPS	Egypt Labor Market Panel Survey
ERF	Economic Research Forum
ESS	European Social Survey
EULFS	European Labor Force Survey
EVS	European Value Survey
FOTP	Freedom of the Press
GSS	General Social Survey
IFES	International Foundation for Electoral Systems
ISPs	Internet Service Providers
IV	Instrumental Variable
IXPs	Internet Exchange Points
LATE	Local Average Treatment Effect
MCA	Multiple Correspondence Analysis
OLF	Out of the Labor Force
OLS	Ordinary Least Squares
PCA	Principal Component Analysis
PSU	Primary Sampling Unit
RoW	Regime of the World
SMCs	Submarine Cables
SSA	Sub-Saharan Africa
US	United States
WVS	World Value Survey

General introduction

In an era characterized by rapid globalization and technological progress, the impact of the Internet and international migration stands out as pivotal in shaping individuals' perceptions, social interactions, and economic outcomes. The Internet, as a borderless hub of information, serves as a virtual gateway surpassing physical boundaries. It facilitates the exchange of ideas and communication, ultimately leading to the world becoming a closely interconnected global village. At the same time, international migration, which involves the movement of people across geographical borders, introduces diversity that affects both host and sending communities. The chapters of this thesis aim to assess distinct political, social, and economic consequences of Internet use and international migration. Its specificity lies in its multidimensional approach, spanning diverse geographical settings, in studying the various impacts of those two distinct yet interconnected globalization forces that have gained increased attention in recent years.

The first chapter focuses on the political repercussions of Internet use as an alternative news source vis-à-vis traditional media, a realm significantly influenced by the spread of the Internet, and especially social media. This chapter, concentrating on the African context where political grievances in recent years are claimed to have been amplified by the use of the Internet, assesses the impact of Internet use on individuals' perception of democracy. In the second chapter, the focus is shifted to Europe which has seen an unprecedented increase in migration with the proliferation of conflicts and unstable economic problems, especially in Africa and the Middle East. The emphasis in this chapter is on the social implications of immigration in terms of cultural diversity, a recurrent concern evident in news headlines and political debates across several European countries. Finally, the last chapter, situated in the context of Egypt, examines the economic impacts of emigration on the sending communities. Specifically, it focuses on male emigration, a significant facet of migration in Egypt, and its effects on the female left behind labor supply.

The three chapters carry out careful empirical investigations, primarily leveraging rich survey data and employing thorough identification strategies to address the aforementioned research questions. The main data sources used include public attitudes surveys represented in the Afrobarometer survey employed in chapter 1 and the European Social Survey (ESS) employed in chapter 2, as well as labor force surveys such as the European Labor Force Survey (EULFS) and the Egypt Labor Market Panel Survey (ELMPS), in

chapter 2 and chapter 3, respectively. Identification challenges constitute mainly omitted variable bias, reverse causality, and selection bias which are addressed thoroughly using recent advancements in the instrumental variable (IV) literature by relying on a shift-share IV strategy in chapter 1 and chapter 2. However, in chapter 3, the panel dimension of the data is exploited to mitigate endogeneity bias.

This general introduction aims to provide an overview of the evolution of Internet use and international migration in recent years. Additionally, it offers a brief, non-exhaustive, overview of the economic literature that underpins the questions studied in this thesis. Furthermore, it aims to highlight the contributions of each chapter relative to the existing literature. The introduction is organized as follows. In section I, I present the evolution of Internet adoption, placing particular emphasis on Africa. I also briefly present the literature on the political effects of the Internet. Moving to section II, the focus shifts to an exploration of international migration, with a special focus on Europe as a host society. I also discuss research advancements related to the impact of immigration on host societies, specifically highlighting social impacts. Finally, in section III, I offer a brief presentation of the Egyptian emigration profile. Additionally, I address the literature on male migration and its repercussions on those left behind. At the end of each section, I underscore the contribution of the chapters of this thesis.

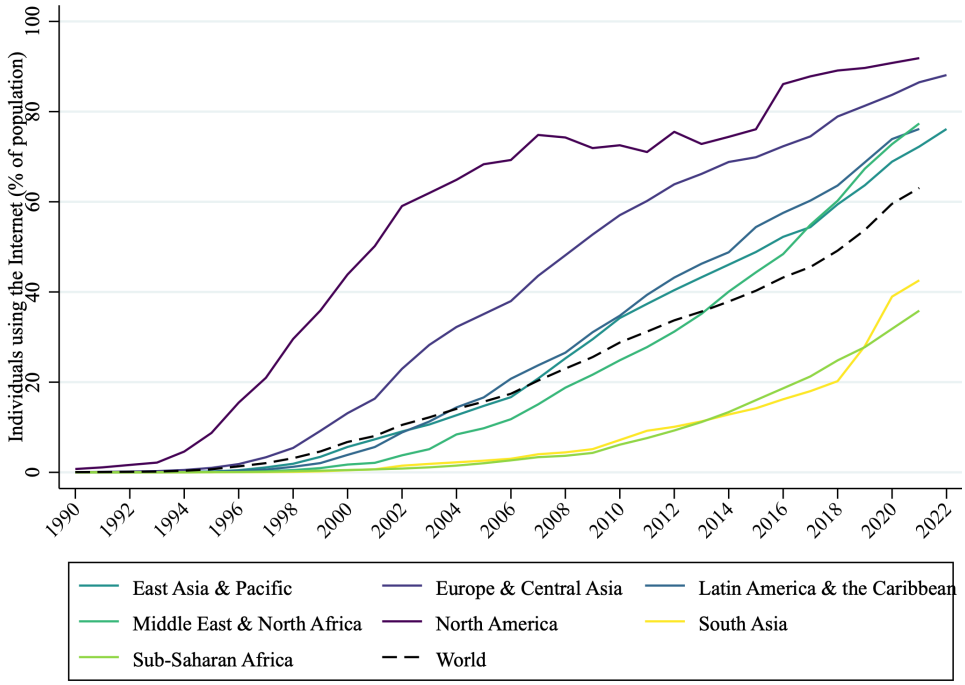
I Internet use and political implications

The evolution of Internet use

From its origin as a decentralized communication system during the Cold War, the Internet has grown to become one of the key information and communication technologies in the world. Today, nearly 5.35 billion are using the Internet worldwide, representing 66.2% of the global population as of January 2024. Internet penetration has not evolved at the same time in all the regions of the world. According to Fig. 1 below, the Internet has started to rise in North America, Europe, and Central Asia from the early 90s, in East Asia & Pacific, Latin America & the Caribbean, and Middle East & North Africa in the early 2000s, and it wasn't until the late 2000s that it has started to thrive in Sub-Saharan Africa and South Asia. Despite being a latecomer to digitization, Internet use has evolved rapidly in Africa growing from 6% in 2010 to 35% in 2022.

The increase in Internet penetration in Sub-Saharan Africa happened as a result of the improvement of the telecommunication and Internet infrastructure with the increase in

Fig. 1. The evolution of Internet use over time.



Note: This figure shows the percentage of Internet users over the period 1990-2022 by region.
Source: Author's elaboration based on data from the World Development Indicators (1990-2022).

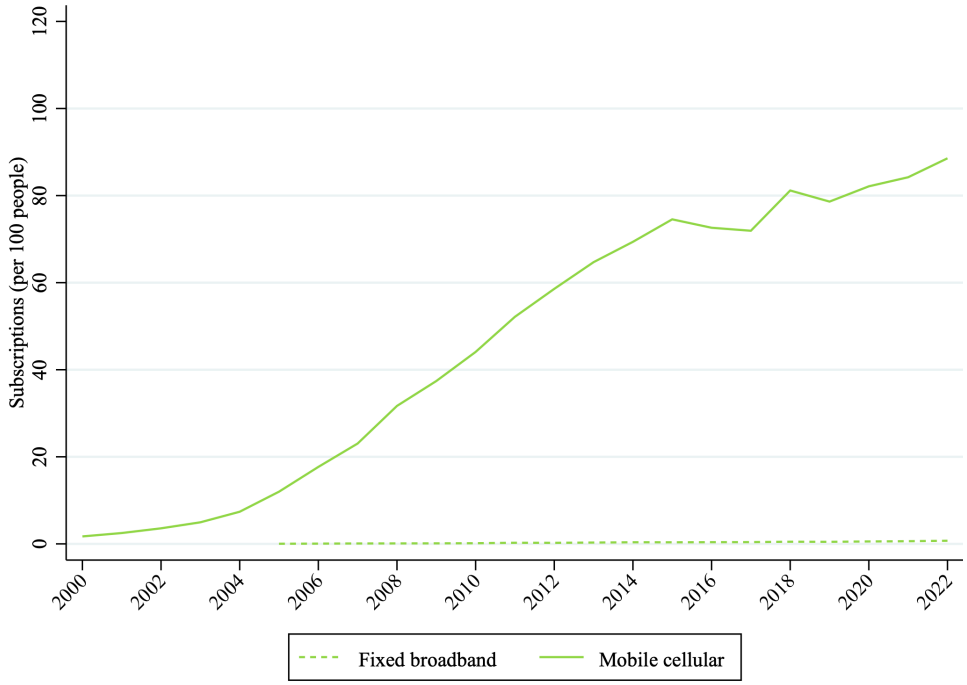
the deployment of submarine cables from the year 2010. There has also been investment in Internet exchange points to keep Internet traffic locally. In addition, satellite systems have been adopted to improve remote areas connection. And while fixed broadband subscriptions are still very limited, the population in Africa mainly relies on mobile cellular subscriptions as a last-mile Internet infrastructure. As depicted in Fig. 2, the number of individuals with mobile cellular network has jumped from 2 in 2000 to 88.5 (per 100 inhabitants) in 2022 in Sub-Saharan Africa, with only less than 1 per 100 inhabitants having a fixed broadband subscription over the whole period.

The increase in mobile subscriptions has been the result of the share of the continent covered with 2G or 3G mobile coverage increasing substantially over the last decade. Despite recent advancements, Internet and telecommunication infrastructure in Africa are still underdeveloped and Internet penetration remains way below the world average.

The political effects of the Internet

With the expansion of the Internet and the use of social media, many scholars have referred to the Internet as being a "liberation technology" (Diamond, 2010). Due to

Fig. 2. The evolution of fixed broadband and mobile cellular subscriptions over time in Sub-Saharan Africa.

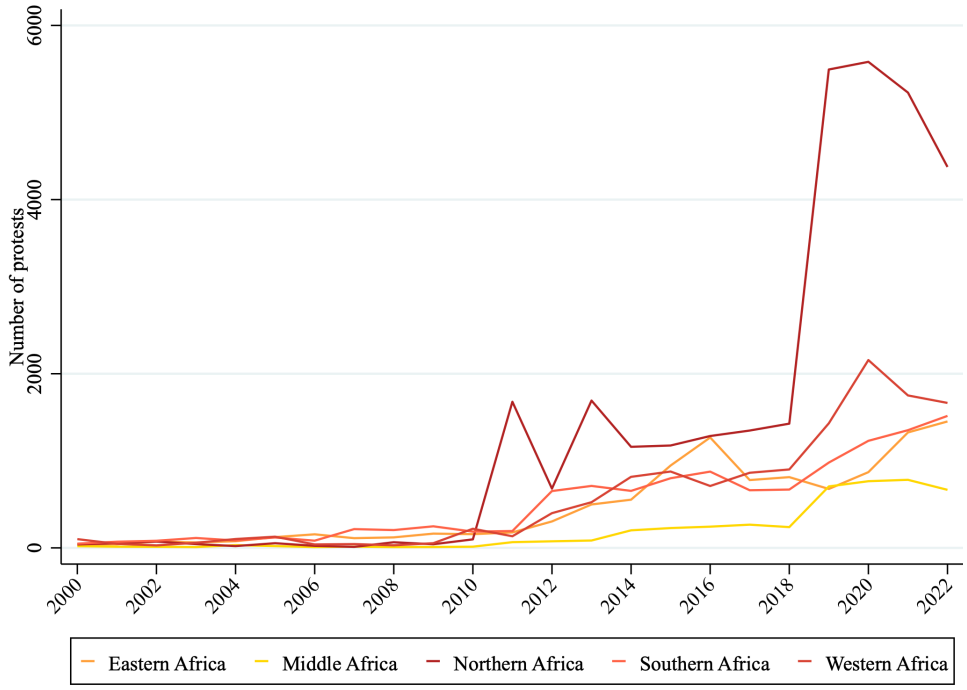


Note: This figure shows the fixed broadband and mobile cellular subscriptions (per 100 people) over the period 2000-2022.
Source: Author's elaboration based on data from the World Development Indicators (2000-2022).

its low cost, open access, and decentralized nature, the Internet may be a freer source of information critical of the government, which makes it harder for political actors to hide harmful information and makes them more vulnerable and accountable (Zhuravskaya et al., 2020). This results in more informed citizens, dissatisfied with their government, and prone to participate in political protests. Compared to traditional media, the Internet is characterized by its low entry barriers and user-generated content which allow easier spread of political information and access to world news and events, especially in countries with limited freedom of speech. Nevertheless, despite being a potential open source of information, the Internet is also prone to being a "misinformation technology". Online media give their users the ability to re-post and re-share content created by others without any fact-checking procedures, which undermines the quality of information spread online. Indeed, some studies report the propagation of false news online (Mocanu et al., 2015; Allcott and Gentzkow, 2017; Grinberg et al., 2019) and that they spread faster and more broadly than true news (Vosoughi et al., 2018). Moreover, the Internet can also be used by autocratic regimes to manipulate information and create propaganda (King et al., 2017).

Several works study the impact of the expansion of the Internet and social media on street protests suggesting that it facilitates the organization and coordination of mass mobilization through the horizontal flow of information which exacerbates the outbreak of protest movements. For instance, Fergusson and Molina (2019) show that new releases of Facebook in a specific language are associated with increased protests in countries where this language is spoken. This has been particularly true in Africa since it has witnessed a rise in public discontent portrayed in a spike in the number of protest movements in the continent during the last decade. Fig. 3 below shows the uprising number of protests from the year 2010 onward. The most notable one is the Arab Spring that took place between 2010 and 2012 in North Africa, and where the Internet and social media played a major organizational role.

Fig. 3. The number of protests over time.



Note: This figure shows the number of protests over the period 2000-2022 in Africa.
Source: Author's elaboration based on data from ACLED (2000-2022).

In that vein, Acemoglu et al. (2018) documents a positive association between Twitter activity and the number of protesters in the Egyptian context, during the Arab Spring, reflecting the role of social media in the coordination of street protests. Steinert-Threlkeld et al. (2015) find similar results using geolocated tweets and protest data from 16 countries during the Arab Spring. Manacorda and Tesei (2020) also find that the expansion of

the 2G mobile network in Africa increases anti-government protests during economic downturns. They highlight the role played by the Internet as both an information and coordination channel. More recently, Guiffard (2022) studies the staggered arrival of submarine cables in Africa and its impact on political mobilization. He documents a positive impact channeled through the enhanced coordination channel. The same pattern is also observed in other parts of the world, including China (Qin et al., 2021), Russia (Enikolopov et al., 2020), and the United States (Amorim et al., 2022).

Beside its impact on the rise of protest movements, there is a vast literature studying the impact of the Internet on other political aspects in both democracies and autocracies. This includes a wide range of studies investigating its impact on voting behavior. In established democracies, authors mostly document a decrease in voting turnout with Internet penetration due to the substitution of political news with entertainment content (Falck et al., 2014; Gavazza et al., 2019; Campante et al., 2018). However, Campante et al. (2018) find that this pattern was reversed in Italy after 2008, presumably with the introduction of social networks. Some authors also find the Internet to be associated with the rise of populism in Europe (Schaub and Morisi, 2020; Guriev et al., 2021; Tabellini et al., 2023). Nevertheless, in autocracies, the Internet was mostly associated with a decreased support for incumbent governments (Miner, 2015; Donati, 2023; Guriev et al., 2021).

In addition, some scholars study the relationship between the Internet and democracy. To date, most of the quantitative research on the topic has focused on the macro relationship between Internet penetration and the level of democracy, documenting a positive association, but with little insights on the underlying mechanisms (Evans, 2019; Jha and Kodila-Tedika, 2020). In the last two decades, democracy scholars started to emphasize how citizens' attitudes toward democracy and access to information technology shape the democratization process. For democracies to emerge and survive, two conditions need to be met. First, citizens need to choose and support democracy as their preferred governance regime (Claassen, 2020). Indeed, increased demand for democracy promotes democratization through the "bottom-up" process where pressure is exerted by citizens over authoritarian regimes (Lei, 2013). Second, they also have to believe that they are getting democracy (Mattes and Bratton, 2007). Studies evaluating individual Internet use and attitudes toward democracy are limited and are largely conceptual works documenting a simple correlation. The latter find that the Internet, in general, and social media, in particular, are associated with lower satisfaction with democracy (Bailard, 2012; Ceron and Memoli, 2016; Chang, 2018).

Contribution of Chapter 1

The contribution of the first chapter of this thesis is to provide a micro-econometric approach running from individuals' Internet use to citizens' attitudes toward democracy. Acemoglu et al. (2021) emphasizes the crucial need to explore the role of media in shaping support for democracy, especially in light of the spread of misinformation from various media outlets and social media platforms, and how this may affect the relationship between successful democratic performance and public support for democracy. To the best of our knowledge, this is the first study to carry out a careful empirical investigation of the causality running from the Internet as a news source to citizens' attitudes and perceptions of democracy in Africa. Specifically, we investigate its impact on their preference for democracy, their perception of the extent of democracy in their countries, and their satisfaction with how democracy works. We also provide an in-depth analysis of the underlying mechanisms exploiting the role played by the Internet as an information or misinformation channel. We study its impact on confidence in governments, perception of corruption, political participation, as well as misinformation. Hence, this chapter adds to the debate on the status of the Internet as being a "liberation" versus a "misinformation" technology.

We rely on the Afrobarometer which is a public attitude survey with a wide variety of questions on democracy, governance, and other related issues in Africa. Data in the Afrobarometer are collected through face-to-face interviews with a random sample of 1,200 or 2,400 in each country. We use rounds 5, 6, and 7 of this survey since previous rounds did not include questions on Internet use. We end up with a repeated cross-section of 99,938 African citizens across 35 countries covering the period between 2011 and 2018.

Our main concern is the endogeneity bias that stems from omitted variables and the bi-directional relationship between Internet news consumption and democracy variables. To address this, we adopt an interacted or composite instrumental variable design following Borusyak et al. (2022) where instruments are constructed as aggregate shocks weighted by lower-level exposure factors. In this setting, the validity of our instrument stems from exogeneous variation in the shocks while allowing for endogenous variation in the exposure factor. This also relates to the economic literature on the shift-share instrument (Borusyak et al., 2022; Goldsmith-Pinkham et al., 2020). More specifically, our instrument consists in combining the number of submarine cables in a country (which is considered as the exogenous shock) with the share of the district covered with a 3G network (which is a potentially endogenous weighting factor). To further ensure that our identification results

from the exogenous shock, we fix the district's 3G coverage to its value at the time of the first Afrobarometer survey wave used in the analysis.

The results indicate that using the Internet to obtain information has a significant negative effect on both the preference for and the perception of the extent of democracy. This negative effect is due to several factors. First, Internet use erodes trust in government institutions, mainly in the parliament and the ruling party. It increases the perception that parliament members are involved in corruption. In addition, the erosion of trust is correlated with more political mobilization, in the form of greater participation in demonstrations and voting. These results echo the existing literature and, in particular, hint at the risks of reversal of nascent democratization processes. Finally, the Internet seems to act as a misinformation channel. On the one hand, Internet users' perception of the extent of democracy and perception of the corruption of legislators diverge from experts' assessments. On the other hand, Internet use increases the likelihood of inconsistency in respondents' stances on their preference for democracy. The Internet is not a neutral information channel: it tends to undermine citizens' preference for democracy while also altering perceptions about political institutions.

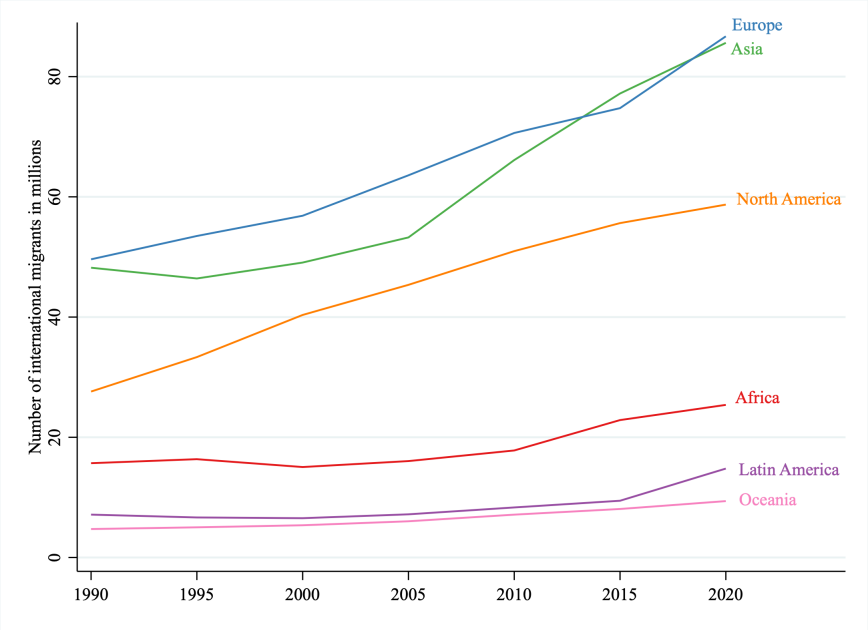
II Immigration and culture

The evolution of immigration rates

With globalization, not only has the population of Internet users expanded significantly in recent decades, but there has also been a notable rise in international migration. According to the latest estimates published by the UN Population Division in 2020, the current number of international migrants stands at 281 million, accounting for approximately 3.6% of the world's population. This surge in migration is evident across all regions of the world, with particularly pronounced increases observed in Asia and Europe. As illustrated in Fig. 4 below, Asia and Europe now host 86 million and 87 million international migrants respectively, collectively representing over 50% of the global international migrant stock in 2020.

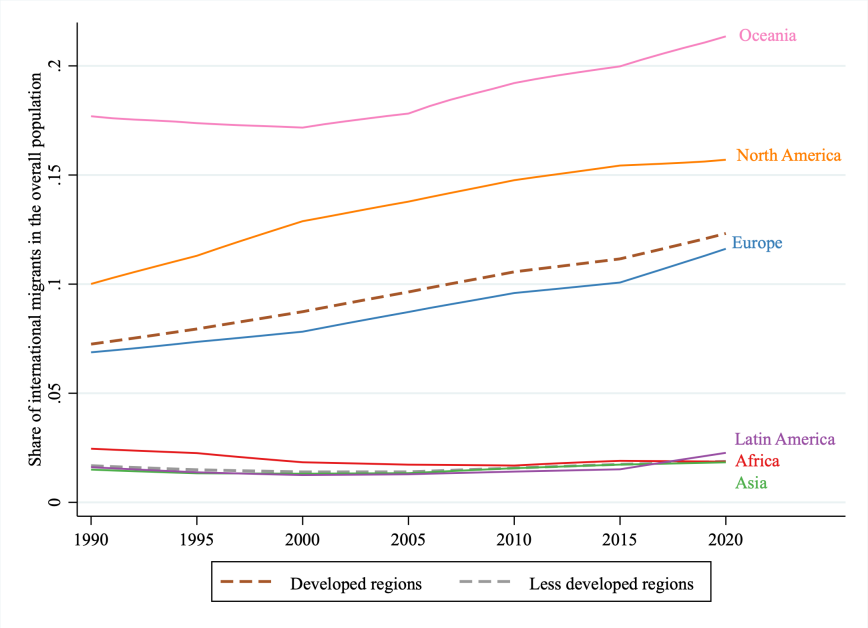
During the period spanning from 2000 to 2020, Asia experienced the most substantial growth, witnessing a remarkable increase of 74%, equivalent to approximately 37 million immigrants. Following closely, Europe saw the second-largest surge, adding 30 million international migrants to its population. However, when examining the proportion of migrants relative to the overall population, as depicted in Fig. 5, distinct patterns emerge.

Fig. 4. The number of international migrants over 1990-2020.



Note: This figure shows the evolution of the number of international migrants by region between 1990 and 2020.
Source: Author's calculation based on data from the United Nations Population Division (1990-2020).

Fig. 5. The share of international migrants over 1990-2020.



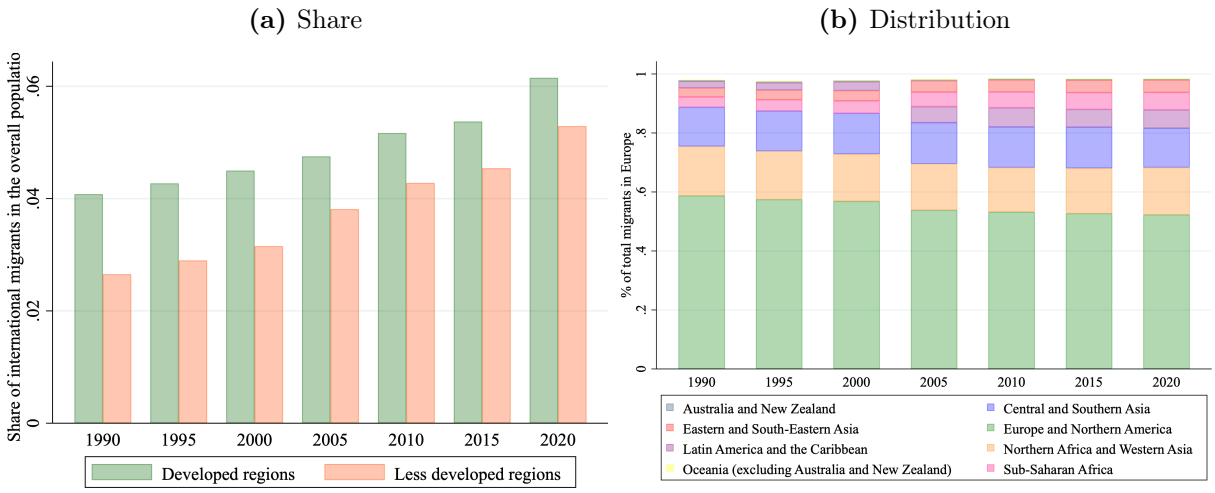
Note: This figure shows the evolution of the share of international migrants by region between 1990 and 2020.
Source: Author's calculation based on data from the United Nations Population Division (1990-2020).

The share of international migrants has consistently risen in the developed world throughout the entire period, in contrast to the developing world where it initially declined before slightly increasing. Notably, Oceania, North America, and Europe boast the highest proportions of international migrants, comprising 22%, 16%, and 12% of their respective populations. Conversely, Latin America, Africa, and Asia exhibit comparatively lower shares, each constituting less than 3% of their populations.

Examining Europe specifically, it's evident that the continent hosts the largest number of international migrants, with the proportion of migrants relative to the overall population steadily increasing from 7% in 1900 to 12% in 2020. This surge can be attributed to various factors, including the Balkan War in 1990, the significant EU enlargement in 2004, and the 2015 refugee crisis, which saw people fleeing poverty and conflict in the Middle East and Africa.

Analysis of the share of migrants by origin group, as shown in Fig. 6(a) reveals an increase in the share of migrants originating from both developed and less developed regions. Notably, the latter accounted for 55% of the increase in the number of international migrants in Europe between 1990 and 2020. Furthermore, examining the distribution of immigrants within Europe by origin region, Fig. 6(b) illustrates a declining share of international migrants coming from Europe and North America, accompanied by a rise in migrants from other regions.

Fig. 6. International migrants in Europe by origin group over 1990-2020.



Notes: This figure shows the evolution of the share of international migrants from developed and less developed regions in Europe between 1990 and 2020 in (a) and their distribution by origin regions over the same period in (b). The developed and less developed regions are based on the UN development group definition.
Source: Author's calculation based on data from the United Nations Population Division (1990-2020).

In summary, Europe in particular, and the rest of the developed world in general, have experienced a significant surge in its migrant population, with individuals now hailing from economically, geographically, and culturally diverse backgrounds, contributing to increased diversity within host communities. This phenomenon has prompted scholars to study the economic, political, and social implications of migration on host societies.

The impact of immigration on host societies

A burgeoning body of literature has emerged studying mainly immigrants' impact on the labor market rising from the natives' concern that immigrants could reduce their wages and limit their employment opportunities. Most studies focus on the impact on natives' wages with little consensus. While some studies document a negative effect (Borjas, 2003; Ortega and Verdugo; Altonji and Card, 2018), others find a null or positive average effect of immigration on natives' wage, particularly in the long run (Ottaviano and Peri, 2012; Manacorda et al., 2012; Borjas, 2014; Edo and Toubal, 2015). Most studies highlight that the skill composition of the immigrant workforce and their degree of substitutability with the natives matter. Moreover, natural experiments studies find that the initial and short-term effect of immigration can be negative depending on the speed of labor market adjustments (Hunt, 1992; Borjas, 2017; Borjas and Monras, 2017; Monras, 2020; Prantl and Spitz-Oener, 2020). In sum, the average effect of immigration on natives' wages, either positive or negative, is found to be negligible (see Edo (2019) for an extensive review of this literature).

Besides resulting natives' economic concerns, immigration has also given rise to cultural concerns. Many studies show the skepticism toward immigration is primarily driven by cultural considerations rather than labor market ones (Dustmann and Preston, 2007a; Card et al., 2012; Poutvaara and Steinhardt, 2018). This has recently been a lively debate in Europe with the recent refugee crisis in 2015 and the growing inflow of migrants coming from geographically, culturally, and economically distant countries. There has been a growing fear that immigration might constitute a threat to the Western culture with migrants bringing in new values and norms affecting host countries' cultural composition. The latter observation has been highlighted by the right-wing movements creating further natives' skepticism toward immigrants and refugees on the basis of cultural considerations. Several studies report a positive link between immigrants' or refugees' shares and voting for anti-immigration and right-wing parties in Europe (Otto and Steinhardt, 2014; Halla et al., 2017; Becker et al., 2016; Edo et al., 2019a). In contrast, some also find heterogeneous effects according to the area of residence (i.e., urban vs rural) (Dustmann

et al., 2019), type of exposure (Steinmayr, 2017), and economic conditions (Tomberg et al., 2021).

Hence, the impact of immigration on host societies' culture has gained increased attention in recent decades. The literature on the latter has identified two main channels through which immigration can affect culture. First, there is a direct effect which is the compositional effect channel where the mere addition of individuals with distinct values to a society affects its composition. The magnitude of this effect would depend on the inter-generational transmission of those values, in other words, how persistent they are and how they are passed on to the next generations (Bisin and Verdier, 2001; Desmet et al., 2017; Galli and Russo, 2019; Desmet and Wacziarg, 2021) and the rate of cultural assimilation (Algan et al., 2012; Abramitzky et al., 2014; Giavazzi et al., 2019; Gonnot and lo Polito, 2021; Fouka et al., 2022; Abramitzky and Boustan, 2022; Gonnot and lo Polito, 2023). Second, there is an indirect effect through the natives' reaction. By interacting with the native population, immigrants might themselves influence the native's culture (Fisman and Miguel, 2007; Schmitz and Weinhardt, 2019; Tabellini, 2020; Giuliano and Nunn, 2021; Miho et al., 2023). Here, the effect on the overall cultural diversity, would depend on how natives react. They could absorb the new values, but they could also have a backlash reaction especially if the migrants are coming from culturally distant countries (Dustmann and Preston, 2007b; Edo et al., 2019a; Steinmayr, 2021; Alesina and Tabellini, 2022; Moriconi et al., 2022; Keita et al., 2023).

Rapoport et al. (2020) combine the aforementioned mechanisms in a unified theoretical framework and test the effect of migration on cultural similarity between origin and destination countries in a global context. They find evidence of a positive correlation between migration and cultural similarity mainly driven by cultural remittances, i.e., immigrants disseminating the host society's culture to their home communities. While their study focuses on overall cultural convergence, it does not tackle the impact of immigration on cultural diversity within host countries.

Contribution of Chapter 2

The contribution of Chapter 2 is thus to investigate immigration's impact on overall cultural heterogeneity in host countries in Europe. Hence, we contribute to two strands of literature. First the literature on how immigration affects culture in host countries. Our study complements the study by (Rapoport et al., 2020) on the impact of immigration on cultural proximity between country pairs at the global level, by focusing on the impact

on within host countries' cultural heterogeneity. It also aligns with the growing literature on the evolution of cultural divides in Western societies which studies whether identity cleavages are good predictors of individuals' attitudes and values (Desmet et al., 2017; Desmet and Wacziarg, 2021). While this literature mainly focuses on age, education, political orientation, race, and ethnicity as identity cleavages, we introduce birthplace as a relevant identity marker in explaining cultural diversity.

To do so, we compare changes in migrants' shares within European regions to changes in cultural heterogeneity. To measure the latter, we adopt Desmet and Wacziarg (2021)'s index of average fractionalization of the population across various cultural traits. It captures the likelihood that two randomly chosen individuals from the entire resident population of a given country hold a different variant of a randomly chosen trait. We combine data from the European Social Survey and the European Labor Force Survey over the period 2004-2018. To address self-selection and unobserved heterogeneity concerns, we rely on a shift-share instrumental variable strategy to predict exogenous stocks of migrants based on the assumption that new inflows of migrants from a given origin are allocated across regions based on the initial distribution of migrants from this same origin.

We find that immigration affects regional cultural heterogeneity in the host countries, an increase of immigration reduces cultural heterogeneity. We find evidence of both a compositional effect with immigrants bringing new cultural values to the host societies increasing cultural heterogeneity. Though, the natives' response outweighs the compositional effect resulting in an overall reduction of cultural heterogeneity. The compositional effect is found to be driven mainly by low-skilled and short-term migrants on issues related to sexual morality, religiosity, and the role of the state, and with migrants being more conservative on these issues compared to natives. However, this disappears within 10 years of their arrival suggesting the idea of cultural assimilation. While on the natives' response side, we find that it is mainly driven by migrants coming from other EU countries and high-skilled migrants on questions related to openness, cultural capital, and migration attitude. We find that natives exhibit more liberal attitudes in general with more migration.

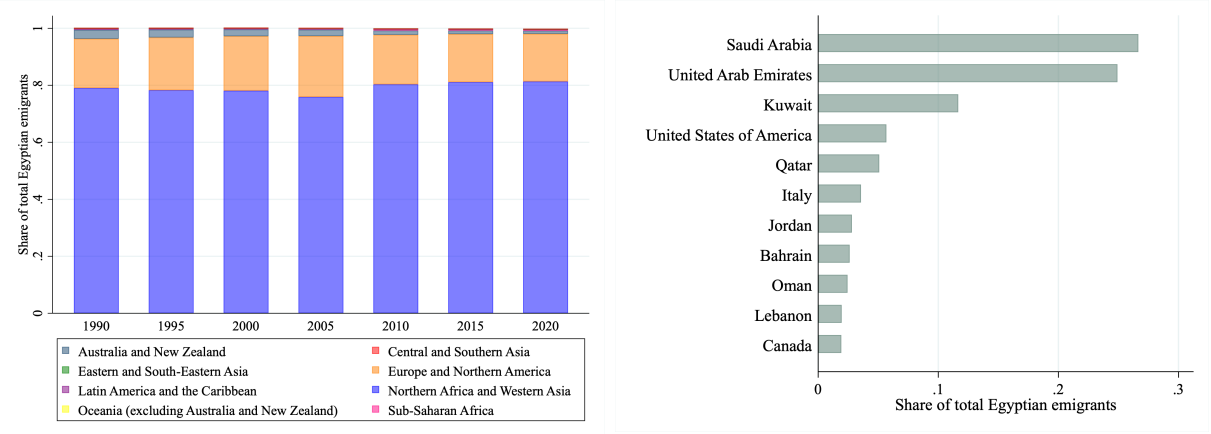
III Male migration and females left behind

Egypt's Migration Profile

While Europe serves as a significant immigration hub, Egypt stands out as the largest migrant-sending nation in the Middle East and North Africa (MENA) region, supplying the most emigrant labor to oil-producing countries in the Middle East (David et al., 2019).[1] The last Egyptian Census carried out in 2017 indicates that approximately 9.4 million Egyptian migrants reside outside the country (CAPMAS, 2017). The distribution of Egyptian emigrants across global regions over time, depicted in Fig. 7(a), reveals a concentration in Arab and Gulf states, where over 80% of emigrants reside in Northern Africa and Western Asia. The second most popular destinations for Egyptians are Northern America and Europe. These patterns have remained the same over the last three decades. Analysis of the primary destination countries for Egyptians in 2020, as presented in Fig. 7(b), highlights Saudi Arabia and the United Arab Emirates as hosting over 50% of all Egyptian emigrants, followed by Kuwait. In Europe and Northern America, the United States, Italy, and Canada emerge as the most sought-after destinations.

Fig. 7. Egyptian emigrants destinations.

(a) Distribution across regions. **(b)** Main destinations in 2020



Note: This figure shows the evolution of the distribution of Egyptian emigrants across regions between 1990 and 2020 in (a) and the main destination countries in 2020 in (b).
Source: Author's calculation based on data from the United Nations Population Division (1990-2020).

Studies indicate that Egyptian emigrants, typically young men from rural areas, exhibit

[1] A brief overview of Egyptian migration history is provided in chapter 3 in section 3.2.

varied educational and occupational statuses depending on their destination countries. Those migrating to the Gulf states often secure employment in white-collar, skilled, and semi-skilled technical fields, whereas those heading to Jordan or Libya are more likely to find low-skilled work in agriculture, construction, and services sectors (De Bel-Air, 2016). Moreover, emigrants to Western countries display differing education levels, with those bound for Southern Europe generally possessing lower levels of education compared to their counterparts heading to Northern America, who tend to be highly educated.

A notable characteristic of Egyptian emigration to the West is its predominantly permanent nature, often involving entire households, in contrast to migration within the Arab region, which is typically male-dominated and primarily temporary due to short-term work contracts and limited citizenship rights (Wahba, 2015b; Tsourapas, 2022). Alongside its temporary nature, various domestic and international factors have contributed to significant waves of return migration from the Gulf, Iraq, and Libya since the late 1980s (Zohry, 2003). These factors include the Iran-Iraq War in 1988 and subsequent declines in oil prices, the First Gulf War in 1990-1991, policies favoring national labor over foreign workers, reduced demand for construction labor in Arab countries, and the Iraq War in 2003. Additionally, there was a notable influx of returning Egyptian migrants from Libya following the collapse of Gaddafi's regime during the Arab Spring (Zohry, 2013; De Bel-Air, 2016). According to the 2018 Egypt Labor Market Panel Survey (ELMPS), approximately 7% of the working-age population were returnees.

Male migration impact on females left behind

As migrants can significantly influence destination countries, they also have profound impacts on their sending communities. Both outward and return migration of male household members can yield significant effects on families left behind, presenting a mix of positive and negative outcomes. Some studies focus on the influence of migrant absence on school attainment and educational performance among children left behind (Yang, 2008; Antman, 2012; Alcaraz et al., 2012). They suggest that the additional income received by households through remittances diminishes the need for child labor and enhances children's access to education, particularly benefiting girls in developing nations. However, other research indicates that the absence of the primary caregiver can elevate the likelihood of children dropping out of school and hinder their educational progression. Limited studies on the health impacts of family members left behind suggest that remittances can enhance families' healthcare and nutrition (Gibson et al., 2011; Mu and De Brauw, 2015). Recent literature on social and family dynamics highlights increased bargaining power

among females left behind (De Haas and Van Rooij, 2010), as well as variations in children's living arrangements (Bertoli et al., 2023). Some studies even suggest that migration of household members influences political (Spilimbergo, 2009; Chauvet and Mercier, 2014; Barsbai et al., 2017), fertility (Beine et al., 2013; Bertoli and Marchetta, 2015), and gender norms (Tuccio and Wahba, 2018; Diabate et al., 2019; Samari, 2021) transmission from destination countries to sending ones.

Moreover, there exists a vast literature on the labor supply responses of family members staying behind. Most studies in this domain focus on the labor supply of wives when their husbands migrate.[2] The majority of these studies indicate a decrease in female labor force participation overall but note an increase in unpaid family work, particularly in rural areas. However, some studies report a slight increase or no effect at all. Nevertheless, there is limited evidence regarding the impact of return migration on female labor supply.

Studies examining the impact of current migration identify two primary channels through which the migration of a male household member may influence female labor force participation. First, there is an income effect where the migration of a household member can increase household income through remittances, potentially leading to an increase in the reservation wage and a lower opportunity cost of leisure, thereby reducing the need for females to work outside the home and acting as an insurance mechanism (Chami et al., 2005). Several studies support this e.g., Amuedo-Dorantes and Pozo (2006), Mendola and Carletto (2012) Binzel and Assaad (2011) and Lenoël and David (2019). Another potential channel is the substitution effect, wherein the absence of the primary breadwinner prompts a redistribution of labor within the household by the individuals left behind to compensate for the lost labor or income of the migrant. Research suggests that this substitution effect primarily manifests as increased unpaid family work, especially in rural areas (Binzel and Assaad, 2011; Lenoël and David, 2019).

In contrast to the extensive literature on the impact of current migrant members, the influence of return migrants in the household has received comparatively less attention. One would expect that return migration could cancel-out the effects that migration had on female labor force participation. On the income effect side, the cessation of remittances increases financial pressure, potentially prompting females to enter the labor market. Regarding the substitution effect, return migrants may reduce female labor force participation if they resume their traditional household duties upon their return, leading to reduced unpaid family work for females left behind. However, return migration can

[2] This literature is detailed in chapter 3 in section 3.1.

also bring back capital that may enable females to pursue self-employment, and it may influence female labor force participation through the transmission of gender norms from destination to origin countries. Return migrants may bring back both human and physical capital to be reallocated or invested by household members (Gubert and Nordman, 2011; Wahba, 2021). Consequently, females may use the capital acquired by men during migration spells to pursue self-employment (Mendola and Carletto, 2012). Additionally, as mentioned above there is a growing body of literature on the transmission of social norms from destination to origin countries, particularly regarding gender norms, which may influence female labor force participation upon return (Tuccio and Wahba, 2018). Overall, the impact of both current and return male migration on female labor force participation among those left behind is complex and relies on various interdependent mechanisms.

Contribution of Chapter 3

The contribution of this chapter is to provide a holistic approach to studying the impact of both current and return male migration on female labor force participation and employment status in Egypt, while considering the geographical context of urban and rural areas. It contributes to the aforementioned literature in several ways. Methodologically, it employs a three-waves panel dataset to better control for unobserved heterogeneity, contrasting with prior cross-sectional analyses. Empirically, it addresses both current and return migration impacts, thus filling a gap in understanding the effects of return migration on female labor supply. And analytically, it explores economic and cultural factors often overlooked in previous studies.

Using data from the last three waves of the Egypt Labor Market Panel Survey (2006, 2012, 2018), I adopt a fixed-effect estimation strategy with a comprehensive set of controls to mitigate self-selection concerns. Time-varying controls account for observable selection, while individual fixed effects address time-invariant unobserved heterogeneity at the individual level, also serving as proxies for couple or family fixed effects. Despite these controls, potential bias from omitted time-varying variables is further assessed using Oster's methodology.

Results reveal that the presence of a current migrant increases the probability of engaging in wage work in urban areas and unpaid family work in rural areas, driven by migrants who do not send remittances, and those with longer stays abroad for the former, and shorter stays abroad for the latter. Conversely, return migrants show no significant impact on female labor supply in urban areas but negatively affect female wage work in rural areas,

particularly among economically better-off migrants, with suggestive evidence of norms transmission. Heterogeneous effects based on the relationship status of the migrant to the female left behind are also observed. These findings underscore the necessity of considering diverse employment categories and relationship statuses within specific geographic settings to fully comprehend females' economic responses to male migration.

Chapter 1

Misinformation Technology: Internet Use and Political Misperceptions in Africa

This chapter is a joint work with Mathilde Maurel (Université Paris 1 Panthéon-Sorbonne, CES, France) and Joël Cariolle (FERDI, Université Clermont Auvergne, CNRS, IRD, CERDI, France).[1] A slightly different version is currently in press at *Journal of Comparative Economics* and available online at `https://doi.org/10.1016/j.jce.2024.01.002`.

1.1 Introduction

The Internet has significantly expanded worldwide, changing our relationship with the world, and the way we communicate, educate, and inform ourselves. Africa, despite having a very low number of fixed-broadband subscriptions for 100 inhabitants, has not escaped the Internet phenomenon, as the number of individuals with Internet access has risen from 2 in 2002 to 39.7 (per 100 inhabitants) in 2022. Similarly, the number of individuals with mobile-cellular telephone subscriptions has jumped from 12.4 in 2002 to 86.3 (per 100 inhabitants) in 2022 (ITU, 2022).

This phenomenon has considerable repercussions, which are widely studied by the social sciences. Our research aims to contribute to the ongoing scholarly discourse on the impact of the Internet on African democracies by examining how the Internet's role in information

[1] We would like to thank Olivia Bertelli, Simone Bertoli, Lisa Chauvet, Axel Dreher, Flore Gubert, Emiliano Grossman, Malka Guillot, Peter Hull, Irène Hu, Xavier Jaravel, Florian Léon, Oyebola Okunogbe, Patricia Schafer, Arthur Silve, Jérôme Valette, numerous participants at the 3rd International Conference on Globalization and Development (GLAD, Göttingen), the Doctoriales Development workshop (Paris), the 2022 European Public Choice Society Meeting (EPCS, Braga), the 21st Journées Louis-André Gérard-Varet (LAGV, Marseille), the 2022 International Conference on Development Economics (ICDE, Clermont-Ferrand), the 5th workshop on Economics & Politics (EP@L, Lille), the CSAE Conference 2023 (CSAE, Oxford), and the CERDI Research Seminar (CERDI, Clermont-Ferrand), as well as the Editor and three anonymous referees at the Journal of Comparative Economics for their valuable comments and suggestions. We are also grateful to Olivier Santoni for his contribution in calculating network coverage using data provided by GSMA-Collins Bartholomew's Mobile Coverage Explorer, and Alban Cornier for calculating the number of protests using the Armed Conflict Location & Event Data Project (ACLED). This work was supported by the Agence Nationale de la Recherche of the French government through the programme 'Investissements d'avenir' (ANR-10-LABX-14-01), through the IDGM initiative led by Ferdi (Fondation pour les études et recherches sur le développement international).

processing is reshaping African citizens' relationship with democracy. This is a key issue since the alleged superiority of democracy depends on individuals' ability to access and process information accurately (Flynn et al., 2017). Our study adds to the debate on the status of the Internet as a technology of liberation, misinformation, or disinformation, which was particularly relevant in the wake of the Arab Spring: while some consider the Internet a "liberation technology" that serves as an alternative to traditional media by providing a more open and freer source of information including blogs and social network sites, especially in countries with limited freedom of speech (Diamond, 2010), others believe that it is a "misinformation technology" used by non-democratic regimes, but also anti-establishment political actors, for fake news dissemination, surveillance, and propaganda (Qin et al., 2017).

The importance of the Internet debate explains why, in the last two decades, democracy scholars have tried to disentangle how access to information technology shapes the democratization process. Information technology has the potential to play a role in either consolidating or undermining democracy (Evans, 2019; Jha and Kodila-Tedika, 2020) by modeling perceptions and shaping preferences and opinions toward it. Our paper contributes to this literature by conducting an empirical analysis of how accessing news through the Internet affects African citizens' preference for, perception of, and satisfaction with democratic governance. We use three rounds of the Afrobarometer survey from 2011 to 2018, consisting of 99,938 individuals living in 1,845 districts across 35 African countries that have recently experienced a surge in democratic movements. Our main empirical concern is the possible endogeneity bias that stems from the bi-directional relationship between Internet news consumption and citizens' preference for and perception of democracy, as well as omitted variable bias.

To mitigate this concern, we rely on an empirical strategy combining an external and exogenous source with an internal source of digital vulnerability as an instrumental variable (IV) for Internet use. Specifically, we use the number of submarine cables (SMCs) as an exogenous and aggregate source of variation in citizens' access to the Internet, as in Cariolle and Le Goff (2023). We then weigh this aggregate connectivity shock by a fixed exposure factor that reflects citizens' access to mobile Internet, namely their district's 3G network coverage at the time of the first Afrobarometer survey wave used in our analysis. To minimize the risk of omitted variable bias, we include year and district fixed effects, as well as respondent and district-level control variables. Our identification strategy, therefore, follows the interacted IV approach emphasized by Borusyak and Hull (2023), and accordingly, proceeds to identifying assumption tests preconized by the latter.

Africa has been digitally isolated from the rest of the world until 2010 when intercontinental submarine infrastructure has widely expanded (Cariolle, 2021). Since then, there has been an increase in the share of Internet users in the African continent over the last decade with the highest share in North Africa where more than half of the population uses the Internet (World Bank, 2018). This rise in connectivity has coincided with the emergence of protest movements in several autocratic regimes in the region, with demands for greater democracy. The Arab Spring in North Africa was the most prominent example, but other significant movements emerged in countries like Burkina Faso, South Africa, Senegal, Nigeria, and Zimbabwe. Thus, Africa offers an interesting case study to examine the impact of Internet use on citizens' preference for and perception of democracy.

We find that Internet use induces a bias toward the belief that "in some circumstances, a non-democratic government can be preferable", away from the belief that "democracy is preferable to any other kind of government". We notice that the preference for democracy may not be consistent, as a significant proportion of respondents express a preference for democracy while approving institutional settings that may not be fully compatible with democratic principles. We also find that Internet users tend to have a more negative perception of the level of democracy in their countries. This may be attributed to a decreased trust in political institutions, notably the parliament and the ruling party, as well as an increased perception of corruption among parliament members. A possible side effect is that Internet users increase their political participation through street demonstrations and voting. Interestingly, this Internet-induced worsening perception of the level of democracy is echoing a widening of the gap between Internet users' perceptions and experts' ratings of the level of democracy. Furthermore, we document a discrepancy between Internet users' perceptions and experts' ratings of the level of corruption among legislators, thereby indicating that the Internet may serve as a source of misinformation and may alter their opinion about democracy's well-functioning.

We make a threefold contribution to the existing literature. First, we provide a micro-level analysis of the relationship between Internet use and citizens' preference for democracy, as well as their perceptions of the level of democracy, in a set of African countries, where Internet penetration has steadily increased over the past decade. Second, we unravel the mechanisms underlying our main findings. We investigate several channels including the variation in trust in government and its institutions induced by Internet exposure, the perceived corruption of different political actors, the political engagement through protests and voting, the extent of Internet-induced misperception, represented by the discrepancy between Internet users' perceptions and experts' assessment, and the consis-

tency in the way individuals understand the concept of democracy. Finally, we propose an original identification strategy that allows us to establish a causality between the use of the Internet to gather information and a negative bias in preferences and opinions toward democracy.

The remainder of this paper is organized as follows. Section 1.2 provides a brief review of the related literature. Section 1.3 presents the data sources and descriptive statistics. In Section 1.4, the empirical strategy is explained, and the main results are presented in Section 1.5. Section 1.6 explores the potential mechanisms, followed by robustness checks in Section 1.7. Finally, Section 1.8 concludes.

1.2 Literature review

The literature has largely focused on the effect of the Internet on political outcomes, and two broad categories of articles can be distinguished: those studying mature democracies and those studying recent democracies or autocracies.[2] In the former case, the Internet seems to shift voters from traditional processes of political participation and foster populism. In the latter case, the focus is on the Internet's ability to mobilize people against authoritarian and corrupt regimes, by providing a means of independent information in environments where information is controlled, by facilitating protests and coordination, and by raising awareness of the corruption of governments in power.

The distinction is based on the censorship and government control of traditional media. While reporting that Facebook is negatively correlated with corruption in a cross-section of more than 150 countries, Jones et al. (2017) argue that social media constitutes an important source of information dissemination when traditional sources are subject to censorship. Enikolopov et al. (2018) found the same negative relationship between social media and corruption in Russia. They provided evidence that blog posts exposing corruption in Russian state-controlled companies reduce their market returns, increase management turnover, and lower shareholder conflicts. Tertytchnaya and Lankina (2020) found that the effect of anti-regime protests on attitudes is hampered by state control of national media. In a similar vein, Guriev et al. (2021) exploited increased Internet penetration through 3G expansion to assess the impact of the Internet on government approval. They found that 3G network access reduces confidence in the government only when the Internet is not censored and that the effect is stronger in countries where tradi-

[2] The contrast in the effect of the Internet in mature and recent democracies was recently established in the systematic study by Lorenz-Spreen et al. (2023).

tional media is under government control and when there is at least some corruption.

Research on mature democracies documented the impact of the Internet on voting behavior and reported overall that the Internet has a significant impact on voting turnout. For instance, Falck et al. (2014) exploited Germany's broadband Internet expansion in 2004-2008 and found that the Internet reduced turnout. Similarly, Gavazza et al. (2019) relied on extreme weather shocks that caused Internet access disruption as their identification strategy and reached the same conclusion in the UK during 2006-2010. They argue that this is due to the substitution of political news with entertainment content online. However, Campante et al. (2018) found that the negative impact of access to broadband Internet on voters' turnout in Italy was only present until 2008, once it reversed with the introduction of social media. Other authors linked the rise of populism in Italy, Germany, and Europe in general to the expansion of the Internet (Schaub and Morisi, 2020; Guriev et al., 2021). A recent study by Tabellini et al. (2023) also showed that the expansion of mobile Internet coverage resulted in a higher vote share for right-wing communitarian parties across twenty European countries from 2007 to 2017, regardless of whether these parties were populist or not.

However, in immature democracies and autocratic regimes, the emphasis was placed on the Internet as a powerful tool to spur political change, which is not necessarily contradictory to the results found in mature democracies. The Internet increases access to political information that is not available through other means due to censorship. Miner (2015), for instance, found that broadband Internet led to a substantial decline in political support for the ruling coalition in Malaysia during the 2004 and 2008 elections. Similarly, Donati (2023) found that the spread of 3G mobile Internet technology led to a decline in the vote share of the ruling party in local elections in South Africa between 2006 and 2016. This negative impact was more pronounced in corrupt localities. Moreover, a recent study by Hatte et al. (2023) found that greater Facebook access was associated with increased election of female candidates in Sub-Saharan Africa due to exposure to content generated in more progressive countries and a greater visibility of female candidates in online campaigns.

Along with the impact of the Internet on voting behavior, another strand of literature reported the crucial role played by the Internet and social media platforms in mobilizing citizens by spreading critical information about the government and facilitating coordination. Fergusson and Molina (2019) showed that Facebook is associated with a higher number of protests worldwide. They found that new releases of Facebook with new languages increased protests in countries where these languages are spoken. This effect was stronger

in countries with wider Internet access and more economic and political grievances, such as China and Russia. Similarly, Qin et al. (2021) found that China's social media Sina Weibo expansion was positively associated with increased protests. Enikolopov et al. (2020) found that penetration of VK, Russia's dominant social media platform, increased the probability of having a protest during 2011. In addition, Manacorda and Tesei (2020) found that the adoption of 2G mobile network technology increased political protests in Africa between 1998 and 2012. More recently, exploiting the gradual arrival of submarine cables on the Sub-Saharan African coast, Guiffard (2022) documented a positive impact of high-speed Internet on participation in protests for 10 countries, highlighting the role of enhanced coordination.

However, the existing empirical literature on the impact of Internet use on attitudes toward democracy has mostly focused on macro-level relationships, giving little insight into their underlying mechanisms (Evans, 2019; Jha and Kodila-Tedika, 2020). Studies evaluating individuals' Internet use and citizens' attitudes toward democracy remain scarce and are largely conceptual works documenting a simple correlation between the Internet and democratic attitudes. For example, using Eurobarometer data, Ceron and Memoli (2016) found that while the Internet per se has no effect on satisfaction with democracy among European citizens, social media news consumption is negatively associated with citizens' satisfaction. Chang (2018) also found that media use in general, and the Internet in particular, have a negative effect on the satisfaction with democracy in 34 countries. However, Bailard (2012) argued that Internet use is correlated with increased satisfaction in advanced democracies and dissatisfaction in weak democracies. Acemoglu et al. (2021) emphasized the crucial need to explore the role of media in shaping support for democracy, especially in light of the spread of misinformation from various media outlets and social media platforms, and how this may affect the relationship between successful democratic performance and public support for democracy. This paper aims to contribute to this research agenda by proposing an in-depth analysis of the causality running from Internet use as an information provider to attitudes toward democracy, anchored in individual-level responses, drawing on several waves of the Afrobarometer, and using an original identification strategy.

1.3 Data

In this section, we introduce the data and some descriptive statistics. The data source is presented in Section 1.3.1. We describe our main outcome variables in Section 1.3.2. Descriptive statistics on the Internet and traditional media use and on submarine cables

are presented in Sections 1.3.3 and 1.3.4, respectively.

1.3.1 Data source

We rely on three recent rounds of the Afrobarometer, a public attitude survey on democracy, governance, corruption, and related issues in African countries.[3] A randomly selected sample of 1,200 or 2,400 individuals is collected through face-to-face interviews in each country. We selected this dataset because it includes a wide variety of questions on citizens' opinions and attitudes, as well as questions on media consumption. Our final sample comprises 99,939 respondents from 1,845 districts across 35 African countries[4,5] surveyed between 2011 and 2018. We report the number of observations by country and by year in Table A.1 and Table A.2, respectively, in Appendix A.1.[6]

1.3.2 Main dependent variables

We analyze the impact of the Internet as a means of accessing news on i) citizens' preference for democracy, ii) their perception of whether they are getting democracy, and iii) their satisfaction with how democracy is functioning.

To measure citizens' preference for democracy, we use respondents' answers to the following question: "Which of these three statements is closest to your own opinion? A) Democracy is preferable to any other form of government; B) In certain situations, a non-democratic government can be preferable; C) To people like me, it doesn't matter what form of government we have". First, we create a binary variable equal to 1 if citizens respond "A) Democracy is preferable to any other form of government" and 0 otherwise. Second, we consider an additional dependent variable that ensures choosing option "A" corresponds to a true preference for democracy, meaning the rejection of real-world alternative regimes with whom African respondents are familiar and to which they can have experience-based responses, namely one-party rule, military government, and presidential dictatorships. Thus, our second outcome of interest reflects a strict preference for democ-

[3] Afrobarometer Data, rounds 5, 6, and 7, available at http://www.afrobarometer.org.

[4] The countries included in our study are: Algeria, Benin, Botswana, Burkina Faso, Burundi, Cameroon, Cape Verde, Côte d'Ivoire, Egypt, Gabon, Ghana, Guinea, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritius, Morocco, Mozambique, Namibia, Niger, Nigeria, Senegal, Sierra Leone, South Africa, Sudan, Tanzania, Togo, Tunisia, Uganda, Zambia, Zimbabwe, and Gambia.

[5] We excluded São Tomé and Príncipe from our sample as it is a two-islands country with only one connected to a SMC, and Swaziland as it is missing one of the variables needed to construct one of our dependent variables.

[6] Additional data sources for variables used in the analysis are reported in Table A.3 in Appendix A.1.

racy taking the value of 1 if citizens prefer democracy to any other form of government and reject all three previously mentioned alternative regimes.

To measure citizens' perception of the actual extent of democracy, we use their response to the following question: "In your opinion, how much of a democracy is your country today? Is it a full democracy, a democracy with minor problems, a democracy with major problems, or not a democracy?". Satisfaction with democracy is measured using the following question: "Overall, how satisfied are you with the way democracy works in your country today? Are you very satisfied, fairly satisfied, not very satisfied, or not at all satisfied?". We recode the categorical responses to these questions into binary values. Specifically, the extent of democracy variable takes the value of 1 if respondents declare their country to be a "full democracy" or a "democracy with minor problems" and 0 otherwise. The satisfaction with democracy variable takes the value of 1 if respondents declare they are "very" or "fairly" satisfied with how democracy works and 0 otherwise.

The correlation between all our dependent variables is reported in Table A.4 in Appendix A.1. Three interesting facts can be pointed out. First, the correlation between satisfaction with how democracy works and the perceived extent of democracy suggests that individuals who are satisfied with democracy also tend to see an extensive democracy. Second, the correlation coefficient between the preference for democracy and the extent and satisfaction with democracy is relatively low, lying around 10%. Lastly, the simple preference and the strict preference for democracy remain moderately correlated, at about 57%, reflecting that a considerable number of individuals have a distorted understanding of democracy as being compatible with one-party rule, military government, or presidential dictatorships, in certain circumstances.

1.3.3 Internet and traditional media use

Africa's media landscape has been changing rapidly during the last decade, with a growing reliance on the Internet as a source of both verified and unverified news. This trend is exemplified by the role of social networks in spurring events like the Arab Spring or the spread of social discontent in Sub-Saharan African countries (Fergusson and Molina, 2019; Bosch et al., 2020). The survey captures this specific information-gathering channel, by asking individuals "How often do you get news from the following sources: radio, television, newspapers, and Internet?". The responses range from "every day" to "never", and we rely on the Internet use ordered categorical variable as our regressor of interest in the analysis. However, for the clarity of the below descriptive analysis, we create a dichotomous variable

that identifies regular Internet users as individuals who use the Internet "every day" or "a few times a week" to get news. Throughout the text, we use the terms "Internet use" and "Internet use to get news" in the same equivalent way.

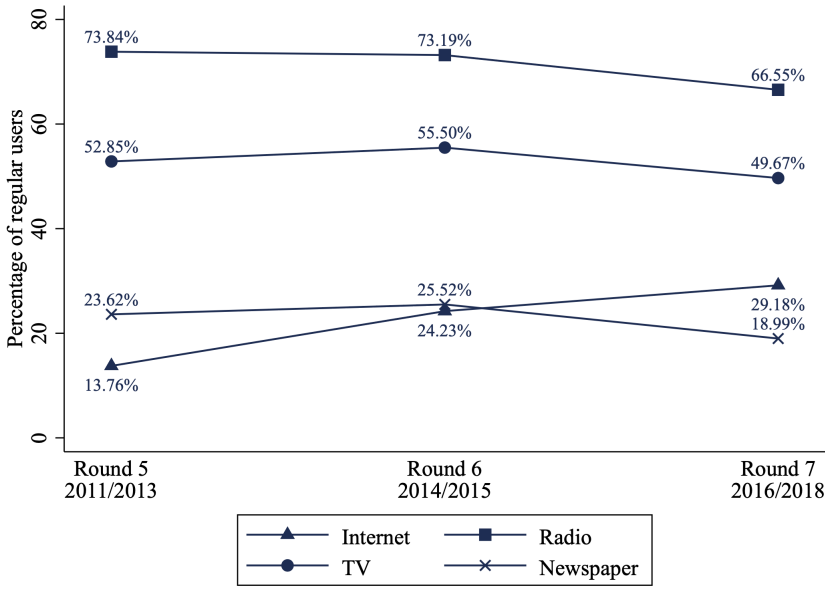
The potential sources of news on the Internet are numerous and include online news websites, social media platforms, blogs, and search engines, among others. Unfortunately, the data at hand does not allow for differentiation between these sources. However, starting from round 6, the survey added social media as an additional source of news in the previous question. Since it is not available for the entire period, we choose to rely on individuals' responses to using the Internet to get news. We note a positive and 1% significant correlation coefficient of 0.88 between Internet and social media news consumption. This suggests that individuals who report using the Internet to get news are highly likely to be referring to using social media platforms.

Based on our baseline sample, the share of individuals who reported using the Internet at least "a few times a week" to get news has nearly doubled. As shown in Fig. 1.1, in the surveyed countries, the share of regular Internet users increased from 13.76% to 29.18% between round 5 and round 7. Conversely, traditional news sources, such as radio, television, and newspapers have been losing ground in the continent. However, radio remains the dominant source of news for most Africans, likely due to its affordability and accessibility. In contrast, newspaper readership in Africa is the lowest, reflecting a limited reading culture and a higher illiteracy rate in the continent.

Furthermore, Fig. 1.2 indicates that there has been a decline in the percentage of people who never use the Internet to get news over time. In round 5, 80% of individuals stated that they "never" use the Internet for news, while in round 7, this percentage decreased to 64%. It is worth noting that this decrease was accompanied by a corresponding increase in the percentage of individuals who report using the Internet "every day" to get news, from 7% in round 5 to 20% in round 7, while the percentage of individuals who use the Internet less frequently remained relatively stable across rounds. This suggests that although the Internet may have not entirely replaced traditional news sources, it is becoming an important complement.

Although the Internet is becoming more widespread in Africa, its diffusion across the continent as a means of accessing news is spatially uneven. Fig. 1.3 displays the distribution of regular Internet users at the district-year level in our baseline sample. We can indeed see the heterogeneous distribution across districts with an over-representation of districts where the share of regular Internet users is between 0% and 20%.

Fig. 1.1. News media consumption across three rounds of the Afrobarometer survey.



Note: This figure shows the percentage of survey respondents who reported using the Internet, radio, TV, or newspapers to get news "every day" or "a few times a week" in rounds 5, 6, and 7 of the Afrobarometer survey.
Source: Authors' calculation on Afrobarometer data based on the baseline sample of 99,938 individuals in 35 countries between 2011 and 2018.

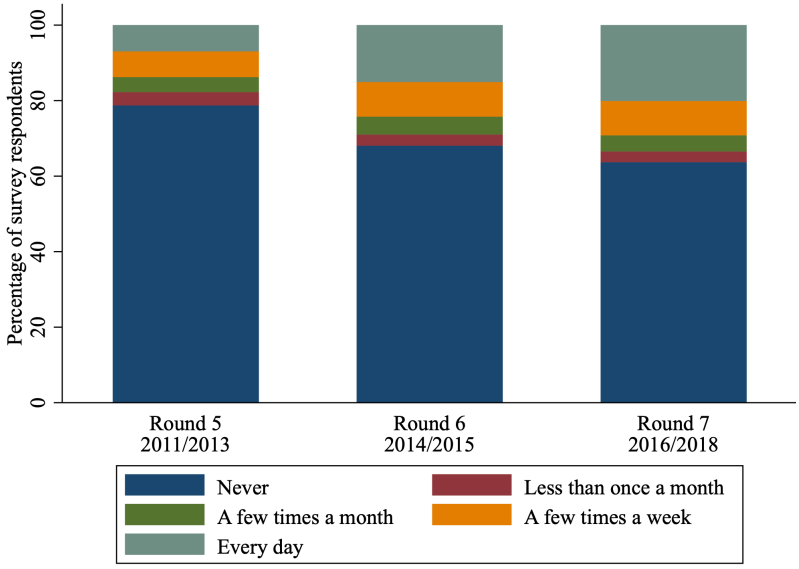
Last, to get a preliminary insight into the Internet use-democracy nexus, we have created binned scatterplots in Fig. 1.4 that depict the simple correlation between district-level incidences of regular Internet use and democracy variables. The plots show a strong and negative relationship between regular Internet use and the share of individuals who (strictly) prefer democracy, perceive their country as a "full democracy" or a "democracy with minor problems", and are "very" or "fairly" satisfied with how democracy works in their country. Overall, this preliminary graphical evidence shows that preference for democracy and satisfaction with the functioning of democracy are lower in districts where Internet use is more widespread.

These trends in media consumption may have political implications, as digital and traditional news sources contrast sharply in the way they shape the political landscape as argued in Zhuravskaya et al. (2020).

1.3.4 Submarine cables

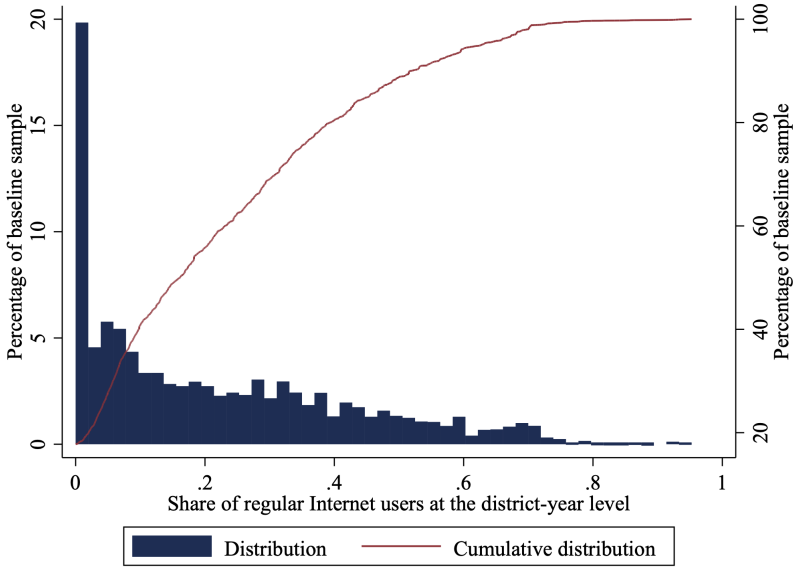
As of late 2021, there are approximately 436 operational fiber-optic SMCs laid down over 1.3 million kilometers and connecting countries around the world (Telegeography, 2021).

Fig. 1.2. Frequency of Internet use to get news across three rounds of the Afrobarometer survey.



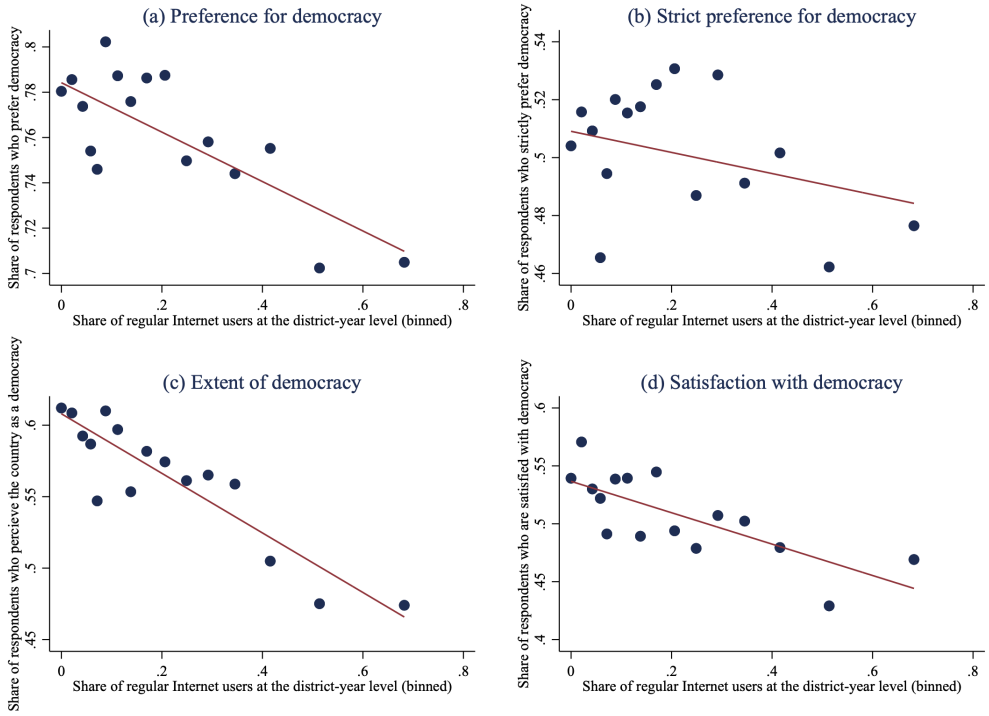
Note: This figure shows the percentage of survey respondents who reported using the Internet to get news, categorized by frequency of use as "never", "less than once a month", "a few times a month", "a few times a week", or "every day" in rounds 5, 6, and 7 of the Afrobarometer survey.
Source: Authors' calculation on Afrobarometer data based on the baseline sample of 99,938 individuals in 35 countries between 2011 and 2018.

Fig. 1.3. Distribution of regular Internet users at the district-year level.



Note: This figure shows the (cumulative) distribution of the share of survey respondents who reported using the Internet to get news "every day" or "a few times a week" at the district-year level.
Source: Authors' calculation on Afrobarometer data based on the baseline sample of 99,938 individuals in 35 countries between 2011 and 2018.

Fig. 1.4. Correlation between regular Internet use and perception of democracy at the district-year level.



Notes: These figures show scatterplots that group the x-axis variable into equal-sized bins, calculate the mean value of the x-axis and y-axis variables within each bin, and plot these data points. The x-axis represents the share of survey respondents at the district-year level who reported using the Internet to get news "every day" or "a few times a week". The y-axis represents the share of respondents at the district-year level who "prefer democracy to any other type of government" in (a), the share who "prefer democracy to any other type of government" and "reject all types of authoritarian rules" in (b), the share who perceive their country as "a full democracy" or "a democracy with minor problems" in (c), and the share who are "very" or "fairly" satisfied with how democracy works in their country in (d).
Source: Authors' calculation on Afrobarometer data based on the baseline sample of 99,938 individuals in 35 countries between 2011 and 2018.

Their deployment is the first step toward accessing the global Internet. Carrying out more than 95% of world Internet traffic, their deployment improves telecommunications network size, capacity, and redundancy (Weller and Woodcock, 2013; Schumann and Kende, 2013; D'Andrea and Limodio, 2023). The greater the number of SMCs, the faster the Internet speed and the larger the Internet bandwidth. In the absence of SMCs, a country has two options to communicate with the rest of the world: buying Internet bandwidth from a neighboring country already connected to SMCs or relying on satellite communication systems. These two solutions are associated with higher costs and lower Internet speed. SMCs are, therefore, the backbone infrastructure of the worldwide telecommunications network.

Today, nearly all coastal African countries are connected to at least one SMC, meaning that anyone who is connected to the Internet, regardless of their device, uses this infrastructure to access the Internet. This reliance on SMCs for Internet access is particularly true in Sub-Saharan Africa (SSA), where the local anchoring of Internet traffic is poor, including traffic between geographically close countries, which is often routed toward data centers located outside the continent before being brought back to the destination country (Fanou et al., 2017). This is also true for landlocked countries, which have been lately connected to the World Wide Web through their connected coastal neighbors' terrestrial fiber-optic network by the beginning of the 2010s. While these countries do not have to pay anymore for the transit of telecommunications through their neighbors' SMC network (Houngbonon et al., 2022), their direct access to the SMC network relies on a few cables, which limits their available bandwidth and rerouting capabilities.[7]

Fig. 1.5 illustrates the distribution of the number of SMCs in our baseline sample, revealing that a significant proportion of the observations (nearly 38%) have access to only one SMC. The percentage of observations drops as the number of SMCs increases, with approximately 14% of the observations having two, three, or four cables. Merely 9% of observations enjoy access to five SMCs, and less than 2% have access to six cables or more. Notably, less than 10% of our sample observations are not connected to any SMC. This distribution pattern points to the unevenness in SMC access across the sample and reflects disparities in Internet connectivity quality across different countries.[8]

Our empirical analysis leverages the staggered deployment of SMCs along African coasts as a plausibly random connectivity shock, with an expected dramatic impact on Internet access for populations covered by the mobile Internet network.

1.4 Empirical strategy

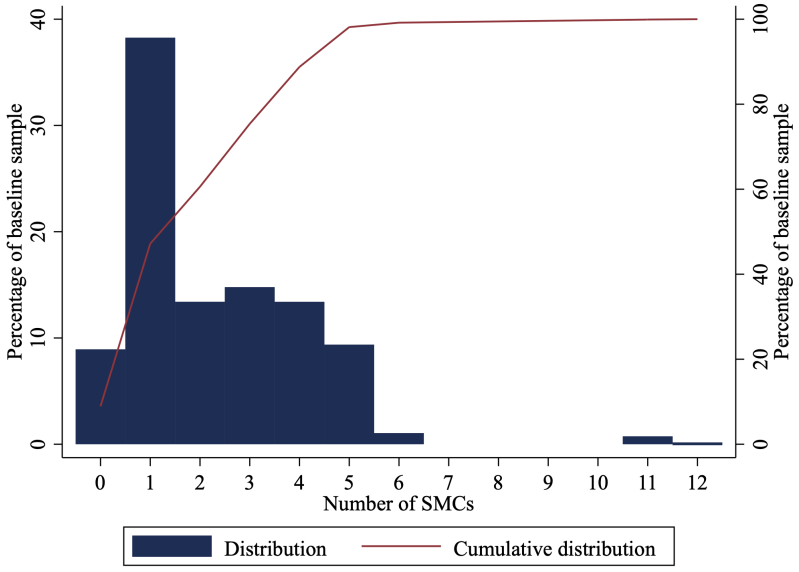
Our empirical strategy builds on the literature using Bartik-type instruments, which are constructed by interacting aggregate shocks with exposure weights. We adopt a design à-la Borusyak and Hull (2023), where the validity of the instrument stems from exogenous variation in the shocks while allowing for endogenous variation in the exposure factor.[9]

[7] According to measurements made by Chavula et al. (2015), on average 75% of the intra-African traffic destined for Africa's National Research and Education Networks was carried through intercontinental links in Europe.

[8] We report the variation in the number of SMCs by country in Fig. A.1 in Appendix A.1.

[9] Our approach does not strictly follow the standard shift-share instrumental variable framework (Borusyak et al., 2022; Goldsmith-Pinkham et al., 2020) since the sum of factors used to weigh the

Fig. 1.5. Distribution of the number of SMCs.



Note: This figure shows the (cumulative) distribution of the number of submarine cables (SMCs) in our baseline sample.
Source: Authors' calculation on Afrobarometer data based on the baseline sample of 99,938 individuals in 35 countries between 2011 and 2018.

1.4.1 Main specification

To estimate the effect of Internet use on citizens' perception of democracy, we estimate the following model:

$$Y_i = \alpha_0 + \beta_0 Internet_i + \boldsymbol{\gamma_0}\mathbf{X_i} + \boldsymbol{\delta_0}\mathbf{W_{d,t}} + \boldsymbol{\sigma_0}\mathbf{Z_{c,t}} + \rho_d + \rho_t + \varepsilon_i \tag{1.1}$$

Where Y_i is four different dependent binary variables representing individual i's (simple or strict) preference for democracy, perception of the extent of democracy, and satisfaction with democracy. Our variable of interest, $Internet_i$, is an ordered categorical variable measuring the frequency of Internet usage to get news, ranging from 0 (never use the Internet to get news) to 4 (use it every day). $\mathbf{X_i}$ is a set of individual characteristics including age, age squared, gender, urban dummy, education, employment status, perception of own living conditions, perception of the country's economic condition, interest in politics, regular TV use dummy, regular radio use dummy, and regular newspaper use dummy.

connectivity shock induced by SMC arrival, i.e., the district coverage (in %) by the 3G network, are not equal to unity.

We also add time-varying district-level ($\mathbf{W_{d,t}}$) and country-level ($\mathbf{Z_{c,t}}$) controls. First, we use nighttime light as a proxy to control for district's economic development. Second, we control for local spillovers possibly induced by the diffusion of the Internet around respondents, by including the district-level share of respondents who declare using the Internet on a weekly or daily basis. Third, we include the 2G and 3G network coverage, expressed as a share of the district's surface area, using data from Collins Bartholomew's Mobile Coverage Explorer. We control for both networks since the 2G network was instrumental to Internet access in the first half of the 2010s, while Internet access mostly relied on the 3G network from the second half.[10] Finally, we control for the logarithm of the distance from the district's centroid to the closest backbone infrastructure node, i.e., SMC landing stations and Internet exchange points (IXPs).[11,12] Finally, our country-level controls include the logarithm of GDP per capita, the overall level of democracy as rated by the Polity2 index, the unemployment rate, and the raw number of SMCs. Table 1.1 below reports summary statistics of the variables used in our regression analysis.

We add district (ρ_d) and year (ρ_t) fixed effects to account for location time-invariant unobserved heterogeneity and yearly common shocks, respectively. ε_i is the error term. In all our specifications, error terms are corrected for heteroscedasticity and clustered at the district-year level. Multi-country sampling weights are applied in all estimations, and observations located in districts with less than 10 non-missing observations are dropped from the baseline sample.

Our main concern is the possible endogeneity bias that stems from omitted variable bias and the bi-directional relationship between Internet news consumption and democracy variables. While Internet use may influence citizens' preference for, perception of, and

[10] 26% of total mobile connections in Sub-Saharan Africa were made through 2G network as of 2021 (GSMA, 2022).

[11] IXPs are national or regional Internet hubs that allow Internet Service Providers (ISPs) to exchange their traffic locally. They constitute a core element of the Internet infrastructure that increases Internet performance and reduces cost by keeping local traffic locally.

[12] Data on SMC landing stations and IXPs status, year of activation, and GPS coordinates are obtained from Telegeography website and completed by the Packet Clearing House and Peering DB databases. If a country does not host any SMC or IXP, the distance is calculated considering the closest infrastructure in neighboring countries. Studies have shown that there is a spatial hierarchy in Internet connectivity favoring Internet access in economic and demographic centers when the telecommunications network capacity is altered (Grubesic et al., 2003; Gorman et al., 2004; Grubesic and Murray, 2006; Malecki, 2009). Populations remote from connectivity infrastructures such as SMC landing stations or IXPs, are indeed more exposed to telecommunication network failures, while populations close to them enjoy better and more stable connectivity. Remote populations are also the last to recover after Internet shutdowns. We assume that individuals closer to telecommunication infrastructures are less exposed to Internet slowdowns or shutdowns than remote ones.

Table 1.1
Summary statistics of main variables.

	Mean	Std. Dev.	Min.	Max.
Democracy variables				
Preference for democracy	0.77	0.42	0	1
Strict preference for democracy	0.53	0.50	0	1
Extent of democracy	0.56	0.50	0	1
Satisfaction with democracy	0.49	0.50	0	1
Media use				
Internet use	0.93	1.53	0	4
TV use	0.53	0.50	0	1
Radio use	0.71	0.45	0	1
Newspaper use	0.23	0.42	0	1
Individual controls				
Age	36.86	14.32	18	106
Male	0.52	0.50	0	1
Urban	0.45	0.50	0	1
No formal education	0.20	0.40	0	1
Primary	0.26	0.44	0	1
Secondary	0.37	0.48	0	1
Post-secondary	0.17	0.37	0	1
Unemployed (not looking for job)	0.39	0.49	0	1
Unemployed (looking for job)	0.26	0.44	0	1
Part-time employee	0.11	0.32	0	1
Full-time employee	0.24	0.42	0	1
(Very) good living conditions	0.34	0.47	0	1
(Very) good country economic condition	0.30	0.46	0	1
Never discuss politics	0.30	0.46	0	1
Ocasionally discuss politics	0.49	0.50	0	1
Frequently discuss politics	0.21	0.41	0	1
District controls				
Nighttime light	11.40	15.96	0	63
Internet incidence	0.98	0.84	0	4
2G coverage	0.68	0.37	0	1
3G coverage	0.17	0.33	0	1
Log distance (in km)	4.71	1.71	-1	7
Country controls				
Log GDP per capita	8.17	0.82	7	10
Polity2 index	4.59	4.11	-4	10
Unemployment rate	8.09	6.95	1	27
Number of SMCs	2.15	1.75	0	12

Note: This table reports the mean, standard deviation, minimum, and maximum of the main variables used in the analysis. Source: Authors' elaboration on Afrobarometer data based on the baseline sample of 99,938 individuals in 35 countries between 2011 and 2018.

