École Doctorale des Sciences Économiques, Juridiques, Politiques et de Gestion
Centre d'Études et de Recherches sur le Développement International (CERDI)
Université Clermont Auvergne, CNRS, IRD, CERDI, F-63000 Clermont-Ferrand, France

Three Essays on Culture and Spatial Mobility in Development Economics

Trois essais sur la culture et la mobilité spatiale en économie du développement

Par
Melchior Clerc

Sous la direction de
Simone Bertoli & Èric Roca Fernández

Soutenance prévue le
3 septembre 2026

Membres du jury

Catherine Guirkinger	Professeure, Université de Namur, CRED	Rapporteuse
Karine Marazyan	Professeure, Université de Rouen Normandie, LERN	Rapporteuse
Marion Mercier	Chargée de recherche CNRS, LEDa-DIAL	Rapporteuse
Simone Bertoli	Professeur, Université Clermont Auvergne, CERDI	Directeur de thèse
Èric Roca Fernández	Professeur assistant, Université Clermont Auvergne, CERDI	Directeur de thèse
Michaël Goujon	Professeur, Université Clermont Auvergne, CERDI	Suffragant

L'Université Clermont Auvergne n'entend donner aucune approbation ni improbation aux opinions émises dans cette thèse. Ces opinions doivent être considérées comme propres à leur auteur.

Remerciements

Mes remerciements vont en premier lieu à mes deux directeurs de thèse. Simone, je mets à votre compte la majeure partie de la responsabilité de m'être embarqué dans cette drôle d'aventure qu'est le doctorat. C'est en 2020, en suivant votre cours de Master « Migration et Développement », que j'ai pour la première fois été piqué par la gymnastique intellectuelle que demande la recherche en économie. J'ai ensuite eu la chance de continuer à bénéficier de vos retours fréquents pendant ce doctorat. Vous vous êtes toujours rendu disponible quand j'allais taper à la porte de votre bureau, et vos retours à mes mails étaient parfois si rapides que je pensais à une réponse automatique. Plus largement, votre rigueur scientifique et votre sens du détail sont aussi exigeants pour le doctorant que précieux pour le chercheur. Au moment où je transitionne de l'un vers l'autre, j'espère en emporter une partie avec moi.

Èric, je te remercie pour ta disponibilité et ton temps pour des échanges et conseils de toute nature : sur mon travail, bien sûr, mais aussi sur la recherche en général, et surtout sur les codes Stata pour sortir de beaux tableaux. Par ailleurs, commencer la thèse en coécrivant avec vous deux ainsi qu'avec Jordan Loper a été très vite formateur, et je souhaitais vous remercier tous les trois pour ces premières expériences de recherche.

Je remercie chaleureusement les membres du jury, Michaël Goujon, Catherine Guirkinger, Karine Marazyan et Marion Mercier, pour avoir accepté d'évaluer ce travail. J'estime beaucoup votre expertise dans les thématiques abordées dans cette thèse. Ce rôle requiert un temps précieux, je suis donc d'autant plus honoré de pouvoir bénéficier de vos retours.

Je suis également très reconnaissant de l'environnement dans lequel j'ai pu évoluer durant cette thèse. L'ambiance chaleureuse et stimulante du CERDI, que ce soit dans les salles de séminaire ou en dehors, m'a vraiment fait me sentir chez moi, et la vue depuis le 4ème me manquera. Un grand merci aux équipes administratives, notamment Chantal et Hassen avec qui les discussions au détour d'un couloir étaient toujours un plaisir.

J'ai bien sûr une pensée particulière pour les doctorants et post-doctorants dont j'ai croisé la route durant ces quatre années. C'est avec eux que j'ai mené au quotidien mon travail de thèse, et je me sens très chanceux qu'il ait toujours pris place dans le rire et la bonne humeur. L'envie de venir au labo malgré un temps parfois très clermontois a souvent tenu à leur compagnie. Adrien, I missed the double or fun as soon as you left, nos savants mélanges de bêtises et d'intenses brainstormings avaient une saveur unique. Aubin, merci pour ces folles folles (folles) soirées, elles ont mis des paillettes dans le quotidien et des sous dans les caisses de l'Other. Pierre, thanks for the jokes when the brain was not braining anymore, and for the tips for my one-handed backhand: all in all, great success. Paul, merci pour tes rires, à mon avis trop généreux, face à mes blagues nulles, et pour être toujours d'attaque pour tout type de plan tant que ça n'implique pas des jeux de société. Eliot, merci d'avoir commencé cette aventure avec moi, et d'avoir montré que suivre sa boussole amène toujours en eaux sûres. Et surtout, merci d'avoir accueilli le fou du bus chez toi l'espace d'une soirée.

J'ai aussi une pensée pour Alfred, Alima, Alistair, Anouck, Arouna, Chanrithy, Frans, Gracia, Habibou, Ibrahim, Idrissa, Ilass, Marine, Marize, Miray, Peter, Rachid, Ulrich, Vincent K., Vincent N. et Yannick. Merci à tous pour votre présence. Plus largement, je remercie tous mes anciens camarades de Magistère d'avoir fait changer d'avis un Parisien sur Clermont-Ferrand.

I also wanted to thank the great people I had the chance to meet during my visit to Gothenburg. First, many thanks to Annika for such a warm welcome. You did not know me at all when I sent you an email asking for a visiting opportunity, yet you immediately made me feel good in the department. Thank you also for your time and for your very valuable feedback on my JMP, together with Ia and Heather; it has been really helpful. Thanks also to all the members of the department who took the time to share their thoughts on my work, especially Joe. Thanks to the organizers of the ASWEDE Annual Conference for the invitation, and for selecting a place with a sauna.

Many thanks to all the PhD students in the department. You have been very welcoming, and I feel lucky to have ended up in your company. In particular, I wanted to thank Cherylynn, Daniel, Eliott, Gustav, Juanita, Stefano, Yan and Yannick. Props to Paula for sharing my office as a fellow foreigner. We managed to organize an after-work I will not forget even though we were clearly living in different time zones. Also, big thanks to team Ullahuset for making this experience far from Blermond-Clerzand so nice, and to Facundo and Roosta for introducing me to so many crazy places.

Un grand merci à mes colocs de Philippe Lebon et des Chaussetiers pour avoir rempli le quotidien de bêtises, de bons plats et de films du dimanche soir. En particulier Petit Pol, Lola, Louise, Florenfe, Rafiki et Soldat Michaud. Couper avec la thèse est, je pense, vital pour ne pas perdre la tête dans le processus, et rentrer tous les soirs auprès d'êtres humains un peu plus normaux qu'un doctorant y a bien contribué. Désolé d'avoir, sur la fin, fonctionné sur un fuseau horaire parallèle, et ramené un peu de force ce quatrième colocataire qu'a été le manuscrit (il reviendra plus, promis). Une pensée particulière pour Alice, avec qui une discussion dans la cuisine s'est transformée en idée de JMP. Merci de m'avoir raconté ta journée en rentrant un soir de semaine.

Merci à mes gens sûrs de Clermont, Toutoune pour sa bonne humeur, DJ Alexis pour accepter de me donner l'heure avant le succès mondial, William pour avoir gardé le sourire de l'autre côté du court en me mettant trempe sur trempe, Aurelle pour ton regard toujours bienveillant, Antoine pour les bagarres en bord d'Aydat, et Geoffrey pour les bières un peu partout dans la ville.

Merci à l'équipe du Clermont Boxe pour m'avoir permis de taper (toujours avec grand respect et bienveillance) sur mon prochain quand la thèse me tapait sur le système, à Lynda et l'équipe du SSU pour l'accompagnement fidèle, et à la boucherie Durif pour les sandwichs au bœuf mariné.

J'ai une pensée toute particulière pour la famille de cœur, ceux qui ont toujours été une bouffée de positif durant ces dernières années. Merci à eux de ne jamais avoir pris personnellement mes délais de réponses dignes de l'époque des pigeons voyageurs. Merci aux Entêtés, Solal, Nans et Vico, avec qui l'alchimie prend toujours quand on se retrouve dans la même pièce, et à Banshee pour les big moments dans l'atelier du dernier étage. Merci à Mathias, pour ta volonté de suivre des chemins un peu sinueux, un peu différents, que j'ai toujours trouvé inspirante, et à Eléonore, qui a elle carrément suivi un chemin qui l'a emmené beaucoup trop loin, mais avec qui paradoxalement rien ne bouge. Merci le chef, aka Dorian, d'être toujours prêt à raconter n'importe quoi avec le plus grand sérieux, et merci Alexandre d'assister à tout ça depuis si longtemps en essayant toujours de trouver une cohérence chez ton pote. Merci à Ma2t pour toutes les bonnes soirées et les discussions qui en ont suivi, merci de rester juste un chouia plus petit que moi, et merci pour le buisson de Karen. Merci Benjam d'être complètement fou, change rien c'est précieux. J'ai grand hâte de ce qui est à venir avec vous tous.

Enfin, merci à ma famille dans son ensemble. J'ai une pensée particulière pour mes couz et mes zines. Solal, Myrto et Joséphine, je sais que vous avez grandi dans mon dos pendant que j'étais loin, mais on va rattraper ça bien assez tôt, promis. Ethan, j'ai hâte de voir ce que tu nous prépares musicalement, de ton dernier devoir maison de *shatta* à tes créations plus personnelles. Rachel, t'es une boss d'être passée me voir de l'Auvergne au grand Nord, merci d'être une présence avec qui on peut parler des choses qui tiennent vraiment à cœur. Étienne, le zinc, merci pour les discussions sur le good job et les angles morts, mais aussi pour ton ouverture et ton soutien sur les chantiers qu'on a en commun. Merci de ramener de la folie dans ma playlist, merci pour les bons plans vers Fort d'Aubervilliers, et en vrai merci d'être un amour tout le temps. Hugo, t'es relou d'avoir été docteur en premier parmi les petits-enfants, mais ça va que j'oublie vite dès qu'on nous met une raquette dans les mains. Merci de tracer la voie en montrant qu'on peut être un brillant chercheur tout en faisant la mariole 90% du reste de son temps de vie. Bryan, merci de me montrer qu'on peut rester drôle, beau gosse et bienveillant malgré son très grand âge, et pour me soutenir sans relâche la fake news selon laquelle mon front s'agrandit.

Merci Pip, pour avoir toujours cru en moi, et pour ta fierté envers ton fiston qui transpire même quand tu veux pas trop le montrer. Merci pour nos discussions sur le défi qu'est la thèse, mais aussi sur tout le reste. Merci Mumu, pour toujours avoir été un inébranlable point de rattachement, pour ta manière de traverser les épreuves, qui inspire forcément à l'heure d'en affronter, et pour ta présence souvent silencieuse mais toujours soutenante.

Merci Caroline, pour ne jamais, jamais t'être arrêtée de me supporter pendant cette aventure, surtout quand j'étais insupportable. Rire de tout (et surtout de rien) avec toi a aidé plus que tu ne peux l'imaginer, et ta patience exemplaire dans les moments où j'étais sous l'eau te rend officiellement éligible au prochain poste de Dalaï-Lama. Merci pour ça, et pour tout le reste.

Pour finir, merci à tous ceux que je n'ai pas pu citer, mais qui ont été d'une façon ou d'une autre à mes côtés durant ces quatre années et qui, même sans comprendre ce sur quoi je me cassais la tête, sont toujours restés soutenants. Ces remerciements longs comme une biblio soulignent à quelle point cette thèse n'aurait pas été la même, ou même pas été tout court,

sans ceux que j'ai eu la chance d'avoir autour.

À Roger, qui voulait que je sois docteur en physique quantique alors que j'y comprenais déjà plus rien en 1ère S, mais qui, j'espère, n'aurait pas trop rougi devant cette thèse. Qu'elle s'intéresse à la migration est peut-être un juste retour des choses.

Résumé

Cette thèse est composée de trois essais empiriques en économie du développement, et porte sur les liens entre culture, mobilité spatiale et normes sociales. Son fil directeur est double. Elle interroge d'une part la manière dont les chercheurs mesurent la culture et rattachent les individus à une origine culturelle lorsque la mobilité rompt le lien entre lieu de résidence et origine ancestrale. Elle analyse d'autre part la mobilité comme outil d'identification et comme réponse aux contraintes de normes sociales locales.

Le premier chapitre propose une réplication approfondie de l'article de Giuliano and Nunn (2021), qui relie la variabilité climatique ancestrale à une moindre importance accordée aujourd'hui à la tradition. Il compare les spécifications décrites dans l'article avec les codes et données de réplication disponibles, et met en évidence plusieurs écarts portant sur les variables, les échantillons et les spécifications estimées. La correction de ces écarts, lorsqu'elle est possible, laisse apparaître une relation entre variabilité climatique ancestrale et importance de la tradition plus nuancée que dans les estimations initiales. Le chapitre met ainsi en évidence un angle mort de la reproductibilité : un code peut reproduire les tables publiées sans correspondre aux spécifications décrites dans le texte.

Le deuxième chapitre étudie l'approche épidémiologique, qui compare des migrants ou descendants de migrants vivant dans un même pays de destination, mais issus d'origines différentes. Le chapitre montre que, dans le cas des États-Unis, l'ascendance étrangère déclarée reflète une origine ancestrale, mais aussi le temps écoulé depuis la migration des ancêtres et la sélection des individus qui continuent à déclarer cette ascendance. Certains résultats attribués à la persistance culturelle peuvent alors refléter des différences de temporalité migratoire ou de sélection.

Le troisième chapitre analyse la mobilité comme réponse concrète à une norme sociale contraignante dans le cas des mutilations génitales féminines en Afrique subsaharienne. Lorsque l'excision est fortement ancrée, des parents opposés à la pratique peuvent subir une pression sociale importante s'ils refusent d'exciser leur fille. À partir des *Enquêtes Démographiques et de Santé* (EDS) menées dans treize pays d'Afrique subsaharienne, le chapitre examine si le confiage des filles dans un autre ménage peut constituer une stratégie de protection face à cela. Les résultats montrent que, dans les groupes à forte prévalence, les filles sont plus souvent confiées que les garçons, surtout lorsque leur mère déclare s'opposer à l'excision. Cette

stratégie est plus marquée lorsque les familles disposent d'options de sortie crédibles. Les filles confiées sont aussi moins susceptibles d'être excisées, et le recul de la pratique après l'interdiction sénégalaise de 1999 s'accompagne d'une baisse du confiage là où l'excision était initialement ancrée.

Ces trois chapitres montrent que l'étude empirique de la culture dépend des choix méthodologiques qui permettent de la mesurer. La mobilité apparaît comme une fragilité d'identification, une ressource méthodologique et une stratégie concrète pour modifier l'exposition aux normes.

Summary

This dissertation is composed of three empirical essays in development economics and focuses on the links between culture, spatial mobility, and social norms. Its guiding thread is twofold. First, it examines how researchers measure culture and assign individuals to a cultural origin when mobility weakens the link between place of residence and ancestral origin. Second, it analyzes mobility both as an identification tool and as a possible response to the constraints imposed by local social norms.

The first chapter provides an in-depth replication of Giuliano and Nunn (2021), who relate ancestral climatic variability to a lower importance attached today to tradition. It compares the specifications described in the article with the available replication codes and data, and identifies several discrepancies related to variables, samples, and estimated specifications. Correcting these discrepancies, when possible, reveals a more nuanced relationship between ancestral climatic variability and the importance of tradition than in the original estimates. The chapter thus highlights a blind spot in reproducibility: code can reproduce published tables without corresponding to the specifications described in the text.

The second chapter studies the epidemiological approach, which compares migrants or descendants of migrants living in the same destination country but originating from different places. The chapter shows that, in the case of the United States, self-reported foreign ancestry reflects an ancestral origin, but also the time elapsed since ancestors migrated and the selection of individuals who still report that ancestry. Some results attributed to cultural persistence may therefore reflect differences in migration timing or selection.

The third chapter analyzes mobility as a concrete response to a binding social norm, in the case of female genital cutting in sub-Saharan Africa. When cutting is deeply entrenched, parents who oppose the practice may face substantial social pressure if they refuse to cut their daughters. Using the *Demographic and Health Surveys* (DHS) conducted in thirteen sub-Saharan African countries, the chapter examines whether fostering girls into another household can serve as a protective strategy. The results show that, in high-prevalence groups, girls are fostered more often than boys, especially when their mother reports opposing cutting. This strategy is more pronounced when families have credible exit options. Fostered girls are also less likely to be cut, and the decline in the practice after the 1999 Senegalese ban is accompanied by a decrease in fostering where cutting was initially entrenched.

Taken together, the three chapters show that the empirical study of culture depends on the methodological choices through which culture is measured. Mobility appears as a source of identification fragility, a methodological resource, and a concrete strategy for changing exposure to norms.

Contents

List of Acronyms

Acronym	Meaning
ACS	American Community Survey
AER	American Economic Review
CPS	Current Population Survey
DHS	Demographic and Health Surveys
ESS	European Social Survey
EVS	European Values Study
FE	Fixed effects
FGC	Female genital cutting
GDP	Gross domestic product
GN	Giuliano and Nunn (2021)
GSS	General Social Survey
HHI	Herfindahl–Hirschman Index
I4R	Institute for Replication
IPUMS	Integrated Public Use Microdata Series
IPV	Intimate partner violence
JDE	Journal of Development Economics
JEL	Journal of Economic Literature
MSA	Metropolitan Statistical Area
NGO	Non-governmental organization
OFPRA	Office Français de Protection des Réfugiés et Apatrides
PUMA	Public Use Microdata Area
QJE	Quarterly Journal of Economics
SSA	Sub-Saharan Africa
TFR	Total fertility rate
UNICEF	United Nations Children's Fund
US/USA	United States / United States of America
WVS	World Values Survey

List of Tables

List of Figures

Introduction Générale

Les débats politiques récents en France, comme dans le reste de l'Europe et aux États-Unis, accordent une place croissante aux questions de culture, d'identité et d'origine. Ces notions sont fréquemment mobilisées dans l'espace public pour expliquer les comportements individuels, les trajectoires d'intégration ou les tensions sociales. Pourtant, elles restent souvent définies de manière imprécise et sont associées, plus ou moins implicitement, à des caractéristiques supposées des individus ou des groupes.

Cette difficulté n'est pas propre au débat public. Elle traverse en réalité tout travail empirique qui cherche à étudier le rôle de la culture dans les comportements économiques : comment définir ce qui relève de la culture ou non, comment la mesurer, et comment rattacher des individus à des normes, croyances ou préférences héritées ? Ces questions deviennent particulièrement saillantes dès lors que les individus se déplacent. La mobilité rompt en effet l'association souvent implicite entre un individu, un lieu et un groupe d'origine : le lieu où l'on vit, le lieu d'où l'on vient, le groupe auquel on se déclare appartenir, ou encore le contexte dans lequel on a été socialisé peuvent alors ne plus coïncider. Cela fait des interrelations entre culture et mobilité spatiale un objet d'étude particulièrement riche, que cette thèse aborde selon trois perspectives complémentaires : l'analyse de la mobilité comme un obstacle à la mesure des normes et préférences héritées, la réflexion critique autour d'une approche qui voit la migration comme une opportunité pour identifier empiriquement l'effet de la culture, et la mise en avant de la mobilité comme réponse concrète à des normes sociales contraignantes.

Une première étape commune à ces trois perspectives est certainement de préciser ce que l'on entend par culture. Objet ancien des sciences sociales, elle demeure pourtant difficile à définir. En économie, une définition influente est celle proposée par Guiso et al. (2006), qui met au cœur l'idée de transmission intergénérationnelle en définissant la culture comme "those customary beliefs and values that ethnic, religious, and social groups transmit fairly unchanged from generation to generation". Cette définition est volontairement restrictive : elle ne prétend pas saisir l'ensemble de ce qui constitue la culture, mais se concentre sur ses dimensions qui sont à la fois mesurables et susceptibles d'influencer les comportements économiques. Dans une perspective complémentaire, Fernández (2011) propose d'appréhender la culture d'abord comme ce qui distingue, en considérant "differences in culture as systematic variation in beliefs and preferences across time, space, or social groups".

Ces définitions ont donné un cadre au développement de l'économie de la culture autour de trois questionnements complémentaires : les origines des normes, croyances et institutions informelles (Alesina et al., 2013; Lowes and Nunn, 2024; Giuliano and Nunn, 2021), leurs effets sur les comportements économiques contemporains (Tabellini, 2010; Algan and Cahuc, 2010; Luttmer and Singhal, 2011), et leur persistance dans le temps (Nunn and Wantchekon, 2011; Michalopoulos and Papaioannou, 2013; Anderson, 2018). L'ensemble de cette littérature repose sur une étape clef : rattacher chaque individu observé à une mesure des normes, préférences ou croyances qui capte le plus fidèlement possible l'environnement culturel auquel il a été exposé.

Cette opération est loin d'être neutre. Les sources de données disponibles pour approcher la culture sont multiples, mais toutes partagent le fait d'être imparfaites : l'enjeu n'est pas de mesurer la culture dans son ensemble, mais plutôt ce qu'elle véhicule comme préférences, croyances, normes ou institutions informelles pertinentes pour comprendre les comportements économiques. Dans un monde idéal, le chercheur observerait pour chaque individu les normes transmises au sein de la famille ou de la communauté, ainsi que les processus de socialisation auxquels il a été exposé. En pratique, ces informations sont rarement observables. Les travaux empiriques s'appuient donc sur des mesures indirectes, ou *proxys*, d'une ou plusieurs dimensions de la culture.

Ces proxys diffèrent par leur nature, leur niveau de construction et la temporalité de la culture qu'ils cherchent à mesurer. Une première famille repose sur des attitudes, croyances ou préférences déclarées dans des enquêtes, comme les *World Values Surveys* (Tabellini, 2010; Alesina et al., 2013), l'*Afrobarometer* (Nunn and Wantchekon, 2011) ou les *Demographic and Health Surveys* (Lindskog et al., 2024). Ces réponses peuvent être utilisées au niveau individuel, souvent comme variables expliquées, mais sont aussi fréquemment agrégées à l'échelle d'un pays, d'une région ou d'un groupe. Elles ont l'avantage de mesurer directement ce que les individus déclarent croire ou préférer, mais restent sujettes aux biais déclaratifs et peuvent refléter l'environnement contemporain dans lequel les individus répondent.

Une deuxième famille de proxys s'appuie sur des comportements agrégés au niveau d'un groupe, par exemple la fécondité ou la participation féminine au marché du travail dans le pays d'ascendance (Fernández and Fogli, 2009). Ces mesures ont l'avantage de reposer sur des comportements observés plutôt que sur des préférences déclarées. Elles sont cependant le fruit d'équilibres contemporains qui traduisent certes des préférences culturelles, mais aussi des contextes institutionnels, politiques et économiques, rendant difficile l'isolement de la composante culturelle d'intérêt.

Une troisième famille repose enfin sur des traits directement codés au niveau de groupes historiques ou ethnographiques, comme les modes de subsistance, les systèmes de parenté ou l'organisation politique préindustrielle dans l'*Ethnographic Atlas* (Murdock, 1967), ou les classifications linguistiques dans l'*Ethnologue* (Eberhard et al., 2023) ou le *Glottolog* (Hammarström et al., 2023). Ces sources permettent de documenter des caractéristiques plus anciennes, moins directement liées aux conditions économiques contemporaines, mais elles supposent d'apparier des groupes historiques ou ethnographiques à des populations contemporaines.

Le point commun de ces mesures est qu'elles nécessitent toutes une opération centrale sur le plan méthodologique : assigner les individus à une unité collective supposée porter une information culturelle pertinente.[1] Cette unité peut être géographique, administrative, linguistique ou déclarative. La règle d'assignation peut ainsi reposer sur le pays d'origine (Fernández and Fogli, 2009; Luttmer and Singhal, 2011), le pays d'ascendance (Putterman and Weil, 2010), ou la région de résidence (Tabellini, 2010). Ces variables sont généralement bien observées et comparables, mais elles peuvent conduire à confondre culture, institutions, conditions économiques et composition locale. Une autre stratégie consiste à utiliser la langue, souvent plus facilement comparable et appariable à des sources historiques ou ethnographiques. La langue constitue un vecteur important de socialisation, mais elle ne recouvre pas nécessairement les frontières pertinentes des normes sociales, notamment dans le cas de langues très largement partagées. Une dernière stratégie reçoit une attention particulière depuis plusieurs années : l'appartenance déclarée, comme l'ethnicité (Michalopoulos et al., 2019; Alesina et al., 2013) ou l'ascendance ancestrale (Giuliano and Nunn, 2021). Ces variables peuvent mieux refléter certaines formes de transmission familiale ou de groupe, mais elles dépendent des catégories proposées par les enquêtes et de la manière dont les individus se définissent.

La première perspective développée dans cette thèse porte précisément sur cette sensibilité des résultats empiriques aux choix de construction des données et d'assignation des individus à une origine. Le premier chapitre propose une reproduction approfondie de Giuliano and

[1] Par construction, cela n'est pas le cas avec les mesures individuelles, mais, dans les faits, celles-ci sont pratiquement toujours agrégées hormis lorsqu'elles servent de variable expliquée.

Nunn (2021), article influent qui étudie le lien entre variabilité climatique ancestrale et persistance culturelle. L'article original avance l'hypothèse selon laquelle les sociétés historiquement exposées à une plus forte variabilité climatique accordent aujourd'hui moins d'importance à la tradition. Le chapitre compare systématiquement les descriptions des spécifications empiriques dans le texte de l'article avec les codes et les données de réplication disponibles. Cette comparaison met en évidence plusieurs écarts portant sur la construction des variables, la définition des échantillons et les spécifications effectivement estimées. Lorsque ces écarts peuvent être corrigés à partir des données disponibles, les résultats deviennent plus nuancés : la relation entre variabilité climatique ancestrale et importance de la tradition apparaît moins robuste qu'elle ne le semblait dans les estimations initiales. La contribution du chapitre est double. D'une part, il participe à la littérature sur la reproductibilité en économie en mettant en évidence un angle mort dans l'évaluation des exercices empiriques : la cohérence entre le texte et le code. En effet, le texte d'un article est examiné par des rapporteurs, tandis que la vérification du code relève souvent d'un processus distinct, généralement mené par le *Data Editor*, et centré sur la capacité du code à reproduire les résultats présentés. Un code peut ainsi tourner correctement et reproduire les tables de l'article (et donc être validé), tout en étant incohérent avec les spécifications décrites dans le texte. D'autre part, il montre que les résultats portant sur la persistance culturelle peuvent être particulièrement sensibles aux choix de construction des données, d'assignation des individus à une origine, et de définition des groupes de comparaison.

La deuxième perspective concerne une stratégie empirique qui transforme la migration en ressource méthodologique. Comparer simplement des individus vivant dans différents pays rend difficilement séparables l'effet des normes culturelles de celui du contexte géographique, institutionnel et économique. En d'autres termes, les différences observées entre deux pays distincts reflètent-elles des différences de pratiques culturelles, ou simplement des contextes contemporains différents ? L'approche épidémiologique (Fernández, 2011) propose de contourner cette difficulté en comparant des migrants ou des descendants de migrants vivant dans un même pays de destination, mais issus d'origines différentes. L'intuition est de maintenir constant, autant que possible, l'environnement institutionnel, économique et géographique contemporain, tout en faisant varier l'origine culturelle. Cette approche repose toutefois sur une condition forte : être capable d'assigner chaque individu à une origine pertinente.

Le deuxième chapitre examine cette condition lorsque l'origine est mesurée par l'ascendance étrangère déclarée dans les recensements américains[2]. Il montre que l'ascendance déclarée ne capture pas seulement une origine ancestrale : elle reflète aussi la sélection des individus qui déclarent encore une ascendance étrangère, ainsi que le temps écoulé depuis la migration de leurs ancêtres. À partir des recensements américains, le chapitre construit plusieurs proxys de cette hétérogénéité, portant notamment sur la présence de migrants de première génération, la concentration spatiale des groupes d'ascendance et l'usage de la langue ancestrale. Ces proxys permettent de documenter les différences entre groupes d'ascendance, puis d'évaluer la sensibilité de résultats existants dans la littérature. Le chapitre montre ainsi que certains résultats attribués à la persistance culturelle peuvent en partie refléter des différences de temporalité migratoire ou de sélection dans la déclaration d'ascendance. Sa contribution principale est de clarifier les conditions d'utilisation de l'approche épidémiologique lorsque l'origine est mesurée par l'ascendance déclarée, et de proposer des diagnostics simples permettant d'évaluer la robustesse des résultats à ces menaces d'identification.

La troisième perspective déplace encore le regard. La mobilité n'est plus seulement envisagée comme un problème de mesure ou comme une ressource méthodologique pour l'étude de la culture, mais comme une réponse possible aux normes sociales elles-mêmes. Dans cette thèse, la mobilité est entendue au sens large comme un déplacement qui modifie le lien entre individus, lieux et groupes d'appartenance. Elle peut désigner la migration internationale des individus ou de leurs ancêtres, centrale dans l'approche épidémiologique, mais aussi des formes plus locales de mobilité, comme la migration interne ou le placement d'un enfant dans un autre ménage. Le troisième chapitre s'ancre dans le contexte des mutilations génitales féminines, une pratique dont la persistance dépend à la fois des préférences des parents et des attentes de la communauté. Des parents opposés à l'excision peuvent rester soumis à une pression sociale forte s'ils vivent dans un environnement où la pratique demeure largement attendue. Le chapitre examine si ces parents peuvent réduire l'exposition de leurs filles à cette pression en les plaçant dans un autre ménage. À partir des *Enquêtes Démographiques et de Santé* (EDS) menées dans treize pays d'Afrique subsaharienne, des mesures de prévalence de l'excision au niveau de groupes ethniques, régionaux et générationnels sont construites puis comparées aux proportions de filles confiées par rapport aux garçons. Les résultats montrent que, dans les groupes à forte prévalence, les filles sont plus souvent confiées que les garçons, en

[2]Ce critère de rattachement est depuis plusieurs années le plus populaire dans la littérature, notamment en raison de sa disponibilité.

particulier lorsque leur mère déclare s'opposer à l'excision. Cette stratégie semble dépendre de l'existence d'options de sortie crédibles : elle est plus marquée lorsque les groupes présentent une forte variation géographique de la prévalence de l'excision, ou lorsque les parents appartiennent à des groupes caractérisés par des niveaux différents de prévalence. Le chapitre montre également que les filles confiées sont moins susceptibles d'être excisées, et que le recul de la pratique après l'interdiction sénégalaise de 1999 s'accompagne d'une baisse du confiage des filles dans les groupes où l'excision était initialement la plus ancrée. Il contribue ainsi à la littérature sur les normes sociales et les pratiques néfastes en montrant que les parents opposés à une norme contraignante peuvent recourir à une stratégie d'*exit* par la mobilité spatiale, le confiage des enfants étant mobilisé comme un moyen concret de réduire l'exposition de leurs filles à cette norme.

Les trois chapitres de la thèse montrent que l'étude empirique de la culture dépend étroitement des choix méthodologiques qui permettent de la mesurer, et qu'elle doit être pensée conjointement à la mobilité des individus. Cette mobilité apparaît successivement comme une source de fragilité pour l'identification empirique, une ressource méthodologique dont il faut discuter les conditions de validité, et une stratégie concrète par laquelle les individus peuvent modifier leur exposition aux normes sociales. Cette thèse vise ainsi à contribuer à une compréhension plus dynamique des liens entre culture, origine et mobilité.

Les fichiers de réplication associés aux deux premiers chapitres sont disponibles en ligne.

Bibliography

Alesina, A., P. Giuliano, and N. Nunn (2013): "On the origins of gender roles: Women and the plough," *The Quarterly Journal of Economics*, 128, 469–530.

Algan, Y. and P. Cahuc (2010): "Inherited trust and growth," *American Economic Review*, 100, 2060–2092.

Anderson, S. (2018): "Legal origins and female HIV," *American Economic Review*, 108, 1407–1439.

Eberhard, D. M., G. F. Simons, and C. D. Fennig, eds. (2023): *Ethnologue: Languages of the World*, Dallas: SIL International, 26 ed.

Fernández, R. (2011): "Does culture matter?" in *Handbook of Social Economics*, ed. by J. Benhabib, A. Bisin, and M. O. Jackson, Elsevier, vol. 1A, 481–510.

Fernández, R. and A. Fogli (2009): "Culture: An empirical investigation of beliefs, work, and fertility," *American Economic Journal: Macroeconomics*, 1, 146–177.

Giuliano, P. and N. Nunn (2021): "Understanding cultural persistence and change," *The Review of Economic Studies*, 88, 1541–1581.

Guiso, L., P. Sapienza, and L. Zingales (2006): "Does culture affect economic outcomes?" *Journal of Economic Perspectives*, 20, 23–48.

Hammarström, H., R. Forkel, and M. Haspelmath (2023): "Bank, S. glottolog/glottolog: Glottolog database 4.8," *Zenodo10*, 5281.

Lindskog, A., H. Congdon Fors, and A.-S. Isaksson (2024): "Harmful Norms: Social Interdependence and Persistence of Female Genital Cutting," *Available at SSRN 4834497.*

Lowes, S. and N. Nunn (2024): "The slave trade and the origins of matrilineal kinship," *Philosophical Transactions of the Royal Society B: Biological Sciences*, 379, 20230032.

Luttmer, E. F. P. and M. Singhal (2011): "Culture, context, and the taste for redistribution," *American Economic Journal: Economic Policy*, 3, 157–179.

Michalopoulos, S. and E. Papaioannou (2013): "Pre-colonial ethnic institutions and contemporary African development," *Econometrica*, 81, 113–152.

Michalopoulos, S., L. Putterman, and D. N. Weil (2019): "Ancestral characteristics of modern populations," *Economic Journal*, 129, 2364–2402.

Murdock, G. P. (1967): "Ethnographic Atlas: A Summary," *Ethnology*, 6, 109–236.

Nunn, N. and L. Wantchekon (2011): "The slave trade and the origins of mistrust in Africa," *American Economic Review*, 101, 3221–3252.

Putterman, L. and D. N. Weil (2010): "Post-1500 population flows and the long-run determinants of economic growth and inequality," *The Quarterly Journal of Economics*, 125, 1627–1682.

Tabellini, G. (2010): "Culture and institutions: Economic development in the regions of Europe," *Journal of the European Economic Association*, 8, 677–716.

Chapter 1

Understanding cultural persistence and change: a replication of Giuliano and Nunn (2021)

This chapter is joint work with Simone Bertoli, Jordan Loper, and Èric Roca Fernández, and published in *Economic Inquiry* (doi:10.1111/ecin.13242).

Replication files: doi:10.3886/E202861V3.

1.1 Introduction

Cultural persistence, the preservation of cultural traits over time, is a topic of profound importance in economics and social sciences. Research has demonstrated the enduring influence of culture on economic outcomes, such as female labor-force participation and fertility decisions. Moreover, harmful practices rooted in tradition may hinder social progress.

The study published in *The Review of Economic Studies* by Giuliano and Nunn (2021) (henceforth GN, the Authors, or the original article) has garnered significant attention in economics due to its novel insights into the determinants of cultural persistence. Drawing inspiration from evolutionary anthropology, GN propose that societies that have historically experienced greater climatic variability exhibit lower persistence of cultural traits. This prediction stems from the idea that intergenerational environmental changes reduce the value of knowledge embedded in tradition, rendering it less relevant to current conditions.

To test this hypothesis, GN employs a combination of theoretical modeling and econometric analyses using multiple data sources. The core of their empirical approach focuses on natives of immigrant origin residing in the United States, who provide a unique opportunity to study the effects of historical climatic variability on cultural persistence.[1] By comparing immigrants from different countries with varying climatic histories but living in the same location, GN can isolate the impact of climatic variability from other confounding factors related to geography. Additionally, the analysis incorporates data on Native Americans to address concerns about unobserved heterogeneity across different countries due to different patterns of self-selection into migration to the United States from different origin countries.

The findings of GN consistently support the evolutionary anthropology prediction. They show that greater ancestral intergenerational climatic variability experienced by one's ancestors is associated with lower levels of cultural persistence, as measured by various indicators of the importance of tradition and the adherence to cultural norms. Empirically, all the results presented in GN point to a consistent negative effect of the ancestral intergenerational climatic

[1]In this way, GN overcome the challenges arising from the correlation between historical and current climatic variability. The literature on the determinants of culture refers to this setup in which individuals from different origins residing in a single location are compared as the "epidemiological approach" (See Fernández, 2011). We refer the reader to a companion paper (Bertoli et al., 2023) which discusses the analytical challenges related to the use of self-reported ancestry to identify the foreign countries of origin of natives, as GN do in a portion of their work.

variability experienced by one's own ancestors on various stated and revealed measures of the importance of tradition as well as on the persistence of cultural norms. These results highlight the crucial role of past environmental factors in shaping the persistence of cultural traits.

GN's theoretical model and empirical findings have become keystones in economics, inspiring multiple articles to explore the implications of their insights. The original article has already been cited extensively, despite its recent publication, with Google Scholar reporting 364 citations and publisher metrics showing that it has received approximately 64 times more citations than the average publication in the same field (information retrieved on March 14, 2024). These citations include influential articles published in leading journals, for instance, Bau (2021), Bazzi et al. (2019, 2020, 2023), Bursztyn et al. (2024), Caprettini and Voth (2022), Enke (2019), Ghosh et al. (2023), Michalopoulos and Xue (2021) and Valencia Caicedo (2018).[2]

Paralleling its bibliometric impact, its scientific contribution is substantial. Indeed, economists are increasingly coming to understand the long-term impact of ancestral cultural traits on contemporary outcomes. In that respect, this literature provides numerous empirical pieces of evidence of the persistence of culture across generations. On the other hand, another vivid literature presents numerous examples of rapid cultural change, such as the Protestant Reformation (see the introduction in GN), or the abandonment of traditional values following the contact with Europeans, or the implementation of national policies (Bau, 2021). GN represents a pioneering effort to answer the central and natural question that arises: when does culture persist, and when does it change?

In this article, we present the results of our "deep reproduction" (Dreber and Johannesson, 2023)[3] of some of the econometric results reported by GN. As defined by Dreber and Johannesson (2023), "deep reproducibility studies, [dig] deeply into the original study including the raw data and the coding decisions to detect potential coding errors or questionable coding decisions."[4]

[2]The literature on persistence and rapid cultural change is vast and multifaceted. For a comprehensive overview, see Bisin and Federico (2021); in particular, Voth (2021) provides a thorough survey of the persistence literature in historical economics, while Acemoglu et al. (2021) offers a conceptual framework for understanding institutional persistence and change. Empirical studies have also shed light on the dynamics of cultural change, such as Alesina and Fuchs-Schündeln (2007), which examines the effect of communism on people's preferences, and Slotwinski and Stutzer (2022), which investigates the impact of religious doctrine on democratization.

[3]Last version accessed: 30 November 2023.

[4]We cite the earlier I4R Discussion Paper version of Dreber and Johannesson (2023), as requested by a referee, but the term "deep reproducibility" appears in a later version of the paper, posted on SSRN on May

This type of reproduction is desirable because, as Krähmer et al. (2023, p. 2) points out, "[o]pen code alleviates two problems that currently compromise the credibility of research: Errors in data preparation and misspecifications of statistical models. While model misspecification should, in principle, be discernible from the body of an article, errors in data preparation are downright impossible to spot without access to authors' code." This approach aligns with Clemens's (2017) definition of replication, where researchers use the "same" specification, population and sample as in the original article.[5] Indeed, within this framework, we conduct a replication test which "include: fixing coding errors so that the code does exactly what the original paper describes (verification)" (Clemens, 2017, p. 329). Moreover, as Huntington-Klein et al. (2021) argue, researcher degrees of freedom, such as decisions made during data collection, preparation, and analysis, can lead to variation in results, even when replicating the same study. Accordingly, we thoroughly verified the correspondence between the specifications described in the original article and the code and estimation sample provided in the replication package of GN. Importantly, this verification was not limited to the analysis code but also extended to the data cleaning and manipulation procedures, which were not included in the replication package. By closely examining the documentation and the data, we were able to reverse-engineer certain aspects of the data processing, even in the absence of access to the original code. For example, we identified instances where variables that should have been constant over time or across subgroups of observations exhibited variations, thus revealing that the actual data manipulation is inconsistent with the description provided by GN.

The replication exercise relies on three main elements: the original article by Giuliano and Nunn (Giuliano and Nunn, 2021), a *corrigendum* published in *The Review of Economic Studies* (Giuliano and Nunn, 2024), and the replication package of the original article, which can be accessed at `https://doi.org/10.5281/zenodo.4026093`. The latter includes the estimation samples used to estimate each equation, together with the corresponding codes.[6] However,

25, 2023, and last revised on November 30, 2023, available at `https://dx.doi.org/10.2139/ssrn.4458153`.

[5]The definition of "same" in the context of Clemens (2017) is crucial in our exercise: "[t]he 'same' specification, population, or sample means the same as reported in the original paper, not necessarily what was contained in the code and data used by the original paper. Thus for example if code used in the original paper contains an error such that it does not run exactly the regressions that the original paper said it does, new code that fixes the error is nevertheless using the 'same' specifications (as described in the paper)." (p. 328).

[6]We successfully verified the computational reproducibility of the results in GN for which the code is included.

it does not include the raw data, the code necessary to transition from the raw data to the various estimation samples, and the codes that generate the results reported in the online Appendix of GN.