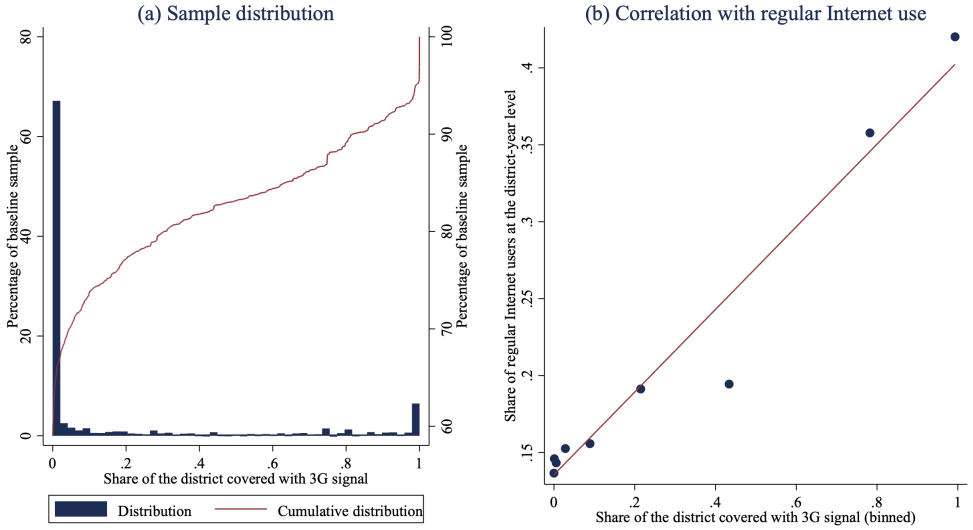
satisfaction with democracy, the causality can work the other way around. For instance, citizens who are dissatisfied with their country's level of democracy may opt to use the Internet as a more accessible and open source of information to express their opinions. To address this endogeneity issue, our empirical strategy is related to the instrumental variable setting emphasized by Borusyak and Hull (2023), and to the econometric literature on shift-share instruments (Borusyak et al., 2022; Goldsmith-Pinkham et al., 2020), in which instruments are constructed as aggregate shocks weighted by lower-level exposure factors.

1.4.2 Identification strategy

SMCs are the backbone of the worldwide telecommunications network. Their number increases Internet speed, capacity, stability, and affordability. The staggered arrival of SMCs in African countries is considered a quasi-experiment providing us with an exogenous source of time and cross-country variation in Internet connectivity, mostly driven by the continent's geography rather than country-specific policy-related factors (Cariolle, 2021; D'Andrea and Limodio, 2023; Eichengreen et al., 2023; Imbruno et al., 2022). In fact, the SAT-3/WASC-WACS-ACE cables on the west coast, the SeaMeWe cables along the Mediterranean and Red Seas and the Aden Gulf, and the EASSy-SEACOM cables on the East and Southern coasts are the principal Internet routes in the region which have all been deployed to serve as many countries as possible along their path and to connect each side of the continent to Europe, Latin America, and Asia (D'Andrea and Limodio, 2023).

Once a SMC has landed on a given country's coast, access to the Internet then relies on the last-mile infrastructure coverage. In Africa, where fixed last-mile infrastructure is lacking, the Internet is mainly accessed through mobile engines. As a result, SMC arrival will be instrumental to populations inasmuch as they are covered by the 3G mobile network. Our identification strategy hence combines an aggregate and exogenous source of Internet connectivity, the number of SMCs, with an internal one, the district's 3G network coverage (in %). Our IV framework, therefore, is close to the interacted IV design emphasized by Borusyak and Hull (2023), where exogeneity proceeds from the random assignment of aggregate shocks, weighted by possibly non-random exposure factors. Fig. 1.6(a) shows the distribution of the 3G signal coverage across districts in our baseline estimation sample. Fig. 1.6(b) plots a binned scatterplot showing a positive correlation between the sample districts' coverage with 3G signal and the share of regular Internet users at the district-year level.

Fig. 1.6. 3G signal: sample distribution and correlation with regular Internet use.



Note: Figure (a) shows the (cumulative) distribution of the share of the district covered with 3G signal, while figure (b) shows a scatterplot that groups the share of the district covered with 3G signal into equal-sized bins, calculates its mean value and plots its correlation with the average share of respondents at the district-year level who reported using the Internet to get news "every day" or "a few times a week" within each bin.
Source: Authors' calculation on Afrobarometer data based on the baseline sample of 99,938 individuals in 35 countries between 2011 and 2018.

To ensure that identification results from a change in aggregate connectivity rather than (endogenous) weighting factors, we interact the SMC number variable with the value of the district's 3G coverage fixed at the time of the first Afrobarometer survey wave used in our analysis.[13] Therefore, the main instrument is the interaction between an aggregate connectivity shock equal to the number of SMCs laid in the country c at time t and the fixed share of the individual's district d of residence covered by the 3G network:

$$IV_{d,c,t} = SMCnumber_{c,t} \times fixed3Gshare_{d,c} \tag{1.2}$$

This leads us to add the following first stage equation to our estimation framework:

$$Internet_i = \alpha_1 + \beta_1 IV_{d,c,t} + \boldsymbol{\gamma_1 X_i} + \boldsymbol{\delta_1 W_{d,t}} + \boldsymbol{\sigma_1 Z_{c,t}} + \rho_d + \rho_t + \epsilon_i \tag{1.3}$$

Where $IV_{d,c,t}$ is our main instrument. The equation is estimated using the LIML estimator.

[13] District's non-random exposure to connectivity shocks, which is a second threat to identification, is tested in the robustness section.

1.5 Main results

In this section, we present our baseline results obtained from estimating equations (1.1) and (1.3). The OLS estimates are presented in Section 1.5.1, while the IV estimates are presented in Section 1.5.2.

1.5.1 OLS results

The OLS results are presented in Table 1.2.[14] First, we find no significant correlation between Internet use to get news and the preference for democratic governments. However, the coefficient on the strict preference is positive and significant. A one-unit increase in the Internet use ordinal variable (i.e., moving up one level of Internet usage frequency) corresponds to a 0.7 percentage points increase in the probability of strictly preferring democracy. Second, we find a negative and significant correlation between Internet use and the perception of the extent of, and satisfaction with, democracy. A one-unit increase in the frequency of using the Internet to get news decreases the probability of perceiving the country as a democracy and being "very" or "fairly" satisfied with the way democracy works in one's own country by 0.8 and 0.9 percentage points, respectively.

1.5.2 IV results

Table 1.3 reports the IV estimates. The first stage estimates indicate that, as expected, the number of SMCs weighted by the fixed 3G network share has a positive impact on the frequency of Internet use. This implies that in countries with at least one SMC, individuals with 3G network coverage are more likely to use the Internet regularly than those without 3G coverage. The F-statistic exceeds the recommended threshold of 10, indicating that our instrument is strong and effective in addressing the issue of endogeneity.[15,16]

The second stage estimates show a significant negative effect of Internet use on the preference for democratic governance and perception of the extent of democracy. A one-unit increase in Internet use frequency as a source of news lowers the probability of (strictly)

[14] Results displaying control variables are reported in Table A.5 in Appendix A.2.

[15] Additionally, we present the reduced form estimations of the direct effect of our IV on democracy variables in Table A.6 in Appendix A.2, and our instrument exhibits similar behavior as in the first stage estimation.

[16] We also adopt a two instruments setting (see section 1.7), using the number of SMCs weighted by the 2G network coverage as an additional instrument, and report the effective F-statistic which is known to be robust to weak instrument bias in multiple instruments setting (Andrews et al., 2019) in Table A.13 in Appendix A.2.

Table 1.2
OLS estimates of the effect of Internet use on perception of democracy.

	(1) Preference	(2) Strict Preference	(3) Extent	(4) Satisfaction
Internet use	−0.002	0.007***	−0.008***	−0.009***
	(0.002)	(0.002)	(0.002)	(0.002)
Individual controls	✓	✓	✓	✓
District controls	✓	✓	✓	✓
Country controls	✓	✓	✓	✓
District FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Observations	99,938	99,938	99,938	99,938
Adjusted R^2	0.080	0.129	0.173	0.183
Mean dependent variable	0.773	0.525	0.566	0.503
Mean Internet use	0.895	0.895	0.895	0.895

Notes: This table reports the OLS results of the effect of Internet use to get news on individuals' perception of democracy. The dependent variable in column (1) is a dummy variable equal to 1 if the individual says "democracy is preferable to any other kind of government" and 0 otherwise. In column (2), it is a dummy variable equal to 1 if the individual says "democracy is preferable to any other kind of government" and "rejects all three authoritarian alternatives" and 0 otherwise. In column (3), it is a dummy variable equal to 1 if the individual perceives his or her country as "a full democracy" or "a democracy with minor problems" and 0 otherwise. In column (4), it is a dummy variable equal to 1 if the individual is "very" or "fairly" satisfied with how democracy works in his or her country and 0 otherwise. Internet use is an ordered categorical variable equal to 0 if the individual never uses Internet, 1 if he or she uses it less than once a month, 2 if he or she uses it a few times a month, 3 if he or she uses it a few times a week, and 4 if he or she uses it every day. Individual controls include age, age squared, gender, urban dummy, education, employment status, perception of own living conditions, perception of the country's economic condition, interest in politics, regular TV use dummy, regular radio use dummy, and regular newspaper use dummy. District controls include nighttime light, local Internet incidence, 2G network coverage, 3G network coverage, and the log of the distance from district's centroid to the closest Internet infrastructure. Country controls include log GDP per capita, Polity2 index, unemployment rate, and the number of SMCs. Standard errors clustered at the district-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' elaboration on Afrobarometer data based on the baseline sample of 99,938 individuals in 35 countries between 2011 and 2018.

preferring democracy by 31.7 (26) percentage points and decreases the probability of perceiving the country as a "full democracy" or a "democracy with minor problems" by 23.3 percentage points. However, we find that a one-unit increase in Internet use frequency is associated with higher satisfaction with democracy functioning.[17,18] While these re-

[17] We report the Anderson and Rubin (1949) weak instrument robust test P-value and confidence set, as recommended by Keane and Neal (2023) in Table A.7 in Appendix A.2. Results indicate that we can reject the null hypothesis $H_0 : \beta_0 = 0$.

[18] We provide alternative IV results using a binary version of the Internet use variable, which is equal to 1 if the individual uses the Internet "every day" or "a few times a week" and 0 otherwise, in Table A.8 in Appendix A.2.

Table 1.3
IV estimates of the effect of Internet use on perception of democracy.

	(1) Preference	(2) Strict Preference	(3) Extent	(4) Satisfaction
First stage regression: Internet use				
SMC number × fixed 3G share	0.594***	0.594***	0.594***	0.594***
	(0.083)	(0.083)	(0.083)	(0.083)
Second stage regression:				
Internet use	−0.317***	−0.260***	−0.233**	0.258**
	(0.107)	(0.081)	(0.114)	(0.104)
Individual controls	✓	✓	✓	✓
District controls	✓	✓	✓	✓
Country controls	✓	✓	✓	✓
District FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Observations	99,938	99,938	99,938	99,938
Mean dependent variable	0.773	0.525	0.566	0.503
Mean Internet use	0.895	0.895	0.895	0.895
KP Wald F-stat	51.050	51.050	51.050	51.050
KP LM P-val	0.000	0.000	0.000	0.000

Notes: This table reports the first and second stages of IV results of the effect of Internet use to get news on individuals' perception of democracy. The dependent variable in column (1) is a dummy variable equal to 1 if the individual says "democracy is preferable to any other kind of government" and 0 otherwise. In column (2), it is a dummy variable equal to 1 if the individual says "democracy is preferable to any other kind of government" and "rejects all three authoritarian alternatives" and 0 otherwise. In column (3), it is a dummy variable equal to 1 if the individual perceives his or her country as "a full democracy" or "a democracy with minor problems" and 0 otherwise. In column (4), it is a dummy variable equal to 1 if the individual is "very" or "fairly" satisfied with how democracy works in his or her country and 0 otherwise. SMC number × fixed 3G share is the interaction between the number of SMCs and the share of the individual's district of residence covered with 3G network in the first Afrobarometer survey wave used in our analysis. Internet use is an ordered categorical variable equal to 0 if the individual never uses Internet, 1 if he or she uses it less than once a month, 2 if he or she uses it a few times a month, 3 if he or she uses it a few times a week, and 4 if he or she uses it every day. Individual controls include age, age squared, gender, urban dummy, education, employment status, perception of own living conditions, perception of the country's economic condition, interest in politics, regular TV use dummy, regular radio use dummy, and regular newspaper use dummy. District controls include nighttime light, local Internet incidence, 2G network coverage, 3G network coverage, and the log of the distance from district's centroid to the closest Internet infrastructure. Country controls include log GDP per capita, Polity2 index, unemployment rate, and the number of SMCs. Standard errors clustered at the district-year level are reported in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' elaboration on Afrobarometer data based on the baseline sample of 99,938 individuals in 35 countries between 2011 and 2018.

sults may seem puzzling, we argue that satisfaction with how democracy works in one's own country is a more complex, blurry, and fluctuating construct to analyze, resulting from idiosyncrasies such as people's frustration after an undesirable electoral outcome or unfavorable government policy that are difficult to observe and control for,[19] compared

[19] All individuals, including those living in well-established democracies and favorable to this political regime, can be in some way and at some point of time dissatisfied with democracy. Moreover, the Hansen J test conducted in the IV estimation with two instruments reported in Table A.12 rejects the over-identification restriction for this particular dependent variable, indicating a potential risk of omitted variable bias.

to the preference for democracy and perception of its level which are relatively more straightforward measures and precise concepts.

When comparing OLS and IV estimates, we observe that they differ in their magnitude and sign. The IV estimates exhibit a significant negative magnitude, which surpasses the OLS estimates, suggesting that the OLS estimates may be influenced by an upward bias. This discrepancy can potentially be explained by considering that our estimates reflect a local average treatment effect (LATE), as underlined by Guriev et al. (2021).[20] The increase in the number of SMCs is likely to predominantly impact Internet usage among individuals residing in regions initially characterized by digital isolation and lower mobile Internet coverage. These regions tend to be less developed and have limited access to information resources, making them more susceptible to the influence of online information. If Internet usage exerts a more substantial influence on "compliers" (individuals whose Internet use is affected by the arrival of SMCs) than on "non-compliers" (those who are well-connected and relatively unaffected by SMC arrival), we can reasonably anticipate that IV estimates will exhibit a larger effect size.

1.6 Mechanisms

In this section, we present potential channels through which Internet use may influence attitudes toward democracy. First, we explore the effect of the Internet on trust and perceived corruption in political institutions in Section 1.6.1, followed by an investigation into its impact on mass mobilization and political participation in Section 1.6.2. Finally, we examine the (mis)information channel by analyzing the discrepancy between Internet users' perceptions and experts' evaluations of a country's level of democracy and the level of corruption among legislators, as well as the likelihood of inconsistent responses regarding individuals' preference for democracy in Section 1.6.3.

1.6.1 Trust and perceived corruption in political institutions

The Internet has proven to be a powerful tool in exposing instances of government misconduct and corruption. This, in turn, can lead to decreased confidence in governance and increased political accountability (Guriev et al., 2021). However, the Internet can also allow for the dissemination of false news that criticizes governments on social media

[20] In their analysis of the impact of 3G Internet on trust in government, Guriev et al. (2021) find that the magnitude of IV estimates is about 2.5 larger than the OLS ones, and argue that this difference is due to the LATE.

platforms, which can influence the public's trust in their regimes. Consequently, the negative impact of Internet use on individuals' preference for and perception of the extent of democracy can be channeled through a decreased confidence in their governments.

To test this assumption, we conduct IV regressions on several dummy variables reflecting individuals' trust in different political institutions, including the president, parliament, electoral commission, local government, ruling party, opposition party, police, army, and courts of law. Each dummy variable is equal to 1 if the individual trusts the entity "a lot" or "somewhat" and 0 otherwise. The importance citizens assign to different political institutions may vary depending on the context of the country, including its form of government and recent political events. To account for this variation, we include a series of individual-, district-, and country-level controls, as well as district- and year-level fixed effects.

Although we find a negative coefficient on trust in all the political institutions presented in Table 1.4, estimated effects are only significant on trust in the parliament and the ruling party. A higher frequency of using the Internet as a source of information is associated with a lower probability of trust in the parliament and the ruling party, by 19.8 and 28.1 percentage points, respectively. Overall, this table is quite consistent with the evidence provided by Guriev et al. (2021) and their analysis of 3G network expansion and trust in government institutions.

In Table 1.5, we complement these results by displaying the impact of the Internet on the perceived corruption of political actors. In this case, the dependent variables are equal to 1 if the individual believes that "all" or "most" of the political actors in question are involved in corruption and 0 otherwise. The significant increase in the perceived corruption of parliament members coincides with the decrease in trust toward the parliament and the ruling party reported above.

1.6.2 Demonstrations and political participation

To further support the evidence that trust in institutions and preference for democracy is altered by the Internet, we examine the nexus between the Internet and the likelihood of attending demonstrations and voting. Trust in government and protests are closely related, as shown by Sangnier and Zylberberg (2017), who found that trust in political leaders and institutions sharply decreases after protests in Africa. Similarly, Ketchley and El-Rayyes (2021) reveal a direct link between protest and popular perceptions of democracy during Egypt's post-Mubarak transition. Moreover, several studies have documented

Table 1.4
Internet use and trust in political institutions.

	(1) President	(2) Parliament	(3) Electoral commission	(4) Local gov.	(5) Ruling party	(6) Opposition	(7) Police	(8) Army	(9) Courts
First stage regression: Internet use									
SMC number × fixed 3G share	0.613***	0.515***	0.621***	0.559***	0.556***	0.590***	0.583***	0.541***	0.587***
	(0.090)	(0.084)	(0.123)	(0.078)	(0.084)	(0.083)	(0.082)	(0.083)	(0.085)
Second stage regression:									
Internet use	−0.056	−0.198**	−0.104	−0.008	−0.281***	−0.039	−0.025	−0.145	−0.012
	(0.146)	(0.086)	(0.224)	(0.089)	(0.086)	(0.074)	(0.096)	(0.105)	(0.125)
Individual controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
District controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Country controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
District FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	98,222	97,139	94,264	95,122	96,548	96,024	99,033	97,749	97,597
Mean dependent variable	0.578	0.508	0.516	0.481	0.482	0.382	0.507	0.675	0.580
Mean Internet use	0.896	0.902	0.893	0.909	0.903	0.907	0.897	0.896	0.903
KP Wald F-stat	46.206	37.323	25.530	51.611	43.668	50.682	50.820	42.741	47.243
KP LM P-val	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000

Notes: This table reports the the first and second stages of IV results of the effect of Internet use to get news on individuals' trust in political institutions. The dependent variables are dummy variables equal to 1 if the individual trusts the political institution in question "a lot" or "somewhat" and 0 otherwise. SMC number × fixed 3G share is the interaction between the number of SMCs and the share of the individual's district of residence covered with 3G network in the first Afrobarometer survey wave used in our analysis. Internet use is an ordered categorical variable equal to 0 if the individual never uses Internet, 1 if he or she uses it less than once a month, 2 if he or she uses it a few times a month, 3 if he or she uses it a few times a week, and 4 if he or she uses it every day. Individual controls include age, age squared, gender, urban dummy, education, employment status, perception of own living conditions, perception of country's economic condition, interest in politics, regular TV use dummy, regular radio use dummy, and regular newspaper use dummy. District controls include nighttime light, local Internet incidence, 2G network coverage, 3G network coverage, and the log of the distance from district's centroid to the closest Internet infrastructure. Country controls include log GDP per capita, Polity2 index, unemployment rate, and the number of SMCs. Standard errors clustered at the district-year level are reported in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' elaboration on Afrobarometer data based on individuals in 35 countries between 2011 and 2018.

the enhanced information and coordination role played by the Internet in organizing collective actions (Fergusson and Molina, 2019; Manacorda and Tesei, 2020; Guiffard, 2022).

We begin by examining citizens' responses to a question about their participation in protests or demonstrations over the past year. We note that the phrasing of the question suggests a loss of trust and dissatisfaction: "Here is a list of actions that people sometimes take as citizens when they are dissatisfied with government performance. For each of these, please tell me whether you, personally, have done any of these things during the past year. If not, would you do this if you had the chance: Participated in a demonstration or protest march". We then create a dummy variable that takes a value of 1 if the respondent reports participating at least once and 0 otherwise.[21] Next, we turn to a more objective measure of protests using the Armed Conflict Location and Event Data Project (ACLED), which is a publicly available dataset that records political violence and protest events across the world, including the location, date, actors involved, and type of event. We assess the relationship between the average frequency of Internet use and the number of protests at

[21] In this case, our benchmark specification is modified to reflect the lagged nature of our dependent variable as follows: $Y_{i,t-1} = \alpha_0 + \beta_0 Internet_i + \boldsymbol{\gamma_0}\mathbf{X_i} + \boldsymbol{\delta_0}\mathbf{W_{d,t}} + \boldsymbol{\sigma_0}\mathbf{Z_{c,t}} + \rho_d + \rho_t + \varepsilon_i$

Table 1.5
Internet use and perception of corruption of political actors.

	(1) Presidency office	(2) Parliament members	(3) Local gov. councilors	(4) Gov. officials	(5) Police	(6) Judges & Magistrates
First stage regression: Internet use						
SMC number × fixed 3G share	0.627***	0.523***	0.557***	0.581***	0.534***	0.553***
	(0.095)	(0.089)	(0.086)	(0.087)	(0.083)	(0.079)
Second stage regression:						
Internet use	0.130	0.266***	0.117	−0.139	−0.020	0.162*
	(0.099)	(0.095)	(0.090)	(0.113)	(0.130)	(0.086)
Individual controls	✓	✓	✓	✓	✓	✓
District controls	✓	✓	✓	✓	✓	✓
Country controls	✓	✓	✓	✓	✓	✓
District FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Observations	89,058	90,885	88,585	93,068	94,982	91,206
Mean dependent variable	0.350	0.377	0.376	0.424	0.515	0.363
Mean Internet use	0.907	0.914	0.930	0.914	0.909	0.910
KP Wald F-stat	43.125	34.575	41.622	44.304	41.609	49.107
KP LM P-val	0.000	0.001	0.000	0.000	0.000	0.000

Notes: This table reports the the first and second stages of IV results of the effect of Internet use to get news on individuals' perception of corruption of political actors. The dependent variables are dummy variables equal to 1 if the individual believes that all or most of the actors in question are involved in corruption and 0 otherwise. SMC number × fixed 3G share is the interaction between the number of SMCs and the share of the individual's district of residence covered with 3G network in the first Afrobarometer survey wave used in our analysis. Internet use is an ordered categorical variable equal to 0 if the individual never uses Internet, 1 if he or she uses it less than once a month, 2 if he or she uses it a few times a month, 3 if he or she uses it a few times a week, and 4 if he or she uses it every day. Individual controls include age, age squared, gender, urban dummy, education, employment status, perception of own living conditions, perception of the country's economic condition, interest in politics, regular TV use dummy, regular radio use dummy, and regular newspaper use dummy. District controls include nighttime light, local Internet incidence, 2G network coverage, 3G network coverage, and the log of the distance from district's centroid to the closest Internet infrastructure. Country controls include log GDP per capita, Polity2 index, unemployment rate, and the number of SMCs. Standard errors clustered at the district-year level are reported in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' elaboration on Afrobarometer data based on individuals in 35 countries between 2011 and 2018.

the district level. Our results are presented in Table 1.6. We find a positive and significant impact of the frequency of Internet use on the probability of attending demonstrations, as well as on the number of protests at the district level.[22,23]

Finally, we measure citizens' voting behavior by relying on the following survey question: "Understanding that some people were unable to vote in the most recent national election in [20xx], which of the following statements is true for you?". First, we consider a dummy variable that takes the value of 1 if the respondent reports voting in the election and 0 otherwise.[24] Then, we consider a second dummy variable that takes the value of 1 if the respondent reports voting in the election or having the intention to vote but couldn't due

[22] The low mean of the number of protests in our sample, relative to the large estimates obtained from regressing the number of protests on the average frequency of Internet use at the district-year level in Column (2) of Table 1.6, can be attributed to the fact that 72% of the districts within our sample have recorded zero protests, thereby pulling the mean downward.

[23] The district-level share of respondents who declare using the Internet on a weekly or daily basis included in our specification allows us to control for the coordination channel, highlighted by Guiffard (2022).

[24] We exclude from the sample those who reported being too young to vote at the time of the most recent national election.

Table 1.6
Internet use and participation in demonstrations.

	Demonstrations	
	(1) Self-reported	(2) ACLED
First stage regression: Internet use		
SMC number × fixed 3G share	0.591***	0.757***
	(0.085)	(0.105)
Second stage regression:		
Internet use	0.249***	43.604***
	(0.050)	(16.708)
Individual controls	✓	✓
District controls	✓	✓
Country controls	✓	✓
District FE	✓	
Year FE	✓	
Observations unit	individuals	districts
Observations	98,878	3,443
Mean dependent variable	0.104	1.922
Mean Internet use	0.895	0.717
KP Wald F-stat	47.870	52.387
KP LM P-val	0.000	0.000
Cluster	district-year	region-year

Notes: This table reports the first and second stages of IV results of the effect of Internet use to get news on attending demonstrations and on the number of protests. The dependent variable in column (1) is a dummy variable equal to 1 if the individual "has ever attended demonstrations" and 0 otherwise. The dependent variable in column (2) is the number of protests at the district-year level according to ACLED data. SMC number × fixed 3G share is the interaction between the number of SMCs and the share of the individual's district of residence covered with 3G network in the first Afrobarometer survey wave used in our analysis. Internet use is an ordered categorical variable equal to 0 if the individual never uses Internet, 1 if he or she uses it less than once a month, 2 if he or she uses it a few times a month, 3 if he or she uses it a few times a week, and 4 if he or she uses it every day. Individual controls include age, age squared, gender, urban dummy, education, employment status, perception of own living conditions, perception of the country's economic condition, interest in politics, regular TV use dummy, regular radio use dummy, and regular newspaper use dummy. District controls include nighttime light, local Internet incidence, 2G network coverage, 3G network coverage, and the log of the distance from district's centroid to the closest Internet infrastructure. Country controls include log GDP per capita, Polity2 index, unemployment rate, and the number of SMCs. Standard errors clustered at the district-year (region-year) level are reported in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' elaboration on Afrobarometer and ACLED datasets based on individuals in 35 countries between 2011 and 2018.

to certain reasons,[25] and 0 if the respondent reports deciding not to vote or not voting for some other unspecified reason. One caveat to using the aforementioned survey question to measure the impact of Internet use on voting behavior is the potential time gap between the survey date and the last election.[26] To address this, we gathered information on each

[25] This includes the following responses: "You were not registered to vote", "You could not find the polling station", "You were prevented from voting", "You did not have time to vote", or "You did not vote because you could not find your name in the voters' register".

[26] It is worth noting that our benchmark specification is modified to take into account that the voting

country's national election dates from the International Foundation for Electoral Systems (IFES) ElectionGuide website and calculated the time elapsed between the survey date and the previous national election. We then introduced an interaction term between Internet use and whether the election took place in the past year, two years ago, or even further in the past, spanning multiple years.[27]

Results reported in Table 1.7 show a positive association between Internet use frequency and voting in the most recent national election, though this association is only significant when considering the interaction with the number of years from the election. This positive effect diminishes as the time elapsed since the last election increases. However, when considering the intention to vote dummy variable, the effect of the Internet becomes significant at the 10% level and exhibits similar patterns when accounting for the number of years from the last election.

Overall, our findings reveal a complex and rich relationship between Internet use, demonstrations, voting, and preference for democracy. While Internet use positively affects participation in demonstrations, maybe as a response to poorly-performing political actors, its impact on voting is less pronounced but still significant when considering individuals' intention to vote and accounting for the time gap between the election and survey dates. This indicates that the Internet can serve as a platform that motivates both spontaneous forms of political engagement, such as demonstrations, and institutional forms like voting. On the one hand, it plays a role in mobilizing individuals, potentially facilitating collective actions, and providing a platform to express dissatisfaction with the political system. On the other hand, it also has the potential to act as a conduit of political information, expanding individuals' knowledge about candidates and elections, thereby fostering increased participation among voters (Tolbert and McNeal, 2003).

One might think that this greater participation through both protests and voting reflects a greater attachment to the intrinsic values of democracy, but our results show the opposite. Internet users are protesting and voting more, but turning away from democracy: how can this apparent puzzle be explained? To address it, we propose a few ideas, based on the interplay of two strands of literature.

The first looks at how protest shapes political attitudes (political preference and voting),

variable reflects a lagged action as follows: $Y_{i,t-n(c)} = \alpha_0 + \beta_0 Internet_i + \boldsymbol{\gamma_0 X_i} + \boldsymbol{\delta_0 W_{d,t}} + \boldsymbol{\sigma_0 Z_{c,t}} + \rho_d + \rho_t + \varepsilon_i$, where n is the time gap between the survey year and the year when the last national election took place.

[27] We instrument the interaction term using our main IV interacted with the number of years elapsed between the survey date and the last election.

Table 1.7
Internet use and voting in national elections.

	Voting		Intention to vote	
	(1)	(2)	(3)	(4)
	First stage regressions			
First stage regression: Internet use				
SMC number × fixed 3G share	0.573***	0.567***	0.573***	0.567***
	(0.085)	(0.126)	(0.085)	(0.126)
SMC number × fixed 3G share × years from election		−0.000		−0.000
		(0.012)		(0.012)
First stage regression: Internet use × years from election				
SMC number × fixed 3G share		0.397		0.397
		(0.509)		(0.509)
SMC number × fixed 3G share × years from election		0.344***		0.344***
		(0.037)		(0.037)
	Second stage regressions			
Internet use	0.023	0.206**	0.128*	0.235***
	(0.076)	(0.089)	(0.073)	(0.073)
Years from election		0.027***		0.019**
		(0.009)		(0.008)
Internet use × years from election		−0.041***		−0.023***
		(0.010)		(0.008)
Individual controls	✓	✓	✓	✓
District controls	✓	✓	✓	✓
Country controls	✓	✓	✓	✓
District FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Observations	92,410	92,410	92,410	92,410
Mean dependent variable	0.762	0.762	0.883	0.883
Mean Internet use	0.862	0.862	0.862	0.862
KP Wald F-stat	45.271	14.128	45.271	14.128
KP LM P-val	0.000	0.000	0.000	0.000

Notes: This table reports the first and second stages of IV results of the effect of Internet use to get news and its interaction with the number of years from last election on voting in the last national elections. The dependent variable in columns (1) and (2) is a dummy variable equal to 1 if the individual "has voted in the last national elections" and 0 otherwise. The dependent variable in columns (3) and (4) is a dummy variable equal to 1 if the individual "has voted in the last national elections" or "could not find the polling station" or "did not have time to vote" or "did not vote because his or her name was not in the registry" or "was not registered to vote" or "was prevented from voting" and 0 otherwise. SMC number × fixed 3G share is the interaction between the number of SMCs and the share of the individual's district of residence covered with 3G network in the first Afrobarometer survey wave used in our analysis. Internet use is an ordered categorical variable equal to 0 if the individual never uses Internet, 1 if he or she uses it less than once a month, 2 if he or she uses it a few times a month, 3 if he or she uses it a few times a week, and 4 if he or she uses it every day. Years from election is the number of years between the survey date and the last national election. Individual controls include age, age squared, gender, urban dummy, education, employment status, perception of own living conditions, perception of the country's economic condition, interest in politics, regular TV use dummy, regular radio use dummy, and regular newspaper use dummy. District controls include nighttime light, local Internet incidence, 2G network coverage, 3G network coverage, and the log of the distance from district's centroid to the closest Internet infrastructure. Country controls include log GDP per capita, Polity2 index, unemployment rate, and the number of SMCs. Standard errors clustered at the district-year level are reported in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' elaboration on Afrobarometer data based on individuals in 35 countries between 2011 and 2018.

and shows that the circumstances surrounding protests are not neutral with regard to their consequences: their size, intensity, form, and the degree of censorship, all shape the efficacy of protests. In the context of the 2006 Latino immigrant rallies in the USA, Wallace et al. (2014) explored how the proximity of small versus large-scale protests has a differential effect on people's perceptions of political efficacy (limited voice in government, politics is complicated). Analyzing the same episode, Branton et al. (2015) used a quasi-experiment setting to demonstrate that the effect of protest on political preferences hinges on the local intensity of street-level activism. Andrews et al. (2016) found that white Southerners have more positive attitudes toward anti-segregation protests in areas that experienced a sit-in during the civil rights movement. Likewise, Mazumder (2018) argues that counties became more politically liberal after having hosted a civil rights demonstration.

The second is a characterization of the Internet generation of protests, based in particular on Tufekci (2017)'s much-cited work: a powerful tool for almost instantaneous coordination and mobilization, the Internet mobilizes large crowds devoid of leadership, with a horizontal structure, and very often proves powerless to structure a reasoned political offer in a context of democratic transition. The structuring of opinions by Facebook through algorithms that prioritize radical ideas, and polarizing debates, is a feature that appears to be transversal to many contexts. Internet-generated protests may have neither the rootedness over time nor the effectiveness that movements like the U.S. civil rights movement were able to achieve, be they in mature democracies (Boyer et al., 2020) or in countries undergoing democratic transition. Tufekci (2017) speaks of tactical freezing, i.e. the inability of these Internet-induced movements to press for tangible political change, as digital technologies enhance their ability to form without too much prior planning, dealing with issues only as they arise, and by the people who emerge ("adhocracy").

Our results therefore find an echo in the literature showing that Internet users' opinions tend to become simplified and radicalized, while the Internet provides them with the means to mobilize through communities that nurture and amplify discontent and expose them to increasingly radicalized content: as Internet-induced protests and content proliferate, distrust toward institutions increases and preference for democracy diminishes. In addition to all this, the vote can increase in favor of conservative parties that advocate for order and stability, rather than the rights attached to democracy and liberalism, from which individuals exposed to street disorder turn away.[28]

[28] While our results focus on the preferences and behaviors of Internet users, the influence of Internet-

1.6.3 Internet as a misinformation technology?

The Internet is often regarded as a "liberation technology" as it provides access to alternative and freer sources of information. However, it is also seen as a "misinformation technology" due to its ability to propagate censored or false information. The liberation versus misinformation technology debate can be apprehended through the lens of two features of the Internet: its low entry barriers cost and its reliance on user-generated rather than expert-generated content. The latter feature gives a voice to marginalized and extremist groups, all the more easily as the absence of safeguarding procedures coupled with the low fact-checking standards lead to a spread of misinformation and fake news, ultimately increasing political misperceptions.[29]

This phenomenon of misinformation can be observed in democracies where the absence of regulation and the principle of press freedom allows false information to spread more easily. Misinformation can also prevail in non-democratic regimes where it is used as a means of propaganda and surveillance (Qin et al., 2017). However, in environments where censorship reigns, the Internet can act as a window to a more open and diverse array of news, from beyond the borders of the country, leading to higher expectations of governments and creating citizens who are prompt to criticize. These benefits of the Internet would make it a "liberation technology". All this is a matter of empirical validation.

To provide evidence of the role played by the Internet as a (mis)information channel, we compare individuals' perception of the level of democracy to experts' ratings using two commonly used indices: the Polity2 index from the Polity5 project and the Regime of the World index (RoW) from the V-Dem dataset. The Polity2 index ranges from -10 (hereditary monarchy) to +10 (consolidated democracy).[30] The RoW index categorizes political regimes into four types: "Closed autocracy", "Electoral autocracy", "Electoral democracy", and "Liberal democracy". The distribution of the Polity2 index and the RoW

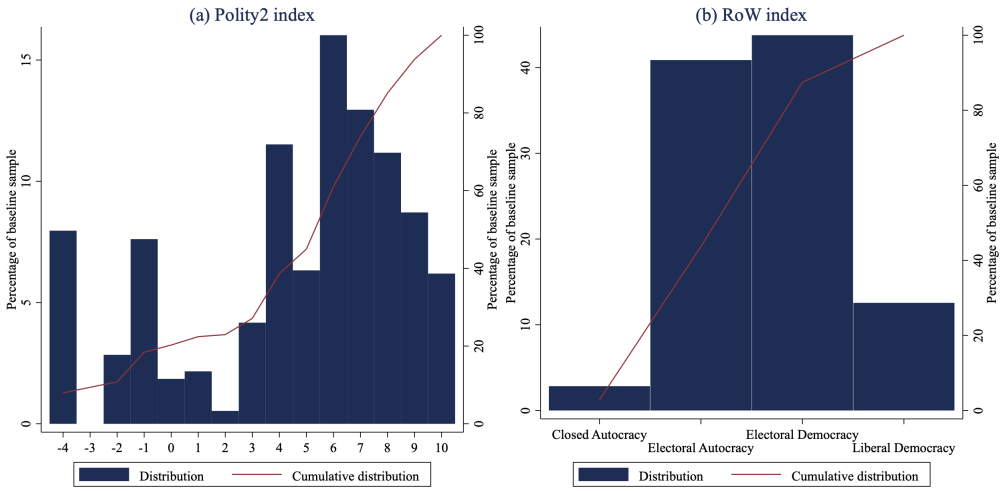
generated protests extends far beyond the circle of the latter and has effects on society as a whole. The Egyptian example is particularly instructive in this respect. In Ketchley and El-Rayyes (2021), the exposure to prolonged and disruptive street protests is shown to have led many Egyptians to equate democracy with the negative externalities of mobilization. By the same logic, greater experience of protests was associated with an increased recognition of the need for order and stability, and a greater readiness to sacrifice human rights in exchange for security (El-Mallakh, 2020). This resulted in a higher share of conservative votes in the regions most affected.

[29] There is a growing literature documenting that political fake news and false information spread online (see, for instance, Mocanu et al. (2015); Allcott and Gentzkow (2017); Grinberg et al. (2019)), and that they spread more rapidly and reach a larger audience than true news (Vosoughi et al., 2018).

[30] Countries scoring between (-10 and -6) are considered "autocracies", those scoring between (-5 and 5) are considered "anocracies", and those scoring 6 or higher (6 to 10) are considered "democracies".

categories in our baseline sample (Fig. 1.7) provides insight into the political landscape. The majority of our observations are in countries scoring 4 or higher on the Polity2 index and classified as either electoral autocracies or democracies according to the RoW index.

Fig. 1.7. Distribution of Polity2 and Regimes of the world (RoW) indices.

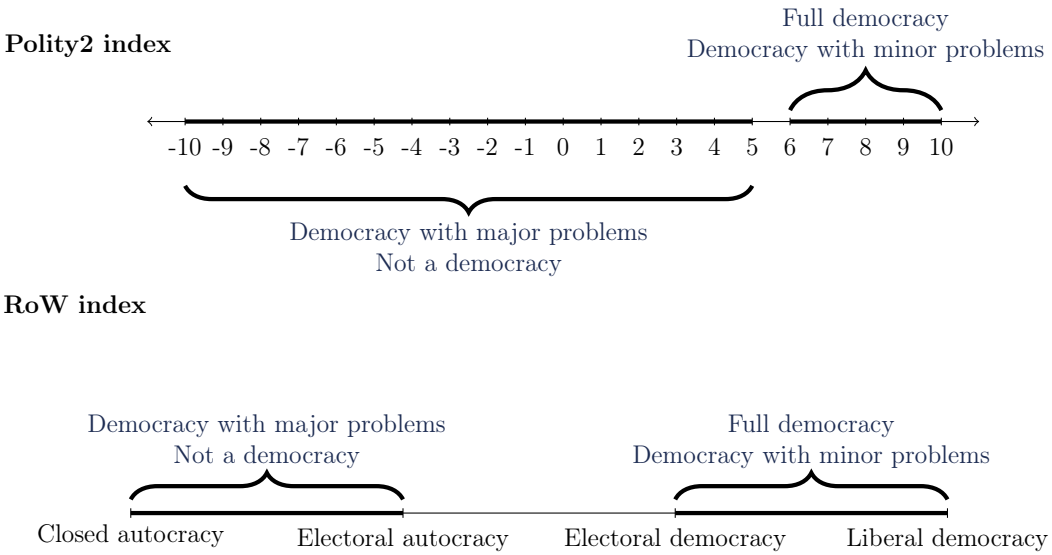


Note: This figure shows the (cumulative) distribution of the Polity2 and RoW indices.
Source: Authors' calculation on Afrobarometer, Polity5, and V-Dem datasets based on the baseline sample of 99,938 individuals in 35 countries between 2011 and 2018.

We create convergence dummies that reflect the convergence between individuals' perceptions and experts' ratings of the level of democracy using each of the two democracy indices. We proceed to a dichotomous split of our convergence dummies as described in Fig. 1.8. They are set equal to 1 if the Polity2 index is greater than or equal to 6, or if the RoW index category is electoral or liberal democracy (Polity2 index is less than or equal to 5, or the RoW index category is closed or electoral autocracy) and the respondent perceives his or her country as a full democracy or a democracy with minor problems (a democracy with major problems or not a democracy).

In addition to comparing the citizens' perceptions and experts' ratings of the level of democracy, we also compare the citizens' perceptions and experts' ratings of the level of corruption among legislators. In Section 1.6.1, our findings indicate that Internet use is associated with a decrease in trust in the parliament and the ruling party, as well as an increase in the perceived corruption of parliament members. We create a convergence dummy variable that measures the alignment between citizens' perceptions of corruption among parliament members on the one hand, and experts' ratings on the other. To this end, we rely on the following question from the V-Dem database, which gauges legislators' involvement in corrupt practices: "Do members of the legislature abuse their position for

Fig. 1.8. Convergence dummies construction based on Polity2 and Regimes of the world (RoW) indices.



Notes: This figure illustrates the construction of convergence dummies based on the Polity2 and RoW indices. The dummy variable equals 1 if the Polity2 index is greater than or equal to 6, or if the RoW index category is electoral or liberal democracy (Polity2 index is less than or equal to 5, or the RoW index category is closed or electoral autocracy) and the respondent perceives his or her country as a full democracy or a democracy with minor problems (a democracy with major problems or not a democracy). Source: Authors' elaboration on Afrobarometer, Polity5, and V-Dem datasets.

financial gain? This includes any of the following: (a) accepting bribes, (b) helping to obtain government contracts for firms that the legislator (or his/her family/friends/political supporters) own, (c) doing favors for firms in exchange for the opportunity of employment after leaving the legislature, (d) stealing money from the state or from campaign donations for personal use".[31] This variable takes the value of 1 if the individual believes that "all" or "most" ("some" or "none") members of parliament are involved in corruption and if experts' assessments indicate that "most" or "many" ("some", "a few", or "none") legislators probably engage in the aforementioned mentioned corrupt activities.

To further test the hypothesis that the Internet is a (mis)information technology, we aim to investigate whether using the Internet to get informed increases or decreases comprehension of the core principles of democracy, such as the separation of executive, legislative, and judiciary power, freedom of speech, free and fair election, the rule of law, and other characteristics that are often absent in non-democratic regimes.

Our variable for strict preference for democracy reflects a true preference for democ-

[31] Information is missing for Egypt (2013, 2015), Guinea (2013), and Mali (2012).

racy, meaning that citizens who respond "democracy is preferable to any other kind of governance" reject all three authoritarian alternatives (one-man rule, one-party rule, and military rule). However, we find some incoherent answers when jointly looking at the preference for democracy and the rejection of authoritarian rule variables. Although some citizens say they prefer democracy, they are willing to tolerate certain types of authoritarian rule. Similarly, among citizens who say that non-democratic governance can be preferable at times, some tend to reject all alternative authoritarian rules.

In Table 1.8, we report the percentage of those who reject none, one, two, or all authoritarian alternatives among those who state "in some circumstances, a non-democratic government can be preferable" and those who claim "democracy is preferable to any other kind of government". We find that 32% of those who prefer democracy do not reject all authoritarian alternatives and around 46% of those who sometimes prefer non-democratic rule reject all of them. These results suggest an inconsistency in respondents' answers.

Table 1.8
Preference for democracy and rejection of authoritarian alternatives.

Authoritarian alternatives	Sometimes non-democratic preferable	Democracy preferable	Total
Reject none	**7.88**	3.61	4.17
Reject one	**14.66**	7.27	8.24
Reject two	**31.66**	21.07	22.46
Reject all	45.79	**68.04**	65.13
Total	100	100	100

Notes: This table reports the percentage of respondents rejecting none, one, two, or all three authoritarian alternative rules (one-man rule, one-party rule, and military rule) among those who say that "democracy is preferable to any other kind of government" and those who say that "in some circumstances, a non-democratic government can be preferable". A coherent response is when an individual says "democracy is preferable to any other kind of government" and "rejects all three authoritarian alternatives" or when he or she says "in some circumstances, a non-democratic government can be preferable" and "does not reject all three authoritarian alternatives". Coherent responses are displayed in bold.
Source: Authors' elaboration on Afrobarometer data based on 89,050 individuals in 35 countries between 2011 and 2018.

To analyze the impact of Internet use on the probability of providing coherent responses to both questions, we consider a dummy variable that takes a value of 1 if citizens either "prefer democracy to any kind of government" and "reject all three authoritarian alternatives", or if they believe that "in some circumstances, a non-democratic government can be preferable" and "reject at most two of the three authoritarian alternatives".[32] We present the IV results of the impact of using the Internet to get news on convergence and

[32] We exclude from the sample those who respond "to people like me, it does not matter what form of government we have" to the preference for democracy question as it does not provide a clear answer to compute the coherent answer dummy variable.

coherence dummies in Table 1.9.

Table 1.9
Internet use, convergence toward experts' ratings, and coherent responses to preference for democracy.

	Convergence			Coherence
	(1) Polity2	(2) RoW	(3) Corruption	(4) Coherence
First stage regression: Internet use				
SMC number × fixed 3G share	0.594***	0.594***	0.514***	0.640***
	(0.083)	(0.083)	(0.089)	(0.079)
Second stage regression:				
Internet use	−0.283**	−0.278**	−0.346***	−0.187***
	(0.128)	(0.123)	(0.111)	(0.071)
Individual controls	✓	✓	✓	✓
District controls	✓	✓	✓	✓
Country controls	✓	✓	✓	✓
District FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Observations	99,938	99,938	88,283	89,050
Mean dependent variable	0.586	0.578	0.580	0.662
Mean Internet use	0.895	0.895	0.928	0.902
KP Wald F-stat	51.050	51.050	33.422	66.075
KP LM P-val	0.000	0.000	0.001	0.000

Notes: This table reports the first and second stages of IV results of the effect of Internet use to get news on individuals' convergence toward experts' ratings and coherent responses to the strict preference for democracy variable. The dependent variables in columns (1), (2), and (3) are dummy variables equal to 1 if the individual convergences toward experts' ratings and 0 otherwise. In column (1), we rely on the Polity2 index from the Polity5 project dataset to compute the convergence dummy. In column (2), we rely on the RoW index from the V-Dem dataset. In column (3), we rely on the legislature corrupt activities index from the V-Dem dataset. The dependent variable in column (4) is a dummy variable equal to 1 if the individual responds coherently to strict preference variable set of questions responding that "democracy is preferable to any other kind of government" and "rejecting all three authoritarian alternatives" or responding that "in some circumstances, a non-democratic government can be preferable" and "does not reject all three authoritarian alternatives" and 0 otherwise. SMC number × fixed 3G share is the interaction between the number of SMCs and the share of the individual's district of residence covered with 3G network in the first Afrobarometer survey wave used in our analysis. Internet use is an ordered categorical variable equal to 0 if the individual never uses Internet, 1 if he or she uses it less than once a month, 2 if he or she uses it a few times a month, 3 if he or she uses it a few times a week, and 4 if he or she uses it every day. Individual controls include age, age squared, gender, urban dummy, education, employment status, perception of own living conditions, perception of the country's economic condition, interest in politics, regular TV use dummy, regular radio use dummy, and regular newspaper use dummy. District controls include nighttime light, local Internet incidence, 2G network coverage, 3G network coverage, and the log of the distance from district's centroid to the closest Internet infrastructure. Country controls include log GDP per capita, Polity2 index, unemployment rate, and the number of SMCs. Standard errors clustered at the district-year level are reported in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' elaboration on Afrobarometer, Polity5, and V-Dem datasets based on individuals in 35 countries between 2011 and 2018.

We find that regardless of the index used to construct the convergence dummy, the effect of the Internet is consistently negative and statistically significant, reflecting a divergence

from experts' ratings, be it on the issue of democracy or corruption.[33] These results support the hypothesis that the Internet may act as a potential source of misinformation.[34] Additionally, we find that Internet use decreases the probability of providing coherent answers, thereby suggesting a lack of consistent understanding of the questions related to the preference for democracy. This finding supplements our previous results on the Internet's role as a misinformation technology. Nonetheless, the inconsistency in responses could be perceived as an indication of a varying comprehension of the term "democracy" among different countries and institutional contexts.[35]

1.7 Robustness checks

We run several robustness checks on our IV baseline estimates reported in Table 1.3. We provide results tables in Appendix A.3.

Alternative instrumentation. First, we test the robustness of our results to alternative instrumentation. As access to 2G mobile network provides basic Internet connectivity, which may allow the dissemination of political information, especially at the beginning of the estimation period or in remote areas, we construct a complementary instrument that weighs the number of SMCs by the fixed district's 2G coverage ($SMC_{c,t} \times fixed2Gshare_{d,c}$). Even though accessing the Internet via the 2G network is not as easy as through the 3G or 4G networks, 2G technology was prevalent to access the Internet in Africa at the beginning of the 2010s. This additional instrument may therefore explain further variation in Internet use. We use both instruments in our IV regressions presented

[33] Given the continuous nature of the Polity2 index, we provide alternative approaches in Fig. A.2 to construct the convergence dummy. These approaches entail varying the threshold used to define our dichotomous variable. Results are reported in Table A.9 in Appendix A.2 and remain consistent.

[34] In a similar vein, a recent paper by Amaral-Garcia et al. (2022) on health procedures showed that mothers with better Internet access had higher rates of elective C-sections, which are chosen by the less informed mothers, rather than emergency C-sections, which are typically recommended by medical experts. This also suggests that the availability of online information may lead to deviations from expert recommendations and sub-optimal decision-making by individuals, depending on the quality of online information and the potential spread of misinformation.

[35] We report the heterogeneous effect of Internet use on convergence and coherence dummies by press censorship and education level in Tables A.10 and A.11 in Appendix A.2, respectively. To measure press censorship, we rely on the Freedom House's Freedom of the Press (FOTP) index which assesses traditional media freedom with a score ranging between 0 (most free) and 100 (less free). We find that using the Internet when the press is most free is associated with a negative effect on convergence toward experts' ratings, which is attenuated as the FOTP increases. These results suggest that the Internet is more likely to act as a misinformation technology when other (traditional) media are free and likely to provide uncensored information. In less democratic environments, where there is no alternative, the Internet provides access to information from the rest of the world and could fill the information gap caused by censorship. We find no significant differences on the coherence dummy nor by education level.

in Table A.12, and find a first stage with the expected signs and a high (effective) F-stat (see Table A.13 for the effective F-stat). The second stage results remain robust, except for satisfaction with democracy, which is no longer significant. We rely on the Hansen J statistic P-value to test the over-identification restriction and find that it is less than 5%, suggesting a potential risk of endogeneity regarding this particular outcome.[36]

Controlling for satisfaction with democracy. Second, we assess citizens' satisfaction with democracy as a potential confounding factor.[37] For instance, citizens (dis)satisfied with an electoral outcome or a government policy might be more prone to interpreting information acquired online in a way that magnifies their (negative) positive perceptions of democracy. Additional estimations are presented in Table A.15 and remain robust to controlling for this potentially omitted factor. We even observe an increase in the magnitude of the coefficient of interest, suggesting satisfaction with democracy dampens the true effect of Internet use on individuals' perceptions of democratic institutions.

Excluding countries one by one. To ensure that our results are not driven by any particular country, we assess their robustness by dropping one country at a time from our estimations. The second stage coefficients on Internet use are presented in Fig. A.3. Results in panel (c) indicate that the negative effect of Internet use on the perception of the extent of democracy is sensitive to the removal of South Africa from the sample, but this sensitivity disappears when the mediating effect of satisfaction with democracy is controlled for as shown in Fig. A.4. Otherwise, our main conclusions remain unchanged.

Excluding districts with less than 30 observations. We then re-estimate the model by excluding districts with less than 30 observations to ensure that our findings are not driven by the small sample size in certain districts. The results, as shown in Table A.16, indicate that our estimates for the Internet variable remain negative and significant, but their magnitude increases considerably. Nevertheless, we observe a loss of statistical significance for satisfaction with democracy.

Excluding districts with no Internet users. We perform additional estimations by excluding districts where no individuals reported using the Internet to ensure the robustness of our findings. The results remain consistent and robust, as shown in Table A.17.

Excluding control variables. We further test the robustness of our findings by removing potential confounding variables, such as satisfaction with economic conditions,

[36] We also report the conditional likelihood ratio (CLR) test in Table A.14.

[37] Previous analysis has shown that this outcome is a fluctuating construct more difficult to analyze, as highlighted in the previous robustness check, and as discussed in subsection 1.5.2.

interest in politics, traditional media use, and country-specific controls, from our estimations. Fig. A.5 illustrates that our results remain robust to the exclusion of these variables.

Falsification test. Finally, we conduct a falsification test following a recent study by Borusyak and Hull (2023). The authors note that omitted variable bias can arise from non-random exposure to an exogenous shock when studying its effects on an outcome variable. In our study, the number of SMCs is as good as randomly assigned, but a district's exposure to their deployment may be non-random. In fact, the observation unit's non-random exposure to shocks may lead to bias that could arise if some units are systematically associated with a higher probability of using the Internet than others, as a consequence of their non-random exposure to connectivity shocks. For instance, even when the deployment of SMCs is as-good-as-randomly assigned across countries, individuals located in economic and geographic centers may, for a given 3G coverage, be more likely to use the Internet than those located in peripheral districts as the former could benefit from a faster or more stable Internet connection. To address this potential bias, we follow the authors' proposed solution. We generate counterfactual shocks by first generating and averaging random normal, Poisson, and uniform draws for the number of SMCs. Next, we regress our initial IV on these randomly generated shocks weighted by the fixed 3G network share (including district and year fixed effects) and we retrieve the resulting residuals to obtain what Borusyak and Hull (2023) refer to as the recentered instrument that purges bias from non-random exposure. We present the impact of the Internet on our dependent variables using the recentered IV in Table A.18. First stage results remain robust, with the recentered IV having a positive and significant impact on Internet use, and our F-stat remains sufficiently high. Second stage results are also robust to this falsification test and consistent with our baseline estimates.

1.8 Conclusions

The widespread use of the Internet as a means of information and communication has fueled ongoing debates on whether it serves as a tool for promoting open and freer access to information or, conversely, as a facilitator of misinformation. All these debates are about the implications and significance of the technology of the Internet for democracy, as argued by Flynn et al. (2017). If the use of the Internet produces biases in perception, beliefs, and judgments, which do not cancel each other out, this can lead to misperceptions at the macro level and to the formulation of erroneous social and economic policies (on immigration, security, public health, public finance), with injurious consequences. Despite

the prevalence of such discussions, there is still a need for careful micro-level analysis on the impact of Internet use on perceptions of democracy, particularly in a developing context (Acemoglu et al., 2021). This paper aims to fill this gap by investigating the role of the Internet as a (mis)information technology through an analysis of the impact of regular Internet news consumption on citizens' perception of democracy in 35 African countries using three rounds of the Afrobarometer survey spanning from 2011 to 2018.

We use citizens' preference for democracy, perception of the level of democracy, and satisfaction with democracy as our main outcome variables. To analyze the impact of Internet use on these variables, we adopt an IV approach, which combines an external and internal source of digital vulnerability as an instrument. Specifically, we consider the number of SMCs as an exogenous connectivity shock and weigh it by districts' 3G signal coverage as an exposure factor to instrument Internet use.

This paper's main finding is that the Internet as an alternative source of news has a negative and significant effect on citizens' preference for and perception of the extent of democracy. A one-unit increase in Internet use frequency decreases the probability of (strictly) preferring democracy by 31.7 (26) and perceiving the country as a democracy by 23.3 percentage points. This suggests that citizens are more likely to prefer non-democratic governance in some circumstances and develop more negative views of the country's level of democracy. However, there was a positive, although not robust, effect on satisfaction with democracy. We believe that satisfaction with the functioning of democracy in one's country is a complex and multifaceted construct that can be challenging to analyze. It may be influenced by various factors, such as frustration after an undesired or unanticipated electoral election outcome or government policies, which are difficult to observe and control for in our analysis. In comparison, measures of preference for democracy and perceptions of its level are relatively more straightforward and precise concepts that can provide a clearer picture of citizens' attitudes toward democracy.

When investigating the potential channels through which Internet news negatively affects citizens' perception of democracy, we found that frequent use of the Internet for news leads to a decrease in trust in the parliament and the ruling party, as well as an increased perception of corruption among parliament members. This aligns with Guriev et al. (2021) on 3G expansion and government approval. We also found that Internet users are more likely to engage in street protests, which is consistent with deteriorated trust in African political leaders and institutions resulting in street protests (Sangnier and Zylberberg, 2017). The Internet-induced increase in voting may correspond to a desire for greater order and stability, values promoted by the conservative parties, and which in

fact coincide with the distrust of democracy that we are highlighting. Additionally, our study suggests that negative attitudes toward democracy and its institutions may stem from a misperception of how it functions. We document that Internet users' perception of the level of democracy and corruption diverges from experts' ratings. Furthermore, we find that Internet users are more likely to give inconsistent answers regarding preference for democracy questions, which may result from an altered understanding of democracy across different countries and institutional settings.

The findings of this article contribute to the wider literature on the role of information technology in consolidating democracy in Africa. They have important policy implications, particularly in the context of developing countries where democratic institutions may be more fragile and the journey toward democratization may experience reversals. By spreading false information, blurring the visibility of government action, reducing accountability and trust in democratic institutions, and leading people to doubt them, the Internet can encourage people to distrust democracy and prefer other regimes. What's more, citizens influenced by the Internet tend to express their discontent and distrust by engaging in street protests, sometimes violent, which people don't like and whose negative externalities they associate with democracy.

Governments need to take steps to ensure that citizens have access to accurate and reliable information while also addressing issues of corruption and political accountability. In addition, efforts should be made to promote civic engagement, as a means of strengthening democracy. Finally, media literacy programs and fact-checking can play an important role in helping citizens critically evaluate the information they find online and form informed opinions about their country's governance and policies, thus limiting the spread of online fake news (Barrera et al., 2020; Henry et al., 2022). Hence, as the Internet continues to play a vital role in shaping public opinion, policymakers and media outlets need to combat misinformation and promote critical thinking among citizens.

A Appendix to Chapter 1

A.1 Additional descriptive statistics

Table A.1
Sample statistics, by country.

Country	Freq.	Percent
Algeria	526	0.53
Benin	2,969	2.97
Botswana	3,131	3.13
Burkina Faso	2,808	2.81
Burundi	1,102	1.10
Cameroon	2,562	2.56
Cape Verde	2,940	2.94
Cote d'Ivoire	2,423	2.42
Egypt	820	0.82
Gabon	2,094	2.10
Ghana	4,216	4.22
Guinea	3,238	3.24
Kenya	4,830	4.83
Lesotho	1,912	1.91
Liberia	2,420	2.42
Madagascar	1,497	1.50
Malawi	4,945	4.95
Mali	3,266	3.27
Mauritius	3,246	3.25
Morocco	2,081	2.08
Mozambique	3,094	3.10
Namibia	2,084	2.09
Niger	2,944	2.95
Nigeria	3,572	3.57
Senegal	2,955	2.96
Sierra Leone	2,739	2.74
South Africa	5,783	5.79
Sudan	2,494	2.50
Tanzania	5,105	5.11
Togo	2,841	2.84
Tunisia	719	0.72
Uganda	4,345	4.35
Zambia	2,762	2.76
Zimbabwe	4,507	4.51
Gambia	968	0.97
Total	99,938	100.00

Note: This table reports the number of observations by country.
Source: Authors' elaboration on Afrobarometer data based on the baseline sample of 99,938 individuals in 35 countries between 2011 and 2018.

Table A.2
Sample statistics, by year.

Year	Freq.	Percent
2011	8,214	8.22
2012	19,986	20.00
2013	7,596	7.60
2014	19,828	19.84
2015	12,813	12.82
2016	4,940	4.94
2017	16,695	16.71
2018	9,866	9.87
Total	99,938	100.00

Note: This table reports the number of observations by year.
Source: Authors' elaboration on Afrobarometer data based on the baseline sample of 99,938 individuals in 35 countries between 2011 and 2018.

Table A.3
Additional data sources.

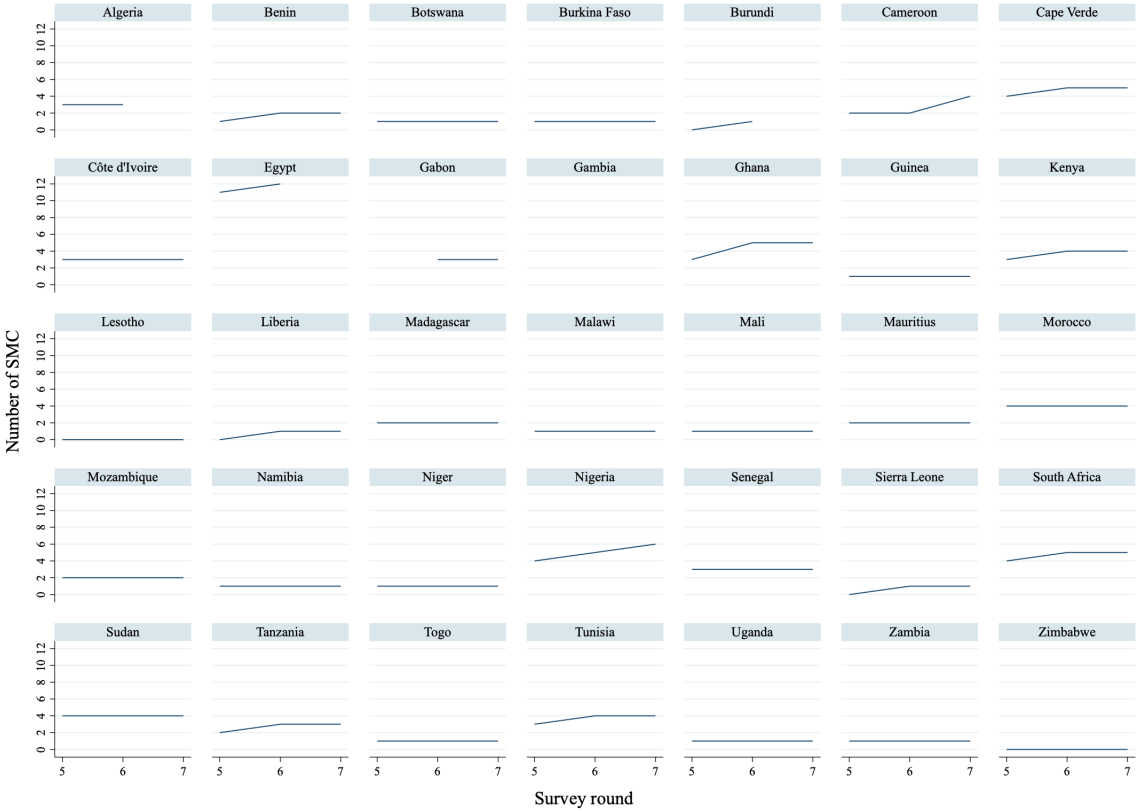
Variable	Source
Nighttime light	Harmonized nighttime light data from the Defense Meteorological Satellite Program (DMSP) and the Visible Infrared Imaging Radiometer Suite (VIIRS) from Li et al. (2020).
Network coverage	GSMA-Collins Bartholomew's Mobile Coverage Explorer.
Distance to infrastructures	Data on SMC landing stations and IXPs GPS coordinates are obtained from Telegeography website and completed by the Packet Clearing House and Peering DB databases. Distance is calculated using QGIS.
Number of SMCs	Telegeography
Polity2 index	Polity5 project
RoW index	V-Dem dataset
Log GDP per capita	World Development Indicators
Unemployment rate	World Development Indicators
Number of protests	The Armed Conflict Location & Event Data Project (ACLED)

Table A.4
Correlation matrix of the dependent variables.

	Preference	Strict preference	Extent	Satisfaction
Preference	1.000			
Strict preference	0.570***	1.000		
Extent	0.096***	0.057***	1.000	
Satisfaction	0.105***	0.059***	0.526***	1.000

Note: This table reports the correlation coefficients between preference for democracy, strict preference for democracy, extent of democracy, and satisfaction with democracy; *** p<0.01, ** p<0.05, * p<0.1.
Source: Authors' elaboration on Afrobarometer data based on the baseline sample of 99,938 individuals in 35 countries between 2011 and 2018.

Fig. A.1. Variation in the number of SMCs by country.



Notes: This figure illustrates the variation in the number of SMCs by country over the three survey rounds.
Source: Authors' elaboration on Telegeography data.

A.2 Additional results

Table A.5
OLS results with control variables.

	(1) Preference	(2) Strict Preference	(3) Extent	(4) Satisfaction
Media use				
Internet use	−0.002	0.007***	−0.008***	−0.009***
	(0.002)	(0.002)	(0.002)	(0.002)
Radio use	0.017***	0.028***	0.010**	0.018***
	(0.004)	(0.004)	(0.004)	(0.004)
Newspaper use	−0.003	−0.003	−0.000	−0.003
	(0.005)	(0.005)	(0.005)	(0.005)
TV use	0.018***	0.034***	0.006	0.022***
	(0.004)	(0.005)	(0.005)	(0.005)
Individual controls				
Age	0.004***	0.008***	−0.002***	−0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Age squared	−0.000***	−0.000***	0.000***	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)
Male	0.014***	0.035***	−0.015***	−0.008**
	(0.003)	(0.003)	(0.004)	(0.003)
Urban	−0.004	0.006	−0.008	−0.017***
	(0.005)	(0.006)	(0.005)	(0.006)
Educational attainment				
Primary	0.014***	0.028***	−0.023***	−0.018***
	(0.006)	(0.007)	(0.006)	(0.006)
Secondary	0.039***	0.088***	−0.035***	−0.042***
	(0.006)	(0.007)	(0.007)	(0.006)
Post-secondary	0.064***	0.144***	−0.059***	−0.049***
	(0.007)	(0.009)	(0.008)	(0.008)
Employment status				
No (looking)	0.001	−0.020***	−0.020***	−0.014***
	(0.005)	(0.005)	(0.005)	(0.005)
Yes, part time	−0.008	−0.024***	−0.008	−0.000

	(0.006)	(0.007)	(0.006)	(0.006)
Yes, full time	0.007	0.000	0.001	0.004
	(0.005)	(0.006)	(0.005)	(0.005)
Discuss politics				
Ocassionally	0.035***	0.053***	0.001	−0.002
	(0.004)	(0.004)	(0.005)	(0.005)
Frequently	0.051***	0.066***	−0.023***	−0.004
	(0.005)	(0.006)	(0.006)	(0.006)
Country economic condition				
(Very) good	0.005	−0.015***	0.137***	0.165***
	(0.004)	(0.005)	(0.005)	(0.005)
Own living conditions				
(Very) good	−0.002	−0.010**	0.053***	0.074***
	(0.004)	(0.005)	(0.004)	(0.005)
District controls				
Nighttime light	−0.003**	−0.003	0.003*	0.002
	(0.002)	(0.002)	(0.002)	(0.002)
2G coverage	0.020	0.056**	0.107***	0.019
	(0.023)	(0.027)	(0.026)	(0.030)
3G coverage	−0.036**	−0.085***	−0.039**	−0.019
	(0.015)	(0.019)	(0.018)	(0.018)
Log distance (in km)	0.006	0.008	−0.014**	−0.002
	(0.007)	(0.009)	(0.007)	(0.006)
Internet incidence	−0.020***	−0.005	−0.015*	−0.005
	(0.007)	(0.009)	(0.009)	(0.009)
Country controls				
Number of SMCs	0.006	−0.010	0.001	0.007
	(0.008)	(0.010)	(0.010)	(0.010)
Polity2 index	0.006*	0.010**	0.027***	0.025***
	(0.004)	(0.005)	(0.003)	(0.003)
Log GDP per capita	−0.267***	−0.147	−0.059	−0.031
	(0.079)	(0.109)	(0.084)	(0.097)
Unemployment rate	−0.010***	0.001	0.006	0.001
	(0.003)	(0.003)	(0.004)	(0.003)
District FE	✓	✓	✓	✓

Year FE	✓	✓	✓	✓
Observations	99,938	99,938	99,938	99,938
Adjusted R^2	0.080	0.129	0.173	0.183
Mean dependent variable	0.773	0.525	0.566	0.503
Mean Internet use	0.895	0.895	0.895	0.895

Notes: This table reports the OLS results of the effect of Internet use to get news on individuals' perception of democracy. The dependent variable in column (1) is a dummy variable equal to 1 if the individual says "democracy is preferable to any other kind of government" and 0 otherwise. In column (2), it is a dummy variable equal to 1 if the individual says "democracy is preferable to any other kind of government" and "rejects all three authoritarian alternatives" and 0 otherwise. In column (3), it is a dummy variable equal to 1 if the individual perceives his or her country as "a full democracy" or "a democracy with minor problems" and 0 otherwise. In column (4), it is a dummy variable equal to 1 if the individual is "very" or "fairly" satisfied with how democracy works in his or her country and 0 otherwise. Standard errors clustered at the district-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' elaboration on Afrobarometer data based on the baseline sample of 99,938 individuals in 35 countries between 2011 and 2018.

Table A.6
Reduced form regressions.

	(1) Preference	(2) Strict Preference	(3) Extent	(4) Satisfaction
SMC number × fixed 3G share	−0.189***	−0.155***	−0.138**	0.153***
	(0.050)	(0.048)	(0.061)	(0.058)
Individual controls	✓	✓	✓	✓
District controls	✓	✓	✓	✓
Country controls	✓	✓	✓	✓
District FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Observations	99,938	99,938	99,938	99,938
Adjusted R^2	0.080	0.129	0.173	0.183
Mean dependent variable	0.773	0.525	0.566	0.503
Wald test P-val	0.000	0.001	0.023	0.008

Notes: This table reports the reduced form results. The dependent variable in column (1) is a dummy variable equal to 1 if the individual says "democracy is preferable to any other kind of government" and 0 otherwise. In column (2), it is a dummy variable equal to 1 if the individual says "democracy is preferable to any other kind of government" and "rejects all three authoritarian alternatives" and 0 otherwise. In column (3), it is a dummy variable equal to 1 if the individual perceives his or her country as "a full democracy" or "a democracy with minor problems" and 0 otherwise. In column (4), it is a dummy variable equal to 1 if the individual is "very" or "fairly" satisfied with how democracy works in his or her country and 0 otherwise. SMC number × fixed 3G share is the interaction between the number of SMCs and the share of the individual's district of residence covered with 3G network in the first Afrobarometer survey wave used in our analysis. Individual controls include age, age squared, gender, urban dummy, education, employment status, perception of own living conditions, perception of the country's economic condition, interest in politics, regular TV use dummy, regular radio use dummy, and regular newspaper use dummy. District controls include nighttime light, local Internet incidence, 2G network coverage, 3G network coverage, and the log of the distance from district's centroid to the closest Internet infrastructure. Country controls include log GDP per capita, Polity2 index, unemployment rate, and the number of SMCs. Standard errors clustered at the district-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' elaboration on Afrobarometer data based on the baseline sample of 99,938 individuals in 35 countries between 2011 and 2018.

Table A.7
Anderson-Rubin (AR) test.

	(1) Preference	(2) Strict Preference	(3) Extent	(4) Satisfaction
AR P-val	0.009	0.022	0.075	0.021
AR Confidence set	[-0.574,-0.103]	[-0.415,-0.061]	[-0.470, 0.022]	[0.050, 0.506]

Note: This table reports the Anderson-Rubin (AR) weak-instrument-robust test P-value and confidence intervals of the coefficient on the endogenous variable (Internet use) in the IV estimations.

Source: Authors' elaboration on Afrobarometer data based on the baseline sample of 99,938 individuals in 35 countries between 2011 and 2018.

Table A.8
IV estimates of the effect of Internet use on perception of democracy - binary Internet use.

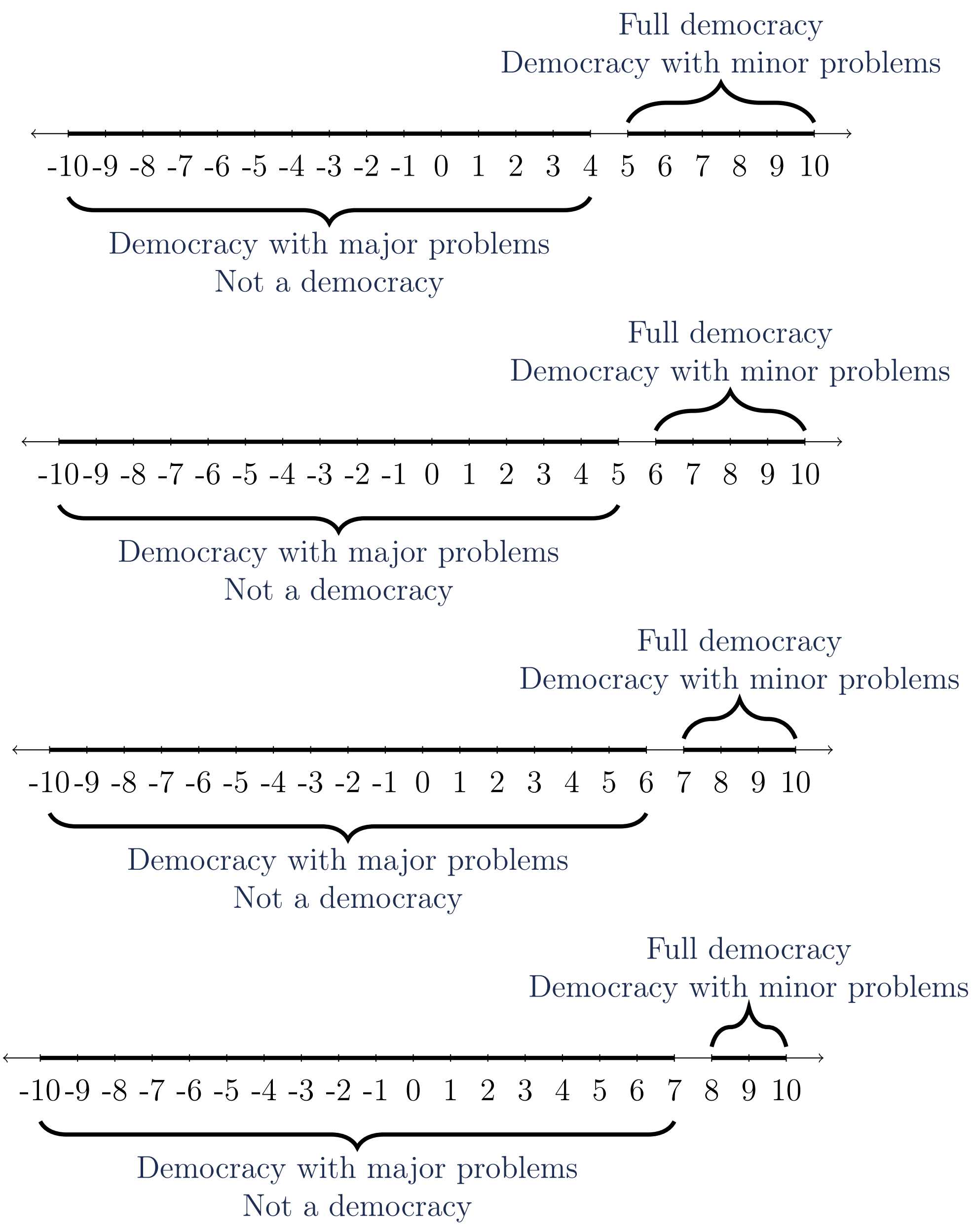
	(1) Preference	(2) Strict Preference	(3) Extent	(4) Satisfaction
First stage regression: Internet use				
SMC number × fixed 3G share	0.146***	0.146***	0.146***	0.146***
	(0.027)	(0.027)	(0.027)	(0.027)
Second stage regression:				
Internet use	−1.288***	−1.057***	−0.945*	1.045**
	(0.499)	(0.346)	(0.503)	(0.416)
Individual controls	✓	✓	✓	✓
District controls	✓	✓	✓	✓
Country controls	✓	✓	✓	✓
District FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Observations	99,938	99,938	99,938	99,938
Mean dependent variable	0.773	0.525	0.566	0.503
Mean Internet use	0.895	0.895	0.895	0.895
KP Wald F-stat	29.908	29.908	29.908	29.908
KP LM P-val	0.000	0.000	0.000	0.000

Notes: This table reports the first and second stages of IV results of the effect of Internet use to get news on individuals' perception of democracy. The dependent variable in column (1) is a dummy variable equal to 1 if the individual says "democracy is preferable to any other kind of government" and 0 otherwise. In column (2), it is a dummy variable equal to 1 if the individual says "democracy is preferable to any other kind of government" and "rejects all three authoritarian alternatives" and 0 otherwise. In column (3), it is a dummy variable equal to 1 if the individual perceives his or her country as "a full democracy" or "a democracy with minor problems" and 0 otherwise. In column (4), it is a dummy variable equal to 1 if the individual is "very" or "fairly" satisfied with how democracy works in his or her country and 0 otherwise. SMC number × fixed 3G share is the interaction between the number of SMCs and the share of the individual's district of residence covered with 3G network in the first Afrobarometer survey wave used in our analysis. Internet use is a binary variable equal to 1 if the individual uses the Internet "every day" or "a few times a week" to get news and 0 otherwise. Individual controls include age, age squared, gender, urban dummy, education, employment status, perception of his or her own living conditions, perception of the country's economic condition, interest in politics, regular TV use dummy, regular radio use dummy, and regular newspaper use dummy. District controls include nighttime light, local Internet incidence, 2G network coverage, 3G network coverage, and the log of the distance from district's centroid to the closest Internet infrastructure. Country controls include log GDP per capita, Polity2 index, unemployment rate, and the number of SMCs. Standard errors clustered at the district-year level are reported in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' elaboration on Afrobarometer data based on the baseline sample of 99,938 individuals in 35 countries between 2011 and 2018.

Fig. A.2. Convergence dummies construction based on Polity2 index.

Polity2 index



Notes: This figure illustrates the construction of convergence dummies based on the Polity2 index. The dummy variable equals 1 if the Polity2 index is $\geq$ ($\leq$) to a certain threshold and the respondent perceives his or her country as a full democracy or a democracy with minor problems (a democracy with major problems or not a democracy).
Source: Authors' elaboration on Afrobarometer and Polity5 datasets.

Table A.9
Internet use and convergence toward experts' ratings: alternative thresholds.

	(1) Polity2 ≥ 5	(2) Polity2 ≥ 6	(3) Polity2 ≥ 7	(4) Polity2 ≥ 8
First stage regression: Internet use				
SMC number × fixed 3G share	0.594***	0.594***	0.594***	0.594***
	(0.083)	(0.083)	(0.083)	(0.083)
Second stage regression:				
Internet use	−0.266**	−0.283**	−0.708***	−0.512***
	(0.126)	(0.128)	(0.151)	(0.148)
Individual controls	✓	✓	✓	✓
District controls	✓	✓	✓	✓
Country controls	✓	✓	✓	✓
District FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Observations	99,938	99,938	99,938	99,938
Mean dependent variable	0.593	0.586	0.551	0.508
Mean Internet use	0.895	0.895	0.895	0.895
KP Wald F-stat	51.050	51.050	51.050	51.050
KP LM P-val	0.000	0.000	0.000	0.000

Notes: This table reports the first and second stages of IV results of the effect of Internet use to get news on individuals' convergence toward experts' ratings using alternative proxies. The dependent variables are dummy variables equal to 1 if the individual convergences toward experts' ratings and 0 otherwise (refer to Fig. A.2). SMC number × fixed 3G share is the interaction between the number of SMCs and the share of the individual's district of residence covered with 3G network in the first Afrobarometer survey wave used in our analysis. Internet use is an ordered categorical variable equal to 0 if the individual never uses Internet, 1 if he or she uses it less than once a month, 2 if he or she uses it a few times a month, 3 if he or she uses it a few times a week, and 4 if he or she uses it every day. Individual controls include age, age squared, gender, urban dummy, education, employment status, perception of own living conditions, perception of the country's economic condition, interest in politics, regular TV use dummy, regular radio use dummy, and regular newspaper use dummy. District controls include nighttime light, local Internet incidence, 2G network coverage, 3G network coverage, and the log of the distance from district's centroid to the closest Internet infrastructure. Country controls include log GDP per capita, Polity2 index, unemployment rate, and the number of SMCs. Standard errors clustered at the district-year level are reported in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' elaboration on Afrobarometer and Polity5 datasets based on the baseline sample of 99,938 individuals in 35 countries between 2011 and 2018.

Table A.10
Internet use, convergence toward experts' ratings, and coherent responses to preference for democracy by censorship status.

	Convergence			Coherence
	(1) Polity2	(2) RoW	(3) Corruption	(4) Coherence
Internet use	−0.0231***	−0.0228***	−0.0320***	0.0045
	(0.005)	(0.005)	(0.006)	(0.005)
FOTP score	0.0010	−0.0013	0.0065***	0.0037***
	(0.001)	(0.001)	(0.001)	(0.001)
Internet use × FOTP score	0.0005***	0.0005***	0.0006***	0.0000
	(0.000)	(0.000)	(0.000)	(0.000)
Individual controls	✓	✓	✓	✓
District controls	✓	✓	✓	✓
Country controls	✓	✓	✓	✓
Observations	99,938	99,938	88,283	89,050
Adjusted R^2	0.120	0.123	0.122	0.076
Mean dependent variable	0.586	0.578	0.580	0.662
Mean Internet use	0.895	0.895	0.928	0.902

Notes: This table reports the OLS results of the effect of Internet use to get news on individuals' convergence towards experts' ratings by press freedom score. The dependent variables in columns (1), (2), and (3) are dummy variables equal to 1 if the individual convergences towards experts' ratings and 0 otherwise. In column (1), we rely on the Polity2 index from the Polity5 project dataset to compute the convergence dummy. In column (2), we rely on the RoW index from the V-Dem dataset. In column (3), we rely on the legislature corrupt activities index from the V-Dem dataset. The dependent variable in column (4) is a dummy variable equal to 1 if the individual responds coherently to strict preference variable set of questions responding that "democracy is preferable to any other kind of government" and "rejecting all three authoritarian alternatives" or responding that "in some circumstances, a non-democratic government can be preferable" and "does not reject all three authoritarian alternatives", and 0 otherwise. Individual controls include age, age squared, gender, urban dummy, education, employment status, perception of own living conditions, perception of the country's economic condition, interest in politics, regular TV use dummy, regular radio use dummy, and regular newspaper use dummy. District controls include nighttime light, local Internet incidence, 2G network coverage, 3G network coverage, and the log of the distance from district's centroid to the closest Internet infrastructure. Country controls include log GDP per capita, Polity2 index, unemployment rate, and the number of SMCs. Standard errors clustered at the district-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' elaboration on Afrobarometer data based on the baseline sample of 99,938 individuals in 35 countries between 2011 and 2018.

Table A.11
Internet use, convergence toward experts' ratings, and coherent responses to preference for democracy by education group.

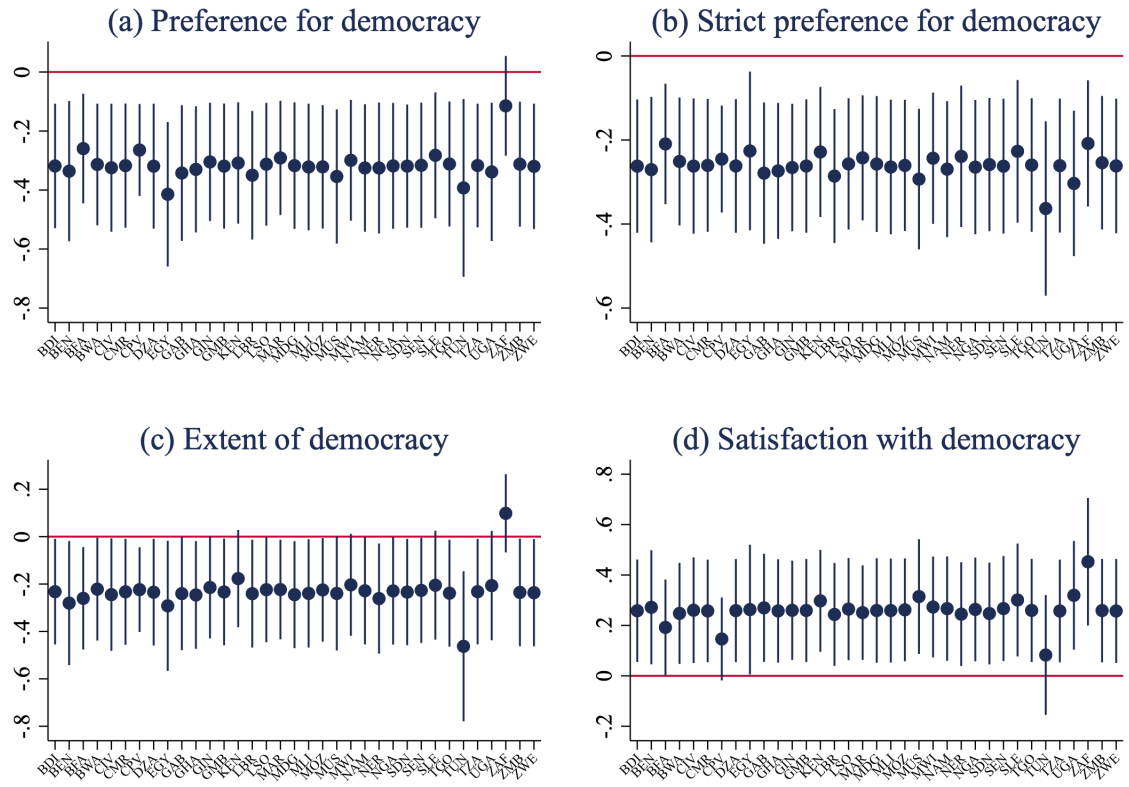
	Convergence			Coherence
	(1) Polity2	(2) RoW	(3) Corruption	(4) Coherence
Internet use	−0.0014	0.0024	−0.0037	0.0084**
	(0.003)	(0.003)	(0.004)	(0.004)
(Post-)secondary education	0.0125**	0.0085*	−0.0077	0.0555***
	(0.005)	(0.005)	(0.005)	(0.005)
Internet use × (Post-)secondary education	0.0049	−0.0010	0.0035	0.0003
	(0.004)	(0.004)	(0.004)	(0.004)
Individual controls	✓	✓	✓	✓
District controls	✓	✓	✓	✓
Country controls	✓	✓	✓	✓
Observations	99,938	99,938	88,283	89,050
Adjusted R^2	0.120	0.123	0.120	0.076
Mean dependent variable	0.586	0.578	0.580	0.662
Mean Internet use	0.895	0.895	0.928	0.902

Notes: This table reports the OLS results of the effect of Internet use to get news on individuals' convergence towards experts' ratings by education group. The dependent variables in columns (1), (2), and (3) are dummy variables equal to 1 if the individual convergences towards experts' ratings and 0 otherwise. In column (1), we rely on the Polity2 index from the Polity5 project dataset to compute the convergence dummy. In column (2), we rely on the RoW index from the V-Dem dataset. In column (3), we rely on the legislature corrupt activities index from the V-Dem dataset. The dependent variable in column (4) is a dummy variable equal to 1 if the individual responds coherently to strict preference variable set of questions responding that "democracy is preferable to any other kind of government" and "rejecting all three authoritarian alternatives" or responding that "in some circumstances, a non-democratic government can be preferable" and "does not reject all three authoritarian alternatives", and 0 otherwise. Individual controls include age, age squared, gender, urban dummy, education, employment status, perception of own living conditions, perception of the country's economic condition, interest in politics, regular TV use dummy, regular radio use dummy, and regular newspaper use dummy. District controls include nighttime light, local Internet incidence, 2G network coverage, 3G network coverage, and the log of the distance from district's centroid to the closest Internet infrastructure. Country controls include log GDP per capita, Polity2 index, unemployment rate, and the number of SMCs. Standard errors clustered at the district-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' elaboration on Afrobarometer data based on the baseline sample of 99,938 individuals in 35 countries between 2011 and 2018.

A.3 Robustness checks

Fig. A.3. Coefficients plots: dropping countries one by one.



Notes: These subfigures illustrate the Internet use coefficients obtained from the second stage IV regressions of the effect of Internet use on preference for democracy in (a), strict preference for democracy in (b), the extent of democracy in (c), and satisfaction with democracy in (d), dropping countries one by one. The country dropped is indicated on the x-axis.
Source: Authors' elaboration on Afrobarometer data.

Table A.12
Two instruments.

	(1) Preference	(2) Strict Preference	(3) Extent	(4) Satisfaction
First stage regression: Internet use				
SMC number × fixed 2G share	0.237***	0.237***	0.237***	0.237***
	(0.047)	(0.047)	(0.047)	(0.047)
SMC number × fixed 3G share	0.511***	0.511***	0.511***	0.511***
	(0.084)	(0.084)	(0.084)	(0.084)
Second stage regression:				
Internet use	−0.315***	−0.250***	−0.176**	0.076
	(0.079)	(0.075)	(0.087)	(0.133)
Individual controls	✓	✓	✓	✓
District controls	✓	✓	✓	✓
Country controls	✓	✓	✓	✓
District FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Observations	99,938	99,938	99,938	99,938
Mean dependent variable	0.773	0.525	0.566	0.503
Mean Internet use	0.895	0.895	0.895	0.895
KP Wald F-stat	42.488	42.488	42.488	42.488
KP LM P-val	0.000	0.000	0.000	0.000
Hansen J P-val	0.967	0.860	0.360	0.001

Notes: This table reports the first and second stages of IV results of the effect of Internet use to get news on individuals' perception of democracy using two IVs. The dependent variable in column (1) is a dummy variable equal to 1 if the individual says "democracy is preferable to any other kind of government" and 0 otherwise. In column (2), it is a dummy variable equal to 1 if the individual says "democracy is preferable to any other kind of government" and "rejects all three authoritarian alternatives" and 0 otherwise. In column (3), it is a dummy variable equal to 1 if the individual perceives his or her country as "a full democracy" or "a democracy with minor problems" and 0 otherwise. In column (4), it is a dummy variable equal to 1 if the individual is "very" or "fairly" satisfied with how democracy works in his or her country and 0 otherwise. SMC number × fixed 2G (3G) share is the interaction between the number of SMCs and the share of the individual's district of residence covered with 2G (3G) network in the first Afrobarometer survey ave used in our analysis. Internet use is an ordered categorical variable equal to 0 if the individual never uses Internet, 1 if he or she uses it less than once a month, 2 if he or she uses it a few times a month, 3 if he or she uses it a few times a week, and 4 if he or she uses it every day. Individual controls include age, age squared, gender, urban dummy, education, employment status, perception of own living conditions, perception of the country's economic condition, interest in politics, regular TV use dummy, regular radio use dummy, and regular newspaper use dummy. District controls include nighttime light, local Internet incidence, 2G network coverage, 3G network coverage, and the log of the distance from district's centroid to the closest Internet infrastructure. Country controls include log GDP per capita, Polity2 index, unemployment rate, and the number of SMCs. Standard errors clustered at the district-year level are reported in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' elaboration on Afrobarometer data based on the baseline sample of 99,938 individuals in 35 countries between 2011 and 2018.

Table A.13
Two instruments: effective F-stat.

	(1) Preference	(2) Strict Preference	(3) Extent	(4) Satisfaction
Effective F-stat	30.089	30.089	30.089	30.089
Confidence level alpha: 5%				
tau=5%	23.494	23.110	23.936	24.266
tau=10%	14.655	14.442	14.900	15.083
tau=15%	9.701	9.580	9.840	9.944
tau=20%	7.848	7.760	7.949	8.025

Notes: This table reports the weak instrument test of Olea and Pflueger (2013). It tests the null hypothesis of weak instruments. The test rejects the null hypothesis when the effective F statistic exceeds a critical value, which depends on the estimator, the significance level, and the desired weak instrument threshold tau.
Source: Authors' elaboration on Afrobarometer data based on the baseline sample of 99,938 individuals in 35 countries between 2011 and 2018.

Table A.14
Two instruments: conditional likelihood ratio (CLR) test.

	(1) Preference	(2) Strict Preference	(3) Extent	(4) Satisfaction
CLR P-val	0.000	0.005	0.106	0.983
CLR Confidence set	[-0.484,-0.169]	[-0.393,-0.094]	[-0.342, 0.030]	[-0.346, 0.289]
J P-val	0.545	0.383	0.292	0.001

Note: This table reports the conditional likelihood ratio (CLR) weak-instrument-robust test P-value, confidence intervals of the coefficient on the endogenous variable (Internet use) in the IV estimations, and the J overidentification test P-value.
Source: Authors' elaboration on Afrobarometer data based on the baseline sample of 99,938 individuals in 35 countries between 2011 and 2018.

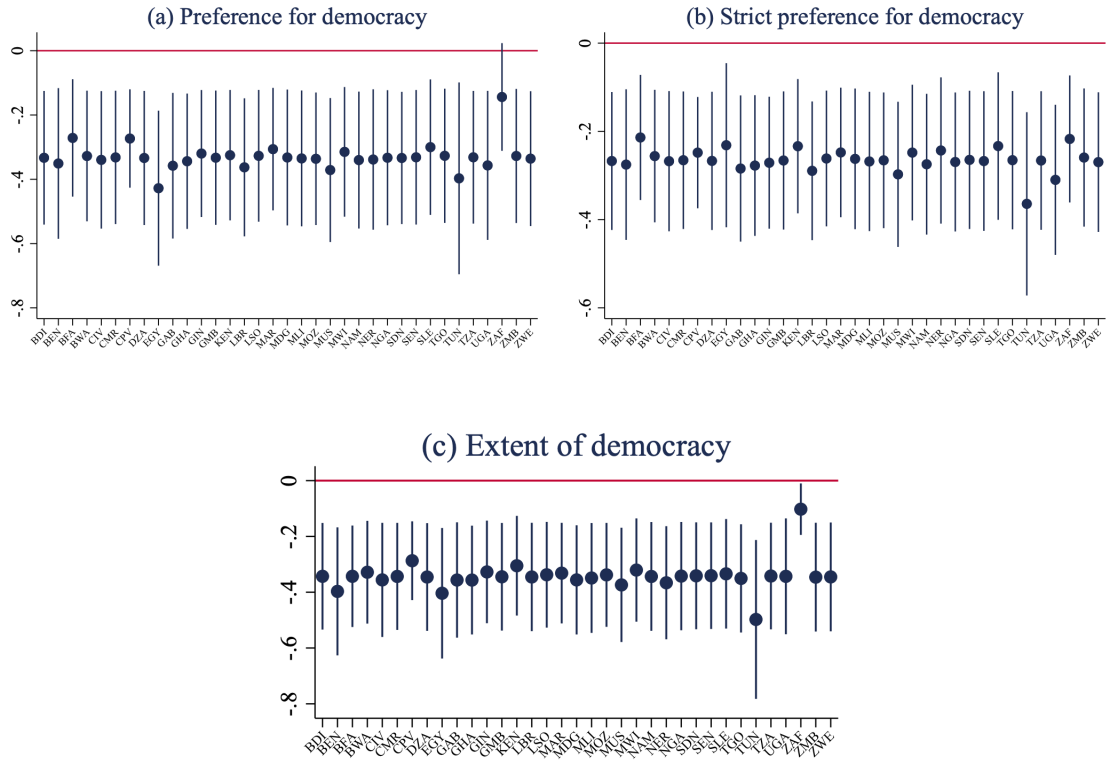
Table A.15
Controlling for satisfaction with democracy.

	(1) Preference	(2) Strict Preference	(3) Extent
First stage regression: Internet use			
SMC number × fixed 3G share	0.603***	0.603***	0.603***
	(0.083)	(0.083)	(0.083)
Second stage regression:			
Internet use	−0.332***	−0.265***	−0.343***
	(0.106)	(0.079)	(0.098)
Individual controls	✓	✓	✓
District controls	✓	✓	✓
Country controls	✓	✓	✓
District FE	✓	✓	✓
Year FE	✓	✓	✓
Observations	99,938	99,938	99,938
Mean dependent variable	0.773	0.525	0.566
Mean Internet use	0.895	0.895	0.895
KP Wald F-stat	52.406	52.406	52.406
KP LM P-val	0.000	0.000	0.000

Notes: This table reports the first and second stages of IV results of the effect of Internet use to get news on individuals' perception of democracy. The dependent variable in column (1) is a dummy variable equal to 1 if the individual says "democracy is preferable to any other kind of government" and 0 otherwise. In column (2), it is a dummy variable equal to 1 if the individual says "democracy is preferable to any other kind of government" and "rejects all three authoritarian alternatives" and 0 otherwise. In column (3), it is a dummy variable equal to 1 if the individual perceives his or her country as "a full democracy" or "a democracy with minor problems" and 0 otherwise. SMC number × fixed 3G share is the interaction between the number of SMCs and the share of the individual's district of residence covered with 3G network in the first Afrobarometer survey wave used in our analysis. Internet use is an ordered categorical variable equal to 0 if the individual never uses Internet, 1 if he or she uses it less than once a month, 2 if he or she uses it a few times a month, 3 if he or she uses it a few times a week, and 4 if he or she uses it every day. Individual controls include age, age squared, gender, urban dummy, education, employment status, perception of own living conditions, perception of the country's economic condition, interest in politics, regular TV use dummy, regular radio use dummy, regular newspaper use dummy, and satisfaction with how democracy works in the country. District controls include nighttime light, local Internet incidence, 2G network coverage, 3G network coverage, and the log of the distance from district's centroid to the closest Internet infrastructure. Country controls include log GDP per capita, Polity2 index, unemployment rate, and the number of SMCs. Standard errors clustered at the district-year level are reported in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' elaboration on Afrobarometer data based on the baseline sample of 99,938 individuals in 35 countries between 2011 and 2018.

Fig. A.4. Coefficients plots: dropping countries one by one and controlling for satisfaction with democracy.



Notes: These subfigures illustrate the Internet use coefficients obtained from the second stage IV regressions of the effect of Internet use on preference for democracy in (a), strict preference for democracy in (b), and the extent of democracy in (c), dropping countries one by one. The country dropped is indicated on the x-axis.
Source: Authors' elaboration on Afrobarometer data.

Table A.16
Excluding districts with less than 30 observations.

	(1) Preference	(2) Strict Preference	(3) Extent	(4) Satisfaction
First stage regression: Internet use				
SMC number × fixed 3G share	0.344***	0.344***	0.344***	0.344***
	(0.126)	(0.126)	(0.126)	(0.126)
Second stage regression:				
Internet use	−0.977**	−0.557**	−1.072**	−0.097
	(0.414)	(0.261)	(0.444)	(0.252)
Individual controls	✓	✓	✓	✓
District controls	✓	✓	✓	✓
Country controls	✓	✓	✓	✓
District FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Observations	55,932	55,932	55,932	55,932
Mean dependent variable	0.782	0.550	0.561	0.496
Mean Internet use	1.073	1.073	1.073	1.073
KP Wald F-stat	7.449	7.449	7.449	7.449
KP LM P-val	0.041	0.041	0.041	0.041

Notes: This table reports the first and second stages of IV results of the effect of Internet use to get news on individuals' perception of democracy excluding districts with less than 30 observations. The dependent variable in column (1) is a dummy variable equal to 1 if the individual says "democracy is preferable to any other kind of government" and 0 otherwise. In column (2), it is a dummy variable equal to 1 if the individual says "democracy is preferable to any other kind of government" and "rejects all three authoritarian alternatives" and 0 otherwise. In column (3), it is a dummy variable equal to 1 if the individual perceives his or her country as "a full democracy" or "a democracy with minor problems" and 0 otherwise. In column (4), it is a dummy variable equal to 1 if the individual is "very" or "fairly" satisfied with how democracy works in his or her country and 0 otherwise. SMC number × 3G share is the interaction between the number of SMCs and the share of the individual's district of residence covered with 3G network in the first Afrobarometer survey wave used in our analysis. Internet use is an ordered categorical variable equal to 0 if the individual never uses Internet, 1 if he or she uses it less than once a month, 2 if he or she uses it a few times a month, 3 if he or she uses it a few times a week, and 4 if he or she uses it every day. Individual controls include age, age squared, gender, urban dummy, education, employment status, perception of own living conditions, perception of the country's economic condition, interest in politics, regular TV use dummy, regular radio use dummy, and regular newspaper use dummy. District controls include nighttime light, local Internet incidence, 2G network coverage, 3G network coverage, and the log of the distance from district's centroid to the closest Internet infrastructure. Country controls include log GDP per capita, Polity2 index, unemployment rate, and the number of SMCs. Standard errors clustered at the district-year level are reported in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' elaboration on Afrobarometer data based on 55,932 individuals in 32 countries between 2011 and 2018.

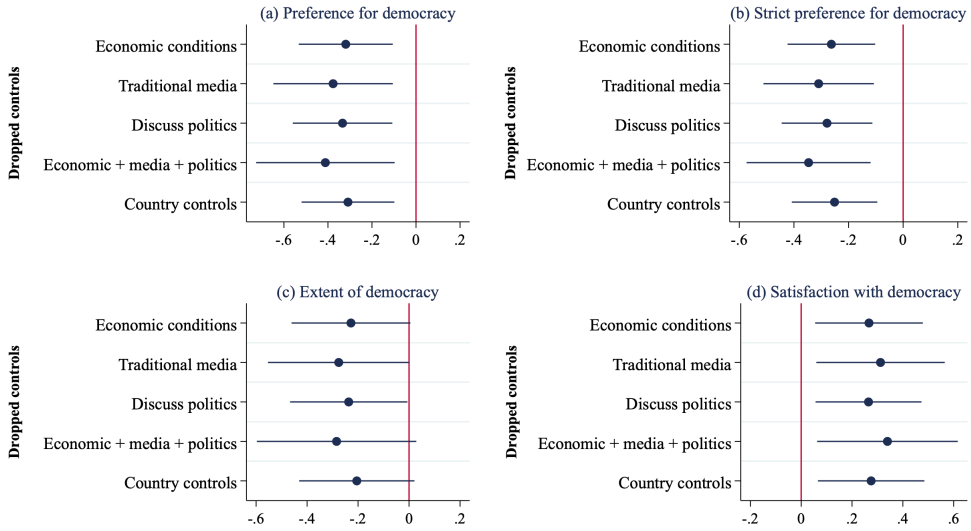
Table A.17
Excluding districts with no Internet users.

	(1) Preference	(2) Strict Preference	(3) Extent	(4) Satisfaction
First stage regression: Internet use				
SMC number × fixed 3G share	0.576***	0.576***	0.576***	0.576***
	(0.087)	(0.087)	(0.087)	(0.087)
Second stage regression:				
Internet use	−0.326***	−0.248***	−0.256**	0.263**
	(0.113)	(0.085)	(0.119)	(0.107)
Individual controls	✓	✓	✓	✓
District controls	✓	✓	✓	✓
Country controls	✓	✓	✓	✓
District FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Observations	90,173	90,173	90,173	90,173
Mean dependent variable	0.773	0.529	0.559	0.498
Mean Internet use	0.992	0.992	0.992	0.992
KP Wald F-stat	43.779	43.779	43.779	43.779
KP LM P-val	0.000	0.000	0.000	0.000

Notes: This table reports the first and second stages of IV results of the effect of Internet use to get news on individuals' perception of democracy excluding districts with no Internet users. The dependent variable in column (1) is a dummy variable equal to 1 if the individual says "democracy is preferable to any other kind of government" and 0 otherwise. In column (2), it is a dummy variable equal to 1 if the individual says "democracy is preferable to any other kind of government" and "rejects all three authoritarian alternatives" and 0 otherwise. In column (3), it is a dummy variable equal to 1 if the individual perceives his or her country as "a full democracy" or "a democracy with minor problems" and 0 otherwise. In column (4), it is a dummy variable equal to 1 if the individual is "very" or "fairly" satisfied with how democracy works in his or her country and 0 otherwise. SMC number × 3G share is the interaction between the number of SMCs and the share of the individual's district of residence covered with 3G network in the first Afrobarometer survey wave used in our analysis. Internet use is an ordered categorical variable equal to 0 if the individual never uses Internet, 1 if he or she uses it less than once a month, 2 if he or she uses it a few times a month, 3 if he or she uses it a few times a week, and 4 if he or she uses it every day. Individual controls include age, age squared, gender, urban dummy, education, employment status, perception of own living conditions, perception of the country's economic condition, interest in politics, regular TV use dummy, regular radio use dummy, and regular newspaper use dummy. District controls include nighttime light, local Internet incidence, 2G network coverage, 3G network coverage, and the log of the distance from district's centroid to the closest Internet infrastructure. Country controls include log GDP per capita, Polity2 index, unemployment rate, and the number of SMCs. Standard errors clustered at the district-year level are reported in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' elaboration on Afrobarometer data based on 90,173 individuals in 35 countries between 2011 and 2018.

Fig. A.5. Coefficients plots: dropping control variables.



Notes: These subfigures illustrate the Internet use coefficients obtained from the second stage IV regressions of the effect of Internet use on preference for democracy in (a), strict preference for democracy in (b), the extent of democracy in (c), and satisfaction with democracy in (d), dropping potential bad controls. Dropped controls are indicated on the y-axis.
Source: Authors' elaboration on Afrobarometer data based on the baseline sample of 99,938 individuals in 35 countries between 2011 and 2018.

Table A.18
Falsification test.

	(1) Preference	(2) Strict Preference	(3) Extent	(4) Satisfaction
Panel A: Normal distribution				
First stage regression: Internet use				
Recentered IV	0.594***	0.594***	0.594***	0.594***
	(0.083)	(0.083)	(0.083)	(0.083)
KP Wald F-stat	51.047	51.047	51.047	51.047
KP LM P-val	0.000	0.000	0.000	0.000
Second stage regression:				
Internet use	−0.317***	−0.260***	−0.233**	0.258**
	(0.107)	(0.081)	(0.114)	(0.104)
Panel B: Poisson distribution				
First stage regression: Internet use				
Recentered IV	0.595***	0.595***	0.595***	0.595***
	(0.083)	(0.083)	(0.083)	(0.083)
KP Wald F-stat	51.183	51.183	51.183	51.183
KP LM P-val	0.000	0.000	0.000	0.000
Second stage regression:				
Internet use	−0.316***	−0.260***	−0.232**	0.257**
	(0.107)	(0.081)	(0.114)	(0.104)
Panel C: Uniform distribution				
First stage regression: Internet use				
Recentered IV	0.594***	0.594***	0.594***	0.594***
	(0.083)	(0.083)	(0.083)	(0.083)
KP Wald F-stat	50.991	50.991	50.991	50.991
KP LM P-val	0.000	0.000	0.000	0.000
Second stage regression:				
Internet use	−0.318***	−0.261***	−0.233**	0.258**
	(0.108)	(0.081)	(0.114)	(0.104)
Individual controls	✓	✓	✓	✓
District controls	✓	✓	✓	✓
Country controls	✓	✓	✓	✓
District FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Observations	99,938	99,938	99,938	99,938
Mean dependent variable	0.773	0.525	0.566	0.503
Mean Internet use	0.895	0.895	0.895	0.895

Notes: This table reports the first and second stages of IV results of the effect of Internet use to get news on individuals' perception of democracy using recentered IV as instrument. The dependent variable in column (1) is a dummy variable equal to 1 if the individual says "democracy is preferable to any other kind of government" and 0 otherwise. In column (2), it is a dummy variable equal to 1 if the individual says "democracy is preferable to any other kind of government" and "rejects all three authoritarian alternatives" and 0 otherwise. In column (3), it is a dummy variable equal to 1 if the individual perceives his or her country as "a full democracy" or "a democracy with minor problems" and 0 otherwise. In column (4), it is a dummy variable equal to 1 if the individual is "very" or "fairly" satisfied with how democracy works in his or her country and 0 otherwise. Recentered IV is obtained by retrieving the residuals of the regression of our initial IV on the average of randomly generated normal (Panel A), poisson (Panel B), and uniform (Panel C) variates of the number of SMCs weighted by the fixed 3G network share for each observation. Internet use is an ordered categorical variable equal to 0 if the individual never uses Internet, 1 if he or she uses it less than once a month, 2 if he or she uses it a few times a month, 3 if he or she uses it a few times a week, and 4 if he or she uses it every day. Individual controls include age, age squared, gender, urban dummy, education, employment status, perception of own living conditions, perception of the country's economic condition, interest in politics, regular TV use dummy, regular radio use dummy, and regular newspaper use dummy. District controls include nighttime light, local Internet incidence, 2G network coverage, 3G network coverage, and the log of the distance from district's centroid to the closest Internet infrastructure. Country controls include log GDP per capita, Polity2 index, unemployment rate, and the number of SMCs. Standard errors clustered at the district-year level are reported in parentheses: *** p<0.01, ** p<0.05, * p<0.1.
Source: Authors' elaboration on Afrobarometer data based on the baseline sample of 99,938 individuals in 35 countries between 2011 and 2018.

Chapter 2

Immigration and Cultural Heterogeneity: Evidence from Two Decades in Europe

This chapter is joint work with Riccardo Turati (Universitat Autonoma de Barcelona, Spain) and Jérôme Valette (CEPII, CES, Université Paris 1 Panthéon-Sorbonne, France).[1]

"[...] the primary threat to our identity is the massive immigration that we have been experiencing for several decades. Previously, the saying "When in Rome, do as the Romans do" rightfully directed immigrants to adhere to French customs. Today, newcomers have become entitled individuals permitted to promote the culture of their origins."

Éric Zemmour, 2022.[2]

2.1 Introduction

Immigration has recently been put at the center of the cultural debate in Europe, and the recent rise of populism has been accompanied by a growing fear that immigration could represent a threat to the Western cultural model and a potential challenge to the cultural homogeneity of host nations (Collier, 2013; Guriev and Papaioannou, 2022; Alesina and Tabellini, 2022). These beliefs rely on the traditional view that birthplace is one of the strongest markers of one's identity and values. Using data from two billion Facebook users across 225 countries, Obradovich et al. (2022) underline the importance of national borders in shaping culture.

Prior research suggests that immigration may affect the cultural heterogeneity of hosting countries through two main channels. A first direct compositional effect from the addition of individuals with plausibly different or similar cultural values to the resident population. In Europe, the growing inflow of foreign-born from increasingly geographically, economically, and culturally distant countries has raised specific concerns, as it has conceivably

[1] It has benefited from valuable suggestions and comments from Simone Bertoli, Flore Gubert, and Hillel Rapoport. We also wish to thank seminar participants at the 13th annual conference on Immigration in OECD countries, at the 2nd Istanbul Applied Micro Workshop, King's College London and Özyegin University, at the University of Salerno, and the DIMIG workshop in Paris for helpful discussions.

[2] Translated by the authors from the original quote in French. Éric Zemmour is a French writer, political commentator, and media personality. He has been running for the French presidential elections in 2022.

brought different cultural values, norms, or different ways of thinking (Docquier et al., 2020b). The magnitude of this direct effect ultimately is expected to depend on immigrants' initial patterns of cultural self-selection at origin (Docquier et al., 2020a), on how their values persist through intergenerational transmission of cultural traits to second and subsequent generations (Bisin and Verdier, 2001; Desmet et al., 2017; Desmet and Wacziarg, 2021), and the rate of cultural assimilation of the foreign-born and their descendants to the native population (Algan et al., 2012; Abramitzky et al., 2014; Gonnot and lo Polito, 2021; Fouka et al., 2022; Abramitzky and Boustan, 2022; Gonnot and lo Polito, 2023). Immigrants may also affect cultural heterogeneity in host societies through a second indirect channel based on their interaction with the native population (Bazzi et al., 2019; Giavazzi et al., 2019; Giuliano and Tabellini, 2021). Immigrants themselves can indeed disseminate their culture to the natives (Fisman and Miguel, 2007; Schmitz and Weinhardt, 2019; Tabellini, 2020; Giuliano and Nunn, 2021; Miho et al., 2023), and may slow down or foster cultural convergence in a society where conformism is a positive function of the relative share of the population that belongs to the majority (Desmet and Wacziarg, 2021). Similarly, immigration may also increase the salience of one's birthplace within society, thereby increasing the pressure to conform if natives perceive benefits in aligning with the predominant values of their group (Bazzi et al., 2019; Desmet and Wacziarg, 2021). Finally, research in social sciences has also shown that immigration may induce strong attitudinal changes within the native population (see Dustmann and Preston, 2007b; Dustmann et al., 2018; Edo et al., 2019b; Steinmayr, 2021; Alesina et al., 2022; Alesina and Tabellini, 2022; Keita et al., 2023, among others) that can extend to cultural values and norms.

This paper investigates the extent to which immigration impacted overall cultural heterogeneity in Europe over the past two decades through these two channels. We take advantage of data from the European Social Surveys (2004 to 2018) to measure the evolution of cultural heterogeneity in Europe along several cultural dimensions. Specifically, we build on Desmet and Wacziarg (2021) to compute an overall index of cultural heterogeneity at the regional level (NUTS 2), which captures the likelihood that two randomly chosen individuals from the entire resident population of a given country hold a different variant of a randomly chosen memetic trait.[3] Following Alesina et al. (2017) and Desmet and Wacziarg (2021), we adopt a broad definition of culture, which encompasses several specific dimensions such as religiosity, sexual morality, the role of the state, cultural capital, political engagement, trust in institutions, attitudes toward immigration,

[3] Memetic traits and cultural values/traits are used interchangeably in the rest of the paper.

and openness. By design, our definition of culture, well-known for its intricate and multifaceted nature, relies also on one of the definitions of culture of the seminal work by Kroeber and Kluckhohn (1952), where culture can be described and proxied by a series of specific values and traits. Then, we leverage immigration data from the European Labor Force Survey, which provides immigration stocks at the regional level, to estimate the impact of immigration on overall cultural heterogeneity.

The benchmark empirical specification includes wave and regional fixed effects as well as a full vector of controls at the regional level.[4] As far as the self-selection and the sorting of immigrants are concerned, we rely on 2SLS estimates with a modified shift-share instrument to predict exogenous immigration stocks by origin based on the initial spatial sorting of immigrants and the growth of their diasporas at the national level over time. To enhance the validity of our identifying assumption based on the exogeneity of the aggregate immigration flows (Borusyak et al., 2022), the stock of immigrants from each origin is predicted using a zero-stage gravity equation that includes exogenous push factors such as conflicts and natural disasters in migrants' origin countries.

The findings indicate that an increase in the share of immigrants is significantly associated with a reduction in overall cultural fractionalization among the resident population. In terms of magnitude, a one percentage point increase in the share of immigrants is associated with a 0.16 percentage point decline in overall cultural heterogeneity. This comprehensive effect is the culmination of the two counterbalancing mechanisms mentioned above, which are under scrutiny in our study through a within- and between-group decomposition of the index, following the approach of Desmet and Wacziarg (2021). On one hand, the inflow of new immigrants tends to increase cultural fractionalization by introducing novel cultural values and norms to the host countries. Focusing on specific groups of immigrants, we find this compositional effect, or direct effect, to be steered by low-skilled immigrants originating from non-European nations, which exhibit large cultural differences with Europe. It vanishes 10 years after the arrival of immigrants, suggesting fast cultural assimilation of first-generation immigrants over time. Additional

[4] It is important to highlight that the use of a panel specification spanning nearly two decades and fixed effects significantly improves the identification. However, this approach does entail a trade-off. While it allows us to deliver insights into the immediate effects of immigration on culture, it also confines our analysis to the short-term horizon. Further research will be therefore needed to comprehensively assess the long-run implications of immigration on the cultural diversity of Europe. In this way, it is worth recognizing that this paper brings a little contribution to the literature on the deep-rooted determinants of culture (See Ashraf and Galor, 2013; Alesina and Giuliano, 2015; Giuliano and Nunn, 2021; Bazzi et al., 2020, among others).

suggestive cross-sectional estimates at the individual level support this interpretation of our results. We find immigrants to systematically differ from natives across almost all dimensions such as gender, sexual morality, the role of the state, and political engagement for instance. Lower educational attainment, distant origins, and shorter duration of stay exacerbate these cultural differences. On the other hand, we find evidence of an indirect effect within the native population, which can be interpreted as natives' response to the presence of immigrants in their region. A rise in the share of immigrants contributes to a decrease in cultural heterogeneity. Focusing on specific immigration inflows, we find this effect to be mainly driven by high-skilled immigrants from other European countries, which suggests contact as a potential channel of cultural diffusion (Allport et al., 1954; Miho et al., 2023). Suggestive cross-sectional evidence at the individual level reports that natives are more likely to hold liberal values, undermining a potential backlash interpretation of the results. Overall, and since the native population overwhelmingly dominates the immigrant population throughout Europe, it is the reduction of cultural diversity among natives that drives the overall negative estimated effect in the benchmark model.

This paper contributes to the fast-growing, and above-cited literature, on how immigration affects culture in recipient countries. The closest paper to our analysis is Rapoport et al. (2020), which develops a theoretical model of migration-based cultural change and studies how immigration drives cultural proximity between home and host countries. Using data from the World Value Surveys (WVS) between 1981 and 2014, they found that migration is significantly associated with higher cultural proximity between pairs of countries. They provide evidence that cultural convergence comes from the adoption of cultural norms from the diasporas (social remittances) in migrants' origin countries, while they find little support for the dissemination of the culture of immigrants after their arrival. Conversely to Rapoport et al. (2020), we are not interested, per se, in cultural convergence at the global level but only in the impact of immigration on recipient countries' cultural heterogeneity. We depart therefore from most of the existing research by aligning our paper with burgeoning research focusing on the evolution of the cultural divide within modern societies, namely whether identity cleavages in society are good predictors of individuals' attitudes and values toward various cultural dimensions. Indeed, as underlined by Rapoport et al. (2020), average effects of immigration on the evolution of culture are "silent" on their effects on within-country cultural heterogeneity.[5] Hence, our study

[5] Particularly, the question on how individual-level shifts contribute to cultural transformations at the aggregate level is complex, especially when averages conceal noteworthy changes pulling in opposing directions (Schneider-Strawczynski and Valette, 2023), or when there exists pronounced heterogeneity in

complements theirs by looking into the effect on within-destination cultural dynamics.

This paper builds therefore on the seminal work by Desmet et al. (2017) who provide both theoretical insights and empirical guidance for exploring the transformation of the cultural divide within countries. Drawing from survey data (WVS) encompassing 76 countries worldwide, their findings reveal that while ethnic identity wields considerable influence over cultural values, the within-group variability is generally more significant than that between-group variability in explaining overall cultural diversity. Similarly, by using data from the General Social Survey (GSS) in the United States spanning 1972 to 2018, Desmet and Wacziarg (2021) confirm that within-group variations largely overstate between-group variations. Importantly, their research underscores that while the overall cultural heterogeneity in the United States did not experience substantial growth during their period of analysis, a nuanced picture emerges when focusing on specific cultural dimensions and identity cleavages. Still, these prior studies often disregard immigration as a relevant identity cleavage to explain the evolution of cultural heterogeneity.[6] Interestingly the most divisive questions in Desmet and Wacziarg (2021) are found to relate to fundamentalism, racial cleavages, government spending to the black population, and police violence, which all report strong connections with traditional concerns toward immigration. Our contribution to this body of literature lies therefore in introducing birthplace as a relevant cleavage and providing novel evidence across European regions. This distinctive approach facilitates the connection of two parallel yet interconnected streams of the literature, namely the investigation of the cultural divide within Western societies, and the impact of immigration on culture within host countries. In addition, while most of this research puts a strong focus on the United States (Bertrand and Kamenica, 2023), our analysis studies the evolution of cultural heterogeneity in Europe. A noticeable exception is Alesina et al. (2017) which reports the evolution of culture in Europe across four waves of the European Values Survey (EVS), between 1980 and 2008. By focusing on a wider set of countries over the last two decades, which includes relevant recent shocks like the 2008 financial crisis and the 2015 refugee crisis, we provide novel and up-to-date evidence on the evolution of cultural heterogeneity across Europe in the light of immigration.

The remainder of this paper is organized as follows. Section 2.2 first describes measures of cultural heterogeneity and derives testable hypotheses from the literature. Section 2.3

these responses.

[6] It is worth noting that these studies still employ race or ethnicity to segment the population into groups. Nevertheless, birthplace stands as another relevant identity cleavage to consider, as it does not entirely overlap with the aforementioned categories.

reports the data that we collected and used in our empirical analysis and Section 2.4 discusses preliminary evidence on the relevance of birthplace to explain cultural fractionalization. Section 2.5 describes our empirical strategy, while Section 2.6 reports the main results, robustness checks and explores potential mechanisms. Finally, Section 2.7 and Section 2.8 offer our heterogeneity analysis and some concluding remarks.

2.2 Theoretical background and measurement

This section first outlines in Section 2.2.1 the definitions of our measures of cultural heterogeneity following the seminal paper by Desmet and Wacziarg (2021). Then, we derive from the literature theoretical predictions regarding the impact of immigration on cultural fractionalization in Section 2.2.2.

2.2.1 Definitions

We follow Desmet and Wacziarg (2021) to construct time-varying measures of cultural heterogeneity across European regions. First, we compute an index of overall heterogeneity in the resident population (CF) which represents the average fractionalization of the whole population across various cultural values named memetic traits m= 1, ... M. It is defined as:

$$CF = \frac{1}{M}\sum_{m=1}^{M} CF^m = 1 - \frac{1}{M}\sum_{m=1}^{M}\sum_{i_m=1}^{I_m}\left(s^{i_m}\right)^2 \tag{2.1}$$

where s^{i_m} is the share of the resident population that holds the i_m variant of the meme m. Thus, CF captures the likelihood that two randomly chosen individuals from the entire resident population hold a different variant of a randomly chosen memetic trait.

To derive a measure of cultural divide that does not focus only on the overall population but highlights the relevance of identity cleavages, we follow Desmet and Wacziarg (2021) by decomposing our index of overall heterogeneity between its within and between components. Conversely to Desmet et al. (2017) or Desmet and Wacziarg (2021), we consider only two groups in the population, namely natives and foreign-born individuals.[7] We

[7] We do exclude from the main sample second-generation immigrants, although we will test the robustness of our results to the second-generation immigrant population as a robustness check. Desmet and Wacziarg (2021) study alternative identity cleavages such as age, education, income, gender, political affiliation, urban status, or religion for instance. It is worth noting that this paper does not specifically focus on ethnicity, or race as a potential driver of cultural heterogeneity while it is likely that these definitions partly overlap with immigration.

measure the average within-origin heterogeneity for meme m such as:

$$CF^{W_m} = \sum_g share_g CF_g^m = \sum_g share_g \left(1 - \sum_{i_m=1}^{I_m} (s_g^{i_m})^2\right) \tag{2.2}$$

where $share_g$ is the share of individuals from group $g \in (\text{natives}; \text{foreign-born})$ in the overall population and CF_g^m the within-group g heterogeneity for meme m. It is important to notice that in our setting, given the unbalanced distribution of natives and immigrants in the resident population, changes within the native population accounts for the majority of the variation of the CF^{W_m} index. As for the previous index, we can average this index over all memes m to obtain the overall within-group heterogeneity:

$$CF^W = \frac{1}{M} \sum_{m=1}^{M} CF^{W_m} \tag{2.3}$$

The between-group component F_{ST} corresponds to a measure of memetic fixation (Wright's fixation index), namely the share of the total population's cultural fractionalization that is not due to within-group fractionalization:

$$F_{ST} = \frac{CF - CF^W}{CF} \tag{2.4}$$

It is worth noting that F_{ST} equals one when there is no within-group heterogeneity, hence there is a perfect overlap between memes and groups. In other words, being a native or a foreign-born would be a perfect predictor of norms and values. On the contrary, F_{ST} equals zero suggests that the place of birth provides no information on cultural norms.

2.2.2 Theoretical background

Before any empirical considerations, the results of the literature on the link between immigration and culture allow us to derive theoretical predictions regarding the impact of immigration on cultural fractionalization in Europe. Prior research suggests that immigration may affect the distribution of values in the host society and by extension its cultural heterogeneity through two main channels: (i) by holding distinctive values and norms compared to the host population, immigrants influence the distribution of values through a direct compositional effect, (ii) the native population, and the already migrant resident population, may also respond to the presence of new individuals and values in

the society, by changing their attitudes and stances.[8] The effect of immigration on overall cultural fractionalization depends on the sum of these two effects.

Based on the previous measures developed in Section 2.2.1, the decomposition of the aggregate index into both within-group and between-group components enables us to disentangle the direct compositional effect of the immigrants from the responses of the native population. Indeed, even if the two measures are not independent, one can read the evolution of F_{ST}, the measure of the cultural divide between immigrants and natives, as the direct effect of immigrants on the distribution of values. On the contrary, focusing on changes in CF^W, which are mainly driven by the native population, can be interpreted as natives' response to the arrival of new immigrants in their region.

The direct effect

The compositional effect of immigration on the cultural fractionalization of the host society results from the mere addition of individuals with plausibly different cultural values with respect to the resident population. Ultimately, it depends on the cultural values brought by the new immigrants in their destination country.

In a society with only two cultural traits, as in the model developed by Desmet and Wacziarg (2021), with one majority and one minority value, where the majority value is mainly held by the group of natives, the compositional effect of immigrants depends on the balance of majority-minority values in the immigration inflows compared to the initial balance of values in the host society. If the balance of values in the immigration inflows is skewed to the minority (majority) value, then the simple arrival of immigrants in the destination country should increase (decrease) the overall cultural fractionalization CF. This should also increase (decrease) F_{ST}, the between-group component by increasing (decreasing) the predictive power of birthplace on one's identity. If the mix of values is similar to the initial population at destination, then one should expect no effect neither on CF nor F_{ST}.

The magnitude of this compositional effect depends on immigrants' initial patterns of cultural self-selection at the origin. In their study on cultural self-selection of emigrants from MENA to OECD countries, Docquier et al. (2020a) find that aspiring emigrants are

[8] It is important to note, that our empirical setting will not allow us to assert anything about the temporal dynamics of these direct and indirect effects, despite the plausible conjecture that the indirect effect stemming from the natives' reaction is more likely to manifest over a longer duration compared to the direct compositional effect triggered by the arrival of immigrants.

culturally self-selected on their religiosity and gender-egalitarian attitudes with respect to aspiring stayers. However, although a small but significant selection was found at origin, aspiring emigrants and actual migrants still exhibit large cultural differences with respect to the destination country's native population (Obradovich et al., 2022).

The extent to which this compositional effect lasts over time is strongly related to the rate of intergenerational transmission of cultural traits and cultural assimilation. Indeed, previous evidence in the literature has found that immigrants tend to assimilate not only economically but also in terms of culture (Algan et al., 2012; Abramitzky et al., 2014; Gonnot and lo Polito, 2021; Fouka et al., 2022; Abramitzky and Boustan, 2022; Gonnot and lo Polito, 2023). This convergence toward native values may be theoretically explained by conformism, wherein individuals derive a premium from coordinating on the same values as the majority, or simply because people do not like to differ from mainstream views (Alba and Nee, 2009; Desmet and Wacziarg, 2021).[9] For instance, studying immigration during the Age of Mass Migration in the United States and using foreign and native-sounding names as a metric to gauge cultural assimilation, Abramitzky et al. (2020) find that immigrants narrow roughly fifty percent of the naming disparity with native-born individuals within 20 years by opting for native-sounding names for their offspring. Notably, the magnitude of this effect persists steadily over time. Over the longer term, which our two-decade analysis may not fully capture, the direct influence of immigrants on cultural fractionalization may also persist or fade through intergenerational transmission of cultural traits to second and subsequent generations (Bisin and Verdier, 2001; Desmet et al., 2017; Rapoport et al., 2020; Desmet and Wacziarg, 2021). The speed (or lack thereof) of this cultural assimilation can be trait-specific (Giavazzi et al., 2019).

Testable hypotheses: *Given the empirical observations that immigrants differ in terms of cultural values from the natives, immigration increases between-group cultural heterogeneity measured with F_{ST}. The intensity of this effect depends on (a) the cultural selection of immigrants from their origin countries, is higher for immigrants from culturally distant countries, and (b) may differ across cultural traits. Eventually, this effect fades over time with the cultural assimilation of immigrants.*

[9] We depart here from Desmet and Wacziarg (2021) assumption, by considering that the utility of conformism can be an increasing function of the overall society instead of group-specific, as also suggested by the sociology literature (Alba and Nee, 2009).

The indirect effect

The indirect effect of immigration on the cultural fractionalization of the host society results from the natives' response to the presence of new individuals and values in the society.[10] Different channels may trigger this response. First, immigrants themselves can influence the culture of the natives (Fisman and Miguel, 2007; Schmitz and Weinhardt, 2019; Tabellini, 2020; Rapoport et al., 2020). For instance, Miho et al. (2023) provide historical evidence that the exogenous forced immigration of German deportees in the Soviet Union increases female labor force participation and pro-gender-equality attitudes in the receiving communities. Similarly, Giuliano and Tabellini (2021) report that US immigrants originating from countries with a stronger welfare state shifted the political preferences of US destination counties toward more social spending and public education. This effect is an increasing function of the likelihood of contact and a decreasing function of the cultural distance between the two groups.

Second, if the pressure to conform is positively influenced by the size of the group holding the majority value, an inflow of immigrants may influence the strength of conformism by changing the relative size of value-specific groups (Desmet and Wacziarg, 2021). An inflow of immigrants holding relatively more majority (minority) value may enhance (dampen) the convergence toward majority values driven by conformism, hence positively (negatively) deviating from its initial trend.[11]

Third, research in social sciences has shown that immigration may induce strong attitudinal changes within the native population (see Dustmann and Preston, 2007b; Dustmann et al., 2018; Edo et al., 2019b; Steinmayr, 2021; Alesina et al., 2022; Alesina and Tabellini, 2022; Keita et al., 2023, among others). For instance, an increase in the salience of immigration has been found to increase polarization by reactivating pre-existing prejudices in the native population (Schneider-Strawczynski and Valette, 2023). If these effects extend to cultural norms, the mere presence of immigrants may also increase the salience of one's birthplace within society, thereby increasing the pressure to conform if, again, individ-

[10] Previous waves of immigrants may also respond to the new arrival of immigrants. We deliberately focus on natives here as driving the indirect effect since it aligns with empirical observations that the share of immigrants relative to the overall population is consistently small across EU regions, such that CF^W is mainly driven by variations in CF among natives.

[11] Similarly, Desmet and Wacziarg (2021) add to their model another factor influencing cultural evolution, namely innovation, defined as the premium placed on specific values. It is conceivable that even if the balance of values in the immigrant inflow perfectly matches that of the initial population at the destination, immigrants may assign varying degrees of importance to holding majority or minority values. Once more, depending on the premium attached to the minority, convergence toward the majority may either be slowed down or accelerated.

uals perceive benefits in aligning with the predominant values of their group (Giavazzi et al., 2019). As a matter of fact, Bazzi et al. (2019) demonstrate that the reallocation of two million voluntary migrants in Indonesia increased the salience of ethnic divisions in polarized communities, and fostered ethnic attachment.

All these three channels ultimately depend on the type of values brought by the immigrants to their host society and the intensity of contact between natives and immigrants. By contact, we mean intergroup contact as defined in Allport et al. (1954), which notably implies communication and cooperation between groups, among other factors. This implies that immigrants closer to the native population either culturally or broadly in terms of individual characteristics (education, language, etc.) may have a higher likelihood of triggering a response from the native population, except in the case where a backlash response is expected.

Testable hypotheses: *The reaction of natives to the arrival of immigrants depends on (a) the values immigrants bring, and (b) the probability of contact between immigrants and natives. Immigrants from culturally similar countries and with a higher probability to interact with natives are likely to accelerate convergence toward majority values, whereas immigrants from culturally distant countries may either have no impact on natives if contact is infrequent, slow down convergence toward the majority, or even trigger a backlash response.*

2.3 Data

This paper combines three main sources of data: (i) the *European Social Survey* (ESS) to compute various indices of cultural heterogeneity; (ii) the *European Labor Force Survey* (EULFS) to measure immigration stocks at the regional level, and (iii) the *Eurostat* data to extract economic and demographic information at NUTS 2 level.[12] The list of regions and countries is reported in Table B.1 in the Appendix. These three sources are sequentially presented in Sections 2.3.1, 2.3.2 , and 2.3.3. Basic descriptive statistics on the final sample of analysis are reported in Section 2.3.4.

[12] The analysis at the regional level relies on the data matched from the European Social Survey (ESS) and the European Labor Force Survey (EULFS). To properly match the data and to be able to track consistent regions over time, we made some methodological choices. These choices mainly relate to the small number of observations of specific regions and the way regional units are defined in the different datasets. These choices are reported in the Appendix B.1. Specifically, we primarily use the NUTS 2 level for most countries. However, due to data limitations, the analysis is conducted at the NUTS 0 level for Cyprus, Estonia, Lithuania, and the Netherlands, and the NUTS 1 level for Austria, Germany, and the United Kingdom.

2.3.1 Cultural heterogeneity data

The European Social Survey (ESS) is a multi-country individual-level survey conducted every two years since 2002 to track the distribution and evolution of values, beliefs, and attitudes across European countries. In each country-wave, the ESS selects a representative sample of approximately 1,500 individuals aged 18 and above. Except for the 2020 wave, the survey has been traditionally administered in respondents' homes by trained interviewers.[13] The survey collects a rich set of personal and household socioeconomic characteristics, such as education, age, birthplace, employment status, and parents' background. Concerning attitudes, values, and beliefs, the ESS elicits respondents' answers on a series of relevant topics in social science included in the core module of the survey. Moreover, every wave includes a rotating section, inquiring about specific topics such as justice, fairness, or welfare attitudes.

The ESS encompasses data from 39 European countries, although not all countries participated in every wave of the survey. To ensure an adequate dataset for panel analysis, we have excluded countries surveyed in fewer than five waves. This exclusion pertains to the following countries: Albania, Croatia, Kosovo, Latvia, Luxembourg, Montenegro, North Macedonia, Romania, Serbia, and Turkey. Given our research focus on the European continent, we further exclude countries not belonging to the European Union such as Iceland, Israel, Norway, the Russian Federation, and Ukraine.

Regarding the selection of cultural traits for constructing the measure of regional cultural heterogeneity presented in Equation (2.1), we adopt a comprehensive approach commonly used in the literature. This approach entails analyzing a broad array of traits and memes that can be linked to culture, as evidenced in previous studies (Alesina et al., 2017; Rapoport et al., 2020; Desmet and Wacziarg, 2021; Jaschke et al., 2022).[14] Two criteria lead our selection of variables. First, whether possible, these traits should overlap the ones used in other studies (e.g., Alesina et al. 2017). Second, the traits should be part of the core module of the ESS, hence asked in every wave. Based on these criteria, we

[13] It is worth noting that the 2020 ESS wave coincided with the COVID-19 pandemic, leading to a change in survey methodology. Several countries transitioned from face-to-face interviews to self-completion of questionnaires. As this change may potentially affect comparability across survey waves, we have taken a cautious approach and excluded the 2020 wave from our benchmark analysis. Table B.5 shows that including the 2020 wave in our main sample does not affect our main results.

[14] As suggested by Kroeber and Kluckhohn (1952), culture is a multifaceted concept with various definitions. However, defining culture precisely falls beyond the scope of this paper, which assumes culture to encompass a series of attitudes, values, and beliefs.

have carefully chosen 46 cultural traits, detailed in Tables B.2 and B.3 in the Appendix. These traits can be categorized into eight comprehensive blocks: (i) *religiosity*, capturing respondent's religiosity and actual religious participation; (ii) *sexual morality*, measuring attitudes related to self-expression and moral values; (iii) *role of the state*, measuring individual political ideology and government involvement in redistribution policies; (iv) *cultural capital*, including aspects related to general trust and importance of a set of values (e.g., obedience); (v) *political engagement*, and participation to politics; (vi) *trust in the institutions*, describing respondent's trust toward national and international institutions; (vii) *attitudes toward migration*, and (viii) *general openness*, complementing our analysis with all the additional items available in the ESS and proposed by Shalom H. Schwartz to capture respondent's moral values (Schwartz, 2012).[15]

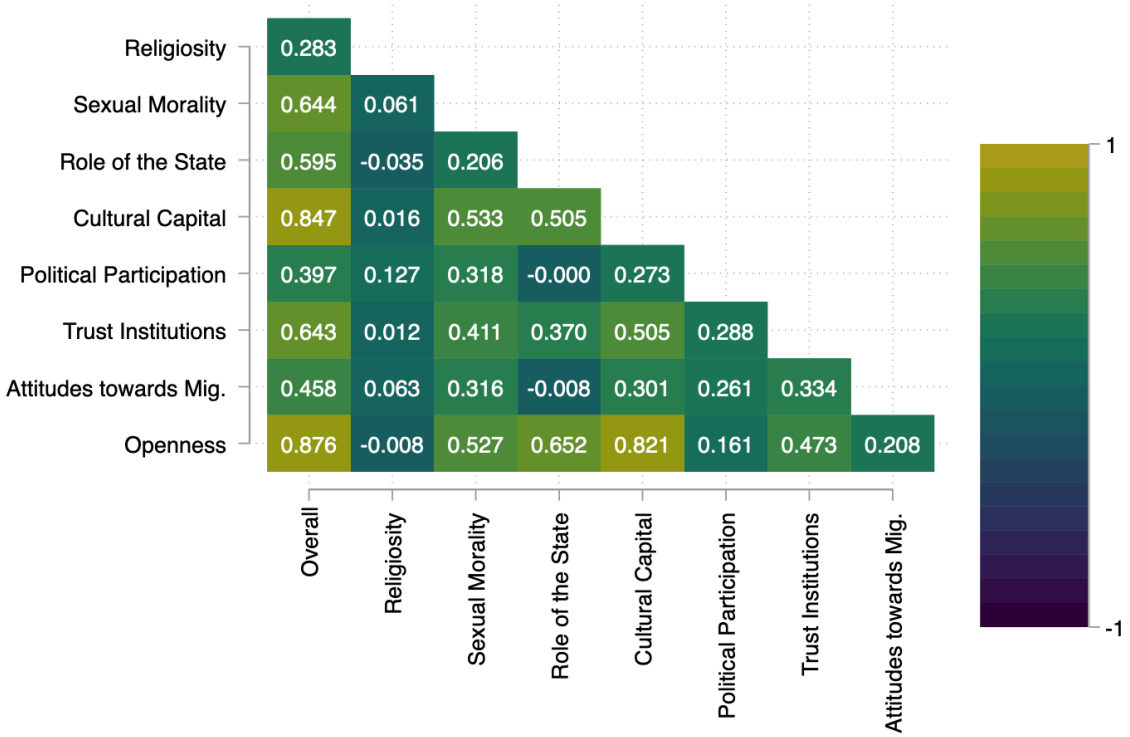
Following the measurement framework presented in the previous section, we compute for each region-year the overall measure of cultural heterogeneity and the within- and between-group components using all cultural traits separately. All the indexes are computed using individual weights provided by the European Social Surveys to make our measures representative. By checking the correlation across different cultural blocks and measures, two main pieces of evidence emerge. First, Fig. 2.1 shows that the correlation between the overall cultural heterogeneity and the ones computed over the different blocks, although always positive, varies substantially across blocks: it is particularly high with the blocks related to cultural capital and overall openness, while it is lower with the blocks related to religiosity and political engagement. Second, regarding the within- and between-group dimensions, the correlation between the within component and the overall heterogeneity in our sample is close to one (0.938), suggesting a strong potential co-movement between the overall and within-origin components. This is not surprising, observing that the within-component is a weighted average of cultural heterogeneity computed among either natives or immigrants, and the weights associated with natives always largely dominate those for immigrants in all regions. The correlation between overall cultural heterogeneity and the between-group dimension stands at -0.343.

2.3.2 Immigration data

We retrieve information on the size and composition of the immigrant population at the regional level from the *European Labor Force Survey* (EULFS). It collects information on

[15] The selection of these various traits and categories aligns with and complements the work of Alesina et al. (2017). The primary exception is gender-equalitarian attitudes, which are omitted due to the lack of questions in different waves, as indicated in Table B.3 and that we only use as a complementary analysis.

Fig. 2.1. Cultural heterogeneity (2004-2018) - cross-correlations across cultural traits.



Note: Cultural heterogeneity across European regions as defined in Eq. 2.1 for different cultural traits.
Source: Authors' elaboration on ESS data (2004-2018).

the age, employment status, and education, of a representative sample of the population above 15 years old. From 2004 on, it provides information on respondents' birthplace over fourteen broad regions[16] and the number of years of residence in the surveyed country. We aggregate this information at the region-by-year level to obtain the stock of foreign-born ($k_{r,t}$), and we then decompose it by migrant population characteristics, such as education (tertiary and not tertiary educated), origin (EU28 origin and not EU28 origin), and length of stay in the host country (less than 5 years, between 5 and 10 years, more than 10 years). We then define the share of foreign-born over the total 2004 population as follows:

$$m_{r,t} = \frac{k_{r,t}}{Pop_{r,2004}} \tag{2.5}$$

[16] The fourteen birthplace regions are: EU15 country, another EU country included with the 2004 expansion, another EU country included with the 2007/2013 expansion, EFTA, Other Europe, North Africa, Other Africa, Near and Middle East, East Asia, South and South East Asia, North America, Central America and Caribbean, South America, and Australia and Oceania.

where $Pop_{r,2004}$ is the 2004 total resident population of the region r. By keeping the population in the denominator fixed at its 2004 value, we ensure that our descriptive evidence and empirical results remain unaffected by potential endogenous changes in native populations (Moriconi et al., 2022). By replacing the total stock of migrants with their decomposed counterparts by education, origin, and length of stay, we provide the share of migrants of different types over the 2004 resident population. Again, all our measures are computed taking into account individual weights provided by the EULFS.

2.3.3 Additional control variables

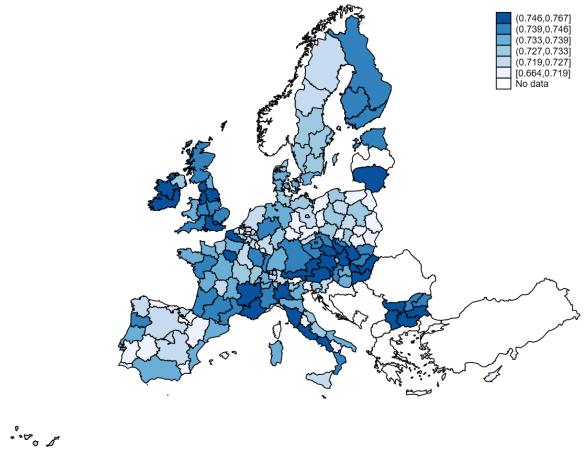
The harmonized *Eurostat* data provides time-varying regional characteristics over time. To capture economic and demographic changes over time, we select the following set of regional characteristics: (i) population density, the number of residents per square kilometer, (ii) the GDP per capita, (iii) the share of high-skilled over the resident population, and (iv) the unemployment rate, as a proxy of labor market conditions.

2.3.4 Sample of analysis and geographical distribution

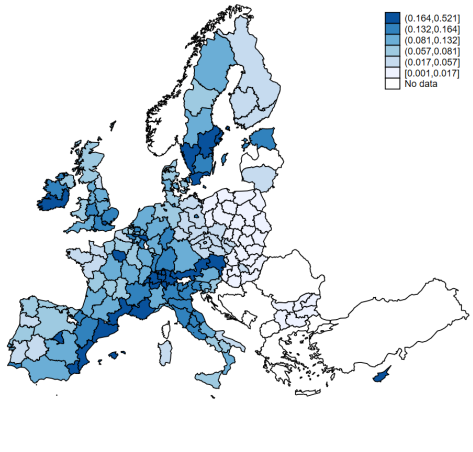
Combining cultural, immigration data, and control variables, we end up with a sample of 1,235 regional-year observations, which corresponds to an unbalanced panel of 175 distinct regions from 23 countries across 8 waves (even years between 2004 and 2018). Summary statistics associated with this sample are reported in Table B.4 in the Appendix. The average degree of overall cultural heterogeneity is around 0.731. Dissecting this measure into its two primary components - within-group and between-group heterogeneity - confirms that a sizeable share of cultural heterogeneity is attributed to the within-group component (0.722), consistently with findings from previous studies on US data (Desmet and Wacziarg, 2021). The average share of migrants is around 10% of the resident population and is mainly driven by low-skilled immigrants (7%) or immigrants coming from outside the European borders (6.1%).

The geographical distribution of our sample is presented in Fig. 2.2, showing the average values of cultural heterogeneity and immigrant populations across European regions from 2004 to 2018. In Fig. 2.2(a), we observe the dispersion of average cultural heterogeneity. Regions in Central-Eastern Europe, such as Eastern Austria (AT1), Central Slovakia (SK03), Ireland, and Île-de-France (FR10), exhibit a high degree of cultural diversity. Conversely, Polish regions and central Spain display the lowest levels of heterogeneity. Fig. 2.2(b) provides a descriptive representation of immigrant distribution across regions.

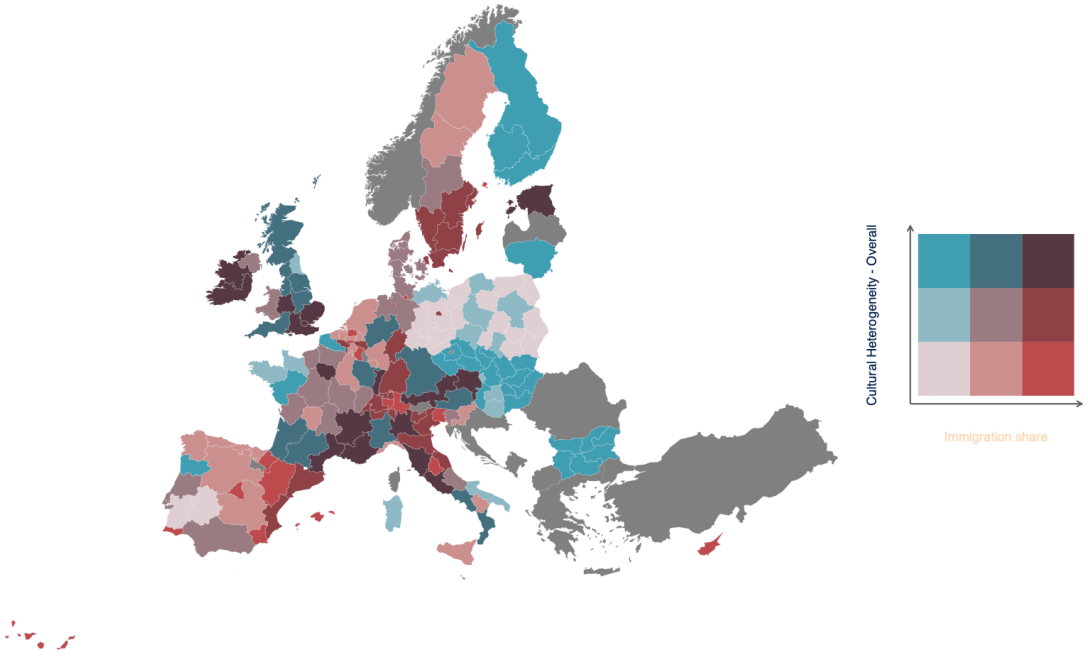
Fig. 2.2. Overall cultural heterogeneity and the share of immigrants - distribution.



(a) Cultural Heterogeneity



(b) Share of Immigrants



(c) Combined Distributions

Note: This figure depicts the average overall cultural heterogeneity across European regions as defined in Eq. (2.1) between 2004 and 2018 and its associated overall share of immigrants as defined in Eq. (2.5). Source: Authors' elaboration on ESS and EULFS data (2004-2018).

Predictably, coastal areas in France, Spain, and Italy exhibit the largest concentration of immigrants, as do regions housing major metropolitan areas like London and Brussels. The overlap between these two distributions is visualized in Fig. 2.2(c), where regions characterized by both a high immigrant population and significant cultural diversity are shaded in dark colors. Once more, coastal regions demonstrate distinctive patterns in

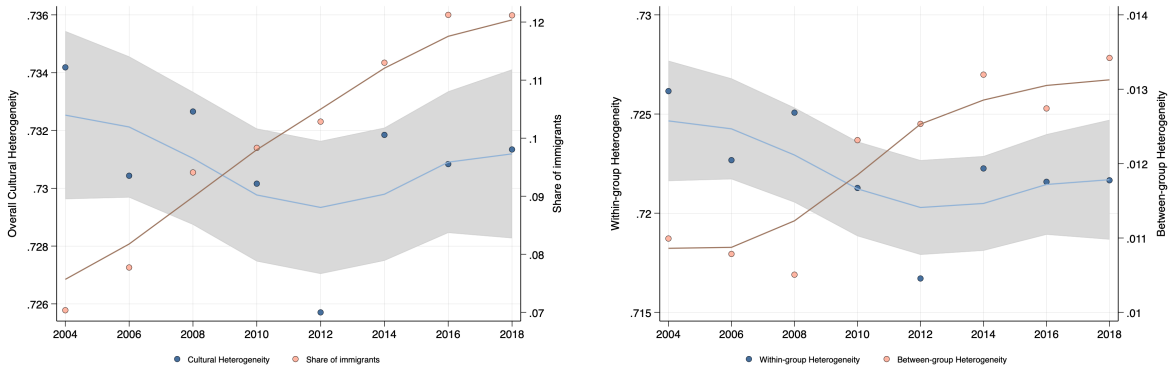
both cultural heterogeneity and immigrant populations, as do regions hosting capital cities. Notably, the Iberian Peninsula is primarily characterized by a high immigrant population and low cultural diversity, while Eastern European countries tend to display the opposite trend.

2.4 Preliminary evidence

This section first reports descriptive statistics on the evolution of cultural heterogeneity and immigration in Section 2.4.1. Then, we provide additional insight on the relevance of birthplace as an identity cleavage in Section 2.4.2.

2.4.1 The evolution of cultural heterogeneity and immigration in Europe

Fig. 2.3. Cultural heterogeneity and immigration - evolution over time.



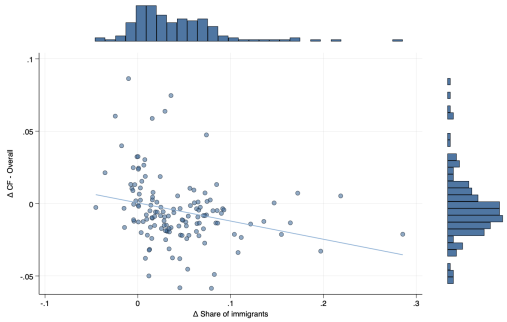
(a) Overall CF and Immigration **(b)** Between and Within Components

Notes: Figure (a) plots the average cultural heterogeneity index and share of immigrants at the NUTS 2 level. Figure (b) plots the average between and within components of cultural heterogeneity.
Source: Authors' calculations on ESS and EULFS data (2004-2018).

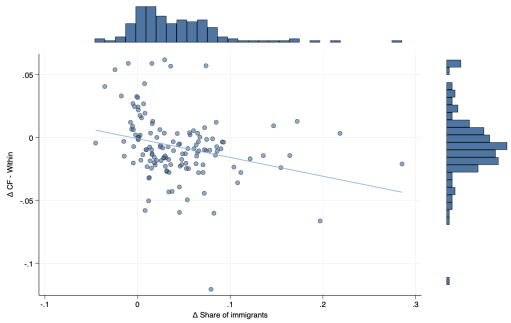
While the figures reported in the previous section offer valuable insights, they do not exactly encompass the variability that we use in the empirical analysis, which focuses on intra-regional variations over time. Thus, Fig. 2.3 illustrates the average trends of our primary variable of interest. Similar to Desmet and Wacziarg (2021) in the US, Fig. 2.3(a) reports that the overall cultural heterogeneity experienced a U-shaped pattern over our period of analysis: a mild decline in the first part of the period analysis, which could be partially explained by the high degree of economic insecurity in the aftermath of the financial crisis, is then followed by a period of positive growth, bringing the overall degree

of cultural heterogeneity closer to its initial level.[17] On the other hand, the share of immigrants evolved with a positive and stable trend, moving from around 7% of the population in 2004 to 12% of the 2004 population in 2018. Concerning the decomposition of the two dimensions of cultural heterogeneity, Fig. 2.3(b) shows that the within-group components experienced a similar trend to the overall cultural heterogeneity; while the average cultural fixation component experienced a positive trend starting from 2008 onward.

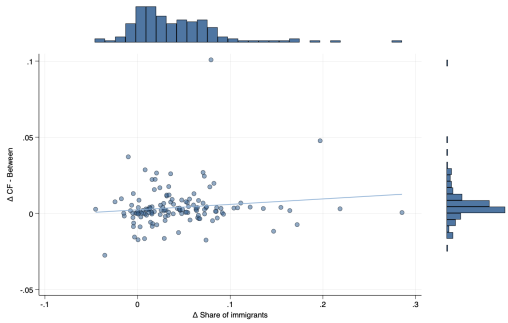
Fig. 2.4. Cultural heterogeneity and immigration - long term variations.



(a) Overall



(b) Within-group



(c) Between-group

Notes: These scatterplots illustrate the relationship between the change in Cultural Heterogeneity (Overall, Within- and Between-group), as defined in Eq. (2.1), and the change in the Share of Immigrants, as defined in Eq. (2.5) between 2004 and 2018 for different regions. Histograms display the distribution of the change in the two variables.
Source: Authors' elaboration on ESS and EULFS data (2004-2018).

[17] Stewart et al. (2020) find an association between economic insecurity and polarization, driven by risk aversion regarding interactions with out-groups. Additionally, several other studies have demonstrated links between economic insecurity in the aftermath of the 2008-2009 financial crisis, the decline in political trust (Wroe, 2016; Algan et al., 2017; Foster and Frieden, 2017; Tormos, 2019), and the rise in demand for populism and far-right voting (Funke et al., 2016; Ausserladscheider, 2019; Guiso et al., 2020; Guriev and Papaioannou, 2022; Ivanov, 2023); all these potentially leading to reduced cultural heterogeneity by fostering a more homogeneous identity.

To get even closer to our empirical analysis presented in the next section, we report in Fig. 2.4 the partial correlation between average long-term variations in our measures of cultural heterogeneity and the share of immigration between 2004 and 2018. Fig. 2.4(a) suggests that regions that experienced larger inflows of immigrants have also experienced a larger decrease in their cultural heterogeneity. At the same time, the decomposition of the within- and between-group components reveals opposite relationships. On the one hand, Fig. 2.4(b) shows that the variation of the within-group component is negatively related to the variation in the share of immigrants. On the other hand, Fig. 2.4(c) depicts a slightly positive correlation between immigration and the extent to which birthplace is a good predictor of cultural differences. Plausibly driven by the non-random allocation of immigrants across regions, this first set of observations calls for a more formal empirical analysis accounting for both unobserved factors and the non-random distribution of immigrants across European regions.

2.4.2 On the relevance of birthplace as cleavage

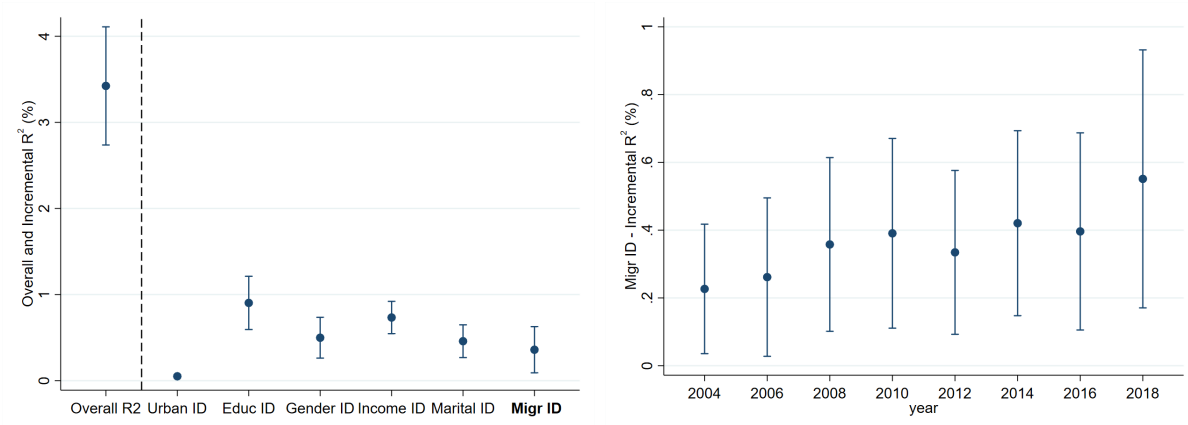
As previously mentioned, one of our main contributions to the literature lies in introducing birthplace as a new cleavage to study the evolution of the cultural divide in modern societies. To motivate further our analysis and to grasp the role played by migration as an identity marker or cleavage, we first estimate a simple linear regression model over the 46 cultural traits, which include a series of dummy regressors highlighted by Desmet and Wacziarg (2021) as potentially relevant identity markers and our birthplace identity marker.[18] We then re-estimate the models excluding one identity marker after the other and record the different R^2 of the estimated models. Finally, we compute the incremental R^2 contribution of each identity marker for each cultural trait i by taking the difference between the estimated overall R^2 once we include all the identity markers and the conditional(s) R^2 once we exclude identity markers one by one. We average these results over the different traits, and report them in Fig. 2.5(a).

The average overall R^2 is small, around 3.5%, in line with previous empirical analysis on the impact of identity markers on individual preferences (Desmet and Wacziarg, 2021; Moriconi et al., 2023). Focusing on the relevance of the different identity markers, income, and tertiary education appear to be the most relevant in explaining individual cultural traits. Migration status fares relatively well among the set of identity markers being as

[18] These identity markers are (i) gender, (ii) college education, (iii) living in an urban area, (iv) belonging to the top two quantiles of the income distribution, (v) marital status. We add immigration status as an additional identity marker.

relevant as marital status or gender, and reporting higher explanatory power than living in an urban area (vs. rural area). Fig. 2.5(b) plots the evolution of the incremental R^2 of migration status as an identity marker. Over time, migration status as an identity marker increases its explanatory power almost threefold. This suggestive evidence confirms the importance of focusing on the role of immigration as a potential contributor to the evolution of the diffusion of cultural values. The importance of birthplace in shaping one's cultural identity can be further highlighted by leveraging individual-level data in the ESS. We provide in Appendix B.5 additional estimates at the individual level that show that first-generation immigrants indeed exhibit strong differences in their cultural traits compared to the native population, even after controlling for other determinants such as age, gender, employment, marital status, education, children, and living in a rural area or not.

Fig. 2.5. Migration status as identity marker - overall and incremental R^2.



(a) Overall R^2 and Incremental by ID **(b)** Incremental R^2 of Migr ID

Notes: Figure (a) plots the average overall R^2 of linear regressions over the 46 cultural traits and also including all the identity cleavage together, and the incremental R^2 due to the inclusion of one identity cleavage at a time. Figure (b) plots the average incremental R^2 of linear regressions over the 46 cultural traits of the migration identity cleavage over time. The figures report the average value and the 95% CI over the 46 cultural traits.
Source: Authors' calculations on ESS data (2004-2018).

2.5 Empirical strategy

This section outlines the empirical strategy of the paper, with the objective of assessing the impact of immigration on the evolution of the cultural heterogeneity of recipient countries. We begin by introducing the benchmark specification in Section 2.5.1. Then, Section 2.5.2 discusses the potential threats to identification and describes our identification strategy

and identifying assumptions, which are based on a Shift-Share (Bartik) instrument.

2.5.1 Benchmark specification

The benchmark specification features CF_{rt} as a measure of cultural heterogeneity in the region r at time t as a dependent variable and m_{rt} as the share of foreign-born over the total 2004 population as the main variable of interest, such that:

$$CF_{r,t} = \alpha + \beta_1 m_{r,t} + \boldsymbol{\beta'} \boldsymbol{X_{r,t}} + \gamma_t + \gamma_r + \varepsilon_{r,t} \tag{2.6}$$

where $\boldsymbol{X_{r,t}}$ is a parsimonious vector of time-varying controls at the regional level including the log of population density, the log of GDP per capita, the share of high-skilled in the population, and the unemployment rate. The parameters γ_t and γ_r stand for year and regional fixed effects, respectively, which control for time-invariant regional characteristics as well as common aggregated change over time. Standard errors are clustered at the regional level since regions are our treated units (Abadie et al., 2023).

2.5.2 Identification strategy

Estimating Equation (2.6) with OLS provides a first insight into the partial correlation between immigration and cultural heterogeneity. Still, immigrants' location choice is not random, therefore this specification may suffer from endogeneity bias and the estimated coefficients cannot be interpreted in causal terms under two conditions: i) time-varying specific regional shocks drive the correlation between immigration and cultural heterogeneity, or ii) immigrants select their locations of residence based on the prevalent cultural heterogeneity. Specifically, the coefficient of interest might be downward biased in absolute terms if immigrants opt to reside in regions with higher levels of multiculturalism.[19] These endogeneity threats, which can be recasted in the form of omitted variable bias and reverse causality, are rather common in empirical studies on immigration, particularly those investigating labor market effects (see Edo (2019) for a review of this literature).

Shift-share approach. To tackle this issue we adopt an instrumental variable (IV) strategy, relying on a shift-share approach (Card, 2001). Such an approach has been widely used in the migration literature (e.g., Ottaviano and Peri, 2006; Docquier et al.,

[19] Immigrants may be a self-selected sample of the origin population (Docquier et al., 2020a). This aspect can be an issue as long as this cultural self-selection drives their destination choice.

2020a; Derenoncourt, 2022) and it builds on the well-documented empirical observation that contemporaneous inflows of migrants from a given origin allocate across different destinations based on the historical geographical distribution of migrants from the same origin. Thus, using information on the initial breakdown of immigrants across regions, one can predict exogenous stocks of immigrants by applying the same allocation scheme to subsequent aggregated inflows. Such an approach then provides a source of variation of immigration that is only driven by the historical distribution of immigrants and by the total inflows by origin, and not by other factors that may drive immigrants' destination selection, such as the region-specific changes in cultural heterogeneity or other unobserved factors. Under the assumption that the predicted immigration flows are orthogonal to omitted characteristics that are correlated with change in cultural fractionalization after 2004, relying on the historical location of the different diasporas across regions, generate therefore a new -as good as random- allocation of the immigrants that can allow for a causal interpretation of our estimates.

Recent developments of the shift-share literature point out that the validity of the instrument relies either on this exogeneity of the initial distribution of immigrants by origin (Goldsmith-Pinkham et al., 2020) or on the exogeneity of the aggregate shocks (Borusyak et al., 2022). Given our empirical setting, our approach matches better the identifying assumption of the exogeneity of the aggregate shocks by origin. To put it differently, we assume the variation of the aggregate inflows of immigrants by origin to be exogenous to the variation of regional-specific cultural heterogeneity.[20] If true, the shift-share approach provides a source of exogenous variation of the immigrant population with respect to cultural fractionalization. To assuage the concerns about the validity of this identifying assumption, we first present the standard shift-share approach and then, propose a modified shift-share approach with predicted aggregate flows through exclusively origin region-specific shocks. By purging out destination-specific pull factors, such an approach is more likely to satisfy our main identifying assumption. Finally, we provide a series of tests to corroborate the validity of this identifying assumption in Section 2.6.3 following Borusyak et al. (2022) suggestions.