The *corrigendum* was published after our exchanges with the Authors and with the editor in charge of the original article. The Authors write "[s]ince the publication of the article, imprecisions in wording and omissions in the text have been pointed out." (p. 597).[7] However, as will become apparent in the following discussion, the *corrigendum*: (*i*) deals with only a minor subset of the inconsistencies that we had signaled to the Authors and to the Editor in charge of the original article, (*ii*) does not describe the implications of these imprecisions for the econometric exercises conducted in the original article, and, in some instances, (*iii*) generates inconsistencies with other portions of the text of the original article that have not been modified. Throughout the following discussion, references to page numbers, equations, and tables correspond to the original article unless otherwise specified. Similarly, references to line numbers pertain to the file `GiulianoNunn_replication.do` included in the replication package.

1.1.1 Our reproduction approach

The characteristics of the replication package have shaped the scope of our replication exercise. Specifically, our focus lies on two key aspects: (*i*) the results reported in the main text and (*ii*) the estimation of all the six equations using individual-level data. Point (*i*) arises from the absence of the code to reproduce the online Appendix within the replication package, limiting our analysis to the results presented in the main text alone. Point (*ii*) stems from the absence of data cleaning and manipulation codes in the replication package, making it considerably more challenging to identify potential inconsistencies between the article's text and the excluded code that pertains to aggregate data.[8] Moreover, according to the Authors,

[7] Giuliano and Nunn (2024) acknowledges our contribution to this *corrigendum*.

[8] It should be noted, however, that there are instances where it is possible to infer inconsistencies between the original article's text and the analysis code, particularly when the aggregate data is directly linked to underlying individual-level data used in other equations. For example, coding errors in the generation of the dependent variable in Eq. (8) extend to Eq. (10), as the latter equation simply aggregates the individual-level data used to estimate the former (see Tables 8 and 10). Additionally, the final point of the *corrigendum* (Giuliano and Nunn, 2024) addresses an inconsistency between the original article's text and the corresponding analysis code using aggregate data.

regressions comparing "individuals with different cultural backgrounds but who live in the same city in the U.S." (Giuliano and Nunn, 2021, p. 1543) are superior.[9] For these reasons, (*iii*) our focus lies on the discrepancies between code and text (i.e., a deep reproducibility) rather than the replication of the estimates because the codes and data provided on the original replication package do not allow to fully do the latter.

Our replication provides detailed citations from the original article to document the precise and unambiguous descriptions of the specifications used in the six equations estimated with individual-level data. We then identify any inconsistencies between these textual descriptions and the actual code or estimation sample used.

We identified two main types of inconsistencies between the code and text. The first type concerns errors in the analysis code included in the replication package, such as deviations from the specified estimation sample or variables that were impossible to construct according to the original article. The second type relates to problems with the data cleaning and manipulation code, including issues with the estimation sample, variables with distributions at odds with the text, and missing or incorrectly constructed variables. These inconsistencies are discussed in more detail in the subsequent sections.

We have identified and documented inconsistencies within the six equations we analyzed, which correspond to four distinct tables and a total of 28 data columns.[10] Discrepancies greatly vary in terms of their empirical relevance (see Figures 1.1–1.3).

The remedies we propose to address the uncovered discrepancies are straightforward: doing so simply needs to adhere to the precise and unambiguous descriptions provided in the original article. However, it is not feasible to solve all the issues due to the lack of essential variables in the replication package.[11] As a result, we are unable to fully re-estimate the equations

[9] Regressions using indigenous population are similarly regarded. Eq. (2) uses individual-level data but only compares "individuals living in the same country but belonging to different ethnic groups." (Giuliano and Nunn, 2021, p. 1543)

[10] We refer the reader to Table 1.A.1 in the online Appendix for a summary of the identified inconsistencies, the corrections we propose and whether we implemented them or not. This Table also associates each inconsistency to the equation in which it appears. Table 1.A.2 indicates the Tables and Columns with errors in the original Appendix, together with the equations that these estimated.

[11] For instance, we mention that cohabiting parents and children could be identified using the variables `poploc` and `momloc`. These variables are not included in the replication package, and since individual identifiers have been removed, it is impossible to merge the data from the replication package with the original information from IPUMS.

in a manner consistent with the original article. Nonetheless, we want to emphasize that the exercise we present is a "deep reproduction" of the original article, and documenting these errors indicates that the original article cannot be replicated, regardless of any possible changes in the estimated coefficients.[12]

It is important to clarify that we do not take a position on whether the specification described in the original article or the one actually implemented in the code is more economically sound. Our thorough reproduction aims to compare the text of the original article, which has undergone peer review, with the corresponding code. The scope of our analysis is limited to this comparison, and it does not evaluate the merits of the possible alternative specifications. Furthermore, we would like to highlight that the published *corrigendum* (Giuliano and Nunn, 2024) has also been validated by *The Review of Economic Studies*, and we refer to its content below. Lastly, we want to clarify that, we never propose or run any specifications that deviate from those described in the paper.

The rest of the paper contains six short sections, each corresponding to one equation. These are followed by a section documenting the impact of the corrections on the estimates in GN, and by a concluding section.

1.2 The six equations using individual-level data

For each equation, we first identify and describe the coding discrepancies. We then compute the share of observations affected by each discrepancy in the estimation sample used in each data column,[13] and propose a method to realign the code with the original article's specifications. In our online Appendix, we document the incidence of each discrepancy and present the regression results obtained by correcting all correctable errors. For each data column except for those corresponding to Eq. (1.2),[14] we provide a figure illustrating the impact of each

[12]For the reader's interest, our online Appendix presents the results obtained when we simultaneously correct all feasible discrepancies. Additionally, Figures 1.1–1.3 illustrate the estimates obtained when addressing different combinations of these discrepancies. These figures serve to evaluate the impact of resolving various issues on the estimated coefficient of the variable of interest.

[13]Eq. (1.2) is estimated in the six data columns of Table 2; Eq. (1.4) is estimated in the four data columns of Table 4; Eq. (1.5) in Cols. (1)-(5) in Table 5; Eq. (1.6) in Cols. (6) and (7) in Table 5; Eq. (1.8) in Cols. (1)-(5) in Table 7, and Eq. (1.9) in Cols. (6) and (7) in Table 7.

[14]We do not provide corrections for Eq. (1.2), because the code that exactly follows the description given in the original article is impossible to estimate, as a variable is simply not available in the data, and because

identified discrepancy (and combinations thereof).

1.2.1 Eq. (2) Importance of tradition

GN write that, when estimating Eq. (1.2) using data from the World Values Survey (WVS):

> "After matching **respondents' self-reported mother tongue from the WVS** with ethnicity from the *Ethnographic Atlas*, we estimate:
>
> $$\text{Tradition}_{i,e,c} = \alpha_c + \beta_2 \text{Climatic Instability}_e + \boldsymbol{X}_i' \Phi + \boldsymbol{X}_e' \Omega + \epsilon_{i,e,c} \tag{1.2}$$
>
> where i denotes an individual who is a member of the historical ethnic group e and lives in country c. $\text{Tradition}_{i,e,c}$ is the self-reported importance of tradition, measured on an $1-6$ integer scale and increasing in the importance of tradition." (p. 1557, bold emphasis added).

Coding mistake

1. (Incidence: undetermined)[15] The WVS does not provide information on mother tongue. Instead, it only asks about the language that the respondent normally speaks at home.

The *corrigendum* reads: "On page 1557, 'mother tongue' should be 'language spoken at home'." (Giuliano and Nunn, 2024). We note that, for individuals of immigrant origin (i.e., foreign-born individuals and natives of foreign ancestry), the language spoken at home is a choice. Indeed, the Authors use this choice as a revealed measure of the attachment to tradition among natives of foreign ancestry in Eqs. (1.5) and (1.6).[16] Individuals of immigrant origin who are, according to the Authors, less attached to tradition are, therefore, more likely to be incorrectly assigned to the same ancestors as the natives of the country in which they reside. We refer the reader to Bertoli et al. (2024), where we propose an approach, departing from a

the *corrigendum* realigns the text with the corresponding code.

[15] The actual incidence depends on the (unobserved) share of each estimation sample for which the mother tongue is not the language normally spoken at home.

[16] "Since children born to immigrant parents in the U.S. are almost always fluent in English, they face the decision of whether to continue speaking their traditional language at home. They face this decision both when they live with their parents as children and when they live on their own with their own family. Thus, as a revealed measure of the importance of maintaining tradition, we examine the extent to which a foreign language is spoken at home among the children of immigrants. Speaking a foreign language at home indicates that the children of the immigrants were taught their origin language, which is a sign of the parents and children valuing their tradition." (p. 1565).

deep reproduction, to gauge the relevance of the coding mistake that we identify for Eq. (1.2) in GN, which cannot be directly corrected.

1.2.2 Eq. (4) Within-group marriage

> "[The] sample comprises all married women in all waves of the March Supplement of the Current Population Survey (CPS) [from 1994 to 2014] with at least one parent who was born outside the US. [...]
>
> $$\mathrm{I}_{i,c,k}^{IngroupMarriage} = \alpha_k + \beta_3 \text{Climatic Instability}_c + \boldsymbol{X}_c \Pi + \boldsymbol{X}_{c,k} \Omega + \boldsymbol{X}_i \Phi + \epsilon_{i,c,k} \tag{1.4}$$
>
> [...] The outcome of interest, $\mathrm{I}_{i,c,k}^{IngroupMarriage}$, is an indicator variable equal to one if **an individual's spouse was born in origin country** c **or if his or her mother or father was born in country** c. α_k denotes the inclusion of residence (i.e. metropolitan-area) fixed effects. The vector of country-level covariates $\boldsymbol{X}_c$ includes the natural log of the current per-capita GDP in the country of origin (measured in the survey year) [...] [$\boldsymbol{X}_{c,k}$] includes **the fraction of the population in the same metropolitan area as the individual who are first- or second-generation immigrants from the same country of origin.**" (p. 1562, bold emphasis added).

Coding mistakes

1. (Dependent variable, incidence between 2.0 and 2.8 percent) The dependent variable in Eq. (1.4) is created in a portion of the code not included in the replication package. Its definition requires information on the country of birth of the spouse and of the parents-in-law. The March Supplement of the CPS provides this information *exclusively* for individuals co-residing with their spouses,[17] but the estimation samples in Table 4 include married individuals *not co-residing* with the spouse, with the dependent variable for them being either equal to 0 or to 1. It is not possible to create the dependent variable in a way consistent with its description provided at p. 1562 for married individuals whose spouse is absent.

2. (Fraction of the population, incidence of 100 percent) The control variable measuring "the fraction of the population in the same metropolitan area as the individual who are

[17] See, for instance, User Note 1 in Appendix G of the Codebook of the 1994 March Supplement of the CPS, available at: https://cps.ipums.org/cps/resources/codebooks/cpsmarapr94.pdf (last accessed: June 25, 2023).

first or second-generation immigrants from the same country of origin" (p. 1562) has actually been computed separately for each year from 1994 to 2014.

3. (GDP 2014 - incidence between 7.9 and 8.7 percent[18]) The "natural log of the current per-capita GDP in the country of origin (measured in the survey year)" is entirely missing for the year 2014. Consequently, all the observations corresponding to this year are incorrectly excluded from the estimation sample.[19]

4. (Missing MSA 1, incidence between 17.0 and 19.2 percent) The estimation samples in the four data columns of Table 4 include individuals who reside in metropolitan areas that are not separately identified. These areas are represented by the `metarea` values 9997 (Other metropolitan areas, unidentified), 9998 (NIU, household not in a metropolitan area) or 9999 (Missing data). The *corrigendum* reads: "On page 1562, the sentence ending in 'denotes the inclusion of residence (i.e. metropolitan-area) fixed effects' should be followed by the clarifying footnote: 'Rural locations within the United States do not have an MSA code. Throughout the paper, we treat a missing MSA as a category in the residence fixed effects.'" (Giuliano and Nunn, 2024). It also should be noted that even after the *corrigendum*, the statement: "[b]y examining individuals who live in the same location, we are able to hold constant the contemporaneous local environment, while examining the effects of an individual's ancestral environment" (p. 1564) remains factually incorrect. This is because the code pools together individuals for whom the variable `metarea` is equal to 9997, 9998 or 9999 irrespective of their actual locations within the United States. Finally, the *corrigendum* does not clarify whether these three different values of the variable `metarea` are treated as distinct categories in the residence fixed effects.

5. (Missing MSA 2, incidence between 17.0 and 19.2 percent) The control variable that measures "the fraction of the population in the same metropolitan area as the individual who are first or second-generation immigrants from the same country of origin" (p. 1562)

[18]For each equation, the incidence have been computed by dividing the number of observations corresponding to countries with a non-missing GDP in 2013, but a missing GDP in 2014 and all other variables that are non-missing, by the total number of observations in the sample

[19]This variable is also *entirely* missing for some countries of origin (e.g., Argentina, Syria), leading to their exclusion from the estimation. However, the exact same control variable exists for these countries for the year 2000 in Eqs. (1.5) and (1.6).

cannot be meaningfully defined for observations where the variable `metarea` is equal to 9997, 9998 or 9999. Additionally, this variable is not constant for individuals associated with the same foreign country c, surveyed in the same year, and for whom `metarea` is equal to (for example) 9997. This inconsistency contradicts the implication of the *corrigendum* that these values of the variable are treated equally to the other non-missing values.

Correcting the coding mistakes

1. (Dependent variable) The observations for which the variable `marst` is equal to 2 (Married, spouse absent) should be excluded from the estimation sample. This correction is feasible.

2. (Fraction of the population) The control variable should be defined in a time-invariant way. This correction is unfeasible (using the replication package).

3. (GDP 2014) The data cleaning and manipulation code could be easily fixed. This correction is unfeasible (using the replication package), as the replication package only mentions the data source but not the exact identifier of this variable.

4. (Missing MSA 1) The observations for which the variable `metarea` is equal to 9997, 9998 or 9999 should be excluded from the sample. This correction is feasible, but no change in the code is needed on the basis of the *corrigendum*.

5. (Missing MSA 2) The control variable should be defined in a way that does not vary, when the variable `metarea` is equal to 9997, 9998 or 9999, across observations corresponding to a unique value of c. This correction is unfeasible (using the replication package).

1.2.3 Eq. (5) Speaking a foreign language at home

GN estimate the following equation on a sample of individuals born in the United States and reporting a foreign ancestry in the 2000 census:

$$\mathrm{I}_{i,c,k}^{ForeignLang} = \alpha_k + \beta \text{Climatic Instability}_c + \boldsymbol{X}_c \Pi + \boldsymbol{X}_{c,k} \Omega + \boldsymbol{X}_i \Phi + \epsilon_{i,c,k} \tag{1.5}$$

"where i denotes an individual, c his/her ancestry, and k a location of residence (metropolitan areas). The dependent variable, $\mathrm{I}_{i,c,k}^{ForeignLang}$, is an indicator that equals to one if a language other than English is spoken at home. α_k denotes the inclusion of residence (i.e. metropolitan areas) fixed effects. [In footnote 24] **We omit from the sample individuals from origin countries with English as an official language**. $\boldsymbol{X}_c$ denotes ancestral country-level covariates: historical distance from the equator, historical economic development, historical political complexity, the GDP in the country of origin measured at the time of the survey, and **the linguistic distance between the country of origin and the U.S.**" (p. 1565, bold emphasis added).

Coding mistakes

1. (English, incidence between 3.7 and 5.8 percent) The sample used to estimate Eq. (1.5) includes natives with ancestry in seven distinct countries or territories (out of 84) in which English is an official language.[20]

   The *corrigendum* reads:

   > "4. On page 1565, for the text 'report their ancestry as being a non-English speaking country,' the word 'country' should be 'group'.
   > 5. In footnote 24, the text 'individuals from origin countries with English as an official language.' should be 'individuals from ancestries with English as an official language and the language of instruction in higher education.' "

   The *corrigendum* amending the original article changes the sample selection criterion, shifting it from the level of the ancestral country to the level of each ancestral group. This modification requires defining the official language and language of instruction in higher education for each ancestral group. However, the *corrigendum* does not specify the data source used to implement this sample selection criterion, resulting in a lack of precise description of the specification of Eq. (1.5), unlike in the original article.

   Through personal communication with the Editor in charge of the original article, it was brought to our attention that, concerning Canada, the estimation only included

[20] The seven countries or territories are Canada, Puerto Rico, the Philippines, India, Samoa, Pakistan, and Eritrea; we identify countries and territories where English is an official language by cross-checking Wikipedia contributors (2023) and Central Intelligence Agency (2021), retaining only the cases in which both data sources report English as an official language.

natives of French Canadian ancestry and disregarded individuals who defined themselves exclusively as being of Canadian ancestry.[21]

Moving the sample selection criterion from the level of the ancestral country to the level of the ancestral group, as Giuliano and Nunn (2024) do, also breaks the symmetry with respect to the construction of the elements in the vector $\boldsymbol{X}_c$, which includes "ancestral country-level covariates". This discrepancy is evident for `lingprox_dominant_a`, a variable used in `Table5.dta` measuring "the linguistic distance between the country of origin and the U.S." is borrowed from "linguistic proximity between dominant groups (Fearon data)". As English is the dominant language in Canada, this implies that observations for *French* Canadians, identified by `isocode=="CAN"`, have a value of `lingprox_dominant_a` at 15 (i.e. the one associated to English), the highest value in the data.[22]

Therefore, according to Giuliano and Nunn (2024), French Canadians are included in the sample as this group does not have English as its official language, even though the value measuring linguistic proximity is determined using English for Canada.

2. (Missing MSA 1, incidence between 25.0 and 32.2 percent) The estimating sample includes observations for which the variable `metaread` is equal to 0 (Not identifiable or not in an MSA). A unique value of the fixed effect α_k is incorrectly assigned to these observations, bundling them together. The *corrigendum* addresses this point, as described with respect to Eq. (1.4) above.

3. (Missing MSA 2, incidence between 25.0 and 32.2 percent) The vector $\boldsymbol{X}_{c,k}$ includes a variable measuring the "fraction of those living in the same metropolitan area who are first-generation immigrants of the same ancestry" (p. 1565). This variable is incorrectly defined for the observations for which the variable `metaread` is equal to 0.

[21]The replication package includes the three-digit ISO code of the country of ancestry, not the name of the ancestral group.

[22]Notice that the linguistic proximity between the United States and France (French) is 1; the only other country of ancestry with a value of 15 (the value corresponding to English) is Samoa, which is also a country with English as an official language, and the second highest value is 8 (for Puerto Rico), a territory with English as an official language, and the average linguistic proximity stands at 2.2 in the Authors' original sample; this contributes to explain why the limited share of observations corresponding to countries where English is an official language play such a pivotal role in the significance of the variable of interest.

4. (Fraction, incidence between 25.0 and 32.2 percent) The variable measuring the "fraction of those living in the same metropolitan area who are first-generation immigrants of the same ancestry" (p. 1565) varies across individuals from the same country of ancestry, and for which the variable `metaread` is equal to 0 (Not identifiable or not in an MSA).

5. (Living with parents, incidence undetermined) Cols. (2) and (3) in Table 5 estimate (1.5) separately for individuals not living and living with parents.[23] The sample split is done, at lines 106–108, using the variable `relate` describing the relationship of individual i to the household head. Thus, the subsample of individuals "Living with parents" includes all the individuals who are children of the household head, while the subsample "Not living with parents" includes all other individuals. For three-generation households, the split is incorrect when, for instance, a grandchild of the household head co-resides with at least one of her parents, or when the household also includes a parent of the household head. In these two cases, the grandchild and the household head are incorrectly classified by GN as "Not living with parents".

Correcting the coding mistakes

1. (English) The correction can be implemented by utilizing the variable `isocode` to identify and exclude observations associated with countries and territories where English is an official language. The correction is feasible.

2. (Missing MSA 1) The observations with a value 0 for the variable `metaread` should be excluded from the sample. This correction is feasible, but no change in the code is needed on the basis of the correction.

3. (Missing MSA 2) The observations with a value 0 for the variable `metaread` should be excluded from the sample because the control variable cannot be meaningfully defined for these observations. The correction is feasible.

4. (Fraction) The fraction of the population in the same metropolitan area that are first or second-generation immigrants from the same country of origin should be treated equally to the other non-missing values. In particular, it should be time-invariant. The correction is unfeasible (using the replication package).

[23]"In columns 2 and 3, we split the samples into those not living with their parents and those living with their parents." (p. 1567).

5. (Living with parents) The correct sample split and the exact identification of a co-residing parent is possible using the variables `momloc` and `poploc`, which are available from IPUMS USA (Ruggles et al., 2023).[24] This correction is unfeasible (using the replication package).

1.2.4 Eq. (6) Speaking a foreign language at home (living with parents)

"We estimate whether the intergenerational transmission of the tradition of speaking one's ancestral language is affected by climatic instability of the ancestral country [...]:

$$\begin{aligned} \mathrm{I}^{ForeignLang}_{i,c,k} = \beta_1 \text{Climatic Instability}_c + \beta_2 \mathrm{I}^{ParentForeignLang}_{i,c,k} + \\ + \beta_3 \mathrm{I}^{ParentForeignLang}_{i,c,k} \times \text{Climatic Instability}_c + \\ + \boldsymbol{X}_c \Pi + \boldsymbol{X}_{c,k} \Omega + \boldsymbol{X}_i \Phi + \alpha_k + \epsilon_{i,c,k} \end{aligned} \tag{1.6}$$

[...] $\mathrm{I}^{ParentForeignLang}_{i,c,k}$, which is an indicator variable that equals one if either the father or the mother speaks a foreign language at home."

Coding mistakes

Eq. (1.6) inherits the five coding mistakes described above with respect to Eq. (1.5):

1. (English, incidence between 5.3 and 5.8 percent)
2. (Missing MSA 1, incidence between 25.0 and 26.1 percent)
3. (Missing MSA 2, incidence between 25.0 and 26.1 percent)
4. (Fraction, incidence between 25.0 and 26.1 percent)
5. (Living with parents, incidence undetermined)

Correcting the coding mistakes

The coding mistakes can be fixed as described with respect to Eq. (1.5).

[24] These variables provide the identifiers of the co-residing mother and father. See, for instance, https://usa.ipums.org/usa-action/variables/MOMLOC#availability_section (last accessed on May 31, 2023). Using these variables also avoids incorrectly identifying a step-father as the father of the child, something that may occur using `relate`.

1.2.5 Eq. (8) Indigenous populations speaking a traditional language

For a subsample of individuals who report a Native American tribal affiliation drawn from the 1930, 1990 and 2000 censuses,

> "[the] estimating equation is:
>
> $$\mathrm{I}_{i,e,k}^{NativeLang} = \alpha_k + \beta \text{Climatic Instability}_e + \boldsymbol{X}_e \Pi + \boldsymbol{X}_i \Phi + \epsilon_{i,e,k} \tag{1.8}$$
>
> where i denotes an individual, e his/her ethnic group, and k a location of residence (metropolitan area). The dependent variable, $\mathrm{I}_{i,e,k}^{NativeLang}$, is an indicator that equals to one if the individual reports speaking an Indigenous language at home. The specification includes location (i.e. metropolitan areas) fixed effects, α_k. Thus, the variation used to estimate β is across individuals from different Native American ethnic groups, but living in the same location. [in footnote 30] **The 1930, 1990 and 2000 US Censuses ask the following question: 'Does the person speak a language other than English at home?' If yes, the person indicates which language**." (p. 1571, bold emphasis added).

Coding mistakes

1. (Language, incidence between 7.3 and 16.9 percent) The questionnaire of the 1930 census did not include the two questions necessary to define the dependent variable in Eq. (1.8).[25] For instance, the sample used in Col. (1) of Table 7 to estimate Eq. (1.8) includes 11,468 obs. (9.0 percent) drawn from this census, and corresponding to 83 Native American ethnic groups.

   Until September 2021, IPUMS USA incorrectly coded the variable `language` for the 1930 census using the information on the mother tongue.[26] In reality, the question about the mother tongue should have been asked in 1930 only to foreign-born individuals.[27] However, it appears that some enumerators may not have followed this restriction.[28]

[25] See https://usa.ipums.org/usa-action/variables/LANGUAGE#availability_section (last accessed on May 31, 2023).

[26] See https://usa.ipums.org/usa-action/revisions#revision_02_09_2021 (last accessed on October 23, 2023); we thank Paola Giuliano and Nathan Nunn for bringing this to our attention.

[27] See https://usa.ipums.org/usa-action/variables/MTONGUE#description_section (last accessed on October 23, 2023).

[28] Private communication with Paola Giuliano and Nathan Nunn, who contacted IPUMS USA for clarification.

The replication file reveals that the dependent variable is always equal to 0 for 68 of these 83 ethnic groups in 1930. Furthermore, two groups, namely Navajo and Hopi, account for the great majority (1,860 out of 1,951) of the observations where the dependent variable in Eq. (1.8) is equal to 1 in 1930.[29] This very strongly suggests that the question about mother tongue was not systematically asked to all Native Americans in the 1930 census.[30]

2. (Children 0 to 4, incidence between 3.0 and 21.4 percent) The language questions used to define $\mathrm{I}_{i,e,k}^{NativeLang}$ are specifically asked to individuals aged 5 and above in the 1990 and 2000 censuses.[31] However, in the estimation sample, there are observations corresponding to children aged 0 to 4 for whom $\mathrm{I}_{i,e,k}^{NativeLang}$ is always equal to 0, although it should have been marked as missing since the language questions are not applicable to individuals in this age group.

3. (Employment, incidence between 0.1 and 61.1 percent) The question regarding employment status, which is used to define an element of the vector $\boldsymbol{X}_i$ in Eq. (1.8), is only asked to individuals aged 16 and above in the 1990 and in the 2000 census.[32] In the estimation of Eq. (1.8), the observations where the variable `empstat` is not available are treated as a separate category of this categorical variable. This differs from what is stated on p. 1572, where the Authors mention that labor-force-status fixed effects have three values: employed, unemployed, and outside of the labor force.

4. (Missing MSA, incidence between 60.2 and 75.3 percent) In the estimation sample, there are observations for which the variable `metarea` is equal to 0, indicating that the location is "Not identifiable or not in an MSA." The *corrigendum* addresses this issue, as explained in relation to Eq. (4) above. We note that the statement "the variation used to estimated β is across individuals from different Native American ethnic groups,

[29] The share of observations where the dependent variable equals 1 is 91.0 for Navajo and 99.2 percent for Hopi in 1930.

[30] It is worth noting that the replication package only includes the binary dependent variable and does not provide information on the language associated with each individual.

[31] See https://usa.ipums.org/usa-action/variables/LANGUAGE#universe_section (last accessed on June 6, 2023).

[32] See https://usa.ipums.org/usa-action/variables/EMPSTAT#universe_section (last accessed on October 23, 2023).

but living in the same location" (p. 1571) is factually incorrect, and not modified by the *corrigendum*.

5. (Living with parents, incidence undetermined) "Estimates of equation (1.9) are reported in columns 1–3 of Table 7 for all individuals (column 1), those not living with their parents (column 2), and those living with their parents (columns 3)." (p. 1572). The split of the sample is done on the basis of the variable `relate` at lines 153 and 155 in the replication file. The remarks related to the fourth mistake for Eqs. (1.5) and (1.6) also apply here.

Correcting the coding mistakes

1. (Language) The variable `year` can be used to exclude the observations corresponding to the 1930 census. The correction is feasible.

2. (Children 0 to 4) The variables `age` and `year` can be used to exclude observations corresponding to children aged 0 to 4 drawn from the 1990 and the 2000 censuses. The correction is feasible.

3. (Employment) The variables `age` and `year` can be used to exclude observations corresponding to children aged 0 to 15 drawn from the 1990 and the 2000 census. The correction is feasible.

4. (Missing MSA) The observations in which the variable `metarea` is equal to 0 should be excluded from the sample. This correction is feasible, but no change in the code is needed on the basis of the *corrigendum*.

5. (Living with parents) The correction for the fourth coding mistakes in Eqs. (1.5) and (1.6) should also be applied to address this issue. The correction is not feasible (using the replication package).

1.2.6 Eq. (9) Speaking a traditional language (living with parents)

> "We again explore the transmission of this cultural practice [speaking an Indigenous language at home] from parents to children by estimating:
>
> $$\begin{aligned} \mathrm{I}_{i,e,k}^{NativeLang} &= \alpha_k + \beta_1 \text{Climatic Instability}_e + \\ +\beta_2 \mathrm{I}_{i,e,k}^{ParentNativeLang} + \beta_3 \mathrm{I}_{i,e,k}^{ParentNativeLang} &\times \text{Climatic Instability}_e + \\ &+ \boldsymbol{X}_e \Pi + \boldsymbol{X}_i \Phi + \epsilon_{i,e,k} \end{aligned} \tag{1.9}$$
>
> where all variables are defined as in Eq. (1.8) and $\mathrm{I}_{i,e,k}^{ParentNativeLang}$ is an indicator variable that equals one if individual i's parents speak an indigenous language at home." (p. 1572).

Eq. (1.9) is estimated for individuals drawn from the 1930, 1990 and 2000 censuses reporting an affiliation to a Native American tribe and "living with parents" (p. 1572).

Coding mistakes

Eq. (1.9) inherits the five coding mistakes described above for Eq. (1.8):

1. (Language, incidence between 13.8 and 16.9 percent)
2. (Children 0 to 4, incidence between 16.0 and 17.0 percent)
3. (Employment, incidence between 59.1 and 59.2 percent)
4. (Missing MSA, incidence between 73.0 and 75.3 percent)
5. (Living with parents, incidence undetermined).

Correcting the coding mistakes

The coding mistakes can be fixed as described with respect to Eq. (1.8).

Impact of the corrections

Our "deep reproduction" exercise reveals significant discrepancies between the original study's reported results and those obtained using the textual definitions provided in the article, even after accounting for the content of the *corrigendum*. We now examine the consequences of these discrepancies for the validity of the original study's findings. In particular, we investigate how the identified issues affect the statistical precision, p-values, and magnitude of the original estimates. By doing so, we provide a comprehensive understanding of the implications of our reproduction failure, shedding light on the robustness of the original study's conclusions.

To facilitate the exposition of the consequences of reproduction failure, we opt for a graphical representation of the results rather than the more traditional tabular format. This choice is motivated by the complexity of the errors identified, which often involve multiple corrections for each equation. Creating a table to analyze the impact of each error separately would imply generating a multitude of tables, akin to the original, but correcting one mistake at a time. Furthermore, we can also correct several errors and combinations of errors, making a tabular presentation cumbersome. Instead, we create figures that provide a concise and intuitive visualization of the results.

For each targeted specification, we generate a figure that compares the original estimation with the corrected results. The figure displays the original estimation, including confidence intervals, in red, alongside the corrected results, in blue. The graphical representation also includes information on sample size and R^2, with a section for comments on the corrections implemented. The corrected results are presented relative to the baseline, original regression. Figure 1.1 provides an overview of the impact of each error in Table 4, Figure 1.2 focuses on Table 5 whereas Figure 1.3 centers on Table 7. Additionally, Tables 1.A.3 and 1.A.4 in our Online Appendix present the results of Table 4 after implementing the (feasible) corrections we identified in previous sections.[33] Similarly, Table 1.A.5 present the corrected estimates corresponding to Table 5. Finally, Tables 1.A.6 and 1.A.7 display the coefficients of Table 7 once the corrections are introduced.[34]

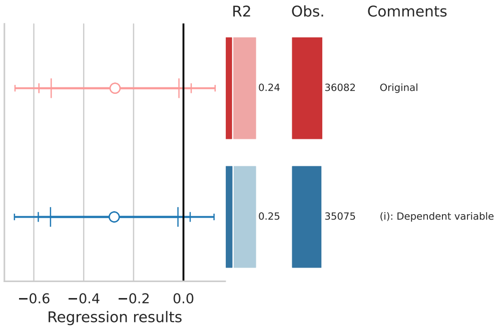
Our "deep reproduction" exercise reveals a total of 26 coding mistakes in the original study. Notably, 53.5 percent of these mistakes (14 out of 26) could be fixed using the codes and data provided in the original replication package. Upon correcting the coding mistakes, we observe substantial changes in the estimates, with some coefficients undergoing dramatic shifts in magnitude and significance. In particular, Table 5 appears to be rather fragile to our proposed corrections, with 5 out of 7 results loosing significance once all corrections are introduced. For instance, the coefficient on the main variable of interest changes from -0.767 to -0.427 in Col. (3), a 38 percent decrease in magnitude, and becomes insignificant. The remaining tables

[33]These two tables differ in their treatment of observations with a missing metropolitan area. The former includes those while the latter does not.

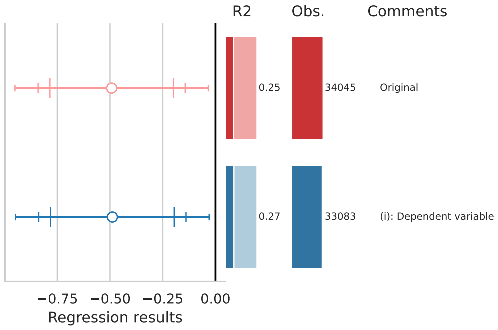
[34]Similarly to Table 4, we present two variants of Table 7. The first includes observations with a missing metropolitan area. The second does not.

Figure 1.1: Estimates of Eq. (1.4)

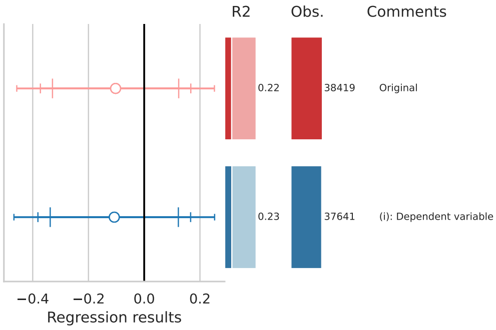
(a) Table 4, Col. (1)



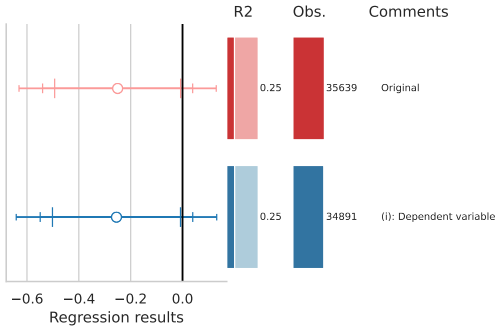
(b) Table 4, Col. (2)



(c) Table 4, Col. (3)



(d) Table 4, Col. (4)

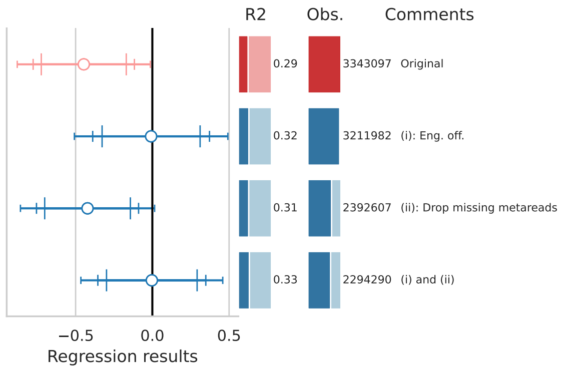


Notes: This figure compares the original estimation (red) with the corrected results (blue) for Eq. (1.4), originally reported in Table 4 of GN. The graphical representation focuses on the main coefficient of interest in each column (interaction term where applicable), and includes confidence intervals, sample size, and R^2, with comments on the corrections implemented.

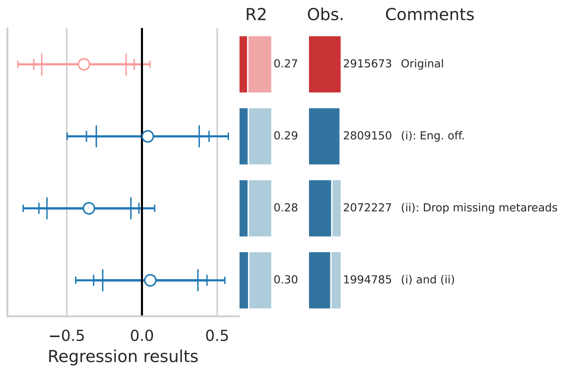
are impacted to a lesser extent. Overall, applying the corrections, 53.0 percent of the main estimates (9 out of 17) become indistinguishable from zero, highlighting the significance of these errors for the validity of the original study's findings.

Figure 1.2: Estimates of Eqs. (1.5) and (1.6)

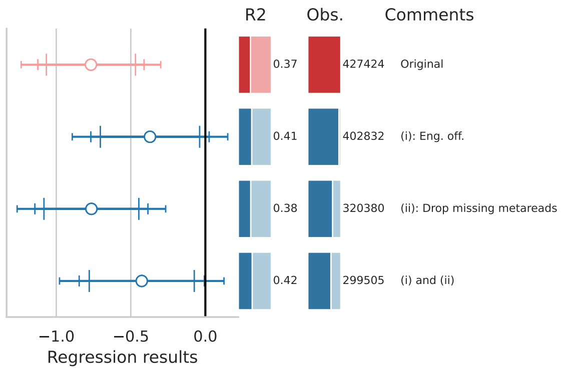
(a) Table 5, Col. (1)



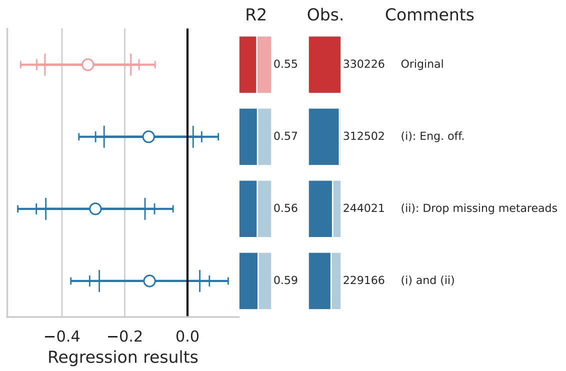
(b) Table 5, Col. (2)



(c) Table 5, Col. (3)



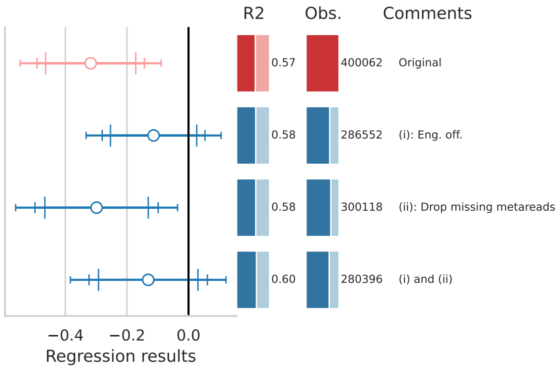
(d) Table 5, Col. (4)



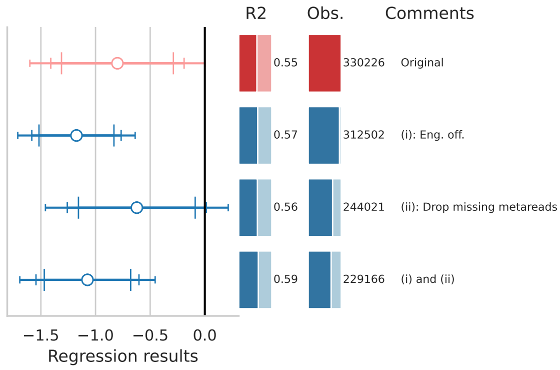
Continued on next page

Figure 1.2 *(continued from previous page)*

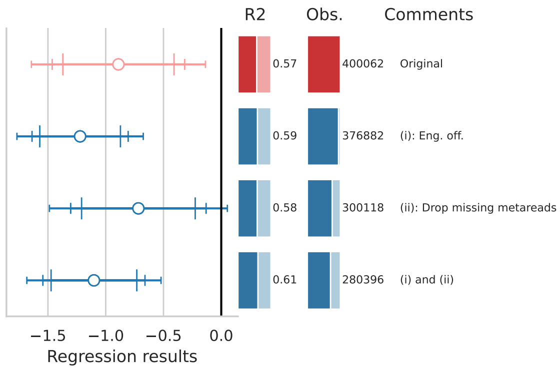
(e) Table 5, Col. (5)



(f) Table 5, Col. (6)



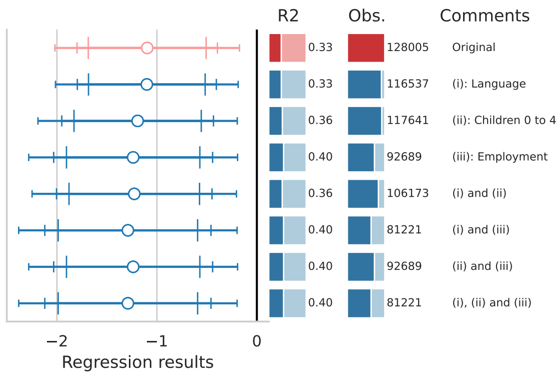
(g) Table 5, Col. (7)



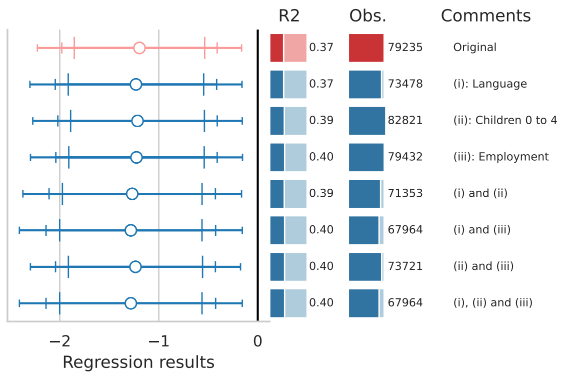
Notes: This figure compares the original estimation (red) with the corrected results (blue) for Eqs. (1.5) and (1.6), originally reported in Table 5 of GN. The graphical representation focuses on the main coefficient of interest in each column (interaction term where applicable), and includes confidence intervals, sample size, and R^2, with comments on the corrections implemented.