Standard shift-share based instrument. We define $Sk_{o,r,2004}$ the initial presence of foreign-born from origin o in the hosting region r in 2004 as the share of the total

[20] For instance, the inflows of immigrants from North Africa in our whole sample of European regions should be orthogonal to the changes in cultural heterogeneity in the Brussel-Capital (B10) region.

immigrants from the same origin country as follows:

$$Sk_{o,r,2004} = \frac{k_{o,r,2004}}{\sum_r k_{o,r,2004}} \tag{2.7}$$

where $k_{o,r,2004}$ is the stock of foreign-born from origin o living in region r in 2004. Our initial year is 2004 since it is the first year in which the EULFS provides the fifteen disaggregated birthplace regions. Then, we compute $Tk_{o,t}$ the total stock of foreign-born for each origin o and year t such as:

$$Tk_{o,t} = \sum_r k_{o,r,t} \tag{2.8}$$

This allows us to construct a predicted stock of foreign-born from origin o in the region r at year t based on their initial distribution in 2004 as the interaction between $Tk_{o,t}$ and $Sk_{o,r,2004}$ such as:

$$\tilde{k}_{o,r,t} = Sk_{o,r,2004} \times Tk_{o,t} \tag{2.9}$$

The aggregate time-variant stocks by origin are then distributed across the regions of our sample based on the 2004 distribution. Finally, we compute the region r and year t predicted migration share ($\tilde{m}_{r,t}$) by simply taking the sum of all $\tilde{k}_{o,r,t}$ predicted stocks across origin, as follows:

$$\tilde{m}_{r,t} = \frac{\sum_o \tilde{k}_{o,r,t}}{Pop_{r,2004}}. \tag{2.10}$$

Modified shift-share based instrument. To enhance the validity of our identifying assumption, which relies on the exogeneity of the aggregate shocks (Borusyak et al., 2022), we modify our shift-share approach by replacing $Tk_{o,t}$ with its predicted version obtained from a zero-stage bilateral migration gravity equation that includes *exclusively* push factors such as conflicts and natural disasters in migrants' origin countries as explanatory variables.[21] This novel approach of combining gravity models with shift-share instruments has gained traction in recent migration literature (See Ortega and Peri, 2014; Docquier et al., 2020b; Orefice et al., 2021, among others). Our gravity equation is as follows:

$$k_{o,d,t} = \alpha_1 \ln(Deaths_{o,t}) + \alpha_2 Disaster_{o,t} + \beta_t \ln(distance_{o,d}) + \theta_{d,t} + \theta_{o,d} + \varepsilon_{o,d,t} \tag{2.11}$$

[21] It is worth to note that our main conclusions remain unchanged when using a simpler version of the shift-share instrument without using immigration stocks estimated from the zero-stage bilateral migration gravity equation.

where $k_{o,d,t}$ is the bilateral stock of immigrants from origin country o to destination d at year t sourced from the United Nations (UN, 2020).[22] The $\ln(Deaths_{o,t})$ and $Disaster_{o,t}$ correspond to origin-specific and time-varying push factors and stand for the logarithm of the cumulative five-year count of total fatalities due to armed conflicts and the cumulative five-year count of natural disasters, respectively.[23] The parameter $\beta_t \ln(distance_{o,d})$ is the time-varying effect of distance on immigration following, which captures the differential impact of changes in technology over time across pairs of countries (Feyrer, 2019; Docquier et al., 2020b). Finally, $\theta_{d,t}$ and $\theta_{o,d}$ stand for destination-year and origin-destination fixed effects, respectively. While destination-year fixed effects are not used to obtain the predicted exogenous stocks, their inclusion in the gravity model allows us to enhance the precision of our estimates.[24] Equation (2.11) is estimated using a Poisson Pseudo Maximum Likelihood estimator (PPML), which performs well under various heteroskedasticity patterns, rounding errors for the dependent variables, and a large number of zeroes (Silva and Tenreyro, 2006, 2010). Standard errors are clustered at the pair level. The total predicted stock of foreign-born for each origin o and year t is such as:

$$\begin{aligned} \widehat{Tk_{o,t}} &= \sum_d \widehat{k_{o,d,t}} \\ &= \sum_d \mathrm{e}^{\widehat{\alpha}_1 \ln(Deaths_{o,t}) + \widehat{\alpha}_2 Disaster_{o,t} + \widehat{\beta}_t \ln(distance_{o,d}) + \widehat{\theta}_{o,d}} \end{aligned} \tag{2.12}$$

Relying predominantly on origin-specific time-varying shocks and purging out the variation generated by destination-specific pull factors, the predicted stocks by origin computed in Equation (2.12) are more likely to satisfy our identifying assumption, which assumes that the variation of the aggregate stocks should be exogenous with respect to changes of the outcome variable and unobserved factors at regional level (Borusyak et al., 2022). We then aggregate them in the fourteen broad origins available in EULFS and use them in

[22] To enhance the precision of our estimates, we maintain the complete 214 × 214 matrix of origin-destination pairs. The gravity model encompasses exclusively 5-year data spanning from 1990 to 2020. For the years in between, immigrant stocks are interpolated before the aggregation of the projected immigration figures.

[23] We sourced data from the UCDP/PRIO Armed Conflict Dataset for deaths (Gleditsch et al., 2002; Davies et al., 2022) and from the Emergency Events Database (EM-DAT) for natural disasters (EMDAT, 2022). Natural disasters include biological (epidemic), climatological (drought, wildfire), geophysical (mass movement, earthquake, volcanic activity), meteorological (storm, fog, extreme temperature), and hydrological events (flood, landslide).

[24] Results for the gravity model are reported in Table B.12 in the Appendix. As predicted by the theory, an increase in the number of natural disasters or fatalities due to conflicts increases international migration while the influence of distance is found to decrease over time.

the construction of the predicted stocks (Eq. 2.9), necessary for the construction of the modified predicted migration share ($\widehat{m}^{Mod}_{r,t}$).

To gain a deeper insight into the underlying variability that our modified shift-share approach leverages, Fig. B.7 in the Appendix illustrates the variation in predicted stocks when aggregated into the fourteen broad origins available in EULFS. Not surprisingly, within-EU immigrants account for the biggest part of the total stock of migrants in our setting. However, upon further examination of group-specific variation, we observe that the predicted growth among different origin groups is rather similar, suggesting that our predicted variation is not driven by any specific origin group. This evidence allows us to interpret our results as being driven by the variation in the immigrant population as a whole, rather than by the variation of specific origin groups.

Finally, Section 2.6.3 provides a series of empirical checks suggested by the literature to support the validity of our identifying assumption. First, we show that our results hold with a leave-one-out version of our shift-share instrument, hence removing from the total stocks those related to each region-year observation, minimizing the presence of any correlation driven by the construction of the instrument (Autor and Duggan, 2003). Second, we provide evidence of no correlation between the growth of predicted total stocks and previous region-specific characteristics, thereby mitigating concerns related to pre-trends (Moriconi et al., 2022). Third, we show that the precision of our estimates is not driven by a similar initial distribution of origin groups across regions, which could potentially bias the estimated error terms (Adão et al., 2019). Although not essential for our identifying assumption, we additionally provide a series of tests to alleviate potential concerns associated with the time closeness between the initial distribution of our historical shares and our period of analysis. Specifically, we show that our results hold by excluding sequentially from our sample those years close to the initial share, hence increasing the gap between our initial distribution by origin and the period of analysis.

2.6 Main results

In this section, we present the results of our empirical analysis. The benchmark OLS and 2SLS estimates of Equation (2.6) relying on the modified shift-share approach are presented in Section 2.6.1. Sections 2.6.2 and 2.6.3 provide robustness checks and empirical tests of the validity of our identifying assumption, respectively. Finally, Section 2.6.4 investigates the mechanisms behind our main effect by decomposing our index of overall heterogeneity between its within and between components.

2.6.1 Benchmark results

Table 2.1
Benchmark table - OLS and 2SLS estimates.

	(1) OLS	(2) OLS	(3) OLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
m_{rt}	-0.000	-0.095***	-0.105***	0.013	-0.138***	-0.162***
	(0.011)	(0.029)	(0.028)	(0.012)	(0.046)	(0.061)
Regional FE	No	Yes	Yes	No	Yes	Yes
Year FE	No	Yes	Yes	No	Yes	Yes
Regional controls	No	No	Yes	No	No	Yes
Observations	1,235	1,235	1,235	1,235	1,235	1,235
Mean Cultural Index	0.731	0.731	0.731	0.731	0.731	0.731
Mean Immig. Share	0.101	0.101	0.101	0.101	0.101	0.101
First-stage				0.740	1.465	1.292
KP F-Test				905.098	82.890	48.528

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variable is CF_{rt} the measure of cultural heterogeneity in the region r at time t. The independent variable m_{rt} is the share of foreign-born in the total 2004 population. Regional controls include the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population.
Sources: Authors' elaboration on ESS and EULFS data (2004-2018).

Table 2.1 presents the results of our benchmark specification (2.6) under various fixed effects structures and control variables. In Column (1), we report the OLS estimate of the relationship between the share of first-generation immigrants and cultural heterogeneity, excluding control variables and fixed effects. This analysis reveals that an increase in the immigrant share does not seem to yield a statistically significant change in cultural heterogeneity. Column (2) introduces regional and year fixed effects, accounting for time-varying common shocks and unobserved regional heterogeneity, and reports that an increase in the share of immigrants is associated with a significant decrease in overall cultural heterogeneity. This negative association holds even when including time-varying regional controls in Column (3) such as population density, GDP per capita, unemployment rate, and the share of high-skilled in the resident population. Columns (4), (5), and (6) employ 2SLS estimation, addressing potential endogeneity concerns through the use of a shift-share-based instrument as described in Section 2.5.2. These results confirm our prior findings that an increase in the immigrant share is associated with reduced cultural heterogeneity in the overall population. Furthermore, compared to OLS estimates, the coefficient with 2SLS becomes bigger in magnitude, aligning with the hypothesis that OLS estimates may be upward biased if immigrants choose their residency based on the

region-specific high degree of multiculturalism. Overall, Table 2.1 consistently reports that a higher first-generation immigrant share is significantly associated with a reduction in overall cultural fractionalization among the resident population. In terms of magnitude, the coefficient of the benchmark specification in Column (6) reveals that a one-percentage-point increase in the share of immigrants is associated with a 0.16 percentage-point decline in overall cultural heterogeneity. To put it in perspective with standard deviations of the two variables, this means that a 10-percentage-point increase in the share of immigrants (one standard deviation of the variable) corresponds to a 1.6-percentage-points decrease in overall cultural heterogeneity (75% of the standard deviation of the variable).

2.6.2 Robustness checks

We conduct a series of robustness checks and tests to check whether our results are sensitive to the sample selected, to the number of observations, to our specification, or to the measurement framework. Detailed results of our analysis are reported in Appendix B.2. We summarize below our main findings.

Robustness to alternative samples. Table B.5 shows that results hold after including the 2020 ESS wave, which adopted a different data collection methodology due to COVID-19. Fig. B.3 (a) and Fig. B.3 (b) show that our main coefficient is largely robust to the exclusion of any waves or countries from our sample, suggesting that our results are not driven by a specific country or wave. Our main coefficient is also robust to the exclusion of any cultural block when computing the cultural heterogeneity index as reported in Fig. B.4, suggesting that the results are not driven by any particular block. Table B.6 includes second-generation immigrants in our analysis and demonstrates that our results are not affected by this change in the population composition, either in terms of significance or magnitude. Table B.7 also reports that our main conclusions remain unchanged when excluding regions with no immigrants in the EULFS, ESS, or both.

Number of observations. Fig. B.5 illustrates the distribution of the average number of observations at the region-year level in the ESS, depicted in Figure (a) for the entire population and Figure (b) specifically for the foreign-born. It is important to highlight that 10% of region-year observations report zero foreign-born individuals according to the European Social Survey (ESS), and the overall distribution skews significantly to the right.[25] These observations raise concerns about the potentially small number of obser-

[25] Similar patterns are observed in Fig. B.6 for EULFS data while the number of absolute zero is

vations within each region-year cell used to compute our dependent variables.[26] We first check in Table B.8 whether our results are not overly sensitive to sequentially excluding region-year where cultural indices are based on fewer than 50 and 100 respondents. While our effect remains robust to the exclusion of regions with less than 50 respondents, we notice a significant decrease in magnitude and precision for the 100-respondents threshold. Still, it is plausible that this coefficient drop reflects a significant sample size change rather than the effect of regions with few observations. Hence, we conduct further additional checks in Table B.9, where we report additional findings interacting our main effect with either a dummy variable for regions with fewer than 50 or 100 respondents. Such a test aims to capture potential heterogeneous effects driven by the number of observations in each region-year cell without modifying the sample of analysis. In both cases, we find that our results remain unaffected by regions with a limited number of observations.[27]

Additional Fixed effects. Table B.10 reports our results providing a different structure of time fixed effects interacted with aggregated regions to capture time-varying regional shocks. Given that our sample includes four countries lacking sub-regional data, namely Cyprus, Estonia, Lithuania (due to the size of the countries), and the Netherlands (due to the availability of data), we first report in Column (2) that our results are not affected by their removal. Then, Column (3) reports that the inclusion of country-year fixed effects removes the statistical significance of our coefficient of interest at conventional levels. This could be attributed to the limited variability in our data and the instrument, as cultural and immigration dynamics over time are largely shared across regions within each country. The incorporation of country-year fixed effects indeed results in an additional reduction of 8 and 7 percentage points in the standard deviation of immigrant shares and overall cultural heterogeneity, respectively and the F-stat of the first-stage equation in the IV is divided by 2. However, reassuringly, the significance of our main effect is restored when time-fixed effects are interacted with broader groups of countries based on geographical regions (Column 4, including Eastern Europe, Central Europe, Southern Europe, Western Europe, and Northern Europe), the 2004 EU enlargement (Column 5, including EU15, NMS10, NMS3, and EFTA), or welfare systems (Column 6, including Nordic, Continental,

substantially lower.

[26] Due to the varying number of respondents to the different cultural traits questions in the ESS, we compute the number of observations at the region-year level using two approaches. First, we compute the number of observations for each cultural trait question in each region-year. Then, we consider both the maximum and the average number of observations across all cultural traits.

[27] We obtain similar results when we consider continuous or categorical measures of observation count by region-year. Our main conclusions remain also robust to weighted estimates using regions' population size despite a slight decrease in the precision of the estimates.

Anglo-Saxon, Mediterranean, and Eastern Europe).

Alternative indices. Although satisfying nice properties and being widely adopted, our measures of cultural heterogeneity are not the only ones available in the literature, as described in Appendix B.1.3. Hence, Table B.11 provides that our results are robust to alternative definitions of our dependent variable. We first adopt a more stringent selection of cultural trait blocks for constructing the cultural heterogeneity measure, retaining only those employed by Alesina et al. (2017).[28] Additionally, we address the possibility that the contribution to overall cultural heterogeneity may vary based on whether an individual's response is close to the prevailing norm in the region. To account for this, we follow Greenberg (1956) and construct an augmented cultural heterogeneity index that assigns higher weights to answers that deviate further from the region-year average answer for a given cultural trait. Another concern involves the varying number of possible answers to each cultural trait question, which could impact the overall heterogeneity measure. While this concern is mitigated by our panel data structure, which explores within-region variation, we also recompute the overall cultural heterogeneity measure using a discretized version of all cultural trait variables. Furthermore, we consider two alternative diversity measures: the Rosenbluth index (Hall and Tideman, 1967) and the Entropy index (Shannon, 1948).[29] The results provided consistently indicate a negative effect of immigration on all cultural diversity indices.

2.6.3 Validity of the instrument

This section addresses several potential concerns associated with the use of a shift-share IV strategy. First, we assess the robustness of our findings by employing two alternative approaches for computing the "shift" component in our shift-share instrument. We then verify the presence or absence of any correlation between pre-existing regional characteristics and the variability in our instrument. We examine the potential bias in our standard errors arising from the correlation in the error term across regions with similar origin-specific shares, as pointed out by Adão et al. (2019). Finally, we provide a robustness analysis excluding years close to the initial distribution of immigrants by origin. We report the results of these tests in Appendix B.3.

Standard & Leave-one-out shift-share. The shift-share instrument described in Sec-

[28] These cultural blocks include religiosity, sexual morality, the role of the state, and cultural capital.
[29] Details about the construction of each index and the correlation between all the different heterogeneity measures are available in Appendix B.1.3.

tion 2.5.2 consists in combining the initial distribution of foreign-born for each origin-region pair in 2004 ($SK_{o,r,2004}$) with the predicted total stock of foreign-born for each origin-year ($\widehat{Tk_{o,t}}$). Our identifying assumption hinges on the exogeneity of these predicted immigrants' stocks (Borusyak et al., 2022). We test the robustness of our benchmark results presented in Table 2.1 with two alternative methods for computing the total immigrants' stocks. First, we consider the more conventional approach, using the actual total stock of foreign-born from origin o in year t in our overall sample of 23 EU countries, as obtained from the EULFS. Second, we implement a leave-one-out version of our primary shift-share instrument, originally proposed by Autor and Duggan (2003). The leave-one-out estimator excludes own-destination i predicted stock of foreign-born when calculating the total predicted stock of foreign-born for each origin o and year t across all destinations d. Hence, we can rewrite Equation (2.12) as follows:

$$\widehat{Tk_{i,o,t}} = \sum_{d-i} \widehat{k_{o,d,t}} \tag{2.13}$$

The rationale behind using this leave-one-out version of the shift-share is to enhance the exogeneity of our instrument by eliminating any remaining mechanical relationships when computing the total predicted stocks for each origin-year observation. The results are presented in Table B.13 and remain robust to using these two alternative versions of the instrument.

Pre-trend analysis. We check that the variation in the predicted immigrant stock is not associated with pre-existing regional trends which could be correlated with cultural heterogeneity. To test that, we estimate the correlation between the growth of several regional indicators over the three years leading up to our initial sample year and the growth of the regional predicted stock of foreign-born over the subsequent three years. In other words, we regress the growth of our shift-share instrument over the period 2004-2007 on the 2000-2003 growth of GDP per capita, population density, unemployment rate, and the share of the tertiary educated population, while controlling for country fixed effects. Results are reported in Table B.14, with Column (1) displaying the results for all immigrants, Column (2) for high-skilled immigrants, Column (3) for low-skilled immigrants, Column (4) for immigrants from EU28 countries, and Column (5) for immigrants from Non-EU28 countries. Our findings indicate that, overall, there is no significant correlation between the pre-2004 trend growth in regional indicators and the variation in predicted immigrant stocks as measured by our instrument. One exception relates to population density's correlation with overall immigration growth, although it is only statistically

significant at the 10% level.

Inference à la Adão et al. (2019). Another concern raised by Adão et al. (2019) in the shift-share setting is the potential spatial correlation of shocks across regions with similar shares. This spatial correlation, if present, could lead to a downward bias in standard errors due to estimation noise in the error terms. To address this concern, we follow the approach proposed by Adão et al. (2019) and compute standard errors that account for the correlation in the error terms between regions with similar initial distribution of immigrants. Table B.15 reports the standard errors, p-values, and confidence intervals obtained using robust standard errors, clustered standard errors (as in our benchmark specification), and the inference procedure described in Adão et al. (2019) (referred to as AKM). Reassuringly, the precision of our estimates remains unaffected when employing any of the aforementioned inference methods.

Historical shares. Although not crucial for our identifying assumption, which is based on the exogeneity of the shocks by origin (Borusyak et al., 2022), one might raise concerns regarding the proximity of the shares we use to construct our instrument to the initial year of estimation in our sample. Therefore, as a robustness check, we adopt a reverse approach by maintaining shares defined in 2004 but sequentially excluding each year in our sample from 2004 onwards. This introduces a time gap between the year our shares are defined in and the initial year of our estimation sample. Fig. B.8 demonstrates that until the exclusion of the 2004-2012 period, when our sample size starts becoming very small, our main conclusions remain unaffected by the exclusion of the preceding years.[30]

2.6.4 Mechanisms

To gain a comprehensive understanding of the driving forces behind the negative association between immigration and cultural heterogeneity within the resident population, we examine whether the impact we observe stems primarily from the pure characteristics of immigrants themselves, constituting a compositional narrative, or if it is predominantly a result of the reactions of the native population to new inflows of immigrants as described in Section 2.2.2.

Table 2.2 disentangles our main effect from our benchmark specification, as presented in

[30] An alternative approach would be to use initial shares obtained before 2004. For instance, Edo and Özgüzel (2023) digitized census data from the early 1990s and merged it with EULFS. Although relevant, this data would cover only 13 countries from our sample, hence generating concerns due to the restriction of our sample of 23 countries.

Column (1) for comparison, into two distinct channels: the compositional consequence and the responses of the native population to new immigrant inflows. Column (2) isolates the effect of an increased immigrant share on cultural fractionalization within the native population. The negative and statistically significant coefficient in this column indicates that as the share of immigrants rises, natives tend to become more culturally homogeneous.[31]

As reported in Section 2.2.2, different mechanisms can drive this effect. On one hand, natives with minority values can absorb the new cultural norms brought by immigrants, hence the change in values would be driven by immigrants with similar values as the native population. Alternatively, natives can simply react homogeneously to the presence of immigrants in their region as a new source of information, and based on the characteristics of immigrants they could either update their beliefs or have a potential backlash toward more conservative values. It is also possible that immigration increases the salience of one's birthplace within society, thereby increasing the pressure to conform if individuals perceive benefits in aligning with the predominant values of their group. A comparison between Columns (1) and (2) also suggests that the direct effect of the immigrant population on cultural heterogeneity is positive, suggesting that immigrants directly contribute to increased cultural heterogeneity by introducing new norms and values to their host country.

This interpretation is further corroborated by Columns (3) and (4), which, in line with Desmet and Wacziarg (2021), decompose our overall index of heterogeneity into its within and between components. It is important to remember that our analysis here considers only two groups within the population, namely natives and foreign-born individuals, excluding second-generation immigrants for simplicity. In line with our earlier findings, we find that within-group heterogeneity (largely driven by natives due to the low share of immigrants in European regions) is inversely associated with an increase in the immigrant share. Conversely, the positive and statistically significant coefficient for between-group heterogeneity suggests that an increase in the immigrant share raises the relevance of place of birth as an indicator of cultural norms. By holding a different set of cultural values and norms, immigrants challenge therefore the cultural homogeneity of their host region by increasing birthplace-driven cultural diversity.

Overall, Table 2.2 confirms, therefore, the existence of the two aforementioned mecha-

[31] We do not find any significant effects when replicating this analysis with immigrants only. These additional results are available upon request from the authors.

Table 2.2
Measures of cultural heterogeneity.

	Overall		Within-group	Between-group
	All	Natives		
	(1)	(2)	(3)	(4)
m_{rt}	-0.162***	-0.203***	-0.186***	0.038**
	(0.061)	(0.067)	(0.067)	(0.019)
Observations	1,235	1,235	1,235	1,235
Mean Cultural Index	0.731	0.728	0.722	0.012
Mean Immig. Share	0.101	0.101	0.101	0.101
First-stage	1.292	1.292	1.292	1.292
KP F-Test	48.528	48.528	48.528	48.528

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variable in Columns (1) and (2) is CF_{rt} the measure of cultural heterogeneity in the region r at time t. The dependent variable in Column (3) is CF^{W}_{rt} the measure of within-group cultural heterogeneity in the region r at time t. The dependent variable in Column (4) is F_{ST} the measure of between-group cultural heterogeneity in the region r at time t. The independent variable m_{rt} is the share of foreign-born in the total 2004 population. We control for the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population. All estimates include regional and year fixed effects.
Sources: Authors' elaboration on ESS and EULFS data (2004-2018).

nisms: immigration influences regional cultural heterogeneity both by triggering a response of the native population and by including in the host societies a distinctive set of values. Nonetheless, the difference in *magnitude* of the two effects is stark: a direct comparison of the two coefficients in Columns (3) and (4) indicates that the compositional effect is largely dominated by a more pronounced reaction of natives.

The reduction in cultural heterogeneity within the native population could be interpreted as either a more concentrated distribution of values (i.e., unimodal distribution) or instead a polarization of values (i.e., bimodal distribution). Indeed, using an additional polarization index (Montalvo and Reynal-Querol, 2005) which describes how closely the overall distribution of values resembles a bimodal distribution, Table B.16 in the Appendix suggests an increase in the polarization of values for the overall population. To disentangle between these two interpretations of the estimated effect, we push further our analysis by examining the impact of immigration on within-group (CF^{W}) and between-group (F_{ST}) measures computed for the native population across various characteristics including ed-

ucation, urbanization, and income groups.[32] Indeed, if natives react differently to the arrival of immigrants, inducing a polarization of their attitudes, this heterogeneity should be rooted in individual characteristics associated with various reactions to immigrants in the economic literature. Results reported in Columns (2) to (4) of Table B.16 indicate a decrease in within-native group heterogeneity, while Columns (5) to (7) reveal no effect on between-native groups heterogeneity. This supports that the increased polarization we primarily observed stems from polarization between the native and foreign-born populations, rather than polarization within the native population.[33] Thus, the decrease in cultural heterogeneity within the native population has to be interpreted as a more concentrated distribution of values, rather than a polarization of values.

2.7 Heterogeneity analysis

This section pushes further the investigation of the mechanisms underpinning the main findings reported so far using a heterogeneity analysis. Indeed, Section 2.2.2 suggests that both the direct and indirect effects evidenced so far should strongly depend on the characteristics of the immigrants. Thus, we investigate in two separate subsections, namely Sections 2.7.1 and 2.7.2, how immigrants' characteristics such as education, duration of stay, or origins influence these two potential channels of transmission.[34] Second, we acknowledge the multi-dimensional nature of the index of overall cultural heterogeneity we have measured, and conjecture that each trait of culture may indeed exhibit a distinct level of responsiveness to immigration. This calls for an in-depth investigation into the specific reactions of various cultural traits to the presence of immigrants as reported in Section 2.7.3. Finally, we summarize the results of a complementary descriptive analysis in Section 2.7.4 that uses observations at the individual level to examine to which extent first-generation immigrants indeed exhibit differences in their cultural traits compared to the native population and the directions in which cultural traits evolve in response to immigration.

[32] The native groups are defined as high-skilled and low-skilled for education, urban and rural for urbanization, and high-income and low-income for income.

[33] The positive correlation between the polarization and the between native and foreign-born heterogeneity measures, as shown in Fig. B.2, further supports this conclusion.

[34] It is worth noting that the issue of a small number of respondents per region-year becomes more pronounced as we narrow our focus to subgroups of immigrants and natives. This consideration must be taken into account and exercised caution when interpreting these results.

2.7.1 Between-group cultural heterogeneity and immigrants' characteristics

To understand which groups of immigrants enhance the compositional effect, in this section we investigate the association between immigration and the between-group cultural fixation index computed over the natives and subsamples of immigrants. Indeed, our theoretical predictions suggest that the greater the cultural distance, the stronger the direct effect should be.

Table 2.3
Immigrants' characteristics and compositional effect.

Dependent variable:	Between-group heterogeneity							
	All	Origin		Education		Duration of stay		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
First gen. group:	All	EU28	NEU28	HS	LS	ST	MT	LT
m_{rt}	0.038**	0.036*	0.043	0.002	0.050**	0.067**	0.042*	0.015
	(0.019)	(0.018)	(0.031)	(0.017)	(0.019)	(0.027)	(0.022)	(0.022)
Observations	1,235	1,235	1,235	1,235	1,235	1,235	1,235	1,235
Mean Cultural Index	0.012	0.008	0.011	0.007	0.012	0.006	0.006	0.010
Mean Immig. Share	0.101	0.101	0.101	0.101	0.101	0.101	0.101	0.101
First-stage	1.292	1.292	1.292	1.292	1.292	1.292	1.292	1.292
KP F-Test	48.528	48.528	48.528	48.528	48.528	48.528	48.528	48.528

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variable is F_{ST} the measure of between-group cultural heterogeneity in the region r at time t for natives and first-generation immigrants' group reported in each column. The independent variable m_{rt} is the share of foreign-born in the total 2004 population. We control for the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population. All estimates include regional and year fixed effects.
Sources: Authors' elaboration on ESS and EULFS data (2004-2018).

Table 2.3 reports the results of new estimates in which we first compute between-group cultural heterogeneity indexes that sequentially isolate specific immigrant sub-groups by different characteristics (origin, education, and duration of stay), and then we estimate our benchmark Equation (2.6). Our modification influences solely our dependent variable, rather than altering the specific immigration flows in our variable of interest. This approach enhances comparability across specifications, as they all share the same first-stage equation. In Columns (2) and (3), we regress the overall percentage of foreign-born individuals within the total 2004 population against the between-group cultural diversity of two distinct hypothetical populations: one composed exclusively of natives and immigrants from other European countries, and the other composed of natives and immigrants from outside European countries, respectively. Overall, Table 2.3 reveals that immigra-

tion increases the cultural fixation of the population with immigrants from distant origin (Not EU28) countries and the native population, although this effect is not precisely estimated in Column (3). When it comes to education, we find that the compositional effect is primarily driven by low-skilled immigrants, as evidenced in Column (5), while no significant effect is observed for college graduates in Column (4). Lastly, Columns (6) to (8) focus on immigrants' duration of stay and provide robust evidence in favor of cultural assimilation over time. Specifically, Column (6) shows that immigration increases the relevance of birthplace as an identity marker for migrants with less than 5 years of residency, but this effect strongly decreases for immigrants between 6 and 10 years of residency and ultimately disappears entirely after 10 years of residency.[35] Overall this set of results aligns perfectly with our theoretical predictions. It reports that immigration enhances the relevance of birthplace to explain cultural heterogeneity mostly among low-educated immigrants and those from outside European countries, and this effect vanishes over time through assimilation.

2.7.2 Within-group cultural heterogeneity and immigrants' characteristics

Since the impact of the within-group component is primarily driven by natives' responses to immigration, this section explores the specific attributes of immigrant inflows—such as education, country of origin, and length of stay—that elicit stronger reactions among natives. As reported in Table 2.2, variations in within-group heterogeneity are entirely driven by the changes in cultural heterogeneity within the native population.

Table 2.4 estimates therefore our benchmark specification by focusing sequentially on different immigrant inflows (m^g_{rt}). Among the different groups g we decompose the share of immigrants by group of origins (EU28 and Not EU28), education, and duration of stay in the host region. It is worth noting that our instrument can be replicated for any migrants' characteristics, as long as we restrict $Tk_{o,t}$ to a given characteristic.[36]

[35] Table B.17 in the Appendix provides additional support for cultural assimilation. Indeed, this table replicates our benchmark analysis replacing first-generation immigrants with second-generations only. One can see that the coefficient for the between-group component is not significant. It supports that being a native of a second-generation immigrant is not informative on cultural differences.

[36] It implies that the variability of such an instrument between high- and low-skilled groups, for instance, would differ mainly due to variations in skill-specific inflows from various origin countries. This difference may be quite low if the initial distribution of both groups across regions in 2004 was fairly similar. Additionally, the instrument's strength may vary across groups making comparisons harder. This is the case for low-skilled immigrants in our analysis, which warrants a cautious interpretation of the 2SLS coefficients associated with this specific group.

Table 2.4
Immigrants' characteristics and natives' response.

Dependent variable:	Within-group heterogeneity							
	All	Origin		Education		Duration of stay		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Immig. charac.:	All	EU28	NEU28	HS	LS	ST	MT	LT
m^g_{rt}	-0.186***	-0.637***	-0.242**	-0.239**	-0.391*	-0.296***	-0.284	-0.074
	(0.067)	(0.231)	(0.104)	(0.096)	(0.234)	(0.091)	(0.172)	(0.064)
Observations	1,235	1,235	1,235	1,235	1,235	1,235	1,235	1,235
Mean Cultural Index	0.722	0.722	0.722	0.722	0.722	0.722	0.722	0.722
Mean Immig. Share	0.101	0.040	0.061	0.027	0.074	0.020	0.018	0.061
First-stage	1.292	1.482	0.926	1.228	0.925	0.971	0.946	1.040
KP F-Test	48.528	18.593	30.725	84.531	5.930	17.022	24.524	45.370

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variable is CF^W_{rt} the measure of within-group cultural heterogeneity in the region r at time t. The independent variable m^g_{rt} is the share of foreign-born, belonging to group g reported in each column, in the total 2004 population. We control for the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population. All estimates include regional and year fixed effects.
Sources: Authors' elaboration on ESS and EULFS data (2004-2018).

This additional set of estimates allows us to discriminate between the aforementioned alternative interpretations of a negative association between immigration and within-group heterogeneity. Results presented in Columns (2) and (3) show that natives' response is stronger with European immigrants, rather than more distant and culturally heterogeneous non-European immigrants. Similarly, the response appears to be more precisely estimated for highly-educated immigrants rather than low-educated ones (Columns (4) and (5)). These results may suggest that the natives' response is stronger with changes in the groups of immigrants that are more likely to interact with natives and to hold values closer to the majority. Therefore, such a response is less likely to be interpreted as a backlash, where natives seek differentiation from out-group members (i.e., immigrants perceived as more distant), but rather a reaction to immigrants to whom natives can relate with, accelerating the convergence toward the majority values. Moreover, results in Column (6) could suggest that these effects are enhanced by newly arrived immigrants.[37]

We push further our analysis and try to unravel which group of natives respond to immigrant inflows. To achieve this, we compute our overall cultural heterogeneity measure for each native subgroup across several characteristics including education (high-skilled and low-skilled), urbanization (urban and rural), and income (high-income and low-income).

[37] Estimates in Columns (7) and (8) are difficult to interpret as they do exploit the variation of stocks of medium and long-term immigrants, which can be hardly interpreted as inflow of immigrants. Instead, they capture the transition of certain cohorts from one category to another over time, without considering the arrival of new immigrants in the destination country.

Table 2.5
Overall cultural heterogeneity within native subgroups.

Dependent variable:	Within-native subgroup heterogeneity					
	Education		Urbanization		Income	
	(1)	(2)	(3)	(4)	(5)	(6)
Native subgroup:	HS	LS	Urban	Rural	High Inc.	Low Inc.
m_{rt}	-0.306*	-0.216***	-0.105	-0.044	-1.068***	-0.159**
	(0.168)	(0.068)	(0.230)	(0.248)	(0.346)	(0.061)
Observations	1,231	1,235	1,196	1,220	1,206	1,229
Mean Cultural Index	0.680	0.725	0.677	0.718	0.682	0.717
Mean Immig. Share	0.101	0.101	0.101	0.101	0.102	0.101
First-stage	1.291	1.292	1.293	1.306	1.303	1.310
KP F-Test	48.420	48.528	47.807	49.174	49.606	50.360

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variables are CF^g_{rt} the measure of cultural heterogeneity within each native subgroup g in the region r at time t. The independent variable m_{rt} is the share of foreign-born in the total 2004 population. We control for the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population at the regional-level. All estimates include regional and year fixed effects.
Sources: Authors' elaboration on ESS and EULFS data (2004-2018).

Table 2.5 reports two main pieces of evidence. First, it shows a consistent decrease in heterogeneity within all native subgroups. Second, there is an interesting heterogeneity, where the estimated effects are bigger in magnitude among high-skilled natives in Column (1), and natives in the high-income group in Column (5) appear to be the most responsive to immigrant inflows. Again, these findings align with our previous interpretation of the results, as high-skilled and high-income native groups are more likely to interact with high-skilled immigrants and immigrants from European countries.

2.7.3 Heterogeneity Analysis on Cultural Traits

Focusing on each memetic trait separately in CF^m, this section investigates which specific cultural values are influenced by immigration within the broader spectrum of overall cultural heterogeneity. To conduct this analysis, Fig. 2.6 reports the coefficients of distinct estimates for all possible combinations of specific immigrant stocks and cultural traits. To ease the interpretation of the differences in magnitudes between these estimates, this figure reports beta coefficients with standardized dependent and independent variables. While each empty square corresponds to an insignificant coefficient, each colored one represents significance at the five percent level. Darker coefficients stand for estimated

coefficients with a larger magnitude.

Fig. 2.6(a) depicts results related to within-group heterogeneity. As shown in previously reported results, the estimated variability here is similar to the one used with overall heterogeneity or focusing on heterogeneity within natives only. It reveals that natives respond to immigration on nearly all dimensions, depending also on the characteristics of the immigrants, with the exception of sexual morality. In line with our previous results, the strongest responses correspond to high-skilled immigration from European countries; and the coefficient is larger for cultural capital, the role of the state, openness, and political engagement.

Fig. 2.6(b) depicts results related to between-group heterogeneity. This set of results emphasizes that between-group heterogeneity is mostly influenced by cultural differences related to religion, sexual morality, the role of the state, and political engagement. However, it is crucial to interpret the findings regarding political engagement cautiously. As it is defined, this specific trait includes questions such as "Did you vote in the last national elections?", which might be interpreted differently by immigrant respondents, especially if they understand the question as pertaining to their destination country where they may not have had the right to vote or if elections were conducted before their arrival.[38] Thus, we provide in Fig. B.4 additional estimates excluding cultural blocks one by one from our benchmark results. Our main conclusions remain unchanged.

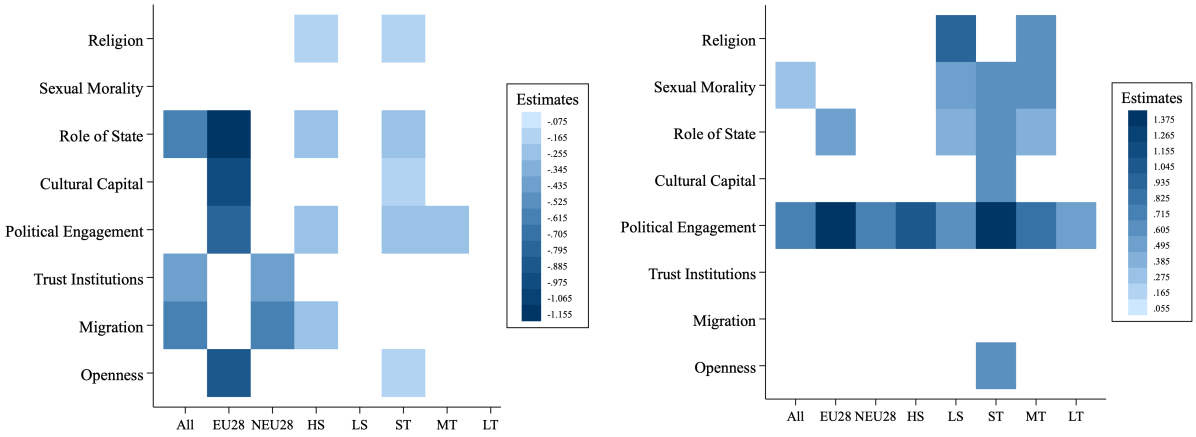
2.7.4 Individual level analysis

Up to this point, our results indicate that cultural heterogeneity increases with the new arrival of low-skilled immigrants, but this effect is outweighed by the response of native populations which became less culturally diverse due to high-skilled immigrants from European countries, resulting in an overall reduction of cultural heterogeneity caused by immigration. A legitimate follow-up question is therefore, beyond heterogeneity, to explore the direction in which immigrants influence the culture of the destination, considering both the compositional effect and the reactions of native residents. This section briefly summarizes the results of a complementary descriptive analysis reported in Appendix B.5 that investigates this question by exploiting the individual-level dimension available in the European Social Survey.

[38] Fig. B.9 in the Appendix reports the same results as Fig. 2.6 with 90% confidence intervals. We also report the within- and between-group heterogeneity for additional cultural blocks in Fig. B.10 and Fig. B.11, namely Schwartz's human values and gender.

Fig. 2.6. Cultural traits and immigrants' characteristics.

(a) Within-group heterogeneity **(b)** Between-group heterogeneity



Notes: Each squared cell in Figure (a) represents the estimated coefficient from regressing the measure of within-group heterogeneity for each cultural block reported on the y-axis on the share of foreign-born belonging to the group reported on the x-axis. Each squared cell in Figure (b) represents the estimated coefficient from regressing the measure of between natives and first-generation immigrants' group reported on the x-axis heterogeneity for each cultural block reported on the y-axis on the share of foreign-born in the total 2004 population. Blank cells represent insignificant coefficients at the 5% level. Each shaded cell corresponds to a significant coefficient at the 5% level, and darker cells mean larger coefficients. We control for the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population. All estimates include regional and year fixed effects. Standard errors are clustered at the regional level.
Source: Authors' elaboration on ESS and EULFS data (2004-2018).

We first examine to which extent first-generation immigrants indeed exhibit differences in their cultural traits compared to the native population. We find that the average immigrant exhibits different cultural traits than the average native across all dimensions. This finding is consistent with the previously estimated increase in between-group heterogeneity after their arrival. Compared to the native population, immigrants tend to be more religious, hold more conservative views toward gays and lesbians, and are more inclined to believe that traditions and customs must be followed. They lean more toward left-wing political views and they report a higher level of trust and more positive attitudes toward immigrants. As expected they are also less likely to be politically engaged at destination.

We then explore the direction in which natives' cultural traits evolve in response to immigration. In line with the indirect effect generated by the reaction of natives to the arrival of immigrants, we find in individual estimates that high-skilled immigration from European countries is correlated with natives becoming more liberal, especially on the role of state cultural block. We also find that natives' attitudes toward immigrants are positively affected as well as their general openness.

2.8 Conclusions

This paper investigates how immigration has affected the cultural heterogeneity of European regions over the last two decades. By combining data from the European Social Surveys (2004 to 2018) to measure the evolution of cultural heterogeneity in Europe across several cultural dimensions, along with immigration data from the European Labor Force Survey, we find that an increase in the share of immigrants is significantly associated with a reduction in overall cultural fractionalization among the resident population. This comprehensive effect results from the interplay of two counterbalancing mechanisms. On the one hand, we find that recent inflows of low-skilled immigrants heighten cultural fractionalization by introducing novel cultural values and norms to the host regions, particularly related to religiosity, sexual morality, and political orientation, for instance. However, this compositional effect tends to disappear within 10 years after immigrants' arrival, supporting the idea of fast cultural assimilation. On the other hand, this effect is counterbalanced by the reaction of native populations to the arrival of high-skilled immigrants from other European countries. This paper supports, therefore, the statement that the mere presence of new immigrants in a given location may increase cultural heterogeneity, but it also demonstrates that this effect is largely compensated by the natives' response to inflows of high-skilled migrants within Europe, which are culturally closer and to which natives can get close to. It is also dampened by the process of immigrants' cultural assimilation over time. Future research will have to explore and test more carefully the mechanisms of why and how the cultural traits of natives react to the arrival of these specific immigrant inflows. Overall, this paper brings to light the importance of focusing on birthplace, among other cleavages, to study the evolution of the cultural divide in modern societies.

B Appendix to Chapter 2

B.1 Data construction

B.1.1 Regional harmonization

Our regional-level analysis hinges on the integration of data sourced from both the European Social Survey (ESS) and the European Labor Force Survey (EULFS). We made a series of methodological decisions to ensure the full comparability of the regions between surveys and across time. These choices were mainly prompted by the relatively limited number of observations associated with particular regions, as well as the distinct manners in which regional entities are defined across the various datasets.

Austria - EULFS provides information only at NUTS 1 level, hence we aggregate the observations available in ESS to match the same NUTS 1 administrative units.

Finland - The NUTS 2 Åland region (FI20) appears in the ESS data only on four waves, given the small size of the region.

France - We exclude from our analysis the *territoire d'outre-mer*. Moreover, ESS does not provide enough observations to have a representative sample of the region FR83 (Corsica).

Germany - EULFS provides information only at NUTS 1 level, hence we aggregate the observations available in ESS to match the same NUTS 1 administrative units.

Ireland - We follow EULFS NUTS 2 classification, which splits Ireland into two regions: the Border, Midland and Western (IE01) and the Southern and Eastern (IE02).

Italy - We merge together the observations belonging to the region of Trento (ITH1) and Sud-Tirol (ITH2). These two areas are part of the same region, named Trentino Alto-Adige, which appears in our dataset only in four waves, compared to the rest of the Italian regions, in which we have over five different waves. Moreover, we merge the region Molise (ITF2) with Abruzzo (ITF1) and the region Valle D'Aosta (ITC2) with Pidemont (ITC1), given the small number of observations associated to these regions ITF2 and ITC2, characterized by a reduced population.

Spain - We merge in one unique region the information associated with the two autonomous cities Ceuta (ES63) and Melilla (ES64), which appears in only eight waves of the ESS, compared to the rest of the regions that are defined over the whole ESS dataset. Moreover, information on La Rioja (ES23) are available only from 2004, hence we merge the few observations associated with this region with the ones from the Aragon region (ES24).

United Kingdom - EULFS provides information only at NUTS 1 level, hence we aggregate the observations available in ESS to match the same NUTS 1 administrative units.

Table B.1. List or regions and countries.

Region	Country	Nb. waves	Region	Country	Nb. waves	Region	Country	Nb. waves
AT1	Austria	7	FR82	France	8	PL61	Poland	8
AT2	Austria	7	DE1	Germany	8	PL62	Poland	8
AT3	Austria	7	DE2	Germany	8	PL63	Poland	8
BE10	Belgium	8	DE3	Germany	8	PT11	Portugal	8
BE21	Belgium	8	DE4	Germany	8	PT15	Portugal	8
BE22	Belgium	8	DE5	Germany	8	PT16	Portugal	8
BE23	Belgium	8	DE6	Germany	8	PT17	Portugal	8
BE24	Belgium	8	DE7	Germany	8	PT18	Portugal	8
BE25	Belgium	8	DE8	Germany	8	SK01	Slovak Republic	6
BE31	Belgium	8	DE9	Germany	8	SK02	Slovak Republic	6
BE32	Belgium	8	DEA	Germany	8	SK03	Slovak Republic	6
BE33	Belgium	8	DEB	Germany	8	SK04	Slovak Republic	6
BE34	Belgium	8	DEC	Germany	8	SI03	Slovenia	8
BE35	Belgium	8	DED	Germany	8	SI04	Slovenia	8
BG31	Bulgaria	5	DEE	Germany	8	ES11	Spain	8
BG32	Bulgaria	5	DEF	Germany	8	ES12	Spain	8
BG33	Bulgaria	5	DEG	Germany	8	ES13	Spain	8
BG34	Bulgaria	5	HU10	Hungary	8	ES21	Spain	8
BG41	Bulgaria	5	HU21	Hungary	8	ES22	Spain	8
BG42	Bulgaria	5	HU22	Hungary	8	ES24	Spain	8
CY00	Cyprus	5	HU23	Hungary	8	ES30	Spain	8
CZ01	Czech Republic	7	HU31	Hungary	8	ES41	Spain	8
CZ02	Czech Republic	7	HU32	Hungary	8	ES42	Spain	8
CZ03	Czech Republic	7	HU33	Hungary	8	ES43	Spain	8
CZ04	Czech Republic	7	IE01	Ireland	8	ES51	Spain	8
CZ05	Czech Republic	7	IE02	Ireland	8	ES52	Spain	8
CZ06	Czech Republic	7	ITC1	Italy	3	ES53	Spain	8
CZ07	Czech Republic	7	ITC3	Italy	3	ES61	Spain	8
CZ08	Czech Republic	7	ITC4	Italy	3	ES62	Spain	8
DK01	Denmark	5	ITF1	Italy	3	ES70	Spain	8
DK02	Denmark	5	ITF3	Italy	3	SE11	Sweden	8
DK03	Denmark	5	ITF4	Italy	3	SE12	Sweden	8
DK04	Denmark	5	ITF5	Italy	3	SE21	Sweden	8
DK05	Denmark	5	ITF6	Italy	3	SE22	Sweden	8
EE00	Estonia	8	ITG1	Italy	3	SE23	Sweden	8
FI18	Finland	8	ITG2	Italy	3	SE31	Sweden	8
FI19	Finland	8	ITH2	Italy	3	SE32	Sweden	8
FI1D	Finland	8	ITH3	Italy	3	SE33	Sweden	8
FI20	Finland	4	ITH4	Italy	3	CH01	Switzerland	6
FR10	France	8	ITH5	Italy	3	CH02	Switzerland	6
FR21	France	8	ITI1	Italy	3	CH03	Switzerland	6
FR22	France	8	ITI2	Italy	3	CH04	Switzerland	6
FR23	France	8	ITI3	Italy	3	CH05	Switzerland	6
FR24	France	8	ITI4	Italy	3	CH06	Switzerland	6
FR25	France	8	LT00	Lithuania	6	CH07	Switzerland	6
FR26	France	8	NL00	Netherlands	8	UKC	United Kingdom	8
FR30	France	8	PL11	Poland	8	UKD	United Kingdom	8
FR41	France	8	PL12	Poland	8	UKE	United Kingdom	8
FR42	France	8	PL21	Poland	8	UKF	United Kingdom	8
FR43	France	8	PL22	Poland	8	UKG	United Kingdom	8
FR51	France	8	PL31	Poland	8	UKH	United Kingdom	8
FR52	France	8	PL32	Poland	8	UKI	United Kingdom	8
FR53	France	8	PL33	Poland	8	UKJ	United Kingdom	8
FR61	France	8	PL34	Poland	8	UKK	United Kingdom	8
FR62	France	8	PL41	Poland	8	UKL	United Kingdom	8
FR63	France	8	PL42	Poland	8	UKM	United Kingdom	8
FR71	France	8	PL43	Poland	8	UKN	United Kingdom	8
FR72	France	8	PL51	Poland	8			
FR81	France	8	PL52	Poland	8			

Source: Authors' elaboration on ESS and EULFS data (2004-2018).

B.1.2 Cultural blocks - definitions and proxies

Table B.2
Cultural blocks - definitions and proxies (I).

Block(s)	Variables & Description	Scale
ESS		
Bl. I - Religiosity	RE1 - Do you belong to a religious group?	0-1
	RE2 - How religious are you?	0-10
	RE3 - How often do you attend religious services?	0-6
	RE4 - How often do you pray?	0-6
Bl. II - Sexual Morality	SM1 - Gays and lesbians free to live the life they wish	0-4
	SM2 - Important to seek fun and things that give pleasure	0-5
	SM3 - Important to follow tradition and customs	0-5
Bl. III - Role of the State	RS1 - Government should reduce income differences	0-4
	RS2 - Self-positioning left-right scale	0-10
	RS3 - Government should be strong and should ensure safety	0-5
Bl. IV - Cultural Capital	CK1 - Most people can be trusted	0-10
	CK2 - Most people try to be fair	0-10
	CK3 - Most of the time people try to be helpful	0-10
	CK4 - Important to make own decision and be free	0-5
	CK5 - Important to be successful and people recognize you	0-5
	CK6 - Important to do what is told and follow the rules	0-5
	CK7 - Important to help people and care for others	0-5

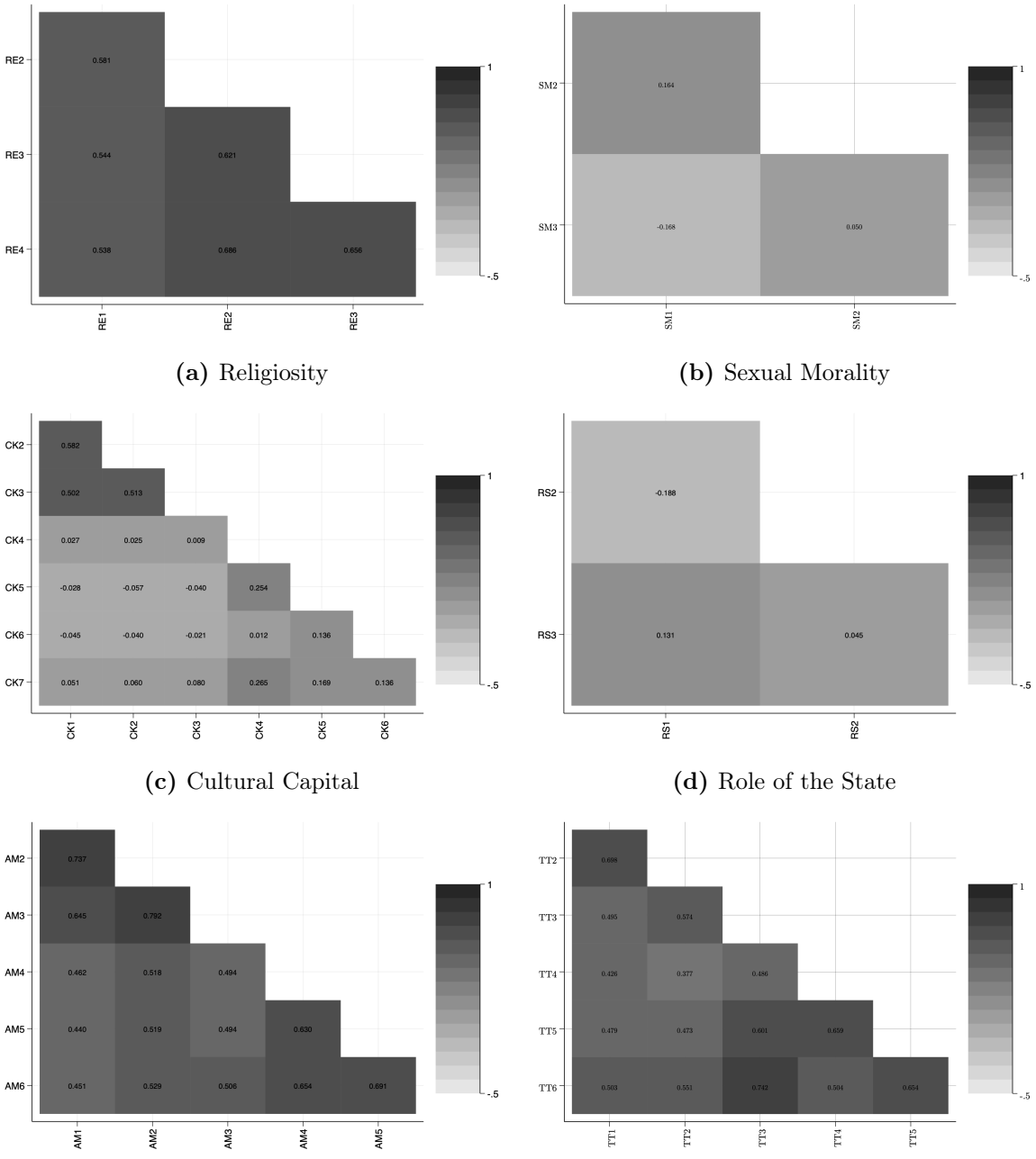
Note: All the questions selected from ESS are available in all the ESS waves.
Source: Authors' elaboration on ESS data (2004-2018).

Table B.3
Cultural blocks - definitions and proxies (II).

Block(s)	Variables & Description	Scale
ESS		
Bl. V - Political Engagement	PP1 - How interested would you say you are in politics?	0-3
	PP2 - Did you vote in the last national election?	0-1
	PP3 - Is there a particular political party you feel closer to than all the other parties?	0-1
Bl. VI - Trust Institutions	TT1 - Do you trust the United Nations?	0-10
	TT2 - Do you trust the European Parliament?	0-10
	TT3 - Do you trust politicians?	0-10
	TT4 - Do you trust the police?	0-10
	TT5 - Do you trust the legal system?	0-10
	TT6 - Do you trust the country's parliament?	0-10
Bl. VII - Attitude toward Migration	AM1 - Do you think [country] should allow people of the same race or ethnic group as most [country] people to come and live here?	0-3
	AM2 - How about people of a different race or ethnic group from most [country] people?	0-3
	AM3 - How about people from the poorer countries outside Europe?	0-3
	AM4 - Would you say it is generally bad or good for [country]'s economy that people come to live here from other countries?	0-10
	AM5 - [country]'s cultural life is generally undermined or enriched by people coming to live here from other countries?	0-10
	AM6 - Is [country] made a worse or a better place to live by people coming to live here from other countries?	0-10
Bl. VIII - Openness	OP1 - Important to think new ideas and being creative	0-5
	OP2 - Important to be rich, have money and expensive things	0-5
	OP3 - Important that people are treated equally and have equal opportunities	0-5
	OP4 - Important to show abilities and be admired	0-5
	OP5 - Important to live in secure and safe surroundings	0-5
	OP6 - Important to try new and different things in life	0-5
	OP7 - Important to understand different people	0-5
	OP8 - Important to be humble, modest and not draw attention	0-5
	OP9 - Important to have a good time	0-5
	OP10 - Important to seek adventures and have an exciting life	0-5
	OP11 - Important to behave properly	0-5
	OP12 - Important to get respect from other	0-5
	OP13 - Important to be loyal to friends and devote to people close	0-5
	OP14 - Important to care for nature and environment	0-5
Bl. IX - Gender Equality	GE1 - Men should have more rights than women to work when jobs are scarce	0-5

Note: All the questions selected from ESS are available in all the ESS waves, excluding GE1, which is available only in a subset of waves (ESS2, ESS4, ESS5, and ESS8).
Source: Authors' elaboration on ESS data (2004-2018).

Fig. B.1. Pairwise correlations.



(a) Religiosity

(b) Sexual Morality

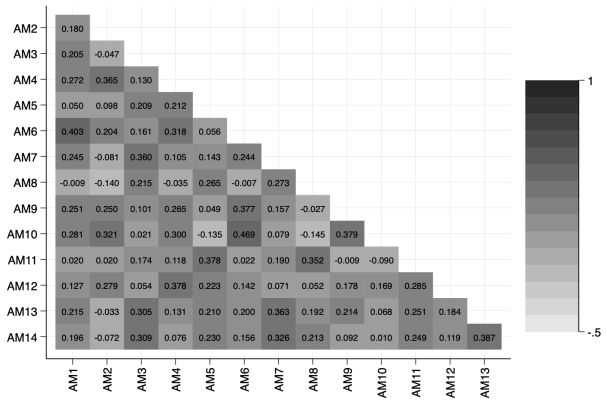
(c) Cultural Capital

(d) Role of the State

(e) Attitudes toward Migration

(f) Trust in Institutions

Fig. B.1. Pairwise correlations [cont'd].



(g) Openness

Note: Pairwise correlations between cultural traits questions within each cultural block.
Source: Authors' elaboration on ESS data (2004-2018).

B.1.3 Construction of alternative indices

In Section 2.2, we measure regional cultural heterogeneity with a cultural fractionalization index. This index is a widely accepted and reliable measure with desirable measurement properties as outlined by Hall and Tideman (1967). Also, it has the advantage that it can be broken down into within and between components once an identity cleavage is identified (Desmet and Wacziarg, 2021). Furthermore, its extensive use in the literature makes it easy to understand and compare with alternative studies (Desmet et al., 2017). Still, its variation can be influenced by the number of questions and available answers for each question, and it does not consider that the contribution of each answer to the overall cultural heterogeneity might differ based on prevailing norms. Therefore, in this appendix, we explore the construction and properties of alternative cultural heterogeneity indices, which we subsequently use as alternative dependent variables in Tables B.11 and B.16. Fig. B.2 displays the correlations between these different indices for reference.

Strict Cultural Fractionalization - Our benchmark measures of cultural fractionalization rely on the 46 cultural traits presented in Tables B.2 and B.3. While these traits encompass a wide range of cultural dimensions, we aim to establish a connection with Alesina et al. (2017), who investigate cultural convergence across European regions. In their study, they use a more condensed set of cultural traits, which are those available in Table B.2. Consequently, to ease comparability, we calculate the "strict" measure of cultural fractionalization using the limited set of questions employed by Alesina et al. (2017). It is important to note that, by construction, using a smaller set of variables (17), the resulting variability is lower compared to our benchmark measure.

Augmented Cultural Fractionalization - As previously noticed, our benchmark index of cultural fractionalization assumes that each answer to each question provides the same degree of contribution to the overall extent of cultural diversity. However, this is not necessarily the case. One way to account for this is to weigh the contribution of each answer based on the distance from the prevailing norm in a given region and year. Hence, following Greenberg (1956) for each trait $m = 1, ..., M$ in region r at year t we compute the augmented cultural fractionalization index as follows:

$$CFA^{m}_{r,t} = \sum_{i_m=1}^{I_m} s^{i_m}_{r,t}(1 - s^{i_m}_{r,t})d^{i_m}_{r,t} \tag{2.14}$$

Compared to the benchmark definition, the augmented version weighs each answer $i_m = 1, ..., I_m$ of trait m by the relative distance from the prevailing norm in the region ($d^{i_m}_{r,t}$),

which is defined as follows:

$$d_{r,t}^{i_m} = \frac{|\overline{i_{m,r,t}} - i_{m,r,t}|}{Max(i_m) - Min(i_m)} \tag{2.15}$$

$\overline{i_{m,r,t}}$ is the average norm computed for each region r and year t. By construction, the measure of distance from the prevailing norm ($d_{r,t}^{i_m}$) spans from 0 to 1, and higher values imply a further distance from the prevailing norm. By computing the augmented cultural fractionalization index across all the traits M and averaging them out, we get the overall augmented cultural fractionalization index.

Discretized Cultural Fractionalization - As described by Tables B.2 and B.3, the span of available answers for each cultural trait is quite heterogeneous, from traits that allow only two answers (e.g., RE1) to traits that allows eleven answers (e.g, TT1). To assess whether the variation in regional cultural heterogeneity primarily arises from the construction of these traits rather than genuine shifts in respondents' views and values, we reduce the dimensionality of the set of available answers for traits with more than two options using the following criteria.[39] First, for traits that offer four answers ("Strongly Agree", "Agree", "Disagree", "Strongly Disagree"), we discretize them by combining responses into two categories: those who "Agree" and those who "Disagree".[40] Second, for traits providing answers on the frequency of certain activities, we discretize them with a dummy equal to one if they do it once or more per month, and zero otherwise.[41] Third, for traits that provide answers on a scale from 0 to 10, we reduce their dimensionality by categorizing responses into three groups: those answering from 0 to 3, those from 4 to 6, and those from 7 to 10.[42] Fourth, for traits asking whether something is important/not important for the respondent, we discretize them with a dummy equal to one if it is important/very important for the respondent, and zero otherwise.[43] Finally, for those traits providing four answers, we discretize them by combining the answers in just two blocks.[44] By computing our overall measure of cultural heterogeneity using these discretized answers, we are less susceptible to capturing variability driven solely by the measurement framework underlying each trait.

[39] It is worth noting that this issue is strongly mitigated by the use of panel data, which compare variations across waves, and thus should be less affected by the definition of the variables.
[40] Traits: SM1 and RS1.
[41] Traits: RE3 and RE4.
[42] Traits: RE2, RS2, CK1, CK2, CK3, TT1, TT2, TT3, TT4, TT5, TT6, AM4, AM5 and AM6.
[43] Traits: SM2, SM3 and all the traits available in the Openness block.
[44] Traits: PP1, AM1, AM2, AM3.

Rosenbluth Index of Cultural Heterogeneity - As noted by Hall and Tideman (1967), in the cultural fractionalization index each answer within each trait is weighted by the share of the population holding that specific answer, implying that the relative share of respondents is more important than the absolute number of available answers in determining the degree of cultural heterogeneity. Nonetheless, the number of available answers is indeed a relevant aspect to take into account. We do partially account for this issue by discretizing the set of available answers for each trait in the Discretized Cultural Fractionalization index. An alternative way to deal with this issue is to rely on the so-called Rosenbluth Index or Hall and Tideman index (Hall and Tideman, 1967). For each cultural trait $m = 1, ..., M$, we construct the Rosenbluth Index as follows:

$$RB^m = \frac{1}{\left(2\sum_{i_m=1}^{I} r^{i_m} s^{i_m}\right) - 1} \tag{2.16}$$

The Rosenbluth index accounts for the rank of each answer (r^{i_m}) from the least used ($r^{i_m} = 1$) to the one that is mostly diffused in our setting. It is important to notice that the ranking is generated from the least to the most diffused answer. We then construct the average overall Rosenbluth index by averaging out the trait-specific Rosenbluth Indices.

Entropy Index of Cultural Heterogeneity - An alternative measure of cultural heterogeneity can be derived from the Entropy Index proposed by Shannon (1948). Such a measure aims to capture the degree of chaos of a specific system: the higher the value, the higher the uncertainty or the complexity of the system. Translating this type of measurement in our setting implies that higher values are associated with more cultural heterogeneity. We then compute the average overall degree of cultural entropy (CE) index across the various cultural traits $m = 1, ..., M$ as follows:

$$CE = \frac{1}{M}\sum_{m=1}^{M} CE^m = \frac{1}{M}\sum_{m=1}^{M}\left(-\sum_{i_m=1}^{I} s^{i_m} ln(s^{i_m})\right) \tag{2.17}$$

As Fig. B.2 shows, our measure of cultural entropy is positive and highly correlated with our measure of cultural fractionalization.

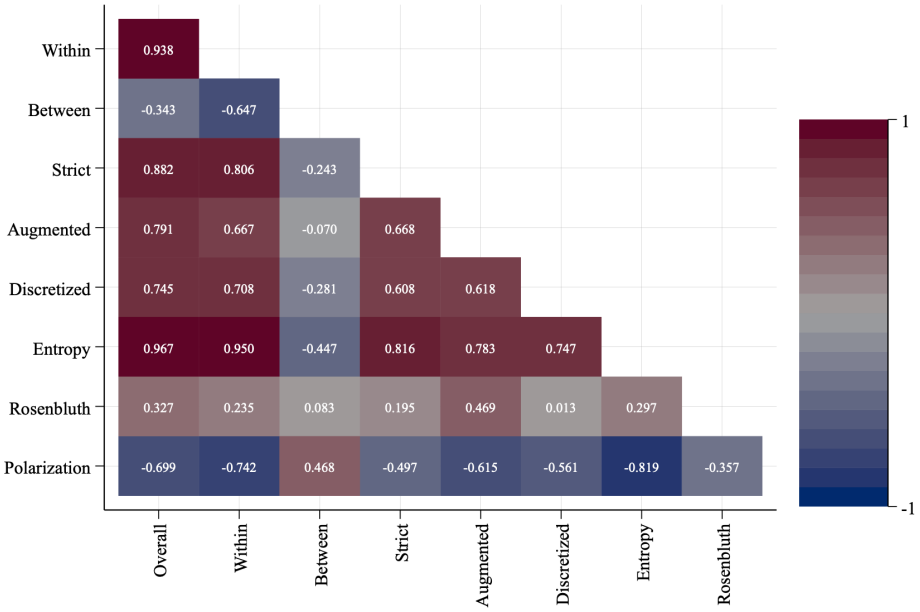
Cultural Polarization - The measure of cultural fractionalization captures the overall degree of cultural heterogeneity within a region. Another relevant index that can be computed is the overall degree of cultural polarization within a region. By relying on Montalvo and Reynal-Querol (2005), we construct a cultural polarization index for each

cultural trait $m = 1, ..., M$ which captures the closeness to a bimodal distribution of the trait m in each region r at year t. The Polarization Index is computed as follows:

$$PL^m = 1 - \sum_{i_m=1}^{I} \left(\frac{0.5 - s^{i_m}}{0.5} \right)^2 s^{i_m} \tag{2.18}$$

By averaging out the cultural polarization indices across our 46 variables, we then get an average overall measure of cultural polarization at the regional level. This index is negatively correlated with the overall index of fractionalization.

Fig. B.2. Cross-correlations across alternative indices.



Source: Authors' elaboration on ESS data (2004-2018).

B.1.4 Summary statistics

Table B.4. Summary statistics.

	Mean	Std. Dev.	Min.	Max.
Cultural heterogeneity:				
Overall (CF)	0.731	0.021	0.574	0.784
Within (CF^W)	0.722	0.025	0.547	0.780
Betwenn (F_{ST})	0.012	0.013	0.000	0.158
Religiosity	0.659	0.060	0.220	0.752
Sexual morality	0.723	0.035	0.463	0.793
Role of the state	0.723	0.046	0.442	0.809
Cultural capital	0.776	0.024	0.580	0.825
Political engagement	0.514	0.038	0.259	0.576
Trust institutions	0.853	0.022	0.636	0.894
Att. Migration	0.745	0.029	0.604	0.814
Openness	0.722	0.031	0.480	0.790
Share of immigrants:				
All (m_{rt})	0.101	0.094	0.000	0.654
High-skilled	0.027	0.034	0.000	0.286
Low skilled	0.074	0.064	0.000	0.407
Outside Europe	0.061	0.058	0.000	0.390
Within Europe	0.040	0.046	0.000	0.322
Less than 6 years	0.020	0.024	0.000	0.193
From 6 to 10 years	0.018	0.021	0.000	0.142
More than 10 years	0.061	0.059	0.000	0.363
Controls:				
ln(Density)	5.620	1.134	1.909	9.612
ln(GDP per capita)	10.688	0.643	8.516	12.114
Unemployment rate	8.955	5.007	1.193	34.800
Share of High-skilled	27.753	8.999	6.800	58.400

Source: Authors' elaboration on ESS, EULFS, and Eurostat data (2004-2018).

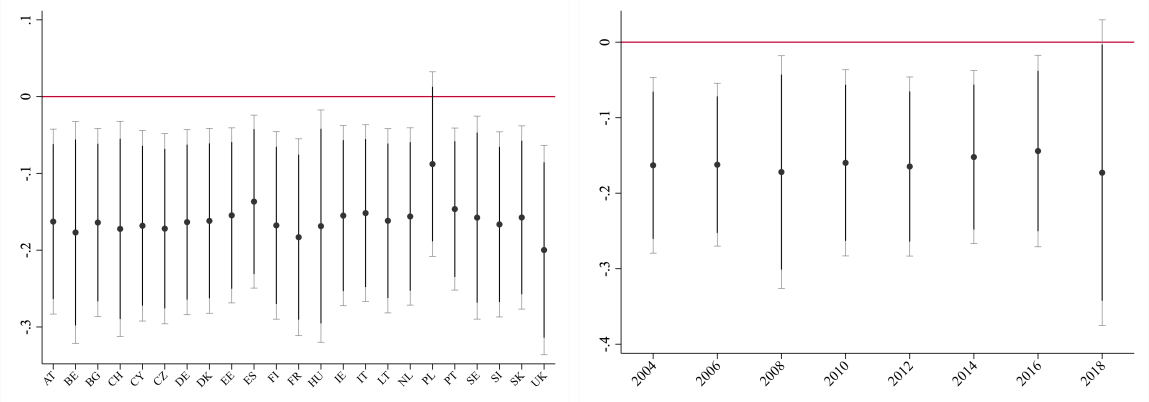
B.2 Robustness checks

Table B.5
OLS and 2SLS estimates including ESS 2020.

	(1) OLS	(2) OLS	(3) OLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
m_{rt}	-0.004	-0.079***	-0.084***	0.009	-0.133***	-0.151**
	(0.011)	(0.025)	(0.025)	(0.011)	(0.046)	(0.061)
Regional FE	No	Yes	Yes	No	Yes	Yes
Year FE	No	Yes	Yes	No	Yes	Yes
Regional controls	No	No	Yes	No	No	Yes
Observations	1,333	1,333	1,333	1,333	1,333	1,333
Mean Cultural Index	0.731	0.731	0.731	0.731	0.731	0.731
Mean Immig. Share	0.103	0.103	0.103	0.103	0.103	0.103
First-stage				0.747	1.444	1.307
KP F-Test				781.968	66.975	40.506

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variable is CF_{rt} the measure of cultural heterogeneity in the region r at time t. The independent variable m_{rt} is the share of foreign-born in the total 2004 population. Regional controls include the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population.
Source: Authors' elaboration on ESS and EULFS data (2004-2020).

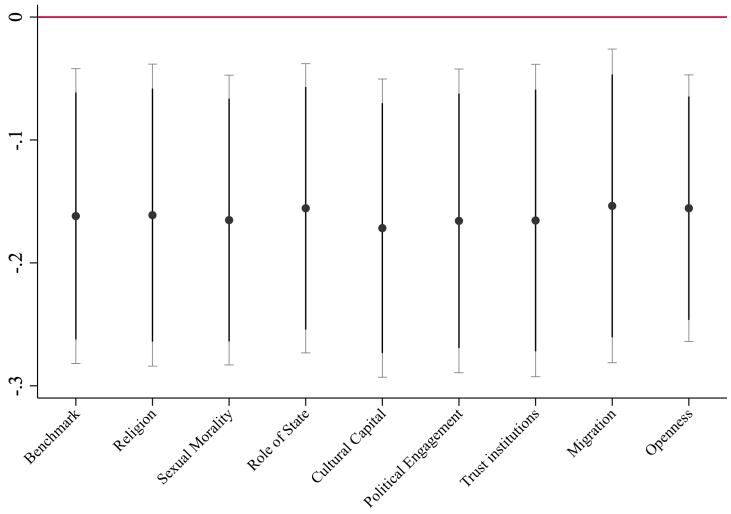
Fig. B.3. 2SLS estimates dropping countries and years one by one.



(a) Dropping countries **(b)** Dropping years

Notes: These figures depict the coefficients obtained from estimating Equation (2.6) dropping countries one by one in Figure (a) and years in Figure (b). The dependent variable is CF_{rt} the measure of cultural heterogeneity. The independent variable is the share of foreign-born in the total 2004 population. We include 95% and 90% confidence intervals around the estimated coefficients. We control for the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population. All estimates include regional and year fixed effects. Standard errors are clustered at the regional level.
Source: Authors' elaboration on ESS and EULFS data (2004-2018).

Fig. B.4. Dropping cultural blocks.



Notes: This figure depicts the coefficients obtained from estimating Equation (2.6) dropping cultural blocks one by one. The dependent variable is CF_{rt} the measure of cultural heterogeneity. The independent variable is the share of foreign-born in the total 2004 population. We include 95% and 90% confidence intervals around the estimated coefficients. We control for the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population. All estimates include regional and year fixed effects. Standard errors are clustered at the regional level.
Source: Authors' elaboration on ESS and EULFS data (2004-2018).

Table B.6
OLS and 2SLS estimates including second-generation immigrants.

	(1) OLS	(2) OLS	(3) OLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
m_{rt}	0.002	-0.098***	-0.109***	0.017	-0.144***	-0.170***
	(0.011)	(0.028)	(0.028)	(0.012)	(0.044)	(0.060)
Regional FE	No	Yes	Yes	No	Yes	Yes
Year FE	No	Yes	Yes	No	Yes	Yes
Regional controls	No	No	Yes	No	No	Yes
Observations	1,235	1,235	1,235	1,235	1,235	1,235
Mean Cultural Index	0.731	0.731	0.731	0.731	0.731	0.731
Mean Immig. Share	0.101	0.101	0.101	0.101	0.101	0.101
First-stage				0.740	1.465	1.292
KP F-Test				905.098	82.890	48.528

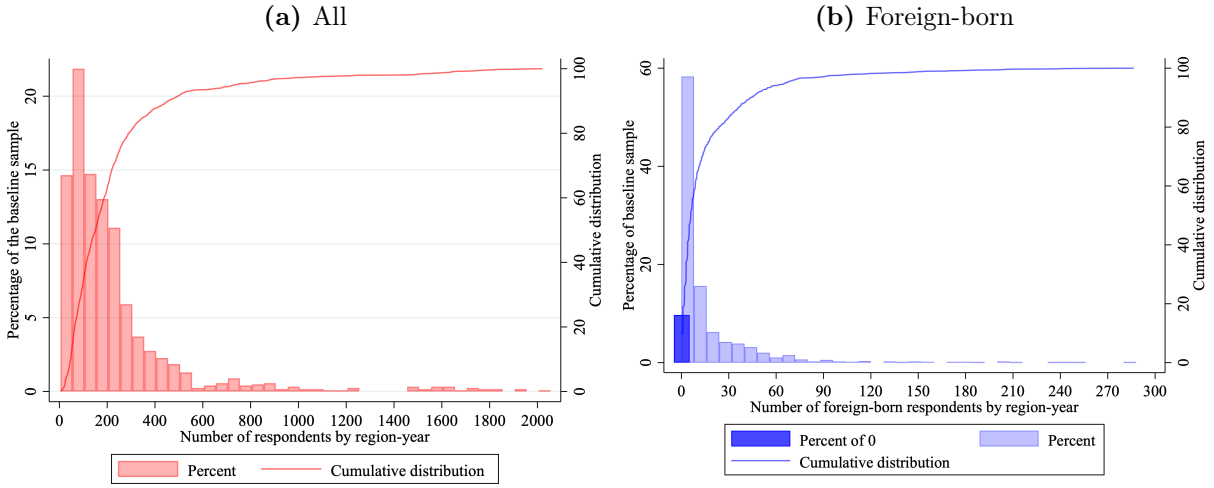
Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variable is CF_{rt} the measure of cultural heterogeneity in the region r at time t. The independent variable m_{rt} is the share of foreign-born in the total 2004 population. Regional controls include the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population. The sample includes natives, first, and second-generation immigrants.
Source: Authors' elaboration on ESS and EULFS data (2004-2018).

Table B.7
2SLS estimates excluding regions with no migrants.

		Excluding no mig.		
	(1) Benchmark	(2) EULFS	(3) ESS	(4) EULFS & ESS
m_{rt}	-0.162***	-0.162***	-0.124**	-0.124**
	(0.061)	(0.061)	(0.061)	(0.061)
Observations	1,235	1,234	1,111	1,111
Mean Cultural Index	0.731	0.731	0.732	0.732
Mean Immig. Share	0.101	0.101	0.110	0.110
First-stage	1.292	1.292	1.265	1.265
KP F-Test	48.528	48.513	36.920	36.920

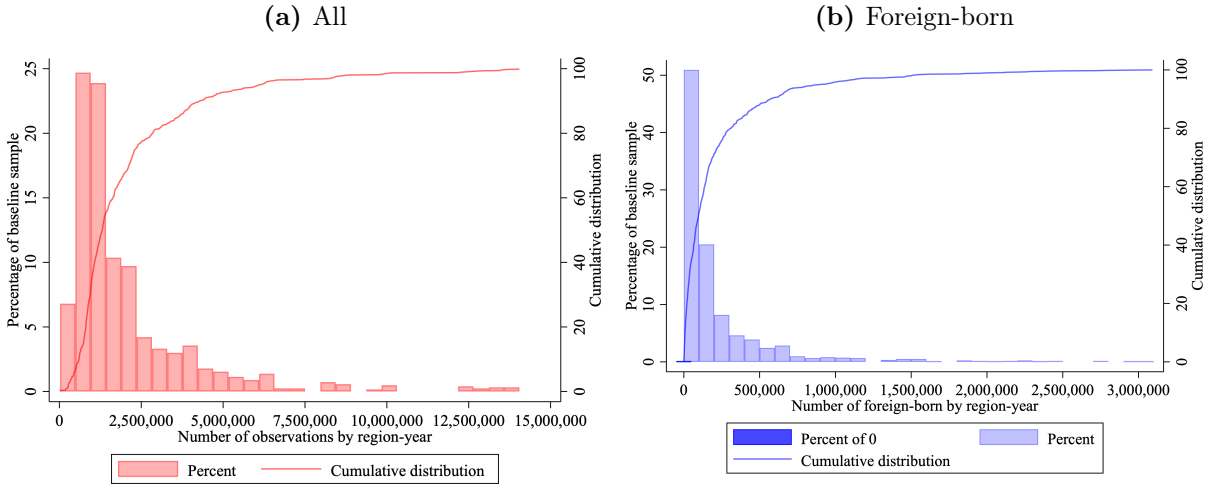
Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variable is CF_{rt} the measure of cultural heterogeneity in the region r at time t. The independent variable m_{rt} is the share of foreign-born in the total 2004 population. Regional controls include the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population. We add regional and year fixed effects.
Source: Authors' elaboration on data from the ESS and the EULFS (2004 to 2018).

Fig. B.5. Distribution of the number of observations - ESS.



Notes: These figures depict the distribution of the number of observations in the European Social Survey (ESS) at the region-year level for the overall population in Figure (a) and the foreign-born population in Figure (b).
Source: Authors' elaboration on ESS data (2004-2018).

Fig. B.6. Distribution of the number of observations - EULFS.



Notes: These figures depict the distribution of the number of observations in the European Labor Force Survey (EULFS) at the region-year level for the overall population in Figure (a) and the foreign-born population in Figure (b).
Source: Authors' elaboration on EULFS data (2004-2018).

Table B.8
2SLS estimates excluding regions with small number of observations.

	Benchmark	Excluding maximum non-missing		Excluding average non-missing	
	(1)	(2)	(3)	(4)	(5)
	All obs.	Obs.<50	Obs.<100	Obs.<50	Obs.<100
m_{rt}	-0.162***	-0.143***	-0.099*	-0.118**	-0.055
	(0.061)	(0.054)	(0.052)	(0.049)	(0.037)
Observations	1,235	1,097	822	1,083	806
Mean Cultural Index	0.731	0.734	0.736	0.734	0.737
Mean Immig. Share	0.101	0.100	0.102	0.100	0.103
First-stage	1.292	1.302	1.303	1.316	1.301
KP F-Test	48.528	46.346	48.014	46.147	46.553

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variable is CF_{rt} the measure of cultural heterogeneity in the region r at time t. The independent variable m_{rt} is the share of foreign-born in the total 2004 population. Regional controls include the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population. We add regional and year fixed effects.
Source: Authors' elaboration on ESS and EULFS data (2004-2018).

Table B.9
2SLS estimates - interaction with small regions dummies.

	Benchmark	Maximum non-missing		Average non-missing	
	(1)	(2)	(3)	(4)	(5)
m_{rt}	-0.162***	-0.166***	-0.166***	-0.162***	-0.164***
	(0.061)	(0.060)	(0.061)	(0.059)	(0.059)
<50 obs.		-0.014*		-0.015***	
		(0.008)		(0.005)	
$m_{rt} \times <50$ obs.		0.012		0.017	
		(0.061)		(0.039)	
<100 obs.			-0.008***		-0.010**
			(0.003)		(0.005)
$m_{rt} \times <100$ obs.			-0.014		0.001
			(0.019)		(0.024)
Observations	1,235	1,235	1,235	1,235	1,235
Mean Cultural Index	0.731	0.731	0.731	0.731	0.731
Mean Immig. Share	0.101	0.101	0.101	0.101	0.101
First-stage (m_{rt})	1.292	1.295	1.280	1.304	1.282
First-stage ($m_{rt} \times obs.$)		0.709	0.761	0.721	0.761
KP F-Test	48.528	24.910	32.866	25.339	33.094

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variable is CF_{rt} the measure of cultural heterogeneity in the region r at time t. The independent variable m_{rt} is the share of foreign-born in the total 2004 population. Regional controls include the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population. We add regional and year fixed effects.
Source: Authors' elaboration on ESS and EULFS data (2004-2018).

Table B.10
2SLS estimates with additional fixed effects.

	(1) Benchmark	(2) Dropping NUTS0	(3) Country-Year FE	(4) Geography-Year FE	(5) Enlargement-Year FE	(6) Welfare-Year FE
m_{rt}	-0.162***	-0.154***	-0.016	-0.136**	-0.158*	-0.141***
	(0.061)	(0.057)	(0.063)	(0.058)	(0.088)	(0.054)
Observations	1,235	1,208	1,208	1,235	1,235	1,234
Mean Cultural Index	0.731	0.731	0.731	0.731	0.731	0.731
Mean Immig. Share	0.101	0.100	0.100	0.101	0.101	0.101
First-stage	1.292	1.378	1.147	1.235	1.118	1.264
KP F-Test	48.528	60.186	24.134	40.013	22.903	41.805

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variable is CF_{rt} the measure of cultural heterogeneity in the region r at time t. The independent variable m_{rt} is the share of foreign-born in the total 2004 population. Regional controls include the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population. We add regional and year fixed effects.
Source: Authors' elaboration on ESS and EULFS data (2004-2018).

Table B.11
2SLS estimates using alternative indices.

	(1) Benchmark	(2) Strict	(3) Augmented	(4) Discretized	(5) Entropy	(6) Rosenbluth
m_{rt}	-0.162***	-0.149**	-0.075***	-0.219***	-0.430***	-0.018**
	(0.061)	(0.069)	(0.027)	(0.072)	(0.158)	(0.009)
Observations	1,235	1,235	1,235	1,235	1,235	1,235
Mean Cultural Index	0.731	0.729	0.161	0.456	1.545	0.136
Mean Immig. Share	0.101	0.101	0.101	0.101	0.101	0.101
First-stage	1.292	1.292	1.292	1.292	1.292	1.292
KP F-Test	48.528	48.528	48.528	48.528	48.528	48.528

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variables are the overall cultural heterogeneity in (1), the strict cultural heterogeneity in (2), the augmented cultural heterogeneity in (3), the discretized cultural heterogeneity in (4), the Entropy diversity index in (5), and the Rosenbluth diversity index in (6). The independent variable m_{rt} is the share of foreign-born in the total 2004 population. Regional controls include the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population. We add regional and year fixed effects.

Source: Authors' elaboration on ESS and EULFS data (2004-2018).

B.3 Validity of the instrument

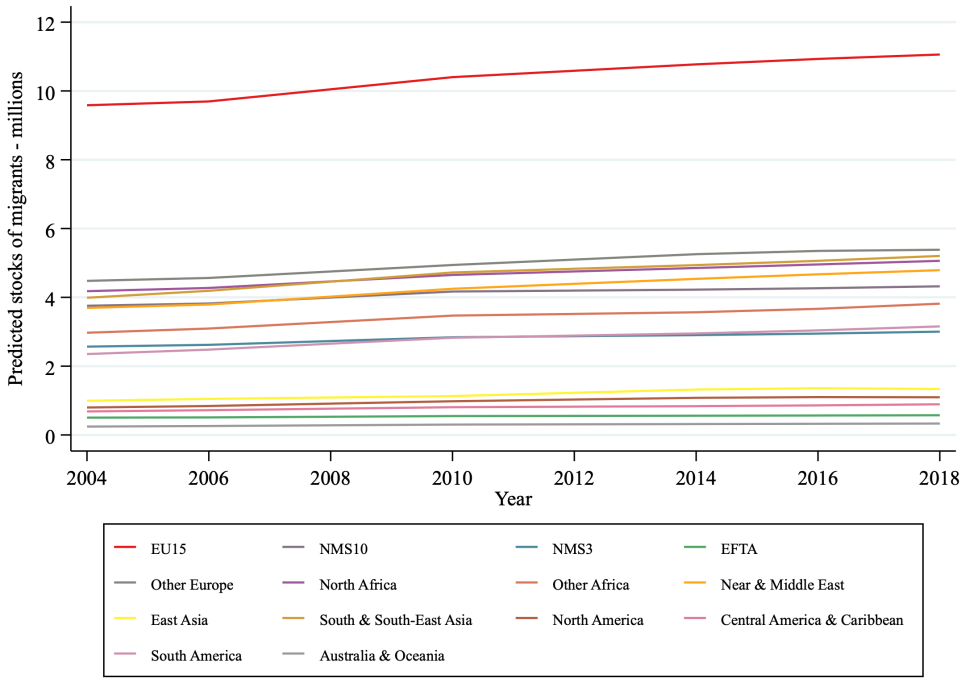
Table B.12
Zero-stage bilateral migration gravity model estimates.

	(1) Stock of immigrants
ln(Deaths)	0.026***
	(0.006)
Disasters	0.004***
	(0.001)
ln(Distance) × 1995	-0.085***
	(0.031)
ln(Distance) × 2000	-0.113***
	(0.036)
ln(Distance) × 2005	-0.059**
	(0.025)
ln(Distance) × 2010	-0.030
	(0.020)
ln(Distance) × 2015	-0.018**
	(0.009)
Year FE	Yes
Destination × Year FE	Yes
Origin × Destination FE	Yes
Observations	78,561
R-squared	0.987

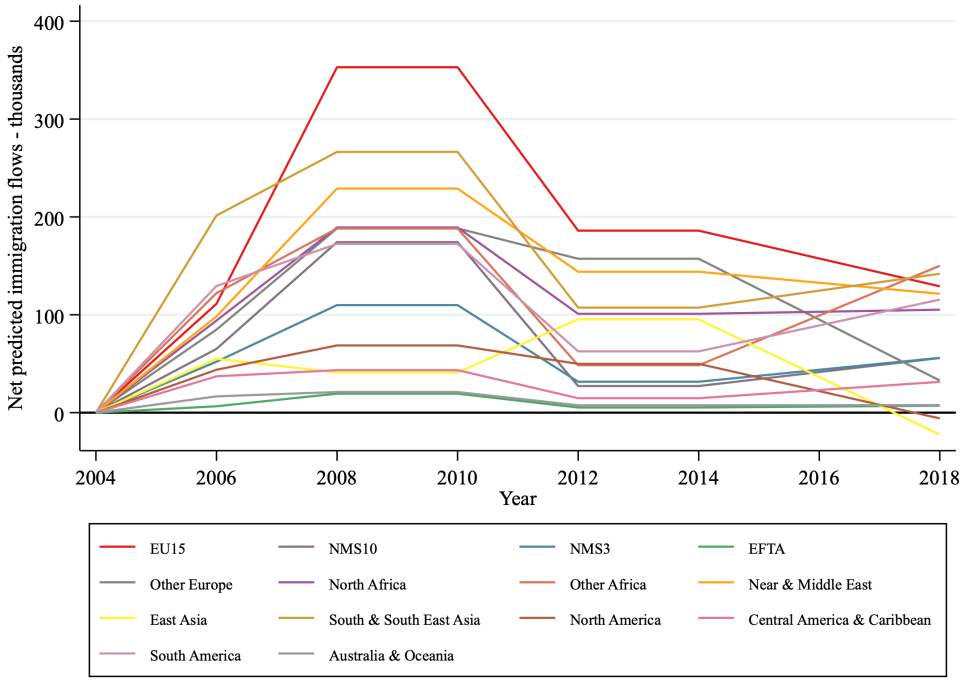
Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the country-pair level. The dependent variable is the stocks of immigrants from origin o in destination d at year t.
Source: Authors' elaboration on data from the United Nations, the UCDP/PRIO Armed Conflict Dataset, and the Emergency Events Database (EM-DAT) (1990-2020).

Fig. B.7. The variation of the predicted stocks of migrants by origin-group.

(a) Predicted immigration stocks



(b) Net predicted immigration flows



Notes: These figures depict the variation of the predicted stocks of migrants for each origin group. The predicted immigration stocks are displayed in (a), while the predicted net immigration flows with two years lagged are displayed in (b).
Source: Authors' elaboration on the United Nations Population Division data (2004-2018).

Table B.13
2SLS estimates using alternative shift-share instruments.

	(1) Standard	(2) Standard	(3) Leave-one-out	(4) Leave-one-out
m_{rt}	-0.163***	-0.207***	-0.147***	-0.178***
	(0.055)	(0.077)	(0.044)	(0.059)
Regional FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Regional controls	No	Yes	No	Yes
Observations	1,235	1,235	1,235	1,235
Mean Cultural Index	0.731	0.731	0.731	0.731
Mean Immig. Share	0.101	0.101	0.101	0.101
First-stage	0.739	0.633	1.754	1.551
KP F-Test	77.818	39.889	130.558	62.971

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variable is CF_{rt} the measure of cultural heterogeneity in the region r at time t. The independent variable m_{rt} is the share of foreign-born in the total 2004 population. Regional controls include the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population.

Source: Authors' elaboration on ESS and EULFS data (2004-2018).

Table B.14
Pre-trend analysis.

	Instrument growth				
	(1) All	(2) HS	(3) LS	(4) EU28	(5) NEU28
GDP per capita growth	-0.004	0.048	-0.021	-0.001	0.006
	(0.008)	(0.050)	(0.017)	(0.004)	(0.006)
P-value	0.623	0.336	0.211	0.758	0.422
t-stat	-0.506	0.956	-1.209	-0.379	0.888
Population density growth	0.058*	-0.062	0.088	0.020	0.037
	(0.031)	(0.137)	(0.067)	(0.013)	(0.034)
P-value	0.293	0.777	0.462	0.329	0.596
t-stat	1.858	-0.450	1.324	1.461	1.112
Unemployment rate growth	0.002	-0.004	0.004	0.000	-0.000
	(0.002)	(0.003)	(0.003)	(0.000)	(0.002)
P-value	0.352	0.159	0.192	0.458	0.802
t-stat	0.993	-1.452	1.439	0.834	-0.225
Tertiary education growth	-0.006	0.005	-0.009	0.001	-0.003
	(0.003)	(0.011)	(0.007)	(0.001)	(0.003)
P-value	0.035	0.866	0.185	0.681	0.233
t-stat	-1.763	0.429	-1.286	0.423	-1.117
Country FE	✓	✓	✓	✓	✓

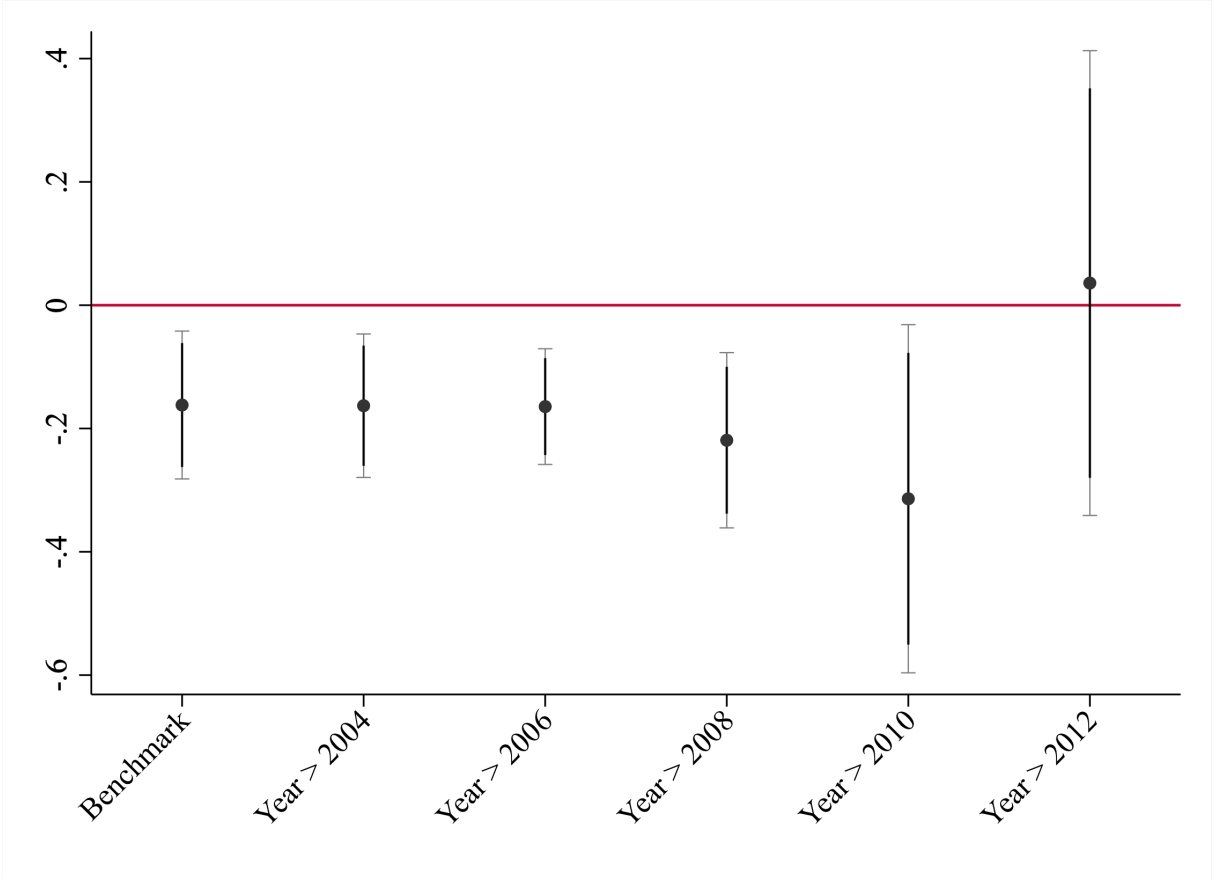
Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the country level. This table shows the coefficients of regressing the regional predicted migration growth (shift-share instrument) over the period 2004 and 2007 on the growth rate of regional economic indicators between 2000 and 2003. We report the p-value and t-stat of the wild cluster bootstrap (999 replications) with Webb weights test.
Source: Authors' elaboration on ESS, EULFS, and Eurostat data (2000-2007).

Table B.15
Adão et al. (2019) inference procedure.

	Coefficient	Std. error	P-value	Confidence Interval
Second-stage				
Robust	-0.162	0.045	0.0004	[-0.251,-0.073]
Cluster	-0.162	0.061	0.0084	[-0.282,-0.042]
AKM	-0.162	0.044	0.0003	[-0.249,-0.075]
First-stage				
Robust	1.292	0.122	0.0000	[1.053,1.531]
Cluster	1.292	0.185	0.0000	[0.928,1.656]
AKM	1.292	0.124	0.0000	[1.048,1.536]

Notes: This table reports the first and second stages benchmark coefficients, standard errors, p-values, and confidence intervals using various inference methods. Robust refers to robust standard errors. Cluster refers to clustered standard errors at the regional level. AKM refers to the inference procedure described in Adão et al. (2019). The dependent variable is CF_{rt} the measure of cultural heterogeneity in the region r at time t. The independent variable is the share of foreign-born in the total 2004 population. We control for the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population. All estimates include regional and year fixed effects.
Source: Authors' elaboration on ESS and EULFS data (2004-2018).

Fig. B.8. 2SLS estimates dropping years sequentially.



Notes: This figure depicts the coefficients obtained from estimating Equation (2.6) dropping years sequentially until 2012. The dependent variable is CF_{rt} the measure of cultural heterogeneity in the region r at time t. The independent variable is the share of foreign-born in the total 2004 population. We include 95% and 90% confidence intervals around the estimated coefficients. We control for the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population. All estimates include regional and year fixed effects. Standard errors are clustered at the regional level.
Source: Authors' elaboration on ESS and EULFS data (2004-2018).

B.4 Additional results

Table B.16
Polarization.

		Within-native heterogeneity			Between-native heterogeneity		
	(1) Polarization	(2) Education	(3) Urbanization	(4) Income	(5) Education	(6) Urbanization	(7) Income
m_{rt}	0.077**	-0.215***	-0.155***	-0.141***	0.018	-0.000	-0.020
	(0.037)	(0.067)	(0.050)	(0.052)	(0.025)	(0.017)	(0.046)
Observations	1,235	1,231	1,181	1,204	1,231	1,181	1,204
Mean Cultural Index	0.671	0.713	0.721	0.711	0.022	0.013	0.025
Mean Immig. Share	0.101	0.101	0.100	0.102	0.101	0.100	0.102
First-stage	1.292	1.291	1.307	1.306	1.291	1.307	1.306
KP F-Test	48.528	48.420	48.361	49.566	48.420	48.361	49.566

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variables are the Polarization index in Column (1), the within-native groups cultural heterogeneity in Columns (2) to (4), and the between-native groups cultural heterogeneity in Columns (5) to (7). The independent variable m_{rt} is the share of foreign-born in the total 2004 population. We control for the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population. All estimates include regional and year fixed effects.

Source: Authors' elaboration on ESS and EULFS data (2004-2018).

Table B.17
Measures of cultural heterogeneity - natives and second-generation immigrants.

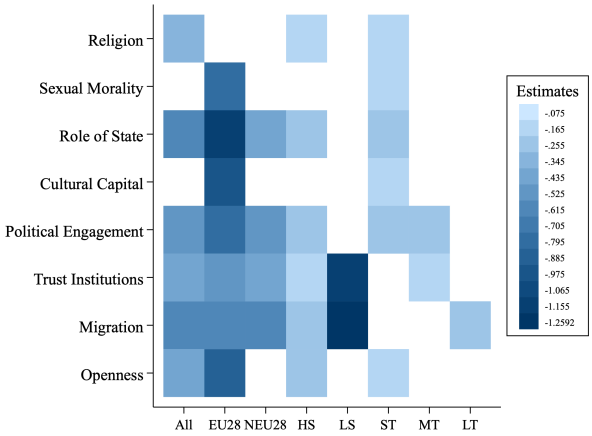
	(1) Overall	(2) Within-group	(3) Between-group
m_{rt}	-0.195***	-0.208***	0.019
	(0.066)	(0.064)	(0.023)
Observations	1,235	1,235	1,235
Mean Cultural Index	0.729	0.722	0.010
Mean Immig. Share	0.101	0.101	0.101
First-stage	1.292	1.292	1.292
KP F-Test	48.528	48.528	48.528

Notes: *** p <0.01, ** p <0.05, * p < 0.1. Standard errors in parentheses are clustered at the regional level. The dependent variable in Column (1) is CF_{rt} the measure of cultural heterogeneity in the region r at time t. The dependent variable in Column (2) is CF_{rt}^{W} the measure of within-group cultural heterogeneity in the region r at time t. The dependent variable in Column (3) is F_{ST} the measure of between-group cultural heterogeneity in the region r at time t. The independent variable m_{rt} is the share of foreign-born in the total 2004 population. We control for the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population. All estimates include regional and year fixed effects. The sample includes natives and second-generation immigrants.

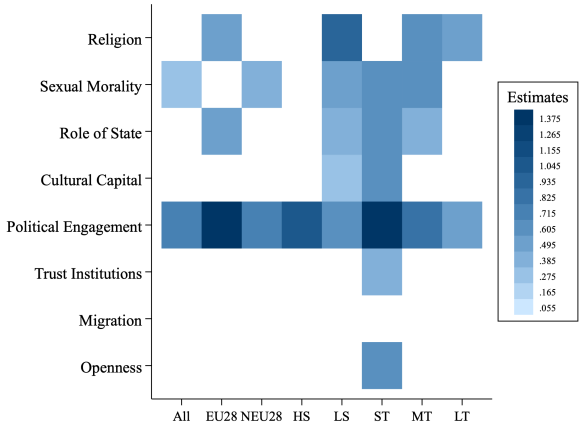
Source: Authors' elaboration on ESS and EULFS data (2004-2018).

Fig. B.9. Cultural traits and immigrants' characteristics.

(a) Within-group heterogeneity

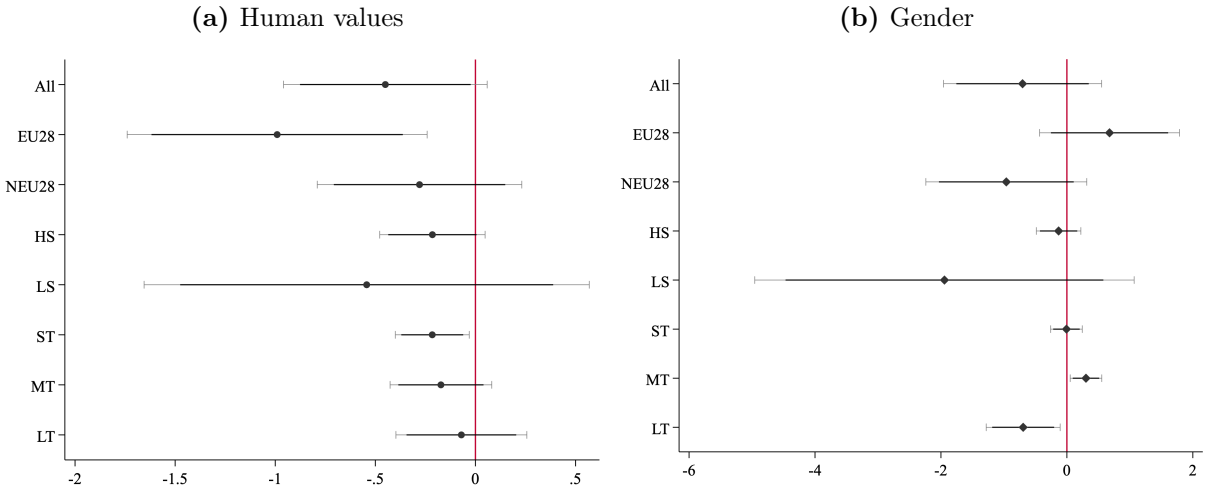


(b) Between-group heterogeneity



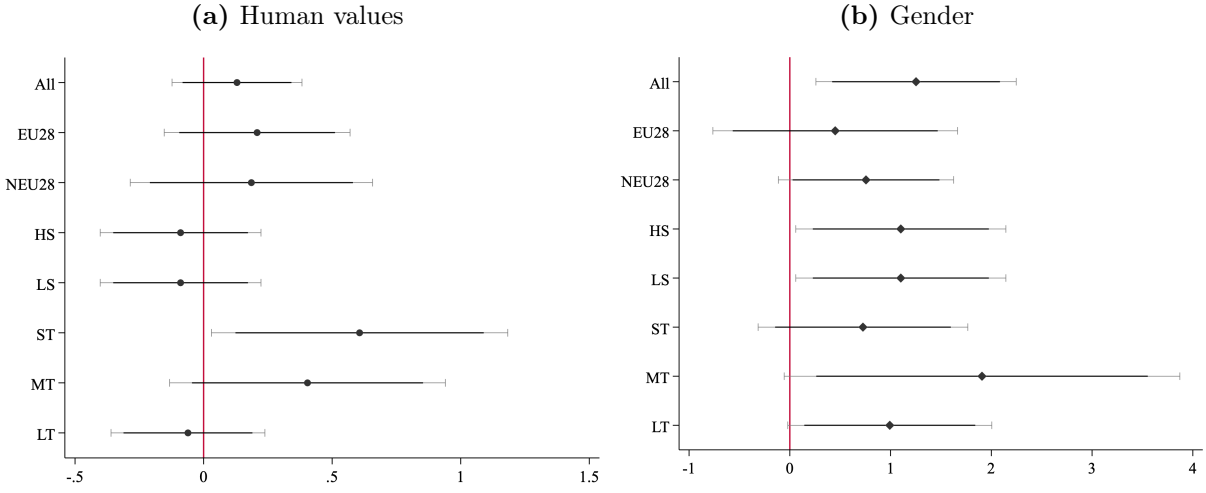
Notes: Each squared cell in Figure (a) represents the estimated coefficient from regressing the measure of within-group heterogeneity for each cultural block reported on the y-axis on the share of foreign-born belonging to the group reported on the x-axis. Each squared cell in Figure (b) represents the estimated coefficient from regressing the measure of between natives and first-generation immigrants' group reported on the x-axis heterogeneity for each cultural block reported on the y-axis on the share of foreign-born in the total 2004 population. Blank cells represent insignificant coefficients at the 10% level. Each shaded cell corresponds to a significant coefficient at the 10% level, and darker cells mean larger coefficients. We control for the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population. All estimates include regional and year fixed effects. Standard errors are clustered at the regional level.
Source: Authors' elaboration on ESS and EULFS data (2004-2018).

Fig. B.10. Within-group heterogeneity - additional cultural blocks.



Notes: These figures depict the estimated coefficients obtained from regressing the measure of within-group cultural heterogeneity for human values cultural block in (a) and gender cultural block in (b) on the share of foreign-born belonging to the group reported on the y-axis. We control for the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population. All estimates include regional and year fixed effects. Standard errors are clustered at the regional level.
Source: Authors' elaboration on ESS and EULFS data (2004-2018).

Fig. B.11. Between-group heterogeneity - additional cultural blocks.



Notes: These figures depict the estimated coefficients obtained from regressing the measure of between natives and first-generation immigrants' group reported on the y-axis cultural heterogeneity for human values cultural block in (a) and gender cultural block in (b) on the share of foreign-born in the total 2004 population. We control for the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population. All estimates include regional and year fixed effects. Standard errors are clustered at the regional level.
Source: Authors' elaboration on ESS and EULFS data (2004-2018).

B.5 Individual analysis

This appendix takes advantage of the individual-level dimension available in the European Social Survey to examine i) to which extent first-generation immigrants indeed exhibit differences in their cultural traits compared to the native population, which would be consistent with the aforementioned compositional mechanism, and ii) the directions in which cultural traits evolve in response to immigration.

It is important to emphasize that these estimates are by no means causal but should be only viewed as suggestive additional evidence supporting the notion of evolving cultural traits associated with specific immigrant characteristics in destination countries. Moreover, as previously mentioned, the question of how individual-level shifts contribute to cultural transformations at the aggregate level is complex and difficult to apprehend without further theoretical insights. Particularly, average change may be silent on aggregated effects when there is strong heterogeneity in responses or changes pulling in opposing directions. Still, estimates at the individual level have the advantage of mitigating the impact of limited observations in the European Social Survey (ESS) when constructing cultural diversity indices at the region-year level.

Keeping these potential caveats in mind, for each cultural trait, we first construct two different indices by averaging the scores of all questions within a cultural block and by computing a composite index through a Principal Component Analysis (PCA) to capture the common underlying factor within each cultural block. It is worth noticing that all separated questions have been recoded such that the correlation of questions within each cultural block is always positive. The interpretation of each dimension is therefore provided in Table B.18.

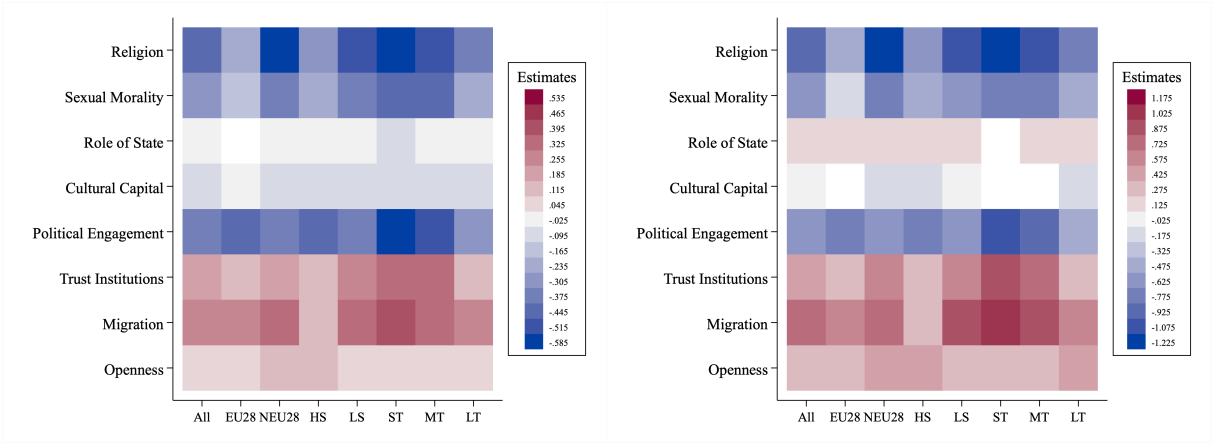
Table B.18. Interpretation of cultural blocks.

Cultural Block	Higher Value	Lower value
Religion	Less religiosity	More religiosity
Sexual Morality	Liberalism	Conservatism
Role of State	Liberalism	Conservatism
Cultural Capital	Liberalism	Conservatism
Political Engagement	Higher engagement	Lower engagement
Trust Institutions	More trust	Less Trust
Migration	Positive attitudes	Negative attitudes
Openness	More Open	Less Open

Source: Authors' elaboration on ESS data (2004-2018). .

Fig. B.12. Immigrants characteristics' and compositional effect - individual analysis.

(a) Mean **(b)** PCA



Notes: Each squared cell represents the estimated coefficient obtained from regressing first-generation immigrants and their characteristics' dummies reported on the x-axis against the mean and PCA of cultural traits in each cultural block reported on the y-axis in Figures (a) and (b), respectively. Blank cells represent insignificant coefficients at the 5% level. Each shaded cell corresponds to a significant coefficient at the 5% level, and darker red cells mean larger coefficients. Individual controls include age, age-squared, gender, employment, educational attainment, marital status, presence of children, and urbanization. Regional controls include the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population. All estimates include regional and year fixed effects. Standard errors are clustered at the regional level.
Source: Authors' elaboration on ESS data (2004-2018).

Our individual-level analysis involves estimating the following specification for each cultural block:

$$Y_{i,r,t} = \alpha + \delta_1 Firstgen^g_{i,r,t} + \boldsymbol{\sigma'} \boldsymbol{Z_{i,r,t}} + \boldsymbol{\phi'} \boldsymbol{X_{r,t}} + \gamma_t + \gamma_r + \varepsilon_{i,r,t} \tag{2.19}$$

where $Y_{i,r,t}$ is the individual cultural traits (mean or PCA) of individual i in region r at year t for a given cultural block. $Firstgen^g_i$ is a dummy variable equal to 1 if the individual is a first-generation immigrant of a subgroup g and 0 if the individual is a native. $\boldsymbol{Z_{i,r,t}}$ is a vector of individual-level controls including age, age-squared, gender, employment, educational attainment, marital status, presence of children, and urbanization. $\boldsymbol{X_{r,t}}$ is a vector of regional-level controls including the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the population. γ_t and γ_r stand for year and regional fixed effects, respectively. Standard errors are clustered at the regional level. To exploit the differential effect of each immigrant characteristic, we also estimate Equation (2.19) separately for each migrant subgroup. Our results are reported in Fig. B.12(a) and Fig. B.12(b) for mean and PCA outcomes, respectively. One can see that the average immigrant exhibits significantly different cultural traits

than the average native across all dimensions. On one hand, it is clear that immigrants, on average, introduce significantly more conservative values to their destination country on sexual morality and religiosity: immigrants tend to be more religious, hold more conservative views on gay rights, and are more inclined to believe that traditions and customs must be followed. As expected, they are also less likely to be politically engaged at destination. On the other hand, immigrants tend to lean more toward left-wing political views compared to the native population, and they report a higher level of trust and more positive attitudes toward immigrants. Once more, these two figures confirm, with a few exceptions, that larger cultural differences are observed among low-skilled immigrants, coming from outside Europe and with recently arrived in the host region.

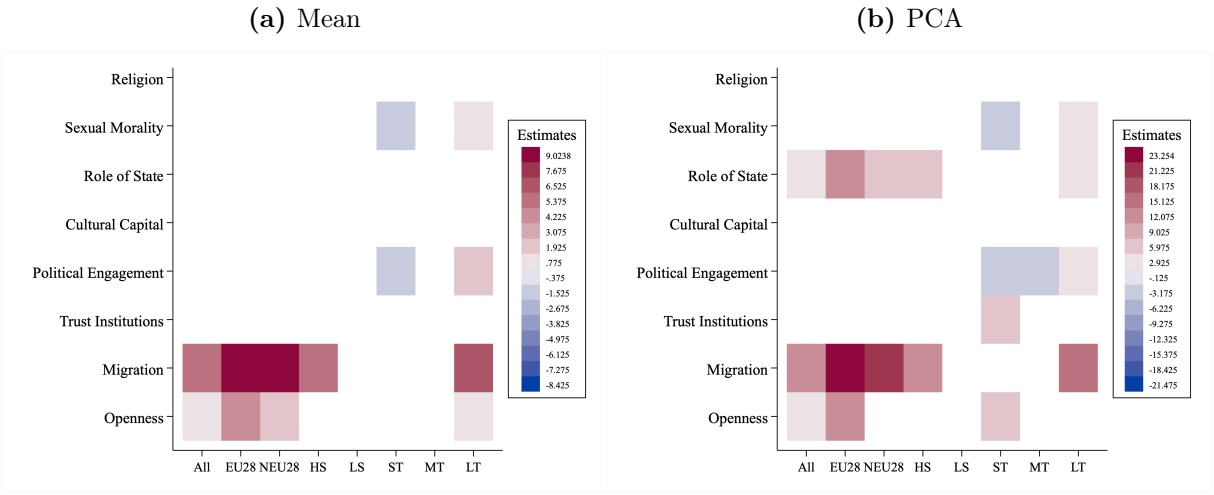
Second, we examine in which direction natives react to the arrival of new immigrants in their country. Focusing only on natives, we estimate the following equation:

$$Y_{i,r,t} = \alpha + \delta_1 m_{r,t} + \boldsymbol{\sigma'} \boldsymbol{Z_{i,r,t}} + \boldsymbol{\phi'} \boldsymbol{X_{r,t}} + \gamma_t + \gamma_r + \varepsilon_{i,r,t} \tag{2.20}$$

This equation mimics Equation (2.19), except that the first-generation indicator is replaced by the share of immigrants at the regional level $m_{r,t}$ defined in Equation (2.5).

Results are reported in Fig. B.13(a) and Fig. B.13(b) for mean and PCA, respectively. On average, these suggestive results indicate a trend of natives becoming more liberal, as evidenced by the positive coefficient associated with the presence of high-skilled immigrants, especially those from European countries on the role of state cultural block. Furthermore, the arrival of high-skilled immigrants appears to positively influence native's attitudes toward immigrants in general. Immigrants from European countries seem also to contribute to greater openness within the native population. Conversely, short-term migrants seem to have the effect of making natives less liberal in matters of sexual morality while decreasing their political engagement. Again, results on medium and long-term immigrants are not interpreted as they do exploit variations of stocks of immigrants which can be hardly interpreted as inflows of immigrants.

Fig. B.13. Immigrants' characteristics and natives' response - individual analysis.



Notes: Each squared cell represents the estimated coefficient obtained from regressing immigrants share reported on the x-axis against the mean and PCA of cultural traits in each cultural block reported on the y-axis in Figures (a) and (b), respectively. Blank cells represent insignificant coefficients at the 5% level. Each shaded cell corresponds to a significant coefficient at the 5% level; darker red cells mean larger coefficients. Individual controls include age, age-squared, gender, employment, educational attainment, marital status, presence of children, and urbanization. Regional controls include the log of population density, the log of GDP per capita, the unemployment rate, and the share of high-skilled in the resident population. All estimates include regional and year fixed effects. Standard errors are clustered at the regional level.
Source: Authors' elaboration on ESS and EULFS data (2004-2018).

Chapter 3

The Double-Edged Sword: Male Migration and Females Left Behind Labor Supply in Egypt*

3.1 Introduction

In the developing world, male-dominated migration is a prevailing trend, driven primarily by the pursuit of better job opportunities, higher incomes, and stable employment. The daunting expenses associated with family migration compel many men to embark on this journey alone leaving behind their family unit. Additionally, migration is a dynamic process, with a substantial proportion of migrants eventually returning to their home countries. According to the OECD, between 20 and 50% of international migrants make the journey back to their countries of origin or move to another country within five years of arriving at their destination (Wahba, 2021). The circular nature of male migration can have profound implications for families left in the homeland, shaping economic and social dynamics.

This paper investigates the impact of male migration, differentiating between current and return migrants, on the labor supply of females left behind in Egypt. The Egyptian context stands out as an ideal case study due to the prevalent phenomenon of men temporarily migrating, particularly to Gulf countries, while their family members remain behind. At the same time, female participation in the labor force remains consistently low. Hence, this study aims to provide a comprehensive analysis of how the male migration experience shapes female labor supply in this setting. Recognizing potential variations in female labor patterns and male migrants' characteristics across urban and rural areas, the analysis separately evaluates the impact within each geographical context.

The analysis relies on data from the last three waves of the Egypt Labor Market Panel Survey (ELMPS) conducted in 2006, 2012, and 2018. It employs a two-way fixed-effects

*I thank Simone Bertoli, Anna Gasten, Flore Gubert, Mathilde Maurel, Mariapia Mendola, Riccardo Turati, Jérôme Valette, and numerous participants at the Development Economics Sorbonne Informal Research Seminar in Paris, the International Conference in Development Economics in Bordeaux, the Department of Applied Economics Lunch Seminar at Universitat Autònoma de Barcelona, and the 9th Joint FESP-Paris 1 Research Workshop for helpful comments and suggestions.

estimation strategy to address concerns related to self-selection into migration and endogeneity bias. Taking the female labor supply when there is no migration experience in the household as a reference point, I first assess their response to the migration of a male household member and subsequently consider their reaction to the presence of a return migrant. When evaluating female labor force participation, two employment definitions are considered. The "market definition" encompasses paid or profit-driven activities, while the "extended definition" includes subsistence work performed for household consumption. Since focusing on labor force participation can obscure heterogeneous effects, the study further explores the impact on various employment statuses, including wage work, self-employment, and unpaid family work.

In urban areas, on average, the presence of a current or return migrant in the household does not significantly impact female labor force participation. Nonetheless, a more detailed analysis of their influence on female employment status reveals that living in a household where a male member is overseas increases the likelihood of females engaging in wage work. A heterogeneous analysis based on the relationship with the migrating individual reveals that the effect stems from husbands who are current migrants. In addition, this increase in wage work is predominantly driven by current migrants who do not send remittances, which reflects the need for the wife to replace the lost labor due to the migrant's absence. Moreover, combining the effects of remittance receipt and migration duration shows that the surge in wage work is not solely influenced by current migrants who do not send remittances; it is also driven by those who send remittances and have been abroad for more than four years. This latter finding suggests that the impact of long-term migration on wage work in urban settings extends beyond remittances, encompassing factors beyond a mere substitution effect. However, the presence of a return migrant has no significant effect on all employment statuses, suggesting an abolition of the migration effect.

In rural areas, the main findings display distinct patterns. Residing in a household where a male member is currently abroad, results in increased female participation in the extended labor force, primarily seen as a rise in unpaid family work. This impact is primarily driven by current migrants who are neither the husband nor the father. Moreover, this increase is largely attributed to current migrants who do not send remittances and have been living abroad for four years or less, reflecting the immediate need to fill the gap caused by their departure. However, living with a return migrant is associated with a decrease in female wage work. This decrease is mostly observed for daughters of returnees who planned their return, spent more than four years abroad, were wage workers themselves, and did

not accumulate savings overseas. In addition, when distinguishing between destination countries based on their adherence to conservative gender norms in comparison to Egypt, the decline in wage work is found to be driven by returnees from countries with more conservative norms. These findings indicate the existence of two distinct channels at play when considering return migration in rural areas: one driven by economic considerations and the other influenced by cultural factors. In summary, the study emphasizes the nuanced effects of male migration on female labor force participation and employment status, highlighting the significance of considering both migration statuses and regional contexts in understanding these dynamics.

A substantial body of literature has investigated the impact of male migration on the labor force participation of females left behind.[1] The prevailing consensus in this literature predominantly underscores a negative impact, primarily attributed to the income effect stemming from remittances. When a family member migrates, it often results in augmented household income through remittances. This increase in income may raise the reservation wage and lower the opportunity cost of leisure, leading to a greater reliance on the income provided by the male migrant. Consequently, remittances may serve as a financial safety net, potentially giving rise to moral hazard issues by discouraging non-migrant labor force participation through an insurance mechanism (Chami et al., 2005). As a result, women may be less inclined to seek employment and choose not to work.[2] One of the seminal works on this topic is the study by Funkhouser (1992), which documents strong negative effects of remittances on the labor market participation of non-migrants. Several subsequent studies corroborate this negative impact on female labor force participation, including Rodriguez and Tiongson (2001) in the Philippines, Amuedo-Dorantes and Pozo (2006) in Mexico, Acosta (2006) in ElSalvador, Lokshin and Glinskaya (2009) in Nepal, Binzel and Assaad (2011) in Egypt, Mendola and Carletto (2012) in Albania, Lenoël and David (2019) in Morocco, and Khan and Valatheeswaran (2016) and Menon and Bhagat (2020) in Kerala.

Beyond the income effect, emigration can also trigger a labor substitution effect where the absence of a male household member, often the primary breadwinner, entails a reallocation of labor within the household to compensate for the migrant's lost contributions.

[1] Notably, many of these studies focus on the labor force participation of wives when their husbands are abroad.

[2] To the best of my knowledge, only one study (Khaled, 1995) documents that husband migration leads to an increase in the labor force participation of wives, primarily driven by financial necessity. However, he found that participation declines among women whose husband's migration duration exceeds 15 years due to long-term financial stability resulting from remittances.

The majority of studies suggest that this reallocation primarily manifests as an increase in unpaid and subsistence work, shifting away from paid employment for females left behind. For instance, in rural areas, females in migrant households are more likely to engage in unpaid family work and subsistence work to maximize the usage of household assets, as observed by Binzel and Assaad (2011). This shift is driven by the need to compensate for the migrant's absence rather than relying on remittances to finance new ventures.[3] Similarly, Acosta (2006), Mendola and Carletto (2012), Karymshakov and Sulaimanova (2017), and Lenoël and David (2019) also identify an increase in females' unpaid family work. Furthermore, the absence of a male household member may result in females assuming additional roles and responsibilities, including household and caregiving duties. This can lead to increased stress and time constraints, limiting their ability to participate in the labor force. The study by Cabegin (2006) reveals that married women experience increased conjugal time and are expected to dedicate more time to child care to compensate for the absence of their migrant husbands. Additionally, Khan and Valatheeswaran (2016) document an increase in female involvement in household duties, particularly in rural areas, at the expense of unpaid family work.[4] This paper deviates from prior studies in terms of methodology. While previous research predominately relies on cross-sectional analysis and instrumental variable approaches, this paper leverages a three-wave panel dataset, substantially enhancing the estimation strategy. By including individual fixed effects, this approach effectively mitigates biases stemming from selection and omitted variables, resulting in more precise and reliable estimates.

In contrast, the literature has paid considerably less attention to the impact of the return migration of male household members on female labor supply. Similar to out-migration, return migration can exert both income and substitution effects, albeit in the opposite direction, potentially counterbalancing the migration impact. On the income effect side, the discontinuation of migrant remittances may increase financial pressure on the house-

[3] However, some studies found that remittances can alleviate budget constraints creating self-employment opportunities, for both men (Acosta, 2006; Khan and Valatheeswaran, 2016) and women (Karymshakov and Sulaimanova, 2017; Cox-Edwards and Rodríguez-Oreggia, 2009).

[4] In contrast to the aforementioned studies, Cox-Edwards and Rodríguez-Oreggia (2009) find that remittances have no significant impact on labor force participation of household members left behind in Mexico, except for a modest increase in participation among women in urban areas. They suggest that migrants, as household members, strive to replace their pre-migration contributions without causing substantial altering in the income levels of recipient households. Similarly, Kan and Aytimur (2019) find no effect of household member migration on female labor force participation in Tajikistan. They argue that the income effect from remittances, which would typically decrease female labor force participation, may be offset by various factors such as labor shortage within the household or the initial financial burden of migration that could result in the left-behind family accumulating debt.

hold, prompting females to enter the labor market to compensate for the loss of non-labor income they previously depended on. This effect becomes more pronounced if migrants return because they were unable to meet their savings goals and acquire new skills or if they faced challenges in securing immediate employment upon return. On the substitution effect side, the return of migrants can lead to a decline in female labor force participation as returnees assume their traditional responsibilities, resulting in a decrease in unpaid family work.[5] Hence, this paper's empirical contribution is to explore the impact of both current and return migration, providing a comprehensive analysis of how males' migration experience affects the labor supply of females left behind. This thorough investigation sheds light on whether the effects of male migration persist even after their return, thus enriching our understanding of this phenomenon.

Furthermore, what adds an intriguing dimension to investigating the impact of return migration lies in the potential additional effects it may entail. Previous migration experiences of household members could involve the return of both human and physical capital (McCormick and Wahba, 2001; Gubert and Nordman, 2011; El-Mallakh and Wahba, 2021; Bensassi and Jabbour, 2022), which can be redeployed or invested by household members to enter self-employment. For example, research by Mendola and Carletto (2012) suggests that females in households with members who have previously migrated are more inclined to engage in self-employment. In addition, there is a rapidly growing body of literature that explores the transmission of social norms from destination to home countries, including those related to fertility (Beine et al., 2013; Bertoli and Marchetta, 2015) and gender norms (Tuccio and Wahba, 2018; Samari, 2021). In Jordan, Tuccio and Wahba (2018) find that females with family members returning from more conservative Arab countries were more likely to adhere to traditional gender norms compared to females in non-migrant households. This adherence, in turn, was associated with lower female labor force participation, higher girls' school dropout, and higher fertility rates.

This paper diverges from Tuccio and Wahba (2018)'s study by primarily focusing on females' labor supply choices, distinguishing different employment statuses, following a male member return migration. Leveraging the rich dataset in hand, it delves into various mechanisms to disentangle the role played by economic and cultural factors, including the transmission of gender norms— an aspect often overlooked in the existing literature on male migration and female labor force participation. In Egypt, where traditional and

[5] Conversely, the return of male migrants can also potentially lead to a redistribution of household duties, thereby potentially freeing up time for females to engage in income-generating activities.

cultural norms significantly influence female labor force participation (Krafft et al., 2019), the impact on female labor supply can vary based on the conservatism or liberalism of the destination countries from which the return migrants come back and the characteristics of the geographical setting to which they return.

Overall, the impact of both current and return migration of male household members on females left behind labor supply is not obvious, involving several interconnected underlying mechanisms. Consequently, establishing a causal link between migration and female labor choices remains a challenging empirical question with inconclusive evidence. This paper aims to contribute to the existing studies by examining both types of migration, in two distinct geographical settings, disentangling the various mechanisms at play.

The remainder of the paper is organized as follows. Section 3.2 describes the Egyptian context. Section 3.3 presents the data source and some descriptive statistics. In Section 3.4 the empirical strategy is explained and the main results are presented in Section 3.5. Section 3.6 provides heterogeneity analysis and potential mechanisms are explored in Section 3.7. Finally, Section 3.8 concludes.

3.2 The Egyptian context

I describe the Egyptian international migration context in Section 3.2.1 and the characteristics of female labor force participation in Section 3.2.2.

3.2.1 International migration

Egypt is considered one of the main emigration countries in the Middle East and North Africa. According to the last Egyptian Census in 2017, there are about 9.4 million Egyptian migrants living abroad (CAPMAS, 2017). Egypt is a country with a long migration history, and it has been a major labor exporter since the beginning of the 70s. Under President Nasser's regime, Egyptian emigration was very limited and tightly regulated (De Bel-Air, 2016). In 1971, all legal restrictions on cross-border movements were lifted as part of President El-Sadat's economic liberalization policy. Since then, emigration has been recognized to be a citizen's right in the Egyptian constitution and a phenomenon of temporary labor migration started (Zohry, 2003). At the same time, following the 1973 oil crisis and the sharp increase in oil prices, the oil-producing Arab countries implemented major development programs and were constrained by the labor shortage, thus turning to neighboring countries to import a large number of workers (Wahba, 2015b). Many Egyptians have migrated to Saudi Arabia and the other Gulf countries, as well as Iraq

and Libya, to work in the oil and construction industries (De Bel-Air, 2016). Egyptian emigration during the latter half of the 20th century was a product, not only of the thriving oil industry in Arab Gulf nations and the demand for labor in neighboring countries, but also of the daunting economic challenges and skyrocketing population growth that beset Egypt during this period (Zohry, 2014). The 2000s witnessed a consistent influx of Egyptian migrants to the Gulf, especially between 2003-2008 with the rise in oil prices during this period. The global financial crisis does not seem to have impacted this trend, with only a moderate impact on migration to Saudi Arabia (De Bel-Air, 2016).

Although Egyptian emigrants are concentrated in the Arab and Gulf states, 30% of all migrant stocks reside in North America and Europe as of the last Egyptian census (CAPMAS, 2017). In the past, Egyptian migration was mainly limited to permanent emigration to the United States (US), which was motivated by political reasons (Wahba, 2015b). However, since the 1990s, Egyptian migrants have been increasingly moving to Western Europe and North America (Wahba, 2015b). Over the last decade, and more so in the wake of the Egyptian revolution, larger numbers of Egyptians turned to the West with popular destinations currently being Italy, France, Germany, the US, and Canada (De Bel-Air, 2016). A salient characteristic of Egyptian emigration to the West is that it is predominately permanent in nature and tends to involve the entire household, as opposed to migration to the Arab region which is male-dominated and largely temporary in nature mainly due to short-term work contracts and no right to citizenship privileges (Wahba, 2015b; Tsourapas, 2022).

In addition to its temporary nature, several domestic and international factors resulted in the contraction of the number of Egyptian emigrants and massive waves of return migration from the Gulf, Iraq, and Libya since the late 1980s (Zohry, 2003). This includes the Iran-Iraq War in 1988 and the resulting decline in oil prices, the First Gulf War in 1990-1991, the policy of replacing foreign labor with a national labor force, the decline in the demand for construction workers, and the Iraq war in 2003. Moreover, there has also been a massive wave of Egyptian migrants returning back from Libya after the fall of Gaddafi's regime during the Arab Spring (Zohry, 2013; De Bel-Air, 2016).

3.2.2 Female labor force participation

Egypt has one of the lowest female labor force participation rates in the world. According to the Global Gender Gap Report in 2022, Egypt is ranked 143rd out of 146 countries in terms of female labor force participation rate (World Economic Forum, 2022). Despite

a significant increase in female education, the female labor force participation rate in Egypt has remained persistently low, consistently below 30%. This has been referred to as the MENA Paradox by the World Bank (Vishwanath, 2012). In fact, recent trends even indicate a substantial decline in participation. This decline is primarily attributed to the reduction in job opportunities within the public sector that has not been balanced by a corresponding growth in opportunities within the formal private sector (Krafft et al., 2019; Assaad et al., 2020).

On the supply side, cultural factors also constitute an important impediment to female employment. Egypt retains its traditional societal structure, where the allocation of time within households follows a gendered pattern: men primarily engage in market work, while the majority, if not all, of the family responsibilities, continue to fall on women's shoulders. This influences their choices regarding employment, especially after marriage. Selwaness and Krafft (2021) even show that women leave work in anticipation of marriage and anticipated domestic work, as prevalent gender norms designate women as the primary caregivers. Thus, marriage, along with the responsibilities of childbearing and household tasks that come with it, contribute in part to the lower rates of female labor force participation (Hendy, 2015).

Education and marital status are significant determinants of the employment patterns observed among Egyptian women. Those with lower levels of education frequently encounter restricted prospects for paid employment, leading them to rely on self-employment activities conducted from home or unpaid family work (Assaad et al., 2020). In contrast, educated married women tend to favor employment in the public sector due to its shorter working hours and favorable maternity and family leave policies, while unmarried women may have the chance to participate in private sector wage work (Hendy, 2015). However, it is common for unmarried women to discontinue such employment upon marriage (Assaad et al., 2022). Furthermore, there exists a disparity in female employment between urban and rural areas, with urban women being more likely to engage in market work, while subsistence work is more prevalent among rural women.

In this paper, the focus is on examining the impact of male migration on females left behind labor force participation and exploring how it affects their engagement in different employment statuses in urban and rural areas distinctly. The study specifically focuses on analyzing the distinctions between current and return migration and evaluating any potential variations in their impacts. Given the significant prevalence of temporary migration and the persistently low female labor force participation rate in Egypt, understanding the relationship between these two phenomena is of utmost importance for informing policy

decisions and interventions aiming to increase female employment.

3.3 Data and descriptive statistics

In this section, I present the main data source in Section 3.3.1 and some descriptive statistics in Section 3.3.2.

3.3.1 Data source

This paper uses three waves of the Egypt Labor Market Panel Survey (ELMPS), carried out in 2006, 2012, and 2018. The ELMPS is a nationally representative longitudinal survey conducted by the Economic Research Forum (ERF) in cooperation with the Egyptian Central Agency for Public Mobilization and Statistics (CAPMAS) since 1998.[6] The survey covers a wide range of topics, including demographic characteristics, labor market dynamics, the status of women, and migration experiences, making it an ideal data source for empirical research on migration and labor market outcomes in Egypt.

Two key features of the ELMPS make it particularly suitable for this study. First, the survey collects data on current international migrants and their characteristics from existing household members. A household is categorized as a current migrant household if it has reported any of its members as living or working abroad at the time of the survey. The most knowledgeable person within the household (often, the household head) then provides information about current migrants' characteristics, including details such as their highest level of education, year of migration, destination country, and occupation overseas. Second, the survey also includes a return migration module where all individuals aged 15 to 59 years old are asked whether they have ever worked abroad for more than six months.[7] A household is considered a return migrant household if at least one member reports having worked abroad for more than six months. This approach, as opposed to relying on retrospective questions about job history and residential mobility, enables more efficient identification of households with return migrants and is less susceptible to recall bias. The return migration module also provides information about various aspects of the migration experience, including the year and country of destination for the first migration episode, the year of final return, and reasons for return.

[6] I exclude the first wave of this longitudinal survey (ELMPS 1998) from my analysis because the variables of interest were not administered during this wave.

[7] It should be noted that the return migration module was only administrated in the 2012 and 2018 waves. For individuals surveyed in 2006 and resurveyed in 2012, the information provided in 2012 was used to determine their return migration status in 2006.

The ELMPS also collects comprehensive data on labor market outcomes. To assess female labor supply, this paper first adopts the prevailing labor force participation definition, which encompasses individuals who were employed or actively seeking employment in the seven days preceding the interview date. Two distinct employment definitions are considered. First, the "market definition" comprises all individuals engaged in an activity for at least one hour per week, with the aim of receiving pay or generating profit. This category includes wage workers, employers, self-employed individuals, and unpaid family workers. However, given that women in Egypt are more likely than men to engage in subsistence work (Hendy, 2015), the "extended definition" is also considered. This broader definition includes those involved in the production or processing of primary commodities for household consumption. Since labor force participation can mask heterogeneous effects, I also consider the impact on distinct employment statuses.

In Egypt, particularly in rural areas, a large share of females participate in farming activities or small-scale businesses as unpaid family workers, as highlighted later in Table 3.2 below. It is worth noting that the contrast between female involvement in paid and unpaid employment does not reflect the same degree of economic independence and autonomy. While wage work and self-employment empower women and enhance their agency, unpaid family labor is often perceived as an extension of their domestic roles and does not lead to increased agency (Lenoël and David, 2019). Nevertheless, in developing contexts, it is crucial to acknowledge the latter category, as it signifies women's active participation in economic activities.

3.3.2 Descriptive statistics

The baseline sample is presented in Section 3.3.2.1, followed by descriptive statistics on the characteristics of migrants in Section 3.3.2.2 and those of females in Section 3.3.2.3.

3.3.2.1 Baseline sample

My baseline sample focuses on working-age females aged 15 to 64 who have lived with male household members at least once and who are not themselves return migrants.[8] To streamline the analysis and reduce irrelevant variability, specific criteria are applied to exclude particular females from the sample. These criteria include the exclusion of females observed in households concurrently hosting both current and return migrants, those in

[8] The share of surveyed females with any migration experience is notably low, accounting for a mere 0.17% of working-age females.

households with multiple current or return migrants, those in households with a female current or return migrant,[9] and those in households with a current or return migrant from a non-Arab country.[10,11] From the overall sample, two sub-samples are examined. The first sub-sample focuses on females observed while residing in a household with either a current male migrant or no migration experience, excluding observations in households with a return migrant. The second considers females observed while residing in a household with either a male return migrant or no migration experience, excluding observations in households with current migrants. This allows for a separate analysis of the impact of having a current migrant or a return migrant compared to having no migration experience within the household.[12] Only females meeting the above-mentioned criteria at least twice over the three waves are retained in my baseline sample forming an unbalanced panel of 11,777 females with a total of 28,315 female-year observations overall.[13] I present the summary statistics for both samples by area of residence in Appendix C.1.4 in Table C.4 and Table C.5, respectively. I observe that 2% of females reside in households with a current migrant in urban areas, while this figure is higher at 5% in rural areas.[14] In the return migrant sample, females living in households with a return migrant constitute 12% and 18% of the urban and rural samples, respectively.

3.3.2.2 Male migrants' characteristics

As highlighted in Section 3.2.1, Egyptian migration is exclusively male-dominated, with the majority of migrants heading primarily toward Middle Eastern countries, particularly those in the Gulf region. Data from the ELMPS reveals that approximately 98% of both current and return migrants are males, with 95% (98%) of current (return) migrants having migrated to Arab countries. Further details on the main destination countries for current and return migrants are outlined in Fig. 3.1.

Saudi Arabia, Kuwait, Jordan, the United Arab Emirates, and Libya, emerge as the main destination countries for current migrants. Migration to non-Arab countries is less

[9] In the 2006 wave, gender and relationship status of current migrants were not reported. Assuming prevalent male dominance in migration to the Arab region, it is presumed that they were male migrants.

[10] I test the robustness of the results to keeping all these observations in the robustness checks in Section 3.5.3.

[11] The percentage of those females is detailed in Appendix C.1.1.

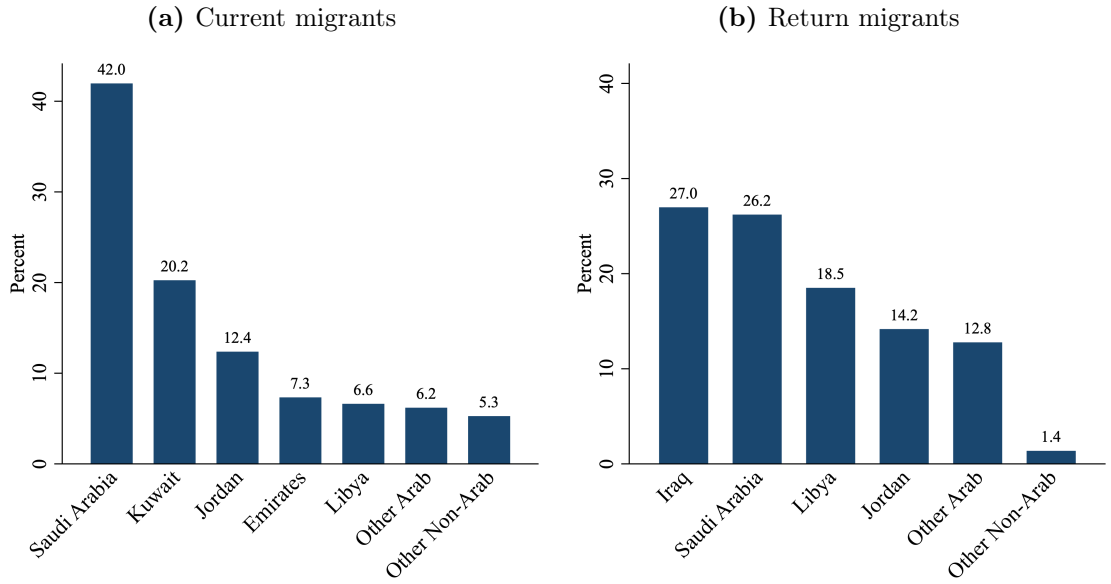
[12] Females' assignment to current and return migrant samples is detailed in Appendix C.1.2.

[13] A detailed description of how the baseline sample and the two sub-samples are formed, along with the variation in the household migration status across waves, is presented in Appendix C.1.3.

[14] One caveat of this study is the low share of households with current migrants, particularly in the urban sample.

common, with only 5% of current migrants choosing Europe and the United States as their destinations.[15] Notably, migration to the Arab region tends to be of a temporary nature, with an average migration duration of approximately four years, while migration to Western countries tends to be more permanent. Return migrants from non-Arab countries constitute a mere 1% of the total returnee population.

Fig. 3.1. Main destination countries.



Note: This figure shows the distribution of current and return migrants' main destination countries.
Source: Author's elaboration on ELMPS (2006, 2012, 2018).

Furthermore, migrant populations exhibit notable differences between urban and rural areas,[16] which requires a careful investigation of their implications in understanding female labor supply. In Table 3.1 below, I compare the characteristics of current and return migrants in urban and rural settings. Current migrants from urban areas tend to come from wealthier households, with higher levels of non-labor income. Additionally, they tend to be more educated and are more likely to work in countries like Saudi Arabia. In contrast, current migrants from rural areas tend to come from households with a higher number of dependents, have lower educational attainment, and are more inclined to migrate to Jordan, Kuwait, and Libya.[17]

[15] Migration to the West often involves the entire household (Wahba, 2015b).

[16] The official definition of urban areas in Egypt is purely administrative, including four main governorates (Cairo, Port Said, Suez, and Alexandria), agglomerations declared as cities, the capitals of other governorates, and district capitals.

[17] This pattern aligns with the fact that in the Gulf States, documented workers are often employed in white-collar, skilled, and semi-skilled technical fields, alongside agriculture and husbandry, whereas Libya, Jordan, and Iraq tend to employ less educated workers in the agriculture, construction, and service sector

Table 3.1
Current and return migrants' characteristics in urban vs rural households.

	Current migrant			Return migrant		
	Urban	Rural	P-value	Urban	Rural	P-value
Household characteristics						
Household wealth	0.611	-0.072	0.000***	0.374	-0.309	0.000***
Non labor income	613.662	364.696	0.043**	400.318	366.856	0.663
Household size	3.641	3.862	0.157	4.713	5.021	0.000***
No. children 0-2	0.354	0.428	0.149	0.246	0.368	0.000***
No. children 3-5	0.283	0.384	0.034**	0.304	0.404	0.000***
No. children 6-12	0.561	0.714	0.040**	0.773	0.863	0.012**
No. elderly ≥ 65	0.076	0.124	0.099*	0.096	0.136	0.002***
Education						
Less than basic education	0.082	0.220	0.000***	0.157	0.331	0.000***
Basic education	0.138	0.119	0.484	0.161	0.154	0.632
Secondary education	0.418	0.496	0.062*	0.393	0.409	0.383
Post-secondary education	0.362	0.166	0.000***	0.288	0.105	0.000***
Main destination countries						
Iraq	0.005	0.002	0.429	0.241	0.272	0.060*
Jordan	0.051	0.155	0.000***	0.108	0.167	0.000***
Kuwait	0.168	0.241	0.033**	0.063	0.046	0.037**
Libya	0.030	0.061	0.099*	0.195	0.210	0.316
Qatar	0.030	0.038	0.608	0.014	0.011	0.511
Saudi Arabia	0.619	0.422	0.000***	0.313	0.231	0.000***
Emirates	0.086	0.061	0.224	0.047	0.036	0.132
Current employment status						
Wage worker				0.622	0.594	0.129
Non-wage worker				0.273	0.319	0.008***
Unemployed or out of labor force				0.105	0.086	0.099*
Number of households	198	573		1,107	2,108	

Note: This table reports the t-test statistics by area of residence for current and return migrants households: *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, and 2018).

When comparing return migrants in urban and rural areas, clear differences among return migrants are still observed. In urban areas, they are also more educated and live in wealthier households with a lower number of dependents. However, rural return migrants demonstrate a higher likelihood of participating in non-wage work compared to their urban counterparts. To account for these variations and explore the factors influencing female labor supply in each context, separate regressions are estimated for urban and rural samples.

(Wahba, 2015b).

3.3.2.3 Females' characteristics

Females' characteristics and labor market dynamics may also vary between urban and rural areas. Table 3.2 below compares females in both settings. It becomes apparent how significantly different they are across all characteristics. In terms of labor supply, urban females are more inclined to be part of the market labor force compared to rural ones. They are significantly more likely to be wage workers and less likely to be unpaid family workers, in contrast to their rural counterparts. Overall, unpaid family work emerges as the predominant employment status in rural areas, reflecting the prevalence of agricultural activities and females' involvement in farm work. This pattern underscores the significance of examining the impact of male migration on different employment statuses, as they do not uniformly represent the same level of female empowerment. Furthermore, urban females in my sample are, on average, older, less likely to be married, but have more years of education. Significant differences also extend to their households. In urban areas, females reside in wealthier households, which are generally smaller in size. These households have fewer children and elderly among their members.

In Table 3.3, I compare females in a current migrant household to females in a non-migrant household in urban and rural areas, separately. In urban areas, females in a current migrant household do not significantly differ from those in non-migrant households across several characteristics. They live in smaller wealthier households with higher non-labor income and they are less likely to have prior work experience. Conversely, in rural areas, females from households with current migrants abroad display distinct traits. They are more likely to be unpaid family workers and are less likely to be wage workers. Additionally, they tend to be younger and have higher levels of education. They live in smaller and wealthier households with a higher number of children aged 0 to 2.

Comparing females in households with a return migrant to those in non-migrant households, as presented in Table 3.4, several differences emerge. Across both urban and rural areas, females in households with a return migrant are more likely to be part of the extended labor force and to be engaged in unpaid family work. They are also more likely to be married and to be students. They tend to live in larger households, they have fewer children aged 0 to 2 but a greater number of children aged 6 to 12. In urban areas, these females tend to have higher levels of education.[18]

[18] It is important to note that the t-tests presented in Table 3.2, Table 3.3 and Table 3.4 are conducted on the pooled sample of females and only reflect average differences in females' characteristics between households with a migration experience and those without.

Table 3.2
Females' characteristics in urban and rural households - pooled cross-section.

	Urban	Rural	P-value
Outcome variables			
Market labor force	0.271	0.227	0.000***
Extended labor force	0.328	0.462	0.000***
Wage work	0.175	0.084	0.000***
Self-employment	0.018	0.034	0.000***
Unpaid family work	0.077	0.306	0.000***
Individual characteristics			
Age 15-24	0.220	0.261	0.000***
Age 25-34	0.304	0.328	0.000***
Age 35-44	0.206	0.197	0.056*
Age 45-54	0.169	0.137	0.000***
Age 55-64	0.100	0.077	0.000***
Married dummy	0.711	0.778	0.000***
Student dummy	0.092	0.069	0.000***
Years of schooling	9.946	6.942	0.000***
Years of schooling squared	124.084	78.011	0.000***
Work experience dummy	0.367	0.342	0.000***
Household characteristics			
Household wealth	0.363	-0.354	0.000***
Non labor income	474.018	365.156	0.000***
Household size	4.512	5.020	0.000***
No. children 0-2	0.272	0.394	0.000***
No. children 3-5	0.284	0.399	0.000***
No. children 6-12	0.603	0.776	0.000***
No. elderly ≥ 65	0.130	0.142	0.010***
Number of female-year obs.	12,687	15,628	

Note: This table reports the t-test statistics for female's characteristics between urban and rural areas: *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, and 2018).

Table 3.3
Females' characteristics in current and non-migrant households by area of residence - pooled cross-section.

	Urban			Rural		
	Current migrant	Non-migrant	P-value	Current migrant	Non-migrant	P-value
Outcome variables						
Market labor force	0.257	0.270	0.665	0.219	0.225	0.704
Extended labor force	0.335	0.323	0.704	0.480	0.457	0.220
Wage work	0.151	0.174	0.340	0.056	0.085	0.008***
Self-employment	0.012	0.019	0.469	0.026	0.035	0.231
Unpaid family work	0.082	0.074	0.650	0.336	0.299	0.041**
Individual characteristics						
Age 15-24	0.245	0.216	0.280	0.285	0.258	0.113
Age 25-34	0.298	0.315	0.580	0.352	0.334	0.331
Age 35-44	0.155	0.201	0.077*	0.162	0.192	0.053*
Age 45-54	0.188	0.166	0.362	0.132	0.134	0.871
Age 55-64	0.114	0.103	0.553	0.069	0.083	0.218
Married dummy	0.714	0.707	0.805	0.789	0.769	0.235
Student dummy	0.102	0.086	0.363	0.069	0.065	0.626
Years of schooling	10.196	9.875	0.323	7.483	6.883	0.005***
Years of schooling squared	131.796	122.782	0.102	84.786	77.208	0.011**
Work experience dummy	0.302	0.369	0.031**	0.288	0.346	0.002***
Household characteristics						
Household wealth	0.669	0.350	0.000***	-0.073	-0.380	0.000***
Non labor income	694.682	470.243	0.023**	373.302	373.325	1.000
Household size	3.935	4.474	0.000***	4.276	5.005	0.000***
No. children 0-2	0.327	0.275	0.125	0.444	0.394	0.034**
No. children 3-5	0.245	0.287	0.220	0.378	0.401	0.332
No. children 6-12	0.522	0.591	0.215	0.726	0.767	0.279
No. elderly ≥ 65	0.098	0.136	0.115	0.139	0.144	0.742
Number of female-year obs.	245	10,844		691	11,983	

Note: This table reports the t-test statistics by migration status for urban and rural households in the current migrant sample: *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, and 2018).

Table 3.4
Females characteristics in return and non-migrant households by area of residence - pooled cross-section.

	Urban			Rural		
	Return migrant	Non-migrant	P-value	Return migrant	Non-migrant	P-value
Outcome variables						
Market LFP	0.284	0.271	0.284	0.242	0.226	0.077*
Extended LFP	0.357	0.324	0.014**	0.485	0.456	0.007***
Wage work	0.184	0.176	0.469	0.083	0.086	0.641
Self-employment	0.019	0.019	0.829	0.037	0.034	0.520
Unpaid family work	0.098	0.074	0.001***	0.334	0.298	0.000***
Individual characteristics						
Age 15-24	0.231	0.217	0.232	0.244	0.261	0.075*
Age 25-34	0.222	0.317	0.000***	0.291	0.336	0.000***
Age 35-44	0.268	0.200	0.000***	0.247	0.190	0.000***
Age 45-54	0.195	0.165	0.005***	0.165	0.132	0.000***
Age 55-64	0.084	0.101	0.044**	0.054	0.081	0.000***
Married dummy	0.751	0.708	0.001***	0.837	0.768	0.000***
Student dummy	0.139	0.084	0.000***	0.085	0.064	0.000***
Years of schooling	10.310	9.891	0.003***	6.917	6.914	0.978
Years of schooling squared	130.169	123.026	0.003***	77.633	77.758	0.940
Work experience dummy	0.361	0.370	0.545	0.351	0.345	0.552
Household characteristics						
Household wealth	0.396	0.348	0.077*	-0.302	-0.382	0.000***
Non labor income	466.324	468.954	0.954	344.005	370.393	0.492
Household size	4.984	4.463	0.000***	5.390	4.995	0.000***
No. children 0-2	0.221	0.278	0.000***	0.369	0.398	0.026**
No. children 3-5	0.266	0.288	0.138	0.402	0.402	0.974
No. children 6-12	0.744	0.588	0.000***	0.875	0.762	0.000***
No. elderly ≥ 65	0.098	0.136	0.000***	0.142	0.142	0.963
Number of female-year obs.	1,444	10,877		2,577	12,068	

Notes: This table reports the t-test statistics by migration status for urban and rural households in the return migrant sample : *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, and 2018).

3.4 Empirical strategy

This section explains the empirical strategy used to identify the impact of male migration on female labor supply. The benchmark specification is presented in Section 3.4.1, while the identification strategy is discussed in Section 3.4.2.

3.4.1 Benchmark specification

To study the effect of male migration on females left behind labor supply, the following regression specification is estimated separately for urban and rural areas:

$$Y_{iht} = \beta_0 + \beta_1 M_{ht} + \boldsymbol{\beta_2 X_{iht}} + \gamma_{dt} + \gamma_i + \epsilon_{iht} \tag{3.1}$$

Y_{iht} captures labor supply outcomes. Specifically, I estimate Equation (3.1), consecutively, for market and extended labor force participation and various types of work including wage work, self-employment, and unpaid family work. It takes the value of 1 if female i within household h during year t is in the labor force, a wage-worker, a self-employed, or an unpaid family worker, and 0 if she is out of the labor force. My variable of interest is the household migration status. Hence, I first estimate Equation (3.1) considering M_{ht} a dummy variable equal to 1 if household h has a current male migrant abroad, and 0 if it is a household with no migration experience. Then, I re-estimate it considering M_{ht} a dummy variable equal to 1 if household h has a male return migrant among its members, and 0 if it is a household with no migration experience.

The vector $\boldsymbol{X_{iht}}$ comprises a set of time-varying individual and household characteristics including age group dummies, a married dummy, a student dummy, years of schooling, years of schooling squared, a work experience dummy, household wealth score, household non-labor income, the number of children aged (0-2), the number of children aged (3-5), the number of children aged (6-12), and the number of elderly (those over 65 years old) residing in the household. γ_{dt} stands for district-year fixed effects which control for time-varying local economic settings and labor demand. γ_i denotes individual fixed effects, which control for time-invariant unobserved characteristics. ϵ_{iht} is the error term. Standard errors are clustered at the household-year level,[19] and the coefficients are estimated

[19] A common guideline is to cluster the data at the level of the regressor of interest when the latter varies at a more aggregate level than the units being observed (Abadie et al., 2023). I test the robustness of the results to alternative clustering schemes in the robustness checks in Section 3.5.3.

using a linear probability model.[20]

3.4.2 Identification strategy

When assessing the impact of male migration on labor-market behavior, two significant concerns related to endogeneity must be taken into account. The primary and most critical issue pertains to self-selection into migration. Migrants, as well as their households, are not randomly chosen from the general population. Instead, their decision to migrate is influenced by both observable and unobservable characteristics. These unobservable factors, such as preferences, risk aversion, and family culture, can simultaneously influence migration decisions and the choice to involve females in the workforce. For instance, households with lower risk aversion may be more inclined to engage in international migration while also promoting female employment. Furthermore, in a patriarchal society like Egypt, men with more progressive attitudes are also more likely to migrate internationally and support female employment, even in the face of societal disapproval. Thus, failure to account for this unobserved heterogeneity can lead to biased estimates. Another potential concern is simultaneity bias, where the decision for a female to enter the labor force and the decision to send a household member abroad are jointly determined within the household. For example, a female may start working to cover the costs associated with a migrating family member, or a male member might decide to migrate in response to a woman leaving the workforce to care for children (Kan and Aytimur, 2019). However, in relatively conservative societies like Egypt, the latter scenario is less likely, as men's choices are less likely to be influenced by the employment decisions of women in their households. In summary, migrant households and non-migrant households differ across various dimensions, leading to variations in their potential outcomes.[21]

Just like in the case of out-migration, it's important to recognize that return migrants may not represent a random sample from the entire group of migrants (Wahba, 2015a). As discussed by (Borjas and Bratsberg, 1996), the direction of self-selection in return migration

[20] Although my dependent variables are all binary, estimating my fixed-effects regressions using logit or probit models is not optimal due to the incidental parameter problem in maximum likelihood methods (Greene, 2004). In this case, a linear probability model is preferable (Angrist and Krueger, 2001).

[21] Another empirical challenge worth mentioning when evaluating the impact of migration on household members left behind, as emphasized by Bertoli et al. (2023), is that the results are silent on individuals who change their living arrangements. When the migrating family member is the husband or the father, females left behind have several choices. They may remain in their current household, have close relatives move in, or relocate to their parents', in-laws', or grandparent's residence (Hoodfar, 1993). In the latter case, this leads to the dissolution of the original household, causing those left behind to drop out of the survey, resulting in attrition.

depends on whether the initial selection of migrants was based on positive or negative attributes. When the initial group of migrants is positively selected, those who return are likely to be negatively selected, and vice versa. Hence, an unsuccessful migration experience can influence both the likelihood of returning home and the decision regarding female household members' participation in the workforce. This influence can manifest as either an attempt to compensate for the unfavorable experience or as a reflection of more negative attitudes toward women (Tuccio and Wahba, 2018).

Furthermore, a final selection concern pertains to the choice of immigration destination. Migrants may not randomly select their destinations; instead, they might choose places with cultural values that align with their own. For example, men from more (less) conservative households may opt for destinations with similar conservative (liberal) values, potentially correlated to more (less) conservative views on female labor participation. However, two factors help alleviate this concern. First, I examine temporary migration by Egyptians to the Arab region, where cultural values are relatively similar, despite obscure differences in the level of conservatism among destination countries. Second, labor migration to the Arab region is primarily driven by demand and influenced by the economic conditions in these destinations. Additionally, many Arab countries have visa entry regulations that require sponsorship from a local resident.

The aforementioned concerns primarily enclose the notion that both observable and unobservable characteristics tend to differ among households with current, return, and non-migrant members. While it's relatively straightforward to manage observable characteristics using a comprehensive set of time-varying controls at the individual, household, and district levels as described in Section 3.4.1, unobserved factors pose a more significant challenge. One potential solution to address these concerns is to employ a fixed effect model, which can help mitigate issues associated with self-selection. In this study, the dataset's panel dimension offers an advantage by allowing me to control for individual-level time-invariant characteristics that may influence the presence of a current or return migrant in the household and the participation of female household members in the labor force. Additionally, these individual-level fixed effects can also account for much of the heterogeneity at the household level, serving as a proxy for couple or family fixed effects. By incorporating these fixed effects, I can address all time-invariant unobserved heterogeneity, including factors like innate abilities, risk preferences, cultural norms, and social networks. These factors are not directly observable but may impact both the decision to send a family member abroad and the decision of female household members to engage in the labor force. While the extensive set of individual and household characteristics,

along with the individual fixed effects, effectively control for a substantial portion of the selection into migration and return migration, there may still be some bias due to omitted variables that vary over time. To further address this concern, I rely on the methodology proposed by Oster (2019) in Section 3.5.2 to assess the significance of these omitted variables in explaining the effects observed in my analysis.[22]

3.5 Main results

In this section, the benchmark results obtained from estimating Equation (3.1) for current and return migration are reported in Section 3.5.1. Then, Section 3.5.2 outlines the findings of the Oster's test regarding the selection on unobservables. Finally, Section 3.5.3 presents various robustness checks.

3.5.1 Benchmark results

Table 3.5 presents the benchmark results of the impact of male migration on female labor force participation in both urban and rural areas obtained from estimating Equation (3.1). Columns (1) and (3) provide insights into the impact on market labor force participation, while Columns (2) and (4) focus on the extended labor force. In Panel (A), I focus on the impact of having a current migrant in the household compared to not having any migration experience. On average, the results suggest that the presence of a current migrant does not yield any statistically significant impact on the labor force participation of females left behind in urban areas. Conversely, in rural areas, the findings indicate a noteworthy impact, particularly on extended labor force participation. Specifically, a female living in a rural household where a male member is currently abroad is 8.5 percentage points more likely to participate in the extended labor force than when residing in a household with no migration experience. In Panel (B), I examine the impact of living in a household where a return migrant is present. On average, I find no statistically significant impact of having a return migrant in the household on either labor force definitions, neither in urban nor rural areas.

Nonetheless, the lack of significance of the impact of household migration status on female labor force participation may conceal heterogeneous effects, particularly in terms of

[22] An alternative ideal approach would have involved employing an instrumental-variable fixed effect estimation strategy to further address any remaining endogeneity. However, due to data limitations and the difficulty in identifying a suitable instrument that varies both at the household level and over time, I opted to leverage the unique characteristics of the panel data structure and its available information.

Table 3.5
Male migration and female labor force participation.

	Urban		Rural	
	(1) Market LFP	(2) Extended LFP	(3) Market LFP	(4) Extended LFP
Panel (A): Current migrant sample				
Current migrant	0.021	0.033	0.012	0.085***
	(0.038)	(0.042)	(0.025)	(0.031)
Observations	11,089	11,089	12,674	12,674
Number of females	4,606	4,606	5,337	5,337
Number of clusters	9,013	9,013	10,472	10,472
Mean outcome	0.269	0.323	0.224	0.458
Mean current migrant	0.022	0.022	0.055	0.055
Panel (B): Return migrant sample				
Return migrant	−0.009	−0.015	−0.019	0.012
	(0.029)	(0.037)	(0.022)	(0.030)
Observations	12,321	12,321	14,645	14,645
Number of females	5,096	5,096	6,127	6,127
Number of clusters	9,966	9,966	12,085	12,085
Mean outcome	0.272	0.328	0.228	0.462
Mean return migrant	0.117	0.117	0.176	0.176
Controls	✓	✓	✓	✓
Individual FE	✓	✓	✓	✓
District × Year FE	✓	✓	✓	✓

Notes: This table reports the baseline results of the impact of male migration on female labor force participation in urban and rural areas. The dependent variable in Columns (1) and (3) is the market labor force participation dummy, while in Columns (2) and (4) it is the extended labor force participation dummy. In Panel (A), current migrant is a dummy variable equal to 1 if there is a current migrant abroad, and 0 if there is no migration experience in the household. In Panel (B), return migrant is a dummy variable equal to 1 if there is a return migrant in the household, and 0 if there is no migration experience in the household. Controls include age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. Standard errors clustered at the household-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.
Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

employment type. While labor force participation offers some insights into women's economic engagement in response to male migration, a more in-depth understanding of their involvement in the workforce is gained by examining the impact on different employment statuses. Consequently, I investigate the impact of male migration on the probability of being a wage worker, self-employed, or unpaid family worker, in comparison to being out of the labor force.[23] Panel (A) in Table 3.6 reveals distinct impacts of current migration on employment status in urban and rural areas.[24] While having a current migrant in-

[23] Using "out of the labor force" as the reference category allows me to assess the impact on the extensive margin of the female labor supply. In Section 3.5.3, I further explore the impact on the intensive margin by considering other employment statuses as the counterfactuals.

[24] The varying number of observations between Table 3.5 and Table 3.6 is due to the use of "out of the labor force" category as the sole counterfactual when assessing the impact of male migration on the

creases the probability of engaging in wage work in urban areas by 5.4 percentage points, it increases unpaid family work in rural areas by 8.1 percentage points. The latter result aligns with the findings of Binzel and Assaad (2011), who observed the same trend in rural areas. There, females tend to participate more in non-wage and subsistence work when a male household member migrates to compensate for his absence. As depicted in Panel (B), return migration does not yield any significant impact on employment statuses in urban areas. However, it is associated with a decrease in the probability of engaging in wage work in rural areas by 3.8 percentage points.

Table 3.6
Male migration and female employment status.