Figure 1.3: Estimates of Eqs. (1.8) and (1.9)

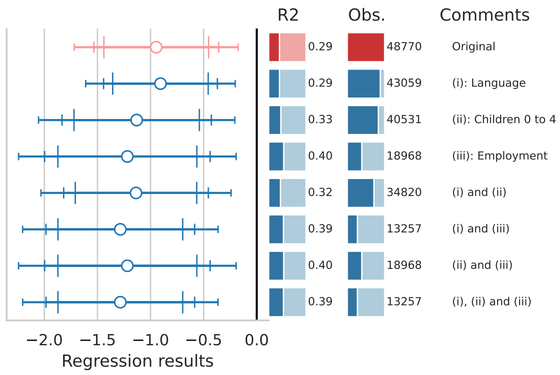
(a) Table 7, Column 1



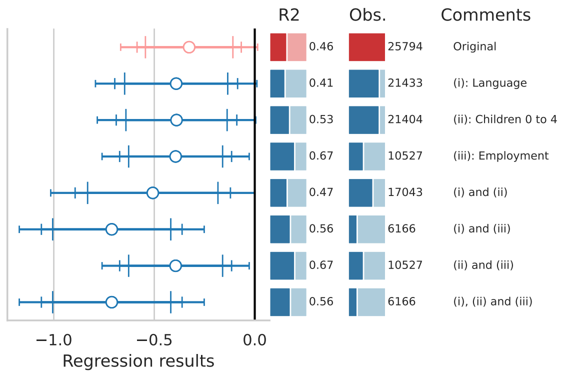
(b) Table 7, Col. (2)



(c) Table 7, Col. (3)



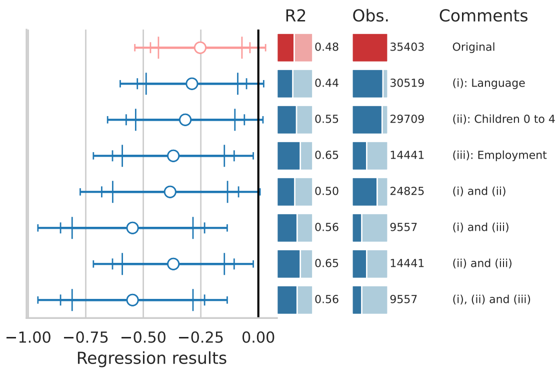
(d) Table 7, Col. (4)



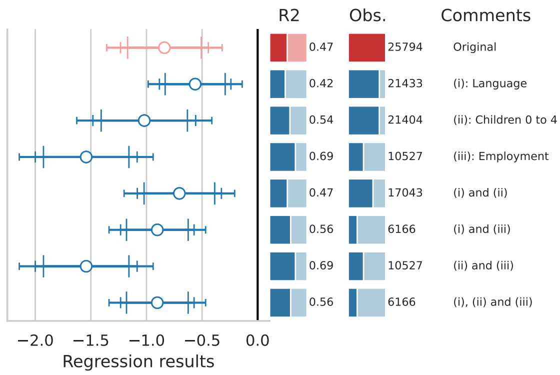
Continued on next page

Figure 1.3 *(continued from previous page)*

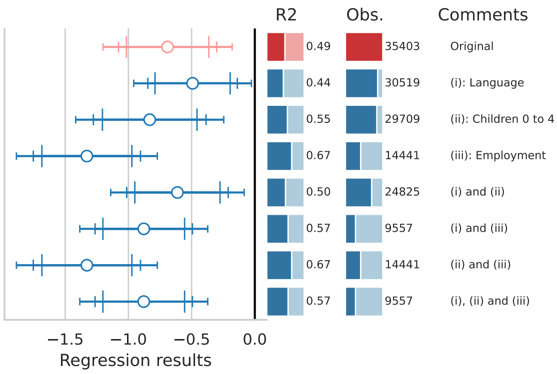
(e) Table 7, Col. (5)



(f) Table 7, Col. (6)



(g) Table 7, Col. (7)



Notes: This figure compares the original estimation (red) with the corrected results (blue) for Eqs. (1.8) and (1.9), originally reported in Table 7 of GN. The graphical representation focuses on the main coefficient of interest in each column (interaction term where applicable), and includes confidence intervals, sample size, and R^2, with comments on the corrections implemented.

1.3 Concluding remarks

Our "deep reproduction" of the individual-level analysis from the study by Giuliano and Nunn (2021) was unsuccessful. Despite the Authors' *corrigendum* (Giuliano and Nunn, 2024), which addresses some imprecision in the original text, we identified several remaining discrepancies between the precise descriptions provided in the article and the corresponding code included in the replication package. These inconsistencies span multiple aspects of the analysis, including variable definitions, sample selection criteria, and regression specifications. Our exercise suggests that correcting for these inconsistencies has a significant impact on the results, which, upon re-examination, reveal a more nuanced relationship between ancestral climatic variability and tradition.

The deep reproduction exercise highlights the importance of ensuring consistency between the descriptions in published work and the underlying analysis code and data. Comprehensive replication packages, including all necessary files, can greatly facilitate verification efforts and enhance research transparency, and the steps taken in this direction by economic journals, such as the *Review of Economic Studies*, are clearly positive.

Data Availability Statement

The data and codes that support the findings of this study are openly available in `https://doi.org/10.3886/E202861V2`, reference number 202861. The data we provide are a snapshot copy of "Replication package for: Understanding Cultural Persistence and Change" by Paola Giuliano and Nathan Nunn, available at `https://zenodo.org/records/4026093`, and distributed under a Creative Commons Attribution 4.0 International license.

Bibliography

Acemoglu, D., G. Egorov, and K. Sonin (2021): "Institutional change and institutional persistence," in *The Handbook of Historical Economics*, ed. by A. Bisin and G. Federico, Elsevier, 365–389.

Alesina, A. and N. Fuchs-Schündeln (2007): "Good-Bye Lenin (or Not?): The Effect of Communism on People's Preferences," *American Economic Review*, 97, 1507–1528.

Bau, N. (2021): "Can policy change culture? Government pension plans and traditional kinship practices," *American Economic Review*, 111, 1880–1917.

Bazzi, S., A. Ferrara, M. Fiszbein, T. Pearson, and P. A. Testa (2023): "The Other Great Migration: Southern Whites and the New Right," *The Quarterly Journal of Economics*, 138, 1577–1647.

Bazzi, S., M. Fiszbein, and M. Gebresilasse (2020): "Frontier culture: The roots and persistence of "rugged individualism" in the United States," *Econometrica*, 88, 2329–2368.

Bazzi, S., A. Gaduh, A. D. Rothenberg, and M. Wong (2019): "Unity in Diversity? How Intergroup Contact Can Foster Nation Building," *American Economic Review*, 109, 3978–4025.

Bertoli, S., M. Clerc, J. Loper, and È. Roca Fernández (2023): "Migration and the epidemiological approach: when history matters," Paper presented at the 17*th* Conference on Migration and Development, Boston University, October 2023.

——— (2024): "Understanding cultural persistence and change: a replication of Giuliano and Nunn (2021)," I4R Discussion Paper Series No. 116, Institute for Replication.

Bisin, A. and G. Federico, eds. (2021): *The Handbook of Historical Economics*, Academic Press, paperback ed.

Bursztyn, L., T. Chaney, T. A. Hassan, and A. Rao (2024): "The Immigrant Next Door," *American Economic Review*, 114, 348–84.

Caprettini, B. and H.-J. Voth (2022): "New Deal, New Patriots: How 1930s Government Spending Boosted Patriotism During World War II," *The Quarterly Journal of Economics*, 138, 465–513.

Central Intelligence Agency (2021): "The World Factbook 2021," [Online; accessed 20-June-2023].

Clemens, M. A. (2017): "The meaning of failed replications: A review and proposal," *Journal of Economic Surveys*, 31, 326–342.

Dreber, A. and M. Johannesson (2023): "A framework for evaluating reproducibility and replicability in economics," I4R Discussion Paper No. 38, Institute for Replication.

Enke, B. (2019): "Kinship, cooperation, and the evolution of moral systems," *The Quarterly Journal of Economics*, 134, 953–1019.

Fernández, R. (2011): "Does culture matter?" in *Handbook of Social Economics*, ed. by J. Benhabib, A. Bisin, and M. O. Jackson, Elsevier, vol. 1, 481–510.

Ghosh, A., S. I. M. Hwang, and M. Squires (2023): "Economic Consequences of Kinship: Evidence From U.S. Bans on Cousin Marriage," *The Quarterly Journal of Economics*, 138, 2559–2606.

Giuliano, P. and N. Nunn (2021): "Understanding cultural persistence and change," *The Review of Economic Studies*, 88, 1541–1581.

——— (2024): "Correction to: Understanding Cultural Persistence and Change," *The Review of Economic Studies*, 91, 597.

Huntington-Klein, N., A. Arenas, E. Beam, M. Bertoni, J. R. Bloem, P. Burli, N. Chen, P. Grieco, G. Ekpe, T. Pugatch, M. Saavedra, and Y. Stopnitzky (2021): "The influence of hidden researcher decisions in applied microeconomics," *Economic Inquiry*, 59, 944–960.

Krähmer, D., L. Schächtele, and A. Schneck (2023): "Care to share? Experimental evidence on code sharing behavior in the social sciences," *PLOS ONE*, 18, e0289380.

Michalopoulos, S. and M. M. Xue (2021): "Folklore," *The Quarterly Journal of Economics*, 136, 1993–2046.

Ruggles, S., S. Flood, M. Sobek, D. Brockman, G. Cooper, S. Richards, and M. Schouweiler (2023): "IPUMS USA: Version 13.0 [dataset]," Minneapolis, MN: IPUMS, https://doi.org/10.18128/D010.V13.0.

Slotwinski, M. and A. Stutzer (2022): "Women Leaving the Playpen: the Emancipating Role of Female Suffrage," *The Economic Journal*, 133, 812–844.

Valencia Caicedo, F. (2018): "The Mission: Human Capital Transmission, Economic Persistence, and Culture in South America," *The Quarterly Journal of Economics*, 134, 507–556.

Voth, H.-J. (2021): "Persistence – myth and mystery," in *The Handbook of Historical Economics*, ed. by A. Bisin and G. Federico, Elsevier, 243–267.

Wikipedia contributors (2023): "List of official languages by country and territory — Wikipedia, The Free Encyclopedia," [Online; accessed 20-June-2023].

1.A Supplementary tables and figures

Table 1.A.1: Synthesis of the corrections

Eq. num.	Inconsistency	Correction	Applied
2	The information used is not on mother tongue, but on the language that the respondent speaks at home	N/A	No
4	The dependent variable is defined only for individuals co-residing with their spouses	Exclude from the estimation sample observations for which the variable marst is equal to 2 (Married, spouse absent)	Yes
	The control variable measuring "the fraction of the population in the same metropolitan area as the individual who are first or second-generation immigrants from the same country of origin" (p. 1562) is computed separately for each year from 1994 to 2014	Compute the variable in a time-invariant way	No
	The "natural log of the current per-capita GDP in the country of origin (measured in the survey year)" is entirely missing for the year 2014	Add the GPD data for 2014	No

Continued on next page

Table 1.A.1 – Continued from previous page

Eq. num.	Inconsistency	Correction	Applied
	The estimation samples include individuals that reside in metropolitan areas that are not separately identified (i.e. "Other metropolitan areas, unidentified"; "NIU, household not in a metropolitan area"; "Missing data")	Exclude observations for which the variable `metarea` is equal to 9997, 9998 and 9999	Yes
	The control variable measuring "the fraction of the population in the same metropolitan area as the individual who are first or second-generation immigrants from the same country of origin" (p. 1562) is not constant, for individuals associated to the same foreign country c, surveyed in the same year	The control variable should be defined in a way that does not vary, when the variable metarea is equal to 9997, 9998 or 9999, across observations corresponding to a unique value of c	No
5–6	The value of the variable `lingprox_dominant_a` associated with Canada corresponds to English and stands at 15. However, the estimation sample excludes natives of Canadian ancestry, but includes natives of French Canadian ancestry	Exclude observations for which isocode is "CAN" or replace the value of lingprox_dominant_a by the one corresponding to French (1)	Yes

Continued on next page

Table 1.A.1 – Continued from previous page

Eq. num.	Inconsistency	Correction	Applied
	The estimation sample include individuals that reside in metropolitan areas that are not separately identified (i.e. "Not identifiable or not in an MSA")	Exclude observation for which he variable metarea is equal to 0	Yes
	The variable measuring the "fraction of those living in the same metropolitan area who are first-generation immigrants of the same ancestry" (p. 1565) is incorrectly defined for the observations for which the variable metaread is equal to 0.	Exclude observation for which the variable metarea is equal to 0	Yes
	The split between living and not living with parents is done using the variable relate, which only describes the relationship of individual i to the household head	Use the variables `momloc` and `poploc`, available from IPUMS USA (Ruggles et al., 2023).	No
8–9	The questionnaire of the 1930 census did not include the two questions necessary to define the dependent variable	Exclude the observations drawn from the 1930 census	Yes

Continued on next page

Table 1.A.1 – Continued from previous page

Eq. num.	Inconsistency	Correction	Applied
	The two questions necessary to define the dependent variable are asked to individuals aged 5 and above	Exclude the observations for children aged 0 to 4	Yes
	The question necessary to define the employment status is asked to individuals aged 16 and above	Exclude the observations for children aged 0 to 15	Yes
	The estimation sample include individuals that reside in metropolitan areas that are not separately identified (i.e. "Not identifiable or not in an MSA").	Exclude observation for which the variable `metarea` is equal to 0.	Yes
	The split between living and not living with parents is done using the variable `relate`, which only describes the relationship of individual i to the household head	Use the variables `momloc` and `poploc`, available from IPUMS USA (Ruggles et al., 2023)	No

Table 1.A.2: Appendix tables associated with equations containing mistakes

Table number	Column number	Equation number
A25	1–5	8
	6	10
A26	1–5	8
A27	1–5	8
A30	1–5	8
A31	1–5	8

Table 1.A.3: Correction of Table 4 - Women and men marrying a spouse from their origin country, using CPS 1994-2014 (not excluding observations for which the metropolitan area is missing)

	Dependant variable: indicator variable for spouse being from the same origin country			
	Sample: married women		Sample: married men	
	Origin country identified from father	Origin country identified from mother	Origin country identified from father	Origin country identified from mother
	(1)	(2)	(3)	(4)
Climatic instability	-0.278*	-0.489***	-0.107	-0.255*
	(0.155)	(0.178)	(0.140)	(0.150)
Country-level controls:				
Distance from equator	-0.006**	-0.005	-0.008***	-0.009***
	(0.003)	(0.003)	(0.003)	(0.003)
Economic complexity	0.006	0.015	-0.013	-0.023
	(0.028)	(0.037)	(0.040)	(0.039)
Political hierarchies	0.094***	0.088***	0.094**	0.088**
	(0.028)	(0.030)	(0.038)	(0.039)
Ln (per capita GDP)	-0.002	-0.019	-0.001	-0.003
	(0.031)	(0.035)	(0.037)	(0.036)
Genetic distance from the US	0.030	0.009	0.012	-0.008
	(0.047)	(0.055)	(0.045)	(0.046)
Fraction of population in location who are 1st or 2nd-generation immigrants from same country of origin	3.405***	3.622***	3.104***	3.430***
	(0.509)	(0.655)	(0.520)	(0.503)
Individual controls	Yes	Yes	Yes	Yes
Number of countries	108	105	110	105
Mean (st.dev.) of dependent variable	0.33 (0.47)	0.32 (0.47)	0.28 (0.45)	0.29 (0.45)
Observations	35,075	33,083	37,641	34,891
Original observations	36,082	34,045	38,419	35,639
R^2	0.251	0.267	0.228	0.251

Notes: Mistake (Dependent variable) is corrected. Mistake (Missing MSA 1) is not corrected. Mistakes (Fraction of the population) and (GDP 2014) cannot be corrected. Except for these corrections, the specification is strictly the same as in GN.

Table 1.A.4: Correction of Table 4 - Women and men marrying a spouse from their origin country, using CPS 1994-2014 (excluding observations for which the metropolitan area is missing)

	Dependant variable: indicator variable for spouse being from the same origin country			
	Sample: married women		Sample: married men	
	Origin country identified from father	Origin country identified from mother	Origin country identified from father	Origin country identified from mother
	(1)	(2)	(3)	(4)
Climatic instability	-0.330*	-0.528**	-0.146	-0.294*
	(0.194)	(0.209)	(0.166)	(0.176)
Country-level controls:				
Distance from equator	-0.006**	-0.005	-0.008**	-0.009***
	(0.003)	(0.003)	(0.003)	(0.003)
Economic complexity	0.008	0.014	-0.015	-0.021
	(0.029)	(0.037)	(0.043)	(0.041)
Political hierarchies	0.094***	0.088***	0.095**	0.091**
	(0.029)	(0.031)	(0.040)	(0.039)
Ln (per capita GDP)	-0.007	-0.023	-0.002	-0.006
	(0.033)	(0.036)	(0.040)	(0.038)
Genetic distance from the US	0.017	-0.002	0.007	-0.017
	(0.048)	(0.054)	(0.047)	(0.046)
Fraction of population in location who are 1st or 2nd-generation immigrants from same country of origin	3.321***	3.564***	2.938***	3.244***
	(0.515)	(0.644)	(0.558)	(0.530)
Individual controls	Yes	Yes	Yes	Yes
Number of countries	106	105	110	105
Mean (st.dev.) of dependent variable	0.35 (0.48)	0.34 (0.48)	0.31 (0.46)	0.31 (0.46)
Observations	28,860	27,427	30,413	28,342
Original observations	36,082	34,045	38,419	35,639
R^2	0.250	0.263	0.229	0.249

Notes: Mistakes (Dependent variable) and (Missing MSA 1) are corrected. Mistakes (Fraction of the population) and (GDP 2014) cannot be corrected. Except for these corrections, the specification is strictly the same as in GN.

Table 1.A.5: Correction of Table 5 - Speaking a foreign language at home, from 2000 Census (excluding all English-speaking countries)

	Dep. variable: Indicator for speaking an Indigenous language at home						
	All 2nd gen+ individuals	Not living with parents	Living with parents				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Climatic instability	-0.003	0.055	-0.427**	-0.121	-0.131	0.132	0.125
	(0.179)	(0.192)	(0.214)	(0.097)	(0.098)	(0.084)	(0.083)
Father speaks a foreign language				0.503***		0.791***	
				(0.029)		(0.058)	
Mother speaks foreign lang.× Climatic instability						-1.074***	
						(0.240)	
Mother speaks a foreign language					0.511***		0.809***
					(0.030)		(0.056)
Father speaks foreign lang.× Climatic instability							-1.102***
							(0.225)
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of countries	77	77	77	77	77	77	77
Mean (st.dev.) of dependent variable	0.13 (0.34)	0.11 (0.32)	0.25 (0.43)	0.25 (0.43)	0.25 (0.43)	0.25 (0.43)	0.25 (0.43)
Observations	2,294,290	1,994,785	299,505	229,166	280,396	229,166	280,396
Original observations	3,343,097	2,915,673	427,424	330,226	400,062	330,226	400,062
R^2	0.332	0.298	0.420	0.588	0.602	0.592	0.606

Notes: Mistakes (Missing MSA 2) and (English) are corrected. Mistake (Living with parents) cannot be corrected. Except for these corrections, the specification is strictly the same as in GN.

Table 1.A.6: Correction of Table 7 - Whether Indigenous populations of the United States speak their traditional language at home: individual-level estimates (not excluding observations for which the metropolitan area is missing)

	Dep. variable: Indicator for speaking a foreign language at home						
	All 2nd gen+ individuals	Not living with parents	Living with parents				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Climatic instability	-1.290***	-1.282***	-1.284***	-0.712***	-0.547***	-0.356**	-0.193
	(0.424)	(0.437)	(0.357)	(0.179)	(0.160)	(0.152)	(0.149)
Father speaks an Indigenous language				0.448***		0.677***	
				(0.051)		(0.063)	
Father speaks an Indigenous lang.× Climatic instability						-0.902***	
						(0.169)	
Mother speaks an Indigenous language					0.471***		0.694***
					(0.046)		(0.051)
Mother speaks an Indigenous lang.× Climatic instability							-0.878***
							(0.196)
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of ethnic groups	19	19	19	19	19	19	19
Number of clusters (grid cells)	18	18	18	18	18	18	18
Mean (st.dev.) of dependent variable	0.21 (0.41)	0.21 (0.41)	0.23 (0.42)	0.31 (0.46)	0.29 (0.46)	0.31 (0.46)	0.29 (0.46)
Observations	81,221	67,964	13,257	6,166	9,557	6,166	9,557
Original observations	128,005	79,235	48,770	25,794	35,403	25,794	35,403
R^2	0.396	0.399	0.395	0.557	0.564	0.564	0.570

Notes: Mistakes (Language), (Children 0 to 4) and (Employment) are corrected. Mistake (Missing MSA) is not corrected. Mistake (Living with parents) cannot be corrected. Except for these corrections, the specification is strictly the same as in GN.

Table 1.A.7: Correction of Table 7 - Whether Indigenous populations of the United States speak their traditional language at home: individual-level estimates (excluding observations for which the metropolitan area is missing)

	Dep. variable: Indicator for speaking a foreign language at home						
	All 2nd gen+ individuals	Not living with parents	Living with parents				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Climatic instability	-0.364	-0.380	-0.140	-0.172	0.064	0.040	0.107
	(0.268)	(0.275)	(0.191)	(0.116)	(0.072)	(0.095)	(0.093)
Father speaks an Indigenous language				0.362^{***}		0.628^{***}	
				(0.062)		(0.097)	
Father speaks an Indigenous lang.× Climatic instability						-1.041^{**}	
						(0.483)	
Mother speaks an Indigenous language					0.408^{***}		0.456^{***}
					(0.058)		(0.110)
Mother speaks an Indigenous lang.× Climatic instability							-0.185
							(0.438)
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of ethnic groups	19	19	19	18	19	18	19
Number of clusters (grid cells)	18	18	18	17	18	17	18
Mean (st.dev.) of dependent variable	0.10 (0.30)	0.10 (0.31)	0.09 (0.29)	0.15 (0.35)	0.13 (0.34)	0.15 (0.35)	0.13 (0.34)
Observations	34,092	29,500	4,592	1,791	2,903	1,791	2,903
Original observations	128,005	79,235	48,770	25,794	35,403	25,794	35,403
R^2	0.381	0.386	0.416	0.533	0.562	0.542	0.562

Notes: Mistakes (Language), (Children 0 to 4), (Employment) and (Missing MSA) are corrected. Mistake (Living with parents) cannot be corrected. Except for these corrections, the specification is strictly the same as in GN.

Chapter 2

Migration and the epidemiological approach: time and self-selection into foreign ancestries matter

This chapter is joint work with Simone Bertoli, Jordan Loper, and Èric Roca Fernández, and published in the *Journal of Development Economics* (doi:10.1016/j.jdeveco.2025.103505).

Replication files: doi:10.3886/E224182V1.

2.1 Introduction

Understanding the deep-rooted determinants of economic development is a fundamental question that has garnered considerable attention among economists. This has sparked a burgeoning literature, revealing how contemporary outcomes such as economic development, inequality, and individual behaviors are influenced by the persistent characteristics of a distant past. Furthermore, this literature has also provided abundant evidence of the persistence of cultural and genetic traits.

Isolating the long-term causal effect of such traits is a vivid challenge due to the likely emergence of numerous credible confounders. To address this, economists have extensively relied on the so-called *epidemiological approach*, an expression introduced by Raquel Fernández.[1]

This method hinges on a criterion that unambiguously assigns each individual to one ancestral origin. The literature commonly employs four main proxies, with this key choice being often subject to binding data constraints: country of birth (Antecol, 2000; Luttmer and Singhal, 2011), maternal or paternal country of birth (e.g., Fernández, 2007; Giuliano, 2007; Galor and Savitskiy, 2018; Galor et al., 2016, 2020; Giuliano and Nunn, 2021; Ek, 2024), foreign ancestry (e.g., Antecol, 2000; Guiso et al., 2006; Fernández and Fogli, 2006; Algan and Cahuc, 2010; Alesina et al., 2015; Galor and Özak, 2016; Giavazzi et al., 2019; Arbatlı et al., 2020; Giuliano and Nunn, 2021; Ek, 2021; Galor et al., 2023), and language (e.g., Alesina et al., 2003; Desmet et al., 2017; Giuliano and Nunn, 2021).[2] Yet, none of these proxies, unequivocally captures an individual's ancestry or identifies a single foreign origin.[3] Parents might be from different countries, a languages can be spoken in several countries, and multiple ancestries

[1]"The essence of what I call the epidemiological approach is the attempt to identify the effect of culture through the variation in economic outcomes of individuals who share the same economic and institutional environment, but whose social beliefs are potentially different." (Fernández, 2011).

[2]To appropriately link these proxies to outcome variables, researchers often need to account for major past population movements, such as those related to the Columbian Exchange; for instance, Giuliano and Nunn (2018, 2021) rely on information on ethnic homelands from the Ethnographic Atlas (Murdock, 1967), languages from Ethnologue and gridded data on the distribution of the population to aggregate data specific to each ethnic homeland at the country level, while Galor et al. (2023) use data from Putterman and Weil (2010).

[3]For example, in the case of repeated migration across different generations, the country of birth of an individual is not necessarily informative about his or her own ancestry; consider, for instance, someone born in Argentina from parents who migrated out of Germany after WWII, and who then moved to the US; this individual would be classified as Argentinian by birth, yet the cultural background could be primarily German.

can be reported. Consequently, these proxies can link an individual to diverse origins—for example, someone born in Spain might speak French at home and report an Italian ancestry. These distinctions are crucial for econometric analysis because different origins often entail different norms, cultural values, or genetic factors.

The multifaceted nature of one's own cultural identity, the fact that migrants from all origin countries are likely to be self-selected along some cultural traits,e.g., long-term orientation and attitude towards risk,and cultural homophily, which tends to make migrants in a given destination more similar than populations at origin are (see, for instance, Antecol, 2000; Alesina et al., 2015, on this), all strongly suggest that the epidemiological approach is exposed to an attenuation bias. Indeed, the widely cited review of the literature by Fernández (2011) observes that:

> "It should be noted explicitly that the epidemiological approach is biased towards finding that culture does not matter. As mentioned previously, the fact that parents are only one source of cultural transmission among many and that they may have cultural attitudes that differ from the average ones in the country of ancestry implies that one is more likely to rule the cultural proxy insignificant."

Thus, the epidemiological approach in comparative development is often presumed to suffer from attenuation bias, as suggested by the quote from Fernández (2011). This presumption, however, may be unwarranted when relying on *self-reported* ancestry data—an approach increasingly common in studies using US census or similar surveys.[4] The initial move towards using ancestry data was motivated by concerns about first-generation immigrants, who constitute a self-selected group unrepresentative of their origin populations, potentially biasing analyses due to unobserved traits.[5] However, with the exception of Galor et al. (2016, 2020), the literature has largely overlooked that similar selection concerns—related to unobservable factors—also apply to natives of foreign ancestry. Importantly, these selection issues in ancestry data can counteract the presumed attenuation bias. Two key factors contribute to this:

[4] The shift towards ancestry data is often driven by data availability: information on paternal and maternal birthplaces is no longer collected in censuses since 1980 nor in surveys like the American Community Survey (ACS) or the Current Population Survey (CPS) except for the subset of respondents co-residing with their own parents, or for the March Supplement of the CPS.

[5] For instance, Alesina and Giuliano (2015) argue that second-generation immigrants form "a more appropriate sample than first-generation immigrants because issues of disruption and selection due to migration are more attenuated." (pp. 903-4).

(i) a large portion of the native population does not report a traceable foreign ancestry, meaning their ancestry cannot be readily linked to a specific foreign country; and (ii) substantial unobserved variation exists in the time elapsed since ancestral migration among those who *do* report foreign ancestry.[6]

Regarding point (i), data from the 2000 census indicate that nearly 20% of the population reported *no ancestry* at all (Brittingham and de la Cruz, 2004). Furthermore, an even larger fraction declared ancestries categorized as native (e.g., African American 8.8%, American Indian 2.8%, American 7.2%) or supranational (e.g., Hispanic 0.9%, European 0.7%), or ancestries with links to multiple current nations (e.g., Basque, Czechoslovakian, Prussian or Yugoslavian). In total, our analysis reveals that for 47.2% of the native population, a connection to a specific foreign country of origin through self-reported ancestry cannot be established. Several factors contribute to this lack of reported foreign ancestry. Natives whose ancestors came from a given foreign country might not report an ancestry in that country simply because they lack, as observed by Farley (1991), the necessary "factual knowledge" about the history of their family, or they might deliberately decide not to do so.[7] A legitimate presumption is that, if we consider a group of natives whose ancestors came to the United States from a given foreign country, those choosing to report the corresponding foreign ancestry have a greater attachment to their ancestral culture than those who chose not to report that ancestry. This, in turn, implies that the individuals entering into the estimation sample are selected in a way that likely increases the chances of finding a significant effect. The intensity of self-selection into reporting a foreign ancestry might also vary across ancestral countries, and be stronger for ancestries towards which natives have, on average, a more negative attitude because of the perception of a greater cultural distance with the United States,[8] but this is not a necessary

[6]As Ek (2021) notes, census data limitations in tracking immigration history beyond the first generation mean that "people whose ancestors arrived 20 or 200 years ago are treated identically." (p. 23).

[7]Related to this, Duncan and Trejo (2011, 2017) and Antman et al. (2016, 2023) provide evidence of the substantial incidence of what they describe as "ethnic attrition", i.e., the propensity of individuals living in the United States who are first to third-generation immigrants from Latin American or Asian countries *not* to report that they are Hispanics or Asians when answering to questions related to race or ethnicity; ethnic attrition increases with the time elapsed since ancestral migration, it greatly varies across origin countries, and across individuals, e.g., higher-educated individuals with a Mexican ancestry are significantly less likely to identify themselves as Hispanics.

[8]As the data do not provide any objective information on the origin of one's own ancestors, e.g., the countries of birth of the grandparents or of the great-grandparents, providing evidence on the origin-specific

condition for a threat to identification.

As far as point (*ii*) is concerned, the different time profiles of past migration flows to the United States from various countries imply that the time elapsed since ancestral migration varies across groups of natives reporting different foreign ancestries.[9] As cultural assimilation can increase with time through horizontal interactions (Bisin and Verdier, 2001), longer elapsed time may lead to greater convergence towards US cultural norms. For example, Guiso et al. (2006) posit that differences in redistribution attitudes among natives of different ancestries may reflect varying times since ancestral migration.[10] Giavazzi et al. (2019) provides supporting empirical evidence across various cultural traits. Crucially, census and American Community Survey data do not capture this time dimension, potentially confounding econometric analyses.[11] This unobserved heterogeneity in migration timing can lead to overestimating cultural persistence if countries with culturally distant norms have more recent migration histories to the US. Similarly, it can bias analyses of determinants of cultural traits if these determinants are correlated with migration timing —a likely scenario given the evolving origins of US immigration (Abramitzky and Boustan, 2017) and the smooth geographical distribution of factors often studied in the epidemiological approach.[12]

To address these issues, our analysis proceeds in four interrelated stages. First, we outline

incidence of not reporting a given ancestry is unfeasible; Fulford et al. (2020) provide an aggregate measure of the ancestral countries of origin for the population of each county in the United States; their approach cannot be deployed at the individual-level.

[9]This temporal heterogeneity is intertwined with point (*i*), as the persistence of foreign ancestry reporting across generations varies by origin; for example, third-generation descendants of Irish immigrants in the US might be less likely to still report and Irish ancestry compared to third-generation descendants of Greek immigrants, due to differing patterns of assimilation and cultural retention across these groups.

[10]As Guiso et al. (2006) argue, “Americans with British, North European or German ancestors derive from earlier immigrants; hence, more generations were raised in the United States and forged by its culture, absorbing the belief that success is mostly determined by individual actions, which makes government intervention highly undesirable” (p. 41).

[11]To our knowledge, only a few studies, notably Algan and Cahuc (2010), Alesina et al. (2015), and Giavazzi et al. (2019), address this by using data from the General Social Survey (GSS), which includes both self-reported ancestry and information on the number of parents and grandparents of the respondent that were foreign-born, which allows differentiating between native individuals of different migrant generations, i.e., second, third and (at least) fourth-generation migrants.

[12]Conley and Kelly (2024) have recently demonstrated that spatial trends and spatial autocorrelation in the data can lead to a substantial underestimation of the standard errors, and thus to an ensuing tendency to unduly reject the null hypothesis of no effect in the literature on cultural persistence.

the canonical epidemiological specification and common methods used to define foreign origin. Second, we develop a theoretical framework to model how self-selection in foreign ancestry reporting and variations in time since migration can bias estimates from this specification. Third, using census data (Ruggles et al., 2023), we empirically document: the extent of non-reporting of foreign ancestry; variations in ancestry reporting over time (in the 1980, 1990 and 2000 Censuses); correlates of non-reporting Mexican ancestry among people that identified as Hispanic-Mexican; and the origin-specific *approximate* values for time since ancestral migration for a set of 116 origin countries using historical Censuses.[13] In addition, we describe and compute proxy measures of the time since ancestral migration using current Censuses. These include the share of individuals claiming an ancestry who were born in the corresponding ancestral country, i.e., first-generation migrants, the spatial concentration of ancestries, and the use of ancestral languages. We argue that such variables should decline with time since migration,[14] something we validate against historical immigration data. Fourth, we illustrate the empirical relevance of the concerns we raise by revisiting findings from two influential studies in the literature: Fernández and Fogli (2006) and Giuliano and Nunn (2021). We show that the cultural persistence of fertility norms in Fernández and Fogli (2006) weakens considerably when considering fourth-generation migrants, who represent the majority of the sample. For Giuliano and Nunn (2021), we find that the estimated impact of historical climate variability on ancestral language use is particularly influenced by natives with Spanish-speaking ancestries, i.e., groups characterized by more recent migration histories. Our aim is here is to highlight the empirical relevance of the potential confounding effects of unobserved heterogeneity when using data on ancestry, and not to call into question the results of the two papers; the findings by Fernández and Fogli (2006) still hold for second and third-generation migrants, and Giuliano and Nunn (2021) analyze data from a variety of sources, which mostly do not rely on self-reported ancestry to identify one's own origin.

[13]Mexican is the only ancestry that allows such exercise due to the conceptual overlap between two questions in the Census. The first is meant to elicit Hispanic origins while the second concerns one own's ancestral identity. Of the 116 origin countries, only 114 are *foreign* and, thus, useful in our analysis. The remaining two are the US and Puerto Rico.

[14]With respect to spatial concentration, this can be due, for instance, to the fact that the importance of migrant networks (see, for instance Patel and Vella, 2013) declines over time, because the relevance of the local labor demand conditions that shaped the initial spatial distribution of the immigrants fades away, or to the fact that natives of immigrant descent move away from more expensive locations where their foreign-born ancestors tended to concentrate (Albert and Monras, 2022).

This paper offers two primary contributions to the literature on comparative development. First, we provide a theoretical framework to analyze the specific identification challenges posed by using self-reported ethnic or ancestral origin in the epidemiological approach. This is crucial given the limited availability of alternative data for natives.[15] Our framework demonstrates that even with origin-invariant non-reporting of ancestry and after controlling for time since migration, the very act of restricting analysis to *those reporting* foreign ancestry can introduce bias, increasing the likelihood of finding significant effects. Second, we empirically document substantial heterogeneity in the US native population across different self-reported ancestries. We show that our readily available proxies can be used as diagnostic tools to assess the sensitivity of results to origin countries that are outliers in terms of time since migration, helping to evaluate potential confounding influences on estimated effects.

The remainder of the paper is structured as follows: Section 2.2 describes the canonical equation and identification challenges related to different origin criteria; Section 2.3 presents our theoretical framework and derives theoretically consistent specifications to address self-selection and time-since-migration biases; Section 2.4 details our data sources; Section 2.5 presents descriptive empirical analyses of census data, documenting patterns in ancestry reporting; Section 2.6 outlines how we build our proxies of unobserved heterogeneities across origins and test their validity; Section 2.7 provides a guideline for researcher interested in addressing biases uncovered in this paper, as well as illustrative examples of our concerns' empirical relevance; and Section 2.8 concludes.

2.2 The epidemiological approach

Consider individuals, denoted by i, residing in location k within a country d. Let o represents the individual's origin, as defined subsequently. Assume cross-sectional data are collected at time T. The canonical regression in the epidemiological approach can be expressed as (omitting subscripts for country and time for conciseness):[16]

$$y_{iok} = \alpha w_o + \boldsymbol{\beta}'\boldsymbol{x}_o + \boldsymbol{\gamma}'\boldsymbol{x}_i[+\boldsymbol{\phi}'\boldsymbol{x}_{ok} + d_k + \lambda f(t_i)] + \epsilon_{iok} \tag{2.1}$$

[15]In contrast to Giavazzi et al. (2019), we specifically explore the implications of self-reported nature of foreign ancestry.

[16]For clarity and brevity, we omit subscripts denoting the specific country of destination (d) and the data collection time (T) as these are held constant.

where y_{iok} is the outcome variable, w_o is the origin-specific variable of interest, $\boldsymbol{x}_o$ and $\boldsymbol{x}_i$ are vectors of origin-level and individual-level control variables, respectively. The terms $\boldsymbol{x}_{ok}$ and d_k represent origin-location controls and location fixed effects; t_i denotes the time since ancestral migration for individual i, possibly transformed by a function $f(t_i)$; and ϵ_{iok} is the error term. Note that the inclusion of $f(t_i)$, $\boldsymbol{x}_{ok}$, and d_k in Eq. (2.1) is not systematic in the literature, something our notation emphasizes.[17] Crucially, the origin-level control vector, $\boldsymbol{x}_o$, is intended to capture characteristics of the population in the origin country o, and thus *excludes* variables derived from individuals of origin o residing in the destination country.

The coefficient of interest in Eq. (2.1) is α, with identification predicated on the assumption that:

$$E(w_o \times \epsilon_{iok} | \boldsymbol{x}_o, \boldsymbol{x}_i, [\boldsymbol{x}_{ok}, d_k, f(t_i)]) = 0.$$

However, this assumption can be violated. For instance, bias arises if: (i) the average time since ancestral migration varies systematically across origins (and is correlated with w_o) when $\lambda \neq 0$, or (ii) unobserved determinants of y_{iok} lead to differential self-selection by origin, and the intensity of this selection correlates with w_o. The relevance of these threats hinges on the chosen criterion for assigning origin. We now consider three common criteria and their associated sample selections: country of birth, parental country of birth, and foreign ancestry.

2.2.1 Country of birth

This criterion, restricting the sample to first-generation immigrants, mitigates concerns related to point (i).[18] However, it conversely amplifies concerns regarding point (ii). This is because immigrants from origin country o likely differ from non-migrants ("stayers") along numerous dimensions (e.g., Antecol, 2000; Alesina et al., 2015; Alesina and Giuliano, 2015).[19] For instance, migrants may exhibit positive self-selection for traits like long-term orientation, and negative self-selection for risk aversion. Consequently, the expected value of the error term ϵ_{iok} in Eq. (2.1) is likely to vary systematically with origin o.

[17] For example, Giuliano and Nunn (2021) and Arbatlı et al. (2020) include location fixed effects d_k, with Giuliano and Nunn (2021) also incorporating origin-location controls $\boldsymbol{x}_{ok}$; however, individual time since ancestral migration t_i is generally not accounted for.

[18] These concerns can be further lessened by including years since migration in the vector $\boldsymbol{x}_i$ in Eq. (2.1).

[19] See, for example, Jaeger et al. (2010) and Beck Knudsen (2022) for empirical evidence of migrant selection on cultural traits.

2.2.2 Parental country of birth

Defining origin o by parental country of birth limits the sample to natives with at least one foreign-born parent, i.e., second-generation immigrants. While this approach increases concerns about unobserved heterogeneity in t_i —the time elapsed since parental migration to country d[20]— it can alleviate concerns related to point (ii). Non-random selection bias may be reduced when moving from first to second-generation migrants (see Giavazzi et al., 2019), especially under imperfect vertical transmission of cultural traits (Bisin and Verdier, 2001). If parental country of birth data is unavailable, this origin criterion requires further restricting the sample to individuals co-residing with parents, as in Galor et al. (2016, 2020), potentially introducing a distinct and new source of sample selection bias.

2.2.3 Country of foreign ancestry

Using self-reported (foreign) ancestry as the origin criterion restricts the estimation sample for Eq. (2.1) to natives who explicitly *declare* a foreign ancestral origin that is valid, in the sense that it can be unambiguously connected to a foreign country. This approach intensifies concerns regarding unobserved heterogeneity in time since ancestral migration, t_i, as census data lacks information on the generational distance between individual i and their immigrant ancestors.