	Urban			Rural		
	(1) Wage work	(2) Self-emp.	(3) Unpaid work	(4) Wage work	(5) Self-emp.	(6) Unpaid work
Panel (A): Current migrant sample						
Current migrant	0.054**	−0.000	−0.006	0.006	−0.007	0.081**
	(0.026)	(0.011)	(0.031)	(0.026)	(0.024)	(0.034)
Observations	8,724	6,643	7,509	6,061	5,337	9,985
Number of females	3,771	2,912	3,223	2,769	2,472	4,305
Number of clusters	7,374	5,883	6,564	5,342	4,750	8,467
Mean outcome	0.214	0.025	0.103	0.160	0.057	0.362
Mean current migrant	0.020	0.020	0.021	0.047	0.051	0.055
Panel (B): Return migrant sample						
Return migrant	−0.011	0.016	−0.019	−0.038**	−0.011	0.035
	(0.027)	(0.018)	(0.036)	(0.016)	(0.020)	(0.034)
Observations	9,670	7,332	8,330	7,002	6,140	11,552
Number of females	4,167	3,206	3,562	3,183	2,835	4,957
Number of clusters	8,126	6,457	7,238	6,139	5,453	9,796
Mean outcome	0.218	0.026	0.107	0.165	0.060	0.367
Mean return migrant	0.112	0.109	0.115	0.162	0.163	0.177
Controls	✓	✓	✓	✓	✓	✓
Individual FE	✓	✓	✓	✓	✓	✓
District × Year FE	✓	✓	✓	✓	✓	✓

Notes: This table reports the baseline results of the impact of male migration on female employment status in urban and rural areas. The dependent variable in Columns (1) and (4) is the wage work dummy, in Columns (2) and (5) is the self-employed dummy, and in Columns (3) and (6) is the unpaid family work dummy. In Panel (A), current migrant is a dummy variable equal to 1 if there is a current migrant abroad, and 0 if there is no migration experience in the household. In Panel (B), return migrant is a dummy variable equal to 1 if there is a return migrant in the household, and 0 if there is no migration experience in the household. Controls include age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. Standard errors clustered at the household-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

In summary, these findings suggest that current migration is associated with increased female labor supply in both urban and rural areas, albeit with varying effects on employ-

specified employment status. This approach omits observations recording employment statuses other than the one analyzed, resulting in a reduction in the number of observations.

ment status. While having a male member abroad exerts a generally positive influence on female wage work in urban areas, it primarily impacts unpaid family work in rural areas. This divergence may translate into heightened economic independence for urban females but may not necessarily have the same effect on their rural counterparts.

Conversely, return migration demonstrates a comparatively limited impact. In urban settings, having a return migrant does not seem to significantly affect the female labor supply. This may suggest that the influence of migration diminishes upon the migrant's return. However, return migration negatively affects wage work in rural areas. This adverse impact might stem from either an economic mechanism, where return migrants improve their economic standing, reducing the need for female household members to engage in paid employment, or a social mechanism, where return migrants may bring back home more conservative norms from their destination countries.

These findings underscore the nuanced effects associated with different male migration statuses on females' employment within distinct regional contexts and highlight the importance of investigating the mechanisms behind these results. Furthermore, they emphasize the significance of considering various employment categories for a more comprehensive understanding of women's economic roles. For the remainder of the analysis, I will focus on the impact on employment status, as it provides a deeper understanding of these dynamics.[25]

3.5.2 Selection on unobservables

As detailed in Section 3.4.2, using fixed-effect estimations and the full set of observed controls helps mitigate the influence of unobserved heterogeneity and omitted variable bias, which could otherwise introduce bias into the estimates of the impact of male migration on female labor supply. However, it is possible that some remaining time-varying omitted factors still affect my estimates. One common approach in empirical research to assess the extent of the omitted variables problem is to observe the sensitivity of the coefficient of interest to the inclusion of the observed controls. A stable coefficient often reflects a limited omitted variable bias. Nevertheless, tracking the stability of the coefficient is only informative when taking the change in the estimated R-squared into account. To assess whether the benchmark results presented earlier are susceptible to omitted variable bias, I employ the methodology proposed by Oster (2019). The latter relies on the aforementioned intuition and develops a methodology to assess the robustness of the results

[25] Benchmark results displaying control variables are reported in Tables C.7 to C.10 in Appendix C.2.

to omitted variable bias by generating bias-adjusted treatment effects. Her approach assumes that unobservable characteristics are correlated with the observed variables in the model.

Oster (2019)'s methodology consists of: (i) measuring how important selection on unobservable factors needs to be, relative to selection on observable ones, denoted as δ, for the coefficient of interest to be zero, and (ii) computing bounds for the coefficient of interest that correct for a reasonable degree of selection on observables (where δ equals one, implying that unobservable factors are as important as observable ones).[26] To compute both of these measures, an assumption on how much of the variation in the dependent variable we aim to explain, denoted as R_{max}, is required. Oster (2019) defines a standard value of R_{max} that is 1.3 times the estimated R-squared ($\tilde{R}$).[27] In Table 3.7 below, I report the significant treatment effects obtained from the benchmark estimations with and without control variables along with their respective estimated R-squared values. I also report the δ and the Oster bounds for each coefficient based on two values of R_{max}, the standard one ($1.3\tilde{R}$) and a stricter one ($2\tilde{R}$).

In Panel (A), I report the test results for the impact of current migration, while in Panel (B), I consider the impact of return migration. The uncontrolled and controlled estimates show that including controls slightly decreases the coefficients of interest in absolute terms but substantially increases the R-squared. This implies that an omitted variable bias is plausible. To test the robustness of the results to this bias, I rely on the Oster's tests results. First, regardless of the assumed R_{max} value, the measure of the relative importance of unobservables to observables (δ) is notably high. It is 11.342 (3.4 with $2\tilde{R}$) when assessing the impact of having a current migrant on being a wage worker in urban areas and 2.665 (1.404 with $2\tilde{R}$) when looking at the impact of being an unpaid family worker in rural areas. This implies that unobservable factors would need to be 11.342 (3.4 with $2\tilde{R}$) and 2.665 (1.404 with $2\tilde{R}$) more important than observables for these coefficients to be equal to zero. Given the high values of δ, which exceed the cut-off of 1, it is unlikely that unobservable factors undermine the validity of the results. Second, even after correcting for selection on unobservables, the Oster bounds do not include zero for both coefficients. This indicates that the coefficients remain statistically different from zero. Similarly, in Panel (B), I obtain a δ value above one and the Oster bounds do not

[26] Oster bounds give the bounding sets of the estimates for when unobservables are not corrected for (δ=0) and when accounting for a reasonable degree of selection on observables (δ=1).

[27] Using Oster (2019)'s methodology with fixed effect estimations assumes individual effects are nuisance parameters and relies on the estimated R-squared within-model.

include zero. Overall, these findings suggest that the impacts identified in Table 3.6 are robust to selection on unobservables, as per the methodology proposed by Oster (2019).

Table 3.7
Selection on unobservables - Oster (2019).

	Benchmark		$R_{max}=1.3\tilde{R}$		$R_{max}=2\tilde{R}$	
	(1) Uncontrolled [R-squared ($\tilde{R}$)]	(2) Controlled [R-squared ($\tilde{R}$)]	(3) δ	(4) Oster bounds	(5) δ	(6) Oster bounds
Panel (A): Current migrant sample						
Urban						
Wage work	0.062** [0.001]	0.054** [0.448]	11.342	[0.051;0.054]	3.400	[0.044;0.054]
Rural						
Unpaid work	0.093** [0.001]	0.081** [0.234]	2.665	[0.059;0.081]	1.404	[0.032;0.081]
Panel (B): Return migrant sample						
Rural						
Unpaid work	-0.051** [0.002]	-0.038** [0.358]	2.832	[-0.038;-0.027]	1.183	[-0.038;-0.007]

Notes: This table reports the Oster (2019)'s tests results. Columns (1) and (2) report the benchmark coefficients without and with controls, respectively, along with the estimated within R-squared ($\tilde{R}$). Columns (3) and (5) report δ the selection on unobservables value which gives β=0. Columns (4) and (6) report the Oster bounds which are the bounding sets of the estimates when unobservables are not corrected for (δ=0) and when accounting for a reasonable degree of selection on observables (δ=1). Benchmark results of the impact of having a current migrant on wage work in urban areas and unpaid work in rural areas are reported in Panel (A). Benchmark results of the impact of having a return migrant on wage work in rural areas are reported in Panel (B). All estimates include individual fixed effects.
Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

3.5.3 Robustness checks

In this Subsection, several robustness checks based on the benchmark results reported in Table 3.6 are conducted. Results are presented in Appendix C.3.

The intensive margin of labor supply. The baseline results, as presented in Table 3.6, primarily focus on evaluating the impact of male migration on the extensive margin of female labor supply. In other words, the analysis assesses whether females are more or less inclined to enter the labor force in response to male migration. However, male migrants can also influence the intensive margin of female labor supply by potentially altering the number of hours worked or their employment status. To explore this, I consider alternative dependent variables. First, I investigate the impact of both current and return migration on the number of hours worked in any employment. The results, reported in Columns (1) and (5) of Table C.11, show no significant impact on the number of hours worked. Second, I redefine the employment status dummies to account for transitions between different employment categories. For instance, I define the wage work dummy as equal to

1 if a female is engaged in wage work and 0 if she is self-employed or an unpaid family worker. While there is a slightly significant negative impact of having a return migrant on being self-employed compared to being a wage worker or an unpaid family worker in rural areas, the overall findings indicate no significant impact on the intensive margin.

Female unemployment. In addition to considering employment, females are also considered active if they are unemployed and actively seeking work. Therefore, I extend the analysis to investigate the impact of male migration on female unemployment, as presented in Table C.12. The results indicate a positive impact of having a current migrant on unemployment in rural areas, but this impact is only statistically significant at the 10% level. Conversely, having a return migrant significantly reduces the probability of being unemployed in rural areas. This suggests an overall negative impact of having a return migrant on female labor force participation in rural areas.

An extended sample. In my benchmark analysis, I excluded some households from my sample to eliminate irrelevant variability. These households included those with female return migrants, those with return or current migrants from non-Arab countries, those with multiple current or return migrants, and those with both return and current migrants simultaneously. In Table C.13, I test the robustness of the results to keeping females in those households in my sample. The results remain robust, except for the impact of current migration on wage work in urban areas, which is found to be not statistically significant.

Alternative clustering. In the benchmark specification, standard errors are clustered at the household-year level to account for within-household correlation in migration and labor market behaviors. To ensure the robustness of the results, I allow for other correlation structures by clustering the error terms at the individual and primary sampling units (PSU) levels, as presented in Tables C.14 and C.15 respectively. Clustering standard errors at the individual level allows to account for correlations across time while clustering at the PSU level accounts for the correlation in labor market opportunities among females within the same geographic area. Results remain robust under both clustering schemes.

Single return migration episode. Return migrants who have experienced multiple migration episodes may exhibit distinct characteristics compared to those with a single migration experience. Within my sample, a large share of return migrants (around 80%) have migrated only once. To ensure the robustness of the findings, I exclude females living in households with return migrants who have experienced two or more migration episodes. The results, available in Table C.16, remain consistent with my earlier findings.

Alternative and additional control variables. Finally, to assess the sensitivity of the coefficients to the set of control variables used in the baseline specification, I introduce alternative and additional control variables. First, I test the robustness of the results by replacing the married dummy, education years, and the number of children and elderly in the household with a categorical marital status variable, educational attainment categories, and dummies indicating the presence of children and elderly in the household. Furthermore, I incorporate additional control variables, including a dummy variable indicating the presence of a family enterprise and a dummy variable indicating whether there is at least one employed male in the household. In addition, I control for the household gender composition using the share of men in the household. Lastly, I replace the district-year fixed effects with governorate-year fixed effects to allow for female mobility between districts within the same governorate.[28] The results of these robustness tests are presented in Table C.17, Table C.18, and Table C.19, respectively. The main conclusions remain unchanged.

3.6 Heterogeneous Analysis

In Section 3.5, the presented findings reflect the general average impact of having a male current or return migrant in the household on female labor supply, irrespective of the relationship of those male migrants to the female left behind. However, this section aims to explore the heterogeneous effects, taking into account the potential variations based on the relationship status. In Section 3.6.1, I focus on the current migrant sample, while Section 3.6.2 focuses on the return migrant sample.

3.6.1 Current migrant relationship

Male migration can have distinct impacts on female labor supply based on the relationship status with the female left behind. Recognizing the variation in family dynamics, I seek to capture nuanced patterns that may have been overlooked in the benchmark results. Conducting this heterogeneous analysis also addresses concerns related to the treatment definition within the current migrant sample. In this sample, the treatment definition hinges on the survey question: "Is there any household member currently working or living abroad?", introducing potential ambiguity associated with cohabitation. Defining a current migrant household hence relies on the head reporting the presence of any members abroad at the time of the survey. The wording of the survey question might result in

[28] Governorates are the first-level administrative divisions in Egypt.

measurement errors if individuals who did not cohabit before migration are still perceived as part of the household by the respondent. For instance, a female might inaccurately answer "yes" if her son, who resides in a separate household following marriage, is employed abroad. Focusing on the relationship of the current migrant to the female left behind, such as being the father or husband, ensures thorough consideration of the cohabitation aspect and improves the precision of the treatment definition.

Table 3.8 above presents the distribution of females living in migrant households according to their relationship with the current migrant in both urban and rural areas.[29] Approximately 47.5% of urban females and 53% of rural females reside in households where their husband is the current migrant. Moreover, around 12% and 7% of females have their father as the current migrant in urban and rural areas, respectively. Additionally, a significant proportion of females have another household member as the current migrant (e.g., a son, a brother, a grandfather, or other), accounting for 40% and 34.5% of females in urban and rural areas, respectively. Given the panel structure of the data, it is also possible to observe females residing in different households across waves, leading to varying current migrants from one wave to the other.[30] However, the share of the latter is negligible in urban areas with only 1% of females observed residing in a household where the husband is the current migrant in one wave and where the current migrant is another household member in another wave. In rural areas, almost 6% of females have been observed in migrant households with distinct current migrants.

To examine the role played by the current migrant relationship status in identifying the effect, I replace the current migrant dummy variable with a categorical one, distinguishing current migrants by their relationship status. Specifically, this variable takes the value of 0 if there is no migration experience in the household, 1 if the current migrant is the spouse, 2 if he is the father, and 3 if he is any other household member.[31] The findings, reported in Table 3.9, indicate that the impact of having a current male migrant abroad depends on the relationship status of that migrant. In urban areas, the increase in female wage work predominantly stems from having a husband who is a current migrant. This suggests that

[29] The relationship of the current migrant with the head of the household was not recorded in the 2006 wave. However, I relied on information available in the survey in 2006 and subsequent waves about females' marital status, the presence of a spouse, and the presence of a father in the household, to infer current migrant's relationship status to the female left behind.

[30] It is noteworthy that the ELMPS attempts to track households included in earlier waves and interview both existing and newly added members. Moreover, it also tries to locate any individuals who may have split from these households between waves, and attempts to interview them, along with other individuals found in the households they formed or joined.

[31] Females observed in different current migrant households are excluded from this analysis.

Table 3.8
Distribution of females in current migrant households based on current migrant's relationship.

	Urban	Rural
Current migrant relationship status		
Husband	47.52%	52.79%
Father	11.88%	6.94%
Other relation	39.60%	34.49%
Father & Husband	0%	1.35%
Husband & Other relation	0.99%	4.05%
Father & Other relation	0%	0.39%

Note: This table reports the distribution of females living in current migrant households based on the relationship status of the current migrant in urban and rural areas.
Source: Author's elaboration on ELMPS data (2006, 2012, and 2018).

wives of male migrants may seek employment opportunities to offset the absence of their partners, possibly driven by economic necessity or the aim to augment household income during their husbands' absence. Conversely, in rural areas, where traditional gender roles and family structures often play a significant role in shaping daily life, the increase in unpaid work when the migrant is neither the husband nor the father suggests a deeper reliance on a broader familial support network within rural communities. Understanding the precise household dynamics in this scenario is challenging due to data limitations hindering the identification of the exact relationships of these other household members.

3.6.2 Return migrant relationship

Similar to the effects of current migration, return migration can have variable impacts on female labor supply depending on the relationship status with the female left behind. Table 3.10 below reports the distribution of females living in households with a return migrant, according to the relationship of the latter, in both urban and rural areas. The data indicates that the majority of females, comprising 63.5% in urban areas and 66.5% in rural areas, reside in households where the return migrant is the husband. A significant proportion of females, accounting for 24.5% in urban areas and 17.6% in rural areas, live in households with a father returnee. Additionally, females living in households with other family members as return migrants constitute a notable portion, representing 10.5% and 12.3% in urban and rural areas, respectively. The share of females observed residing with distinct return migrants is relatively small accounting for only 1.5% and 3.5% of the urban and rural samples, respectively.

To investigate the potential heterogeneous impact on female labor supply based on relationship status, I consider again a categorical variable, distinguishing return migrants by

Table 3.9
Current male migrant relationship and female employment status.

	Urban			Rural		
	(1) Wage work	(2) Self-emp.	(3) Unpaid work	(4) Wage work	(5) Self-emp.	(6) Unpaid work
Panel (A): Current migrant sample						
Husband current migrant	0.059**	0.010	−0.016	0.034	0.001	−0.030
	(0.029)	(0.014)	(0.056)	(0.030)	(0.029)	(0.046)
Father current migrant	0.039	−0.039	−0.166	−0.027	0.030	0.088
	(0.044)	(0.040)	(0.118)	(0.050)	(0.035)	(0.106)
Other current migrant	0.052*	0.005	−0.062	−0.027	−0.045	0.122***
	(0.031)	(0.013)	(0.052)	(0.030)	(0.036)	(0.046)
Observations	8,721	6,640	7,506	6,027	5,304	9,918
Number of females	3,770	2,911	3,222	2,753	2,456	4,274
Number of clusters	7,372	5,880	6,561	5,316	4,723	8,421
Mean outcome	0.214	0.025	0.103	0.160	0.058	0.362
Mean husband current migrant	0.010	0.011	0.011	0.027	0.030	0.030
Mean father current migrant	0.002	0.002	0.002	0.004	0.004	0.004
Mean other current migrant	0.008	0.008	0.009	0.013	0.012	0.017
Controls	✓	✓	✓	✓	✓	✓
Individual FE	✓	✓	✓	✓	✓	✓
District × Year FE	✓	✓	✓	✓	✓	✓

Notes: This table reports the results of the impact of current male migrant relationship status on female employment status in urban and rural areas. The reference category is no migration experience. The dependent variable in Columns (1) and (4) is the wage work dummy, in Columns (2) and (5) is the self-employed dummy, and in Columns (3) and (6) is the unpaid family work dummy. Controls include age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. Standard errors clustered at the household-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

Table 3.10
Distribution of females in return migrant households based on return migrant's relationship.

	Urban	Rural
Return migrant relationship status		
Husband	63.49%	66.49%
Father	24.53%	17.66%
Other relation	10.53%	12.32%
Father & Husband	1.01%	2.78%
Husband & other relation	0.29%	0.68%
Father & other relation	0.14%	0.08%

Note: This table reports the distribution of females living in return migrant households based on the relationship status of the return migrant in urban and rural areas.

Source: Author's elaboration on ELMPS data (2006, 2012, and 2018).

their relationship status, and excluding females who resided with different return migrants across the waves. Table 3.11 shows that the impact on female labor supply in urban areas remains modest, exhibiting a negative effect, albeit only significant at the 10% level, on the probability of females with a father who is a return migrant to be self-employed or engaged in unpaid family work. The decrease in wage work in rural areas observed in the benchmark results primarily stems from fathers who are return migrants. The potential mechanisms behind this result are explored in Section 3.7.2. In contrast to the benchmark findings, where no significant impacts on other employment statuses were identified, the categorization of return migrants by relationship status reveals that rural females whose husbands are return migrants are more likely to be unpaid family workers. However, the latter finding requires further investigation, as the timing of the migration episode relative to the timing of marriage is crucial for understanding this effect.

Table 3.11
Return male migrant relationship and female employment status.

	Urban			Rural		
	(1) Wage work	(2) Self-emp.	(3) Unpaid work	(4) Wage work	(5) Self-emp.	(6) Unpaid work
Panel (A): Return migrant sample						
Husband return migrant	−0.000	0.012	−0.023	−0.008	−0.031	0.117**
	(0.027)	(0.040)	(0.057)	(0.033)	(0.030)	(0.050)
Father return migrant	0.005	−0.034*	−0.103*	−0.068***	−0.006	0.001
	(0.043)	(0.020)	(0.053)	(0.024)	(0.024)	(0.049)
Other return migrant	−0.048	0.054	0.057	−0.068*	−0.025	−0.047
	(0.057)	(0.038)	(0.080)	(0.035)	(0.037)	(0.055)
Observations	9,649	7,314	8,312	6,938	6,083	11,459
Number of females	4,158	3,198	3,554	3,154	2,808	4,917
Number of clusters	8,116	6,447	7,229	6,091	5,409	9,733
Mean outcome	0.218	0.026	0.107	0.165	0.060	0.367
Mean husband return migrant	0.080	0.076	0.083	0.120	0.119	0.136
Mean father return migrant	0.026	0.026	0.025	0.031	0.031	0.026
Mean other return migrant	0.010	0.010	0.012	0.015	0.016	0.020
Controls	✓	✓	✓	✓	✓	✓
Individual FE	✓	✓	✓	✓	✓	✓
District × Year FE	✓	✓	✓	✓	✓	✓

Notes: This table reports the results of the impact of return male migrant relationship status on female employment status in urban and rural areas. The dependent variable in columns (1) and (4) is the wage work dummy, in columns (2) and (5) is the self-employed dummy, and in columns (3) and (6) is the unpaid family work dummy. The reference category is no migration experience. Controls include age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. Standard errors clustered at the household-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

Marriage is a potential confounding factor when examining the impact of return migration on female labor supply, particularly considering that a considerable proportion of returnees (16.65%) report marriage as the main reason behind their return. Additionally, existing

literature highlights how marriage can influence female labor force participation (Lee et al., 2008; Nazier and Ramadan, 2018; Assaad et al., 2022). Therefore, it is crucial to explore this aspect further to understand its implications fully.

Table 3.12
Husband's return migration, marriage timing, and female employment status.

	Urban			Rural		
	(1) Wage work	(2) Self-emp.	(3) Unpaid work	(4) Wage work	(5) Self-emp.	(6) Unpaid work
Panel (A): Migration prior to marriage						
Husband return migrant	0.048	0.016	−0.044	0.032	0.039	0.335***
	(0.056)	(0.035)	(0.074)	(0.057)	(0.065)	(0.119)
Observations	8,516	6,475	7,323	5,749	5,015	9,336
Number of females	3,678	2,834	3,141	2,621	2,318	4,012
Number of clusters	7,222	5,748	6,421	5,080	4,475	7,967
Mean outcome	0.216	0.026	0.103	0.168	0.059	0.361
Mean husband return migrant	0.009	0.009	0.010	0.020	0.015	0.017
Panel (B): Migration after marriage						
Husband return migrant	−0.052	−0.041	−0.062	−0.001	−0.033	0.089
	(0.037)	(0.054)	(0.081)	(0.044)	(0.040)	(0.064)
Observations	8,850	6,714	7,624	6,042	5,357	10,098
Number of females	3,813	2,933	3,261	2,757	2,479	4,341
Number of clusters	7,556	5,991	6,727	5,364	4,803	8,725
Mean outcome	0.219	0.026	0.107	0.161	0.062	0.367
Mean husband return migrant	0.043	0.041	0.046	0.060	0.070	0.084
Controls	✓	✓	✓	✓	✓	✓
Individual FE	✓	✓	✓	✓	✓	✓
District × Year FE	✓	✓	✓	✓	✓	✓

Notes: This table reports the results of the impact of husband's return migration on female employment status in urban and rural areas. Results for return migrants who migrated before marriage are reported in Panel (A), while results for return migrants who migrated after marriage are reported in Panel (B). The reference category is no migration experience. The dependent variable in Columns (1) and (4) is the wage work dummy, in Columns (2) and (5) is the self-employed dummy, and in Columns (3) and (6) is the unpaid family work dummy. Controls include age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. Standard errors clustered at the household-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

To investigate whether the observed association between husband return migration and female unpaid family work is attributed to the migration event itself or changes in marital status and household dynamics, I distinguish females in households where the husband's migration episode occurred before marriage and the male returned to get married, from those whose husband migrated during marriage and returned for other reasons. To do that, I rely on survey questions pertaining to the migration year, final return year, and marriage year. The results are presented in Table 3.12. I find that the impact of husband's return migration on females' unpaid work is predominantly driven by returnees who came back to get married rather than those who migrated after marriage. This suggests that marriage may indeed be a significant confounding factor influencing the outcome rather

than the impact being solely driven by the return migration episode itself.[32]

3.7 Mechanisms

In this section, I aim to provide deeper insights into the various potential mechanisms underlying the diverse impact of male migration on female labor supply identified in Section 3.5. The channels through which male out- and return migration influence female employment are investigated in Subsections 3.7.1 and 3.7.2, respectively.[33]

3.7.1 Current migration

The mechanisms through which current migrants influence the economic engagement of females left behind are essential to understanding the dynamics of migration's impact on sending communities. Given the data in hand, two main channels are investigated. First, the role of remittances as a pivotal channel is examined in Section 3.7.1.1. Then, the current migrant's duration of stay in the destination country is explored as a complementary factor in Section 3.7.1.2.

3.7.1.1 The role of remittances

The literature looking into the impact of male migration on the labor supply of household members left behind has extensively studied the role of remittances as a crucial mechanism. Remittances can have two distinct effects. On the one hand, they can serve as a form of compensation for the lost labor of the migrant, or an insurance mechanism, leading to a reduced labor supply within the household (Amuedo-Dorantes and Pozo, 2006; Cox-Edwards and Rodríguez-Oreggia, 2009; De Haas and Van Rooij, 2010). On the other hand, remittances can significantly affect the economic well-being of the household and potentially alleviate financial constraints on family enterprises. Consequently, this may lead to an increase in non-wage work, particularly in rural areas (Arouri and Nguyen, 2018). However, households with a current migrant and not receiving remittances may only need to replace the lost income of the migrating individual (Binzel and Assaad, 2011). The increase in wage work observed in urban areas aligns with the latter hypothesis. Nevertheless, the increase in unpaid family work in rural areas could be attributed to either the need to replace migrant labor or to a reduction in financial constraints. It is

[32] Some studies have noted a slight increase in unpaid family work for females after marriage (Hendy, 2015; Ragui Assaad and Selwaness, 2022).
[33] All results tables are reported in Appendix C.4.

anticipated that if the former scenario is the primary factor, the increase in unpaid family work will be more pronounced among females in households that do not receive remittances. Conversely, if the latter scenario is at play, it is expected to be more prominent among those in households that report having received remittances.

To shed light on these dynamics, I rely on information related to household remittances income received over the 12 months preceding the survey wave. I estimate Equation (3.1) replacing the current migrant dummy (M_{ht}) with a categorical variable that further distinguishes households with a current migrant and receiving remittances from households with a current migrant and not receiving remittances.[34,35] It is worth noting that remittance receipt is an endogenous variable. On the one hand, there may be a bias arising from self-selection into sending remittances. Non-remitters may be negatively self-selected compared to remitters, resulting in lower wages at their destination and thus decreasing their likelihood of remitting. Conversely, non-remitters could also be positively self-selected, coming from wealthier households and therefore less inclined to remit.[36] While the panel structure of the data and the use of fixed effects partially mitigate this, there's still a potential for reverse causality. For instance, female employment status might influence emigrants' decisions to send remittances, or females may choose not to work if they expect to receive remittances. Hence, estimates obtained from regressing female employment status on remittance receipt status presented in Fig. 3.2 could be biased. Nonetheless, these estimates offer insights into the role played by remittances.

Fig. 3.2(a) shows that financial constraints seem to be driving the results in urban areas. The observed increase in wage work is primarily driven by females living in households with current migrants abroad and not receiving remittances. This suggests the presence of a substitution effect, wherein females in urban areas tend to engage in wage work to compensate for the lost income of the male migrant. In contrast, in rural areas, irrespective of remittances receipt by the household left behind, females are more likely to engage in unpaid family work as shown in Fig. 3.2(b). This can be due to either the need to replace the migrating members in agricultural activities,[37] or to remittances loosening

[34] This categorical variable takes the value of 0 if there is no migrant in the household, 1 if there is a current migrant and the household receives remittances, and 2 if there is a current migrant and the household does not receive remittances.

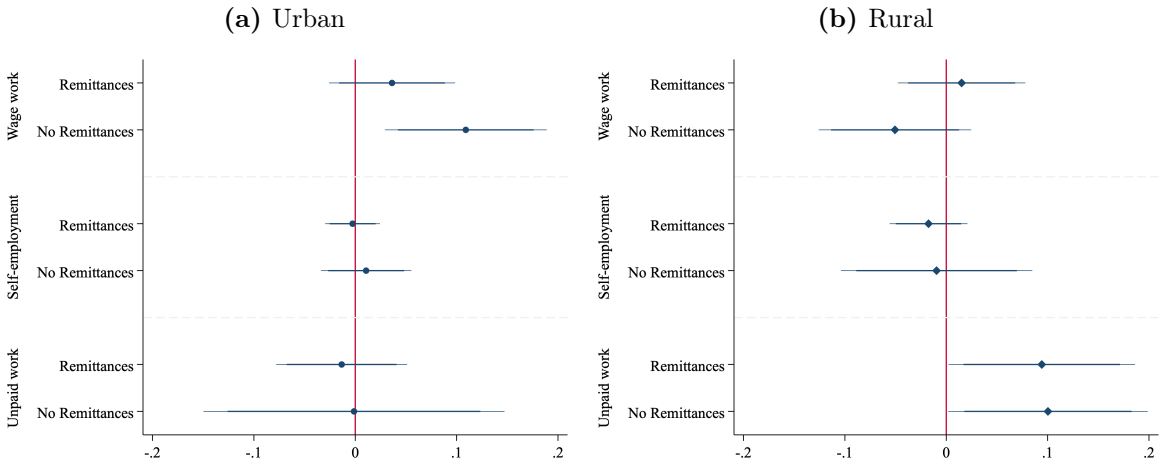
[35] I exclude non-migrant households receiving remittances. They represent a minority of the non-migrant households (2% in the urban sample and 3% in the rural sample).

[36] See Table C.6 in Appendix C.6 for a t-test on the characteristics of current migrant households with and without remittances in urban and rural areas, pointing out a negative self-selection of remitters.

[37] Around 96% of female unpaid family workers in my sample report working in the agriculture sector.

financial constraints on family enterprise. The prevalence of better job opportunities and formal employment options in urban areas, as opposed to rural regions, might account for the increase in wage work observed in urban settings. Furthermore, the more entrenched traditional gender norms and cultural expectations in rural areas may influence women's labor force decisions, leading to their engagement in unpaid family work, such as agricultural activities or household chores, aligning with traditional roles (Taylor, 1984).

Fig. 3.2. Remittances and female employment status.



Notes: These figures plot the coefficients of the impact of current migrant household remittance receipt status on female employment status in urban areas in (a) and rural areas in (b). The reference category is households with no migration experience. I control for age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children aged (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. All estimates include individual and district-year fixed effects. Standard errors are clustered at the household-year level. 90% and 95% confidence intervals are provided.
Source: Author's elaboration on ELMPS (2006, 2012, 2018).

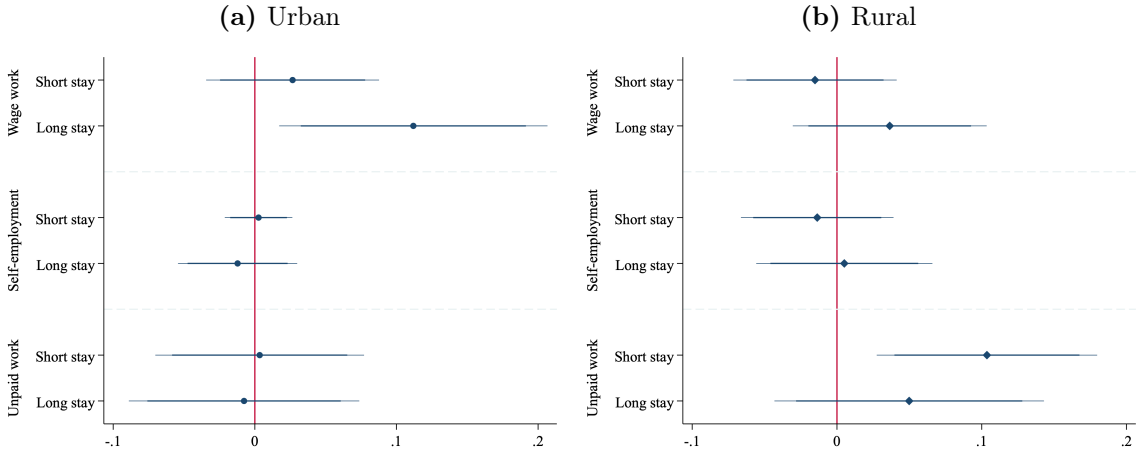
3.7.1.2 The migration duration

Another important factor, often overlooked, is the duration of current migrants' stay in the destination country. Short and long migration stays can have varying impacts on female labor supply and can provide additional insights into the dynamics at play. A longer migration duration may lead to enhanced socioeconomic integration of the migrating male member, increased access to resources and opportunities, and potentially larger remittance amounts. This can result in higher dependence on remittances and reduced female labor supply. However, it can also lead to lesser attachment over time to left-behind household members due to prolonged physical separation and a lower likelihood of remitting (Lucas and Stark, 1985; Merkle and Zimmermann, 1992). This, in turn, can translate into greater

female financial independence. I explore this aspect by using information on the migration year to calculate the number of years current migrants have been living abroad. I rely on the median migration period in my sample to identify short-term (four years or less) and long-term (more than four years) migration periods. I define a categorical variable that distinguishes households with short-term current migrants, long-term current migrants, and no migrants.

As illustrated in Fig. 3.3(a), the rise in wage work among urban females is primarily attributed to current migrants who have been abroad for more than four years. This potentially points to a higher degree of economic independence rather than increased dependence on remittances. Conversely, Fig. 3.3(b) shows that the increase in female unpaid family work observed in rural areas is attributed to having migrants who have only been living abroad for four years or less. This suggests the need to take up roles in the family business or farm shortly after one of the members migrates, potentially filling the gap left by the latter. This is even more likely to occur when the household possesses agricultural land or livestock that can not be left unattended.

Fig. 3.3. Current migrant's migration duration and female employment status.



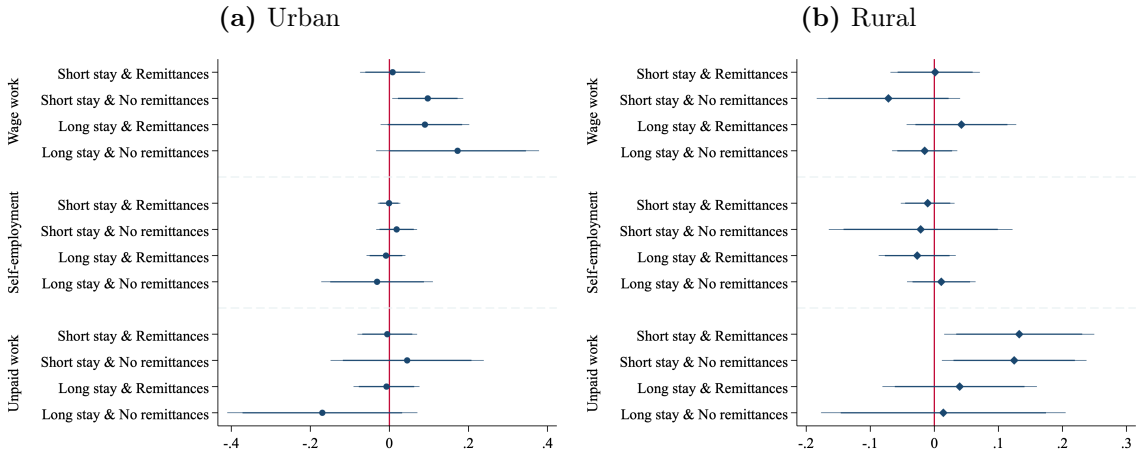
Notes: These figures plot the coefficients of the impact of current migrant's migration duration on female employment status in urban areas in (a) and rural areas in (b). The reference category is households with no migration experience. I control for age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children aged (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. All estimates include individual and district-year fixed effects. Standard errors are clustered at the household-year level. 90% and 95% confidence intervals are provided.
Source: Author's elaboration on ELMPS (2006, 2012, 2018).

To gain deeper insights into the role of migration duration, I combined information on migration duration with household remittance receipt status, resulting in four distinct

categories: households with a short-term migrant and receiving remittances, households with a short-term migrant but not receiving remittances, households with a long-term migrant and receiving remittances, and households with a long-term migrant but not receiving remittances.

As depicted in Fig. 3.4(a) below, the results indicate that current migrants who have been living abroad for a short or long period of time increase urban females' wage work only if they do not send remittances. This suggests that short-term and long-term migrants who do not send remittances may leave their families in a more financially vulnerable position. In such cases, female household members may need to seek employment to support themselves financially, leading to an increase in wage work. In rural areas, the increase in unpaid family work is driven by current migrants who have been living abroad for less than four years, irrespective of whether the household is receiving remittances as seen in Fig. 3.4(b). The latter observation confirms the need to immediately fill the labor gap left by the migrating member.

Fig. 3.4. Current migrant's migration duration, remittances, and female employment status.



Notes: These figures plot the coefficients of the impact of current migrant's migration duration and household remittance receipt status on female employment status in urban areas in (a) and rural areas in (b). The reference category is households with no migration experience. I control for age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children aged (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. All estimates include individual and district-year fixed effects. Standard errors are clustered at the household-year level. 90% and 95% confidence intervals are provided.
Source: Author's elaboration on ELMPS (2006, 2012, 2018).

In summary, these findings suggest that the increase in urban wage work can be attributed to a combination of factors. The influence of short-term migrants seems primarily motivated by economic necessities, while long-term migrants exhibit a broader impact extended

beyond remittances. As for the increase of unpaid family work in rural areas, it is predominantly driven by the immediate need to fill the labor gap left by the migrating family member.

3.7.2 Return migration

Return migrants can influence females left behind labor supply through two main channels: an economic channel and a cultural channel. The economic channel operates through return migrants' financial contributions and resource allocation within the household upon their return. Meanwhile, the cultural channel involves the potential transmission of social norms acquired in destination countries. The economic channel is explored in Section 3.7.2.1 and the cultural channel in Section 3.7.2.2.

3.7.2.1 The economic channel

Migration is often seen as an investment decision (Sjaastad, 1962), with migrants often relocating abroad to accumulate savings and acquire new skills that would benefit them in their home country (Dustmann and Glitz, 2011). The decision to return can be either planned or unplanned (Wahba, 2021). Planned returns are the consequence of migrants having reached their target, allowing them to bring back their savings or human capital accumulated abroad for use in their home country. Conversely, unplanned returns stem from either the failure to meet their goals abroad or from unforeseen events. These factors can have diverse effects on females left behind labor supply. Successful migration experiences, which entail return migrants' reintegrating into the domestic labor market or using their savings to establish businesses at home, can lead to reduced female employment. In contrast, unsuccessful migration experiences may require increased female labor force participation to offset the costs associated with migration.

I start by investigating the impact of the reasons for return on female employment. The ELMPS collects information from return migrants regarding the reasons behind their decision to return from abroad. The distribution of return reasons in my sample is presented in Table 3.13 below. I classify return migrants into two broad categories based on their return reason: (i) planned, or (ii) unplanned. A planned return occurs when the migrant willingly chooses to return to Egypt for family or economic reasons, while an unplanned return is prompted by unforeseen circumstances such as job loss, health issues, or security concerns. The categorization of planned and unplanned reasons is detailed in Table C.23 in Appendix C.4.

Table 3.13
Distribution of return reasons.

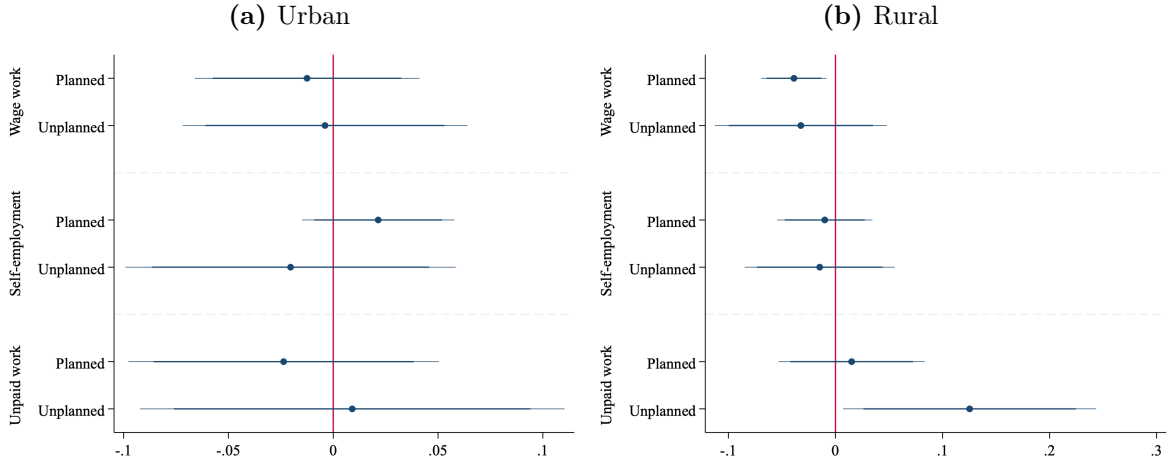
Reason	Frequency	Percent
Left work due to poor working conditions	1120	27.91
Contract ended	738	18.39
To get married	668	16.65
Security problems (wars and revolutions)	294	7.33
Family problems/care for family	275	6.85
Sudden termination by employer	209	5.21
To start up business at home country	182	4.54
Health problems	167	4.16
To look after family business or farm	126	3.14
Appointment	56	1.40
Voluntarily resign	53	1.32
Find better job	38	0.95
Deported	19	0.47
Recruitment	18	0.45
Retired	17	0.42
Army	16	0.40
Accident or illness	10	0.25
Death	4	0.10
Study	3	0.07
Total	4013	100.00

Note: This table reports the distribution of return migrants' return reasons.
Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

Since planned returns are typically driven by economic motivation or the achievement of specific financial goals, return migrants within this category are more likely to have experienced a successful migration episode. Consequently, their return might ease the financial burden on females in the household, leading to a reduced need for female labor force participation. In contrast, unplanned returns may require increased female involvement in the labor force to mitigate the sudden loss of income. Therefore, the negative impact of return migration on wage work in rural areas is more consistent with planned returns, a conclusion supported by Fig. 3.5(b), which indicates that the presence of a return migrant who planned his return is associated with a decreased probability of female engaging in wage work. In addition, I find that having a return migrant whose return was unplanned increases the probability of females working as unpaid family workers. However, there is still no significant impact on any employment status in urban areas.

Another potential factor providing insights into the return migrant's migration experience is his duration of stay abroad. Longer migration episodes are typically associated with a higher likelihood of accumulating savings and human capital abroad. Thus, a longer

Fig. 3.5. Return migrant's return reason and female employment status.

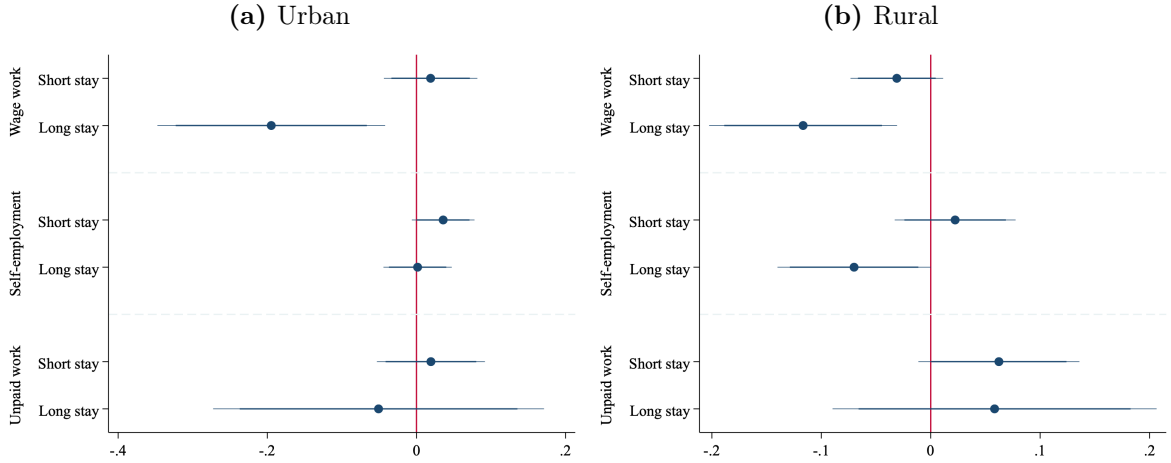


Notes: These figures plot the coefficients of the impact of return migrant's return reason on female employment status in urban areas in (a) and rural areas in (b). The reference category is households with no migration experience. I control for age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children aged (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. All estimates include individual and district-year fixed effects. Standard errors are clustered at the household-year level. 90% and 95% confidence intervals are provided.
Source: Author's elaboration on ELMPS (2006, 2012, 2018).

stay might reflect a more prosperous and successful migration experience compared to a shorter migration episode. If the decrease in the probability of female wage work is influenced by the return migrant's capacity to provide for the household, it is expected that this outcome will be more pronounced among females living with return migrants who spent a longer period abroad. Thus, I consider a categorical variable that takes the value of 0 for females living in households with no migration experience, 1 for those living in households with a return migrant who spent four years or less abroad (short stay), and 2 for those in households with a return migrant who spent more than four years abroad (long stay). As shown in Fig. 3.6(b), the decline in the probability of wage work in rural areas is mainly associated with return migrants who spent more than four years abroad, emphasizing the relationship between a longer stay abroad and female employment. The same effect is also observed in urban areas in Fig. 3.6(a).

To further investigate whether the observed decline in female wage work is linked to the ability of return migrants to financially support their households upon return, I examine their economic integration by looking at their current employment status. Return migrants who successfully secure employment are more likely to financially contribute to their households, potentially reducing the need for female labor force participation. To

Fig. 3.6. Return migrant's migration duration and female employment status.



Notes: These figures plot the coefficients of the impact of return migrant's migration duration on female employment status in urban areas in (a) and rural areas in (b). The reference category is households with no migration experience. I control for age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children aged (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. All estimates include individual and district-year fixed effects. Standard errors are clustered at the household-year level. 90% and 95% confidence intervals are provided.
Source: Author's elaboration on ELMPS (2006, 2012, 2018).

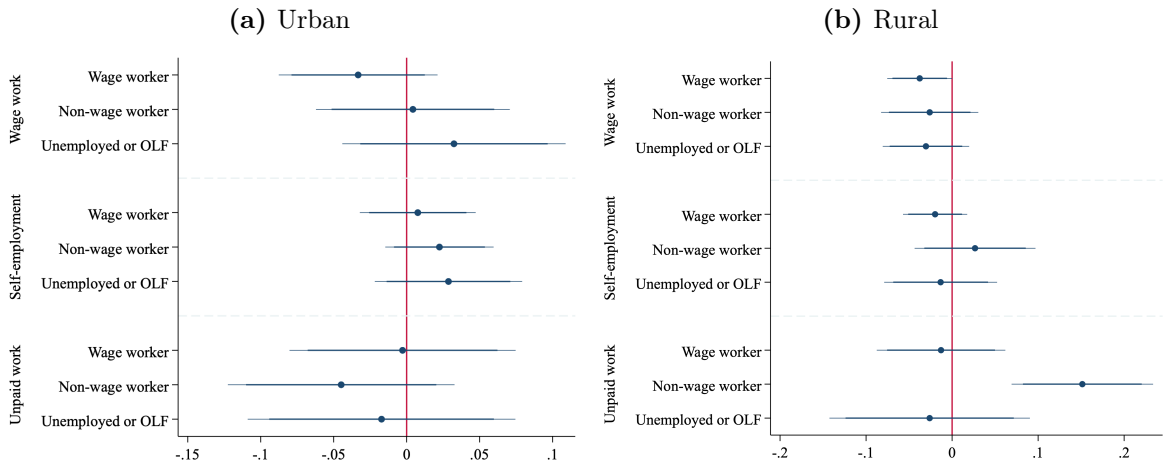
test this hypothesis, I explore the varying impact of return migrants' employment status on the employment of females left behind by introducing a categorical variable taking the value of 0 for households with no migration experience, 1 for households with a return migrant who is a wage worker, 2 for households with a return migrant who is a non-wage worker (i.e., self-employed or unpaid family worker), and 3 for households with a return migrant who is unemployed or out of the labor force (OLF). Fig. 3.7(b) shows that the decline in female wage work is predominantly driven by return migrants who are themselves wage workers, implying that return migrants who provide income to the household are the primary contributors to the decrease in female wage work.[38] Moreover, an additional impact is observed on females' unpaid family work. Interestingly, females living in a household with return migrants who are non-wage workers are more likely to engage in unpaid family work, indicating that they might be assisting their return migrant relative.[39] However, in Section 3.6.2, I found that the increase in unpaid family work is driven

[38] El-Mallakh and Wahba (2021) find that return migration of Egyptian migrants increases the probability of upward occupational mobility, especially among highly educated return migrants.

[39] Many scholars have documented that return migrants exhibit a higher tendency to become entrepreneurs (Kilic et al., 2009; Piracha and Vadean, 2010; Hamdouch and Wahba, 2015; Batista et al., 2017; Mahé, 2022) and sustain this entrepreneurial status (Marchetta, 2012; Bensassi and Jabbour, 2017).

by females' husbands who returned to get married and thus less likely to be exclusively due to the return migration episode. In contrast, no significant impact is observed on urban females' employment status.

Fig. 3.7. Return migrant's employment status and female employment status.



Notes: These figures plot the coefficients of the impact of return migrant's employment status on female employment status in urban areas in (a) and rural areas in (b). The reference category is households with no migration experience. I control for age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children aged (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. All estimates include individual and district-year fixed effects. Standard errors are clustered at the household-year level. 90% and 95% confidence intervals are provided.
Source: Author's elaboration on ELMPS (2006, 2012, 2018).

In a final assessment, I leverage information on return migrants' savings accumulated while abroad and how those savings are employed, examining their potential association with female employment status. Return migrants often allocate their savings to cover personal and family expenses or invest them in income-generating activities or businesses (Wahba, 2015b). The distribution of the use of savings by return migrants is displayed in Table 3.14 below. The majority of return migrants use their savings to buy shares (31%) or build or acquire housing (26%), followed by those who spend their savings on marriage (9%).

Return migrants' savings can affect female labor supply in various ways, including reducing the household's financial burden, offering an alternative income source that may decrease the need for female employment, or enabling females to use the capital for self-employment ventures (Mendola and Carletto, 2012). Hence, I consider a categorical variable that takes a value of 0 if the female lives in a household with no migration experience,

Table 3.14
Distribution of savings use.

Use	Frequency	Percent
Bought shares	676	30.81
Built or acquired housing	575	26.21
Spent on marriage	206	9.39
Deposited in banks	192	8.75
Established economic project	143	6.52
Saving	97	4.42
Bought agriculture land	78	3.56
Expenses at home	73	3.33
Bought appliances	72	3.28
Bought gold	29	1.32
Family Expenses	29	1.32
Bought non-agriculture land	11	0.50
Debt repayment	6	0.27
Deposited in financial company	2	0.09
Aid for parents or children	2	0.09
Pay a loan	2	0.09
Treatment or care for family	1	0.05
Total	2194	100.00

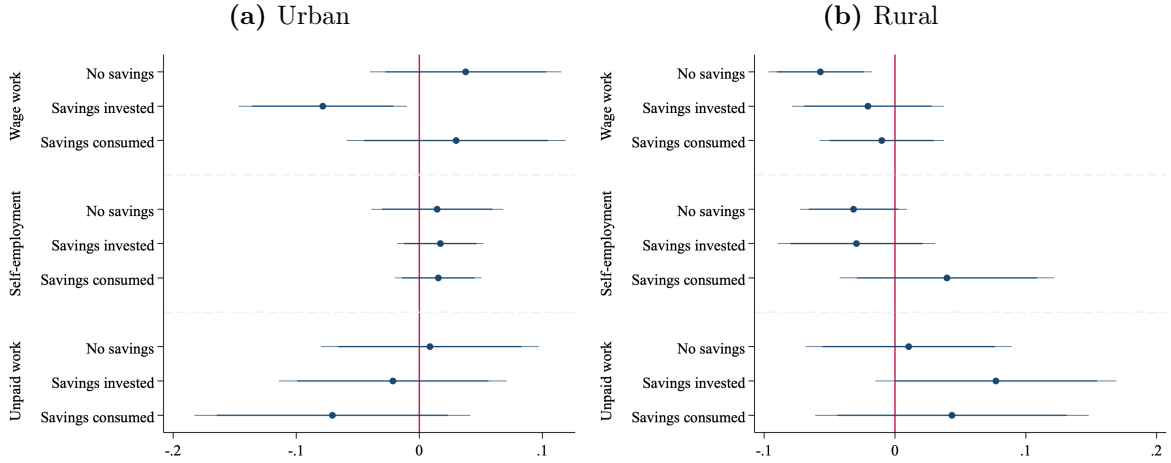
Note: This table reports the distribution of return migrants' savings use.
Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

1 if she lives in a household with a return migrant who did not accumulate any savings, 2 if she lives in a household with a return migrant who accumulated savings abroad and invested them in income-generating activities, 3 if she lives with a return migrant who accumulated savings but used them for consumption purposes.[40] The results are reported in Fig. 3.8. Unexpectedly, Fig. 3.8(b) shows that the decrease in wage work in rural areas is driven by return migrants who did not save. However, living in a household with a return migrant who invested his savings is associated with an increase in unpaid family work, further suggesting female members' assistance in family businesses. In urban areas, females in households where the return migrant invested his savings are less likely to be wage workers.

In summary, the decline in wage work in rural areas is primarily associated with return migrants who planned their return, had an extended stay abroad, are employed as wage workers, and had no savings. This implies the presence of an economic factor triggering the decline in female wage work following a male member return migration compared to

[40] Table C.27 in Appendix C.4 details the savings uses considered as an investment and those considered as a consumption.

Fig. 3.8. Return migrant's savings accumulation abroad and female employment status.



Notes: This figure plots the coefficients of the impact of return migrant's savings accumulated abroad on female employment status in urban and rural areas. The reference category is households with no migration experience. I control for age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children aged (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. All estimates include individual and district-year fixed effects. Standard errors are clustered at the household-year level. 90% and 95% confidence intervals are provided.
Source: Author's elaboration on ELMPS (2006, 2012, 2018).

when there is no migration experience in the household. However, it does not entirely rule out the potential influence of other factors. Although not consistently significant, the presence of a return migrant, regardless of their characteristics, continues to have a negative impact on female wage work. This suggests that other factors, including a potential cultural channel as discussed in Section 3.7.2.2 below, may also be at play.

3.7.2.2 The cultural channel

In Egypt, where traditional gender roles and conservative norms prevail, female labor force participation faces significant constraints. These societal norms often designate men as the primary breadwinners and women as the primary caregivers (Sieverding and Hassan, 2016; UN Women, 2017; Assaad et al., 2022). The exposure of return migrants to different gender norms and attitudes during their stay abroad can lead to potential changes in perceptions about left-behind women's roles and capabilities. As a result, return migrants may either reinforce traditional gender norms or bring back more progressive views. Consequently, return migrants may influence female labor force participation, not only through economic channels but also through their impact on social norms.

While financial remittances have traditionally monopolized economists' attention, there is

a growing interest in the analysis of what is referred to as "social remittances".[41] Migrants and returnees bring back not only economic resources but also novel ideas and perspectives that can potentially reshape opinions and behaviors within their home communities. Previous studies have documented transmission of political (Spilimbergo, 2009; Chauvet and Mercier, 2014; Barsbai et al., 2017; Tuccio et al., 2019) fertility (Beine et al., 2013; Bertoli and Marchetta, 2015), and gender norms (Tuccio and Wahba, 2018; Diabate et al., 2019; Samari, 2021).[42]

As a consequence, return migrants may influence female labor supply through the transfer of gender norms. The level of conservatism or liberalism prevalent in the host countries plays a pivotal role in this dynamic. Return migrants who have experienced more equitable gender dynamics in their host countries can advocate for similar changes in their home communities. Consequently, they can contribute to a more equitable distribution of household labor and caregiving responsibilities, thus increasing the participation of females in other productive activities, such as wage work. In contrast, return migrants from more conservative countries, where traditional gender roles are strongly emphasized, might bring back home the conservative gender norms they experienced abroad reinforcing and even further entrenching traditional gender divisions of labor within the household.

To investigate this gender norms transmission channel, I follow Tuccio and Wahba (2018) by relying on two sets of variables available in the ELMPS reflecting women's status in the household. The first set of variables measures women's freedom of mobility by asking them whether they need permission to go to the market, to go to the doctor, to visit relatives, friends, or neighbors, or to take children to the doctor. The other set reflects women's decision-making power by asking them who has the final say in making various decisions including household purchases, cooking, going to the doctor, personal purchases, taking children to the doctor, sending children to school, or buying them clothes.[43] Each variable has been re-coded to take a binary value as demonstrated in Table C.29 in Appendix C.4. These variables are aggregated to construct a composite indicator as a proxy for the underlying gender norms using three approaches: the mean, the Multiple Correspondence Analysis (MCA), and the Principal Components Analysis (PCA), with higher values indicating more liberal norms.

[41] The term social remittances was first introduced in Levitt (1998)'s seminal work that defined them as "*the ideas, behaviors, identities, and social capital that flow from receiving to sending country communities*".

[42] See Tuccio and Wahba (2020) for an extensive review of the economic literature on social remittances.

[43] There exists a third set of variables reflecting gender-egalitarian attitudes but unfortunately, it was not administrated in ELMPS2012 for me to use it in computing the gender norms indicators.

Despite a positive trend in gender gaps, the latest Arab Barometer survey conducted between 2021 and 2022 reveals that the majority of respondents in the surveyed countries still uphold traditional gender norms. However, even among Arab countries, varying degrees of conservatism still exist. I rely on the Women's Civil Liberties Index from the V-DEM dataset, which assesses the extent of women's decision-making abilities in critical aspects of their lives, as a proxy for the prevailing gender norms in the destination countries.[44] I choose to rely on the Women's Civil Liberties Index for two reasons. First, it offers a comprehensive assessment of women's freedoms and rights across multiple dimensions. Second, its extensive historical coverage, spanning from 1789 to 2021, enables me to ascertain the degree of conservatism prevalent in destination countries during the initial periods of migrants' stay.

Leveraging the information provided by the ELMPS, including data on the destination country, migration year, and the final return year of the individuals, I compute the average index score for return migrants' host countries over their period of stay. Subsequently, I compare this average with the corresponding index for Egypt over the same period. Based on this comparison, females in my sample are categorized into two groups: those living in households with a returnee from a more conservative country, characterized by a lower average index score compared to Egypt, and those living with a returnee from a less conservative country, marked by either a higher or similar average score. This classification system allows for a nuanced assessment of how the gender norms experienced in host countries might influence the returnees' potential impact on the female labor supply upon their return to Egypt. Hence, my variable of interest is a categorical variable that takes the value of 0 if there is no migration experience in the household, 1 if there is a return migrant from a more conservative country, and 2 if there is a return migrant from a less conservative country.[45]

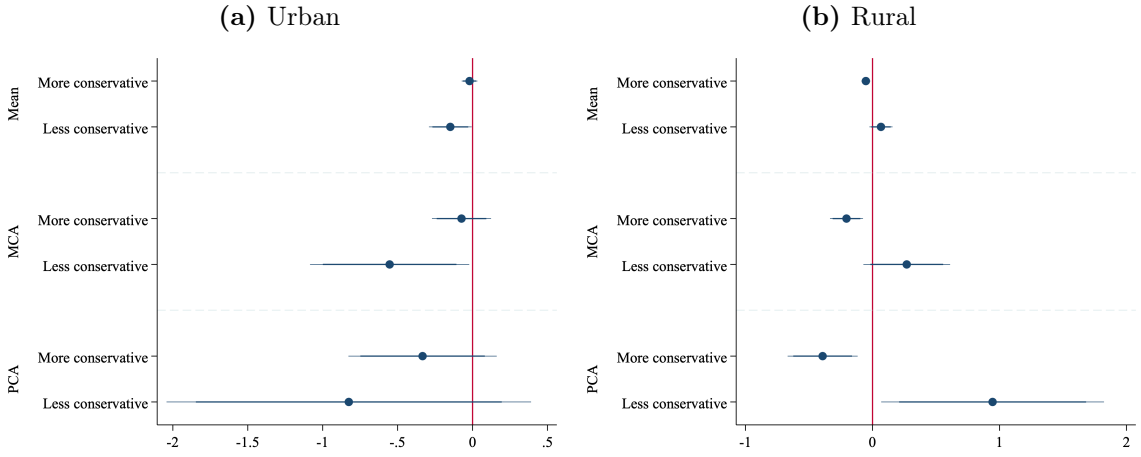
First, results on the impact of the return migrant's destination country group on the gender norms measures are reported in Fig. 3.9. Fig. 3.9(b) shows that, in rural areas, having a return migrant from a more conservative destination leads to females having a more conservative attitude, while having a return migrant from a similar or less conservative country results in more liberal attitudes as measured by the MCA and PCA.

[44] Women's civil liberties index includes measures for the freedom of domestic movement, the right to private property, the freedom from forced labor, and access to justice.

[45] As the ELMPS dataset exclusively provides information about the first destination country, I have opted to focus solely on female respondents residing in households with return migrants who experienced a single migration episode for the purpose of this analysis.

Nevertheless, this pattern is not observed in urban areas. This suggests that the degree of conservatism in the host countries may be influencing gender norms only in rural areas.

Fig. 3.9. Return migrant's destination country and female gender norms.



Notes: These figures plot the coefficients of the impact of return migrant's destination country's conservatism level compared to Egypt on gender norms in urban areas in (a) and rural areas in (a), with higher values indicating more liberal norms. The reference category is households with no migration experience. I control for age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children aged (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. All estimates include individual and district-year fixed effects. Standard errors are clustered at the household-year level. 90% and 95% confidence intervals are provided.
Source: Author's elaboration on ELMPS (2006, 2012, 2018).

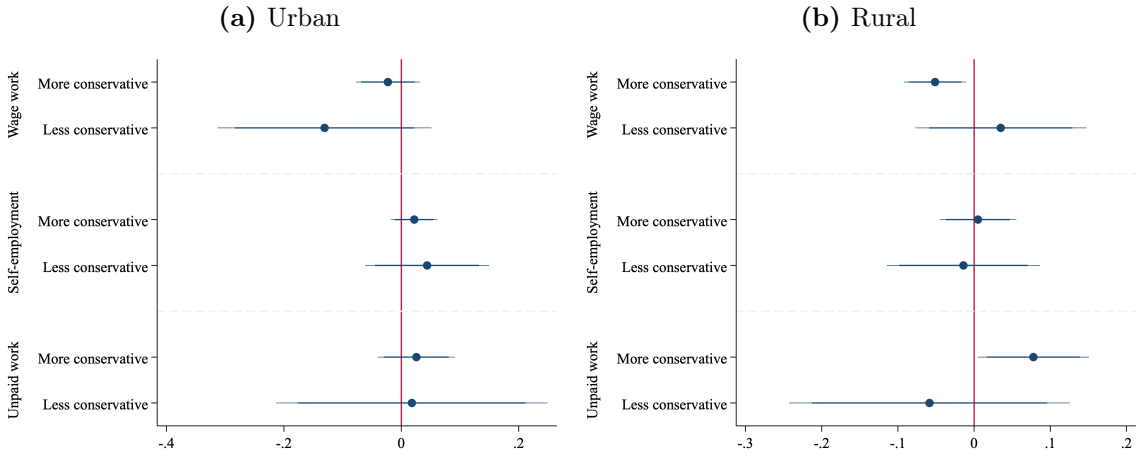
To test whether this reflects on female labor supply, I look into the impact of return migrant's destination country group on female employment status in Fig. 3.10. Results in Fig. 3.10(b) indicate that the decrease in wage work and the increase in unpaid family work observed in rural areas are mainly driven by returnees from more conservative countries, pointing to norms transmission potentially hindering female engagement in income-generating activities. No significant impacts are observed in urban areas as depicted in Fig. 3.10(a)

Moreover, given that migrants' adoption of the host country's values and norms is more likely to occur with prolonged periods spent abroad, we can anticipate that if there is a transmission of gender norms mechanism that manifests as an impact on labor supply, returnees who have spent a more extended duration overseas will likely exert a more pronounced influence.

To test that, I split the rural sample between those with returnees who spent four years or less abroad (short stay) and those with returnees who spent more than four years

abroad (long stay). In addition, I control for the economic channel variables including return migrants' employment status, savings abroad, and return reasons. The decrease in wage work seems to be driven by returnees from more conservative countries who spent more than four years abroad as observed in Fig. 3.11(b). Thus, further reinforcing the suggestive evidence of norms transmission. However, the increase in unpaid family work does not hold.

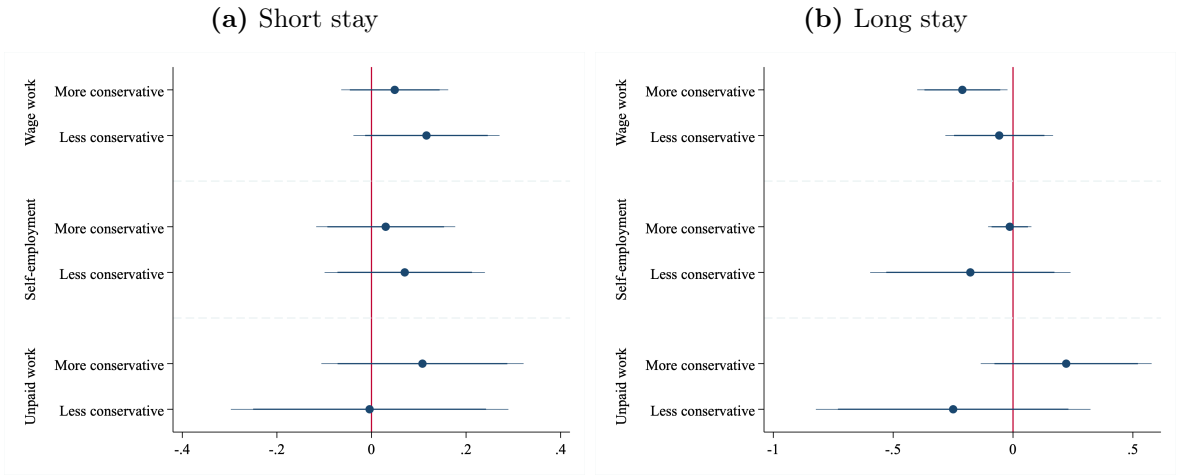
Fig. 3.10. Return migrant's destination country and female employment status.



Notes: These figures plot the coefficients of the impact of return migrant's destination country's conservatism level compared to Egypt on female employment status in urban areas in (a) and rural areas in (b). The reference category is households with no migration experience. I control for age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children aged (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. All estimates include individual and district-year fixed effects. Standard errors are clustered at the household-year level. 90% and 95% confidence intervals are provided.
Source: Author's elaboration on ELMPS (2006, 2012, 2018).

Overall, the analysis points to a complex interplay of economic and cultural factors in explaining the decrease in female wage work in rural areas following the return migration of a male household member. While economic factors play a notable role, even after accounting for those factors, the transmission of gender norms remains a significant influential factor. This suggests that cultural expectations and traditional roles continue to exert a strong pull on female labor force decisions. In contrast, the increase in unpaid work seems to be mostly driven by economic considerations.

Fig. 3.11. Return migrant's destination country, migration duration, and female employment status in rural areas.



Notes: These figures plot the coefficients of the impact of return migrant's destination country's conservatism level compared to Egypt and migration duration on female employment status in rural areas. The reference category is households with no migration experience. I control for age, age squared, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children aged (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old, return migrants' employment status, return migrants' savings use, and return migrants' return reason. All estimates include individual and district-year fixed effects. Standard errors are clustered at the household-year level. 90% and 95% confidence intervals are provided.
Source: Author's elaboration on ELMPS (2006, 2012, 2018).

3.8 Conclusions

This paper investigates the role of male-dominated migration in shaping females left behind labor supply in Egypt, where female labor force participation remains relatively low, and male migration is a prevalent phenomenon. Leveraging data from three waves of the Egypt Labor Market Panel Survey (ELMPS) conducted in 2006, 2012, and 2018, I explore whether distinct migration statuses exert varying impacts on female labor supply in different geographical contexts. Employing a fixed effect estimation strategy, I find varying impacts of male current and return migration on female employment status in urban and rural areas.

First, the presence of a male member currently abroad, compared to having no migration experience in the household, increases the probability of females participating in wage work in urban areas and in unpaid family work in rural ones. The former effect is driven by husbands' current migration, while the latter is associated with other household members' migration. The effects are primarily driven by current migrants who do not send

remittances, indicating the necessity to replace the labor lost due to the migrant's absence. However, an analysis of migration duration reveals that the increase in urban female wage work may extend beyond mere economic factors.

Conversely, having a return migrant, particularly the father, tends to reduce the probability of women engaging in wage work in rural areas while having no significant impact on labor supply in urban settings. This suggests that the migration effect may not be sustained in urban areas as migrants return, but in rural regions, return migrants may introduce additional complexities. I find that the decline in rural female wage work is not solely attributed to return migrants' ability to provide for their households upon return; it is also influenced by the potential transmission of gender norms.

In light of these findings, it appears that male out-migration can provide an opportunity for women to join income-generating activities, particularly in urban areas. This can enhance their economic independence, self-esteem, and bargaining power within the household, aligning with the idea that women's participation in paid employment outside the home is a crucial pathway to their economic empowerment (Kabeer, 2021). Nonetheless, this impact might not be sustained with the migrant return. In rural areas, the impact of male out-migration may differ, where it leads to increased involvement in unpaid family work which does not necessarily translate into greater female autonomy. Moreover, the return of male migrants in rural areas is associated with reduced economic engagement of females left behind, either because the returnee brings back resources or reintroduces traditional gender norms that hinder women's economic participation.

In essence, male migration of household members emerges as a double-edged sword with both positive and negative impacts on female labor force participation, contingent on migration status, migrant relationship status, migration experience, and geographic context. From a policy perspective, these findings underscore the significance of examining the impact of both current and return migration on the employment patterns of non-migrating females. This insight is crucial for understanding the contribution of temporary migration experiences to the well-being of sending households and economic growth at the origin. Increased female labor force participation is not only linked to poverty reduction but also improves the living standards of other household members, particularly children (Kabeer, 2021; Thomas, 1990). Therefore, development strategies should consider these gendered and geographic dimensions of migration episodes for a more holistic approach to socio-economic development.

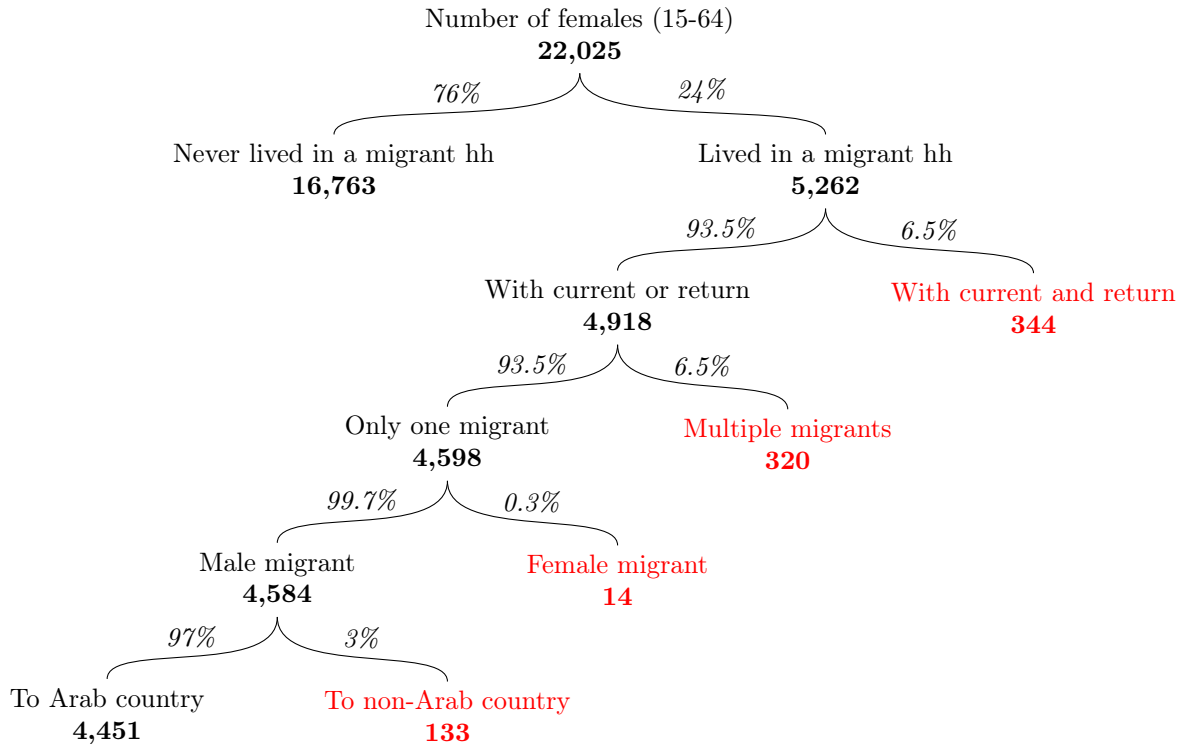
C Appendix to Chapter 3

C.1 Baseline sample description

C.1.1 Females excluded from the baseline sample

This paper uses three waves of the Egypt Labor Market Panel Survey (ELMPS 2006, ELMPS 2012, and ELMPS 2018) to investigate the impact of male migration on the labor supply of females left behind. My baseline sample focuses on working-age females aged 15 to 64 with no migration experience, who have been surveyed at least once while residing with a male household member. After accounting for missing variables, this consists of a total of 22,025 females across the three waves of the ELMPS. To enhance the precision of my analysis, certain criteria are applied to exclude specific females from the sample. Fig. C.1 provides a summary of the applied criteria, with those highlighted in red representing females excluded from the baseline sample. Among the 22,025 females considered, 5,262 have lived at least once in a migrant household (either current or return migrant), while 16,763 have never lived in a household with a migration experience. This constitutes approximately 24% and 76% of the total number of females, respectively. Of those who have lived in a migrant household, 344 (6.5%) females have resided in a household where both a male member was abroad and a male member was a return migrant simultaneously, while 4,918 (93.5%) were observed in either a current migrant or a return migrant household. Among the latter group, 4,598 (93.5%) females had only one current or return migrant at a time, while 320 (6.5%) were observed in households with multiple migrants. Among females in households with only one current or return migrant, the vast majority, 4,584 (99.7%) had a male current or return migrant, while only 14 (0.3%) had a female migrant. Finally, among females living with only one male migrant, 4,451 (97%) resided in a household with a current migrant in or a return migrant from an Arab country, compared to only 133 (3%) with a migrant who moved to a non-Arab country. The females I exclude from my baseline sample are those residing in households hosting both return and current migrants simultaneously, those in households with multiple current or return migrants, those in households with a female current or return migrant, and those in households with a current or return migrant from a non-Arab country. In Section 3.5.3, I test the robustness of the results to keeping these observations in my sample and my main conclusions remain unchanged.

Fig. C.1. Females excluded from the baseline sample.



Note: This figure illustrates the criteria based on which females are excluded from the baseline sample, along with the percentage of females excluded.
Source: Author's elaboration on ELMPS (2006, 2012, 2018).

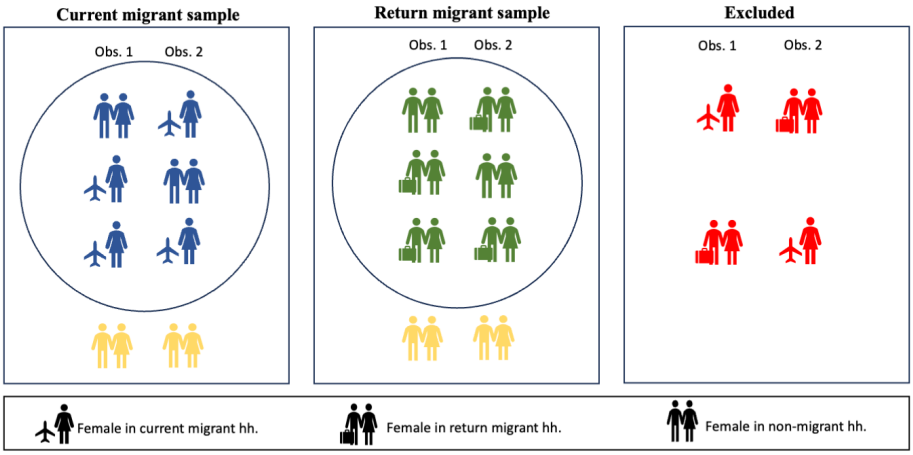
C.1.2 Females' Assignment to current and return migrant samples

The aim of this study is to differentiate the impact of having a current migrant in the household from that of having a return migrant, relative to not having any migration experience on female labor supply patterns in urban and rural areas. To achieve this, my two variables of interest are dummy variables that represent household migration status. The current migrant dummy variable takes the value of 1 if the female lives in a household where a male member is currently abroad, and 0 if she lives in a household where there is no migration experience (neither a current nor a return migrant). The return migrant dummy variable takes the value of 1 if the female lives in a household with a male return migrant and 0 if she lives in a household where there is no migration experience. Since I use a fixed-effect estimation strategy, I only keep females who appear at least twice in my baseline sample and my estimates represent a within-female impact. Given the definition of my variables of interest and the panel structure of my data, it is not obvious how females are assigned to the current and return migrant samples. Hence, Fig. C.2 presents females' assignment to each sample for those who appear twice across the three

waves as an illustrative example. Knowing that female household migration status can vary from one status to another in any wave, all possible migration status combinations are expressed in the figure below.

Females belonging to the current migrant sample are denoted in blue, those belonging to the return migrant sample are denoted in green, those included in both samples are denoted in yellow, and those excluded from both samples are denoted in red. Each pair of icons on the same line reflects the household migration status of the same female when first observed and when observed in a subsequent wave. If a female is observed living at least once in a current migrant household and never in a return migrant one, she is part of the current migrant sample, while if she is observed living at least once in a return migrant household and never in a current migrant one, she is part of the return migrant sample. However, females who have never lived in a migrant household will be included in both current and return migrant samples. In contrast, females observed once in a current migrant household and once in a return migrant household are excluded from my baseline sample since the aim is not to assess the impact of having a current migrant versus a return migrant on females' labor supply but rather the impact of having a current or a return migrant compared to not having any migration experience in the household. The same logic applies to females observed in the three waves.

Fig. C.2. Females' assignment to current and return migrant samples.

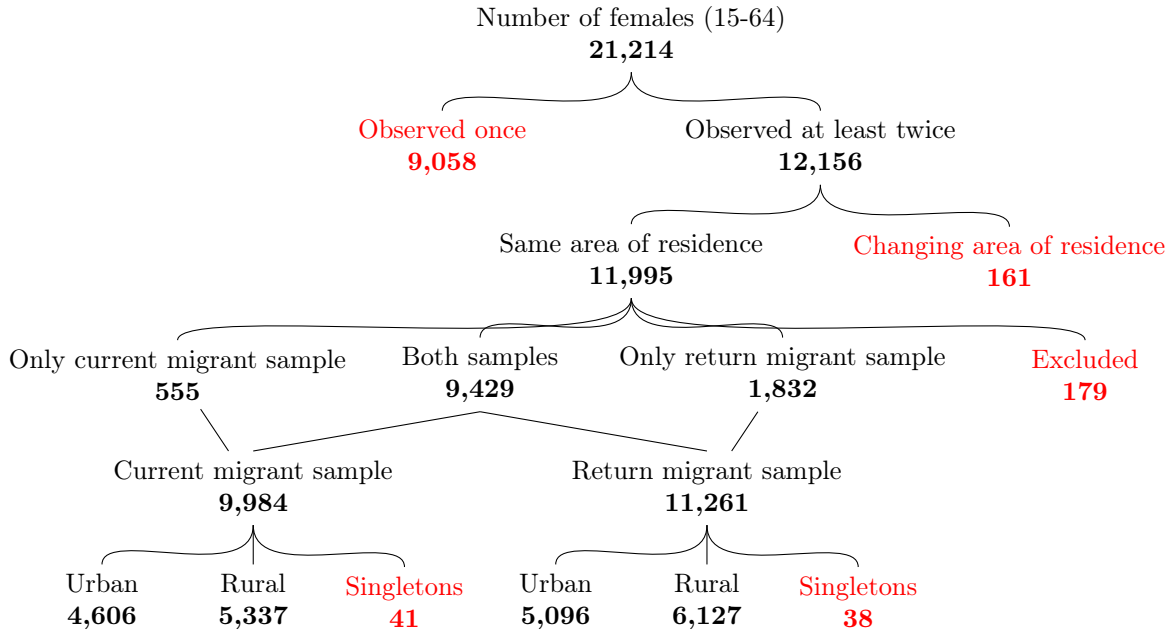


Notes: This figure illustrates the assignment of females observed twice across the three survey waves, indicating their inclusion in either the current migrant sample (depicted by blue icons) or the return migrant sample (depicted by green icons). Yellow icons represent female observations appearing in both samples, while red icons represent female observations excluded from both samples. Each pair of icons on the same line denotes the household migration status of the same female when first observed and when observed in a subsequent wave.
Source: Author's elaboration on ELMPS (2006, 2012, 2018).

C.1.3 Final current and return migrant baseline samples

To sum up, after excluding females specified in Section C.1.1, the total number of remaining females is 21,214. In light of the fixed effect model, I retain only those females observed at least twice, resulting in 12,156 individuals. Given the focus on the impact within urban and rural areas separately, I further filtered them to include only those observed at least twice in either setting, yielding 11,995 females and excluding 161 individuals who moved between urban and rural areas. Among the 11,995 females, 179 are excluded (refer to Fig. C.2). This process resulted in 555 females exclusively in the current migrant sample, 1,832 exclusively in the return migrant sample, and 9,429 appearing in both samples. After dropping singletons, the final baseline samples include 4,606 urban and 5,337 rural females in the current migrant sample, and 5,096 urban and 6,127 rural females in the return migrant sample.

Fig. C.3. Final baseline samples.



Notes: This figure illustrates the number of females included in the final current and return migrant baseline samples by area of residence.
Source: Author's elaboration on ELMPS (2006, 2012, 2018).