While seemingly alleviating concerns related to point (ii), as longer times since migration might dilute initial migrant selection bias,[21] this dilution may not hold for those who *do* report foreign ancestry who are likely more attached to their ancestral identity.

Indeed, for natives of immigrant origin, ancestry reporting reflects a choice based on perceived costs and benefits of identity (Akerlof and Kranton, 2000) and parental efforts in identity transmission (Bisin and Verdier, 2001).[22] Consequently, individuals with weaker ties to their origin may choose to omit ancestry reports or declare "American" or "general heritage"

[20]For example, a survey might include an 80-year-old born in the US whose parents migrated 50 years prior, and a young individual born to recently arrived parents; the difference in time since parental migration can thus be substantial and close to a century.

[21]Indeed, as t_i increases, horizontal cultural transmission within country d becomes more influential than vertical transmission of ancestral culture from origin o.

[22]Antman and Duncan (2023) present a similar framework for analyzing identity choice among Native Americans in the United States.

ancestries (e.g., Hispanic) instead of their specific foreign ancestry, thus being excluded from the analysis. Negative attitudes among natives towards certain origins strongly drive self-selection into the sample, particularly among recent immigrant groups. Historically, many immigrant groups faced prejudice in the US. Consider these examples regarding Irish and Chinese immigrants from Fasani et al. (2019):

> "Just outside the US borders were 'hordes of Wild Irishmen ... the turbulent and disorderly of all the world [who come to the United States in order to] distract our tranquillity.' (Massachusetts Representative Harrison Gray Otis, 1797). [...] Chinese immigrants were 'morally the most debased people on the face of the earth' (Connecticut Senator Orville Platt, 1882), who 'bring every character of vice ... [and would be] injurious in every sense of the world' (Texas Senator Samuel Bell Maxey, 1882)."

Historically, Italians also faced long-lasting negative attitudes (see, for instance, Fouka et al., 2021), as individuals originating from Latin American countries still experience in the United States (see, for instance, Chavez, 2013). Therefore, natives of immigrant origin who *do* report a foreign ancestry —despite potential negative perceptions— likely possess a stronger attachment to this (potentially costly) identity compared to those who do not. Consequently, their degree of nonrandom selection bias is likely greater than that of individuals reporting ancestries favorably regarded by natives.[23]

As the implications of using country of foreign ancestry data remain underexplored in the existing literature, the following section develops a simplified theoretical framework. This framework aims to isolate and analyze the combined influence of self-selection in ancestry reporting and the time elapsed since ancestral migration on estimates derived from Eq. (2.1).

2.3 A simple theoretical framework

To illustrate our concerns, consider two distinct foreign countries of origin, $o \in \{A, B\}$, each characterized by different baseline values for the cultural trait of interest, w_o. The two countries might be, for instance, Germany and Mexico, two of the main countries of foreign ancestry for the native population in the US, and that often represent large and shares of the various estimation samples. For simplicity, and without loss of generality, let w_o represent the average

[23] Duncan and Trejo (2011) and Duncan and Trejo (2017) show how "ethnic attrition" varies across origin countries.

value of the outcome y in the origin country o (i.e., $w_o = \overline{y}_o$), and assume that $\overline{y}_A < \overline{y}_B$. For analytical tractability, we further simplify the population of natives with ancestry from either country A or B to consist of only two generations j of immigrants: second-generation (natives with foreign-born parents) and third-generation (natives with native-born parents and foreign-born grandparents). Thus, $j \in \{2, 3\}$. Let $s_o \in [0, 1]$ represent the proportion of the native population with ancestors in country o that is second-generation.

Let y_{io}^j be the outcome for native individual i with ancestry from country o, where $j \in \{2, 3\}$ indicates the generation (second or third). We assume that for second-generation immigrants ($j = 2$), the outcome is directly determined by the origin norm: $y_{io}^2 = \overline{y}_o$.[24] For third-generation immigrants ($j = 3$), we posit a probabilistic cultural transmission: with probability p_o, individual i adopts the destination country norm, $y_{io}^3 = \overline{y}$; and with probability $(1 - p_o)$, maintains the origin country norm, $y_{io}^3 = \overline{y}_o$.[25] To capture this probabilistic adoption, define a dummy variable D_{io}^3 that equals 1 if individual i (third-generation, ancestry o) adopts the destination country norm, and 0 otherwise.

Similarly, let F_{io}^j be a dummy variable equal to 1 if native individual i with ancestry in country o reports this ancestry, and 0 otherwise.[26] We assume that second-generation immigrants consistently report their foreign ancestry ($F_{io}^2 = 1$). For third-generation immigrants, let $q_o \in [0, 1]$ denote the probability that an individual of ancestry o will report this ancestry ($F_{io}^3 = 1$).

To illustrate the implications of our framework, let us first assume that $q_A = q_B = 1$. This means all natives of immigrant origin, regardless of generation, report their foreign ancestry.

Consider a regression of y_{io}^2 on w_o (approximated by $\overline{y}_o$) using only *second-generation* immigrant data. The estimated coefficient, $\widehat{\alpha}^2$, will be:

$$\widehat{\alpha}^2 = \frac{\overline{y}_B^2 - \overline{y}_A^2}{\overline{y}_B - \overline{y}_A} = 1 \tag{2.2}$$

This result is derived from the assumption that second-generation immigrants fully maintain

[24] This simplification is made to isolate the effects of heterogeneity in time since ancestral migration and self-selection in ancestry reporting by initially abstracting from direct migrant selection effects.

[25] A zero-mean error term could be added to y_{io}^2 and y_{io}^3 to incorporate individual-level unobserved heterogeneity, but is omitted here for simplicity.

[26] For simplicity, we consider a binary choice: reporting the actual ancestry or reporting no ancestry. In survey data, individuals may misreport ancestry, which our model does not account for. However, reporting a supranational or USA ancestry is treated as equivalent to not reporting any foreign ancestry, as these responses typically lead to exclusion from estimation samples.

the origin country norm, i.e., $\overline{y}_o^2 = \overline{y}_o$. Therefore, the regression directly recovers the expected unit relationship when using data solely from second-generation immigrants.

Now consider a regression of y_{io}^3 on w_o using only *third-generation* immigrant data. The estimated coefficient, $\widehat{\alpha}^3$, will be:[27]

$$\widehat{\alpha}^3 = \frac{\overline{y}_B^3 - \overline{y}_A^3}{\overline{y}_B - \overline{y}_A} = \frac{[p_B\overline{y} + (1-p_B)\overline{y}_B] - [p_A\overline{y} + (1-p_A)\overline{y}_A]}{\overline{y}_B - \overline{y}_A} \tag{2.3}$$

where $\overline{y}_o^3 \equiv p_o\overline{y} + (1-p_o)\overline{y}_o$ represents the average outcome for third-generation immigrants from country o. Here, $\widehat{\alpha}^3$ reflects a potentially attenuated relationship due to the partial convergence of third-generation immigrants towards the destination country norm.

Instead, consider pooling data from both second and third-generation immigrants *without* controlling for generation. The estimated coefficient, $\widehat{\alpha}^{\text{pooled}}$, from a regression of y_{io} on w_o in this pooled sample will be a weighted average:

$$\widehat{\alpha}^{\text{pooled}} = \frac{s_B\overline{y}_B^2 + (1-s_B)\overline{y}_B^3 - [s_A\overline{y}_A^2 + (1-s_A)\overline{y}_A^3]}{\overline{y}_B - \overline{y}_A} \tag{2.4}$$

In the absence of self-selection in ancestry reporting (i.e., if we still assume $q_A = q_B = 1$), $\widehat{\alpha}^{\text{pooled}}$ simply reflects a compositional effect determined by the generational shares, s_A and s_B, within each ancestry group.[28]

Under the plausible assumption that the destination country norm, $\overline{y}$, lies between the origin country norms, i.e., $\overline{y} \in [\overline{y}_A, \overline{y}_B]$[29], it follows that $\widehat{\alpha}^3$ from Eq. (2.3) is bounded above by $\widehat{\alpha}^2 = 1$, meaning $\widehat{\alpha}^3 \leq \widehat{\alpha}^2 = 1$. Similarly, $\widehat{\alpha}^{\text{pooled}}$ from Eq. 2.4 is bounded between $\widehat{\alpha}^3$ and $\widehat{\alpha}^2$, i.e., $\widehat{\alpha}^{\text{pooled}} \in [\widehat{\alpha}^3, \widehat{\alpha}^2]$.

If we then consider regressions limited to *third-generation* immigrants *who report foreign ancestry*, the estimated coefficient becomes:

$$\widehat{\alpha}_F^3 = \frac{\overline{y}_{B,F}^3 - \overline{y}_{A,F}^3}{\overline{y}_B - \overline{y}_A} = \frac{[p_B^F\overline{y} + (1-p_B^F)\overline{y}_B] - [p_A^F\overline{y} + (1-p_A^F)\overline{y}_A]}{\overline{y}_B - \overline{y}_A} \tag{2.5}$$

[27] If the probability of adopting the destination norm is uniform across origins, i.e., $p_A = p_B = p$, then $\widehat{\alpha}^3 = (1-p)\widehat{\alpha}^2 = 1-p$.

[28] Specifically, if the second-generation shares are equal across origins ($s_A = s_B = s$), then $\widehat{\alpha}^{\text{pooled}}$ is a simple weighted average of the generation-specific coefficients: $\widehat{\alpha}^{\text{pooled}} = s\widehat{\alpha}^2 + (1-s)\widehat{\alpha}^3$.

[29] This assumption seems reasonable for the US context. The initial US population comprised diverse groups migrants (e.g., British, Irish, German). Subsequent immigration likely shifted societal norms to a point intermediate between those of earlier settlers and later arrivals.

where $\overline{y}^3_{o,F} \equiv p^F_o \overline{y} + (1 - p^F_o)\overline{y}_o$ represents the average outcome for third-generation immigrants from country o *who report foreign ancestry*. From the bounds established earlier, it follows that $\widehat{\alpha}^3_F \in [\widehat{\alpha}^3, \widehat{\alpha}^2]$.

In the extreme case where the correlation between assimilation and ancestry reporting is perfectly negative, i.e., $\text{corr}(D^3_{io}, F^3_{io}) = -1$, then $p^F_o = 0$. This means that all third-generation individuals who adopt the destination country norm *cease* reporting foreign ancestry. In this scenario, the estimated coefficient $\widehat{\alpha}^3_F$ becomes equal to $\widehat{\alpha}^2 = 1$, and importantly, the coefficient from a pooled regression *also restricted to individuals reporting foreign ancestry*, denoted $\widehat{\alpha}^{\text{pooled}}_F$, also equals $\widehat{\alpha}^2 = 1$. Thus, self-selection in foreign ancestry reporting can lead to an *overestimation* of cultural persistence, both when analyzing third-generation immigrants specifically and in pooled samples. This overestimation occurs because individuals who assimilate towards the destination norm are systematically excluded from the estimation sample when conditioning on reporting foreign ancestry. In the epidemiological approach, conditioning on reported ancestry is necessary, as origin-country norms and values can only be meaningfully attributed to individuals identified as having a foreign ancestry. Consequently, the upward bias we describe may counteract any inherent attenuation bias in the epidemiological methodology.

Furthermore, if an origin country's norm closely aligns with the destination country's norm (e.g., if $\overline{y}_A = \overline{y}$),[30] then even without self-selection in ancestry reporting among individuals of origin A, we would find that $\widehat{\alpha}^3_F = \widehat{\alpha}^{\text{pooled}}_F = \widehat{\alpha}^2$. This is because for individuals of origin A, the origin and destination norms are effectively identical, resulting in $\overline{y}^3_{A,F} = \overline{y}^3_A$. Consequently, self-selection in foreign ancestry reporting poses a greater concern for origin countries whose norms diverge more significantly from those of the destination. In other words, if the two countries A and B are Germany and Mexico, self-selection into reporting a Mexican ancestry would be a more pressing concern.

A theoretically consistent specification, addressing the limitations discussed, would modify Eq. (2.1) to include time since ancestral migration, t_i, as follows:[31]

$$y_{iok} = \alpha_1 w_o + \alpha_2(w_o \times f(t_i)) + \boldsymbol{\beta}'\boldsymbol{x}_o + \boldsymbol{\gamma}'\boldsymbol{x}_i + \boldsymbol{\phi}'\boldsymbol{x}_{ok} + d_k + \lambda f(t_i) + \epsilon_{iok} \tag{2.6}$$

The key distinction between Eq. (2.1) and our theoretically refined Eq. (2.6) is the incorpora-

[30]For many cultural traits, US norms were shaped by early European migrants. Thus, cultural distance between the US and countries like the UK, Germany, or Ireland may be minimal.

[31]We thank an anonymous referee for suggesting this specification.

tion of $f(t_i)$ both as an additive term and interacted with w_o. Estimating Eq. (2.6), however, presents data challenges. Direct measurement of t_i is rare; to our knowledge, only the GSS provides this information, unlike larger datasets such as the census, the ACS or the CPS. The previous problem can be partially mitigated with geographically aggregated controls, such as continental fixed effects, which help reduce origin-level heterogeneity. However, this approach cannot address individual-level variation in migration timing. Additionally, Eq. (2.6) does not account for potential biases stemming from self-selection in foreign ancestry reporting. Data limitations remain a significant constraint, and the inherent sample selection must be considered when interpreting econometric results.

2.4 Data Sources

Our analysis uses data primarily from the US population censuses, focusing on the 5% sample from 2000 (Ruggles et al., 2023), with supplementary data from the 5% samples of the 1980 and 1990 censuses for specific analyses.[32] Earlier censuses (1850-1970) are also used to validate our proxies for time since ancestral migration, constructed using 2000 census data.

2.4.1 US Population Censuses

US population censuses are a rich source of individual-level data, encompassing demographic and socioeconomic characteristics of all US residents, including undocumented immigrants. Along with the ACS, they are a cornerstone of epidemiological research on cultural persistence and the long-term impact of ancestry (e.g., Fernández and Fogli, 2006; Giuliano and Nunn, 2021; Ek, 2021; Galor et al., 2020, 2023). The US, with its history of diverse immigration, provides an ideal setting to study these phenomena.

For identifying individual origin, US census data offers four key variables: birthplace, ancestry, language spoken at home (other than English), and parental country of birth. Birthplace is universally available in the census. Parental country of birth is limited to individuals co-residing with parents, resulting in a smaller and potentially self-selected sample. Home language is less suitable for origin identification due to endogeneity and ambiguity.[33] Therefore,

[32]While other datasets, like the March Supplement of the CPS and include parental country of birth, their smaller sample sizes limit their utility for constructing reliable origin-specific variables, which are crucial for our analysis. The ACS includes similar variables to the census.

[33]Home language is endogenous to migration and does not uniquely identify origin countries (e.g., Spanish

we primarily focus on birthplace and ancestry as our main origin indicators for the subsequent analysis, selectively utilizing the remaining variables when appropriate.

2.4.2 Main variables used in the analysis

Our analysis employs four key variables used in the literature to determine individual origin: respondent's birthplace, self-reported ancestry, parental birthplace (maternal or paternal), and language (other than English) spoken at home. In the following subsections, we detail each variable and how we constructed crosswalks to standardize origin classifications across these measures.

Birthplace

The birthplace variable (`bpld`), available in censuses, identifies either the state of birth for natives or the country of birth for foreign-born individuals. It is more detailed than some aggregated birthplace codes (`bpl`), providing sub-regional identifiers where available. For instance, in the 2000 census, distinct codes exist for birthplaces like Aruba or Madeira, offering enhanced geographical specificity.

Parental Birthplace

Parental birthplace, specifically maternal and paternal, is derived by linking the respondent's birthplace information to that of co-resident mothers and fathers within the census data. Census variables identify co-resident parents (variables `momloc` and `poploc`), allowing construction of parental birthplace for the subset of co-residing respondents.

Self-Reported Ancestries

The ancestry question was introduced in the 1980 census, replacing parental birthplace and mother tongue questions. It asks respondents to specify the "ancestry group with which [she] identifies." Question wording evolved slightly from 1980 ("What is this person's ancestry?") to 1990 and 2000 ("What is this person's ancestry or ethnic origin?").[34] Participation is

speakers can originate from various countries). Census data on mother tongue is available only until 1970.

[34] See questionnaire text at https://usa.ipums.org/usa-action/variables/ANCESTR1#questionnaire_text_section.

voluntary, and respondents may choose not to report any ancestry. Census questionnaires include examples of ancestries (e.g., "Italian, Jamaican, African Am., Cambodian...").[35]

Respondents could report up to two ancestries in 1990, 2000, and the ACS, recorded in variables `ancestr1` and `ancestr2`. The Census Bureau provides no explicit ordering criteria. In 1980, up to three ancestries were recorded. When reporting three ancestries, the Census Bureau recorded them in alphabetical order in `ancestr1`, leaving `ancestr2` empty.[36] Thus, reporting a third ancestry in 1980 could alter the alphabetical order of the first two in `ancestr1`.[37] Reported ancestries encompass native (e.g., American, African American, American Indian) and foreign origins. Foreign ancestries can specify a country (e.g., Italian), a sub-national region (e.g., Sicilian), or a supranational group (e.g., European, Hispanic).[38] The "American" ancestry is uniquely treated by the Census Bureau: it is valid only if no other ancestry is reported and is otherwise discarded. Consequently, the first reported ancestry should be interpreted as the first *valid* ancestry.[39]

As Farley (1991) notes, ancestry questions are complex, relying on "factual knowledge" of family history and individual identity choices:

> "The ancestry question is more complicated in that it depends to a large degree on factual knowledge about the history of one's family, but also requires many people to make a decision about identification: Which one or two of several possible ethnicities does a person report? Undoubtedly, some individuals identify very strongly with a particular ancestry, and will do so regardless of clues on the census form or contemporary political events. Many others, however, may not identify strongly, so their answers may depend on ephemeral events." (Farley, 1991, p. 414).

"Ephemeral events" influencing ancestry reporting can include questionnaire context. Farley (1991) suggests that in 1980, the ancestry question's placement after an English proficiency

[35]The example list has changed over time, influencing ancestry distributions (see below).

[36]See comparability details at https://usa.ipums.org/usa-action/variables/ANCESTR1#comparability_section.

[37]For instance, when an individual reported only "German" and "English" ancestry, the coding placed "German" —reflecting the primary identification— into `ancestr1` and English in `ancestr2`. However, upon reporting a third ancestry, like "French", this potential preference signal vanished: `ancestr1` was overwritten with the alphabetized string "English-French-German", while `ancestr2` became void.

[38]These supranational ancestries are termed "general heritages" by the Census Bureau.

[39]For instance, reporting "American" and "German" ancestries leads to "German" being code in `ancestr1` while "American" is discarded.

question might have biased responses towards "English" ancestry. Similarly, changes in example lists between 1980 and 1990 significantly altered ancestry distributions (Rosenwaike, 1993). Therefore, ancestry data reflects subjective identity judgments, not merely objective factual information.

Language Spoken at Home

The variable `language` records languages other than English spoken at home by respondents aged five and older. Those speaking only English at home are coded as `English`. This variable offers limited coverage of indigenous languages from foreign countries.

To facilitate comparative analysis across origin measures, we developed a crosswalk linking `bpld` (birthplace), `ancestr1` (ancestry 1), `ancestr2` (ancestry 2), and `language`. Specifically, we mapped each reported ancestry to corresponding birthplace codes (from `bpld`). For example, "Belgian" and "Flemish" ancestries were linked to the `bpld` code for Belgium. Subsequently, for each ancestry, we associated any language that is either an official language of the mapped country or spoken by at least 20% of its population, drawing on language data from Mayer and Zignago (2011). This crosswalk allows us to identify, for example, respondents of a specific foreign ancestry who were born in their ancestral country or speak a language associated with that ancestry at home.

2.5 Empirical Analysis of Census Data

This section outlines the steps of our empirical analysis using US census data, focusing on how we identify individuals' origins and document patterns of ancestry reporting.

2.5.1 Identifying Individual Origin

Following standard practice, we define individual origin based on the first reported ancestry only.[40] We exclude "general heritage" ancestries (e.g., European, Latin American) that lack clear country links, and aggregate detailed ancestries to the country level (e.g., English,

[40] Using the second reported ancestry for individuals whose first ancestry is non-informative (native or general heritage) is empirically inconsequential; few such individuals have a usable second ancestry.

Scottish, Welsh are all associated to Great Britain; Acadian, Canadian, French Canadian to Canada).

This process allows us to identify 114 distinct foreign ancestries, predominantly European and American, but with limited African representation. In the 2000 census, approximately 52.8% of native respondents reported a usable foreign ancestry.[41] The remaining 47.2% are roughly equally divided between those reporting no ancestry and those reporting non-usable native or general heritage ancestries.

Data limitations prevent origin-specific analysis of non-reporting for all 114 ancestries, as we lack ancestor birthplace data. However, an exception arises with Mexican ancestry, due to a questionnaire overlap. Census Question 5 asks about "Spanish/Hispanic/Latino" identity, including "Mexican, Mexican American, Chicano". These categories directly correspond to our Mexican ancestry group and the comparison is meaningful because Mexicans represent the main non-European ancestry.

Analyzing the 589,358 natives identifying as "Mexican, Mexican American, Chicano" in Question 5 of the 2000 census (5% sample) reveals that only 72.7% reported a Mexican-linked first ancestry (Mexican: 60.2%, Mexican American: 12.0%, Chicano: 0.5%). Of the remainder, 10.9% reported no ancestry, 6.2% "Hispanic", 1.6% "Spanish", and 1.3% "American". Thus, conservatively assuming all "Hispanic (Mexican, Mexican American, Chicano)" individuals have Mexican ancestors, over 27% chose not to report a Mexican ancestry.

Aggregating to the PUMA level reveals a negative correlation (-0.319) between the propensity *not* to report Mexican ancestry and the local Hispanic population share. Thus, in areas with larger Hispanic populations, non-reporting of Mexican ancestry is less frequent — indicating a contextual influence on identity reporting. Furthermore, we examined whether non-reporting of Mexican ancestry is systematically related to cultural attachment, proxied by Spanish home language use. Among "Hispanic (Mexican, Mexican American, Chicano)" natives *not* reporting Mexican ancestry, 44.3% speak Spanish at home, compared to 56.6% among those *reporting* Mexican ancestry —a 27.8% difference. This disparity, observed for a major ancestry group, suggests non-random selection in ancestry reporting, linked to cultural attachment.

[41] The proportion of native respondents used in prior studies is not consistently reported, precluding direct comparisons.

2.5.2 Main Foreign Ancestries in 2000

In 2000, the top five first ancestries reported by natives were: German (12.3% of all natives, 22.1% of those reporting an ancestry), Great Britain (9.7% and 17.6%), Irish (7.6% and 13.6%), Italian (4.7% and 8.5%), and Mexican (3.9% and 7.0%). Together, these five ancestries account for 38.4% of all natives (69.0% among ancestry reporters). Only seven additional countries represent at least 1.0% of native ancestries. Figure 2.1 maps the global distribution of natives reporting each country as their first ancestry.[42]

Figure 2.2 shows the proportion of natives reporting a second ancestry, conditioned on their first ancestry.[43] Countries with longer migration histories to the US, particularly European nations, exhibit a higher prevalence of reported multiple ancestries. This pattern likely reflects intermarriage across different ancestries in earlier generations.

Figure 2.3 maps the most frequent ancestry among natives in each Public Use Microdata Area (PUMA),[44] revealing 23 distinct ancestries as dominant across various PUMAs. As expected, the geographical distribution of these dominant ancestries closely reflects historical settlement patterns. For instance, Mexican ancestry predominates in PUMAs along the US-Mexico border (Texas, California, Arizona, New Mexico), mirroring the historical immigration and settlement of Mexicans in these border regions. Similarly, French ancestry is concentrated in Louisiana PUMAs, consistent with French colonial history in that state. Norwegian ancestry is most prevalent in northern North Dakota and Minnesota PUMAs, reflecting 19th-century Norwegian immigration to the Midwest, drawn to areas with climates similar to Norway (Obolensky et al., 2024).[45]

PUMA-level ancestry data vividly illustrates the substantial variation in time elapsed since

[42]Figure A.1 in the online Appendix reports the share of individuals who report each foreign country as either their first or second ancestry; as, with the only exception of the "American" ancestry, ancestries are no longer mutually exclusive, the sum of shares across countries exceeds 100%; Figures A.2 and A.3 in the online Appendix provide logarithmically scaled versions of Figure 2.1 and A.1 for better visualization of less prevalent ancestries.

[43]Figure A.4 in the online Appendix expands this by showing the proportion reporting *another* ancestry, regardless of whether we consider their first or second reported ancestry as the conditioning origin.

[44]For this mapping, "United States", "American Indian", "Nuevo Mexicano", and "Hawaiian" ancestries are categorized as "United States".

[45]Figures A.5 and A.6 in the online Appendix illustrate the PUMA-level distribution of German and Mexican ancestry prevalence, even when not the dominant ancestry, further detailing the geographic concentration of these two significant ancestry groups.

migration across different reported ancestries —a crucial factor for epidemiological frameworks to consider. The prevalence of German ancestry, the most reported in 1990 and often tracing back their arrival for over a century, provides a stark example. In contrast to more recent immigrant ancestries, descendants of 19th-century German migrants have had generations to disperse geographically across the US. This greater dispersion is evident in the PUMA data: German ancestry shares are widely distributed across diverse regions. Conversely, more recent ancestries, like Mexican, exhibit more geographically concentrated patterns, such as near the southern border. This contrast underscores the importance of accounting for time since migration when analyzing ancestry data in epidemiological studies.

2.5.3 Cohort Variation in Self-Reported Ancestry

Brittingham and de la Cruz (2004) documents significant shifts in reported foreign ancestries between the 1990 and 2000 censuses. Notably, German ancestry, the most common, decreased from 23.3% to 15.2% of the population —an absolute drop from 57.9 to 42.8 million individuals.[46] This decline may be attributed to demographic shifts, such as mortality in older German ancestry cohorts from 1990, limited German immigration, and reduced propensity for newer cohorts to report German ancestry. Indeed, cohort analysis reveals substantial census-to-census variation in German ancestry reporting among natives born in the same year.

Figure 2.4 illustrates this for natives born 1940-1979, showing German first-ancestry reporting across the 1980, 1990, and 2000 censuses. Limiting birth years to pre-1980 minimizes demographic effects, as no individual analyzed was over 60 in 2000. For each cohort, German ancestry reporting increased sharply from 1980 to 1990, then declined abruptly by 2000. While changes in multiple ancestry recording between 1980 and 1990 might explain the initial rise (see Section 2.4.1), they do not account for the subsequent decline. This decline may be linked to questionnaire design: German ancestry was an example in 1980 (fourth in the list) and 1990 (first in the list), but absent in 2000, and these wording changes are known to influence responses (Rosenwaike, 1993).

Figure 2.5 shows parallel trends for Mexican ancestry. While overall Mexican ancestry reporting increased across birth cohorts due to rising Mexican immigration throughout the 20th century, within each birth cohort, a consistent *decline* in Mexican ancestry reporting

[46]The difference with previously discussed figures is due to our focus on native population, along with our exclusion of native or general heritage first ancestries.

is evident across census years, mirroring the German pattern, though with smaller absolute magnitudes but similar *relative* declines. For instance, Mexican ancestry reporting among natives born in 1940 decreased from 1.9% in 1980 to 1.5% in 2000 —a proportional drop exceeding 40%.

These trends caution against directly comparing origin data across census waves without acknowledging potential inconsistencies, and emphasize the need for robustness checks when using different waves separately in epidemiological regressions. More broadly, these examples —German and Mexican, representing major European and Latin American ancestries— highlight a key point: self-reported ancestry is not a fixed, objective trait, but a dynamic, subjective identity marker that can evolve over an individual's lifetime and be influenced by seemingly ephemeral factors like census questionnaire design. The econometric implications of this subjective and time-varying nature of ancestry reporting warrant careful consideration.

2.6 Proxying Unobserved Heterogeneity

This section outlines how we build our proxies capturing heterogeneity in time since ancestral migration as well as attachment to the culture of origin. We then test the validity of our proxies.

2.6.1 Building proxies

We describe here how we build each of the three proxies, namely the share of the (total) population of a given ancestry that is born in the ancestral country, the spatial concentration of natives of a given ancestry, and the share of natives with an ancestry in a non-English speaking country speaking the language of the ancestral country at home.

Share Born in the Ancestral Country

The first proxy, v_o^1, represents the proportion of individuals of a given ancestry (native or foreign-born) who were actually born in their reported ancestral country—e.g., the share of Italian ancestry individuals born in Italy. This variable reflects the prevalence of first-generation immigrants within each ancestry group. Even when restricting analysis to natives, v_o^1 remains informative about the average time elapsed since ancestral migration: ancestries with higher v_o^1 values generally indicate more recent migration histories. Our data reveals significantly lower v_o^1 values for European ancestries compared to Latin American ones, and lower still for Mexican and Central American ancestries within the Latin American group, consistent with longer migration histories from Europe and Mexico (see Figure 2.6).

For example, v_o^1 is very low for European ancestries: 2.1% for Germany, 0.7% for Ireland, 2.1% for Great Britain, and 3.5% for Italy. In stark contrast, it is substantially higher for Latin American ancestries: 47.2% for Mexico and 72.7% for El Salvador. These differences align with documented migration patterns (see Figure 2 in Abramitzky and Boustan, 2017), and further highlight regional variations in migration timing within Latin America itself. While v_o^1 technically measures first-generation shares in the *overall* ancestry group, it reliably proxies time since migration even for *natives*.[47]

This ancestry-specific variation in migration timing introduces potential confounding factors in epidemiological analyses. Consider, for example, studies of English proficiency and ancestral homeland characteristics. Estimates could be biased because ancestries with longer migration histories (lower v_o^1, e.g., German) have had more generational time to assimilate to English, independent of any inherent ancestral trait affecting language acquisition.

Spatial Concentration of Natives

We calculate a spatial concentration index, v_o^2, for each ancestry using a Herfindahl-Hirschman Index (HHI) based on PUMA-level native population shares (s_{ok}). Here, s_{ok} is the proportion of natives of ancestry o residing in PUMA k relative to the total native population of ancestry o. This HHI, computed at the State level for sufficient observation counts, represents the probability that two randomly selected natives of the same ancestry reside in the same state.

Index values range widely, from 3.5% (Great Britain) to 72.5% (Anguilla),[48] with an average of 17.8% and a median of 13.0%. Figure 2.7 shows significantly lower v_o^2 values for European ancestries compared to Latin American ancestries. For instance, v_o^2 is 28.2% for Mexican ancestry and 10.5% for Bolivian ancestry (the lowest among Spanish-speaking Latin American countries). Conversely, German ancestry, representative of older migration waves, exhibits a

[47] Figure A.7 in the online Appendix confirms this: children, who co-reside with their parents, with ancestries linked to long-term migration patterns tend to have parents *not* born in the ancestral country, whereas those with more recent ancestries often have at least one foreign-born parent; The strong correlation (0.843) of v_o^1 with this intergenerational measure further validates its use as a time-since-migration proxy.

[48] The high value for Anguilla is largely due to the very small number of natives reporting that ancestry (95 in the 5% sample).

low concentration (4.8%). These origin-specific HHI variations reflect differing times since migration, with higher spatial concentration for more recent migrant groups due to stronger initial network effects on location choices.

Share of Natives Speaking Ancestral Language

Our third proxy, v_o^3, measures the share of natives with a given foreign ancestry who report speaking their ancestral language at home. Following Giuliano and Nunn (2021), this reveals individuals' attachment to their foreign ancestry identity and cultural tradition. Meaningfully, v_o^3 can only be computed for ancestries originating from countries where English is *not* an official language or where less than 20% of the population speaks English (Mayer and Zignago, 2011). This restricts its definition to 82 of our 114 foreign ancestral countries.

Figure 2.8 displays v_o^3 values for first ancestries from non-English speaking countries. Globally, European ancestries (excluding Spanish) show very low v_o^3 values —typically below 10%. In contrast, Latin American ancestries exhibit significantly higher values, often at or above 50%. These pronounced differences again underscore the varying degrees of assimilation and cultural persistence across ancestry groups, likely linked to differing times since migration and varying assimilation pressures.

2.6.2 Validating Proxies

We analyze the pattern of correlation between the three proxies, and verify their ability to provide information on the unobserved origin-specific average time elapsed since ancestry migration using information on the timing of origin-specific past migration flows to the United States.

Correlations and Spatial Patterns

We assessed correlations between our proxies, weighting by the number of natives per ancestry to align with individual-level econometric analyses and to minimize the influence of small ancestry groups on correlation estimates.[49] Correlations are high: 0.843 between v_o^1 and v_o^2; 0.867 between v_o^1 and v_o^3; and 0.834 between v_o^2 and v_o^3. These strong inter-correlations reinforce the validity of our proxies for capturing unobserved heterogeneity related to ancestry.

[49] Unweighted correlations, relevant for ancestry-level analyses, yield similar results.

Consistently, the ten lowest values for v_o^1, v_o^2, and v_o^3 are predominantly associated with European countries—reflecting their longer migration histories to the US. Conversely, Latin American countries tend to exhibit among the highest values. This confirms that countries with more distant migration histories generally have lower values for all three v_o proxies: their ancestries are primarily composed of natives, exhibit greater spatial dispersion across states, and demonstrate lower rates of ancestral language use at home.

Regressing each v_o on continent dummies further underscores this continental pattern. Weighted regressions of v_o^1, v_o^2, and v_o^3 on continental dummies yield high R^2 values: 0.805, 0.698, and 0.877, respectively. These high R^2s indicate that the majority of the variability in our proxies occurs *between* continents, rather than within them, reinforcing the strong continental dimension of these ancestry-related heterogeneity measures.

Evidence from Earlier Censuses (1850-1990)

To further validate our ancestry proxies, we use historical US censuses (1850-1990) to estimate the average census year of enumeration for immigrants from each origin country. Since 1850, censuses have recorded birthplace for foreign-born individuals, providing insights into origin-specific average time since migration for natives self-reporting different ancestries in later censuses.[50]

However, this validation exercise has inherent limitations. Our proxies aim to capture time since ancestral migration for those *reporting* specific ancestries, not all descendants of immigrants from a given country. These averages can differ due to: (*i*) self-selection in ancestry reporting and (*ii*) historical return migration, which was substantial during the era of Mass Migration (Bandiera et al., 2013), meaning that not all observed immigrants in older censuses have US-resident descendants in current data. Furthermore, differential fertility rates across ancestry groups also complicate direct historical comparisons.

Despite these caveats, we analyzed 1850-1990 census data to assess our proxies' validity. For each origin country o, we pooled foreign-born individuals across census years 1850-1990. We then computed the origin-specific average census year, $\bar{t}_o$, using individuals aged 30-35 at survey time. This age restriction minimizes double-counting individuals across censuses[51] and targets an age likely to maximize descendant presence in later generations. Note that $\bar{t}_o$

[50] See https://usa.ipums.org/usa-action/variables/bpl#availability_section (last accessed on June 18, 2024).

[51] Older censuses are full count.

overestimates the true average immigration year, as the immigration year itself is unobserved in most census years.[52] Furthermore, our measure inherently underestimates time since migration, as no immigrant could have arrived before 1815 in our dataset. We also use the 1850 census immigrant stock per origin to gauge earlier immigration flows.

Table 2.1 presents $\bar{t}_o$ and 1850 immigrant stock for the 25 largest 2000-census ancestries. For example, $\bar{t}_{DEU} = 1908$ and $\bar{t}_{MEX} = 1971$. The table also highlights timing differences within Europe: Irish immigrants are observed on average in 1887 with a large 1850 stock (984,851), while Italian immigrants appear later (average year 1930) with a much smaller 1850 stock (2,684).

Figure 2.10 visually confirms the strong positive relationship between our 2000-census ancestry proxies and the average census year of immigration, $\bar{t}_o$. As the figure illustrates, higher values of $\bar{t}_o$ are systematically associated with higher values of v_o^1 (share born in ancestral country), v_o^2 (spatial concentration), and v_o^3 (ancestral language use). This consistency, observed between historical immigration patterns and our contemporary ancestry proxies, significantly strengthens our confidence in using v_o^1, v_o^2, and v_o^3 as valid indicators of unobserved time since ancestral migration.

2.7 Using Origin-Specific Proxies

2.7.1 A Guideline

The proxies v_o^1, v_o^2, and v_o^3 serve as simple diagnostic tools to identify ancestry groups that may exhibit distinct patterns in terms of average time since ancestral migration. If specific ancestry groups stand out as outliers on these proxies, it becomes crucial to test the sensitivity of epidemiological regression estimates, described in Eq. (2.1), to the exclusion of these groups.

Specifically, if estimates change significantly upon sample restriction in ways consistent with expected effects based on w_o values and origin-specific proxies, then the original estimates may be confounded by unobserved heterogeneity in time since ancestral migration. Importantly, we caution against directly incorporating these proxies as controls in Eq. (2.1). Replacing unobserved $f(t_i)$ with v_o^1, v_o^2, or v_o^3 would fail to capture within-ancestry heterogeneity and

[52] Immigration year is recorded only 1900-1930 and since 1970; missing in other census years (see https://usa.ipums.org/usa-action/variables/YRIMMIG#availability_section, last accessed June 18, 2024).

could introduce "bad controls".[53] For example, historical US immigration policies favored groups perceived as culturally proximate, meaning our proxies likely reflect cultural distance to the US. Similarly, higher ancestry attachment may correlate with spatial concentration (v_o^2), further confounding their use as simple controls.

We propose a three-step procedure to assess the potential exposure of epidemiological regressions to the biases described in this paper.

Step 1 - Diagnostic Analysis. Analyze country-level correlations between the treatment variable and our proposed proxies (v_o^1, v_o^2, v_o^3), which capture self-selection in ancestry reporting and unobserved heterogeneity in time since migration. We recommend presenting descriptive statistics and graphical analyses to understand data structure and potential biases.

Step 2 - Address Influential Observations. Evaluate regression robustness by excluding observations identified as outliers in Step 1 or by including fixed effects to down-weight their influence. This helps isolate the impact of potentially biased observations on main results.

Step 3 - Add Relevant Controls. Consider adding controls for self-selection and time-since-migration heterogeneity, such as the proxies discussed in this paper. If Step 1 and 2 diagnostics suggest minimal bias, adding further proxies may be unnecessary and could inadvertently introduce "bad controls". Transparently discuss this tradeoff.

2.7.2 Illustrative Examples

To illustrate this procedure, we revisit two influential papers: Fernández and Fogli (2006) and Giuliano and Nunn (2021). Our choice of Fernández and Fogli (2006) is based on: (*i*) its seminal contribution to the epidemiological approach (605 Google Scholar citations as of Febr. 18, 2025); (*ii*) its clear identification strategy, linking native fertility to ancestral country fertility (1950 TFR), which helps explain time-since-migration bias; (*iii*) its use of GSS data, enabling generational distinctions (second, third and at least fourth generation immigrants) to explore unobserved heterogeneity; and (*iv*) our successful reconstruction of their sample and replication of findings despite data unavailability.[54]

Our choice of Giuliano and Nunn (2021) is based on: (*i*) its influential contribution to the

[53] We thank an anonymous referee for highlighting this point.

[54] To our knowledge, replication data are not available online; we did not request access to data as we successfully replicated original findings using original data sources.

epidemiological approach (461 Google Scholar citations as of Febr. 18, 2025); (*ii*) its analysis of the effect of historical climate variability on ancestral language use in the US, illustrating concerns about time-since-migration correlation; and (*iii*) the availability of replication data.

Fernández and Fogli (2006)

Fernández and Fogli (2006) analyze data from nine GSS waves (1977-1987) to estimate the impact of ancestral fertility norms on fertility of US natives. Their sample comprises 1,145 married native women aged 29-50 reporting foreign ancestries from 14 countries.[55] They estimate the following core equation:

$$y_{iokt} = \alpha \overline{y}_o + [\boldsymbol{\gamma}' \boldsymbol{x}_i] + d_k + d_t + \epsilon_{iokt} \tag{2.7}$$

Here, y_{iokt} is the number of children for woman i of ancestry o in region k at time t; $\overline{y}_o$ is the Total Fertility Rate (TFR) in origin country o in 1950; $\boldsymbol{x}_i$ includes controls for siblings, a quadratic polynomial in age, education dummies (high school, some college, college), and parental education.[56] Region fixed effects (d_k) control for geographical location (9 US regions), and d_t are survey year fixed effects.

Crucially, their sample includes all generational cohorts of natives reporting foreign ancestry, without controlling for time since migration. The authors find a positive and significant effect of $\overline{y}_o$ on y_{iokt}, interpreting this as evidence that:

> "[I]t is not only personal experience ... that matters to a woman's fertility, but also that her culture ... plays a role. ... even after controlling for various characteristics and family background of a woman, both her own personal experience and her culture play a role in influencing her fertility." (Fernández and Fogli, 2006, p. 561)

Step 1 - Diagnostic. Fernández and Fogli (2006) acknowledge data limitations in distin-

[55] They exclude 6 former socialist economies and 8 countries with very small GSS sample sizes.