Considering the panel structure of the data and the use of individual fixed effects, it is important to display the share of females who experienced a change in their household migration status, as they play a crucial role in identifying the effect. Table C.1 presents the percentage of females with both constant and varying household migration statuses in both the current and return migrant samples, categorized by area of residence. In the current migrant sample, 3.82% and 7.68% of females experienced a change in urban and rural areas, respectively. In the return migrant sample, these percentages rise to 4.04% and 8.29%.

Table C.1
Percentage of females with constant and variable household migration status.

	Urban		Rural	
	Constant	Variable	Constant	Variable
Panel (A): Current migrant sample				
Percentage of females	96.18%	3.82%	92.32%	7.68%
Number of females	4,430	176	4,927	410
Panel (B): Return migrant sample				
Percentage of females	95.96%	4.04%	91.71%	8.29%
Number of females	4,890	206	5,619	508

Note: This table reports the percentage of females with constant and varying household migration status in the current and return migrant samples by area of residence.
Source: Author's elaboration on ELMPS data (2006, 2012, and 2018).

More specifically, Table C.2 and Table C.3 are two-way tables presenting household migration status at two distinct time points, namely year t and year t+6, for females observed twice and three times, respectively. The rows of the table categorize females based on their household migration status in year t (classified as either "no migrant" or "migrant"), while the columns represent the corresponding female migration status in year t+6. The cells within the table contain percentages, revealing the distribution of females transitioning between these categories over the specified time span. It is noteworthy that the majority of females with a constant household migration status are those who have resided in households that never underwent any migration.

Table C.2
Two-way table of household migration status of females observed two times in t and t+6.

	HH. Migration status in t+6			
	Urban		Rural	
HH. Migration status in t	No migrant	Migrant	No migrant	Migrant
Panel (A): Current migrant sample				
No migrant	94.98%	2.02%	88.76%	3.78%
Migrant	2.09%	0.92%	4.44%	3.03%
Panel (B): Return migrant sample				
No migrant	87.33%	1.01%	78.85%	2.52%
Migrant	2.19%	9.47%	4.90%	13.73%

Note: This table reports the percentage of females at the intersection of two household migration statuses in t and t+6 (2006/2012 or 2012/2018) for females observed twice in the baseline sample.
Source: Author's elaboration on ELMPS data (2006, 2012, and 2018).

Table C.3
Two-way table of household migration status of females observed three times in t and t+6.

	HH. Migration status in 2012			
	Urban		Rural	
HH. Migration status in 2006	No migrant	Migrant	No migrant	Migrant
Panel (A): Current migrant sample				
No migrant	97.55%	0.96%	94.40%	2.20%
Migrant	0.91%	0.59%	2.00%	1.40%
Panel (B): Return migrant sample				
No migrant	86.05%	1.69%	78.46%	3.18%
Migrant	1.36%	10.90%	3.47%	14.89%
	HH. Migration status in 2018			
	Urban		Rural	
HH. Migration status in 2012	No migrant	Migrant	No migrant	Migrant
Panel (A): Current migrant sample				
No migrant	97.34%	1.12%	94.35%	2.05%
Migrant	1.28%	0.27%	1.90%	1.70%
Panel (B): Return migrant sample				
No migrant	86.43%	0.99%	80.80%	1.13%
Migrant	1.46%	11.13%	2.34%	15.73%

Note: This table reports the percentage of females at the intersection of two household migration statuses in t and t+6 for females observed three times in the baseline sample.

Source: Author's elaboration on ELMPS data (2006, 2012, and 2018).

C.1.4 Additional description statistics

Table C.4
Summary statistics for the current migrant sample - pooled cross-section.

	Urban				Rural			
	Mean	Std. dev.	Min.	Max.	Mean	Std. dev.	Min.	Max.
Migration status								
Current migrant	0.02	0.15	0	1	0.05	0.23	0	1
Outcome variables								
Market LFP	0.27	0.44	0	1	0.22	0.42	0	1
Extended LFP	0.32	0.47	0	1	0.46	0.50	0	1
Wage work	0.17	0.38	0	1	0.08	0.28	0	1
Self-employment	0.02	0.13	0	1	0.03	0.18	0	1
Unpaid family work	0.07	0.26	0	1	0.30	0.46	0	1
Individual characteristics								
Age 15-24	0.22	0.41	0	1	0.26	0.44	0	1
Age 25-34	0.31	0.46	0	1	0.33	0.47	0	1
Age 35-44	0.20	0.40	0	1	0.19	0.39	0	1
Age 45-54	0.17	0.37	0	1	0.13	0.34	0	1
Age 55-64	0.10	0.30	0	1	0.08	0.27	0	1
Married dummy	0.71	0.46	0	1	0.77	0.42	0	1
Student dummy	0.09	0.28	0	1	0.07	0.25	0	1
Years of schooling	9.88	5.03	0	23	6.92	5.46	0	20
Years of schooling squared	122.98	85.22	0	529	77.62	76.19	0	400
Work experience dummy	0.37	0.48	0	1	0.34	0.47	0	1
Household characteristics								
Household wealth	0.36	0.96	-3	5	-0.36	0.73	-3	4
Non labor income	475.20	1532.38	0	53000	373.32	1707.32	0	101700
Household size	4.46	1.72	1	17	4.97	2.26	1	21
No. children 0-2	0.28	0.52	0	4	0.40	0.60	0	4
No. children 3-5	0.29	0.54	0	3	0.40	0.62	0	5
No. children 6-12	0.59	0.85	0	7	0.76	0.95	0	6
No. elderly ≥ 65	0.14	0.37	0	3	0.14	0.38	0	3
Number of females	4,606				5,337			
Number of female-year obs.	11,089				12,674			

Note: This table reports the mean, standard deviation, minimum, and maximum of the main outcome and explanatory variables in the current migrant sample.
Source: Author's elaboration on ELMPS data (2006, 2012, and 2018).

Table C.5
Summary statistics for the return migrant sample - pooled cross-section.

	Urban				Rural			
	Mean	Std. dev.	Min.	Max.	Mean	Std. dev.	Min.	Max.
Migration status								
Return migrant	0.12	0.32	0	1	0.18	0.38	0	1
Outcome variables								
Market LFP	0.27	0.45	0	1	0.23	0.42	0	1
Extended LFP	0.33	0.47	0	1	0.46	0.50	0	1
Wage work	0.18	0.38	0	1	0.09	0.28	0	1
Self-employment	0.02	0.14	0	1	0.03	0.18	0	1
Unpaid family work	0.08	0.27	0	1	0.30	0.46	0	1
Individual characteristics								
Age 15-24	0.22	0.41	0	1	0.26	0.44	0	1
Age 25-34	0.31	0.46	0	1	0.33	0.47	0	1
Age 35-44	0.21	0.41	0	1	0.20	0.40	0	1
Age 45-54	0.17	0.37	0	1	0.14	0.34	0	1
Age 55-64	0.10	0.30	0	1	0.08	0.27	0	1
Married dummy	0.71	0.45	0	1	0.78	0.41	0	1
Student dummy	0.09	0.29	0	1	0.07	0.25	0	1
Years of schooling	9.94	5.01	0	23	6.91	5.47	0	23
Years of schooling squared	123.86	84.92	0	529	77.74	76.30	0	529
Work experience dummy	0.37	0.48	0	1	0.35	0.48	0	1
Household characteristics								
Household wealth	0.35	0.95	-3	5	-0.37	0.72	-3	4
Non labor income	468.65	1638.21	0	53000	365.75	1768.35	0	101700
Household size	4.52	1.73	1	17	5.06	2.26	1	25
No. children 0-2	0.27	0.51	0	4	0.39	0.60	0	4
No. children 3-5	0.29	0.54	0	3	0.40	0.62	0	5
No. children 6-12	0.61	0.86	0	7	0.78	0.95	0	7
No. elderly ≥ 65	0.13	0.37	0	3	0.14	0.38	0	3
Number of females	5,096				6,127			
Number of female-year obs.	12,321				14,645			

Note: This table reports the mean, standard deviation, minimum, and maximum of the main outcome and explanatory variables in the return migrant sample.
Source: Author's elaboration on ELMPS data (2006, 2012, and 2018).

Table C.6
Characteristics of current migrants households with and without remittances' in urban and rural areas.

	Urban			Rural		
	With remittance	Without remittance	P-value	With remittance	Without remittance	P-value
Household characteristics						
Household wealth	0.758	0.472	0.073*	-0.155	-0.043	0.161
Non labor income	1255.596	274.017	0.007***	511.858	303.206	0.038**
Household size	3.702	3.534	0.526	4.220	3.794	0.030**
No. children 0-2	0.211	0.364	0.086*	0.305	0.496	0.003***
No. children 3-5	0.263	0.280	0.844	0.277	0.434	0.008***
No. children 6-12	0.404	0.678	0.039**	0.638	0.783	0.125
No. elderly ≥ 65	0.140	0.042	0.032**	0.113	0.124	0.775
Education						
Less than basic	0.071	0.103	0.511	0.199	0.223	0.559
Basic education	0.018	0.188	0.002***	0.092	0.127	0.281
Secondary education	0.500	0.385	0.152	0.468	0.501	0.504
Post-secondary education	0.411	0.325	0.271	0.241	0.149	0.015**
Main destination countries						
Iraq	0.018	0.000	0.147	0.007	0.000	0.113
Jordan	0.036	0.059	0.514	0.149	0.146	0.945
Kuwait	0.125	0.186	0.312	0.199	0.279	0.065*
Libya	0.036	0.017	0.443	0.092	0.048	0.062*
Qatar	0.018	0.034	0.557	0.043	0.037	0.757
Saudi Arabia	0.643	0.602	0.605	0.418	0.423	0.934
Emirates	0.107	0.093	0.774	0.071	0.054	0.457
Number of households	57	118		141	355	

Note: This table reports the t-test statistics by remittance receipt status for current migrants households in urban and rural areas: *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, and 2018).

C.2 Baseline results with control variables

Table C.7
Current male migration and female labor force participation.

	Urban		Rural	
	(1) Market LFP	(2) Extended LFP	(3) Market LFP	(4) Extended LFP
Current migrant	0.021	0.033	0.012	0.085***
	(0.038)	(0.042)	(0.025)	(0.031)
Age 25-34	0.044**	0.043*	0.016	0.054**
	(0.020)	(0.022)	(0.016)	(0.022)
Age 35-44	0.110***	0.113***	0.018	0.057
	(0.033)	(0.036)	(0.028)	(0.037)
Age 45-54	0.157***	0.144***	0.018	0.023
	(0.043)	(0.050)	(0.040)	(0.054)
Age 55-64	0.149***	0.116*	−0.049	−0.063
	(0.055)	(0.063)	(0.053)	(0.071)
Married	−0.194***	−0.204***	−0.069***	−0.073***
	(0.028)	(0.034)	(0.019)	(0.026)
Years of schooling	0.008	0.008	−0.005	0.002
	(0.011)	(0.012)	(0.009)	(0.013)
Years of schooling squared	0.001***	0.001*	0.002***	0.001
	(0.000)	(0.001)	(0.000)	(0.001)
Student dummy	−0.326***	−0.328***	−0.237***	−0.311***
	(0.027)	(0.028)	(0.022)	(0.032)
Work experience dummy	0.524***	0.500***	0.650***	0.418***
	(0.025)	(0.027)	(0.021)	(0.024)
Household wealth	0.012	0.026**	0.003	0.009
	(0.009)	(0.010)	(0.009)	(0.012)
Non labor income	−0.000	−0.000	−0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Household size	0.005	−0.001	0.002	0.019***
	(0.006)	(0.007)	(0.003)	(0.005)
No. children 0-2	−0.019	−0.019	−0.022**	−0.041***
	(0.013)	(0.015)	(0.009)	(0.012)
No. children 3-5	−0.025**	−0.009	−0.010	−0.030**
	(0.011)	(0.015)	(0.008)	(0.012)
No. children 6-12	0.012	0.014	0.004	−0.004
	(0.009)	(0.010)	(0.006)	(0.008)
No. elderly ≥ 65	−0.008	−0.012	0.017	0.008
	(0.017)	(0.019)	(0.015)	(0.020)
Observations	11,089	11,089	12,674	12,674
Number of females	4,606	4,606	5,337	5,337
Number of clusters	9,013	9,013	10,472	10,472
Mean outcome	0.269	0.323	0.224	0.458
Mean current migrant	0.022	0.022	0.055	0.055

Notes: This table reports the baseline results of the impact of current male migration on female labor force participation in urban and rural areas. The dependent variable in Columns (1) and (3) is the market labor force participation dummy, while in Columns (2) and (4) it is the extended labor force participation dummy. Current migrant is a dummy variable equal to 1 if there is a current migrant abroad, and 0 if there is no migration experience in the household. All estimates include individual and district-year fixed effects. Standard errors clustered at the household-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

Table C.8
Return male migration and female labor force participation.

	Urban		Rural	
	(1) Market LFP	(2) Extended LFP	(3) Market LFP	(4) Extended LFP
Return migrant	−0.009	−0.015	−0.019	0.012
	(0.029)	(0.037)	(0.022)	(0.030)
Age 25-34	0.041**	0.043**	0.027*	0.058***
	(0.019)	(0.021)	(0.015)	(0.020)
Age 35-44	0.112***	0.113***	0.031	0.060*
	(0.030)	(0.034)	(0.026)	(0.034)
Age 45-54	0.169***	0.153***	0.038	0.035
	(0.040)	(0.045)	(0.038)	(0.050)
Age 55-64	0.163***	0.125**	−0.031	−0.076
	(0.051)	(0.058)	(0.049)	(0.066)
Married	−0.193***	−0.201***	−0.083***	−0.074***
	(0.026)	(0.032)	(0.018)	(0.025)
Years of schooling	0.006	0.005	−0.011	−0.007
	(0.010)	(0.012)	(0.008)	(0.012)
Years of schooling squared	0.001***	0.001*	0.002***	0.001**
	(0.000)	(0.001)	(0.000)	(0.001)
Student dummy	−0.332***	−0.331***	−0.220***	−0.280***
	(0.026)	(0.027)	(0.022)	(0.031)
Work experience dummy	0.523***	0.497***	0.639***	0.425***
	(0.024)	(0.026)	(0.019)	(0.023)
Household wealth	0.016*	0.029***	0.005	0.026**
	(0.008)	(0.010)	(0.008)	(0.011)
Non labor income	−0.000*	−0.000	−0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Household size	0.006	−0.000	0.001	0.016***
	(0.005)	(0.007)	(0.003)	(0.004)
No. children 0-2	−0.020	−0.019	−0.013	−0.028**
	(0.012)	(0.014)	(0.009)	(0.012)
No. children 3-5	−0.023**	−0.009	−0.003	−0.024**
	(0.010)	(0.014)	(0.008)	(0.011)
No. children 6-12	0.014*	0.014	0.003	−0.001
	(0.008)	(0.009)	(0.006)	(0.008)
No. elderly ≥ 65	−0.004	−0.010	0.012	0.016
	(0.016)	(0.018)	(0.014)	(0.019)
Observations	12,321	12,321	14,645	14,645
Number of females	5,096	5,096	6,127	6,127
Number of clusters	9,966	9,966	12,085	12,085
Mean outcome	0.272	0.328	0.228	0.462
Mean return migrant	0.117	0.117	0.176	0.176

Notes: This table reports the baseline results of the impact of return male migration on female labor force participation in urban and rural areas. The dependent variable in Columns (1) and (3) is the market labor force participation dummy, while in Columns (2) and (4) it is the extended labor force participation dummy. Return migrant is a dummy variable equal to 1 if there is a return migrant in the household, and 0 if there is no migration experience in the household. All estimates include individual and district-year fixed effects. Standard errors clustered at the household-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

Table C.9
Current male migration and female employment status.

	Urban			Rural		
	(1) Wage work	(2) Self-emp.	(3) Unpaid work	(4) Wage work	(5) Self-emp.	(6) Unpaid work
Current migrant	0.054**	−0.000	−0.006	0.006	−0.007	0.081**
	(0.026)	(0.011)	(0.031)	(0.026)	(0.024)	(0.034)
Age 25-34	0.050***	0.008	−0.010	0.020	0.036**	0.047*
	(0.019)	(0.013)	(0.017)	(0.015)	(0.014)	(0.024)
Age 35-44	0.120***	−0.000	−0.003	0.038	0.041	0.037
	(0.029)	(0.021)	(0.029)	(0.027)	(0.028)	(0.042)
Age 45-54	0.149***	0.002	−0.019	0.046	0.022	−0.024
	(0.039)	(0.035)	(0.042)	(0.041)	(0.045)	(0.062)
Age 55-64	0.131***	−0.035	−0.022	0.047	−0.059	−0.099
	(0.049)	(0.046)	(0.051)	(0.050)	(0.059)	(0.079)
Married	−0.118***	−0.068**	−0.005	−0.062***	−0.001	−0.001
	(0.026)	(0.027)	(0.019)	(0.021)	(0.016)	(0.028)
Years of schooling	0.007	0.020**	0.001	−0.011	0.009	0.020
	(0.010)	(0.010)	(0.009)	(0.010)	(0.008)	(0.014)
Years of schooling squared	0.001	−0.001*	−0.000	0.001***	−0.000	−0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Student dummy	−0.148***	−0.017	−0.053***	−0.108***	−0.016	−0.178***
	(0.022)	(0.011)	(0.020)	(0.019)	(0.014)	(0.035)
Work experience dummy	0.567***	0.203***	0.121***	0.501***	0.366***	0.250***
	(0.027)	(0.039)	(0.024)	(0.032)	(0.036)	(0.032)
Household wealth	0.009	0.008*	0.012*	−0.001	0.014	0.003
	(0.008)	(0.005)	(0.007)	(0.008)	(0.009)	(0.013)
Non labor income	−0.000	−0.000	−0.000	−0.000	−0.000	0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Household size	0.007	−0.005	−0.001	0.003	0.002	0.020***
	(0.006)	(0.004)	(0.005)	(0.003)	(0.003)	(0.005)
No. children 0-2	−0.014	0.009	0.002	−0.014*	−0.005	−0.038***
	(0.013)	(0.006)	(0.009)	(0.008)	(0.009)	(0.013)
No. children 3-5	−0.015	−0.000	0.006	−0.012	−0.005	−0.023*
	(0.010)	(0.007)	(0.011)	(0.008)	(0.008)	(0.013)
No. children 6-12	−0.002	0.010	0.010	−0.001	0.005	−0.008
	(0.008)	(0.007)	(0.007)	(0.006)	(0.006)	(0.010)
No. elderly ≥ 65	−0.008	−0.000	−0.023*	−0.000	0.007	0.011
	(0.015)	(0.009)	(0.013)	(0.015)	(0.014)	(0.022)
Observations	8,724	6,643	7,509	6,061	5,337	9,985
Number of females	3,771	2,912	3,223	2,769	2,472	4,305
Number of clusters	7,374	5,883	6,564	5,342	4,750	8,467
Mean outcome	0.214	0.025	0.103	0.160	0.057	0.362
Mean current migrant	0.020	0.020	0.021	0.047	0.051	0.055

Notes: This table reports the baseline results of the impact of current male migration on female employment status in urban and rural areas. The dependent variable in Columns (1) and (4) is the wage work dummy, in Columns (2) and (5) is the self-employed dummy, and in Columns (3) and (6) is the unpaid family work dummy. Current migrant is a dummy variable equal to 1 if there is a current migrant abroad, and 0 if there is no migration experience in the household. All estimates include individual and district-year fixed effects. Standard errors clustered at the household-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

Table C.10
Return male migration and female employment status.

	Urban			Rural		
	(1) Wage work	(2) Self-emp.	(3) Unpaid work	(4) Wage work	(5) Self-emp.	(6) Unpaid work
Return migrant	−0.011	0.016	−0.019	−0.038**	−0.011	0.035
	(0.027)	(0.018)	(0.036)	(0.016)	(0.020)	(0.034)
Age 25-34	0.053***	0.006	0.004	0.019	0.039***	0.052**
	(0.018)	(0.012)	(0.016)	(0.013)	(0.013)	(0.022)
Age 35-44	0.122***	0.000	0.003	0.031	0.049*	0.043
	(0.028)	(0.020)	(0.026)	(0.024)	(0.026)	(0.039)
Age 45-54	0.154***	−0.000	−0.014	0.042	0.033	−0.015
	(0.036)	(0.032)	(0.037)	(0.037)	(0.041)	(0.057)
Age 55-64	0.143***	−0.037	−0.022	0.021	−0.051	−0.119
	(0.046)	(0.043)	(0.046)	(0.047)	(0.055)	(0.074)
Married	−0.127***	−0.063**	−0.009	−0.066***	−0.017	0.009
	(0.025)	(0.026)	(0.019)	(0.019)	(0.017)	(0.027)
Years of schooling	0.003	0.019**	0.004	−0.011	0.007	0.008
	(0.009)	(0.009)	(0.010)	(0.010)	(0.008)	(0.013)
Years of schooling squared	0.001*	−0.001*	−0.000	0.001***	−0.000	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Student dummy	−0.148***	−0.018*	−0.050***	−0.102***	−0.017	−0.145***
	(0.021)	(0.010)	(0.019)	(0.018)	(0.014)	(0.034)
Work experience dummy	0.572***	0.202***	0.129***	0.511***	0.374***	0.265***
	(0.026)	(0.037)	(0.023)	(0.028)	(0.032)	(0.030)
Household wealth	0.012	0.011**	0.014**	−0.004	0.016*	0.022*
	(0.007)	(0.005)	(0.007)	(0.007)	(0.008)	(0.013)
Non labor income	−0.000	−0.000	−0.000	−0.000	−0.000	0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Household size	0.004	−0.004	−0.000	0.000	0.002	0.018***
	(0.005)	(0.004)	(0.004)	(0.003)	(0.003)	(0.005)
No. children 0-2	−0.016	0.006	0.004	−0.012	−0.001	−0.030**
	(0.012)	(0.006)	(0.009)	(0.008)	(0.008)	(0.013)
No. children 3-5	−0.020**	−0.001	0.001	−0.004	0.001	−0.023*
	(0.009)	(0.006)	(0.010)	(0.008)	(0.008)	(0.012)
No. children 6-12	−0.003	0.008	0.006	0.002	0.004	−0.007
	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)	(0.009)
No. elderly ≥ 65	−0.019	−0.005	−0.022*	−0.005	0.003	0.022
	(0.015)	(0.009)	(0.013)	(0.014)	(0.014)	(0.021)
Observations	9,670	7,332	8,330	7,002	6,140	11,552
Number of females	4,167	3,206	3,562	3,183	2,835	4,957
Number of clusters	8,126	6,457	7,238	6,139	5,453	9,796
Mean outcome	0.218	0.026	0.107	0.165	0.060	0.367
Mean return migrant	0.112	0.109	0.115	0.162	0.163	0.177

Notes: This table reports the baseline results of the impact of return male migration on female employment status in urban and rural areas. The dependent variable in Columns (1) and (4) is the wage work dummy, in Columns (2) and (5) is the self-employed dummy, and in Columns (3) and (6) is the unpaid family work dummy. Return migrant is a dummy variable equal to 1 if there is a return migrant in the household, and 0 if there is no migration experience in the household. All estimates include individual and district-year fixed effects. Standard errors clustered at the household-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

C.3 Robustness checks tables

Table C.11
Male migration and female employment status at the intensive margin.

	Urban				Rural			
	(1) Hours worked	(2) Wage work	(3) Self-emp.	(4) Unpaid work	(5) Hours worked	(6) Wage work	(7) Self-emp.	(8) Unpaid work
Panel (A): Current migrant sample								
Current migrant	−2.867	−0.064	−0.010	0.074	2.693	0.017	−0.038	0.022
	(3.067)	(0.056)	(0.036)	(0.062)	(2.664)	(0.026)	(0.032)	(0.038)
Observations	1,881	1,915	1,915	1,915	3,315	3,344	3,344	3,344
Number of females	824	836	836	836	1,489	1,503	1,503	1,503
Number of clusters	1,801	1,834	1,834	1,834	3,089	3,120	3,120	3,120
Mean outcome	35.776	0.749	0.054	0.196	24.316	0.248	0.082	0.670
Mean current migrant	0.018	0.018	0.018	0.018	0.049	0.050	0.050	0.050
Panel (B): Return migrant sample								
Return migrant	−1.637	0.000	−0.015	0.015	3.471	0.020	−0.069**	0.049
	(2.282)	(0.035)	(0.022)	(0.037)	(3.664)	(0.032)	(0.034)	(0.044)
Observations	2,183	2,219	2,219	2,219	3,935	3,957	3,957	3,957
Number of females	957	967	967	967	1,763	1,774	1,774	1,774
Number of clusters	2,088	2,123	2,123	2,123	3,673	3,697	3,697	3,697
Mean outcome	35.703	0.743	0.057	0.201	24.656	0.251	0.082	0.667
Mean return migrant	0.139	0.139	0.139	0.139	0.195	0.194	0.194	0.194
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Individual FE	✓	✓	✓	✓	✓	✓	✓	✓
District × Year FE	✓	✓	✓	✓	✓	✓	✓	✓

Notes: This table reports the baseline results of the impact of male migration on female employment status in urban and rural areas. The dependent variable in Columns (1) and (4) is the wage work dummy, in Columns (2) and (5) it is the self-employed dummy, and in Columns (3) and (6) it is the unpaid family work dummy. In Panel (A), current migrant is a dummy variable equal to 1 if there is a current migrant abroad, and 0 if there is no migration experience in the household. In Panel (B), return migrant is a dummy variable equal to 1 if there is a return migrant in the household, and 0 if there no migration experience in the household. Controls include age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. Standard errors clustered at the household-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

Table C.12
Male migration and female unemployment.

	Unemployed	
	(1) Urban	(2) Rural
Panel (A): Current migrant sample		
Current migrant	0.000	0.062*
	(0.033)	(0.032)
Observations	7,289	5,469
Number of females	3,172	2,525
Number of clusters	6,317	4,861
Mean outcome	0.078	0.068
Mean current migrant	0.022	0.055
Panel (B): Return migrant sample		
Return migrant	−0.001	−0.049**
	(0.030)	(0.022)
Observations	8,012	6,232
Number of females	3,474	2,872
Number of clusters	6,917	5,521
Mean outcome	0.076	0.065
Mean return migrant	0.109	0.160
Controls	✓	✓
Individual FE	✓	✓
District × Year FE	✓	✓

Notes: This table reports the baseline results of the impact of male migration on female unemployment in urban and rural areas. The dependent variable is the unemployed dummy. In Panel (A), current migrant is a dummy variable equal to 1 if there is a current migrant abroad, and 0 if there is no migration experience in the household. In Panel (B), return migrant is a dummy variable equal to 1 if there is a return migrant in the household , and 0 if there no migration experience in the household. Controls include age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. Standard errors clustered at the household-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

Table C.13
Male migration and female employment status - extended sample.

	Urban			Rural		
	(1) Wage work	(2) Self-emp.	(3) Unpaid work	(4) Wage work	(5) Self-emp.	(6) Unpaid work
Panel (A): Current migrant sample						
Current migrant	0.007	−0.001	−0.020	0.018	0.006	0.063**
	(0.030)	(0.010)	(0.027)	(0.023)	(0.022)	(0.030)
Observations	9,104	6,882	7,761	6,330	5,601	10,558
Number of females	3,936	3,019	3,337	2,894	2,597	4,565
Number of clusters	7,662	6,093	6,788	5,552	4,968	8,913
Mean outcome	0.220	0.028	0.103	0.159	0.060	0.365
Mean current migrant	0.025	0.024	0.026	0.053	0.059	0.063
Panel (B): Return migrant sample						
Return migrant	0.042	0.013	−0.043	−0.039**	−0.018	0.017
	(0.045)	(0.018)	(0.037)	(0.015)	(0.019)	(0.031)
Observations	10,089	7,590	8,608	7,392	6,495	12,333
Number of females	4,352	3,321	3,685	3,361	2,998	5,298
Number of clusters	8,443	6,688	7,486	6,436	5,735	10,372
Mean outcome	0.224	0.029	0.108	0.166	0.062	0.371
Mean return migrant	0.114	0.111	0.118	0.168	0.168	0.186
Controls	✓	✓	✓	✓	✓	✓
Individual FE	✓	✓	✓	✓	✓	✓
District × Year FE	✓	✓	✓	✓	✓	✓

Notes: This table reports the baseline results of the impact of male migration on female employment status in urban and rural areas. The dependent variable in Columns (1) and (4) is the wage work dummy, in Columns (2) and (5) it is the self-employed dummy, and in Columns (3) and (6) it is the unpaid family work dummy. In Panel (A), current migrant is a dummy variable equal to 1 if there is a current migrant abroad, and 0 if there is no migration experience in the household. In Panel (B), return migrant is a dummy variable equal to 1 if there is a return migrant in the household, and 0 if there no migration experience in the household. Controls include age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. Standard errors clustered at the household-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

Table C.14
Male migration and female employment status - individual clusters.

	Urban			Rural		
	(1) Wage work	(2) Self-emp.	(3) Unpaid work	(4) Wage work	(5) Self-emp.	(6) Unpaid work
Panel (A): Current migrant sample						
Current migrant	0.054**	−0.000	−0.006	0.006	−0.007	0.081**
	(0.024)	(0.011)	(0.032)	(0.026)	(0.022)	(0.034)
Observations	8,724	6,643	7,509	6,061	5,337	9,985
Number of females	3,771	2,912	3,223	2,769	2,472	4,305
Number of clusters	3,771	2,912	3,223	2,769	2,472	4,305
Mean outcome	0.214	0.025	0.103	0.160	0.057	0.362
Mean current migrant	0.020	0.020	0.021	0.047	0.051	0.055
Panel (B): Return migrant sample						
Return migrant	−0.011	0.016	−0.019	−0.038**	−0.011	0.035
	(0.026)	(0.017)	(0.034)	(0.016)	(0.020)	(0.033)
Observations	9,670	7,332	8,330	7,002	6,140	11,552
Number of females	4,167	3,206	3,562	3,183	2,835	4,957
Number of clusters	4,167	3,206	3,562	3,183	2,835	4,957
Mean outcome	0.218	0.026	0.107	0.165	0.060	0.367
Mean return migrant	0.112	0.109	0.115	0.162	0.163	0.177
Controls	✓	✓	✓	✓	✓	✓
Individual FE	✓	✓	✓	✓	✓	✓
District × Year FE	✓	✓	✓	✓	✓	✓

Notes: This table reports the baseline results of the impact of male migration on female employment status in urban and rural areas. The dependent variable in Columns (1) and (4) is the wage work dummy, in Columns (2) and (5) it is the self-employed dummy, and in Columns (3) and (6) it is the unpaid family work dummy. In Panel (A), current migrant is a dummy variable equal to 1 if there is a current migrant abroad, and 0 if there is no migration experience in the household. In Panel (B), return migrant is a dummy variable equal to 1 if there is a return migrant in the household, and 0 if there no migration experience in the household. Controls include age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. Standard errors clustered at the individual level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

Table C.15
Male migration and female employment status - primary sampling units clusters.

	Urban			Rural		
	(1) Wage work	(2) Self-emp.	(3) Unpaid work	(4) Wage work	(5) Self-emp.	(6) Unpaid work
Panel (A): Current migrant sample						
Current migrant	0.054**	−0.000	−0.006	0.006	−0.007	0.081**
	(0.027)	(0.010)	(0.032)	(0.028)	(0.027)	(0.034)
Observations	8,724	6,643	7,509	6,061	5,337	9,985
Number of females	3,771	2,912	3,223	2,769	2,472	4,305
Number of clusters	803	759	767	837	741	952
Mean outcome	0.214	0.025	0.103	0.160	0.057	0.362
Mean current migrant	0.020	0.020	0.021	0.047	0.051	0.055
Panel (B): Return migrant sample						
Return migrant	−0.011	0.016	−0.019	−0.038**	−0.011	0.035
	(0.026)	(0.017)	(0.040)	(0.016)	(0.021)	(0.035)
Observations	9,670	7,332	8,330	7,002	6,140	11,552
Number of females	4,167	3,206	3,562	3,183	2,835	4,957
Number of clusters	823	771	778	889	809	1,012
Mean outcome	0.218	0.026	0.107	0.165	0.060	0.367
Mean return migrant	0.112	0.109	0.115	0.162	0.163	0.177
Controls	✓	✓	✓	✓	✓	✓
Individual FE	✓	✓	✓	✓	✓	✓
District × Year FE	✓	✓	✓	✓	✓	✓

Notes: This table reports the baseline results of the impact of male migration on female employment status in urban and rural areas. The dependent variable in Columns (1) and (4) is the wage work dummy, in Columns (2) and (5) it is the self-employed dummy, and in Columns (3) and (6) it is the unpaid family work dummy. In Panel (A), current migrant is a dummy variable equal to 1 if there is a current migrant abroad, and 0 if there is no migration experience in the household. In Panel (B), return migrant is a dummy variable equal to 1 if there is a return migrant in the household, and 0 if there no migration experience in the household. Controls include age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. Standard errors clustered at the primary sampling unit (PSU) level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

Table C.16
Male migration and female employment status - single migration episode.

	Urban			Rural		
	(1) Wage work	(2) Self-emp.	(3) Unpaid work	(4) Wage work	(5) Self-emp.	(6) Unpaid work
Panel (A): Return migrant sample						
Return migrant	−0.025	0.016	0.018	−0.044**	−0.002	0.050
	(0.027)	(0.020)	(0.033)	(0.019)	(0.023)	(0.035)
Observations	9,467	7,186	8,145	6,717	5,884	11,020
Number of females	4,079	3,141	3,483	3,051	2,714	4,730
Number of clusters	7,957	6,327	7,073	5,890	5,225	9,345
Mean outcome	0.217	0.026	0.106	0.166	0.060	0.365
Mean return migrant	0.095	0.093	0.097	0.133	0.133	0.143
Controls	✓	✓	✓	✓	✓	✓
Individual FE	✓	✓	✓	✓	✓	✓
District × Year FE	✓	✓	✓	✓	✓	✓

Notes: This table reports the baseline results of the impact of male migration on female employment status in urban and rural areas. The dependent variable in Columns (1) and (4) is the wage work dummy, in Columns (2) and (5) it is the self-employed dummy, and in Columns (3) and (6) it is the unpaid family work dummy. In Panel (A), current migrant is a dummy variable equal to 1 if there is a current migrant abroad, and 0 if there is no migration experience in the household. In Panel (B), return migrant is a dummy variable equal to 1 if there is a return migrant in the household, and 0 if there no migration experience in the household. Controls include age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. Standard errors clustered at the household-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

Table C.17
Male migration and female employment status - alternative controls.

	Urban			Rural		
	(1) Wage work	(2) Self-emp.	(3) Unpaid work	(4) Wage work	(5) Self-emp.	(6) Unpaid work
Panel (A): Current migrant sample						
Current migrant	0.052**	0.002	−0.006	0.007	−0.006	0.069**
	(0.025)	(0.012)	(0.030)	(0.027)	(0.024)	(0.034)
Observations	8,719	6,641	7,507	6,056	5,334	9,973
Number of females	3,769	2,911	3,222	2,767	2,471	4,302
Number of clusters	7,369	5,881	6,562	5,340	4,748	8,457
Mean outcome	0.214	0.025	0.103	0.160	0.057	0.362
Mean current migrant	0.020	0.020	0.021	0.047	0.051	0.055
Panel (B): Return migrant sample						
Return migrant	−0.013	0.020	−0.022	−0.045***	−0.007	0.018
	(0.027)	(0.018)	(0.036)	(0.017)	(0.020)	(0.032)
Observations	9,663	7,328	8,327	6,995	6,135	11,537
Number of females	4,164	3,204	3,561	3,180	2,833	4,953
Number of clusters	8,120	6,454	7,235	6,136	5,450	9,784
Mean outcome	0.218	0.026	0.107	0.165	0.060	0.366
Mean return migrant	0.112	0.109	0.115	0.162	0.163	0.176
Controls	✓	✓	✓	✓	✓	✓
Individual FE	✓	✓	✓	✓	✓	✓
District × Year FE	✓	✓	✓	✓	✓	✓

Notes: This table reports the baseline results of the impact of male migration on female employment status in urban and rural areas. The dependent variable in Columns (1) and (4) is the wage work dummy, in Columns (2) and (5) it is the self-employed dummy, and in Columns (3) and (6) it is the unpaid family work dummy. In Panel (A), current migrant is a dummy variable equal to 1 if there is a current migrant abroad, and 0 if there is no migration experience in the household. In Panel (B), return migrant is a dummy variable equal to 1 if there is a return migrant in the household, and 0 if there no migration experience in the household. Controls include age categories, marital status, student dummy, educational attainment, work experience dummy, household wealth score, household non-labor income, household size, children less than 12 years old living in the household dummy, and elderly individuals over 65 years old living in the household dummy. Standard errors clustered at the household-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

Table C.18
Male migration and female employment status - additional controls.

	Urban			Rural		
	(1) Wage work	(2) Self-emp.	(3) Unpaid work	(4) Wage work	(5) Self-emp.	(6) Unpaid work
Panel (A): Current migrant sample						
Current migrant	0.046*	−0.006	−0.008	0.003	−0.017	0.092***
	(0.027)	(0.013)	(0.031)	(0.026)	(0.024)	(0.034)
Observations	8,724	6,643	7,509	6,061	5,337	9,985
Number of females	3,771	2,912	3,223	2,769	2,472	4,305
Number of clusters	7,374	5,883	6,564	5,342	4,750	8,467
Mean outcome	0.214	0.025	0.103	0.160	0.057	0.362
Mean current migrant	0.020	0.020	0.021	0.047	0.051	0.055
Panel (B): Return migrant sample						
Return migrant	−0.010	0.006	−0.020	−0.038**	−0.018	0.032
	(0.027)	(0.018)	(0.036)	(0.017)	(0.019)	(0.034)
Observations	9,670	7,332	8,330	7,002	6,140	11,552
Number of females	4,167	3,206	3,562	3,183	2,835	4,957
Number of clusters	8,126	6,457	7,238	6,139	5,453	9,796
Mean outcome	0.218	0.026	0.107	0.165	0.060	0.367
Mean return migrant	0.112	0.109	0.115	0.162	0.163	0.177
Controls	✓	✓	✓	✓	✓	✓
Individual FE	✓	✓	✓	✓	✓	✓
District × Year FE	✓	✓	✓	✓	✓	✓

Notes: This table reports the baseline results of the impact of male migration on female employment status in urban and rural areas. The dependent variable in Columns (1) and (4) is the wage work dummy, in Columns (2) and (5) it is the self-employed dummy, and in Columns (3) and (6) it is the unpaid family work dummy. In Panel (A), current migrant is a dummy variable equal to 1 if there is a current migrant abroad, and 0 if there is no migration experience in the household. In Panel (B), return migrant is a dummy variable equal to 1 if there is a return migrant in the household , and 0 if there no migration experience in the household. Controls include age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children (0-2), (3-5), and (6-12) years old in the household, the number of elderly individuals over 65 years old, family enterprise dummy, the share of male in the household, and a dummy for whether there is at least one employed male in the household. Standard errors clustered at the household-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

Table C.19
Male migration and female employment status - gov. × year FE.

	Urban			Rural		
	(1) Wage work	(2) Self-emp.	(3) Unpaid work	(4) Wage work	(5) Self-emp.	(6) Unpaid work
Panel (A): Current migrant sample						
Current migrant	0.037	0.003	−0.007	0.008	−0.013	0.090***
	(0.023)	(0.007)	(0.024)	(0.027)	(0.022)	(0.034)
Observations	8,724	6,643	7,509	6,061	5,337	9,985
Number of females	3,771	2,912	3,223	2,769	2,472	4,305
Number of clusters	7,374	5,883	6,564	5,342	4,750	8,467
Mean outcome	0.214	0.025	0.103	0.160	0.057	0.362
Mean current migrant	0.020	0.020	0.021	0.047	0.051	0.055
Panel (B): Return migrant sample						
Return migrant	−0.009	0.027	−0.014	−0.033**	−0.009	0.031
	(0.025)	(0.037)	(0.032)	(0.015)	(0.018)	(0.033)
Observations	9,670	7,332	8,330	7,002	6,140	11,552
Number of females	4,167	3,206	3,562	3,183	2,835	4,957
Number of clusters	8,126	6,457	7,238	6,139	5,453	9,796
Mean outcome	0.218	0.026	0.107	0.165	0.060	0.367
Mean return migrant	0.112	0.109	0.115	0.162	0.163	0.177
Controls	✓	✓	✓	✓	✓	✓
Individual FE	✓	✓	✓	✓	✓	✓
Governorate × year FE	✓	✓	✓	✓	✓	✓

Notes: This table reports the baseline results of the impact of male migration on female employment status in urban and rural areas. The dependent variable in Columns (1) and (4) is the wage work dummy, in Columns (2) and (5) it is the self-employed dummy, and in Columns (3) and (6) it is the unpaid family work dummy. In Panel (A), current migrant is a dummy variable equal to 1 if there is a current migrant abroad, and 0 if there is no migration experience in the household. In Panel (B), return migrant is a dummy variable equal to 1 if there is a return migrant in the household, and 0 if there no migration experience in the household. Controls include age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. Standard errors clustered at the household-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

C.4 Mechanisms tables

Table C.20
Remittances and female employment status.

	Urban			Rural		
	(1) Wage work	(2) Self-emp.	(3) Unpaid work	(4) Wage work	(5) Self-emp.	(6) Unpaid work
Current migrant sample						
Remittances	0.036	−0.003	−0.013	0.015	−0.017	0.094**
	(0.032)	(0.014)	(0.033)	(0.032)	(0.020)	(0.047)
No Remittances	0.109***	0.011	−0.001	−0.051	−0.010	0.100**
	(0.041)	(0.023)	(0.076)	(0.038)	(0.048)	(0.050)
Observations	8,365	6,347	7,165	5,681	4,957	9,356
Number of females	3,625	2,789	3,082	2,597	2,299	4,048
Number of clusters	7,080	5,624	6,269	5,004	4,413	7,946
Mean outcome	0.217	0.025	0.102	0.163	0.057	0.362
Mean remittances	0.012	0.013	0.013	0.028	0.031	0.033
Mean no remittances	0.006	0.006	0.006	0.011	0.011	0.014
Controls	✓	✓	✓	✓	✓	✓
Individual FE	✓	✓	✓	✓	✓	✓
District × Year FE	✓	✓	✓	✓	✓	✓

Notes: This table reports the results of the impact of remittances status on female employment status in urban and rural areas. The dependent variable in Columns (1) and (4) is the wage work dummy, in Columns (2) and (5) is the self-employed dummy, and in Columns (3) and (6) is the unpaid family work dummy. The independent variable is a categorical variable equal to 1 if the female lives in a household with a current migrant and receives remittances,2 if she lives in a household with a current migrant and does not receive remittances, and 0 if there is no migration experience in the household. Controls include age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. Standard errors clustered at the household-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

Table C.21
Current migrant's migration duration and female employment status.

	Urban			Rural		
	(1) Wage work	(2) Self-emp.	(3) Unpaid work	(4) Wage work	(5) Self-emp.	(6) Unpaid work
Current migrant sample						
Short stay	0.027	0.003	0.003	−0.015	−0.014	0.104***
	(0.031)	(0.012)	(0.038)	(0.029)	(0.027)	(0.039)
Long stay	0.112**	−0.012	−0.008	0.036	0.005	0.050
	(0.048)	(0.021)	(0.041)	(0.034)	(0.031)	(0.047)
Observations	8,714	6,633	7,495	6,044	5,321	9,958
Number of females	3,767	2,908	3,217	2,762	2,465	4,293
Number of clusters	7,367	5,876	6,554	5,325	4,734	8,446
Mean outcome	0.215	0.025	0.103	0.160	0.058	0.362
Mean short stay	0.010	0.011	0.011	0.023	0.024	0.029
Mean long stay	0.009	0.008	0.009	0.022	0.025	0.025
Controls	✓	✓	✓	✓	✓	✓
Individual FE	✓	✓	✓	✓	✓	✓
District × Year FE	✓	✓	✓	✓	✓	✓

Notes: This table reports the results of the impact of current migrant's duration of stay abroad on female employment status in urban and rural areas. The dependent variable in Columns (1) and (4) is the wage work dummy, in Columns (2) and (5) is the self-employed dummy, and in Columns (3) and (6) is the unpaid family work dummy. The independent variable is a categorical variable equal to 1 if the female lives in a household with a current migrant who has been living abroad for four years or less (short stay), 2 if she lives in a household with a current migrant who has been living abroad for more than four years (long stay), and 0 if there is no migration experience in the household. Controls include age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. Standard errors clustered at the household-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

Table C.22
Current migrant's migration duration, remittances, and female employment status.

	Urban			Rural		
	(1) Wage work	(2) Self-emp.	(3) Unpaid work	(4) Wage work	(5) Self-emp.	(6) Unpaid work
Current migrant sample						
Short stay & Remittances	0.008	−0.001	−0.005	0.001	−0.010	0.133**
	(0.042)	(0.015)	(0.038)	(0.036)	(0.021)	(0.060)
Short stay & No remittances	0.097**	0.018	0.045	−0.072	−0.021	0.125**
	(0.046)	(0.026)	(0.099)	(0.057)	(0.073)	(0.058)
Long stay & Remittances	0.090	−0.009	−0.007	0.042	−0.027	0.040
	(0.057)	(0.025)	(0.043)	(0.044)	(0.031)	(0.062)
Long stay & No remittances	0.173*	−0.031	−0.170	−0.015	0.011	0.014
	(0.105)	(0.072)	(0.123)	(0.026)	(0.027)	(0.097)
Observations	8,357	6,339	7,153	5,670	4,947	9,335
Number of females	3,622	2,786	3,077	2,593	2,295	4,039
Number of clusters	7,075	5,619	6,261	4,993	4,403	7,929
Mean outcome	0.217	0.025	0.102	0.163	0.057	0.363
Mean short stay & remittances	0.005	0.006	0.006	0.013	0.014	0.015
Mean short stay & no remittances	0.005	0.004	0.004	0.007	0.006	0.010
Mean long stay & remittances	0.006	0.006	0.006	0.014	0.016	0.016
Mean long stay & no remittances	0.002	0.001	0.001	0.004	0.004	0.004
Controls	✓	✓	✓	✓	✓	✓
Individual FE	✓	✓	✓	✓	✓	✓
District × Year FE	✓	✓	✓	✓	✓	✓

Notes: This table reports the results of the impact of current migrant's duration of stay abroad and remittances status on female employment status in urban and rural areas. The dependent variable in Columns (1) and (4) is the wage work dummy, in Columns (2) and (5) is the self-employed dummy, and in Columns (3) and (6) is the unpaid family work dummy. The independent variable is a categorical variable equal to 1 if the female lives in a household with a current migrant who has been living abroad for four years or less and receives remittances, 2 if she lives in a household with a current migrant who has been living abroad for four years or less and does not receive remittances, 3 if she lives in a household with a current migrant who has been living abroad for more than four years and receives remittances, 4 if she lives in a household with a current migrant who has been living abroad for more than four years and does not receive remittances, and 0 if there is no migration experience in the household. Controls include age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. Standard errors clustered at the household-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

Table C.23
Return reasons.

Planned return	Unplanned return
Left work due to poor working conditions	Security problems (wars and revolutions)
To get married	Sudden termination by employer
Contract ended	Health problems
Family problems/care for family	Deported
To start up a business in home country	Accident or illness
To look after family business or farm	Army
Appointment	Death
Voluntarily resign	
Find a better job	
Recruitment	
Retired	
Study	

Note: This table reports the categorization of the reasons for return migration into planned and unplanned reasons.
Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

Table C.24
Return migrant's return reason and female employment status.

	Urban			Rural		
	(1) Wage work	(2) Self-emp.	(3) Unpaid work	(4) Wage work	(5) Self-emp.	(6) Unpaid work
Return migrant sample						
Planned	−0.013	0.021	−0.024	−0.039**	−0.010	0.015
	(0.027)	(0.019)	(0.038)	(0.016)	(0.023)	(0.035)
Unplanned	−0.004	−0.020	0.009	−0.032	−0.015	0.125**
	(0.035)	(0.040)	(0.052)	(0.041)	(0.036)	(0.060)
Observations	9,670	7,332	8,330	7,000	6,138	11,550
Number of females	4,167	3,206	3,562	3,182	2,834	4,956
Number of clusters	8,126	6,457	7,238	6,137	5,451	9,794
Mean outcome	0.218	0.026	0.107	0.165	0.060	0.367
Mean planned	0.095	0.090	0.095	0.134	0.132	0.143
Mean unplanned	0.017	0.019	0.021	0.028	0.030	0.033
Controls	✓	✓	✓	✓	✓	✓
Individual FE	✓	✓	✓	✓	✓	✓
District × Year FE	✓	✓	✓	✓	✓	✓

Notes: This table reports the results of the impact of return migrant's reason of return on female employment status in urban and rural areas. The dependent variable in Columns (1) and (4) is the wage work dummy, in Columns (2) and (5) is the self-employed dummy, and in Columns (3) and (6) is the unpaid family work dummy. The independent variable is a categorical variable equal to 1 if the female lives in a household with a return migrant whose return was planned, equal to 2 if she lives in a household with a return migrant whose return was unplanned, and 0 if there is no migration experience in the household. Controls include age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. Standard errors clustered at the household-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

Table C.25
Return migrant's migration duration and female employment status.

	Urban			Rural		
	(1) Wage work	(2) Self-emp.	(3) Unpaid work	(4) Wage work	(5) Self-emp.	(6) Unpaid work
Return migrant sample						
Short stay	0.019	0.036*	0.019	−0.031	0.022	0.062*
	(0.032)	(0.021)	(0.037)	(0.022)	(0.028)	(0.038)
Long stay	−0.195**	0.002	−0.051	−0.117***	−0.070**	0.058
	(0.078)	(0.023)	(0.113)	(0.044)	(0.036)	(0.076)
Observations	9,482	7,203	8,166	6,725	5,901	11,046
Number of females	4,085	3,149	3,493	3,055	2,721	4,742
Number of clusters	7,971	6,343	7,092	5,907	5,248	9,369
Mean outcome	0.217	0.026	0.106	0.165	0.060	0.365
Mean short stay	0.070	0.070	0.075	0.102	0.101	0.108
Mean long stay	0.027	0.025	0.024	0.033	0.035	0.037
Controls	✓	✓	✓	✓	✓	✓
Individual FE	✓	✓	✓	✓	✓	✓
District × Year FE	✓	✓	✓	✓	✓	✓

Notes: This table reports the results of the impact of return migrant's migration duration on female employment status in urban and rural areas. The dependent variable in Columns (1) and (4) is the wage work dummy, in Columns (2) and (5) is the self-employed dummy, and in Columns (3) and (6) is the unpaid family work dummy. The independent variable is a categorical variable equal to 1 if the female lives in a household with a return migrant who spent four years or less abroad (short stay), equal to 2 if she lives in a household with a return migrant who spent more than four years abroad (long stay), and 0 if there is no migration experience in the household. Controls include age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. Standard errors clustered at the household-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

Table C.26
Return migrant's employment status and female employment status.

	Urban			Rural		
	(1) Wage work	(2) Self-emp.	(3) Unpaid work	(4) Wage work	(5) Self-emp.	(6) Unpaid work
Return migrant sample						
Wage worker	−0.033	0.008	−0.003	−0.038*	−0.020	−0.013
	(0.028)	(0.020)	(0.040)	(0.019)	(0.019)	(0.038)
Non-wage worker	0.004	0.022	−0.045	−0.026	0.027	0.151***
	(0.034)	(0.019)	(0.040)	(0.029)	(0.036)	(0.042)
Unemployed or OLF	0.032	0.029	−0.017	−0.030	−0.013	−0.026
	(0.039)	(0.026)	(0.047)	(0.026)	(0.033)	(0.059)
Observations	9,613	7,290	8,277	6,917	6,050	11,436
Number of females	4,147	3,189	3,541	3,146	2,793	4,910
Number of clusters	8,080	6,423	7,195	6,062	5,369	9,697
Mean outcome	0.218	0.026	0.107	0.165	0.060	0.368
Mean wage worker	0.067	0.063	0.066	0.102	0.096	0.094
Mean non-wage worker	0.028	0.029	0.032	0.037	0.038	0.060
Mean unemp. or OLF	0.012	0.012	0.012	0.014	0.017	0.016
Controls	✓	✓	✓	✓	✓	✓
Individual FE	✓	✓	✓	✓	✓	✓
District × Year FE	✓	✓	✓	✓	✓	✓

Notes: This table reports the results of the impact of return migrant's employment status on female employment status in urban and rural areas. The dependent variable in Columns (1) and (4) is the wage work dummy, in Column (2) and (5) is the self-employed dummy, and in Columns (3) and (6) is the unpaid family work dummy. The independent variable is a categorical variable equal to 1 if the female lives in a household with a return migrant who is wage worker, equal to 2 if she lives in a household with a return migrant who is a non-wage worker, equal to 3 if she lives in a household with a return migrant who is unemployed or out of the labor force, and 0 if there is no migration experience in the household. Controls include age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. Standard errors clustered at the household-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

Table C.27
Savings use.

Savings invested	Savings consumed
Bought shares	Built or acquired housing
Deposited in banks	Spent on marriage
Bought agriculture land	Bought appliances
Bought non-agriculture land	Expenses at home
Established economic project	Family Expenses
Deposited in financial company	Debt repayment
Saving	Pay a loan
Bought gold	

Note: This table reports the categorization of savings into savings invested and savings consumed.

Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

Table C.28
Return migrant's savings accumulation abroad and female employment status.

	Urban			Rural		
	(1) Wage work	(2) Self-emp.	(3) Unpaid work	(4) Wage work	(5) Self-emp.	(6) Unpaid work
Return migrant sample						
No savings	0.038	0.015	0.009	−0.057***	−0.032	0.011
	(0.040)	(0.027)	(0.045)	(0.020)	(0.021)	(0.040)
Savings invested	−0.078**	0.017	−0.021	−0.021	−0.029	0.077
	(0.035)	(0.018)	(0.047)	(0.030)	(0.031)	(0.047)
Savings consumed	0.030	0.015	−0.071	−0.010	0.040	0.044
	(0.045)	(0.018)	(0.057)	(0.024)	(0.042)	(0.053)
Observations	9,670	7,332	8,330	7,000	6,138	11,550
Number of females	4,167	3,206	3,562	3,182	2,834	4,956
Number of clusters	8,126	6,457	7,238	6,137	5,451	9,794
Mean outcome	0.218	0.026	0.107	0.165	0.060	0.367
Mean no saving	0.043	0.044	0.047	0.081	0.083	0.086
Mean savings invested	0.048	0.044	0.046	0.043	0.043	0.048
Mean savings consumed	0.021	0.022	0.022	0.037	0.037	0.043
Controls	✓	✓	✓	✓	✓	✓
Individual FE	✓	✓	✓	✓	✓	✓
District × Year FE	✓	✓	✓	✓	✓	✓

Notes: This table reports the results of the impact of return migrant's savings accumulated abroad on female employment status in urban and rural areas. The dependent variable in Columns (1) and (4) is the wage work dummy, in Columns (2) and (5) is the self-employed dummy, and in Columns (3) and (6) is the unpaid family work dummy. The independent variable is a categorical variable equal to 1 if the female lives in a household with a return migrant who did not have any savings, equal to 2 if she lives in a household with a return migrant who invested his savings, equal to 3 if she lives in a household with a return migrant who consumed his savings, and 0 if there is no migration experience in the household. Controls include age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. Standard errors clustered at the household-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

Table C.29
Gender norms proxy variables.

Variable	Categories
Freedom of mobility	
You can go to the market without permission	Agree/disagree
You can go to the doctor for treatment without permission	Agree/disagree
You can visit the home of relatives, friend, or neighbors without permission	Agree/disagree
You can take the children to health center or doctor without permission	Agree/disagree
Decision making power	
You usually have the final say in purchasing major household items	Agree/disagree
You usually have the final say in making household purchases for daily needs	Agree/disagree
You usually have the final say in visiting family, friends, or relatives	Agree/disagree
You usually have the final say in choosing what to cook	Agree/disagree
You usually have the final say in going to the doctor for treatment	Agree/disagree
You usually have the final say in buying personal clothes	Agree/disagree
You usually have the final say in taking children to the doctor	Agree/disagree
You usually have the final say in sending children to school on a daily basis	Agree/disagree
You usually have the final say in buying clothes for children	Agree/disagree

Notes: This table reports the women's status variables and their response categories.
Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

Table C.30
Return migrant's destination country and female gender norms.

	Urban			Rural		
	(1) Mean	(2) MCA	(3) PCA	(4) Mean	(5) MCA	(6) PCA
Return migrant sample						
More conservative	−0.020	−0.074	−0.333	−0.052***	−0.205***	−0.393***
	(0.027)	(0.101)	(0.252)	(0.018)	(0.066)	(0.141)
Less conservative	−0.148**	−0.553**	−0.826	0.067	0.270	0.946**
	(0.073)	(0.271)	(0.620)	(0.047)	(0.174)	(0.448)
Observations	10,559	10,559	8,493	12,672	12,672	10,332
Number of females	4,425	4,425	3,660	5,368	5,368	4,482
Number of clusters	8,863	8,863	7,526	10,829	10,829	9,251
Mean outcome	0.248	-0.104	0.030	0.230	-0.175	-0.118
Mean more conservative	0.093	0.093	0.095	0.124	0.124	0.122
Mean less conservative	0.009	0.009	0.010	0.022	0.022	0.023
Controls	✓	✓	✓	✓	✓	✓
Individual FE	✓	✓	✓	✓	✓	✓
District × Year FE	✓	✓	✓	✓	✓	✓

Notes: This table reports the impact of return migrant's destination country level of conservatism compared to Egypt on female gender norms in urban and rural areas. The dependent variable in Columns (1) and (6) is the mean of gender norms variables, in Columns (2) and (5) is the MCA, and in Columns (3) and (6) is the PCA. The independent variable is a categorical variable equal to 1 if the female lives in a household with a return migrant from a more conservative country than Egypt, equal to 2 if she lives in a household with a return migrant from a country less conservative than Egypt, and 0 if there is no migration experience in the household. Controls include age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. Standard errors clustered at the household-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

Table C.31
Return migrant's destination country and female employment status.

	Urban			Rural		
	(1) Wage work	(2) Self-emp.	(3) Unpaid work	(4) Wage work	(5) Self-emp.	(6) Unpaid work
Return migrant sample						
More conservative	−0.023	0.022	0.026	−0.051**	0.005	0.078**
	(0.028)	(0.020)	(0.034)	(0.021)	(0.026)	(0.037)
Less conservative	−0.131	0.044	0.018	0.035	−0.014	−0.059
	(0.093)	(0.054)	(0.118)	(0.057)	(0.051)	(0.094)
Observations	9,454	7,174	8,135	6,690	5,864	10,985
Number of females	4,073	3,136	3,479	3,038	2,704	4,715
Number of clusters	7,948	6,317	7,064	5,876	5,214	9,321
Mean outcome	0.218	0.026	0.106	0.166	0.060	0.365
Mean more conservative	0.088	0.086	0.089	0.114	0.113	0.120
Mean less conservative	0.006	0.006	0.007	0.017	0.018	0.021
Controls	✓	✓	✓	✓	✓	✓
Individual FE	✓	✓	✓	✓	✓	✓
District × Year FE	✓	✓	✓	✓	✓	✓

Notes: This table reports the impact of return migrant's destination country level of conservatism compared to Egypt on female employment status in urban and rural areas. The dependent variable in Columns (1) and (4) is the wage work dummy, in Columns (2) and (5) is the self-employment dummy, and in Columns (3) and (6) is the unpaid family work dummy. The independent variable is a categorical variable equal to 1 if the female lives in a household with a return migrant from a more conservative country than Egypt, equal to 2 if she lives in a household with a return migrant from a country less conservative than Egypt, and 0 if there is no migration experience in the household. Controls include age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. Standard errors clustered at the household-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

Table C.32
Return migrant's destination country, migration duration, and female employment status in rural areas.

	Urban			Rural		
	(1) Wage work	(2) Self-emp.	(3) Unpaid work	(4) Wage	(5) Self-emp.	(6) Unpaid work
Panel (A): Short stay						
More conservative	0.080	0.063	0.087	0.049	0.030	0.108
	(0.074)	(0.047)	(0.106)	(0.058)	(0.075)	(0.109)
Less conservative	0.074	0.092	0.128	0.116	0.070	−0.004
	(0.081)	(0.063)	(0.134)	(0.079)	(0.086)	(0.150)
Observations	9,150	6,962	7,899	6,409	5,599	10,503
Number of females	3,948	3,047	3,382	2,916	2,587	4,516
Number of clusters	7,723	6,157	6,888	5,638	4,984	8,936
Mean outcome	0.216	0.026	0.106	0.167	0.060	0.365
Mean more conservative	0.062	0.062	0.066	0.087	0.085	0.092
Mean less conservative	0.005	0.005	0.006	0.013	0.012	0.015
Panel (B): Long stay						
More conservative	0.081	−0.142	−0.051	−0.211**	−0.014	0.222
	(0.093)	(0.131)	(0.110)	(0.096)	(0.046)	(0.182)
Less conservative	−0.926***	−1.385***	−1.167***	−0.058	−0.178	−0.250
	(0.115)	(0.311)	(0.225)	(0.115)	(0.213)	(0.292)
Observations	8,692	6,598	7,445	5,843	5,115	9,555
Number of females	3,754	2,888	3,194	2,662	2,360	4,104
Number of clusters	7,332	5,829	6,496	5,140	4,544	8,120
Mean outcome	0.217	0.026	0.102	0.163	0.058	0.361
Mean more conservative	0.024	0.022	0.022	0.027	0.026	0.028
Mean less conservative	0.001	0.001	0.001	0.004	0.004	0.006
Controls	✓	✓	✓	✓	✓	✓
Individual FE	✓	✓	✓	✓	✓	✓
District × Year FE	✓	✓	✓	✓	✓	✓

Notes: This table reports the impact of return migrant's destination country level of conservatism compared to Egypt on female employment status in rural areas by migration duration. The dependent variable in Columns (1) and (4) is the wage work dummy, in Columns (2) and (5) is the self-employment dummy, and in ColumnS (3) and (6) is the unpaid family work dummy. The independent variable is a categorical variable equal to 1 if the female lives in a household with a return migrant from a more conservative country than Egypt, equal to 2 if she lives in a household with a return migrant from a country less conservative than Egypt, and 0 if there is no migration experience in the household. Panel (A) reports the results for return migrants with a short stay abroad, while Panel (B) reports the results for return migrants with a long stay abroad. Controls include age categories, married dummy, student dummy, years of schooling, years of schooling squared, work experience dummy, household wealth score, household non-labor income, household size, the number of children (0-2), (3-5), and (6-12) years old in the household, and the number of elderly individuals over 65 years old. Standard errors clustered at the household-year level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's elaboration on ELMPS data (2006, 2012, 2018).

Résumé de la thèse

À une époque caractérisée par une mondialisation rapide et le progrès technologique, l'impact d'Internet et des migrations internationales se distingue par son rôle essentiel dans la formation des perceptions, des interactions sociales et des résultats économiques des individus. L'Internet, en tant que centre d'information sans frontières, sert de passerelle virtuelle dépassant les frontières physiques. Il facilite l'échange d'idées et la communication, ce qui fait du monde un village planétaire étroitement interconnecté. Parallèlement, les migrations internationales impliquent le déplacement de personnes au- delà des frontières géographiques, introduisant une diversité qui a un impact à la fois sur les communautés d'accueil et d'origine. Les chapitres de cette thèse visent à évaluer les conséquences politiques, sociales et économiques distinctes de l'utilisation d'Internet et des migrations internationales. Sa spécificité réside dans son approche multidimensionnelle, couvrant divers contextes géographiques, dans l'étude des divers impacts de ces deux forces de mondialisation distinctes mais interconnectées qui ont attiré une attention accrue ces dernières années.

Le premier chapitre se concentre sur les répercussions politiques de l'utilisation d'Internet en tant que source d'information alternative par rapport aux médias traditionnels, un domaine considérablement influencé par la diffusion d'Internet, et en particulier des médias sociaux. Ce chapitre se concentre sur le contexte africain, où les griefs politiques de ces dernières années auraient été amplifiés par l'utilisation d'Internet, et évalue l'impact de l'utilisation d'Internet sur la perception de la démocratie par les individus. Dans le deuxième chapitre, l'accent est mis sur l'Europe, qui a connu une augmentation sans précédent des migrations avec la prolifération des conflits et des problèmes économiques instables, en particulier en Afrique et au Moyen-Orient. Ce chapitre met l'accent sur les implications sociales de l'immigration en termes de diversité culturelle, une préoccupation récurrente qui fait la une des journaux et des débats politiques dans plusieurs pays européens. Enfin, le dernier chapitre, qui s'inscrit dans le contexte de l'Égypte, examine les impacts économiques de l'émigration sur les communautés d'origine. Il se concentre plus particulièrement sur l'émigration masculine, une facette importante de la migration en Égypte, et sur ses effets sur l'offre d'emplois des femmes restée au pays.

Les trois chapitres mènent des investigations empiriques minutieuses, en s'appuyant principalement sur de riches données d'enquête et en employant des stratégies d'identification approfondies pour répondre aux questions de recherche susmentionnées. Les principales

sources de données utilisées comprennent des enquêtes sur les attitudes du public, telles que l'enquête Afrobaromètre employée dans le chapitre 1 et l'European Social Survey (ESS) employée dans le chapitre 2, ainsi que des enquêtes sur la main-d'œuvre telles que l'European Labor Force Survey (EULFS) et l'Egypt Labor Market Panel Survey (ELMPS), respectivement dans le chapitre 2 et le chapitre 3. Les défis d'identification sont principalement le biais de la variable omise, la causalité inverse et le biais de sélection, qui sont traités en profondeur en utilisant les progrès récents de la littérature sur les variables instrumentales (IV) en s'appuyant sur une stratégie IV de shift-share dans le chapitre 1 et le chapitre 2. Cependant, dans le chapitre 3, la dimension panel des données est exploitée pour atténuer le biais d'endogénéité.

Cette introduction générale vise à donner un aperçu de l'évolution de l'utilisation de l'Internet et de la migration au cours des dernières années. En outre, elle offre un bref aperçu, non exhaustif, de la littérature économique qui sous-tend les questions étudiées dans cette thèse. Enfin, elle vise à mettre en évidence les contributions de chaque chapitre par rapport à la littérature existante. L'introduction est organisée comme suit. Dans la section I, je présente l'évolution de l'adoption d'Internet, en mettant particulièrement l'accent sur l'Afrique. Je présente également brièvement la littérature sur les effets politiques de l'Internet. Dans la section II, l'accent est mis sur l'exploration des migrations internationales, avec une attention particulière pour l'Europe en tant que société d'accueil. J'aborde également les progrès de la recherche concernant l'impact de l'immigration sur les sociétés d'accueil, en soulignant plus particulièrement les impacts sociaux. Enfin, dans la section III, je présente brièvement le profil de l'émigration égyptienne. En outre, j'aborde la littérature sur la migration masculine et ses répercussions sur les personnes restées au pays. À la fin de chaque section, je souligne la contribution des chapitres de cette thèse.

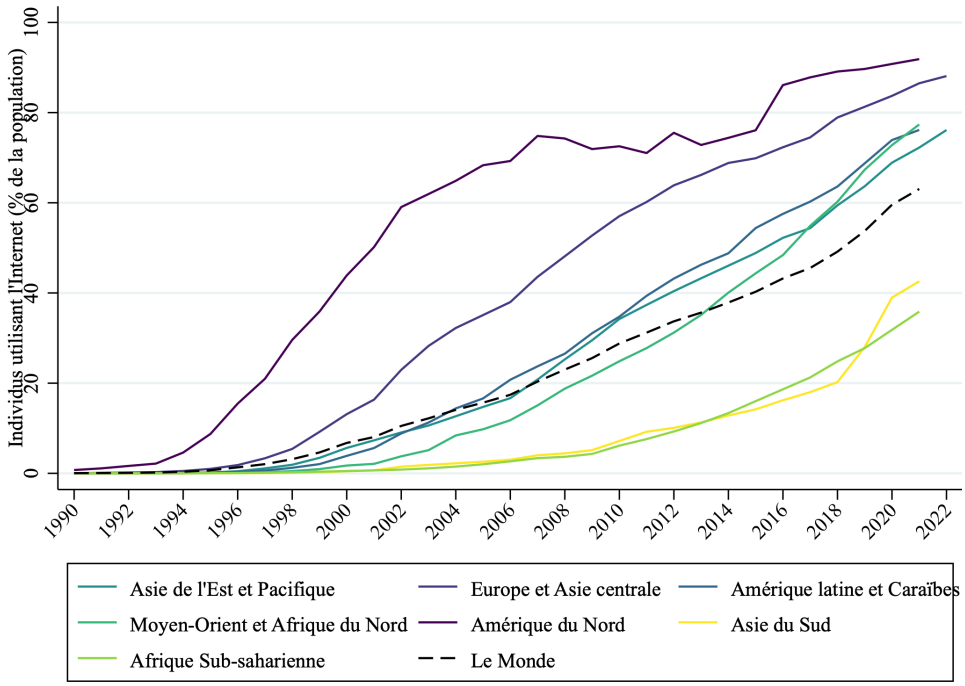
I Utilisation d'Internet et implications politiques

L'évolution de l'utilisation d'Internet

Depuis son origine en tant que système de communication décentralisé pendant la guerre froide, l'Internet est devenu l'une des principales technologies de l'information et de la communication dans le monde. Aujourd'hui, près de 5,35 milliards de personnes utilisent l'Internet dans le monde, ce qui représente 66.2% de la population mondiale en janvier 2024. La pénétration de l'Internet n'a pas évolué au même rythme dans toutes les régions du monde. Selon la Fig. 1 ci-dessous, l'Internet a commencé à se développer en Amérique

du Nord, en Europe et en Asie centrale au début des années 90, en Asie de l'Est et dans le Pacifique, en Amérique latine et dans les Caraïbes, au Moyen- Orient et en Afrique du Nord au début des années 2000, et ce n'est qu'à la fin des années 2000 qu'il a commencé à prospérer en Afrique Sub-saharienne et en Asie du Sud. Bien que l'Afrique soit un retardataire en matière de numérisation, l'utilisation d'Internet y a évolué rapidement, passant de 6% en 2010 à 35% en 2022.

Fig. 1. L'évolution de l'utilisation de l'Internet au fil du temps.

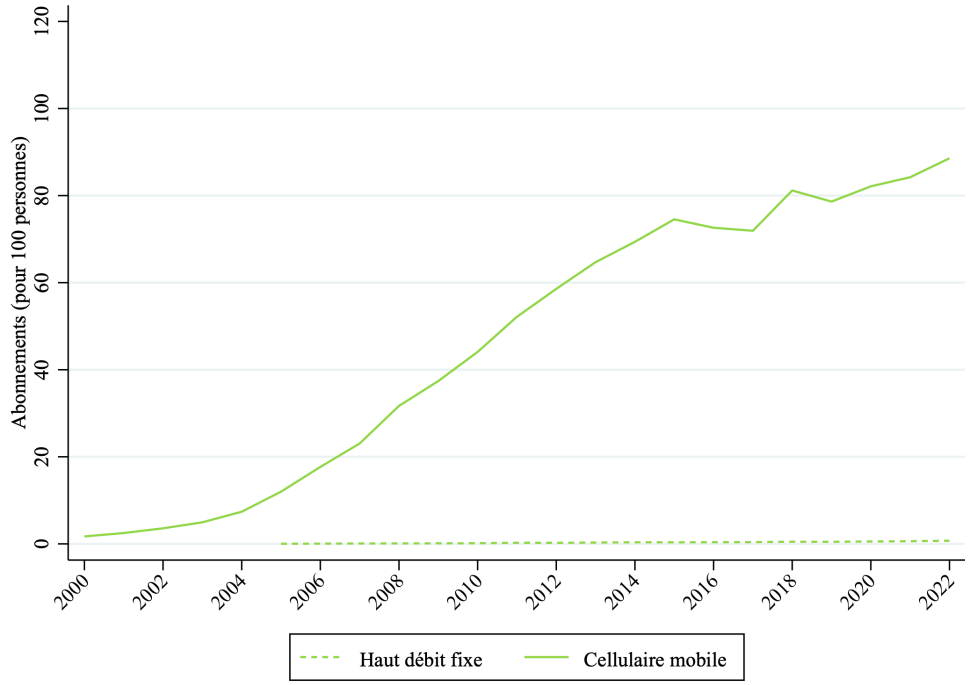


Note: Cette figure montre le pourcentage d'utilisateurs d'Internet sur la période 1990-2022 par région.
Source: Élaboration de l'auteur sur la base des données des Indicateurs du développement dans le monde (1990-2022).

L'augmentation de la pénétration de l'Internet en Afrique Sub-saharienne résulte de l'amélioration de l'infrastructure des télécommunications et de l'Internet avec l'augmentation du déploiement des câbles sous-marins à partir de l'année 2010. Il y a également eu des investissements dans les points d'échange Internet pour maintenir le trafic Internet au niveau local. En outre, des systèmes satellitaires ont été adoptés pour améliorer la connexion des zones reculées. Alors que les abonnements au haut débit fixe sont encore très limités, la population africaine s'appuie principalement sur les abonnements cellulaire mobile comme infrastructure d'Internet. Comme le montre la Fig. 2, le nombre de personnes disposant d'un réseau cellulaire mobile est passé de 2 en 2000 à 88,5 (pour 100 habitants) en 2022 en Afrique Sub-saharienne, avec seulement moins de 1 pour 100

habitants disposant d'un abonnement au haut débit fixe sur l'ensemble de la période.

Fig. 2. Evolution dans le temps des abonnements au haut débit fixe et au cellulaire mobile en Afrique Sub-saharienne.



Note: Cette figure montre les abonnements au haut débit fixe et au cellulaire mobile (pour 100 personnes) au cours de la période 2000-2022.
Source: Élaboration de l'auteur sur la base des données des Indicateurs du développement dans le monde (2000-2022).

La part du continent couverte par la connection mobile 2G ou 3G a également augmenté de manière substantielle au cours de la dernière décennie. Malgré les progrès récents, l'infrastructure de l'Internet et des télécommunications est encore sous-développée et le taux de pénétration de l'Internet en Afrique reste bien inférieur à la moyenne mondiale.

Les effets politiques de l'Internet

Avec l'expansion d'Internet et l'utilisation des médias sociaux, de nombreux chercheurs ont qualifié Internet de "technologie de libération" (Diamond, 2010). En raison de son faible coût, de son accès libre et de sa nature décentralisée, l'Internet peut être une source plus libre d'informations critiques à l'égard du gouvernement, ce qui rend les acteurs politiques plus difficiles à dissimuler des informations nuisibles et les rend plus vulnérables et responsables (Zhuravskaya et al., 2020). Il en résulte des citoyens mieux informés, mécontents de leur gouvernement et enclins à participer à des manifestations politiques.

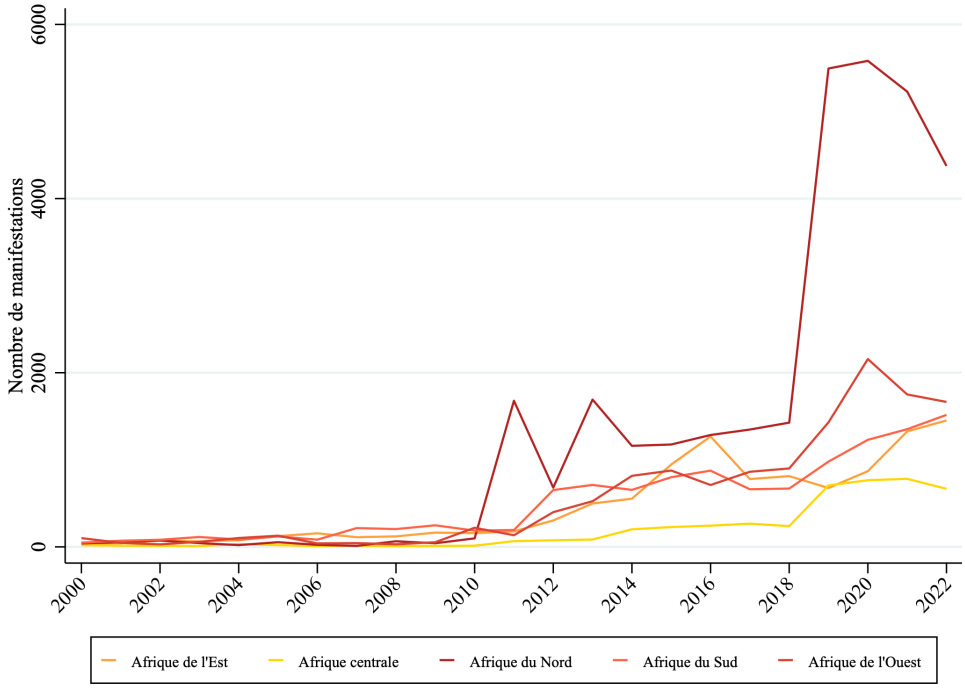
Par rapport aux médias traditionnels, Internet se caractérise par de faibles barrières à l'entrée et un contenu généré par les utilisateurs, ce qui facilite la diffusion d'informations politiques et l'accès aux nouvelles et événements mondiaux, en particulier dans les pays où la liberté d'expression est limitée. Néanmoins, bien qu'il soit une source d'information ouverte potentielle, l'Internet est également susceptible d'être une "technologie de mésinformation". Les médias en ligne donnent à leurs utilisateurs la possibilité de réafficher et de partager des contenus créés par d'autres sans aucune procédure de vérification des faits, ce qui nuit à la qualité des informations diffusées en ligne. En effet, certaines études font état de la propagation de fausses nouvelles en ligne (Mocanu et al., 2015; Allcott and Gentzkow, 2017; Grinberg et al., 2019) et du fait qu'elles se propagent plus rapidement et plus largement que les vraies nouvelles (Vosoughi et al., 2018). En outre, Internet peut également être utilisé par des régimes autocratiques pour manipuler l'information et créer de la propagande (King et al., 2017).

Plusieurs travaux étudient l'impact de l'expansion d'Internet et des médias sociaux sur les manifestations de rue, suggérant qu'ils facilitent l'organisation et la coordination des mobilisations de masse grâce au flux horizontal d'informations qui exacerbent l'éclosion de mouvements de protestation. Par exemple, Fergusson and Molina (2019) montrent que les nouvelles versions de Facebook dans une langue spécifique sont associées à une augmentation des protestations dans les pays où cette langue est parlée. Cela s'est avéré particulièrement vrai en Afrique puisqu'elle a été témoin d'une montée du mécontentement public dépeint dans un pic du nombre de mouvements de protestation sur le continent au cours de la dernière décennie. La Fig. 3 ci-dessous montre l'augmentation du nombre de manifestations à partir de 2010. Le plus notable est le printemps arabe qui s'est déroulé entre 2010 et 2012 en Afrique du Nord et où l'Internet et les médias sociaux ont joué un rôle organisationnel majeur.

Dans cette veine, Acemoglu et al. (2018) documente une association positive entre l'activité de Twitter et le nombre de manifestants dans le contexte égyptien, pendant le printemps arabe, reflétant le rôle des médias sociaux dans la coordination des manifestations de rue. Steinert-Threlkeld et al. (2015) obtiennent des résultats similaires en utilisant des tweets géolocalisés et des données sur les manifestations dans 16 pays pendant le printemps arabe. Manacorda and Tesei (2020) constatent également que l'expansion du réseau mobile 2G en Afrique augmente les manifestations antigouvernementales pendant les périodes de ralentissement économique. Ils soulignent le rôle joué par Internet en tant que canal d'information et de coordination. Plus récemment, Guiffard (2022) étudie l'arrivée échelonnée des câbles sous-marins en Afrique et son impact sur la mobilisation politique. Il

documente un impact positif canalisé par le canal de la coordination améliorée. Le même schéma est également observé dans d'autres parties du monde, notamment en Chine (Qin et al., 2021), en Russie (Enikolopov et al., 2020) et aux États-Unis (Amorim et al., 2022).

Fig. 3. Nombre de manifestations au fil du temps.



Note: Cette figure indique le nombre de manifestations sur la période 2000-2022 en Afrique.
Source: Élaboration de l'auteur sur la base des données de l'ACLED (2000-2022).

Outre son impact sur la montée des mouvements de protestation, il existe une vaste littérature étudiant l'impact d'Internet sur d'autres aspects politiques, tant dans les démocraties que dans les autocraties. Il s'agit notamment d'un large éventail d'études portant sur l'impact de l'Internet sur le comportement électoral. Dans les démocraties établies, les auteurs documentent généralement une baisse de la participation électorale avec la pénétration d'Internet en raison du remplacement des informations politiques par des contenus divertissants (Falck et al., 2014; Gavazza et al., 2019; Campante et al., 2018). Cependant, Campante et al. (2018) constatent que cette tendance s'est inversée en Italie après 2008, vraisemblablement avec l'introduction des réseaux sociaux. Certains auteurs estiment également qu'Internet est associé à la montée du populisme en Europe (Schaub and Morisi, 2020; Guriev et al., 2021; Tabellini et al., 2023). Néanmoins, dans les autocraties, Internet a surtout été associé à une diminution du soutien aux gouvernements en place (Miner, 2015; Donati, 2023; Guriev et al., 2021).

En outre, certains chercheurs étudient la relation entre Internet et la démocratie. Jusqu'à présent, la plupart des recherches quantitatives sur le sujet se sont concentrées sur la macro-relation entre la pénétration de l'Internet et le niveau de démocratie, documentant une association positive, mais avec peu d'idées sur les mécanismes sous- jacents (Evans, 2019; Jha and Kodila-Tedika, 2020). Au cours des deux dernières décennies, les spécialistes de la démocratie ont commencé à mettre l'accent sur la manière dont les attitudes des citoyens à l'égard de la démocratie et l'accès aux technologies de l'information façonnent le processus de démocratisation. Pour que les démocraties émergent et survivent, deux conditions doivent être remplies. Premièrement, les citoyens doivent choisir et soutenir la démocratie comme leur régime de gouvernance préféré (Claassen, 2020). En effet, une demande accrue de démocratie favorise la démocratisation par le biais d'un processus "ascendant" dans lequel les citoyens exercent une pression sur les régimes autoritaires (Lei, 2013). Deuxièmement, les citoyens doivent également croire qu'ils obtiennent la démocratie (Mattes and Bratton, 2007). Les études évaluant l'utilisation d'Internet par les individus et leurs attitudes à l'égard de la démocratie sont limitées et sont en grande partie des travaux conceptuels documentant une simple corrélation. Les dernières constatent qu'Internet en général et les médias sociaux en particulier sont associés à une moindre satisfaction à l'égard de la démocratie (Bailard, 2012; Ceron and Memoli, 2016; Chang, 2018).