[56] Parental education controls, while included, substantially reduce sample size (from 1145 to 921) and introduce potential selection biases. GSS parental education data depends on co-residence patterns at age 16, leading to non-random missingness and potential bias related to family structure and fertility outcomes. For instance, paternal education may take the value of a step-father or any co-residing male relative for women whose father was not present. Women in that category have more children (195 women, 2.48 children compared to 2.32) and recorded paternal education —corresponding to the co-residing male relative when fathers are missing— is also lower (61 women, 9.79 years compared to 10.54). Clearly, the addition of paternal education generates a non-random selection with respect to unobserved determinants of fertility.

guishing immigrant generations, but GSS data *does* allow this.[57] GSS data indicates parental nativity and counts foreign-born grandparents.[58] Following Giavazzi et al. (2019), we categorize natives of foreign ancestry as: second-generation (at least one foreign-born parent), third-generation (native parents, at least two foreign-born grandparents), and fourth-generation or more (native parents, at most one foreign-born grandparent).

Applying these established generational definitions to the estimation sample reveals a notable generational distribution: second-generation immigrants constitute just 8.6%, third-generation 22.5%, while a substantial 64.8% are of (at least)fourth-generation immigrants; a small fraction (4.2%) possess missing generational information. Two key observations emerge from these figures. First, a precise generational status, directly linked to either parents or grandparents, can only be definitively assigned to a minority —merely 31.1%— of the sample, specifically those of the second or third generations. The generational depth for the primarily native fourth generation and beyond, representing nearly two-thirds of the sample, remains less precisely defined within the GSS framework. Second, and perhaps more critically, we observe marked generational heterogeneity across different ancestries themselves (detailed in Table A.1 in the online Appendix). Strikingly, for the majority —nine out of fourteen— of ancestral origins, fourth-generation immigrants and beyond constitute the largest group. In contrast, third-generation immigrants are most numerous for four ancestries (Finnish, Italian, Russian, and Swedish), and second-generation immigrants are most prevalent only in the case of Mexican ancestry. This uneven generational distribution immediately raises a diagnostic flag: individuals of Mexican origin emerge as potentially influential observations, possessing characteristics that might introduce bias into the overall estimation.

This diagnostic concern is further reinforced by considering our independent measure of time since ancestral migration, $\bar{t}_o$.[59] Indeed, visual inspection of Figure 2.11 provides compelling corroboration. Individuals of Mexican origin are starkly revealed as outliers within this visual representation, characterized by a more recent migration history (reflected in a higher $\bar{t}_o$), coupled with both higher Total Fertility Rates in their ancestral origin (1950 TFR, $\bar{y}_o$),

[57]The authors write: "unfortunately we cannot distinguish second-generation Americans from those who have been in the US for longer" (p. 554)", but the necessary information is available in the currently available version of the dataset from the waves of the GSS used by Fernández and Fogli (2006).

[58]GSS data indicates whether parents are native/foreign-born, and counts foreign-born grandparents, but does *not* list their birth countries.

[59]To reiterate, individual-level generational status is derived from GSS data, while the country-level measure $\bar{t}_o$ is constructed from the analysis of historical census data spanning 1850-1990.

and higher observed fertility (y_{iokt}). This conspicuous clustering and outlier status strongly suggests a potential for upward bias, wherein the relationship discerned within the original analysis may be, at least in part, spuriously driven by the unobserved generational heterogeneity confounded with these particularly influential observations of Mexican origin.

Step 2 - Address Influential Observations. To assess the influence of Mexican-origin individuals, identified as potential outliers, we test the sensitivity of the original core findings by systematically excluding them from their sample. Table 2.2 replicates the original specifications reported in Table 3a of Fernández and Fogli (2006), serving as a baseline. In contrast, Table 2.3 presents estimates from six key specifications —those including the 1950 TFR as a regressor— *after* removing Mexican-ancestry observations.

The impact of this sample restriction is notable. The estimated coefficient on origin-country TFR in 1950 demonstrably *weakens*. In Col. (1), for example, the coefficient decreases from a robustly significant 0.166 (s.e. 0.025) to a less precisely estimated 0.135 (s.e. 0.050). This seemingly minor sample alteration —dropping just 38 observations, a mere 3.3% of the original sample— precipitates a substantial 18.7% reduction in the estimated coefficient's magnitude. Furthermore, the standard error *doubles*.

Step 3 - Add Relevant Controls. Table 2.4 expands the baseline specification in Eq. (2.7) by incorporating generational controls. We introduce dummies for immigrant generation (third and fourth, with second-generation as the reference) and interact these with origin-country TFR ($\overline{y}_o$).[60] In these specifications, the baseline TFR coefficient (α) now reflects the effect for second-generation immigrants. The results reveal a significant TFR effect on fertility specifically for second and third-generation women (comprising 31.1% of the sample in the first data column), but no significant effect for fourth-generation immigrants.

This generational pattern aligns with our theoretical framework. However, alternative interpretations warrant consideration. First, lower TFR variability in the fourth-generation sample, potentially reducing statistical power; Mexico's high TFR is less represented in this group. Second, 1950 TFR may be a less accurate cultural norm proxy for longer-descent generations. Third, the composition of the estimation sample across ancestral countries varies significantly, implying that the coefficient may reflect differences across countries rather than over time elapsed since ancestral migration. Fourth, attenuation bias, inherent to the epi-

[60]This table revisits specifications with 1950 TFR (Table 3a in Fernández and Fogli (2006)). Table A.4 in the Appendix confirms that the original specification yields similar estimates when restricted to the generationally-defined subsample (96% of original sample).

demiological approach and amplified by mixed ancestries more common among native women who are fourth-generation immigrants, could also explain the non-significant finding.[61]

Our theoretical framework suggests interpreting the core TFR coefficient as capturing the effect primarily within the subsample who *reports* foreign ancestry. GSS data reveals that a substantial fraction of natives with foreign-born ancestors *do not* report foreign ancestry. If, as our Mexican ancestry analysis in Section 2.5.1 suggests, these non-reporting individuals are *less* culturally attached to their origin, their exclusion could lead to an upward bias in the estimated TFR effect in the baseline specification.

Giuliano and Nunn (2021)

Giuliano and Nunn (2021) use 2000 Census data to analyze the influence of historical climatic variability in the ancestral country on natives' likelihood of speaking a non-English language at home—a proxy for tradition. Their main hypothesis is that higher climatic variability reduces tradition's importance, thus decreasing ancestral language use. Their sample includes 3,343,097 US natives with foreign ancestries from 84 non-English-official-language countries. They estimate:

$$y_{iok} = \alpha w_o + \boldsymbol{\beta}'\boldsymbol{x}_o + \boldsymbol{\gamma}'\boldsymbol{x}_i + \phi x_{ok} + d_k + \epsilon_{iok} \tag{2.8}$$

Here, y_{iok} is a dummy for speaking a non-English language at home, for native i of ancestry o in US metro area k. The key variable, w_o, measures historical climatic variability in origin country o. Origin-country controls ($\boldsymbol{x}_o$) include historical distance from equator, historical development/political complexity measures, linguistic distance to English, and the logarithm of GPD in 2000. Metro-area control (x_{ok}) is the fraction of metro area k's population born in country o. Individual controls ($\boldsymbol{x}_i$) include age (a quadratic polynomial), sex, marital status, education/labor-force-status fixed effects, the logarithm of income, and urban/rural indicator.[62]

[61] See Fernández (2011) and Ek (2021). Appendix Table A.5 shows qualitatively similar results using $\bar{t}_o$ (in-sample terciles of country of origin's average year of enumeration) as an alternative time-since-migration measure, both as a control and interacted with TFR. Specifications with parental education controls (columns 5-6) are not emphasized due to their inherent selection biases.

[62] Bertoli et al. (2024) note discrepancies between the textual descriptions and actual specifications, including English-official-language countries in the sample and "Not in MSA" observations treated as a separate metro area. Paola Giuliano and Nathan Nunn describe these as "imprecisions and omissions" in a corrigendum (Giuliano and Nunn, 2024); we do not address these inconsistencies in this exercise.

In their benchmark specification, i.e., the first data column in Table 5 in Giuliano and Nunn (2021), the estimated α is -0.447. A one standard deviation increase in climatic variability reduces predicted foreign language speaking by 3.2 percentage points —a notable effect given the 12% baseline rate of speaking a foreign language at home.

Step 1 - Diagnostic. A significant portion of the original sample —11.6%— reports ancestry from American countries, with 21 out of 84 ancestral origins being American-based. Critically, *all but Canada* among these 21 American countries share a recent migration history to the US. This recency, reflected in our v_o^1 proxy (share born in ancestral country), might inflate ancestral language use. Moreover, 19 of these 21 American ancestries are Spanish-speaking.

Notably, Spanish-ancestry natives in the sample exhibit dramatically higher rates of non-English use at home: 65.4% versus only 6.0% for the rest of the sample —over ten times greater.[63] Because Spanish is a shared language among multiple origins, controlling for co-ethnic presence at the metro-area may be insufficient to capture the incentive to speak Spanish. For instance, someone of Spanish ancestry might speak Spanish not because Spaniards are present but because individuals from Equator are in the same metro-area.

Thus, the estimated effect of climatic variability may confounded by: (*i*) shorter time since migration for Latin American Spanish-speaking ancestries, and (*ii*) inadequate controls for the general prevalence and incentives to speak Spanish in the US. Both factors likely inflate observed ancestral language use for Spanish-speaking groups, and *both are correlated with* (not counteracted by) higher historical climatic variability.

As Figure 2.12 visually confirms, Spanish-speaking ancestry countries indeed appear as outliers in the original sample, characterized by both higher climatic variability and, crucially, more recent average year of enumeration, reinforcing concerns about potential confounding upward bias in the estimated climatic variability effect.

Step 2 - Address Influential Observations. To gauge the impact of Spanish-speaking ancestries, potentially influential due to their recency, we re-estimate the original benchmark specification in Table 2.6, excluding the 362,050 observations from these origins (10.8% of the sample). The estimated coefficient α on climatic variability, initially -0.447 (s.e. 0.168), decreases to -0.202 (s.e. 0.062). While still statistically significant, despite the modest sample reduction, the magnitude of the climatic variability effect is substantially reduced —by 54.8%.

[63] This figure includes people reporting their ancestry in Spain.

A one standard deviation increase in climatic variability now implies a 1.4 percentage point reduction in predicted foreign language speaking, down from 3.2 percentage points in the original estimate.

This sensitivity analysis does *not* invalidate the article's core finding, nor do we claim our restricted estimate as the "true" effect. Instead, it highlights the disproportionate influence of Spanish-speaking ancestries in shaping the estimated coefficient in the baseline specification. Spanish, spoken at home by 28.1 million individuals out of 47.0 speaking a foreign language at home in 2000 (Shin and Bruno, 2003), holds a unique position reflecting large recent Latin American immigration waves. This analysis empirically underscores our concern: unobserved heterogeneity across ancestry groups, particularly in time since migration and Spanish language incentives unique to the US context, may lead to *overestimation* of cultural persistence in epidemiological regressions, motivating the addition of adequate control variables.

Step 3 - Add Relevant Controls. Table 2.5 presents original estimates (Table 5 in Giuliano and Nunn (2021)) alongside those incorporating our time-since-migration proxy, $\bar{t}_o$ (average year of enumeration). Due to data limitations ($\bar{t}_o$ unavailable for one ancestry), Col. (3) first focuses on the reduced 83-country sample; the climatic variability coefficient remains largely unchanged. Col. (4) then adds $\bar{t}_o$ as a control, reducing the main coefficient by 42.6%, to -0.257 (s.e. 0.154), while $\bar{t}_o$ is positively and significantly associated with foreign language speaking, as expected: more recent migration histories (higher $\bar{t}_o$) correlate with greater language persistence.

Col. (2) explores ancestry reporting self-selection by adding the percentage of individuals in the metro-area that identify with the respondent's ancestry. This reduces the climatic variability coefficient from -0.447 to -0.387, consistent with self-selection bias inflating estimates of cultural persistence due to non-random ancestry reporting. The ancestry population control itself is positively and significantly associated with foreign language speaking, indicating localized network effects.[64] Beyond its possible role in controlling for non-random selection in reporting an ancestry, this variable may also capture language use incentives, similar to one of the original controls. This could explain why our control is not negative —the expected sign for non-random selection—: a large community increases the value of one's ancestral language.

[64]This control captures within-city ancestry population variation beyond metro-area fixed effects, complementing the original controls for metropolitan area.

Importantly, our theoretical framework suggests the baseline original estimates reflect the effect of climate variability on the *selected subsample* of natives who *choose to report* foreign ancestry. If, as our Mexican ancestry analysis suggests, these ancestry-reporting natives are systematically more attached to their origin culture, the baseline estimate may still exhibit upward bias, even with added controls, highlighting the fundamental challenge of selection on unobservables in this methodological approach.

2.8 Concluding Remarks

A substantial body of influential research in comparative development increasingly utilizes natives' first self-reported foreign ancestry in US data to define origin. This methodological choice largely reflects data constraints, as the post-1980 Census and ACS lack alternative origin identifiers.

This paper argues and empirically demonstrates that relying on foreign ancestry introduces specific identification threats to the epidemiological approach, challenging the widespread assumption of inherent attenuation bias. Contrary to this presumption, estimating canonical specifications specifically on ancestry-reporting natives may *amplify*, rather than attenuate, biases. The primary threats stem from: (*i*) the significant proportion of natives not reporting foreign ancestry and (*ii*) unobserved heterogeneity in time since ancestral migration.

We therefore recommend that researchers rigorously assess the sensitivity of their findings to the inclusion of ancestries with recent US migration histories —even if these constitute a small fraction of the overall sample. Furthermore, we emphasize that even with such robustness checks, interpretations must acknowledge the inherently selected nature of ancestry-based samples. Estimates are necessarily conditional on a population subgroup —those who *choose* to report foreign ancestry— potentially magnifying the likelihood of detecting significant cultural or genetic persistence due to the likely stronger cultural attachment within this selected sample. Current data limitations unfortunately preclude fully overcoming this inherent sample selection bias.

Bibliography

ABRAMITZKY, R. AND L. BOUSTAN (2017): "Immigration in American Economic History," *Journal of Economic Literature*, 55, 1311–45.

AKERLOF, G. A. AND R. E. KRANTON (2000): "Economics and identity," *Quarterly Journal of Economics*, 115, 715–753.

ALBERT, C. AND J. MONRAS (2022): "Immigration and Spatial Equilibrium: The Role of Expenditures in the Country of Origin," *American Economic Review*, 112, 3763–3802.

ALESINA, A., Y. ALGAN, P. CAHUC, AND P. GIULIANO (2015): "Family Values and the Regulation of Labor," *Journal of the European Economic Association*, 13, 599–630.

ALESINA, A., A. DEVLEESCHAUWER, W. EASTERLY, S. KURLAT, AND R. WACZIARG (2003): "Fractionalization," *Journal of Economic Growth*, 8, 155–194.

ALESINA, A. AND P. GIULIANO (2015): "Culture and Institutions," *Journal of Economic Literature*, 53, 898–944.

ALGAN, Y. AND P. CAHUC (2010): "Inherited trust and growth," *American Economic Review*, 100, 2060–2092.

ANTECOL, H. (2000): "An examination of cross-country differences in the gender gap in labor force participation rates," *Labour Economics*, 7, 409–426.

ANTMAN, F., B. DUNCAN, AND S. J. TREJO (2016): "Ethnic attrition and the observed health of later-generation Mexican Americans," *American Economic Review*, 106, 467–471.

ANTMAN, F. M. AND B. DUNCAN (2023): "American Indian Casinos and Native American Self-Identification," *Journal of the European Economic Association*, 21, 2547–2585.

ANTMAN, F. M., B. DUNCAN, AND S. J. TREJO (2023): "Hispanic Americans in the Labor Market: Patterns over Time and across Generations," *Journal of Economic Perspectives*, 37, 169–198.

ARBATLI, C. E., Q. H. ASHRAF, O. GALOR, AND M. KLEMP (2020): "Diversity and conflict," *Econometrica*, 88, 727–797.

Bandiera, O., I. Rasul, and M. Viarengo (2013): "The Making of Modern America: Migratory Flows in the Age of Mass Migration," *Journal of Development Economics*, 102, 23–47.

Beck Knudsen, A. S. (2022): "Those Who Stayed: Selection and Cultural Change in the Age of Mass Migration," Mimeo, https://annesofiebeckknudsen.com/wp-content/uploads/2022/08/thosewhostayed.pdf.

Bertoli, S., M. Clerc, J. Loper, and E. Roca Fernández (2024): "Understanding Cultural Persistence and Change: a replication exercise of Giuliano and Nunn (2021)," *Economic Inquiry*, https://doi.org/10.1111/ecin.13242.

Bisin, A. and T. Verdier (2001): "The economics of cultural transmission and the dynamics of preferences," *Journal of Economic Theory*, 97, 298–319.

Brittingham, A. and G. P. de la Cruz (2004): "Ancestry: 2000," US Census Bureau, 2000 Census Brief, available at https://www.census.gov/history/pdf/ancestry.pdf.

Chavez, L. R. (2013): *The Latino Threat: Constructing Immigrants, Citizens and the Nation*, Stanford University Press, second edition.

Conley, T. G. and M. Kelly (2024): "The standard errors of persistence," *Journal of International Economics*, 104027.

Desmet, K., I. Ortuño-Ortín, and R. Wacziarg (2017): "Culture, Ethnicity, and Diversity," *American Economic Review*, 107, 2479–2513.

Duncan, B. and S. J. Trejo (2011): "Tracking intergenerational progress for immigrant groups: The problem of ethnic attrition," *American Economic Review*, 101, 603–608.

——— (2017): "The complexity of immigrant generations: Implications for assessing the socioeconomic integration of Hispanics and Asians," *ILR Review*, 70, 1146–1175.

Ek, A. (2021): "Cross-country differences in preferences for leisure," *Labour Economics*, 72, 102054.

——— (2024): "Cultural Values and Productivity," *Journal of Political Economy*, 124, 295–335.

Farley, R. (1991): "The New Census Question about Ancestry: What Did It Tell Us?" *Demography*, 28, 411–429.

Fasani, F., G. Mastrobuoni, E. G. Owens, and P. Pinotti (2019): *Does immigration increase crime?*, Cambridge University Press.

Fernández, R. (2011): "Does culture matter?" in *Handbook of Social Economics*, ed. by J. Benhabib, A. Bisin, and M. O. Jackson, Elsevier, vol. 1, 481–510.

Fernández, R. (2007): "Women, Work, and Culture," *Journal of the European Economic Association*, 5, 305–332.

Fernández, R. and A. Fogli (2006): "Fertility: The Role of Culture and Family Experience," *Journal of the European Economic Association*, 4, 552–561.

Fouka, V., S. Mazumder, and M. Tabellini (2021): "From Immigrants to Americans: Race and Assimilation during the Great Migration," *Review of Economic Studies*, 89, 811–842.

Fulford, S. L., I. Petkov, and F. Schiantarelli (2020): "Does it matter where you came from? Ancestry composition and economic performance of US counties, 1850–2010," *Journal of Economic Growth*, 25, 341–380.

Galor, O., M. Klemp, and D. Wainstock (2023): "Roots of Inequality," NBER Working Paper 31580.

Galor, O. and O. Özak (2016): "The Agricultural Origins of Time Preference," *American Economic Review*, 106, 3064–3103.

Galor, O., Ö. Özak, and A. Sarid (2016): "Geographical Origins and Economic Consequences of Language Structures," IZA Discussion Paper No. 10379.

——— (2020): "Linguistic traits and human capital formation," *AEA Papers and Proceedings*, 110, 309–313.

Galor, O. and V. Savitskiy (2018): "Climatic Roots of Loss Aversion," NBER Working Paper 25273.

Giavazzi, F., I. Petkov, and F. Schiantarelli (2019): "Culture: Persistence and evolution," *Journal of Economic Growth*, 24, 117–154.

Giuliano, P. (2007): "Living arrangements in Western Europe: Does cultural origin matter?" *Journal of the European Economic Association*, 5, 927–952.

Giuliano, P. and N. Nunn (2018): "Ancestral characteristics of modern populations," *Economic History of Developing Regions*, 33, 1–17.

——— (2021): "Understanding cultural persistence and change," *The Review of Economic Studies*, 88, 1541–1581.

——— (2024): "Correction to: Understanding Cultural Persistence and Change," *The Review of Economic Studies*, 91, 597.

Guiso, L., P. Sapienza, and L. Zingales (2006): "Does culture affect economic outcomes?" *Journal of Economic Perspectives*, 20, 23–48.

Jaeger, D. A., T. Dohmen, A. Falk, D. Huffman, U. Sunde, and H. Bonin (2010): "Direct evidence on risk attitudes and migration," *The Review of Economics and Statistics*, 92, 684–689.

Luttmer, E. F. P. and M. Singhal (2011): "Culture, context, and the taste for redistribution," *American Economic Journal: Economic Policy*, 3, 157–179.

Mayer, T. and S. Zignago (2011): "Notes on CEPII's distances measures: The GeoDist database," Working Papers 2011-25, CEPII.

Murdock, G. (1967): *Ethnographic Atlas*, Pittsburgh: University of Pittsburgh Press.

Obolensky, M., M. Tabellini, and C. Taylor (2024): "Homeward Bound: How Migrants Seek Out Familiar Climates," NBER Working Paper 32035.

Patel, K. and F. Vella (2013): "Immigrant networks and their implications for occupational choice and wages," *Review of Economics and Statistics*, 95, 1249–1277.

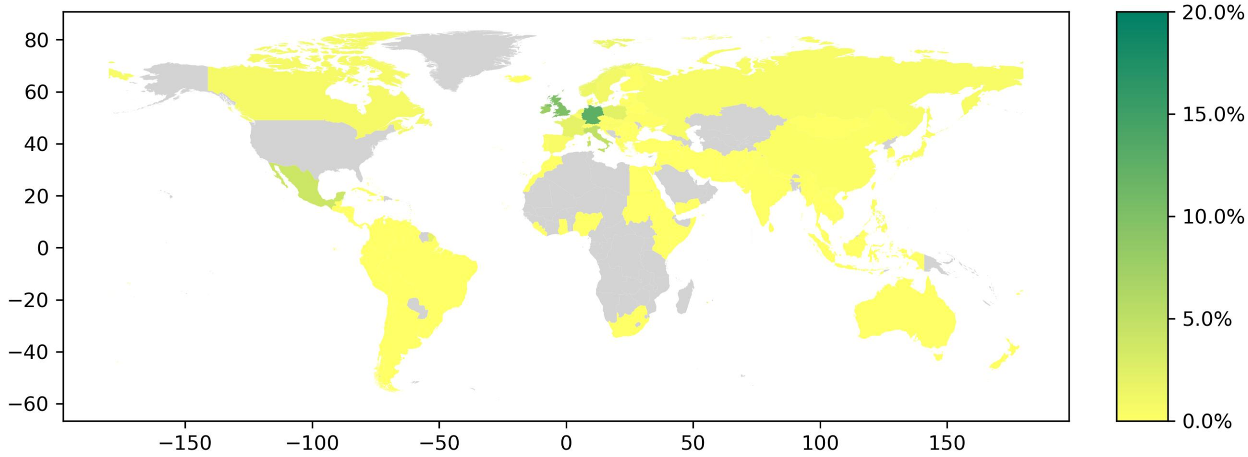
Putterman, L. and D. N. Weil (2010): "Post-1500 population flows and the long-run determinants of economic growth and inequality," *The Quarterly Journal of Economics*, 125, 1627–1682.

Rosenwaike, I. (1993): "Ancestry in the United States Census, 1980-1990," *Social Science Research*, 22, 383–390.

Ruggles, S., S. Flood, M. Sobek, D. Brockman, G. Cooper, S. Richards, and M. Schouweiler (2023): "IPUMS USA: Version 13.0 [dataset]," Minneapolis, MN: IPUMS. https://doi.org/10.18128/D010.V13.0.

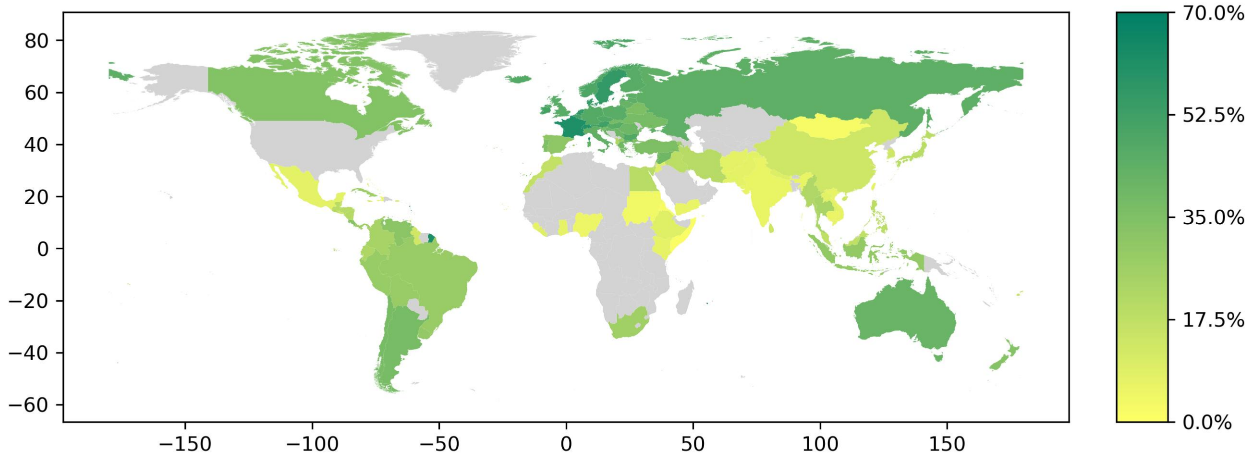
Shin, H. B. and R. Bruno (2003): "Language Use and English-Speaking Ability: 2000," Census 2000 Brief, US Census Bureau, available at https://www2.census.gov/library/publications/decennial/2000/briefs/c2kbr-29.pdf.

Figure 2.1: Proportional Distribution of Native Ancestry by Country in 2000, considering first ancestry



Data sources: Authors' elaboration on the 2000 Census (Ruggles et al., 2023).
Notes: For each country, we assign a value representing the proportion of individuals who identify that country as their first ancestry.

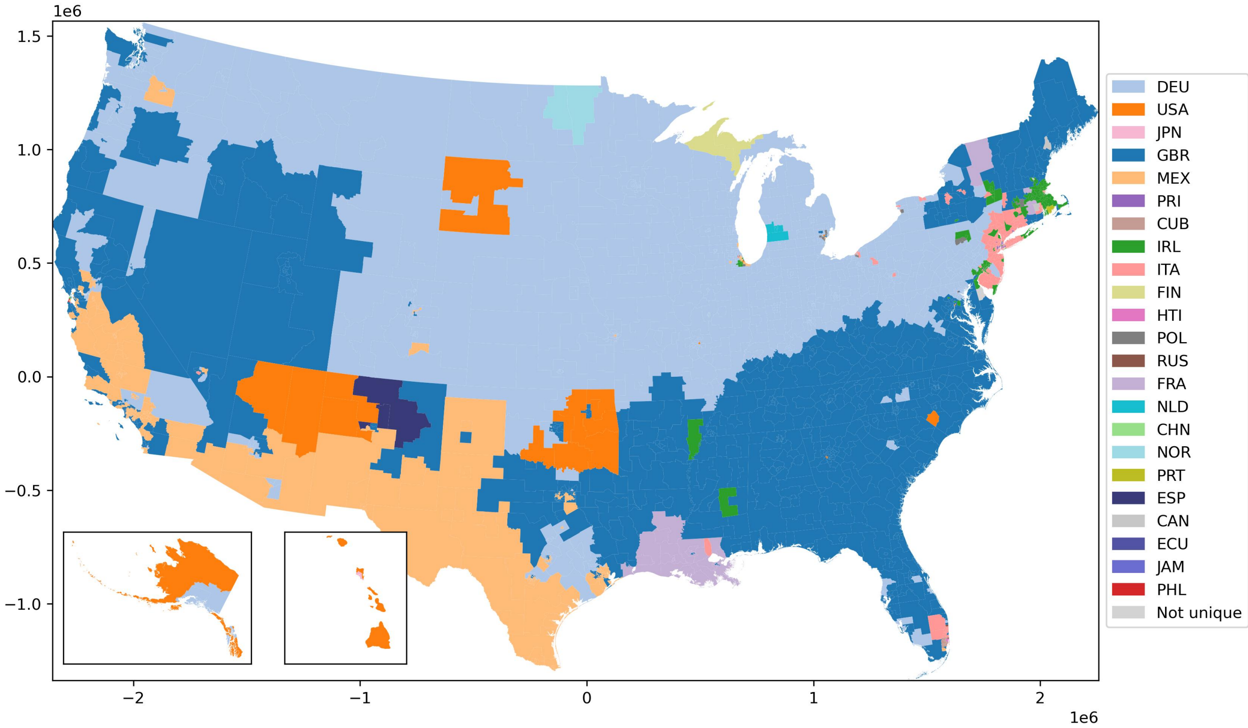
Figure 2.2: Incidence of multiple ancestries



Data sources: Authors' elaboration on the 2000 Census (Ruggles et al., 2023).

Notes: We examine the percentage of individuals with a specific ancestral background who identify as having a secondary ancestry. For example, we analyze the proportion of individuals with Italian as their primary ancestry who also report having a secondary ancestry.

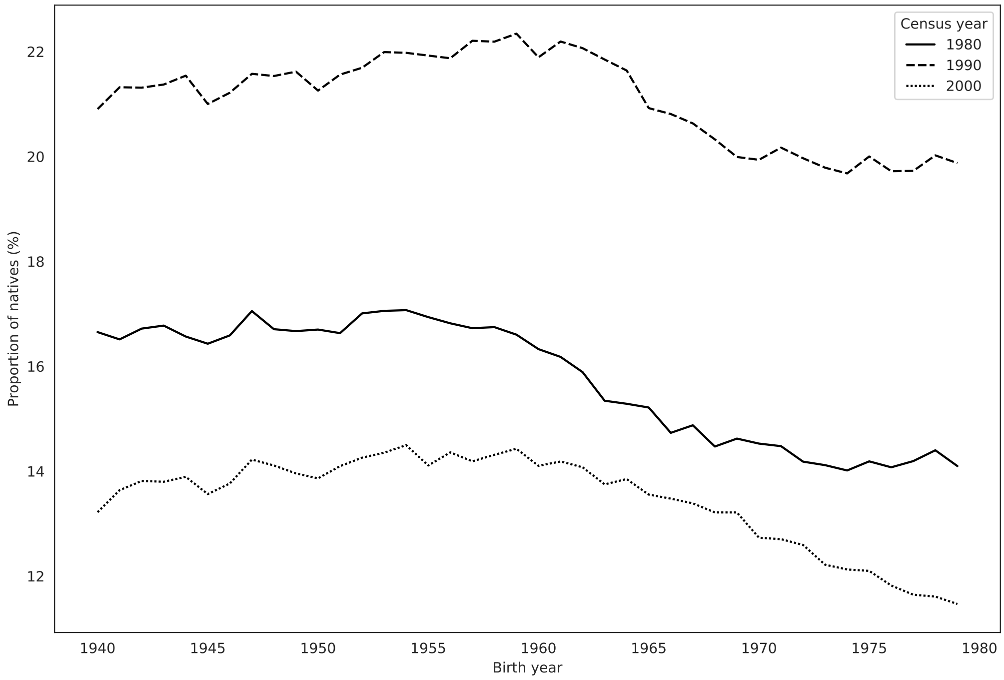
Figure 2.3: Most prevalent ancestry in each PUMA



Data sources: Authors' elaboration on the 2000 Census (Ruggles et al., 2023).

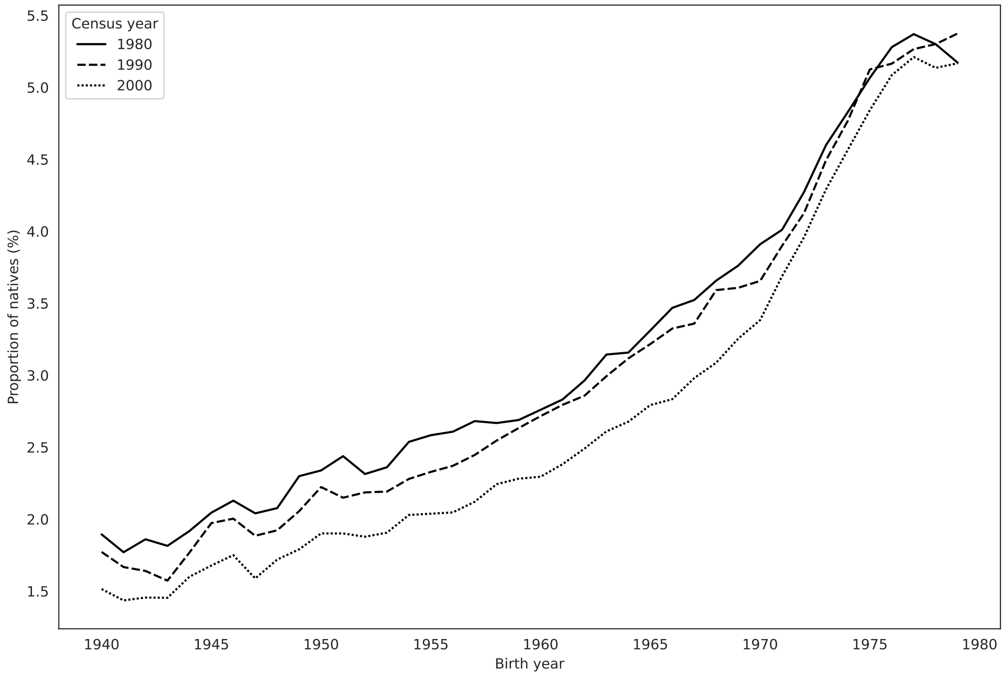
Notes: For each Public Use Microdata Area (PUMA), we identify and assign the most prevalent ancestral heritage among native residents.

Figure 2.4: Natives with German ancestry by birth cohort, different census years



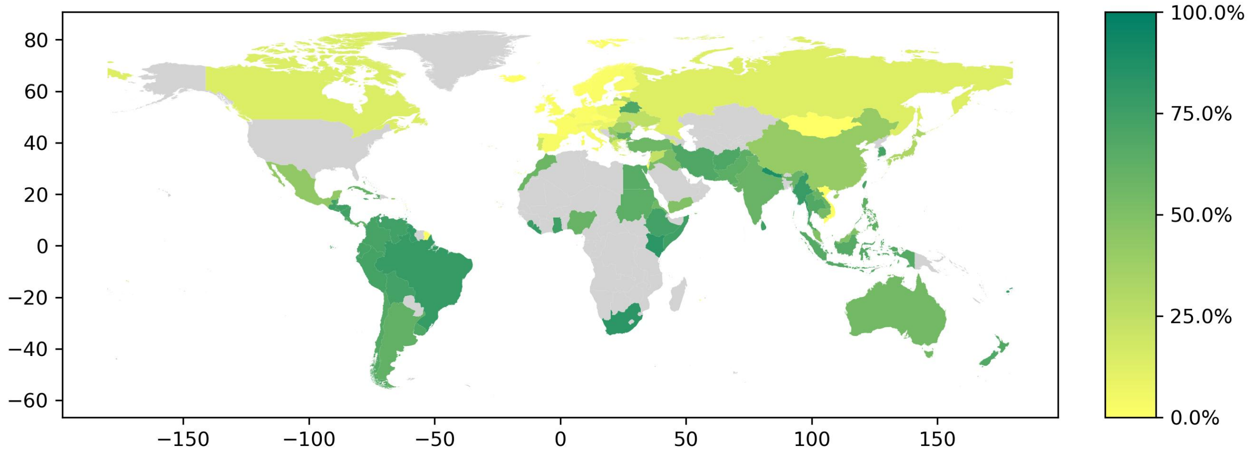
Data sources: Authors' elaboration on the 1980, 1990 and 2000 Census (Ruggles et al., 2023). Notes: This Figure plots the share of natives born between 1940 and 1979 reporting German as a first ancestry in the 1980, 1990 and 2000 census. The restrictions on the year of birth are meant to minimize the influence of demographic events, as no individual in our analysis is aged above 60 in 2000.

Figure 2.5: Natives with Mexican ancestry by birth cohort, different census years



Data sources: Authors' elaboration on the 1980, 1990 and 2000 Census (Ruggles et al., 2023). Notes: This Figure plots the share of natives born between 1940 and 1979 reporting Mexican as a first ancestry in the 1980, 1990 and 2000 census. The restrictions on the year of birth are meant to minimize the influence of demographic events, as no individual in our analysis is aged above 60 in 2000.

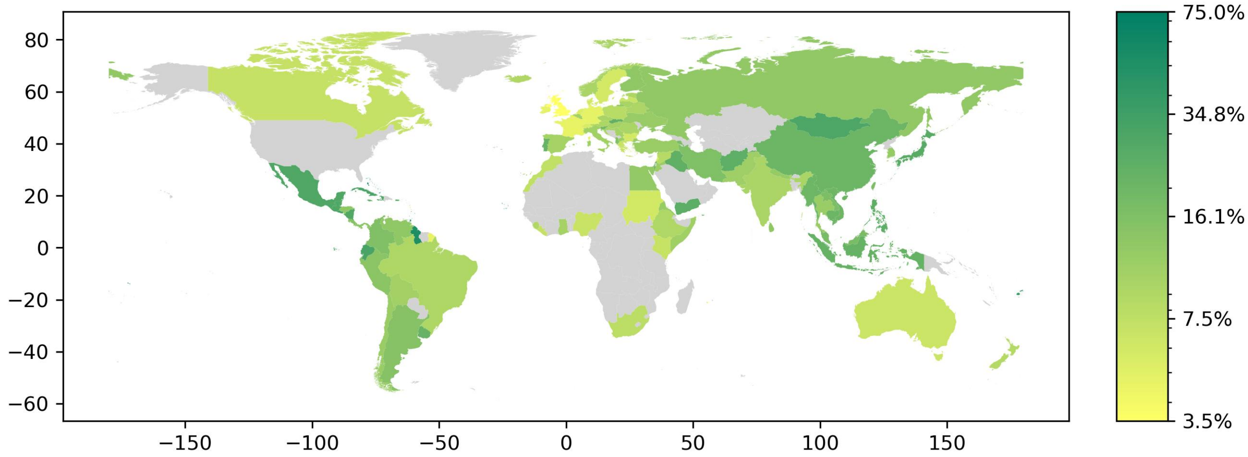
Figure 2.6: Share of individuals born in the ancestral country



Data sources: Authors' elaboration on the 2000 Census (Ruggles et al., 2023).

Notes: For each country, we assign a value representing the proportion of individuals who were born there among those declaring a first ancestry linked to that country.

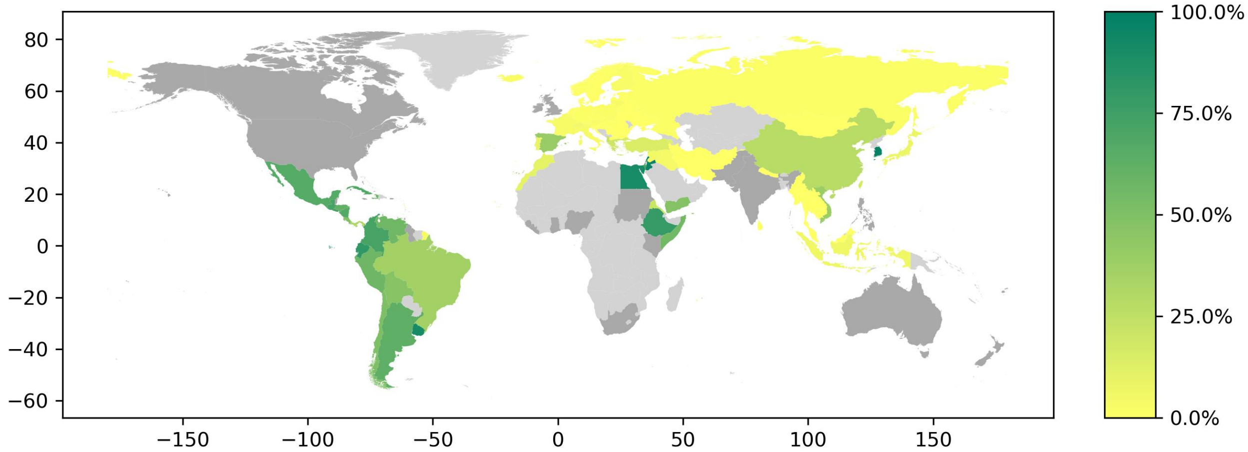
Figure 2.7: Herfindahl-Hirschman Index of spatial concentration of natives of foreign ancestry



Data sources: Authors' elaboration on the 2000 Census (Ruggles et al., 2023).

Notes: This Figure displays, for each country of ancestry, the Herfindahl-Hirschman of spatial concentration at the State level. It represents the proportion of natives residing in a given State declaring a first ancestry linked to that country relative to the total native population of ancestry.

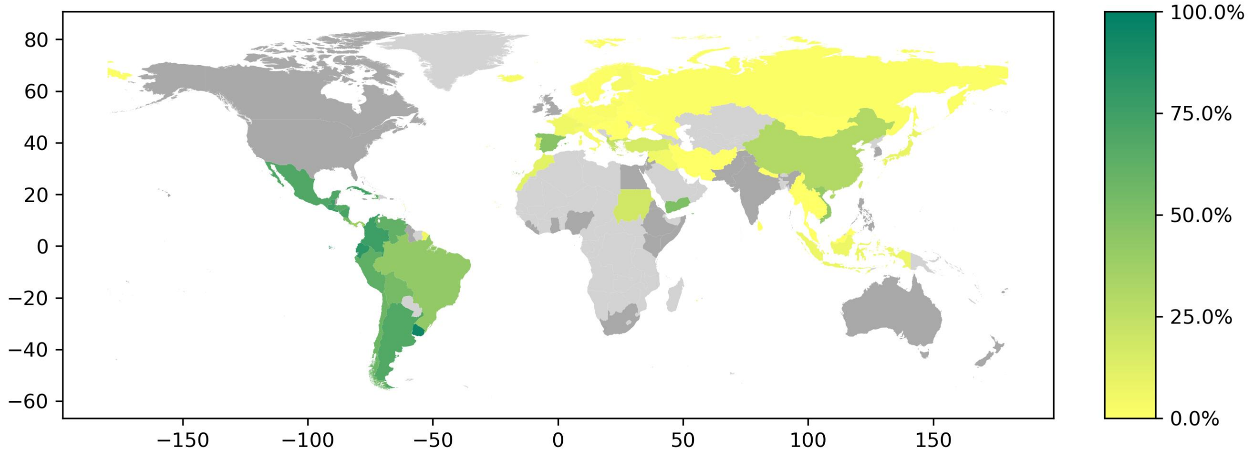
Figure 2.8: Share of natives speaking the ancestral language, by country of ancestry



Data sources: Authors' elaboration on the 2000 Census (Ruggles et al., 2023).

Notes: This figure displays the percentage of individuals who claim ancestry in a particular country (first ancestry only) and also speak an official language of that country. The analysis includes only individuals with at least one reported ancestry, and countries where English is an official language are represented in light gray.

Figure 2.9: Share of natives speaking the ancestral language, by country of ancestry (first and second ancestry)



Data sources: Authors' elaboration on the 2000 Census (Ruggles et al., 2023).

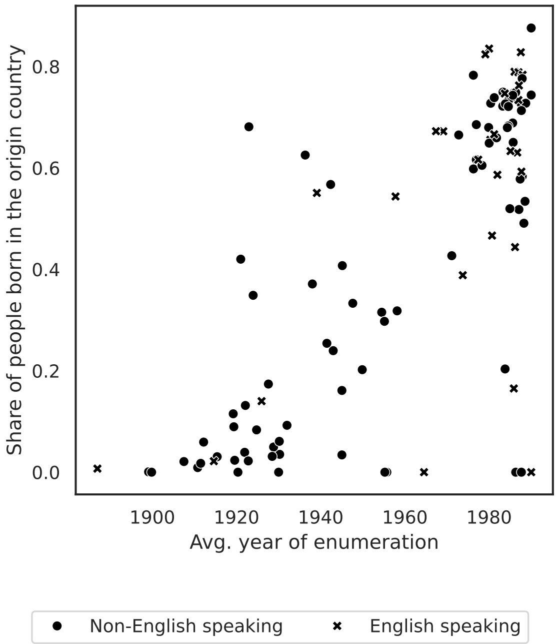
Notes: This figure displays the percentage of individuals who claim ancestry in a particular country (first or second ancestry) and also speak an official language of that country. The analysis includes only individuals with at least one reported ancestry, and countries where English is an official language are represented in light gray.

Table 2.1: Average year of enumeration in a US Census for different countries

Country	Native-born %	Average year of enumeration	Residents in 1850
Germany	22.06%	1908	596,181
United Kingdom	17.56%	1915	387,710
Ireland	13.65%	1887	984,851
Italy	8.50%	1930	2,684
Mexico	7.03%	1971	15,038
Poland	4.10%	1930	3,208
France	3.43%	1923	61,433
Norway	2.39%	1911	12,749
Netherlands	1.81%	1928	11,662
Sweden	1.74%	1911	3,440
Canada	1.47%	1926	135,300
Russia	1.05%	1922	716
Puerto Rico	0.95%	1974	-
Spain	0.91%	1945	2,634
Denmark	0.61%	1915	2,080
Hungary	0.56%	1925	100
Greece	0.51%	1945	92
Portugal	0.49%	1950	1,848
Philippines	0.49%	1977	-
China	0.46%	1945	671
Japan	0.45%	1954	-
Switzerland	0.38%	1912	14,546
Slovakia	0.34%	1920	-
Finland	0.33%	1922	-
Ukraine	0.30%	1941	-

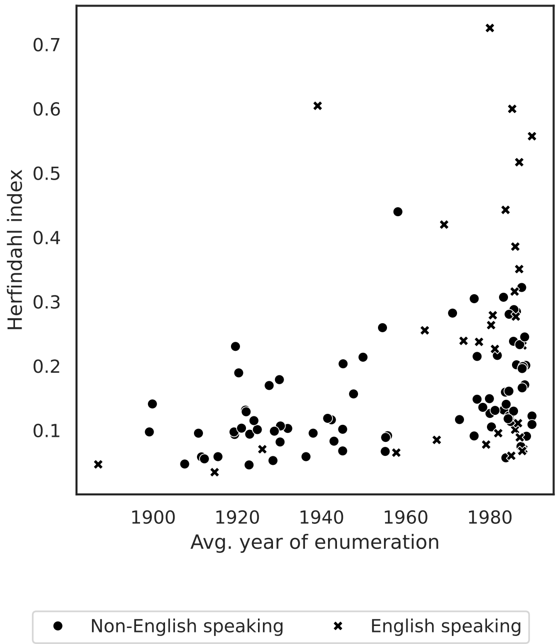
Notes: This table reports, for each of the 25 main foreign countries of ancestry for the native population in the 2000 census, the average year $\bar{t}_o$ in which the individuals aged 30 to 35 born in each foreign country o were observed in the data (1850 to 1990 census), and the (total) stock of individuals born in country c in the 1850 census.

Figure 2.10: Average year of enumeration and our proxies



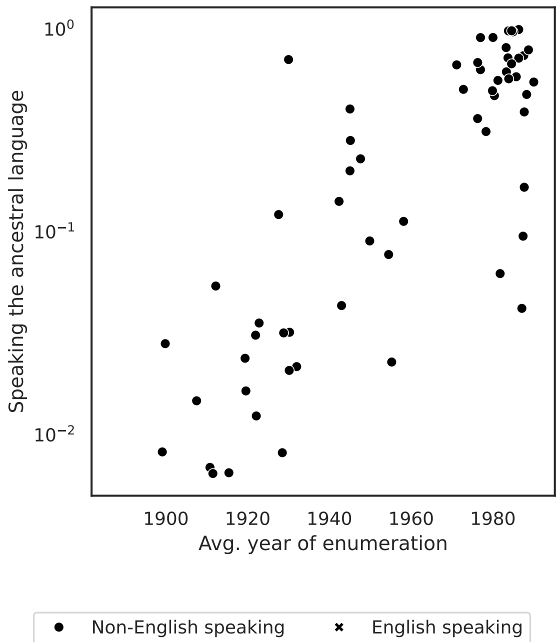
(a) % of ind. born in the ancestral country

Notes: The bivariate cross-country relationship between the share of people born in the origin country and the average year of enumeration.



(b) Herfindahl index

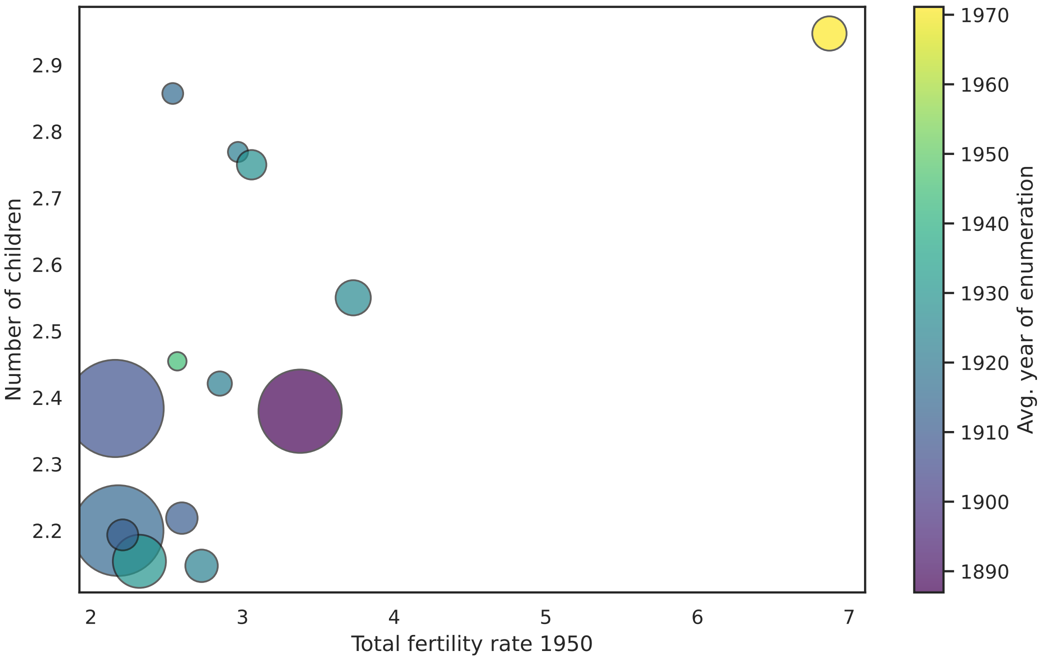
Notes: The bivariate cross-country relationship between the Herfindahl index of the origin country and the average year of enumeration.



(c) Speaking the ancestral language

Notes: The bivariate cross-country relationship between share of natives speaking the ancestral language and the average year of enumeration.

Figure 2.11: Number of children, country-of-origin total fertility rate and average year of enumeration



Data sources: Authors' elaboration on data from the General Social Survey, on the origin-specific data on the TFR in Fernández and Fogli (2006), and on the 1850 to 1990 census (Ruggles et al., 2023).

Notes: This bubble plot displays the unconditional relationship between origin-country total fertility rate and the number of children. The size of each bubble represents the population size of the ancestry group, while the color indicates the average year of enumeration in a US Census. Each bubble represents a different ancestry group.

Table 2.2: Replication of Fernández and Fogli (2006): Fertility, culture, and siblings

	Dependant variable is Children								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
TFR 1950	0.166***		0.101***	0.117***		0.097***	0.135***		0.118**
	(0.025)		(0.026)	(0.016)		(0.020)	(0.036)		(0.042)
SIBS		0.093***	0.086***		0.044***	0.039**		0.045**	0.039*
		(0.012)	(0.014)		(0.012)	(0.014)		(0.017)	(0.019)
Age				0.281**	0.304***	0.299***	0.345***	0.364***	0.359***
				(0.098)	(0.095)	(0.095)	(0.079)	(0.083)	(0.083)
Age sq				−0.003*	−0.003**	−0.003**	−0.003***	−0.004***	−0.004***
				(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
High School				−0.814***	−0.764***	−0.738***	−0.629***	−0.574***	−0.564***
				(0.128)	(0.122)	(0.124)	(0.152)	(0.170)	(0.162)
Some College				−0.921***	−0.865***	−0.822***	−0.619***	−0.553**	−0.532**
				(0.063)	(0.072)	(0.055)	(0.181)	(0.194)	(0.178)
College				−1.362***	−1.292***	−1.250***	−1.143***	−1.071***	−1.057***
				(0.100)	(0.106)	(0.108)	(0.128)	(0.147)	(0.134)
Mother's education							−0.060**	−0.063***	−0.056**
							(0.024)	(0.019)	(0.021)
Father's education							0.027**	0.027**	0.028**
							(0.010)	(0.010)	(0.010)
Obs.	1145	1144	1144	1144	1143	1143	922	921	921
Adj. R^2	0.037	0.060	0.062	0.203	0.206	0.208	0.223	0.225	0.229
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table replicates Table 3a in Fernández and Fogli (2006). It reports the results of estimating Eq. 2.7 on individual-level data from the GSS (1977-1987). An observation is a married women of foreign ancestry, born in the US and between 29 and 50 years of age. The dependent variable is the number of children born to a woman. TFR 1950 if the total fertility rate in the woman's country of ancestry in 1950. In-sample countries of ancestry are Canada, Denmark, UK, Finland, France, Germany, Ireland, Italy, Mexico, Netherlands, Norway, Russia, Spain, and Sweden. SIBS is the number of siblings a women has. The individual controls include: age, age squared, and a set of dummy variables to capture the level of education (below high school [omitted], high school degree ("High School"), some college ("Some College"), and at least a college degree ("college"). The parental controls include mother's education and father's education in number of years. Estimations also include region of residence fixed effects, as well as year of survey fixed effects. Standard errors are clustered at country of ancestry level. $^{*}\,p < 0.1$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

Table 2.3: Replication of Fernández and Fogli (2006), dropping native women of Mexican ancestry

Dependant variable is Children						
	(1)	(2)	(3)	(4)	(5)	(6)
TFR 1950	0.126**	0.074	0.101***	0.085***	0.194***	0.181***
	(0.056)	(0.045)	(0.020)	(0.020)	(0.040)	(0.040)
SIBS		0.089***		0.041**		0.039*
		(0.015)		(0.015)		(0.020)
Age			0.266**	0.285***	0.352***	0.368***
			(0.094)	(0.091)	(0.091)	(0.093)
Age sq			−0.002*	−0.003**	−0.004**	−0.004***
			(0.001)	(0.001)	(0.001)	(0.001)
High School			−0.788***	−0.717***	−0.627***	−0.574***
			(0.107)	(0.113)	(0.112)	(0.117)
Some College			−0.864***	−0.769***	−0.577***	−0.502**
			(0.075)	(0.067)	(0.174)	(0.167)
College			−1.299***	−1.188***	−1.073***	−0.996***
			(0.119)	(0.128)	(0.125)	(0.123)
Mother's education					−0.073**	−0.069**
					(0.025)	(0.023)
Father's education					0.029**	0.029***
					(0.010)	(0.009)
Obs.	1062	1061	1061	1060	867	866
Adj. R^2	0.030	0.055	0.191	0.196	0.220	0.226
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

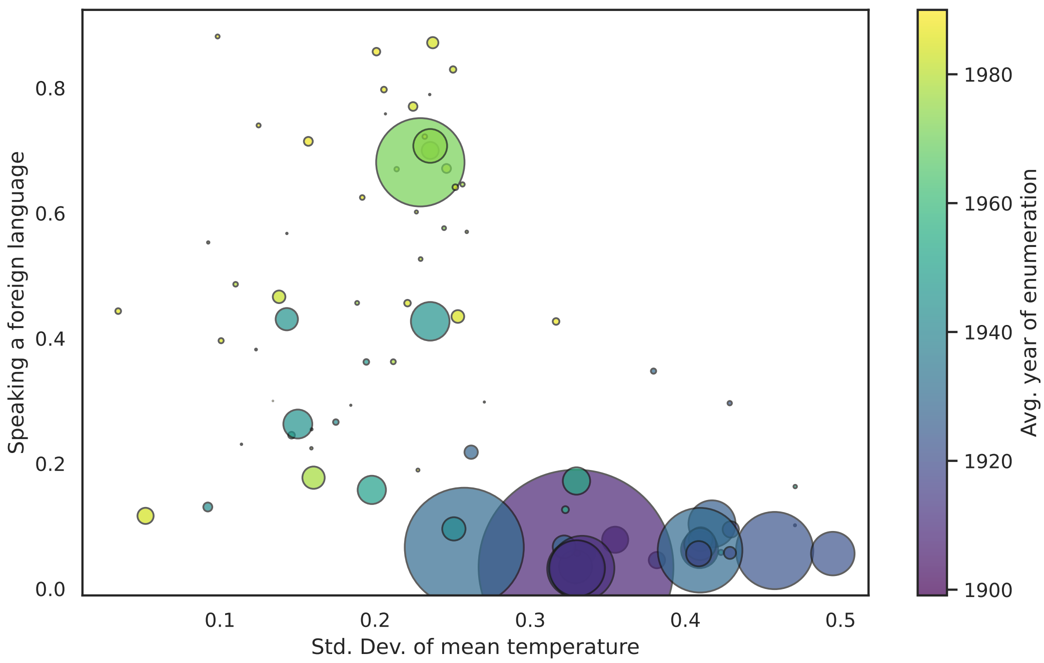
Notes: This table replicates Table 3a in Fernández and Fogli (2006) (excluding specifications without "TFR 1950"), dropping women of Mexican ancestry. $^{*}\,p < 0.1$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

Table 2.4: Cultural convergence in Fernández and Fogli (2006): heterogeneity by generations of migrants

	Dependant variable is Children					
	(1)	(2)	(3)	(4)	(5)	(6)
TFR 1950	0.193***	0.143***	0.215***	0.198***	0.199**	0.184**
	(0.037)	(0.043)	(0.045)	(0.046)	(0.069)	(0.072)
SIBS		0.079***		0.033**		0.031
		(0.015)		(0.014)		(0.019)
TFR 1950 × 3rd gen. migrant	0.033	0.035	−0.013	−0.010	0.017	0.017
	(0.047)	(0.043)	(0.035)	(0.036)	(0.049)	(0.049)
TFR 1950 × 4th gen. migrant	−0.143**	−0.154**	−0.210**	−0.210**	−0.125	−0.121
	(0.056)	(0.056)	(0.087)	(0.085)	(0.083)	(0.082)
TFR 1950 + TFR 1950 × 3rd gen. mig.	0.225***	0.178**	0.202***	0.188***	0.215**	0.201*
	(0.058)	(0.060)	(0.054)	(0.054)	(0.097)	(0.098)
TFR 1950 + TFR 1950 × 4th gen. mig.	0.050	-0.012	0.005	-0.012	0.074*	0.063
	(0.041)	(0.040)	(0.051)	(0.053)	(0.036)	(0.039)
Mean outcome 2nd gen. mig.	2.633	2.633	2.598	2.598	2.537	2.537
Mean outcome 3rd gen. mig.	2.265	2.265	2.265	2.265	2.265	2.265
Mean outcome 4th gen. mig.	2.307	2.306	2.307	2.306	2.296	2.295
Obs.	1097	1096	1096	1095	893	892
Adj. R^2	0.036	0.057	0.194	0.198	0.215	0.220
Individual controls	No	No	Yes	Yes	Yes	Yes
Parental controls	No	No	No	No	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table explores heterogeneities by generation of migration in Table 3a in Fernández and Fogli (2006). We work on Fernández and Fogli (2006) initial sample defined in Table 2.2, while keeping women for who we are able to retrieve information about their generation of migration. “3rd gen. migrant” is a dummy indicating a woman being a third-generation migrant, “4th gen. migrant” is a dummy indicating a women being (at least) a fourth-generation migrant. These two last definitions are based on Giavazzi et al. (2019). These two dummies are also separately included in the regressions. The dependent variable is the number of children born to a woman (we report in the second part of the table the mean dependent variable of each generation of migrants in our regression samples). TFR 1950 if the total fertility rate in the woman’s country of ancestry in 1950. SIBS is the number of siblings a women has. We report in the second part of the table the estimates and standard errors of the total effect of TFR 1950 for third-generation and fourth-generation migrants. The individual controls include: age, age squared, and a set of dummy variables to capture the level of education (below high school [omitted], high school degree (“High School”), some college (“Some College”), and at least a college degree (“college”). The parental controls include mother’s education and father’s education in number of years. Estimations also include region of residence fixed effects, as well as year of survey fixed effects. Standard errors are clustered at country of ancestry level. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure 2.12: Speaking a foreign language at home, cross-generational climatic instability and average year of enumeration



Data sources: Authors' elaboration on the replication files of Giuliano and Nunn (2021), and on data from the 1850 to 1990 censuses (Ruggles et al., 2023).

Notes: This bubble plot displays the unconditional relationship between cross-generational climatic instability and speaking a foreign language at home. The size of each bubble represents the population size of the ancestry group, while the color indicates the average year of enumeration in a US Census. Each bubble represents a different ancestry group.

Table 2.5: Giuliano and Nunn (2021) original sample - Controlling for average year of enumeration and local fraction of the same ancestry

	Sample: All countries, GN sample				
	(1)	(2)	(3)	(4)	(5)
Climatic variability	-0.447^{***}	-0.387^{**}	-0.447^{***}	-0.257^{*}	-0.165
	(0.168)	(0.165)	(0.168)	(0.154)	(0.139)
Frac. of same anc.		0.152^{***}			0.213^{***}
		(0.057)			(0.041)
Av. year of enum.				0.007^{***}	0.007^{***}
				(0.002)	(0.002)
R^2	0.294	0.296	0.294	0.324	0.328
Av. year of enum. defined	Yes	Yes	Yes	Yes	Yes
Metarea fixed effect	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes
Mean of dep. var.	0.12	0.12	0.12	0.12	0.12
Number of countries	84	84	83	83	83
Number of Metareas	298	298	298	298	298
Observations	3,343,097	3,343,097	3,343,073	3,343,073	3,343,073

Notes: This table replicates results from Table 5 of Giuliano and Nunn (2021), adding the fraction of metropolitan area population of the same ancestry and the average year of enumeration in a US Census as controls. The unit of observation is a person born in the U.S with ancestry from a non-English speaking country. The dependent variable is an indicator that equals one if the person speaks a foreign language at home. All specifications include the following control variables: the fraction of metropolitan area population of the same ancestry, the fraction of metropolitan area population who is first-generation immigrants of the same ancestry, historical distance from the equator, historical economic development, historical political complexity, the GDP in the country of origin measured at the time of the survey, the linguistic distance between the country of origin and the U.S, age, age squared, sex, marital status, education level, labor force status, natural log of annual income, a rural/urban indicator variable. and metarea fixed effects. Standard errors clustered at the ancestry-country level are in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 2.6: Giuliano and Nunn (2021) original sample without natives with an ancestry in a Spanish-speaking country - Controlling for average year of enumeration and local fraction of the same ancestry

	Sample: All countries, GN sample				
	(1)	(2)	(3)	(4)	(5)
Climatic variability	-0.202^{***}	-0.201^{***}	-0.202^{***}	-0.194^{***}	-0.187^{***}
	(0.062)	(0.065)	(0.062)	(0.064)	(0.068)
Frac. of same anc.		0.002			0.015
		(0.017)			(0.020)
Av. year of enum.				0.001	0.001
				(0.001)	(0.001)
R^2	0.069	0.069	0.069	0.069	0.069
Av. year of enum. defined	Yes	Yes	Yes	Yes	Yes
Metarea fixed effect	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes
Mean of dep. var.	0.06	0.06	0.06	0.06	0.06
Number of countries	65	65	64	64	64
Number of Metareas	298	298	298	298	298
Observations	2,981,047	2,981,047	2,981,023	2,981,023	2,981,023

Notes: This table replicates results from Table 5 of Giuliano and Nunn (2021), adding the fraction of metropolitan area population of the same ancestry and the average year of enumeration in a US Census as controls. $^{*}\, p < 0.1$, $^{**}\, p < 0.05$, $^{***}\, p < 0.01$.

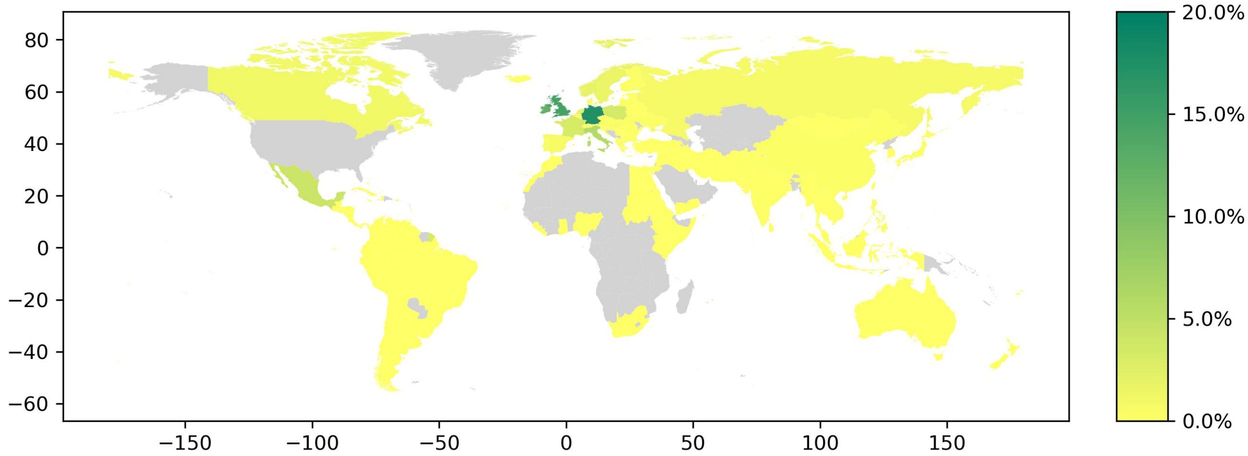
Table 2.7: Replication of Fernández and Fogli (2006), adding Mexican origin fixed effect

Dependant variable is Children						
	(1)	(2)	(3)	(4)	(5)	(6)
TFR 1950	0.127**	0.081	0.101***	0.090***	0.197***	0.187***
	(0.057)	(0.048)	(0.020)	(0.021)	(0.041)	(0.041)
Mexican ancestry	0.144	0.089	0.092	0.070	−0.394	−0.414
	(0.246)	(0.212)	(0.121)	(0.116)	(0.327)	(0.329)
SIBS		0.079***		0.032**		0.033
		(0.016)		(0.014)		(0.019)
Age			0.272**	0.286***	0.347***	0.359***
			(0.092)	(0.089)	(0.091)	(0.092)
Age sq			−0.003*	−0.003**	−0.004**	−0.004**
			(0.001)	(0.001)	(0.001)	(0.001)
High School			−0.757***	−0.699***	−0.564***	−0.513***
			(0.109)	(0.110)	(0.114)	(0.123)
Some College			−0.845***	−0.768***	−0.536***	−0.465***
			(0.070)	(0.065)	(0.157)	(0.152)
College			−1.280***	−1.191***	−1.037***	−0.966***
			(0.121)	(0.125)	(0.119)	(0.119)
Mother's education					−0.066**	−0.064**
					(0.024)	(0.022)
Father's education					0.027**	0.027**
					(0.010)	(0.009)
Obs.	1097	1096	1096	1095	893	892
Adj. R^2	0.035	0.055	0.193	0.196	0.216	0.221
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table replicates Table 3a in Fernández and Fogli (2006), adding Mexican origin fixed effect ("Mexican origin" in the table) as a control. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

2.A Additional results

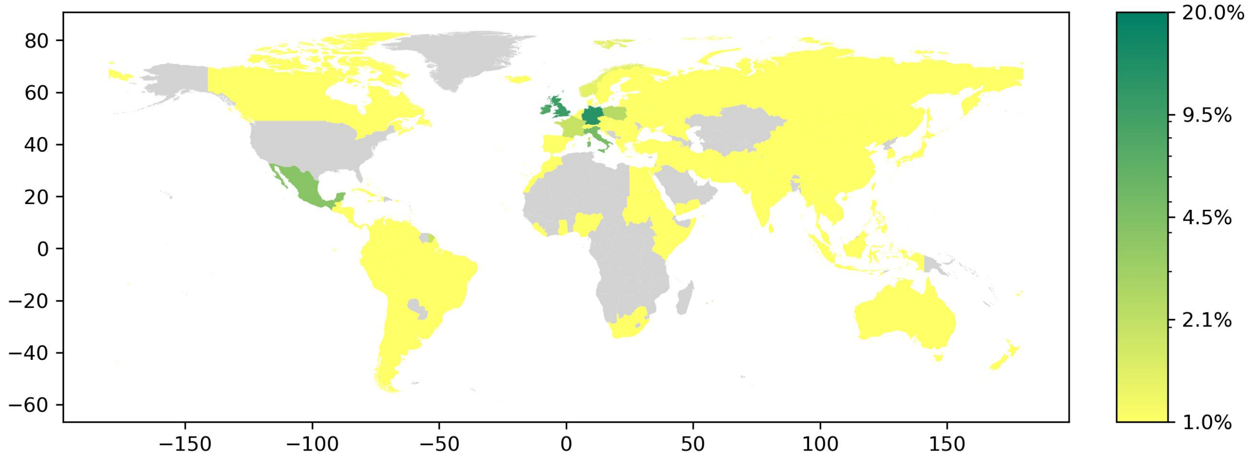
Figure 2.A.1: Proportional Distribution of Native Ancestry by Country in 2000, considering first and second ancestry



Data sources: Authors' elaboration on the 2000 Census (Ruggles et al., 2023).

Notes: For each country, we assign a value representing the proportion of individuals who identify that country as their first or second ancestry. By construction, the sum of these shares exceed 100 percent.

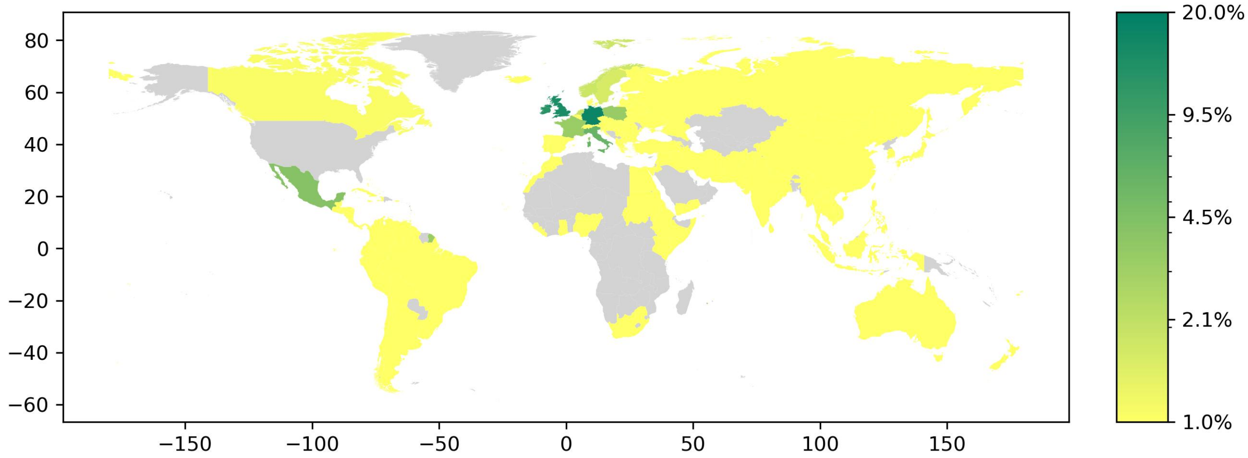
Figure 2.A.2: Proportional Distribution of Native Ancestry by Country in 2000 (logarithmic scale)



Data sources: Authors' elaboration on the 2000 Census (Ruggles et al., 2023).

Notes: For each country, we assign a value representing the proportion of individuals who identify that country as their first ancestry.

Figure 2.A.3: Proportional Distribution of Native Ancestry by Country in 2000, considering first and second ancestry (logarithmic scale)

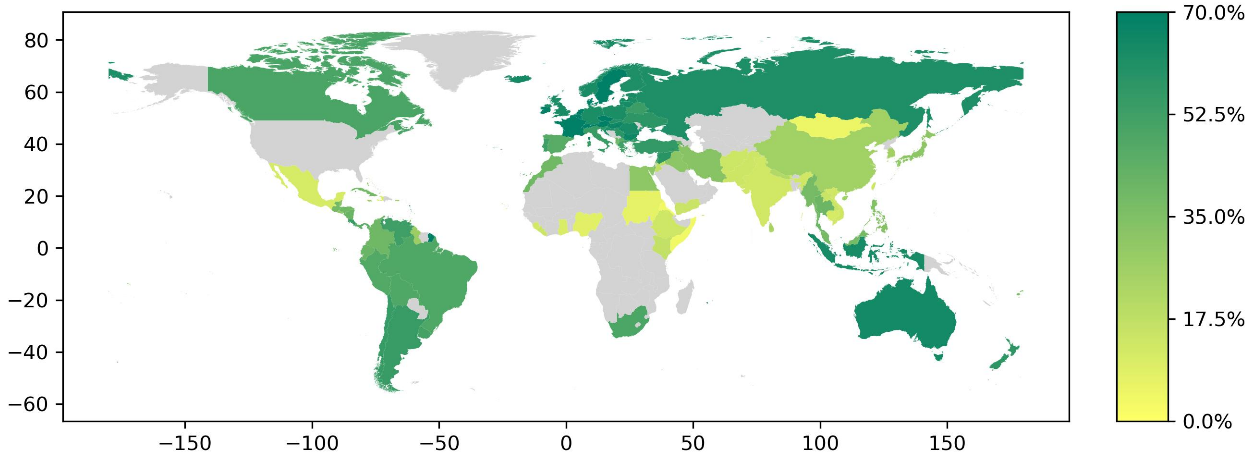


Data sources: Authors' elaboration on the 2000 Census (Ruggles et al., 2023).

Notes: For each country, we assign a value representing the proportion of individuals who identify that country as their first or second ancestry. By construction, the sum of these shares exceed 100 percent.

Figure 2.A.4: Incidence of multiple ancestries

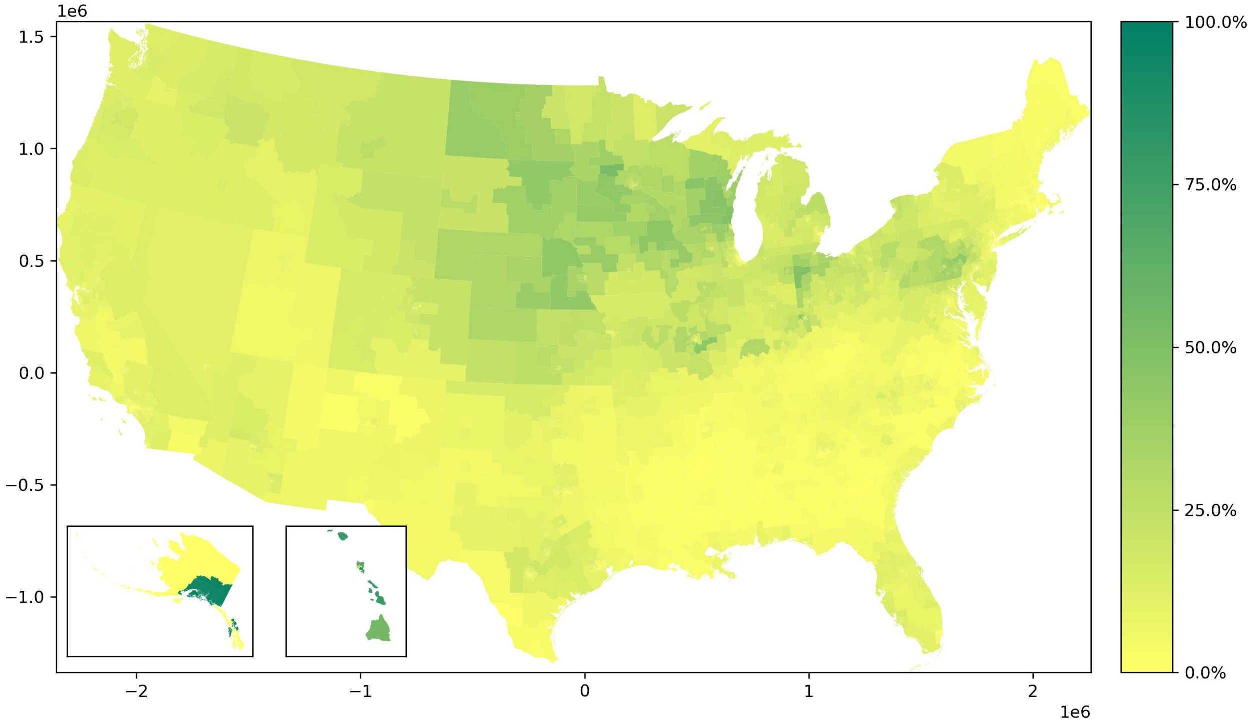
(first and second ancestry)



Data sources: Authors' elaboration on the 2000 Census (Ruggles et al., 2023).

Notes: For each ancestry, we compute the share of natives who also report a second ancestry, e.g., we compute the share of natives with Italian first or second ancestry that also report another ancestry.

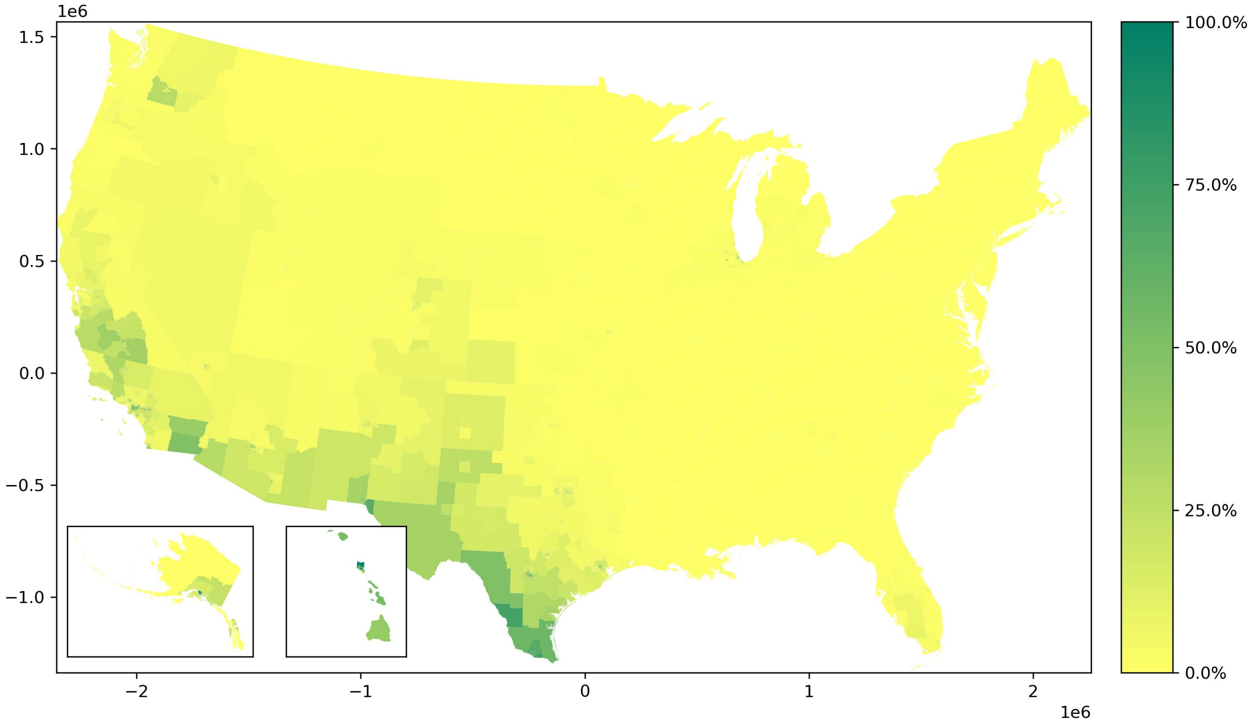
Figure 2.A.5: German ancestry among natives in each PUMA



Data sources: Authors' elaboration on the 2000 Census (Ruggles et al., 2023).

Notes: This figure present the percentage of native-born population that identify with the German ancestry as first ancestry. The sample is composed of individuals born in the USA that report at least one ancestry.

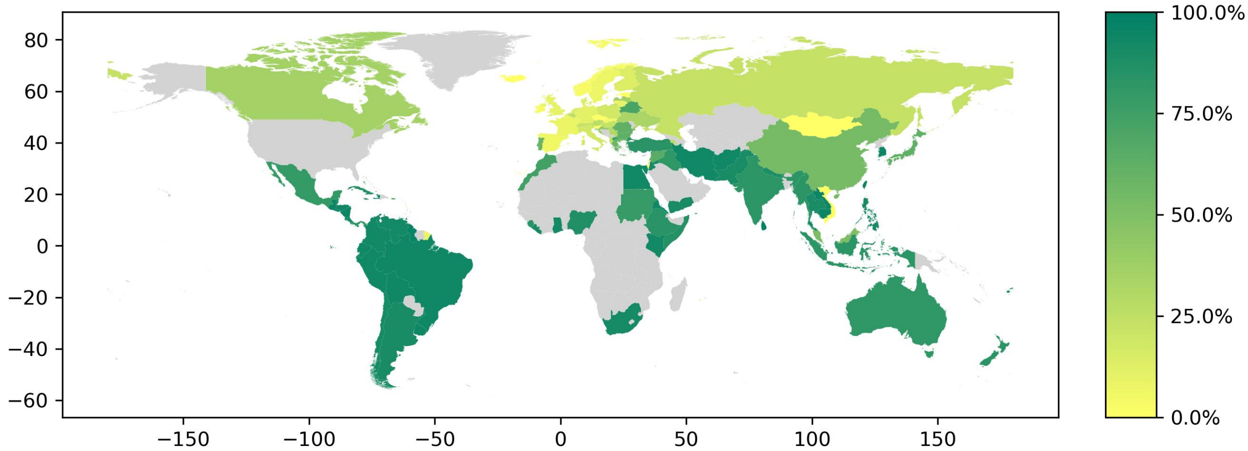
Figure 2.A.6: Mexican ancestry among natives in each PUMA



Data sources: Authors' elaboration on the 2000 Census (Ruggles et al., 2023).