Contribution du Chapitre 1

La contribution du premier chapitre de cette thèse est de fournir une approche micro-économétrique allant de l'utilisation d'Internet par les individus aux attitudes des citoyens envers la démocratie. Acemoglu et al. (2021) soulignent le besoin crucial d'explorer le rôle des médias dans la formation du soutien à la démocratie, en particulier à la lumière de la diffusion de fausses informations par divers médias et plateformes de médias sociaux, et la manière dont cela peut affecter la relation entre les performances démocratiques réussies et le soutien du public à la démocratie. À notre connaissance, il s'agit de la première étude à mener une enquête empirique minutieuse sur le lien de causalité entre Internet en tant que source d'information et les attitudes et perceptions des citoyens à l'égard de la démocratie en Afrique. Plus précisément, nous étudions son impact sur leur préférence pour la démocratie, leur perception de l'étendue de la démocratie dans leur pays et leur satisfaction quant au fonctionnement de la démocratie. Nous fournissons également une analyse approfondie des mécanismes sous-jacents qui exploitent le rôle joué par Internet en tant que canal d'information ou de mésinformation. Nous étudions

son impact sur la confiance dans les gouvernements, la perception de la corruption, la participation politique, ainsi que la mésinformation. Ce chapitre contribue donc au débat sur le statut d'Internet en tant que technologie de "libération" ou de "mésinformation".

Nous nous appuyons sur l'Afrobaromètre, qui est une enquête sur les attitudes du public comportant un large éventail de questions sur la démocratie, la gouvernance et d'autres sujets connexes en Afrique. Les données de l'Afrobaromètre sont collectées par le biais d'entretiens en face à face avec un échantillon aléatoire de 1,200 ou 2,400 personnes dans chaque pays. Nous utilisons les cycles 5, 6 et 7 de cette enquête, car les cycles précédents ne comportaient pas de questions sur l'utilisation d'Internet. Au final, nous disposons d'un échantillon transversal répété de 99,938 citoyens africains dans 35 pays, couvrant la période entre 2011 et 2018.

Notre principale préoccupation est le biais d'endogénéité qui découle des variables omises et de la relation bidirectionnelle entre la consommation de nouvelles sur Internet et les variables de la démocratie. Pour y remédier, nous adoptons un modèle de variable instrumentale composite ou interactif, à l'instar de Borusyak et al. (2022), où les instruments sont construits comme des chocs globaux pondérés par des facteurs d'exposition de niveau inférieur. Dans ce cadre, la validité de notre instrument découle de la variation exogène des chocs tout en permettant une variation endogène du facteur d'exposition. Ceci est également lié à la littérature économique sur les instruments de shift-share (Borusyak et al., 2022; Goldsmith-Pinkham et al., 2020). Plus précisément, notre instrument consiste à combiner le nombre de câbles sous-marins dans un pays (qui est considéré comme le choc exogène) avec la part du district couverte par le réseau 3G (qui est un facteur de pondération potentiellement endogène). Afin de s'assurer que notre identification résulte du choc exogène, nous fixons la couverture 3G du district à son taux de couverture au moment de la première vague d'enquête Afrobaromètre utilisée dans l'analyse.

Les résultats indiquent que l'utilisation d'Internet pour obtenir des informations a un effet négatif significatif à la fois sur la préférence et la perception de l'étendue de la démocratie. Cet effet négatif est dû à plusieurs facteurs. Premièrement, l'utilisation d'Internet érode la confiance dans les institutions gouvernementales, principalement dans le parlement et le parti au pouvoir. Elle augmente la perception que les membres du parlement sont impliqués dans la corruption. En outre, l'érosion de la confiance est corrélée à une plus grande mobilisation politique, sous la forme d'une plus grande participation aux démonstrations et au vote. Ces résultats font écho à la littérature existante et, en particulier, soulignent les risques d'inversion des processus de démocratisation naissants. Enfin, Internet semble agir comme un canal de mésinformation. D'une part, la perception qu'ont les internautes

de l'étendue de la démocratie et de la corruption des législateurs diverge des évaluations des experts. D'autre part, l'utilisation d'Internet augmente la probabilité d'incohérence dans les positions des personnes interrogées sur leur préférence pour la démocratie. Internet n'est pas un canal d'information neutre : il tend à saper la préférence des citoyens pour la démocratie tout en modifiant les perceptions des institutions politiques.

II Immigration et culture

L'évolution des taux d'immigration

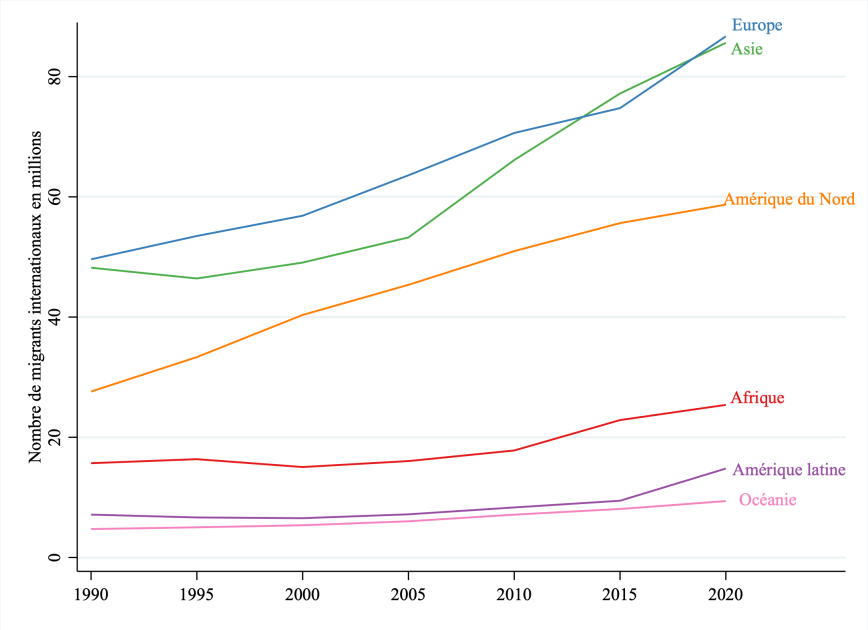
Avec la mondialisation, non seulement la population des utilisateurs d'Internet s'est considérablement accrue au cours des dernières décennies, mais les migrations internationales ont également connu une hausse notable. Selon les dernières estimations publiées par la Division de la population des Nations unies en 2020, le nombre actuel de migrants internationaux s'élève à 281 millions, ce qui représente environ 3.6% de la population mondiale. Cette poussée migratoire est évidente dans toutes les régions du monde, avec des augmentations particulièrement prononcées observées en Asie et en Europe. Comme le montre la Fig. 4 ci-dessous, l'Asie et l'Europe accueillent aujourd'hui respectivement 86 millions et 87 millions de migrants internationaux, représentant collectivement plus de 50% du stock mondial de migrants internationaux en 2020.

Au cours de la période allant de 2000 à 2020, c'est l'Asie qui a connu la croissance la plus importante, avec une augmentation remarquable de 74%, ce qui équivaut à environ 37 millions d'immigrants. L'Europe suit de près, avec 30 millions de migrants internationaux supplémentaires. Toutefois, si l'on examine la proportion de migrants par rapport à l'ensemble de la population, comme le montre la Fig. 5, des tendances distinctes se dessinent.

La part des migrants internationaux n'a cessé d'augmenter dans les pays développés au cours des dernières années. Le nombre de migrants internationaux a augmenté pendant toute la période, contrairement aux pays en développement où il a d'abord diminué avant d'augmenter légèrement. Notamment, l'Océanie, l'Amérique du Nord et l'Europe affichent les plus fortes proportions de migrants internationaux, soit 22%, 16% et 12% de leurs populations respectives. À l'inverse, l'Amérique latine, l'Afrique et l'Asie affichent des parts comparativement plus faibles, représentant chacune moins de 3% de leur population.

Si l'on examine l'Europe en particulier, il est évident que le continent accueille le plus grand nombre de migrants internationaux, la proportion de migrants par rapport à la

Fig. 4. Nombre de migrants internationaux sur la période 1990-2020.



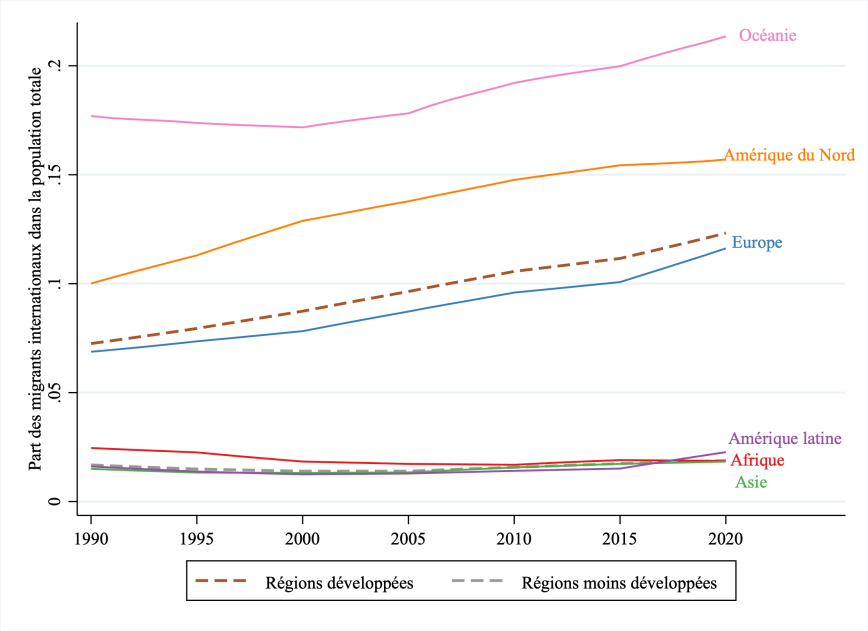
Note: Cette figure montre l'évolution du nombre de migrants internationaux par région entre 1990 et 2020.
Source: Calculs de l'auteur basés sur les données de la Division de la population des Nations Unies (1990-2020).

population totale augmentant régulièrement, passant de 7% en 1900 à 12% en 2020. Cette augmentation peut être attribuée à divers facteurs, notamment la guerre des Balkans en 1990, l'élargissement important de l'UE en 2004 et la crise des réfugiés de 2015, qui a vu des personnes fuir la pauvreté et les conflits au Moyen-Orient et en Afrique.

L'analyse de la part des migrants par groupe d'origine, comme le montre la Fig. 6 (a), révèle une augmentation de la part des migrants originaires des régions développées et moins développées. Ces dernières représentent notamment 55% de l'augmentation du nombre de migrants internationaux en Europe entre 1990 et 2020. En ce qui concerne la répartition des immigrants en Europe par région d'origine, la Fig. 6 (b) montre une diminution de la part des migrants internationaux originaires d'Europe et d'Amérique du Nord, accompagnée d'une augmentation des migrants originaires d'autres régions.

En résumé, l'Europe en particulier, et le reste du monde développé en général, ont connu une augmentation significative de leur population migrante, avec des individus provenant désormais d'horizons économiques, géographiques et culturels divers, contribuant à une diversité accrue au sein des communautés d'accueil. Ce phénomène a incité les chercheurs à étudier les implications économiques, politiques et sociales de la migration sur les sociétés d'accueil.

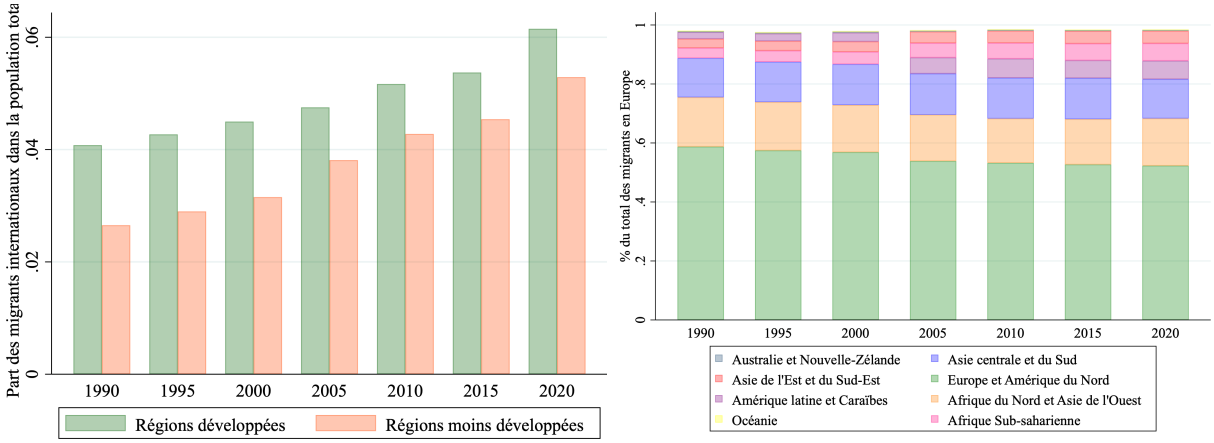
Fig. 5. Part des migrants internationaux sur la période 1990-2020.



Note: Cette figure montre l'évolution de la part des migrants internationaux par région entre 1990 et 2020.
Source: Calculs de l'auteur basés sur les données de la Division de la population des Nations Unies (1990-2020).

Fig. 6. Migrants internationaux en Europe par groupe d'origine sur la période 1990-2020.

(a) Proportion **(b)** Distribution



Notes: Cette figure montre l'évolution de la part des migrants internationaux des régions développées et moins développées en Europe entre 1990 et 2020 en (a) et leur répartition par région d'origine au cours de la même période en (b). Les régions développées et moins développées sont basées sur la définition des groupes de développement des Nations Unies.
Source: Calculs de l'auteur basés sur les données de la Division de la population des Nations Unies (1990-2020).

L'impact de l'immigration sur les sociétés d'accueil

L'impact des immigrants sur le marché du travail a fait l'objet d'un nombre croissant d'études, les natifs craignant que les immigrants ne réduisent leurs salaires et ne limitent leurs possibilités d'emploi. La plupart des études se concentrent sur l'impact sur les salaires des natifs, sans grand consensus. Alors que certaines études documentent un effet négatif (Borjas, 2003; Ortega and Verdugo; Altonji and Card, 2018), d'autres trouvent un effet moyen nul ou positif de l'immigration sur les salaires des natifs, en particulier à long terme (Ottaviano and Peri, 2012; Manacorda et al., 2012; Borjas, 2014; Edo and Toubal, 2015). La plupart des études soulignent l'importance de la composition des compétences de la main-d'œuvre immigrée et de son degré de substituabilité avec les natifs. En outre, les études sur les expériences naturelles montrent que l'effet initial et à court terme de l'immigration peut être négatif en fonction de la rapidité des ajustements du marché du travail (Hunt, 1992; Borjas, 2017; Borjas and Monras, 2017; Monras, 2020; Prantl and Spitz-Oener, 2020). En résumé, l'effet moyen de l'immigration sur les salaires des natifs, qu'il soit positif ou négatif, s'avère négligeable.

Outre les préoccupations économiques des natifs, l'immigration a également suscité des préoccupations culturelles. De nombreuses études montrent que le scepticisme à l'égard de l'immigration est principalement motivé par des considérations culturelles plutôt que par des considérations liées au marché du travail (Dustmann and Preston, 2007a; Card et al., 2012; Poutvaara and Steinhardt, 2018). Cette question a récemment fait l'objet d'un débat animé en Europe avec la récente crise des réfugiés en 2015 et l'afflux croissant de migrants venant de pays géographiquement, culturellement et économiquement éloignés. On craint de plus en plus que l'immigration ne constitue une menace pour la culture occidentale, les migrants apportant de nouvelles valeurs et normes qui affectent la composition culturelle des pays d'accueil. Cette dernière observation a été mise en évidence par les mouvements de droite, ce qui a renforcé le scepticisme des natifs à l'égard des immigrés et des réfugiés sur la base de considérations culturelles. Plusieurs études font état d'un lien positif entre la part des immigrés ou des réfugiés et le vote pour les partis anti-immigration et de droite en Europe (Otto and Steinhardt, 2014; Halla et al., 2017; Becker et al., 2016; Edo et al., 2019a). En revanche, certains constatent également des effets hétérogènes en fonction de la zone de résidence (c'est-à-dire urbaine ou rurale) (Dustmann et al., 2019), du type d'exposition (Steinmayr, 2017) et des conditions économiques (Tomberg et al., 2021).

C'est pourquoi l'impact de l'immigration sur la culture des sociétés d'accueil a fait l'objet d'une attention accrue au cours des dernières décennies. La littérature sur ce sujet a

identifié deux canaux principaux par lesquels l'immigration peut affecter la culture. Tout d'abord, il existe un effet direct, à savoir le canal de l'effet de composition, où le simple fait d'ajouter à une société des individus ayant des valeurs distinctes affecte sa composition. L'ampleur de cet effet dépend de la transmission intergénérationnelle de ces valeurs, c'est-à-dire de leur persistance et de la manière dont elles sont transmises aux générations suivantes (Bisin and Verdier, 2001; Desmet et al., 2017; Galli and Russo, 2019; Desmet and Wacziarg, 2021) et le taux d'assimilation culturelle (Algan et al., 2012; Abramitzky et al., 2014; Giavazzi et al., 2019; Gonnot and lo Polito, 2021; Fouka et al., 2022; Abramitzky and Boustan, 2022; Gonnot and lo Polito, 2023). Deuxièmement, il existe un effet indirect à travers la réaction des natifs. En interagissant avec la population natifs, les immigrants peuvent eux-mêmes influencer la culture des natifs (Fisman and Miguel, 2007; Schmitz and Weinhardt, 2019; Tabellini, 2020; Giuliano and Nunn, 2021; Miho et al., 2023). Dans ce cas, l'effet sur la diversité culturelle globale dépendrait de la manière dont les natifs réagissent. Ils pourraient absorber les nouvelles valeurs, mais aussi avoir une réaction de rejet, surtout si les migrants viennent de pays culturellement éloignés (Dustmann and Preston, 2007b; Edo et al., 2019a; Steinmayr, 2021; Alesina and Tabellini, 2022; Moriconi et al., 2022; Keita et al., 2023).

Rapoport et al. (2020) combinent les mécanismes susmentionnés dans un cadre théorique unifié et testent l'effet de la migration sur la similarité culturelle entre les pays d'origine et de destination dans un contexte mondial. Ils constatent une corrélation positive entre la migration et la similarité culturelle, principalement due aux transferts de fonds culturels, c'est-à-dire au fait que les immigrés diffusent la culture de la société d'accueil dans leurs communautés d'origine. Alors que leur étude se concentre sur la convergence culturelle globale, elle n'aborde pas l'impact de l'immigration sur la diversité culturelle dans les pays d'accueil.

Contribution du chapitre 2

La contribution du chapitre 2 est donc d'étudier l'impact de l'immigration sur l'hétérogénéité culturelle globale dans les pays d'accueil en Europe. Nous contribuons ainsi à deux courants de littérature. Tout d'abord, la littérature sur la manière dont l'immigration affecte la culture dans les pays d'accueil. Notre étude complète celle de (Rapoport et al., 2020) sur l'impact de l'immigration sur la proximité culturelle entre paires de pays au niveau mondial, en se concentrant sur l'impact sur l'hétérogénéité culturelle au sein des pays d'accueil. Elle s'aligne également sur la littérature croissante sur l'évolution des fractures culturelles dans les sociétés occidentales, qui étudie si les clivages identitaires

sont de bons prédicteurs des attitudes et des valeurs des individus (Desmet et al., 2017; Desmet and Wacziarg, 2021). Alors que cette littérature se concentre principalement sur l'âge, l'éducation, l'orientation politique, la race et l'ethnicité en tant que clivages identitaires, nous présentons le lieu de naissance comme un marqueur identitaire pertinent pour expliquer la diversité culturelle.

Pour ce faire, nous comparons l'évolution de la part des migrants dans les régions européennes à l'évolution de l'hétérogénéité culturelle. Pour mesurer cette dernière, nous adoptons l'indice de Desmet and Wacziarg (2021) de fractionnement moyen de la population entre divers traits culturels. Cet indice mesure la probabilité que deux individus choisis au hasard dans l'ensemble de la population résidente d'un pays donné possèdent une variante différente d'un trait choisi au hasard. Nous combinons les données de l'Enquête sociale européenne et de l'Enquête européenne sur les forces de travail pour la période 2004-2018. Pour répondre aux problèmes d'autosélection et d'hétérogénéité non observée, nous nous appuyons sur une stratégie de variable instrumentale shift-share pour prédire les stocks exogènes de migrants en partant de l'hypothèse que les nouveaux flux de migrants d'une origine donnée sont répartis entre les régions sur la base de la distribution initiale des migrants de cette même origine.

Nous constatons que l'immigration affecte l'hétérogénéité culturelle régionale dans les pays d'accueil et qu'une augmentation de l'immigration réduit l'hétérogénéité culturelle. Nous trouvons des preuves d'un effet de composition, les immigrants apportant de nouvelles valeurs culturelles aux sociétés d'accueil, ce qui accroît l'hétérogénéité culturelle. Cependant, la réponse des natifs l'emporte sur l'effet de composition, ce qui entraîne une réduction globale de l'hétérogénéité culturelle. L'effet de composition est principalement le fait des migrants peu qualifiés et des migrants à court terme sur les questions liées à la moralité sexuelle, à la religiosité et au rôle de l'État, les migrants étant plus conservateurs sur ces questions que les natifs. Toutefois, cette tendance disparaît dans les 10 ans suivant leur arrivée, ce qui suggère l'idée d'une assimilation culturelle. En ce qui concerne la réponse des natifs, nous constatons qu'elle est principalement le fait de migrants originaires d'autres pays de l'UE et les migrants hautement qualifiés sur des questions relatives à l'ouverture, au capital culturel et à l'attitude à l'égard de la migration. Nous constatons que les natifs affichent des attitudes plus libérales en général avec plus de migration.

III Migration masculine et femmes restées au pays

Profil migratoire de l'Égypte

Alors que l'Europe constitue une importante plaque tournante de l'immigration, l'Égypte se distingue comme étant le plus grand pays d'origine des migrants dans la région du Moyen-Orient et de l'Afrique du Nord (MENA), fournissant le plus de main-d'œuvre émigrée aux pays producteurs de pétrole du Moyen-Orient (David et al., 2019).[1] Le dernier recensement égyptien de 2017 indique qu'environ 9,4 millions de migrants égyptiens résident en dehors du pays (CAPMAS, 2017). La répartition des émigrants égyptiens dans les différentes régions du monde au fil du temps, illustrée à la Fig. 7(a), révèle une concentration dans les États arabes et les États du Golfe, où plus de 80% des émigrants résident en Afrique du Nord et en Asie de l'Ouest. Les destinations les plus populaires pour les Égyptiens sont ensuite l'Amérique du Nord et l'Europe. Ces tendances sont restées les mêmes au cours des trois dernières décennies. L'analyse des principaux pays de destination des Égyptiens en 2020, présentée à la Fig. 7(b), montre que l'Arabie saoudite et les Émirats arabes unis accueillent plus de 50% de tous les émigrants égyptiens, suivis par le Koweït. En Europe et en Amérique du Nord, les États-Unis, l'Italie et le Canada apparaissent comme les destinations les plus recherchées.

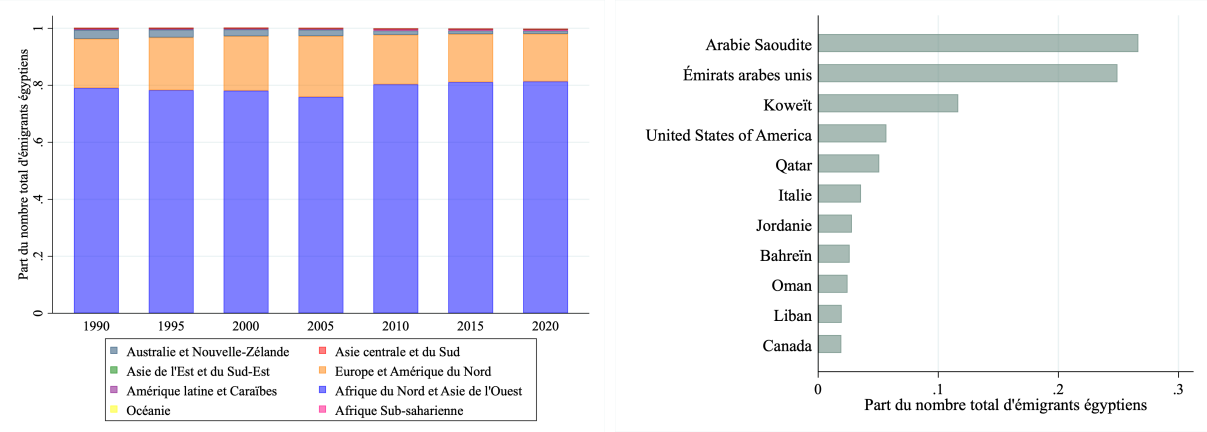
Des études indiquent que les émigrants égyptiens, généralement de jeunes hommes originaires de zones rurales, présentent des statuts éducatifs et professionnels variés en fonction de leur pays de destination. Ceux qui migrent vers les États du Golfe trouvent souvent un emploi dans des domaines techniques spécialisés, qualifiés et semi-qualifiés, tandis que ceux qui se rendent en Jordanie ou en Libye sont plus susceptibles de trouver un emploi peu qualifié dans les secteurs de l'agriculture, de la construction et des services (De Bel-Air, 2016). En outre, les émigrants vers les pays occidentaux affichent des niveaux d'éducation différents, ceux qui se rendent en Europe du Sud ayant généralement un niveau d'éducation inférieur à celui de leurs homologues qui se rendent en Amérique du Nord, lesquels ont tendance à avoir un niveau d'éducation élevé.

Une caractéristique notable de l'émigration égyptienne vers l'Occident est sa nature essentiellement permanente, impliquant souvent des ménages entiers, contrairement à la migration dans la région arabe, qui est généralement dominée par les hommes et principalement temporaire en raison des contrats de travail à court terme et des droits de

[1] Un bref aperçu de l'histoire des migrations égyptiennes est fourni dans le chapitre 3 en section 3.2.

Fig. 7. Destinations des émigrants égyptiens.

(a) Répartition entre les régions. **(b)** Principales destinations en 2020.



Note: Cette figure montre l'évolution de la répartition des émigrants égyptiens entre les régions entre 1990 et 2020 en (a) et les principaux pays de destination en 2020 en (b).
Source: Calculs de l'auteur basés sur les données de la Division de la population des Nations Unies (1990-2020).

citoyenneté limités (Wahba, 2015b; Tsourapas, 2022). Malgré leur nature temporaire, divers facteurs nationaux et internationaux ont contribué à d'importantes vagues de migration de retour en provenance du Golfe, d'Irak et de Libye depuis la fin des années 1980 (Zohry, 2003). Ces facteurs comprennent la guerre Iran- Irak en 1988 et les baisses subséquentes des prix du pétrole, la première guerre du Golfe en 1990-1991, les politiques favorisant la main-d'œuvre nationale au détriment des travailleurs étrangers, la réduction de la demande de main-d'œuvre dans le secteur de la construction dans les pays arabes, et la guerre d'Irak en 2003. En outre, il y a eu un afflux notable de migrants égyptiens de retour de Libye après l'effondrement du régime de Kadhafi pendant le printemps arabe (Zohry, 2013; De Bel-Air, 2016). Selon l'enquête 2018 du panel sur le marché du travail égyptien (ELMPS), environ 7% de la population en âge de travailler étaient des migrants de retour.

L'impact de la migration masculine sur les femmes restées au pays

Si les migrants peuvent avoir une influence considérable sur les pays de destination, ils ont également de profondes répercussions sur les communautés d'origine. L'émigration et le retour des membres masculins du ménage peuvent avoir des effets significatifs sur les familles restées au pays, présentant un mélange de résultats positifs et négatifs. Certaines études se concentrent sur l'influence de l'absence des migrants sur le niveau scolaire et les résultats scolaires des enfants restés au pays (Yang, 2008; Antman, 2012; Alcaraz

et al., 2012). Elles suggèrent que le revenu supplémentaire reçu par les ménages grâce aux envois de fonds diminue le besoin de travail des enfants et améliore l'accès des enfants à l'éducation, ce qui profite particulièrement aux filles dans les pays en développement. Cependant, d'autres recherches indiquent que l'absence de la personne qui s'occupe de l'enfant au premier chef peut augmenter la probabilité que les enfants abandonnent l'école et entraver leur progression dans l'éducation. Les études sur l'impact sur la santé des membres de la famille restés au pays suggèrent que les envois de fonds peuvent améliorer les soins de santé et la nutrition des familles (Gibson et al., 2011; Mu and De Brauw, 2015). La littérature récente sur la dynamique sociale et familiale met en évidence le pouvoir de négociation accru des femmes restées au pays (De Haas and Van Rooij, 2010), ainsi que les variations dans les conditions de vie des enfants (Bertoli et al., 2023). Certaines études suggèrent même que la migration des membres du ménage influence la transmission politique (Spilimbergo, 2009; Chauvet and Mercier, 2014; Barsbai et al., 2017), la fertilité (Beine et al., 2013; Bertoli and Marchetta, 2015) et les normes de genre (Tuccio and Wahba, 2018; Diabate et al., 2019; Samari, 2021) des pays de destination vers les pays d'origine.

En outre, il existe une vaste littérature sur les réponses de l'offre de travail des membres de la famille restés au pays. La plupart des études dans ce domaine se concentrent sur l'offre de travail des épouses lorsque leurs maris émigrent à l'étranger.[2] La majorité de ces études indiquent une diminution de la participation des femmes au marché du travail en général, mais notent une augmentation du travail familial non rémunéré, en particulier dans les zones rurales. Néanmoins, les preuves de l'impact de la migration de retour sur l'offre de main-d'œuvre féminine sont limitées.

Les études examinant l'impact de la migration actuelle identifient deux canaux principaux par lesquels la migration d'un membre masculin du ménage peut influencer la participation des femmes au marché du travail. Premièrement, il y a un effet de revenu où la migration d'un membre du ménage peut augmenter le revenu du ménage par le biais des envois de fonds, ce qui peut conduire à une augmentation du salaire de réserve et à une baisse du coût d'opportunité des loisirs, réduisant ainsi la nécessité pour les femmes de travailler en dehors du foyer et agissant comme un mécanisme d'assurance (Chami et al., 2005). Plusieurs études vont dans ce sens, par exemple Amuedo-Dorantes and Pozo (2006), Mendola and Carletto (2012) Binzel and Assaad (2011) et Lenoël and David (2019). Un autre canal potentiel est l'effet de substitution, dans lequel l'absence du principal soutien

[2] Cette littérature est détaillée dans le chapitre 3 en section 3.1.

de famille entraîne une redistribution du travail au sein du ménage par les individus restés au pays pour compenser la perte de travail ou de revenu du migrant. La recherche suggère que cet effet de substitution se manifeste principalement par une augmentation du travail familial non rémunéré, en particulier dans les zones rurales (Binzel and Assaad, 2011; Lenoël and David, 2019).

Contrairement à l'abondante littérature sur l'impact des membres actuels de la migration, l'influence des migrants de retour dans le ménage a reçu comparativement moins d'attention. On pourrait s'attendre à ce que la migration de retour annule les effets de la migration sur la participation des femmes au marché du travail. En ce qui concerne l'effet sur le revenu, l'arrêt des transferts de fonds augmente la pression financière, ce qui pourrait inciter les femmes à entrer sur le marché du travail. En ce qui concerne l'effet de substitution, les migrants de retour peuvent réduire la participation des femmes au marché du travail s'ils reprennent leurs tâches ménagères traditionnelles à leur retour, ce qui entraîne une réduction du travail familial non rémunéré pour les femmes restées au pays. Cependant, la migration de retour peut également rapporter du capital qui peut permettre aux femmes de poursuivre une activité indépendante, et elle peut influencer la participation des femmes au marché du travail par la transmission des normes de genre des pays de destination vers les pays d'origine. Les migrants de retour peuvent ramener du capital humain et physique qui sera réaffecté ou investi par les membres du ménage (Gubert and Nordman, 2011; Wahba, 2021). Par conséquent, les femmes peuvent utiliser le capital acquis par les hommes pendant les périodes de migration pour exercer une activité indépendante (Mendola and Carletto, 2012). En outre, comme indiqué plus haut, il existe un nombre croissant de publications sur la transmission des normes sociales des pays de destination aux pays d'origine, en particulier en ce qui concerne les normes de genre, qui peuvent influer sur la participation des femmes au marché du travail à leur retour (Tuccio and Wahba, 2018). Dans l'ensemble, l'impact de la migration masculine actuelle et de retour sur la participation des femmes au marché du travail parmi celles qui sont restées au pays est complexe et repose sur divers mécanismes interdépendants.

Contribution du Chapitre 3

La contribution de ce chapitre est de fournir une approche holistique dans l'étude de l'impact de la migration masculine actuelle et de retour sur la participation des femmes au marché du travail et sur leur statut professionnel en Égypte, tout en considérant le contexte géographique des zones urbaines et rurales. Cette étude contribue à la littérature susmentionnée de plusieurs manières. Sur le plan méthodologique, elle utilise un ensemble

de données de panel en trois vagues pour mieux contrôler l'hétérogénéité non observée, contrairement aux analyses transversales antérieures. Sur le plan empirique, elle traite à la fois des impacts de la migration actuelle et de la migration de retour, comblant ainsi une lacune dans la compréhension des effets de la migration de retour sur l'offre de main-d'œuvre féminine. Sur le plan analytique, elle explore les facteurs économiques et culturels souvent négligés dans les études précédentes.

En utilisant les données des trois dernières vagues de l'enquête par panel sur le marché du travail en Égypte (2006, 2012, 2018), j'adopte une stratégie d'estimation à effets fixes avec un ensemble complet de contrôles pour atténuer les problèmes d'autosélection. Les contrôles variables dans le temps tiennent compte de la sélection observable, tandis que les effets fixes individuels tiennent compte de l'hétérogénéité non observée variable dans le temps au niveau individuel, et servent également de substituts aux effets fixes de couple ou de famille. Malgré ces contrôles, le biais potentiel dû à l'omission de variables temporelles est évalué de manière plus approfondie à l'aide de la méthodologie d'Oster.

Les résultats révèlent que la présence d'un migrant actuel augmente la probabilité d'exercer un travail salarié dans les zones urbaines et un travail familial non rémunéré dans les zones rurales, sous l'impulsion des migrants qui n'envoient pas de fonds et ceux qui séjournent plus longtemps à l'étranger pour les premiers et moins longtemps pour les seconds. À l'inverse, les migrants de retour n'ont pas d'impact significatif sur l'offre de main-d'œuvre féminine dans les zones urbaines, mais affectent négativement le travail salarié des femmes dans les zones rurales, en particulier chez les migrants économiquement plus aisés, ce qui laisse supposer une transmission des normes. Des effets hétérogènes basés sur le statut de la relation entre le migrant et la femme restée au pays sont également observés. Ces résultats soulignent la nécessité de prendre en compte les diverses catégories d'emploi et les statuts des relations dans des contextes géographiques spécifiques pour comprendre pleinement les réponses économiques des femmes à la migration masculine.

Bibliography

Abadie, A., Athey, S., Imbens, G.W., Wooldridge, J.M., 2023. When should you adjust standard errors for clustering? The Quarterly Journal of Economics 138, 1–35.

Abramitzky, R., Boustan, L., 2022. Streets of gold: America's untold story of immigrant success. PublicAffairs.

Abramitzky, R., Boustan, L., Eriksson, K., 2020. Do immigrants assimilate more slowly today than in the past? American Economic Review: Insights 2, 125–41. doi:10.1257/aeri.20190079.

Abramitzky, R., Boustan, L.P., Eriksson, K., 2014. A nation of immigrants: Assimilation and economic outcomes in the age of mass migration. Journal of Political Economy 122, 467–506.

Acemoglu, D., Ajzenman, N., Aksoy, C.G., Fiszbein, M., Molina, C.A., 2021. (Successful) Democracies Breed Their Own Support. Working Paper 29167. National Bureau of Economic Research. doi:10.3386/w29167.

Acemoglu, D., Hassan, T.A., Tahoun, A., 2018. The power of the street: Evidence from egypt's arab spring. The Review of Financial Studies 31, 1–42. doi:https://doi.org/10.1093/rfs/hhx086.

Acosta, P., 2006. Labor supply, school attendance, and remittances from international migration: the case of el salvador. World Bank Policy Research Working Paper doi:10.1596/1813-9450-3903.

Adão, R., Kolesár, M., Morales, E., 2019. Shift-Share Designs: Theory and Inference. The Quarterly Journal of Economics 134, 1949–2010. URL: https://doi.org/10.1093/qje/qjz025, doi:10.1093/qje/qjz025, arXiv:https://academic.oup.com/qje/article-pdf/134/4/1949/32666168/qjz025.pdf.

Alba, R.D., Nee, V., 2009. Remaking the American mainstream: Assimilation and contemporary immigration. Harvard University Press.

Alcaraz, C., Chiquiar, D., Salcedo, A., 2012. Remittances, schooling, and child labor in mexico. Journal of Development Economics 97, 156–165. doi:https://doi.org/10.1016/j.jdeveco.2010.11.004.

Alesina, A., Giuliano, P., 2015. Culture and institutions. Journal of Economic Literature 53, 898–944.

Alesina, A., Miano, A., Stantcheva, S., 2022. Immigration and Redistribution. The Review of Economic Studies 90, 1–39. doi:10.1093/restud/rdac011.

Alesina, A., Tabellini, G., Trebbi, F., 2017. Is europe an optimal political area? Brookings Papers on Economic Activity , 169–213doi:http://www.jstor.org/stable/90013171.

Alesina, A., Tabellini, M., 2022. The political effects of immigration: Culture or economics? NBER No. 30079 .

Algan, Y., Bisin, A., Manning, A., Verdier, T., 2012. Cultural integration of immigrants in Europe. Oxford University Press.

Algan, Y., Guriev, S., Papaioannou, E., Passari, E., 2017. The european trust crisis and the rise of populism. Brookings papers on economic activity 2017, 309–400.

Allcott, H., Gentzkow, M., 2017. Social Media and Fake News in the 2016 Election. Journal of Economic Perspectives 31, 211–236. doi:10.1257/jep.31.2.211.

Allport, G.W., Clark, K., Pettigrew, T., 1954. The nature of prejudice .

Altonji, J.G., Card, D., 2018. The effects of immigration on the labor market outcomes of less-skilled natives, in: The New Immigrant in the American Economy. Routledge, pp. 137–170.

Amaral-Garcia, S., Nardotto, M., Propper, C., Valletti, T., 2022. Mums go online: Is the Internet changing the demand for health care? Review of Economics and Statistics 104, 1157–1173. doi:10.1162/rest_a_01033.

Amorim, G., Lima, R.C., Sampaio, B., 2022. Broadband internet and protests: Evidence from the occupy movement. Information Economics and Policy 60, 100982. doi:https://doi.org/10.1016/j.infoecopol.2022.100982.

Amuedo-Dorantes, C., Pozo, S., 2006. Migration, remittances, and male and female employment patterns. American Economic Review 96, 222–226. doi:10.1257/000282806777211946.

Anderson, T.W., Rubin, H., 1949. Estimation of the parameters of a single equation in a complete system of stochastic equations. The Annals of mathematical statistics 20, 46–63. doi:10.1214/aoms/1177730090.

Andrews, I., Stock, J.H., Sun, L., 2019. Weak instruments in instrumental variables regression: Theory and practice. Annual Review of Economics 11, 727–753. doi:10.1146/annurev-economics-080218-025643.

Andrews, K.T., Beyerlein, K., Tucker Farnum, T., 2016. The legitimacy of protest: explaining white southerners' attitudes toward the civil rights movement. Social Forces 94, 1021–1044. doi:10.1093/sf/sov097.

Angrist, J.D., Krueger, A.B., 2001. Instrumental variables and the search for identification: From supply and demand to natural experiments. Journal of Economic perspectives 15, 69–85. doi:10.1257/jep.15.4.69.

Antman, F.M., 2012. Gender, educational attainment, and the impact of parental migration on children left behind. Journal of Population Economics 25, 1187–1214. doi:https://doi.org/10.1007/s00148-012-0423-y.

Arouri, M., Nguyen, C.V., 2018. Does international migration affect labor supply, nonfarm diversification and welfare of households? evidence from egypt. International Migration 56, 39–62. doi:10.1111/imig.12384.

Ashraf, Q., Galor, O., 2013. Genetic diversity and the origins of cultural fragmentation. American Economic Review 103, 528–533.

Assaad, R., Hendy, R., Lassassi, M., Yassin, S., 2020. Explaining the mena paradox: Rising educational attainment, yet stagnant female labor force participation. Demographic Research 43, 817. doi:10.4054/demres.2020.43.28.

Assaad, R., Krafft, C., Selwaness, I., 2022. The impact of marriage on women's employment in the middle east and north africa. Feminist Economics 28, 247–279. doi:10.1080/13545701.2021.2007415.

Ausserladscheider, V., 2019. Beyond economic insecurity and cultural backlash: Economic nationalism and the rise of the far right. Sociology Compass 13, e12670.

Autor, D.H., Duggan, M.G., 2003. The rise in the disability rolls and the decline in unemployment. The Quarterly Journal of Economics 118, 157–206.

Bailard, C.S., 2012. Testing the Internet's Effect on Democratic Satisfaction: A Multi-Methodological, Cross-National Approach. Journal of Information Technology & Politics 9, 185–204. doi:10.1080/19331681.2011.641495.

Barrera, O., Guriev, S., Henry, E., Zhuravskaya, E., 2020. Facts, alternative facts, and

fact checking in times of post-truth politics. Journal of Public Economics 182, 104123. doi:10.1016/j.jpubeco.2019.104123.

Barsbai, T., Rapoport, H., Steinmayr, A., Trebesch, C., 2017. The effect of labor migration on the diffusion of democracy: evidence from a former soviet republic. American Economic Journal: Applied Economics 9, 36–69. doi:10.1257/app.20150517.

Batista, C., McIndoe-Calder, T., Vicente, P.C., 2017. Return migration, self-selection and entrepreneurship. Oxford Bulletin of Economics and Statistics 79, 797–821. doi:10.1111/obes.12176.

Bazzi, S., Fiszbein, M., Gebresilasse, M., 2020. Frontier culture: The roots and persistence of "rugged individualism" in the united states. Econometrica 88, 2329–2368.

Bazzi, S., Gaduh, A., Rothenberg, A.D., Wong, M., 2019. Unity in diversity? how intergroup contact can foster nation building. American Economic Review 109, 3978–4025.

Becker, S.O., Fetzer, T., et al., 2016. Does migration cause extreme voting. Center for Competitive Advantage in the Global Economy and The Economic & Social Research Council , 1–54.

Beine, M., Docquier, F., Schiff, M., 2013. International migration, transfer of norms and home country fertility. Canadian Journal of Economics/Revue canadienne d'économique 46, 1406–1430. doi:10.1111/caje.12062.

Bensassi, S., Jabbour, L., 2017. Return migration and entrepreneurial success: an empirical analysis for Egypt. Technical Report. GLO Discussion Paper.

Bensassi, S., Jabbour, L., 2022. Beyond experience and capital. is there a return to return migration? The Journal of Development Studies 58, 730–751. doi:10.1080/00220388.2021.1988076.

Bertoli, S., Gautrain, E., Murard, E., 2023. Left behind but not immobile: Living arrangements of mexican transnational households. Economic Development and Cultural Change 71, 000–000. doi:10.1086/717282.

Bertoli, S., Marchetta, F., 2015. Bringing it all back home–return migration and fertility choices. World Development 65, 27–40. doi:10.1016/j.worlddev.2013.08.006.

Bertrand, M., Kamenica, E., 2023. Coming apart? cultural distances in the united states over time. American Economic Journal: Applied Economics .

Binzel, C., Assaad, R., 2011. Egyptian men working abroad: Labour supply responses by the women left behind. Labour Economics 18, S98–S114. doi:10.1016/j.labeco.2011.03.002.

Bisin, A., Verdier, T., 2001. The economics of cultural transmission and the dynamics of preferences. Journal of Economic Theory 97, 298–319.

Borjas, G.J., 2003. The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. The quarterly journal of economics 118, 1335–1374. doi:https://doi.org/10.1162/003355303322552810.

Borjas, G.J., 2014. Immigration economics. Harvard University Press.

Borjas, G.J., 2017. The wage impact of the marielitos: A reappraisal. Ilr Review 70, 1077–1110. doi:https://www.jstor.org/stable/26944704.

Borjas, G.J., Bratsberg, B., 1996. Who leaves? the outmigration of the foreign-born. doi:10.2307/2109856.

Borjas, G.J., Monras, J., 2017. The labour market consequences of refugee supply shocks. Economic Policy 32, 361–413. doi:https://doi.org/10.1093/epolic/eix007.

Borusyak, K., Hull, P., 2023. Nonrandom exposure to exogenous shocks. Econometrica 91, 2155–2185. doi:10.3982/ECTA19367.

Borusyak, K., Hull, P., Jaravel, X., 2022. Quasi-experimental shift-share research designs. The Review of Economic Studies 89, 181–213.

Bosch, T.E., Admire, M., Ncube, M., 2020. Facebook and politics in africa: Zimbabwe and kenya. Media, Culture & Society 42, 349–364. doi:10.1177/0163443719895194.

Boyer, P.C., Delemotte, T., Gauthier, G., Rollet, V., Schmutz, B., 2020. Social Media and the Dynamics of Protests. CESifo Working Paper Series 8326. doi:10.2139/ssrn.3618851.

Branton, R., Martinez-Ebers, V., Carey Jr, T.E., Matsubayashi, T., 2015. Social protest and policy attitudes: The case of the 2006 immigrant rallies. American Journal of Political Science 59, 390–402. doi:10.1111/ajps.12159.

Cabegin, E., 2006. The effect of filipino overseas migration on the non-migrant spouse's market participation and labor supply behavior. IZA Discussion Paper No. 2240 doi:10.2139/ssrn.924567.

Campante, F., Durante, R., Sobbrio, F., 2018. Politics 2.0: The Multifaceted Effect of Broadband Internet on Political Participation. Journal of the European Economic Association 16, 1094–1136. doi:10.1093/jeea/jvx044.

Card, D., 2001. Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. Journal of Labor Economics 19, 22–64. URL: http://www.jstor.org/stable/10.1086/209979.

Card, D., Dustmann, C., Preston, I., 2012. Immigration, wages, and compositional amenities. Journal of the European Economic Association 10, 78–119. doi:https://doi.org/10.1111/j.1542-4774.2011.01051.x.

Cariolle, J., 2021. International connectivity and the digital divide in Sub-Saharan Africa. Information Economics and Policy 55, 100901. doi:10.1016/j.infoecopol.2020.100901.

Cariolle, J., Le Goff, M., 2023. Spatial internet spillovers in manufacturing. The Journal of Development Studies 59, 1163—1186. doi:10.1080/00220388.2023.2204177.

Ceron, A., Memoli, V., 2016. Flames and Debates: Do Social Media Affect Satisfaction with Democracy? Social Indicators Research 126, 225–240. doi:10.1007/s11205-015-0893-x.

Chami, R., Fullenkamp, C., Jahjah, S., 2005. Are immigrant remittance flows a source of capital for development? IMF Staff papers 52, 55–81. doi:https://www.jstor.org/stable/30035948.

Chang, W.C., 2018. Media Use and Satisfaction with Democracy: Testing the Role of Political Interest. Social Indicators Research 140, 999–1016. doi:10.1007/s11205-017-1806-y.

Chauvet, L., Mercier, M., 2014. Do return migrants transfer political norms to their origin country? evidence from mali. Journal of Comparative Economics 42, 630–651. doi:10.1016/j.jce.2014.01.001.

Chavula, J., Feamster, N., Bagula, A., Suleman, H., 2015. Quantifying the Effects of Circuitous Routes on the Latency of Intra-Africa Internet Traffic: A Study of Research and Education Networks, in: Nungu, A., Pehrson, B., Sansa-Otim, J. (Eds.), E-Infrastructure and e-Services for Developing Countries, Springer International Publishing, Cham. pp. 64–73. doi:10.1007/978-3-319-16886-9_7.

Claassen, C., 2020. Does public support help democracy survive? American Journal of Political Science 64, 118–134. doi:https://doi.org/10.1111/ajps.12452.

Collier, P., 2013. Exodus: Immigration and multiculturalism in the 21st century. Penguin UK.

Cox-Edwards, A., Rodríguez-Oreggia, E., 2009. Remittances and labor force participation in mexico: an analysis using propensity score matching. World Development 37, 1004–1014. doi:10.1016/j.worlddev.2008.09.010.

David, A., El-Mallakh, N., Wahba, J., 2019. Internal versus international migration in egypt: together or far apart, Economic Research Forum (ERF).

Davies, S., Pettersson, T., Öberg, M., 2022. Organized violence 1989–2021 and drone warfare. Journal of Peace Research 59, 593–610.

De Bel-Air, F., 2016. Migration profile: Egypt, migration policy centre. Policy Briefs 1. doi:10.2870/363523.

De Haas, H., Van Rooij, A., 2010. Migration as emancipation? the impact of internal and international migration on the position of women left behind in rural morocco. Oxford development studies 38, 43–62. doi:https://doi.org/10.1080/13600810903551603.

Derenoncourt, E., 2022. Can you move to opportunity? evidence from the great migration. American Economic Review 112, 369–408.

Desmet, K., Ortuño-Ortín, I., Wacziarg, R., 2017. Culture, ethnicity, and diversity. American Economic Review 107, 2479–2513. URL: https://www.aeaweb.org/articles?id=10.1257/aer.20150243, doi:10.1257/aer.20150243.

Desmet, K., Wacziarg, R., 2021. The cultural divide. The Economic Journal 131, 2058–2088.

Diabate, I., Mesplé-Somps, S., et al., 2019. Female genital mutilation and migration in mali: do return migrants transfer social norms? Journal of Population Economics 32, 1125–1170. doi:10.1007/s00148-019-00733-w.

Diamond, L., 2010. Liberation Technology. Journal of Democracy 21, 69–83. doi:10.1353/jod.0.0190.

Docquier, F., Tansel, A., Turati, R., 2020a. Do emigrants self-select along cultural traits? evidence from the mena countries. International Migration Review 54, 388–422.

Docquier, F., Turati, R., Valette, J., Vasilakis, C., 2020b. Birthplace diversity and economic growth: evidence from the US states in the Post-World War II period. Journal of Economic Geography 20, 321–354. doi:10.1093/jeg/lbz016.

Donati, D., 2023. Mobile internet access and political outcomes: Evidence from South Africa. Journal of Development Economics 162, 103073. doi:10.1016/j.jdeveco.2023.103073.

Dustmann, C., Glitz, A., 2011. Migration and education, in: Handbook of the Economics of Education. Elsevier. volume 4, pp. 327–439. doi:10.1016/B978-0-444-53444-6.00004-3.

Dustmann, C., Preston, I.P., 2007a. Racial and economic factors in attitudes to immigration. The BE Journal of Economic Analysis & Policy 7. doi:https://doi.org/10.2202/1935-1682.1655.

Dustmann, C., Preston, I.P., 2007b. Racial and economic factors in attitudes to immigration. The BE Journal of Economic Analysis & Policy 7, 1–41.

Dustmann, C., Vasiljeva, K., Piil Damm, A., 2018. Refugee Migration and Electoral Outcomes. The Review of Economic Studies 86, 2035–2091.

Dustmann, C., Vasiljeva, K., Piil Damm, A., 2019. Refugee migration and electoral outcomes. The Review of Economic Studies 86, 2035–2091. doi:https://doi.org/10.1093/restud/rdy047.

D'Andrea, A., Limodio, N., 2023. High-speed Internet, financial technology, and banking. Management Science doi:10.1287/mnsc.2023.4703.

Edo, A., 2019. The impact of immigration on the labor market. Journal of Economic Surveys 33, 922–948.

Edo, A., Giesing, Y., Öztunc, J., Poutvaara, P., 2019a. Immigration and electoral support for the far-left and the far-right. European Economic Review 115, 99–143. doi:https://doi.org/10.1016/j.euroecorev.2019.03.001.

Edo, A., Giesing, Y., Öztunc, J., Poutvaara, P., 2019b. Immigration and electoral support for the far-left and the far-right. European Economic Review 115, 99 – 143.

Edo, A., Toubal, F., 2015. Selective immigration policies and wages inequality. Review of International Economics 23, 160–187. doi:https://doi.org/10.1111/roie.12163.

Edo, A., Özgüzel, C., 2023. The impact of immigration on the employment dynamics of

european regions. Labour Economics 85, 102433. doi:https://doi.org/10.1016/j.labeco.2023.102433.

Eichengreen, B., Lafarguette, R., Mehl, A., Minesso, M.F., 2023. Technology and the geography of the foreign exchange market. Journal of International Money and Finance 131, 102802. doi:10.1016/j.jimonfin.2023.102802.

El-Mallakh, N., 2020. How do protests affect electoral choices? Evidence from Egypt. Journal of Economic Behavior & Organization 179, 299–322. doi:10.1016/j.jebo.2020.09.005.

El-Mallakh, N., Wahba, J., 2021. Upward or downward: Occupational mobility and return migration. World Development 137, 105203. doi:10.1016/j.worlddev.2020.105203.

EMDAT, 2022. The Emergency Events Database, Université catholique de Louvain (UCL), Centre for Research on the Epidemiology of Disasters (CRED). Technical Report. URL: https://www.emdat.be.

Enikolopov, R., Makarin, A., Petrova, M., 2020. Social Media and Protest Participation: Evidence From Russia. Econometrica 88, 1479–1514. doi:10.3982/ECTA14281.

Enikolopov, R., Petrova, M., Sonin, K., 2018. Social Media and Corruption. American Economic Journal: Applied Economics 10, 150–174. doi:10.1257/app.20160089.

Evans, O., 2019. Digital politics: Internet and democracy in Africa. Journal of Economic Studies 46, 169–191. doi:10.1108/JES-08-2017-0234.

Falck, O., Gold, R., Heblich, S., 2014. E-lections: Voting Behavior and the Internet. American Economic Review 104, 2238–2265. doi:10.1257/aer.104.7.2238.

Fanou, R., Valera, F., Dhamdhere, A., 2017. Investigating the causes of congestion on the African IXP substrate, in: Proceedings of the 2017 Internet Measurement Conference, Association for Computing Machinery, New York. p. 57–63. doi:10.1145/3131365.3131394.

Fergusson, L., Molina, C., 2019. Facebook causes protests. Documento CEDE No.41. doi:10.2139/ssrn.3553514.

Feyrer, J., 2019. Trade and income—exploiting time series in geography. American Economic Journal: Applied Economics 11, 1–35.

Fisman, R., Miguel, E., 2007. Corruption, norms, and legal enforcement: Evidence from diplomatic parking tickets. Journal of Political Economy 115, 1020–1048.

Flynn, D.J., Nyhan, B., Reifler, J., 2017. The nature and origins of misperceptions: Understanding false and unsupported beliefs about politics. Political Psychology 38, 127–150. doi:10.1111/pops.12394.

Foster, C., Frieden, J., 2017. Crisis of trust: Socio-economic determinants of europeans' confidence in government. European Union Politics 18, 511–535.

Fouka, V., Mazumder, S., Tabellini, M., 2022. From immigrants to americans: Race and assimilation during the great migration. The Review of Economic Studies 89, 811–842.

Funke, M., Schularick, M., Trebesch, C., 2016. Going to extremes: Politics after financial crises, 1870–2014. European Economic Review 88, 227–260.

Funkhouser, E., 1992. Mass emigration, remittances, and economic adjustment: the case of el salvador in the 1980s, in: Immigration and the work force: Economic consequences for the United States and source areas. university of Chicago Press, pp. 135–176.

Galli, F., Russo, G., 2019. Immigration restrictions and second-generation cultural assimilation: Theory and quasi-experimental evidence. Journal of Population Economics 32, 23–51.

Gavazza, A., Nardotto, M., Valletti, T., 2019. Internet and Politics: Evidence from U.K. Local Elections and Local Government Policies. The Review of Economic Studies 86, 2092–2135. doi:10.1093/restud/rdy028.

Giavazzi, F., Petkov, I., Schiantarelli, F., 2019. Culture: Persistence and evolution. Journal of Economic Growth 24, 117–154.

Gibson, J., McKenzie, D., Stillman, S., 2011. What happens to diet and child health when migration splits households? evidence from a migration lottery program. Food Policy 36, 7–15. doi:https://doi.org/10.1016/j.foodpol.2010.08.003.

Giuliano, P., Nunn, N., 2021. Understanding cultural persistence and change. The Review of Economic Studies 88, 1541–1581.

Giuliano, P., Tabellini, M., 2021. The Seeds of Ideology: Historical Immigration and Political Preferences in the United States. Harvard Business School BGIE Unit Working Paper .

Gleditsch, N.P., Wallensteen, P., Eriksson, M., Sollenberg, M., Strand, H., 2002. Armed conflict 1946-2001: A new dataset. Journal of Peace Research 39, 615–637.

Goldsmith-Pinkham, P., Sorkin, I., Swift, H., 2020. Bartik Instruments: What, When,

Why, and How. American Economic Review 110, 2586–2624. doi:10.1257/aer.20181047.

Goldsmith-Pinkham, P., Sorkin, I., Swift, H., 2020. Bartik instruments: What, when, why, and how. American Economic Review 110, 2586–2624.

Gonnot, J., lo Polito, F., 2021. The political assimilation of immigrants:migrant-to-native differences in western europe. Robert Schuman Centre for Advanced Studies Migration Policy Centre No. 2021-87 .

Gonnot, J., lo Polito, F., 2023. Cultural Transmission and Political Attitudes: Explaining Differences between Natives and Immigrants in Western Europe. Technical Report. CEPII, Documents de Travail.

Gorman, S.P., Schintler, L., Kulkarni, R., Stough, R., 2004. The Revenge of Distance: Vulnerability Analysis of Critical Information Infrastructure. Journal of Contingencies and Crisis Management 12, 48–63. doi:10.1111/j.0966-0879.2004.00435.x.

Greenberg, J.H., 1956. The measurement of linguistic diversity. Language 32, 109–115.

Greene, W., 2004. The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects. The Econometrics Journal 7, 98–119. doi:10.1111/j.1368-423X.2004.00123.x.

Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., Lazer, D., 2019. Fake news on twitter during the 2016 U.S. presidential election. Science 363, 374–378. doi:10.1126/science.aau2706.

Grubesic, T.H., Murray, A.T., 2006. Vital Nodes, Interconnected Infrastructures, and the Geographies of Network Survivability. Annals of the Association of American Geographers 96, 64–83. doi:10.1111/j.1467-8306.2006.00499.x.

Grubesic, T.H., O'Kelly, M.E., Murray, A.T., 2003. A geographic perspective on commercial Internet survivability. Telematics and Informatics 20, 51–69. doi:10.1016/S0736-5853(02)00003-5.

GSMA, 2022. The Mobile Economy 2022. Technical Report. GSMA Association. URL: https://www.gsma.com/mobileeconomy/wp-content/uploads/2022/02/280222-The-Mobile-Economy-2022.pdf.

Gubert, F., Nordman, C.J., 2011. Return migration and small enterprise development in the maghreb. Diaspora for development in Africa 3, 103–126.

Guiffard, J.B., 2022. The Impact of Broadband Internet on Political Mobilization in Sub-Saharan Africa. Working paper doi:10.2139/ssrn.4247083.

Guiso, L., Herrera, H., Morelli, M., Sonno, T., et al., 2020. Economic insecurity and the demand of populism in europe. Einaudi Institute for Economics and Finance 3.

Guriev, S., Melnikov, N., Zhuravskaya, E., 2021. 3G Internet and Confidence in Government. The Quarterly Journal of Economics 136, 2533–2613. doi:10.1093/qje/qjaa040.

Guriev, S., Papaioannou, E., 2022. The political economy of populism. Journal of Economic Literature 60, 753–832.

Hall, M., Tideman, N., 1967. Measures of concentration. Journal of the american statistical association 62, 162–168.

Halla, M., Wagner, A.F., Zweimüller, J., 2017. Immigration and voting for the far right. Journal of the European economic association 15, 1341–1385. doi:https://www.jstor.org/stable/90023600.

Hamdouch, B., Wahba, J., 2015. Return migration and entrepreneurship in morocco. Middle East Development Journal 7, 129–148. doi:10.1080/17938120.2015.1072696.

Hatte, S., Loper, J., Taylor, T., 2023. Connecting the Unconnected: Facebook Access and Female Political Representation in Sub-Saharan Africa. Post-Print. HAL. https://EconPapers.repec.org/RePEc:hal:journl:hal-04310234.

Hendy, R., 2015. Women's participation in the egyptian labor market: 1998-2012. The Egyptian labor market in an era of revolution , 147–161doi:10.1093/acprof:oso/9780198737254.003.0008.

Henry, E., Zhuravskaya, E., Guriev, S., 2022. Checking and sharing alt-facts. American Economic Journal: Economic Policy 14, 55–86. doi:10.1257/pol.20210037.

Hoodfar, H., 1993. The impact of egyptian male migration on urban families left behind:'feminization of the egyptian family'or a re-affirmation of traditional gender roles. Sociological Bulletin (India) 42, 115–135. doi:23620250.

Houngbonon, G.V., Mensah, J.T., Traore, N., 2022. The Impact of Internet Access on Innovation and Entrepreneurship in Africa. World Bank Policy Research Paper doi:hdl.handle.net/10986/37059.

Hunt, J., 1992. The impact of the 1962 repatriates from algeria on the french labor market. Ilr Review 45, 556–572. doi:https://doi.org/10.1177/001979399204500310.

Imbruno, M., Cariolle, J., De Melo, J., 2022. Digital connectivity and firm participation in foreign markets: An exporter-based bilateral analysis. CEPR Press Discussion Paper No. 17318 doi:10.2139/ssrn.4169512.

ITU, 2022. World telecommunication/ict indicators database 2022. URL: https://www.itu.int/en/ITU-D/Statistics/Pages/publications/wtid.aspx.

Ivanov, D., 2023. Economic insecurity, institutional trust and populist voting across europe. Comparative Economic Studies , 1–22.

Jaschke, P., Sardoschau, S., Tabellini, M., 2022. Scared Straight? Threat and Assimilation of Refugees in Germany. Technical Report. National Bureau of Economic Research.

Jha, C.K., Kodila-Tedika, O., 2020. Does social media promote democracy? Some empirical evidence. Journal of Policy Modeling 42, 271–290. doi:10.1016/j.jpolmod.2019.05.010.

Jones, J.J., Bond, R.M., Bakshy, E., Eckles, D., Fowler, J.H., 2017. Social influence and political mobilization: Further evidence from a randomized experiment in the 2012 U.S. presidential election. PLOS ONE 12, e0173851. doi:10.1371/journal.pone.0173851.

Kabeer, N., 2021. Gender equality, inclusive growth, and labour markets. Women's Economic Empowerment: Insights from Africa and South Asia , 13–48.

Kan, S., Aytimur, R.E., 2019. Labor force participation of women left behind in tajikistan. Oxford Development Studies 47, 1–28. doi:10.1080/13600818.2018.1484899.

Karymshakov, K., Sulaimanova, B., 2017. Migration impact on left-behind women's labour participation and time-use: Evidence from Kyrgyzstan. Technical Report. WIDER working paper. doi:10.35188/UNU-WIDER/2017/343-1.

Keane, M., Neal, T., 2023. Instrument strength in IV estimation and inference: A guide to theory and practice. Journal of Econometrics 235, 1625–1653. doi:10.1016/j.jeconom.2022.12.009.

Keita, S., Renault, T., Valette, J., 2023. The Usual Suspects: Offender Origin, Media Reporting and Natives' Attitudes Towards Immigration. The Economic Journal 134, 322–362. doi:10.1093/ej/uead059.

Ketchley, N., El-Rayyes, T., 2021. Unpopular protest: Mass mobilization and attitudes to democracy in post-mubarak Egypt. The Journal of Politics 83, 291–305. doi:10.1086/709298.

Khaled, L., 1995. Migration and women's status: the jordan case. International Migration (Geneva, Switzerland) 33, 235–250. doi:10.1111/j.1468-2435.1995.tb00027.x.

Khan, M.I., Valatheeswaran, C., 2016. International migration, remittances and labour force participation of left-behind family members: A study of kerala. Margin: The Journal of Applied Economic Research 10, 86–118. doi:10.1177/0973801015612669.

Kilic, T., Carletto, C., Davis, B., Zezza, A., 2009. Investing back home return migration and business ownership in albania 1. Economics of Transition 17, 587–623. doi:10.1111/j.1468-0351.2009.00357.x.

King, G., Pan, J., Roberts, M.E., 2017. How the chinese government fabricates social media posts for strategic distraction, not engaged argument. American political science review 111, 484–501. doi:https://doi.org/10.1017/S0003055417000144.

Krafft, C., Assaad, R., Keo, C., 2019. The evolution of labor supply in egypt from 1988-2018: A gendered analysis. Economic Research Forum (ERF) Working Papers .

Kroeber, A.L., Kluckhohn, C., 1952. Culture: A critical review of concepts and definitions. Papers. Peabody Museum of Archaeology & Ethnology, Harvard University .

Lee, B.S., Jang, S., Sarkar, J., 2008. Women's labor force participation and marriage: The case of korea. Journal of Asian Economics 19, 138–154.

Lei, Y.W., 2013. The political consequences of the rise of the internet: Political beliefs and practices of chinese netizens, in: Political communication in China. Routledge, pp. 31–62.

Lenoël, A., David, A., 2019. Leaving work behind? the impact of emigration on female labor force participation in morocco. International Migration Review 53, 122–153. doi:10.1177/0197918318768553.

Levitt, P., 1998. Social remittances: Migration driven local-level forms of cultural diffusion. International migration review 32, 926–948.

Li, X., Zhou, Y., Zhao, M., Zhao, X., 2020. A harmonized global nighttime light dataset 1992–2018. Scientific Data 7, 168. doi:10.1038/s41597-020-0510-y.

Lokshin, M., Glinskaya, E., 2009. The effect of male migration on employment patterns of women in nepal. The World Bank Economic Review 23, 481–507. doi:10.1093/wber/lhp011.

Lorenz-Spreen, P., Oswald, L., Lewandowsky, S., Hertwig, R., 2023. A systematic review

of worldwide causal and correlational evidence on digital media and democracy. Nature human behaviour 7, 74–101. doi:10.1038/s41562-022-01460-1.

Lucas, R.E., Stark, O., 1985. Motivations to remit: Evidence from botswana. Journal of political Economy 93, 901–918. doi:1833062.

Mahé, C., 2022. Return migration and self-employment: is there a 'jack-of-all-trades' effect? Oxford Economic Papers 74, 62–84. doi:10.1093/oep/gpab013.

Malecki, E.J., 2009. The Economic Geography of the Internet's Infrastructure. Economic Geography 78, 399–424. doi:10.1111/j.1944-8287.2002.tb00193.x.

Manacorda, M., Manning, A., Wadsworth, J., 2012. The impact of immigration on the structure of wages: theory and evidence from britain. Journal of the European economic association 10, 120–151. doi:https://doi.org/10.1111/j.1542-4774.2011.01049.x.

Manacorda, M., Tesei, A., 2020. Liberation Technology: Mobile Phones and Political Mobilization in Africa. Econometrica 88, 533–567. doi:10.3982/ECTA14392.

Marchetta, F., 2012. Return migration and the survival of entrepreneurial activities in egypt. World Development 40, 1999–2013. doi:10.1016/j.worlddev.2012.05.009.

Mattes, R., Bratton, M., 2007. Learning about democracy in africa: Awareness, performance, and experience. American Journal of Political Science 51, 192–217. doi:https://doi.org/10.1111/j.1540-5907.2007.00245.x.

Mazumder, S., 2018. The persistent effect of U.S. civil rights protests on political attitudes. American Journal of Political Science 62, 922–935. doi:10.1111/ajps.12384.

McCormick, B., Wahba, J., 2001. Overseas work experience, savings and entrepreneurship amongst return migrants to ldcs. Scottish journal of political economy 48, 164–178. doi:10.1111/1467-9485.00192.

Mendola, M., Carletto, C., 2012. Migration and gender differences in the home labour market: Evidence from albania. Labour Economics 19, 870–880. doi:10.1016/j.labeco.2012.08.009.

Menon, R.R., Bhagat, R., 2020. Emigration and its effect on the labour force participation of women in the left-behind household, in: India Migration Report 2020. Routledge India, pp. 162–176.

Merkle, L., Zimmermann, K.F., 1992. Savings, remittances, and return migration. Economics Letters 38, 77–81. doi:10.1016/0165-1765(92)90165-U.

Miho, A., Jarotschkin, A., Zhuravskaya, E., 2023. Diffusion of gender norms: Evidence from stalin's ethnic deportations. Journal of the European Economic Association .

Miner, L., 2015. The unintended consequences of internet diffusion: Evidence from Malaysia. Journal of Public Economics 132, 66–78. doi:10.1016/j.jpubeco.2015.10.002.

Mocanu, D., Rossi, L., Zhang, Q., Karsai, M., Quattrociocchi, W., 2015. Collective attention in the age of (mis) information. Computers in Human Behavior 51, 1198–1204. doi:10.1016/j.chb.2015.01.024.

Monras, J., 2020. Immigration and wage dynamics: Evidence from the mexican peso crisis. Journal of Political Economy 128, 3017–3089. doi:https://doi.org/10.1086/707764.

Montalvo, J.G., Reynal-Querol, M., 2005. Ethnic polarization, potential conflict, and civil wars. American economic review 95, 796–816.

Moriconi, S., Peri, G., Turati, R., 2022. Skill of the immigrants and vote of the natives: Immigration and nationalism in european elections 2007–2016. European Economic Review 141, 103986. doi:https://doi.org/10.1016/j.euroecorev.2021.103986.

Moriconi, S., Peri, G., Turati, R., 2023. Are Immigrants More Left Wing than Natives? IZA Discussion Paper No. 16164, 1–64.

Mu, R., De Brauw, A., 2015. Migration and young child nutrition: evidence from rural china. Journal of Population Economics 28, 631–657. doi:https://doi.org/10.1007/s00148-015-0550-3.

Nazier, H., Ramadan, R., 2018. Ever married women's participation in labor market in egypt: constraints and opportunities. Middle East Development Journal 10, 119–151. doi:10.1080/17938120.2018.1443605.

Obradovich, N., Özak, Ö., Martín, I., Ortuño-Ortín, I., Awad, E., Cebrián, M., Cuevas, R., Desmet, K., Rahwan, I., Cuevas, Á., 2022. Expanding the measurement of culture with a sample of two billion humans. Journal of the Royal Society Interface 19, 20220085.

Olea, J.L.M., Pflueger, C., 2013. A robust test for weak instruments. Journal of Business & Economic Statistics 31, 358–369. doi:10.1080/00401706.2013.806694.

Orefice, G., Rapoport, H., Santoni, G., 2021. How do immigrants promote exports? Networks, knowledge, diversity CESifo Working Paper No. 9288.

Ortega, F., Peri, G., 2014. Openness and income: The roles of trade and migration.

Journal of International Economics 92, 231–251. doi:https://doi.org/10.1016/j.jinteco.2013.11.008.

Ortega, J., Verdugo, G., . Moving up or down? immigration and the selection of natives across occupations and locations doi:http://dx.doi.org/10.2139/ssrn.2861088.

Oster, E., 2019. Unobservable selection and coefficient stability: Theory and evidence. Journal of Business & Economic Statistics 37, 187–204. doi:10.1080/07350015.2016.1227711.

Ottaviano, G.I., Peri, G., 2006. The economic value of cultural diversity: evidence from us cities. Journal of Economic Geography 6, 9–44.

Ottaviano, G.I., Peri, G., 2012. Rethinking the effect of immigration on wages. Journal of the European economic association 10, 152–197. doi:https://www.jstor.org/stable/41426727.

Otto, A.H., Steinhardt, M.F., 2014. Immigration and election outcomes—evidence from city districts in hamburg. Regional Science and Urban Economics 45, 67–79. doi:https://doi.org/10.1016/j.regsciurbeco.2014.01.004.

Piracha, M., Vadean, F., 2010. Return migration and occupational choice: Evidence from albania. World Development 38, 1141–1155. doi:10.1016/j.worlddev.2009.12.015.

Poutvaara, P., Steinhardt, M.F., 2018. Bitterness in life and attitudes towards immigration. European Journal of Political Economy 55, 471–490. doi:https://doi.org/10.1016/j.ejpoleco.2018.04.007.

Prantl, S., Spitz-Oener, A., 2020. The impact of immigration on competing natives' wages: Evidence from german reunification. Review of Economics and Statistics 102, 79–97. doi:https://doi.org/10.1162/rest_a_00853.

Qin, B., Strömberg, D., Wu, Y., 2017. Why Does China Allow Freer Social Media? Protests versus Surveillance and Propaganda. Journal of Economic Perspectives 31, 117–140. doi:10.1257/jep.31.1.117.

Qin, B., Strömberg, D., Wu, Y., 2021. Social Media and Collective Action in China. CEPR working paper, DP16731 doi:10.2139/ssrn.3976832.

Ragui Assaad, C.K., Selwaness, I., 2022. The impact of marriage on women's employment in the middle east and north africa. Feminist Economics 28, 247–279. doi:10.1080/13545701.2021.2007415.

Rapoport, H., Sardoschau, S., Silve, A., 2020. Migration and cultural change. Technical Report. IZA working paper No. 14772. IZA working paper No. 14772.

Rodriguez, E.R., Tiongson, E.R., 2001. Temporary migration overseas and household labor supply: evidence from urban philippines. International Migration Review 35, 709–725. doi:10.1111/j.1747-7379.2001.tb00037.x.

Samari, G., 2021. Coming back and moving backwards: return migration and gender norms in egypt. Journal of Ethnic and Migration Studies 47, 1103–1118. doi:10.1080/1369183X.2019.1669437.

Sangnier, M., Zylberberg, Y., 2017. Protests and trust in the state: Evidence from African countries. Journal of Public Economics 152, 55–67. doi:10.1016/j.jpubeco.2017.05.005.

Schaub, M., Morisi, D., 2020. Voter mobilisation in the echo chamber: Broadband internet and the rise of populism in Europe. European Journal of Political Research 59, 752–773. doi:10.1111/1475-6765.12373.

Schmitz, S., Weinhardt, F., 2019. Immigration and the evolution of local cultural norms. IZA Discussion Paper .

Schneider-Strawczynski, S., Valette, J., 2023. Media coverage of immigration and the polarization of attitudes. IZA Discussion Paper .

Schumann, R., Kende, M., 2013. Lifting barriers to Internet development in Africa: suggestions for improving connectivity. Analysys Mason Limited, London 9, 30–5. https://www.internetsociety.org/resources/doc/2013/lifting-barriers-to-internet-development-in-africa-suggestions-for-improving-connec

Schwartz, S.H., 2012. An overview of the schwartz theory of basic values. Online readings in Psychology and Culture 2, 2307–0919.

Selwaness, I., Krafft, C., 2021. The dynamics of family formation and women's work: what facilitates and hinders female employment in the middle east and north africa? Population Research and Policy Review 40, 533–587.

Shannon, C.E., 1948. A mathematical theory of communication. The Bell system technical journal 27, 379–423.

Sieverding, M., Hassan, R., 2016. 'her future is marriage': Young people's attitudes

towards gender roles and the gender gap in egypt. Population Council, Cairo doi:10.31899/pgy9.1014.

Silva, J.S., Tenreyro, S., 2006. The log of gravity. The Review of Economics and Statistics 88, 641–658.

Silva, J.S., Tenreyro, S., 2010. On the existence of the maximum likelihood estimates in poisson regression. Economics Letters 107, 310–312.

Sjaastad, L.A., 1962. The costs and returns of human migration. Journal of political Economy 70, 80–93. doi:10.1086/258726.

Spilimbergo, A., 2009. Democracy and foreign education. American economic review 99, 528–543. doi:10.1257/aer.99.1.528.

Steinert-Threlkeld, Z.C., Mocanu, D., Vespignani, A., Fowler, J., 2015. Online social networks and offline protest. EPJ Data Science 4, 1–9. doi:https://doi.org/10.1140/epjds/s13688-015-0056-y.

Steinmayr, A., 2017. Did the refugee crisis contribute to the recent rise of far-right parties in europe? ifo DICE Report 15, 24–27.

Steinmayr, A., 2021. Contact versus Exposure: Refugee Presence and Voting for the Far Right. The Review of Economics and Statistics 103, 310–327. doi:10.1162/rest_a_00922.

Stewart, A.J., McCarty, N., Bryson, J.J., 2020. Polarization under rising inequality and economic decline. Science advances 6, eabd4201.

Tabellini, G., Manacorda, M., Tesei, A., 2023. Mobile Internet and the Rise of Communitarian Politics. CEPR Discussion Paper No. 18063 https://cepr.org/publications/dp18063.

Tabellini, M., 2020. Gifts of the immigrants, woes of the natives: Lessons from the age of mass migration. The Review of Economic Studies 87, 454–486.

Taylor, E., 1984. Egyptian Migration and Peasant Wives. MERIP Reports , 3–10doi:10.2307/3011611. publisher: Middle East Research and Information Project (MERIP).

Telegeography, 2021. Submarine cable frequently asked questions, submarine cable 101. URL: https://www2.telegeography.com/submarine-cable-faqs-frequently-asked-questions.

Tertytchnaya, K., Lankina, T., 2020. Electoral protests and political attitudes under electoral authoritarianism. The Journal of Politics 82, 285–299. doi:10.1086/705815.

Thomas, D., 1990. Intra-household resource allocation: An inferential approach. Journal of human resources , 635–664.

Tolbert, C.J., McNeal, R.S., 2003. Unraveling the effects of the internet on political participation? Political research quarterly 56, 175–185. doi:10.2307/3219896.

Tomberg, L., Stegen, K.S., Vance, C., 2021. "the mother of all political problems"? on asylum seekers and elections. European Journal of Political Economy 67, 101981. doi:https://doi.org/10.1016/j.ejpoleco.2020.101981.

Tormos, R., 2019. Measuring personal economic hardship and its impact on political trust during the great recession. Social Indicators Research 144, 1209–1232.

Tsourapas, G., 2022. Migration and development in egypt—a holistic view. International Development Policy| Revue internationale de politique de développement doi:10.4000/poldev.4803.

Tuccio, M., Wahba, J., 2018. Return migration and the transfer of gender norms: evidence from the middle east. Journal of Comparative Economics 46, 1006–1029. doi:10.1016/j.jce.2018.07.015.

Tuccio, M., Wahba, J., 2020. Social Remittances, in: Zimmermann, K.F. (Ed.), Handbook of Labor, Human Resources and Population Economics. Springer International Publishing, Cham, pp. 1–13. doi:10.1007/978-3-319-57365-6_112-1.

Tuccio, M., Wahba, J., Hamdouch, B., 2019. International migration as a driver of political and social change: evidence from morocco. Journal of Population Economics 32, 1171–1203. doi:10.1007/s00148-019-00734-9.

Tufekci, Z., 2017. Twitter and tear gas: The power and fragility of networked protest. Yale University Press. doi:10.25969/mediarep/14848.

UN, 2020. United Nations Department of Economic and Social Affairs, Population Division, International Migrant Stock 2020. Technical Report POP/DB/MIG/Stock/Rev.2020. URL: https://www.un.org/development/desa/pd/content/international-migrant-stock.

UN Women, 2017. Understanding masculinities, results from the international men and gender equality study in the middle east and north africa. United Nations .

Vishwanath, T., 2012. Opening doors: gender equality in the middle east and north africa. World Bank, Washington, DC doi:hdl.handle.net/10986/10844.

Vosoughi, S., Roy, D., Aral, S., 2018. The spread of true and false news online. Science 359, 1146–1151. doi:10.1126/science.aap9559.

Wahba, J., 2015a. Selection, selection, selection: the impact of return migration. Journal of Population Economics 28, 535–563. doi:10.1007/s00148-015-0541-4.

Wahba, J., 2015b. Through the Keyhole: International Migration in Egypt, in: Assaad, R., Krafft, C. (Eds.), The Egyptian Labor Market in an Era of Revolution. Oxford University Press, p. 198–217. doi:10.1093/acprof:oso/9780198737254.003.0011.

Wahba, J., 2021. Who benefits from return migration to developing countries? IZA World of Labor .

Wallace, S.J., Zepeda-Millán, C., Jones-Correa, M., 2014. Spatial and temporal proximity: Examining the effects of protests on political attitudes. American Journal of Political Science 58, 433–448. doi:10.1111/ajps.12060.

Weller, D., Woodcock, B., 2013. Internet traffic exchange: Market developments and policy challenges. OECD Digital Economy Papers, No. 207. doi:10.1787/5k918gpt130q-en.

World Bank, 2018. World development indicators, individuals using the internet (% of population). URL: https://data.worldbank.org/indicator/IT.NET.USER.ZS?name_desc=false.

World Economic Forum, 2022. Global gender gap report 2022. URL: https://www3.weforum.org/docs/WEF_GGGR_2022.pdf. retrieved from the World Economic Forum website.

Wroe, A., 2016. Economic insecurity and political trust in the united states. American Politics Research 44, 131–163.

Yang, D., 2008. International migration, remittances and household investment: Evidence from philippine migrants' exchange rate shocks. The Economic Journal 118, 591–630. doi:https://doi.org/10.1111/j.1468-0297.2008.02134.x.

Zhuravskaya, E., Petrova, M., Enikolopov, R., 2020. Political Effects of the Internet and Social Media. Annual Review of Economics 12, 415–438. doi:10.1146/annurev-economics-081919-050239.

Zohry, A., 2003. The place of egypt in the regional migration system as a receiving country. Revue européenne des migrations internationales 19, 129–149. doi:10.4000/remi.2664.

Zohry, A., 2013. Egypt's international migration after the revolution: Is there any change? Confluences Méditerranée , 47–54doi:10.3917/come.087.0047.

Zohry, A., 2014. 2.migration and development in egypt, in: Sievers, W., Bommes, M., Fassman, H. (Eds.), Past Developments, Current Status and Future Potentials. Amsterdam University Press, Amsterdam, pp. 75–98. doi:10.1515/9789048523177-005.