Notes: This figure present the percentage of native-born population that identify with the Mexican ancestry as first ancestry. The sample is composed of individuals born in the USA that report at least one ancestry.

Figure 2.A.7: Share of natives co-residing with at least one parent whose first ancestry coincides with the maternal or paternal country of birth



Data sources: Authors' elaboration on the 2000 Census (Ruggles et al., 2023).

Notes: For each country, we assign a value representing the proportion of individuals who live with at least one parent born in an ancestral country that matches their first or second reported ancestry, among those declaring a first ancestry linked to that country.

Table 2.A.1: Countries of origin and generations of migrants in Fernández and Fogli (2006)

	Generation			
	Second	Third	Fourth	Total
DEU	19	52	218	289
GBR	9	28	222	259
IRL	4	34	174	212
ITA	23	56	11	90
CAN	9	13	17	39
MEX	14	12	9	35
FRA	2	6	24	32
NOR	3	10	18	31
SWE	2	16	13	31
NLD	5	7	14	26
RUS	5	12	1	18
DNK	0	4	9	13
FIN	2	6	3	11
ESP	1	1	9	11
Total	98	257	742	1097

Notes: This table reports the distribution (number of observations) of generations of migrants by country of origin in GSS 1977, 1978, 1980 and 1982-1987, following Fernández and Fogli (2006). The sample comprises married women aged between 29 and 50 years old, born in the U.S. and reporting a foreign ancestry. Generations of migrants follow the definition of Giavazzi et al. (2019). "Second" are second-generation migrants (i.e., respondents born in the US and at least one of their parents was born abroad). "Third" are third-generation migrants (i.e., respondents born in the US, all of their parents are born in the US and at least two of their grandparents were born abroad). "Fourth" are fourth (and more) generation migrants (i.e., born in the US, all their parents are born in the US and at most one their grandparent was born abroad). This sample is used in the estimations reported in column 1 in Tables 4 and 2.A.3.

Table 2.A.2: Giuliano and Nunn (2021) original sample - Controlling for average year of enumeration, local fraction of the same ancestry and Spanish-speaking country fixed effect

	Sample: All countries, GN sample				
	(1)	(2)	(3)	(4)	(5)
Climatic variability	−0.197***	−0.185**	−0.197***	−0.170**	−0.144*
	(0.072)	(0.075)	(0.072)	(0.080)	(0.082)
Frac. of same anc.		0.034			0.066*
		(0.026)			(0.034)
Av. year of enum.				0.002*	0.002*
				(0.001)	(0.001)
Spanish-sp. country FE	0.402***	0.399***	0.402***	0.361***	0.351***
	(0.037)	(0.038)	(0.037)	(0.046)	(0.047)
R^2	0.358	0.358	0.358	0.360	0.360
Av. year of enum. defined	Yes	Yes	Yes	Yes	Yes
Metarea fixed effect	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes
Mean of dep. var.	0.12	0.12	0.12	0.12	0.12
Number of countries	84	84	83	83	83
Number of Metareas	298	298	298	298	298
Observations	3,343,097	3,343,097	3,343,073	3,343,073	3,343,073

Notes: This table replicates results from Table 5 of Giuliano and Nunn (2021), adding the fraction of metropolitan area population of the same ancestry, the average year of enumeration in a US Census as controls and Spanish-speaking country FE as controls. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 2.A.3: Replication of Fernández and Fogli (2006), sample of individuals with information on their generation of migration

Dependant variable is Children						
	(1)	(2)	(3)	(4)	(5)	(6)
TFR 1950	0.149***	0.094***	0.115***	0.100***	0.139***	0.127**
	(0.027)	(0.031)	(0.019)	(0.024)	(0.039)	(0.044)
SIBS		0.079***		0.032**		0.033
		(0.016)		(0.014)		(0.019)
Age			0.273**	0.287***	0.343***	0.355***
			(0.092)	(0.089)	(0.090)	(0.092)
Age sq			−0.003*	−0.003**	−0.003**	−0.004**
			(0.001)	(0.001)	(0.001)	(0.001)
High School			−0.757***	−0.699***	−0.567***	−0.516***
			(0.109)	(0.110)	(0.115)	(0.124)
Some College			−0.847***	−0.769***	−0.539***	−0.469***
			(0.070)	(0.064)	(0.160)	(0.155)
College			−1.280***	−1.191***	−1.047***	−0.977***
			(0.121)	(0.125)	(0.120)	(0.119)
Mother's education					−0.065**	−0.063**
					(0.023)	(0.021)
Father's education					0.028**	0.028***
					(0.010)	(0.009)
Obs.	1097	1096	1096	1095	893	892
Adj. R^2	0.036	0.056	0.193	0.197	0.217	0.221
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table replicates Table 3a in Fernández and Fogli (2006) (excluding specifications without "TFR 1950"), keeping women for who we are able to retrieve information about their generation of migration. The dependent variable is the number of children born to a woman (we report in the second part of the table the mean dependent variable of each generation of migrants in our regression samples). TFR 1950 if the total fertility rate in the woman's country of ancestry in 1950. SIBS is the number of siblings a women has. We report in the second part of the table the estimates and standard errors of the total effect of TFR 1950 for third-generation and fourth-generation migrants. The individual controls include: age, age squared, and a set of dummy variables to capture the level of education (below high school [omitted], high school degree ("High School"), some college ("Some College"), and at least a college degree ("college"). The parental controls include mother's education and father's education in number of years. Estimations also include region of residence fixed effects, as well as year of survey fixed effects. Standard errors are clustered at country of ancestry level. $^{*}\,p < 0.1$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

Table 2.A.4: Cultural convergence in Fernández and Fogli (2006): heterogeneity by average year of first-generation migrants enumeration

Dependant variable is Children						
	(1)	(2)	(3)	(4)	(5)	(6)
TFR 1950	0.083	0.039	0.085***	0.074**	0.141***	0.132***
	(0.056)	(0.049)	(0.025)	(0.028)	(0.022)	(0.027)
SIBS		0.079***		0.032**		0.033
		(0.016)		(0.014)		(0.019)
TFR 1950 × 2nd tert. avg. yr of enumeration	0.026	0.056	−0.151	−0.134	0.039	0.053
	(0.254)	(0.238)	(0.229)	(0.221)	(0.224)	(0.214)
TFR 1950 × 3rd tert. avg. yr of enumeration	0.089	0.076	0.050**	0.043**	−0.010	−0.019
	(0.059)	(0.044)	(0.019)	(0.017)	(0.036)	(0.037)
TFR 1950 + TFR 1950 × 2nd tert. avg. yr of enumeration	0.109	0.094	-0.066	-0.060	0.180	0.185
	(0.251)	(0.238)	(0.231)	(0.223)	(0.226)	(0.217)
TFR 1950 + TFR 1950 × 3rd tert. avg. yr of enumeration	0.173***	0.114***	0.135***	0.117***	0.131**	0.114*
	(0.032)	(0.033)	(0.024)	(0.028)	(0.049)	(0.054)
Mean outcome 1st tert. avg. yr of enumeration	2.296	2.295	2.291	2.290	2.266	2.264
Mean outcome 2nd tert. avg. yr of enumeration	2.425	2.425	2.425	2.425	2.537	2.537
Mean outcome 3rd tert. avg. yr of enumeration	2.414	2.414	2.414	2.414	2.355	2.355
Obs.	1097	1096	1096	1095	893	892
Adj. R^2	0.033	0.053	0.191	0.194	0.215	0.219
Individual controls	No	No	Yes	Yes	Yes	Yes
Parental controls	No	No	No	No	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table explores heterogeneities in Table 3a in Fernández and Fogli (2006), by respondent's country of origin's first-generation migrants average year of enumeration in a US Census. "2nd tert. avg. of enumeration" is a dummy indicating a woman originating from a country being in the 2nd tercile of the country-level in-sample distribution of average year of enumeration of first generation migrants in Fernández and Fogli (2006) sample. "3rd tert. avg. of enumeration" is a dummy indicating a woman originating from a country being in the 3rd tercile of the country-level in-sample distribution of average year of enumeration of first generation migrants in Fernández and Fogli (2006) sample. These two dummies are also separately included in the regressions. The dependent variable is the number of children born to a woman (we report in the second part of the table the mean dependent variable of each generation of migrants in our regression samples). TFR 1950 if the total fertility rate in the woman's country of ancestry in 1950. SIBS is the number of siblings a women has. We report in the second part of the table the estimates and standard errors of the total effect of TFR 1950 for third-generation and fourth-generation migrants. The individual controls include: age, age squared, and a set of dummy variables to capture the level of education (below high school [omitted], high school degree ("High School"), some college ("Some College"), and at least a college degree ("college"). The parental controls include mother's education and father's education in number of years. Estimations also include region of residence fixed effects, as well as year of survey fixed effects. Standard errors are clustered at country of ancestry level. $^{*}\,p < 0.1$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

Chapter 3

Facing Pressure, Choosing Distance: Female Genital Cutting and Child Fostering in Sub-Saharan Africa

3.1 Introduction

What can parents do when what they want for their child conflicts with what their community expects? Parents rarely make decisions alone: choices about children are embedded in social norms and local expectations. When deviating from these expectations is costly, parents may be pressured to follow a norm they oppose to avoid the costs of non-compliance. One way out of this dilemma is to remove the child from the social environment in which the norm is enforced. This paper studies such a strategy in the context of female genital cutting (FGC henceforth): the displacement of daughters through child fostering, a common arrangement in which children are sent to live away from their mother.[1]

FGC is a prime example of a socially enforced practice: the decision to cut a daughter is the product of both parental beliefs and what others in the community do and expect. Consistent with this, community FGC rates remain highly predictive of daughters' cutting status even after controlling for private preferences (Lindskog et al., 2024; Shell-Duncan et al., 2011). FGC is also a policy-relevant setting because of the scale of the practice: UNICEF (2016) estimates that more than 200 million girls and women have been cut in the 30 countries in Africa, Asia and the Middle East where FGC is concentrated. Moreover, 3 million girls are estimated to be at risk each year. The United Nations classifies FGC as a human rights violation, and its eradication was included as a target in the Sustainable Development Goals adopted by all UN member states in 2015.[2] Despite criminalization in most countries, as well as numerous information campaigns and social programs, the practice persists. If current trends continue, the absolute number of affected girls and women is expected to rise due to population growth (UNICEF, 2016).

Yet, opposition to female genital cutting is increasing: in a sample of 338,444 women from 13 sub-Saharan African countries described below, 44% of those who are cut actually want the practice to stop. Such opposition is correlated with behaviours: mothers who oppose FGC are substantially less likely to cut their daughters.[3] However, private preferences do not fully

[1]FGC is a practice referring to the "partial or total removal of external female genitalia or other injury to the female genital organs for non-medical reasons". See `https://www.who.int/news-room/fact-sheets/detail/female-genital-mutilation`, last accessed on May 27, 2026.

[2]See `https://sdgs.un.org/goals`, last accessed on May 27, 2026.

[3]As fathers' attitudes are not available in most survey waves, I use mothers' beliefs as a proxy for parental beliefs throughout the paper. This choice is supported by evidence that spouses rarely hold opposing beliefs about FGC and that women often play a central role in decisions regarding daughters' cutting (Cappa et al.,

determine cutting decisions: nearly 10% of cut daughters have a mother who is against FGC. This gap between stated beliefs and actual practice, also observed in the literature (Cappa et al., 2020; Gichohi et al., 2025; Lindskog et al., 2024), is explained by FGC being collectively enforced: openly deviating can be costly even for mothers who privately oppose it,[4] which can drive some mothers to cut their daughters against their own preferences. These mothers opposed to cutting while living in communities where it is the norm thus face a tough choice: complying with a norm they dislike or bearing the costs of non-compliance. By moving their daughters to contexts where the social cost of remaining uncut is lower or non-existent, they can avoid this trade-off and remain aligned with their beliefs while limiting the immediate costs of non-compliance borne by their daughters.

This situation can be analysed through the lens of the work of Hirschman (1972), which postulates three alternative options, namely "exit", "voice", or "loyalty" for an individual facing a dissatisfying situation. "Exit" translates into leaving, while "voice" refers to loudly expressing discontent. "Loyalty" describes the option of staying without voicing explicitly any particular discontent, or even supporting the situation. Applied to the context of FGC, parents who oppose the practice are left with these three options: (i) choose loyalty by complying with the prevailing norm and accepting the physical and psychological costs of cutting for their daughter; (ii) choose voice by refusing to cut their daughter, thus bearing the high social and economic penalties of non-compliance; or (iii) choose exit by relocating their daughter to a place where the social costs of being uncut are minimal or non-existent, while facing arguably smaller costs associated with mobility.[5]

Building on this framework, I explore whether the exit option (i.e., the spatial mobility of

2020; Bjalkander et al., 2012). References to parental beliefs should therefore be understood as based on mothers' reported beliefs.

[4]Uncut girls are associated with a lower bride price (Khalifa, 2022; Chesnokova and Vaithianathan, 2010), and a higher dowry (Gibson et al., 2023). They also face intense social stigma, humiliation and rejection by their peers (Vissandjée et al., 2014). Some newly cut girls even report a sense of relief, knowing that their psychological torment has ended and that they have gained social acceptance and full symbolic capital through being cut (Omigbodun et al., 2022). These social sanctions often extend to the family as a whole (Hernlund and Shell-Duncan, 2007), giving parents additional incentives to comply with the norm.

[5]Note that in some contexts, non-compliance is synonymous with being excluded from the community, leaving only the "loyalty" and "exit" options available: "[...] women were forced to choose between coerced compliance with gender role norms, however discordant with their own views and experiences, and fleeing their homeland." (Kahn, 2016)

girls) is a strategy used by mothers who oppose FGC. While it is already well-known that international migration through asylum seeking is a tool used by parents trying to protect their daughters from cutting[6] (Vogt et al., 2017; Kahn, 2016; Käkelä, 2022), much less is known about other forms of mobility, and more specifically child fostering. This practice refers to cases where non-orphaned children live apart from their mother; about 25% of mothers in Sub-Saharan Africa foster-out at least one child (Cotton, 2021).

To study whether parents opposed to cutting leverage child fostering to protect their daughters, I rely on two key features of FGC which enable me to isolate parental decisions: FGC is (i) a gendered practice that only directly involves girls,[7] and (ii) a practice typically performed at ages when girls have no agency over the decision, so that observed outcomes reflect parental choices. This enables me to go beyond comparisons between individuals who oppose vs. support cutting, and to examine the reaction of the *same* parents when they quasi-randomly face a situation where the norm is almost immaterial (the birth of a boy) or highly salient (the birth of a girl). In this sense, the sex of the child serves as a natural experiment that shifts exposure to the norm without directly affecting parental preferences. This design helps overcome the potential bias that arises when comparing the behaviour of people who oppose a practice with those who do not, as these groups may differ on other characteristics. Instead, I compare how the same opposed mothers respond to variation in exposure to cutting risk.

Using a sample from the *Demographic and Health Surveys* (DHS) covering 54 waves in 13 Sub-Saharan African countries, I find that mothers from high-prevalence contexts are more likely to out-foster girls than boys while they are still young enough to be cut. This gender difference is concentrated entirely among mothers who oppose FGC, while mothers who support the practice foster girls and boys at similar rates. I rule out several alternative explanations and explore two factors that could make fostering effective. First, the results are concentrated among mothers from ethnic groups with high within-country geographic variation in cutting prevalence. This suggests that when an ethnic group is spread across areas with

[6]Since the 1990s, several European countries have granted asylum to women and girls at risk of cutting, and the 2011 Qualification Directive officially recognized this as a valid ground for refugee status. Europe is seeing an increasing number of applicants seeking asylum from FGC-practising countries of origin (Middelburg and Balta, 2016), with France receiving the largest number of female asylum seekers from these countries (Sonobe, 2025): as of December 31, 2023, a total of 20,936 minor children are benefiting from international protection on this basis in France according to the OFPRA (2023).

[7]Though boys can be indirectly impacted through marriage markets; see for example Khalifa (2022).

different prevalence levels, mothers may be able to rely on co-ethnic networks located in lower-risk environments to secure a viable "exit option" for daughters. Second, results are driven primarily by inter-ethnic couples, suggesting that having a partner with access to a distinct ethnic network also provides an "exit option" for daughters at risk. Overall, the exit strategy appears effective: fostered girls from high-prevalence groups are significantly less likely to be cut than their non-fostered counterparts. Finally, I exploit the 1999 law banning cutting in Senegal as a natural experiment: if girls' mobility is a reaction to the cutting risk, it should decline when the ban reduces prevalence. My results support this hypothesis, as I find that fostering of girls (but not boys) declines after the ban, especially in higher-prevalence groups.

This paper contributes to three strands of the literature. First, it connects to the extensive literature on the motivations for child fostering. Existing work highlights a variety of reasons for this practice: education and/or work (Serra, 2009); coping with negative income shocks (Akresh, 2009; Marazyan, 2015); reaching an ideal family size (Cotton, 2024); managing gender imbalance (Akresh, 2009); or as a consequence of parental divorce or remarriage (Grant and Yeatman, 2014). To my knowledge, no existing study has examined whether child fostering can serve as protection against harmful customs.[8]

Second, this paper speaks to the literature exploring strategies for protection against harmful norms, such as norm replacement (Gulesci et al., 2023), education (Corno and de Walque, 2007), and legislation (García-Hombrados and Salgado, 2023). By demonstrating that child fostering serves as a protective strategy for populations at risk of cutting, this paper highlights the importance of expanding mobility options for vulnerable groups.

Third, and more speculatively, the results of this paper may also have implications for the persistence of harmful norms. Under the hypothesis that a substantial share of daughters do not return once they are beyond the age at risk of cutting, protection through child fostering could itself contribute to norm persistence. Indeed, the seminal work of Mackie (1996) presents FGC as a self-enforcing social convention sustained by a coordination failure: expectations regarding the marriage market generate strong incentives to comply with the norm, thereby maintaining an equilibrium in which girls are cut. This framework implies the existence of a tipping point in the proportion of individuals complying with the norm, a concept empirically

[8]However, another aspect of the relationship between mobility and FGC has been studied: Diabate and Mesplé-Somps (2019) examine whether return migrants in Mali contribute to reducing the practice, and find that returnees from countries where FGC is less common indeed have a negative and significant influence on FGC practices.

explored by Novak (2020) and approached through mixed methods by Shell-Duncan et al. (2011). When daughters of non-compliers leave the community, the composition of the latter changes, potentially keeping communities below this tipping point and thereby reinforcing persistence. The persistence of harmful norms has been linked to dominant social conventions (Lindskog et al., 2024; Khalifa, 2022), lack of information, or pluralistic ignorance (Bursztyn et al., 2020; Ferreira et al., 2024), but rarely to selective mobility. From a policy perspective, this suggests that community-scale interventions that shift beliefs collectively may be more effective than episodic national campaigns that change beliefs for a few individuals per community, who might then leave. The former successfully eradicated foot binding in China in one generation (Mackie, 1996),[9] and are at the heart of the strategy of Tostan, the most well-known NGO working on FGC. However, I cannot empirically observe whether daughters return once they are beyond the age at risk of cutting, or whether they affect norm evolution through cultural remittances (Barsbai et al., 2017), and therefore cannot formally test this channel.

The rest of the paper is structured as follows: Section 3.2 presents the data and some descriptive statistics, Section 3.3 outlines testable predictions about what should be observed in the data if daughters are displaced for protection, Section 3.4 presents the results, Section 3.5 assesses their robustness, and Section 3.6 concludes.

3.2 Data

This analysis draws on publicly available, nationally representative, cross-sectional datasets from the *Demographic and Health Surveys* (DHS henceforth). I keep only countries for which at least one FGC module is available,[10] and ethnicity of respondents is asked. The final sample covers 13 sub-Saharan African countries: Benin, Burkina Faso, Chad, Côte d'Ivoire, Ethiopia, Guinea, Kenya, Mali, Niger, Nigeria, Senegal, Sierra Leone, and Togo. The detailed step-by-step construction of the estimation sample is described in Appendix A. For all survey waves, the FGC module provides extensive information on the female respondent, including

[9] The core of the policy was to create natural-foot societies, whose members pledged not to bind their daughters' feet nor to let their sons marry women with bound feet.

[10] Note that the content of the FGC module is not consistent across country-year waves. For instance, Ghana (in 2003) and Eswatini (in 2006) technically have an FGC module, but it solely includes a question on the cutting status of the surveyed women, and thus cannot be used for the analysis.

her cutting status, age at cutting, and beliefs about the practice. However, information on the cutting status of daughters varies across waves. The methodology used to build the daughters' cutting status is described in detail in Appendix B.1. An important point relates to the timing of the question on beliefs about FGC relative to the daughter's cutting decision. Beliefs are elicited at the time of the survey, whereas the decision about a daughter's cutting may have taken place long before, leaving room for changes in attitudes over time or even *ex post* rationalization. I return to this issue when considering the daughter's cutting status in Section 3.4.4.

I use the cutting status and age at cutting of mothers to construct a continuous measure of risk for their daughters, defined as the prevalence of cutting within the group. I do not use daughters for this measure because they may be not-yet-cut, so their observed uncut status may be temporary. In addition, daughters' age at cutting is not observed in earlier survey waves, whereas mothers' age at cutting is consistently available. Following the literature (Lindskog et al., 2024), I define a group as individuals who share the same ethnicity, reside in the same region, and belong to the same 10-year birth cohort.[11] This requires assigning each child to the mother's self-reported ethnicity. This step is conceptually similar to the epidemiological approach, in which individual outcomes are related to group-level characteristics inherited from ancestry or origin (Arbatlı et al., 2020; Fernández, 2007; Giavazzi et al., 2019). As in that literature, measurement error may arise if group identity is imperfectly observed or selectively reported. In this setting, ethnicity is self-declared and may therefore not be reported, or reported imprecisely, especially when respondents declare residual categories, national-level identities, or socially desirable groups. Appendix E provides a definition of non-usable ethnicity categories, gauges their prevalence, and discusses their implications for the results. I restrict the analysis to groups with at least 50 observations, in order to reduce sampling noise and measurement error arising from idiosyncratic variation in small cells. Groups excluded by this restriction account for 0.5% of the sample. I then compute a group-specific "age limit", defined as the age by which 90% of ever-cut women within a given ethnicity-region-cohort group report having undergone cutting. This measure captures the age by which the majority of girls in that group have already undergone the practice. Formally, it corresponds to the 9$^{\text{th}}$ decile of the age-at-cutting distribution among ever-cut women. Results are robust

[11] For daughters, ethnicity is defined as that of the mother, as the daughter's ethnicity is not collected in the survey.

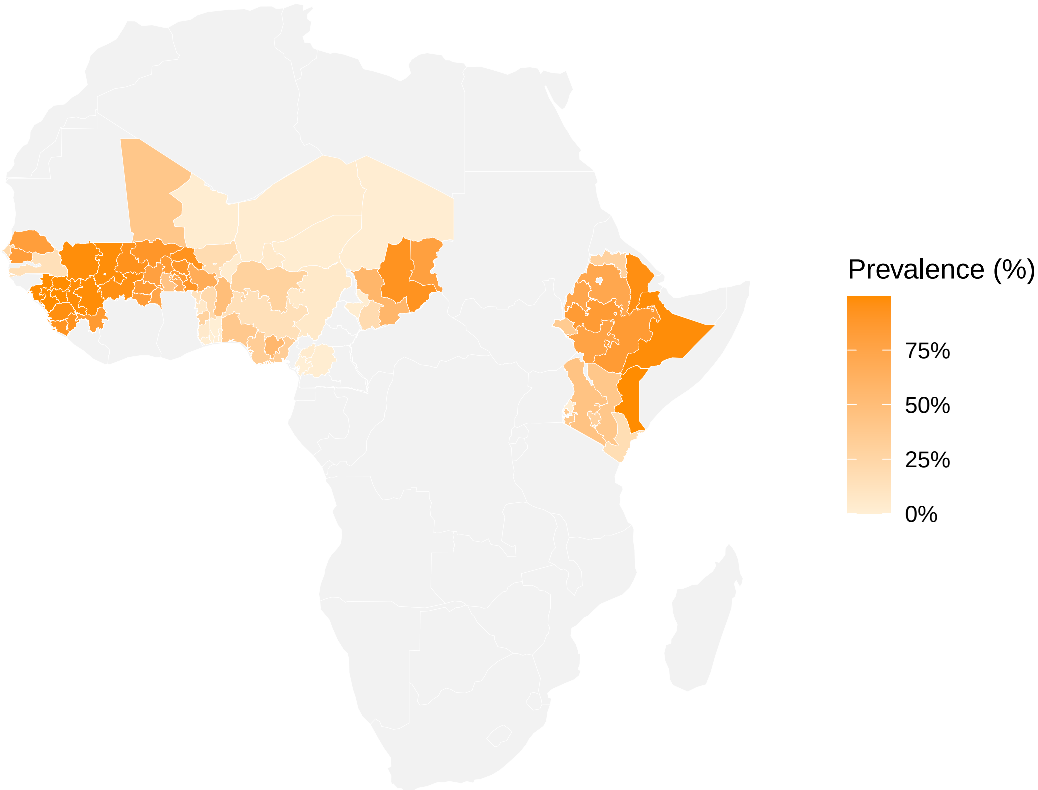
to alternative cut-offs, as shown in the robustness checks. The age limit varies substantially across groups (see Table D.2), with a modal value of 13 years. A child is considered to be at risk if he or she is younger than the age limit defined for his/her group, and not at risk if he or she is older. I then use women older than the age limit to compute group-level prevalence, defined as the proportion of women who have been cut within each ethnicity–region–cohort group. This group-level prevalence cannot easily be translated onto a map, since members of a given ethnic group may reside outside their ethnic homeland.[12] Figure 3.1 therefore presents regional prevalence rates of cutting, which already display substantial variation. As expected, the highest prevalence rates are observed in Guinea, Sierra Leone, and parts of Mali and Ethiopia, while countries such as Niger, Nigeria, and Togo exhibit lower prevalence. The variation exploited in this study is, in fact, much larger, as ethnic groups within the same region may have different prevalence levels, which may also change over time.

With this measure of exposure in hand, I now turn to the measurement of child fostering. Ideally, I would observe individual-level information tracking every child of each female respondent, including the location and composition of both origin and destination households, as well as the timing of fostering. However, no survey simultaneously provides detailed information on cutting and precise tracking of fostered children, and the DHS remains widely used to study child fostering. The literature relies on two main approaches to measure child fostering, each with its own advantages and limitations. The first exploits information from the mother's questionnaire in the household of *origin*, while the second relies on information from the household questionnaire in the household of *destination*. Appendix B.2 describes the information available in the DHS, details the methodology underlying each approach, discusses their respective advantages and limitations, and provides a sanity check to assess the quality of the measure. Since the household-of-origin approach is the only one that allows me to recover information on the child's exposure to cutting risk at origin, notably through the ethnic group, and to observe variation in fostering status across siblings, it constitutes my preferred measure of child fostering.

One alternative form of mobility would be for mothers to move with their daughters rather than foster them out. Because the DHS provides only limited information on the timing and

[12]For instance, groups such as *Fon*, *Dogon*, and *Kalenjin* are widely dispersed within their respective countries (Benin, Mali, and Kenya), while groups such as *Peuhl*, *Hausa*, and *Malinke* even spread across several countries.

Figure 3.1: Regional prevalence of FGC



Data source: Author's elaboration on the DHS.

destination of maternal migration, this channel cannot be studied with the same precision as child fostering. I nevertheless examine this channel in Appendix C and find no evidence that mothers respond to cutting risk by migrating more with daughters than with sons.

I now turn to a descriptive analysis of observable characteristics across four groups: fostered children at origin, non-fostered at origin, fostered children at destination, and non-fostered children at destination. The objective is not to be exhaustive, but to focus on dimensions most relevant for the analysis. Because the information collected in the mother's and household questionnaires differs, I cannot systematically construct identical variables for the two approaches. In the origin household, education and FGC-related variables refer to the mother, whereas in the destination household (where the biological parents of fostered children are unknown), I use averages among all eligible women residing in the household with at least one child. For these variables, I compute means, standard deviations, and gender differences using t-tests, separately for boys and girls. Results are presented for fostered versus non-fostered children identified from the origin household (Table 3.1) and from the destination household (Table 3.2). Among fostered children identified at origin, girls are fostered at younger ages than boys and are more likely to originate from rural households. Their mothers are, on average, older and less educated, and reside in areas with higher cutting prevalence, both at the regional and ethnicity–region levels. Consistent with this, these mothers are more likely to be cut themselves and to support the continuation of the practice, although they are equally aware of FGC as mothers fostering boys. Table 3.2 mirrors Table 3.1 using fostered children at destination. Fostered-in girls are more likely than boys to reside in urban areas, in households with younger and more educated women, and where the household head is female.[13] They are *less* likely to have one of their grandparents as the household head, but more likely to have someone who is not a relative as so. This is noteworthy, as grandparents often carry more traditional attitudes, and protection outside family kinship may be even more effective. Women in these households are less likely to be cut and less supportive of the practice, and they reside in regions with lower cutting prevalence.[14] Comparing origin and destination households for fostered children reveals that girls tend to move to regions with lower prevalence and into households with more educated women who are less supportive of cutting. In contrast, boys are, on average, fostered into regions with similar or higher prevalence and into households with women who are similarly likely to be cut and to support the practice. Differences between non-fostered boys and girls are smaller and mostly relate to beliefs or practices around cutting. These pre-existing differences are unlikely to drive the results, as they are substantially smaller in magnitude and often opposite in sign relative to those observed among fostered children, possibly reflecting a composition effect.

As discussed earlier, measuring fostering from the household of origin is my preferred approach, as it allows me to observe exposure to cutting risk and sibling living arrangements. This measure is therefore used in the remainder of the analysis.

Nevertheless, I conduct a final exercise using the household-of-destination approach to

[13]Note that this does not necessarily imply additional protection, as a female household head usually signals that the household experienced a negative shock (Fafchamps and Quisumbing, 2007).

[14]I do not compute ethnicity–region prevalence here, as it would implicitly assume that fostered children share the ethnicity of household members.

Table 3.1: Descriptive Statistics – Fostered vs. Non-fostered children in origin household

	Fostered			Non-fostered		
Variable	Girls	Boys	Diff.	Girls	Boys	Diff.
Age of the child	8.403	8.594	-0.191***	5.673	5.709	-0.036***
	(0.020)	(0.021)	(0.029)	(0.009)	(0.009)	(0.013)
Urban status	0.287	0.316	-0.029***	0.269	0.268	0.002
	(0.003)	(0.003)	(0.004)	(0.001)	(0.001)	(0.002)
Education	0.407	0.444	-0.036***	0.418	0.410	0.008***
	(0.004)	(0.005)	(0.007)	(0.002)	(0.002)	(0.002)
Mother's age	33.112	32.779	0.334***	32.354	32.420	-0.066***
	(0.043)	(0.046)	(0.063)	(0.018)	(0.017)	(0.025)
Heard about FGC	0.924	0.924	0.000	0.917	0.915	0.001
	(0.002)	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)
FGC should continue	0.519	0.492	0.027***	0.506	0.513	-0.007***
	(0.004)	(0.004)	(0.006)	(0.002)	(0.002)	(0.002)
Woman cut	0.695	0.654	0.041***	0.691	0.697	-0.007***
	(0.003)	(0.004)	(0.005)	(0.001)	(0.001)	(0.002)
Regional prevalence of FGC	0.469	0.452	0.017***	0.464	0.468	-0.004***
	(0.002)	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)
Ethnic-region prev. of FGC	0.437	0.413	0.025***	0.436	0.441	-0.004***
	(0.003)	(0.003)	(0.004)	(0.001)	(0.001)	(0.001)
Number of observations	24,552	21,433	45,985	164,110	171,156	335,266

Notes: The unit of observation is a child (boy or girl) below the minimum of three thresholds: the age limit for cutting, the age limit for marriage, and 15 years of age. Children are identified through the mother's survey, which means that mothers with multiple children appear multiple times in the dataset. Standard errors are reported in parentheses. $^{*}\, p < 0.1$, $^{**}\, p < 0.05$, $^{***}\, p < 0.01$.

Table 3.2: Descriptive Statistics – Fostered vs. Non-fostered children in destination household

	Fostered			Non-fostered		
Variable	Girls	Boys	Diff.	Girls	Boys	Diff.
Age of the child	8.367	8.419	-0.052*	6.120	6.180	-0.061***
	(0.021)	(0.024)	(0.031)	(0.009)	(0.009)	(0.013)
Urban status	0.420	0.355	0.065***	0.279	0.277	0.001
	(0.003)	(0.003)	(0.005)	(0.001)	(0.001)	(0.001)
Education	0.751	0.695	0.055***	0.486	0.480	0.006**
	(0.005)	(0.006)	(0.008)	(0.002)	(0.002)	(0.002)
Head is a female	0.300	0.279	0.021***	0.154	0.148	0.006***
	(0.003)	(0.003)	(0.004)	(0.001)	(0.001)	(0.001)
Head's age	50.371	51.013	-0.643***	46.020	46.005	0.015
	(0.099)	(0.112)	(0.149)	(0.031)	(0.030)	(0.044)
Head is grandparent	0.315	0.365	-0.050***	0.094	0.090	0.004***
	(0.003)	(0.003)	(0.005)	(0.001)	(0.001)	(0.001)
Head is not relative	0.077	0.065	0.012***	0.005	0.004	0.000*
	(0.002)	(0.002)	(0.002)	(0.000)	(0.000)	(0.000)
Heard about FGC	0.924	0.924	-0.001	0.904	0.902	0.002
	(0.002)	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)
FGC should continue	0.453	0.490	-0.037***	0.502	0.507	-0.006***
	(0.004)	(0.004)	(0.005)	(0.001)	(0.001)	(0.002)
Woman cut	0.626	0.651	-0.025***	0.661	0.666	-0.005***
	(0.003)	(0.004)	(0.005)	(0.001)	(0.001)	(0.002)
Regional prevalence of FGC	0.463	0.482	-0.019***	0.465	0.466	-0.002
	(0.002)	(0.003)	(0.004)	(0.001)	(0.001)	(0.001)
Number of observations	24,998	19,582	44,580	174,101	184,331	358,432

Notes: The unit of observation is a child (boy or girl) below the minimum of two thresholds: the age limit for cutting and 15 years of age. Children are identified through the household survey. Education, FGC knowledge, FGC beliefs and cutting status are calculated as the mean of all women in the household with at least one child. Standard errors are reported in parentheses. $^{*}\, p < 0.1$, $^{**}\, p < 0.05$, $^{***}\, p < 0.01$.

explore more formally household characteristics by child sex and fostering status. I regress each variable in Table 3.2 on an interaction between a fostered-in dummy and child sex, including child age fixed effects. I do not include household-level controls, as doing so would mechanically absorb the variation of interest—namely, whether households receiving fostered children differ systematically from others. While these regressions cannot identify causal effects, as information about the cutting risk at origin is unknown, they offer first insights on the differences between groups. Results are presented in Figure D.1 in the Appendix, and are consistent with the previous ones: fostered girls are more likely than boys to reside in urban households in lower-prevalence regions, with more educated women who are more opposed to cutting, and who are themselves less likely to be cut. Fostered girls are *less* likely to have a grandparent as household head, whereas they are more likely to have a non-relative as so. These results, although descriptive, are broadly consistent with the use of fostering as a protective strategy to displace girls from risky contexts.

Finally, although DHS data do not allow connecting fostered children from origin to destination, I explore the extent of mobility by comparing the number of children fostered out of and into each ethnicity–region pair. Similar numbers suggest either no fostering or fostering within the same ethnicity-region, whereas discrepancies indicate potential mobility across ethnic groups, across regions, or both. Knowing what kind of mobility occurs is central, as child fostering is unlikely to serve a protective role if the child is sent within the same region and ethnic group (which is plausible given that fostering often occurs within extended family networks, see Marazyan, 2015; Beck et al., 2015; Akresh, 2009).

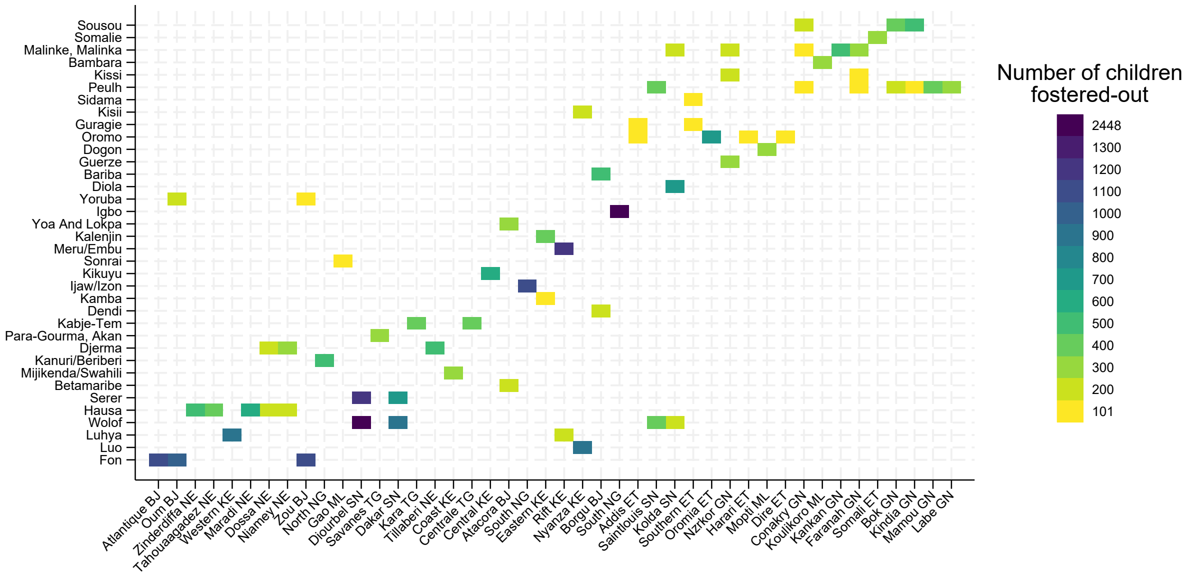
Exploring these patterns requires substantial preliminary work to harmonize ethnic group names across countries. Indeed, the options for self-declaring an ethnicity vary from country to country, and sometimes multiple ethnicities are bundled under the same name, or the same ethnicity is coded under different names.[15] Figure 3.2 investigates the mobility of fostered children by plotting the number of fostered-out and fostered-in children within each region–ethnic group pair. For readability, only region–ethnic group pairs with at least 100 fostered-in chil-

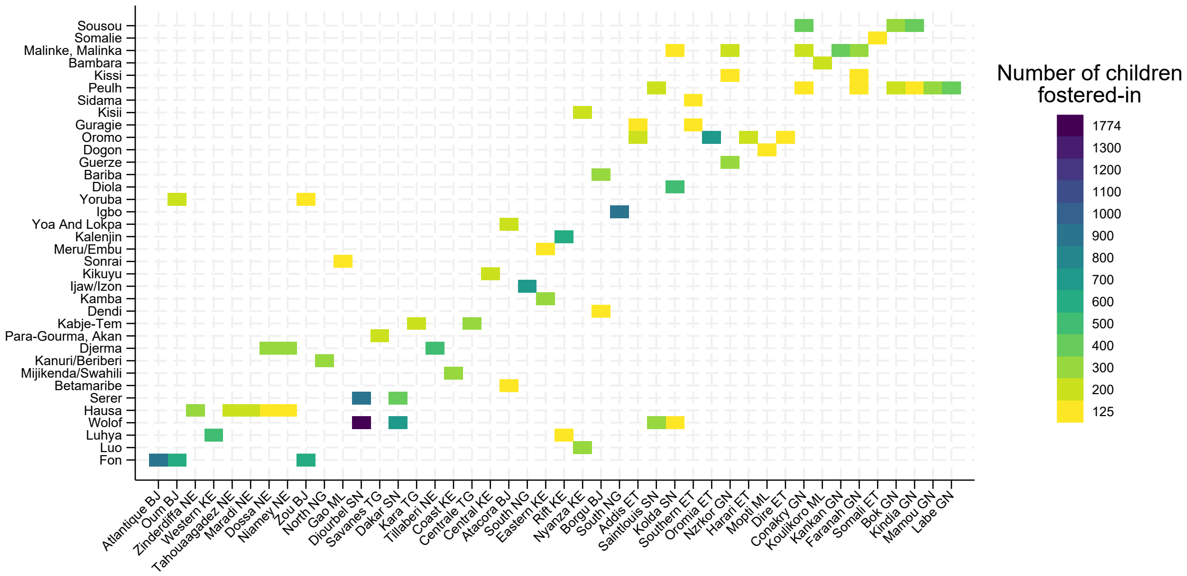
[15] For instance, *Hausa* in Nigeria refers to the same group as *Haoussa* in Niger and Cameroon, but some cases are less obvious: *Kissi* and *Kisii* refer to two totally different ethnic groups respectively from Guinea and Kenya (see https://en.wikipedia.org/wiki/Kissi_people and https://en.wikipedia.org/wiki/Gusii_people, last accessed on May 27, 2026), while the *Kanembou* in Chad and the *Kanuri* in Niger and Cameroon are closely related and bundled together (see https://en.wikipedia.org/wiki/Kanembu_people and https://en.wikipedia.org/wiki/Kanuri_people, last accessed on May 27, 2026).

dren (which account for 80% of cases) are included in the graphs. Regions and ethnic groups are ordered in ascending order of FGC prevalence, meaning that the top-right corners feature ethnicities where the practice is most frequent residing in regions with the highest prevalence. The mobility of fostered children varies substantially across ethnic groups. For instance, the number of *Peuhl* children fostered-out is roughly equal to the number of fostered-in children in each region, indicating little to no mobility outside the ethnicity-region groups. In contrast, the number of *Hausa* fostered-in children is generally lower than the number of fostered-out children, suggesting potential ethnic mobility. The *Malinke* appear to foster-out roughly the same number of children as they fostered-in, but with a different regional distribution, suggesting potential geographic mobility. These observations provide a sense of the mobility patterns of fostered children; however, they do not indicate whether fostering differs systematically between boys and girls.

Figure D.2 in the Appendix explores this further by presenting, for each region–ethnic group pair the proportion of fostered girls relative to fostered boys. Three interesting patterns emerge. First, girls are on average fostered more frequently than boys. This is in line with the literature (Marazyan, 2015; McGavock, 2018; Beck et al., 2015), and may be explained by multiple factors, beyond the risk of cutting alone. Second, groups with lower FGC prevalence and located in lower-prevalence regions tend to foster-in a larger proportion of girls than they foster-out (e.g., the *Hausa* and the *Wolof*), whereas the reverse pattern is observed for groups with higher prevalence and located in higher-prevalence regions (e.g., the *Oromo* and the *Kissi*). This is consistent with the idea of displacing girls away from riskier contexts. Third, the proportion of girls fostered-in is substantially higher than the proportion of fostered-out in regions containing capital cities. This is particularly visible in Niamey in Niger, Addis Ababa in Ethiopia, and Dakar in Senegal. Again, this is not conclusive evidence, as such patterns may be explained by multiple factors, but are not inconsistent with the idea of fostering as a protective strategy.

Figure 3.2: Number of children fostered-out and fostered-in, by ethnic group and region





Data sources: Author's elaboration on the DHS.

Notes: In the first (second) graph, each square represents the number of children under 15 fostered-out (fostered-in) by ethnicity–region pairs. Only ethnicity–region pairs with at least 100 children fostered-in are kept in the sample. Regions and ethnic groups are ordered in ascending order of FGC prevalence.

3.3 Conceptual framework

This section formulates testable hypotheses derived from the patterns one would expect to observe if mothers used fostering as a strategy to protect their daughters from cutting. The first, certainly most intuitive, prediction is the following:

H1: The proportion of fostered girls relative to boys is higher in ethnic groups with a higher prevalence of cutting.

Holding everything else constant, if mothers respond to cutting risk by fostering their daughters, an increase in group-level cutting prevalence should be associated with a higher share of fostered girls relative to boys. However, cutting risk is not the only dimension along which social norms vary across groups. Other group-specific norms—such as gender preferences, schooling patterns, or norms governing the marriage market—may also affect the gender composition of fostered children. For instance, the coexistence of cutting practices and norms encouraging the fostering of girls for elderly care could generate patterns consistent with *H1* even in the absence of a protective motive. Moreover, an alternative interpretation is that mothers may foster their daughters not to prevent them from being cut, but rather to ensure that they undergo the practice. This concern is particularly relevant given that a substantial share of fostered children are sent to grandparents' households, where attachment to traditional practices may be stronger. Finally, a higher prevalence could signal a greater acceptance of the practice, i.e., a context in which almost every family supports cutting and almost none protects daughters. For these reasons, *H1* should be viewed as a necessary but not sufficient condition for identifying a protection mechanism. To address these alternative explanations, I exploit heterogeneity in beliefs about cutting within groups. If the fostering of girls is motivated by protection, the gender differential in mobility should be driven by mothers who oppose the practice. Mothers who support cutting provide a natural placebo group: although they face the same group-level norms and constraints, they should not exhibit differential fostering of girls relative to boys. This comparison also rules out the possibility that girls are fostered in order to be cut. This leads to the second hypothesis.

H2: The differential fostering of girls relative to boys is driven by mothers who oppose cutting, while mothers who support the practice exhibit no such differential behaviour.

Beyond group-level prevalence, the broader normative environment may further shape parental responses. Even within a given ethnic group, the social costs of deviating from the norm may be higher in regions where cutting is widespread. In such contexts, remaining uncut may carry greater penalties in terms of social approval, marriage prospects, or access to community support, thereby strengthening incentives to displace girls away from risk.[16] I therefore hypothesize that regional exposure amplifies group-level effects. I thus postulate the following:

H3: The differential fostering of girls relative to boys is driven by households located in regions where the regional prevalence of cutting is higher than the group-level prevalence.

Next, for fostering to protect daughters against cutting, it must modify the social environment in which the cutting decision becomes relevant. In my sample covering 13 countries and more than one hundred ethnic groups, this is unlikely to operate through a single mechanism. Indeed, the practice of FGC varies substantially across contexts: it may be made visible through public ceremonies or performed more discreetly within the family; it may be associated with religious or moral obligations, or linked to marriageability and expectations on the marriage market (Mackie, 1996; Shell-Duncan et al., 2011; Hernlund and Shell-Duncan, 2007; Khalifa, 2022). This heterogeneity matters for understanding when fostering can effectively protect.

In contexts where cutting status is publicly known or directly verified before marriage, temporarily sending a daughter away is unlikely to be sufficient: if she later returns to the same community or marries within the same group, the sanctions associated with remaining uncut may still apply. In such cases, fostering can only protect if it is associated with a durable change in the relevant social environment, for instance by increasing the probability that the girl marries into a lower-prevalence group or region, or by connecting her to a network where remaining uncut is less costly.

In other contexts, however, fostering may protect even without permanently severing ties with the origin community. This may occur when cutting is decided within the household or extended family, and when temporary exposure to relatives who support the practice, such as grandparents, is decisive. In these cases, residing elsewhere during the relevant age window may be sufficient to reduce the risk of cutting. I therefore do not assume that fostering is mechanically protective *per se*. Rather, it can operate as a protective strategy when it changes

[16] However, it may not matter that much in contexts with very strong ethnic differentiation and very limited inter-ethnic marriage.

the network, location, or marriage-market prospects that determine the costs and benefits of remaining uncut.

In any case, if fostering is indeed used as a protective strategy, it should translate into observable differences in cutting outcomes. In particular, fostered girls should be less likely to have undergone cutting than non-fostered ones, especially in high-prevalence contexts where the protective motive is strongest.

H4: Fostered girls are on average less likely to be cut than non-fostered ones, particularly in high-prevalence groups.

Finally, if girls' fostering is indeed a response to cutting risk, it should decline as that risk decreases. Senegal provides a relevant case study in this respect: the 1999 ban on cutting was followed by a substantial decline in prevalence. If fostering reflects a protective response, this policy-induced change should reduce the displacement of girls (but not boys) particularly in groups where cutting was initially more prevalent:

H5: In Senegal, the decline in cutting prevalence following the ban reduced the fostering of girls (but not boys) particularly in higher pre-ban prevalence groups.

Taken together, these hypotheses move from documenting gendered fostering patterns (*H1*), to identifying a protection mechanism (*H2* and *H3*), its effects (*H4*), and dynamic responses to policy-induced risk reductions (*H5*).

3.4 Results

3.4.1 FGC and girls' mobility

The first set of hypotheses (*H1* and *H2*) predicts that mothers respond to cutting risk by differentially fostering daughters compared to sons, and that this response is concentrated among mothers who oppose the practice. To test its empirical relevance, I estimate the following equation:

$$Y_i = \alpha G_i \times P_{g(i)} + \beta G_i + \gamma P_{g(i)} + \delta_{a(i)} + \delta_{m(i)} + \epsilon_i \tag{3.1}$$

where Y_i is equal to one if child i belonging to the ethnicity-cohort-region group g was fostered while being at risk. G_i is a dummy equal to one if the child is a girl, and $P_{g(i)}$ captures the group-level (i.e., ethnicity-cohort-region) cutting prevalence. The coefficient α associated with the interaction between these two variables is the coefficient of interest. It tests *H1* by capturing the differential likelihood of fostering a girl versus a boy, depending on the cutting prevalence within the group. Boys serve as a placebo, as their mobility should not be influenced by the risk of cutting. $\delta_{a(i)}$ denotes child-age fixed effects, and $\delta_{m(i)}$ mother fixed effects. The inclusion of mother fixed effects makes it possible to drop all time-invariant mother-level controls, as they absorb all characteristics that do not vary within a mother. In effect, it compares siblings with one another.

An interpretative issue is that group-level prevalence $P_{g(i)}$ may proxy not only for the risk of cutting, but also for the broader normative environment, including the intensity of sanctions faced by uncut daughters and their families. This is not necessarily a threat to the main interpretation, since these sanctions are part of the risk that parents may seek to avoid. However, it implies that prevalence should be interpreted as a broad measure of normative pressure rather than as the mechanical probability of being cut. If sanction intensity increases with prevalence, then groups with similar prevalence should also face similar levels of social pressure. If, however, groups with similar prevalence differ in the intensity of sanctions, $P_{g(i)}$ measures the relevant exposure with error. This would tend to attenuate the estimated relationship and bias the coefficient on α toward zero, unless sanction intensity is systematically correlated with gender-specific fostering norms. I return to this issue in Section 3.4.3, where I discuss whether the contexts in which exit options are more available may also be those in which sanctions are weaker.

Results of Eq. (3.1) are presented in Table 3.3. In line with *H1*, the coefficient of interested is positive and significant. It indicates that mothers from higher-prevalence groups tend to foster girls more than boys, compared to mothers from lower-prevalence groups. More precisely, a 10 percentage point increase in the group's prevalence increases the likelihood of fostering a girl rather than a boy by 0.17 ($0.017 \times 0.10 = 0.0017$) percentage points (Column 1). This effect is even larger when restricting the sample to mothers whose attitudes about cutting are known (Column 2). More importantly, it holds only for mothers who oppose the continuation of cutting (Column 3), while no significant effect appears for mothers who support it (Column 4), in line with *H2*.

Table 3.3: FGC prevalence and child fostering

	Dep. var.: Child is fostered			
	All Mothers	Att. known	Anti-FGC	Pro-FGC
	(1)	(2)	(3)	(4)
Child girl × Prev.	0.017***	0.027***	0.033***	0.011
	(0.003)	(0.005)	(0.007)	(0.010)
FGC prevalence	-0.134***	-0.131***	-0.293***	-0.043
	(0.018)	(0.025)	(0.035)	(0.035)
Child is a girl	0.017***	0.009***	0.006*	0.022**
	(0.002)	(0.003)	(0.003)	(0.009)
R^2	0.587	0.605	0.602	0.608
Mother FE	Yes	Yes	Yes	Yes
Mean of dep. var.	0.14	0.16	0.15	0.16
Number of countries	13	13	13	13
Number of ethnic gr.	114	104	101	98
Observations	306,384	171,791	85,698	86,093

Notes: The unit of observation is a child younger than the age limit for cutting and under 18. The dependent variable is a binary indicator equal to one if the child does not coreside with his or her mother. All specifications include the following control variables: child's age fixed effects and mother fixed effects. Standard errors clustered at the mother level are in parentheses. $^{*}\,p < 0.1$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

This is the most precise econometric specification, but it restricts the sample to mothers with at least one boy and one girl. Mothers with a single child or only daughters/sons therefore become immaterial for the analysis. To keep these observations in the sample, I thus run a less demanding specification without mother fixed effects. Results are presented in Table D.3 in the Appendix. I replace mother fixed effects with individual controls (mother's age fixed effects, her level of education, the number of births she has had, and a dummy for whether she resides in a rural or urban area), as well as ethnicity and year of enumeration fixed effects (Columns 1 to 4), and I further add cluster fixed effects (Columns 5 to 8). In both specifications, results remain qualitatively similar to the main ones.

Beyond group-level prevalence, *H3* emphasizes the role of regional prevalence in shaping fostering behaviour: the protective response should be stronger when girls reside in areas where cutting is more widespread relative to their ethnic group. To explore this, I re-estimate Eq. (3.1) while stratifying the sample by relative regional prevalence, assessing whether mothers opposed to cutting foster girls more intensively when local exposure to cutting risk is high. I investigate this in Table 3.4, where the left panel shows mothers belonging to groups where the prevalence at the time their child was born is larger than the regional prevalence at that time, and the right panel shows the opposite. Columns 1 to 4 show no fostering response when the regional prevalence is lower than the group prevalence. However, Columns 5 to 8 show that the results supporting *H1* and *H2* are driven by mothers residing in areas where regional prevalence is higher than their group-level prevalence, in particular those opposed to the practice. In Column 7, a 10 percentage point increase in group prevalence is associated with a 0.52 percentage point increase in the likelihood of fostering girls compared to boys. These results suggest that mothers do not respond only to the average prevalence of cutting in their own ethnic group, but also to the local environment in which the child grows up. When mothers opposed to cutting live in regions where the practice is more widespread than within their own group, keeping a daughter uncut may be especially difficult. In such contexts, fostering may allow mothers to draw on lower-prevalence networks within their own ethnic group, consistent with *H3*.

3.4.2 Alternative explanations

This section examines three alternative explanations that could explain the previous results: child marriage, women's bargaining power and unobserved heterogeneity in social norms. First, FGC and child marriage are known to be correlated (Boyden et al., 2013; Ahinkorah et al., 2023). Consequently, the lower propensity of mothers from high-prevalence groups to coreside with young daughters might reflect early marriage rather than fostering. Although this is less likely since the results are driven by mothers opposed to cutting, who should *not* be the ones supporting early marriage if both are positively correlated, I formally assess this possibility. I compute a group-specific *lower* age limit for child marriage risk in the same way I previously constructed an *upper* age limit for cutting. Specifically, I define among ever-married women the age below which 90% had not yet entered marriage ("marriage floor") and compare it with the age above which 90% of ever-cut women had already undergone cutting ("cutting ceiling"). Ambiguous cases, where early marriage might plausibly explain mobility,

Table 3.4: FGC prevalence and child fostering, by regional prevalence

	Dep. var.: Child is fostered							
	Sample							
	Group prev. $\geq$ Regional prev.				Group prev. $<$ Regional prev.			
	All	Att. known	Anti-FGC	Pro-FGC	All	Att. known	Anti-FGC	Pro-FGC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Child girl × Prev.	-0.016**	0.006	0.002	0.002	0.032***	0.041***	0.052***	0.025
	(0.007)	(0.010)	(0.015)	(0.017)	(0.006)	(0.007)	(0.014)	(0.015)
FGC prevalence	-0.108***	-0.188***	-0.266***	-0.138***	-0.156***	-0.073*	-0.329***	0.012
	(0.025)	(0.029)	(0.045)	(0.040)	(0.028)	(0.039)	(0.059)	(0.055)
Child is a girl	0.041***	0.023***	0.024**	0.028*	0.013***	0.006*	0.005	0.017
	(0.005)	(0.008)	(0.010)	(0.015)	(0.002)	(0.003)	(0.004)	(0.011)
R^2	0.597	0.599	0.608	0.596	0.586	0.607	0.600	0.615
Mother FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep. var.	0.12	0.13	0.12	0.13	0.15	0.18	0.17	0.22
Number of countries	13	13	13	13	13	13	13	13
Number of ethnic gr.	71	61	58	58	85	78	75	71
Observations	118,986	81,956	27,245	54,711	184,330	88,335	58,021	30,314

Notes: The unit of observation is a child younger than the age limit for cutting and under 18. The left (right) panel presents results from the subsample of mothers living in regions where the prevalence is lower (higher) than their group prevalence. The dependent variable is a binary indicator equal to one if the child does not coreside with his or her mother. All specifications include the following control variables: child's age fixed effects and mother fixed effects. Standard errors clustered at the mother level are in parentheses. $^{*}\, p < 0.1$, $^{**}\, p < 0.05$, $^{***}\, p < 0.01$.

occur when the marriage floor is below the cutting ceiling; this applies to slightly fewer than one-quarter of girls, and in half of these cases the difference is only one or two years (see Figure D.3 in the Appendix). Many children from these groups fall outside this ambiguous range. To rule out child marriage as a driver, I exclude children older than the cutting age limit as well as those old enough to be married. The modest reduction in sample size between Tables 3.3 and D.4 confirms that relatively few girls are at risk of both cutting and child marriage, and results remain very similar.

Second, the results could reflect differences in gender norms: mothers opposed to cutting in high-prevalence groups may simply hold more bargaining power, which could independently influence their propensity to foster daughters. The fact that the results are robust to the inclusion of mother fixed effects is reassuring, as these absorb all time-invariant maternal characteristics, including gender norms.[17] To further assess the validity of the results for children excluded from the mother fixed-effects estimation sample (that is, children without a sibling of the other sex), I add a composite index of women's influence in household decision-making to the specification without mother fixed effects presented in Table D.3. This index is intended to proxy bargaining power. It ranges from 0 to 3 and counts affirmative answers to three questions regarding who has the final say over large household purchases, visits to relatives, and the woman's own healthcare.[18] Results including this index as an indicator variable are shown in Table D.5. Because these questions are available only in more recent waves, the sample size decreases slightly. Nonetheless, the results remain qualitatively similar to those in Table 3.3 when restricting to these waves (Columns 1–3), and they continue to hold when the index is added as a control (Columns 4–6).

Third, beyond mother's bargaining power, other gender-norm-related omitted variables could explain both the beliefs about cutting and the propensity to foster girls rather than boys, generating a spurious correlation between being against cutting and fostering more girls. Although I cannot fully rule out this concern, I assess the likelihood that such bias is driving the results by conducting a placebo test. I replace beliefs about cutting with an attitude expected to be correlated with the same omitted variables but that should not

[17]Gender norms may evolve over a mother's life course. However, since each mother is surveyed only once, mother-level characteristics are *de facto* fixed at the time of observation and are absorbed by mother fixed effects.

[18]The wording of the questions varies slightly across survey waves, but the response options remain consistent.

directly affect girls' fostering decisions: beliefs about intimate partner violence (IPV). The DHS elicits attitudes toward IPV through the following questions: "In your opinion, is a husband justified in hitting or beating his wife/partner in the following situations: (i) if she goes out without telling him; (ii) if she neglects the children; (iii) if she argues with him; (iv) if she refuses to have sex with him; (v) if she burns the food?" I construct a dummy variable equal to one if the mother reports that IPV is justified in at least one of these situations and zero otherwise. As expected, this measure correlates quite well with beliefs about cutting ($\rho = 0.22$). Results are presented in Table D.6. I first restrict the sample to mothers for whom both beliefs about IPV and FGC are observed (Column 1). I then split the sample, first between mothers opposed to and supportive of FGC, and second between mothers who consider IPV as never justified and those who consider it sometimes justified. Regarding FGC, the pattern mirrors the results previously presented in Table D.3: only mothers opposed to cutting foster girls more than boys (Columns 2 and 3). In contrast, with respect to IPV, mothers tend to foster girls more than boys as group prevalence increases, regardless of their beliefs about IPV (Columns 4 and 5). This echoes the results in Column 1 for the full sample and suggests that splitting the sample according to IPV beliefs does not reveal meaningful differences in fostering patterns. I then further restrict the sample to assess whether fostering more girls is more closely related to beliefs about FGC or to beliefs about IPV. The results remain large in magnitude and statistically significant when mothers are opposed to cutting, irrespective of their beliefs about IPV (Columns 6 and 7). By contrast, the estimates are not statistically significant (Column 8) or only marginally significant at the 10% level (Column 10) when mothers support cutting, again regardless of their beliefs about IPV. This set of results strongly suggests that beliefs about FGC drive the excess fostering of girls, whereas beliefs about IPV do not. Showing that an attitudinal measure plausibly shaped by similar omitted variables as beliefs about FGC fails to explain the results is reassuring regarding the relevance of beliefs about FGC as a driver of fostering behaviour.

Overall, neither early marriage nor unobserved gender norm–related factors appear to account for the higher propensity of mothers from high-prevalence groups to foster their daughters.

3.4.3 Exit options

The evidence presented so far is consistent with a protection narrative: only mothers who oppose cutting and reside in higher-prevalence areas are more likely to displace girls than boys. However, a central piece of the puzzle remains missing: what are the "exit options" available to girls at risk? In other words, how might a mother living in a high-prevalence area foster her daughter in a household located in a less-affected area? As it is generally acknowledged that most child fostering occurs within the extended family or among relatives (Beck et al., 2015; Alber, 2004), this may initially appear unlikely. However, such a strategy may be feasible in two distinct settings.

First, such protection may occur when ethnic groups are characterized by substantial geographic variation in cutting prevalence. In such contexts, parents may reduce their daughters' exposure to cutting risk through relocation alone, even while remaining within the same ethnic network. This situation is more likely for groups that are geographically dispersed across regions and characterized by heterogeneous local prevalence rates. In these cases, fostering a daughter to another location within the same ethnic group may provide a viable "exit option" by moving her into an environment where the social pressure to cut is lower.

To explore the role of this geographic dimension, I construct a measure of within-country geographic variation in cutting prevalence for each ethnic group by computing the standard deviation of ethnicity–region prevalence rates within each ethnic group *and country.*[19] This measure ranges from 0 for groups present in only one region[20] to 0.39, with a median value of 0.05. I then split the sample between ethnic groups with above- and below-median geographic variation. Table D.7 presents the results. The pattern observed in Table 3.3 (i.e., only mothers opposed to cutting foster girls more than boys) appears only among respondents belonging to high-variation ethnic groups (Columns 5 to 8), and not among mothers belonging to low-variation groups (Columns 1 to 4). These results suggest that within-ethnicity interregional fostering may constitute a viable protection strategy against cutting.

Second, when both parents belong to the same ethnic group (or to two groups with simi-

[19]This means that I do not consider child fostering across borders. Although possible, such cases remain rare, and allowing for between-country variation would bias the measure toward very large groups spread across multiple countries (such as the *Peulh* or *Arabic* groups), even though ties within these groups may weaken with geographic dispersion.

[20]Such groups represent less than 7% of children.

larly high prevalence) they are less likely to have access to family networks located in lower-prevalence contexts. By contrast, inter-ethnic couples, in which each parent is associated with a different level of cutting risk, may reside in a high-prevalence environment while still retaining exit options in lower-prevalence ones through their family ties. If this mechanism is at play, such households should be driving the observed effect. I test this hypothesis and present the results in Table D.8. The first column replicates the baseline specification reported in the first column of Table 3.3, using the full sample of children. The second column restricts the sample to children whose mother's husband is likely to be their biological father and for whom the father's ethnicity is observed. This restriction substantially reduces the sample size for two reasons. First, while women are interviewed in all surveyed households, only a subset of households includes the men's individual questionnaire. Second, because the DHS does not explicitly indicate whether the interviewed man is the (step-)father of the child, I assume this to be the case only when the number of years the man has resided in the household is greater than or equal to the child's age. Children whose mother's husband has lived in the household for a shorter period are therefore excluded when moving from column 1 to column 2. Despite the reduction in the number of observations, the results remain quantitatively similar: higher group-level prevalence is associated with a greater probability of fostering girls rather than boys. I then examine whether inter-ethnic couples foster girls more frequently than intra-ethnic couples. The results indicate that the effect is driven by couples of mixed ethnicity (columns 3 and 4), who account for roughly one quarter of the total sample. More interestingly, the effect is stronger when the *gap* in prevalence between the two parents' ethnic groups is above the median (column 6), and weaker when the gap is below the median (column 5),[21] even if this result must be taken with a grain of salt because of the smaller size of the sample.

Inter-ethnic couples and geographically dispersed ethnic groups may thus provide access to lower-prevalence networks, making fostering a more feasible exit strategy. A possible concern, however, is that these same contexts may also be those in which the cutting norm is less strictly enforced, which would reduce the need for protective fostering. Conversely, in groups where sanctions are strongest, exit may be most needed but least feasible. If this is the case, the lower need for protection in these contexts would work against finding stronger

[21] Cases in which the gap equals zero are excluded. These mostly correspond to intra-ethnic couples, but may also include couples from different ethnic groups with identical prevalence rates (likely both equal to zero).

fostering responses where exit options are more available: mothers would have less reason to rely on fostering precisely in contexts where fostering is more feasible. The fact that the fostering response is nevertheless stronger in these contexts suggests that the availability of exit options dominates this possible differential in sanction intensity. Under this interpretation, the estimated heterogeneity can be viewed as a lower bound on the role of exit-option availability.

Taken together, these findings suggest that the availability of exit options is central to daughters' mobility: only households with access to alternative geographic or ethnic networks are more likely to foster girls rather than boys.

3.4.4 Daughters' fostering and cutting status

This section examines whether fostered girls are less likely to be cut than non-fostered girls, directly testing the prediction of *H4*. In particular, *H4* predicts that if child mobility is used as a protective strategy, displaced girls should face a lower risk of cutting, especially in higher-prevalence contexts. To test this, I estimate on the population of girls only an equation similar to Eq. (3.1), but taking the cutting status of the daughter as the outcome and the interaction between being fostered and the prevalence of cutting as the variable of interest:

$$Y_i = \alpha F_i \times P_{g(i)} + \beta F_i + \gamma P_{g(i)} + \eta X_{m(i)} + \Phi_{a(i)} + \epsilon_i \tag{3.2}$$

where Y_i is equal to one if daughter i is cut,[22] F_i is a dummy equal to one if the daughter is fostered, and $P_{g(i)}$ captures the group-level (i.e., ethnicity-cohort-region) cutting prevalence to which the daughter is exposed (or was, if she is fostered). The interaction between these two variables is the main variable of interest. It captures the differential likelihood of being cut for fostered versus non-fostered girls, depending on the cutting prevalence within the group. The vector $X_{m(i)}$ is the same as the one used in the specification without mother fixed effects in Section 3.4.1 (mother's age fixed effects, her level of education, the number of births she has had, and a dummy for whether she resides in a rural or urban area). In this equation, I do *not* control for maternal attitudes toward FGC nor stratify the sample along this dimension, because attitudes are likely to be endogenous and potentially affected by the daughter's status. This is due to the fact that attitudes are likely measured after the cutting decision and are thus subject to *ex post* rationalization to avoid cognitive dissonance. I replace this control with the mother's cutting status, which is still informative but has occurred before the daughter's

[22]The cutting status of each daughter is the one as reported by the mother, not the daughter herself.

cutting decision. Finally, the term $\Phi_{a(i)}$ denotes the set of age-related fixed effects included in each specification and will be described throughout the section.

The results are presented in Table 3.5. I begin with a simple regression of a dummy for being cut on a dummy for being fostered, controlling only for group-level cutting prevalence. As expected, the coefficient on prevalence is strongly positive, confirming the validity of the group-level risk measure. In contrast, the coefficient for being fostered is different from zero only at the 10% level and small in magnitude (Column 1), suggesting little systematic difference in cutting risk between fostered and non-fostered girls on average. This likely reflects two compositional factors: fostered girls are, on average, older than non-fostered girls, and they often come from higher-risk contexts—both of which mechanically increase their probability of having been cut. To account for these differences, I add age fixed effects and baseline controls: mother's education, mother's cutting status, year of enumeration, and a dummy for urban residence (Column 2). In this specification, the coefficient on being fostered becomes larger in magnitude and turns significant at the 1% level, indicating a 2.5 percentage-point reduction in the probability of being cut. Columns 3 and 4 include more demanding fixed effects to account for systematic differences in the timing of cutting and fostering across areas and groups. Column 3 adds age-by-cluster fixed effects, while Column 4 further adds age-by-ethnicity fixed effects. In these specifications, the coefficient remains significant but smaller, suggesting that fostering has a modest average effect across the full sample. I do not include mother fixed effects in this equation, as it would restrict the estimation sample to mothers who have both cut and uncut daughters, which is far from being the norm.

To test the key prediction of *H4* regarding higher-prevalence contexts, Columns 5–8 introduce the interaction between being fostered and cutting prevalence. The results show that the protective effect of fostering increases with the prevalence of cutting: once age-by-cluster or age-by-ethnicity fixed effects are included, the main effect of fostering becomes insignificant, while the interaction term becomes large and significant. This indicates that fostering reduces the probability of being cut primarily in contexts where the risk is highest, in line with the conceptual framework linking displacement to protective motives. These results are similar when excluding girls at risk of child marriage (Table D.9 in the Appendix), and robust to several tests presented in Section 3.5.

Table 3.5: Daughter's fostering and cutting status

	Dep. var.: Girl is cut							
	Sample: Girls							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Girl is fostered	-0.007*	-0.025***	-0.013***	-0.013***	-0.004	-0.018***	-0.000	0.001
	(0.004)	(0.004)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
Fostered × Prev.					-0.005	-0.011	-0.023**	-0.025**
					(0.009)	(0.009)	(0.010)	(0.010)
FGC prevalence	0.668***	0.505***	0.198***	0.105*	0.668***	0.507***	0.201***	0.108*
	(0.004)	(0.007)	(0.028)	(0.058)	(0.004)	(0.007)	(0.028)	(0.058)
R^2	0.291	0.307	0.660	0.665	0.291	0.307	0.660	0.665
Basic controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Child age FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
C. age x Cluster FE	No	No	Yes	Yes	No	No	Yes	Yes
C. age x Ethnic gr. FE	No	No	No	Yes	No	No	No	Yes
Mean of dep. var.	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
Number of countries	12	12	12	12	12	12	12	12
Number of ethnic gr.	103	103	103	103	103	103	103	103
Observations	92,039	92,039	92,039	92,039	92,039	92,039	92,039	92,039

Notes: The unit of observation is a girl younger than the age limit for cutting and under 15. The dependent variable is a binary indicator equal to one if the girl is cut. Basic controls include year of enumeration, mother's level of education, mother's cutting status, and a dummy for urban residence. Additional fixed effects are indicated in the table. Standard errors clustered at the mother level are in parentheses. $^{*}\,p < 0.1$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

While these results are consistent with a protective role of fostering, they must be interpreted cautiously. First, the daughter's cutting status is not systematically asked in the survey, and is thus subject to measurement error. Second, selection into fostering may not be random: mothers who are more educated, wealthier, or more opposed to cutting could be more likely to foster their daughters, potentially biasing the estimates. By including detailed age, cluster, and ethnicity fixed effects, as well as mother-level controls, I aim to mitigate these concerns. Nevertheless, unobserved factors influencing both the decision to foster and the child's risk of cutting may remain.

3.4.5 FGC ban in Senegal

Lastly, *H5* predicts that fostering should respond to changes in cutting risk over time: as prevalence declines, the protective motive for relocating girls should weaken. In this section, I exploit the natural experiment offered by the introduction of an FGC ban to assess how an arguably exogenous decline in cutting risk affects fostering. Although all but three Sub-Saharan African countries have officially enacted such bans, evidence on their effectiveness remains mixed. First, because assessing the effect of a ban on the prevalence of the practice is difficult when the ban can also increase under-reporting (García-Hombrados and Salgado, 2023), and second, because bans have, in certain cases, proven to have little to no effectiveness (Cetorelli et al., 2020). However, García-Hombrados and Salgado (2023) show that, in the case of Senegal, the ban introduced in 1999 was followed by a decrease in the prevalence of the practice, unlikely explained by under-reporting. If fostering is a response to the risk of cutting, it should decline as this same risk declines.

To investigate this, I estimate the following equation:

$$Y_i = \alpha BornAfterLaw_i \times Z_i + \beta BornAfterLaw_i + \gamma Z_i + \delta_{a(i)} + \delta_{m(i)} + \epsilon_i \qquad (3.3)$$

where $BornAfterLaw_i$ is a dummy equal to one if the child was born after the implementation of the law, and Z_i denotes the dimension along which the post-ban change is compared. In Column 1, Z_i is a dummy equal to one if the child is a girl, so that α captures whether girls born after the ban became less likely to be fostered relative to boys. In Columns 2 and 3, Z_i is replaced by the pre-ban prevalence of FGC[23]. Because the six major ethnic groups in Senegal

[23] In this case, the construction of this variable of interest becomes similar to that of the reduced form used in García-Hombrados and Salgado (2023)

can be clearly divided into non-practicing (*Wolof* and *Serer*, with prevalence below 1%) and practicing groups (*Diola*, *Mandingue*, *Poular*, and *Soninké*, with prevalence between 63% and 85%), I replace in Columns 4 and 5 Z_i by a dummy equal to one for practicing groups. These specifications in Columns 2 to 5 are estimated separately for boys and girls, so that α captures whether the post-ban change in fostering varies with pre-ban cutting risk within each sex. $\delta_{a(i)}$ denotes child-age fixed effects,[24] and $\delta_{m(i)}$ mother fixed effects. A recurrent concern when studying a ban is the increase in under-reporting of cutting after the ban. However, if mothers still cut their daughters after the ban while simply not reporting it, I should not observe any post-ban effect on fostering. The potential under-reporting bias thus sets a lower bound on the results rather than challenging their validity.

If fostering is a protective response to cutting risk, then it should decrease in proportion to the risk *for girls only*. Table 3.6 indicates that girls born after the ban are relatively less likely to be fostered compared to boys (Column 1). More specifically, boys born after the ban in a higher-prevalence group are *as likely* to be fostered as boys born before, compared to lower-prevalence groups (Column 2). However, this is not the case for girls: within higher-prevalence groups, girls born after the ban are less likely to be fostered than those born before (Column 3). The results remain similar when considering a binary variable for FGC practice (Columns 4–5).

In contrast to regressions estimating Eq. (3.1), splitting the sample by FGC support is not feasible here: the ban may itself influence actual and/or reported beliefs, and trimming the sample on this variable would therefore bias the estimates. For this reason, I do not distinguish between supporters and opponents in Table 3.6, which may slightly attenuate the estimated effects. Finally, regressions with mother fixed effects are demanding, as they require mothers to have at least one child born before and one born after the ban. In regressions restricted to girls (boys), this further implies that mothers included in the sample must have daughters (sons) born both before and after the ban, which selects the sample. I thus re-estimate the equation replacing mother fixed effects by the same mother-level controls as in Eq. 3.1, as well as cluster and ethnic-group fixed effects. The estimates are presented in Table D.10, and remain in line with *H5*. Broadly, results are consistent with the use of fostering as a protective strategy: as the risk fades away, the fostering of girls decreases as well.

[24] Because the Senegal sample includes multiple survey waves (see Table E.1, which displays the available country-year waves), child-age fixed effects do not absorb the $BornAfterLaw_i$ dummy: children of the same age can be born either before or after the ban depending on the survey year.

Table 3.6: FGC ban and child fostering in Senegal

	Dep. var.: Child is fostered				
	Sample: children of mothers interviewed in Senegal				
	All	Boys	Girls	Boys	Girls
	(1)	(2)	(3)	(4)	(5)
Born after ban × Girl	-0.026***				
	(0.006)				
Born after ban × Prev.		-0.064	-0.143**		
		(0.062)	(0.072)		
Born after ban × Practicing gp.				-0.049	-0.106**
				(0.043)	(0.048)
R^2	0.520	0.636	0.522	0.636	0.522
Mother FE	Yes	Yes	Yes	Yes	Yes
Mean of dep. var.	0.12	0.12	0.12	0.12	0.12
Number of ethnic gr.	6	6	6	6	6
Observations	63,789	22,363	21,573	22,363	21,573

Notes: The unit of observation is a child younger than the age limit for cutting and under 18. The dependent variable is a binary indicator equal to one if the child does not coreside with his or her mother. All specifications include the following control variables: child's age fixed effects and mother fixed effects. The sum of the sample sizes in Columns 2 and 3 being lower than the sample size in Column 1 reflects the more demanding identifying variation: with mother fixed effects, these specifications rely on mothers who have children of the same sex born both before and after the ban. Standard errors clustered at the mother level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

3.5 Robustness

This section assesses the robustness of the main results, namely Tables 3.3, 3.5 and 3.6. First, I ensure that age heaping does not drive the results. Second, I consider alternative definitions of the risk of cutting faced by daughters. Third, I conduct placebo tests on populations of daughters for whom no fostering response is expected. Finally, I discuss in Appendix E methodological issues revolving around the self-declared ethnicity.

Age heaping in age at experiencing FGC. Age heaping can be defined as the extent to which recorded ages in a dataset show disproportionate preference for certain digits (typically "round" ages ending in 0 or 5), suggesting that the exact age may be unknown. It is common in various contexts (including the self-declared current age), but it has been shown to be particularly acute for age at cutting (Weny et al., 2025). This could pose problems in two ways: first, age at cutting is central for building the age limit at cutting in this analysis; second, the extent of age heaping could be endogenous, for instance due to attachment to tradition. It is therefore essential to ensure that age heaping does not drive the results. I first explore the extent of age heaping in mothers' reports of their own age at cutting in Figure D.4, by computing the distribution of reported ages at cutting. We observe a typical age-heaping pattern, where FGC at age 10 is more than twice as likely as at neighbouring ages. Similar, though less pronounced, spikes appear around ages 5 and 15. Figure D.5 shows the distribution of age limits at cutting under different thresholds, revealing spikes at 5, 10, and 15. These spikes appear symmetrical, indicating that misreporting may occur both below and above the round ages. This is reassuring, as it suggests that the cumulative distribution may naturally smooth out. Figure D.6 empirically explores this by comparing the cumulative distribution of age limits using all declared ages versus excluding round ages (5, 10, and 15). The two distributions remain close, and are nearly identical after age 10, which is reassuring for the results. Even if the impact of age heaping is likely limited, I replicate the main results after recomputing age limits excluding round ages. The results, presented in Tables D.11, D.16 and D.20 are consistent with those obtained using all ages at cutting.

Alternative definition of cutting prevalence. While FGC is a widespread practice in Sub-Saharan Africa, 8.5% of female respondents report that they have never heard of the practice. These women are mostly located in countries with lower overall prevalence[25]

[25]For example, the share of women who have never heard of FGC is 64% in Niger and 44% in Uganda,

Because these women are not administered the FGC module, their cutting status remains unknown and they are excluded from the computation of prevalence rates. This may lead to an overestimation of prevalence, as they are likely to be uncut. I therefore recompute prevalence by assuming that all women who have never heard of FGC are uncut. The results, presented in Table D.12, D.17 and D.21 show very little difference compared to the main estimates.

Mother's cutting status. Since FGC is a tradition primarily perpetuated by mothers and grandmothers (Shell-Duncan et al., 2018; Alradie-Mohamed et al., 2020; Yount, 2002), one could hypothesize that daughters of uncut mothers face a significantly lower risk of being cut, and therefore of being fostered. To test this, I replicate the main results replacing group-level prevalence with the individual cutting status of the mother. Results are presented in Table D.13, D.18 and D.22 and are broadly consistent with those obtained using group-level prevalence.

Threshold for the age limit. The age threshold beyond which "most" girls of a given group are considered to be cut is subjective, and is initially set arbitrarily at the 90% mark. To ensure that the results are not sensitive to this choice, I alternatively use the 85^{th}, 95^{th}, and 99^{th} percentiles of the age distribution to define the threshold. The results, presented in Table D.15, D.19 and D.24 show that the results remain robust to these variations.

Heterogeneity by mother's FGC knowledge. As described above, fewer than 10% of mothers have never heard about cutting. These women were not administered the FGC module: no information is available regarding their own cutting status, that of their daughters (if any), or their beliefs about the practice. However, being unaware of the existence of cutting likely reflects a relatively safe environment, implying a low or negligible risk of the practice.[26] Daughters of such mothers should therefore not be more likely than boys to be fostered in high-prevalence groups compared to low-prevalence ones, making this group suitable for a placebo test. Results are presented in Table D.14. Column 1 estimates Eq. (3.1) on all mothers, replicating the main results, while Column 2 estimates the model on mothers for whom information on FGC knowledge is available. Column 3 restricts the sample to the (much smaller) group of mothers who have never heard of the practice, and Column 4 to those who have. As expected, Columns 2 and 4 show a significant and positive association between higher group prevalence and a greater likelihood of fostering girls. In contrast, the

while it falls below 0.5% in Guinea and Sierra Leone.

[26] Note that it could also reflect questionnaire fatigue, or a reluctance to recall the procedure.

same relationship is negative and not statistically significant among mothers unaware of FGC. This suggests that fostering is concentrated among daughters most at risk, consistent with the hypothesis that fostering serves as a protective response against the practice. Similarly, I reproduce Table 3.6 using only mothers who have never heard about cutting. The results are largely non-significant, which could be explained by the smaller number of observations. More importantly, the coefficient has the opposite sign compared to Table 3.6. This suggests that mothers of daughters who are unlikely to be at risk of cutting do not respond to cutting risk by fostering their daughters. I cannot replicate this analysis for the test of *H4*, as the cutting status of daughters is unknown for mothers who have never heard about cutting.

Taken together, these robustness checks confirm that the main results are not driven by specific assumptions or sample definitions, while also providing further support for the hypothesis that fostering may reflect a protective strategy. For completeness, all possible combinations of robustness checks (e.g., testing the alternative definition of cutting using a 95% threshold, or examining the impact of the mother's cutting status on the fostering of children older than the age limit) are available upon request.

3.6 Conclusion

In this paper, I use a large sample of DHS data to examine whether mothers of daughters at risk of FGC tend to relocate them in order to prevent the practice. I find that, in communities characterized by a higher risk of cutting, girls are fostered more frequently than boys. More precisely, a 10 percentage point increase in the group's prevalence increases the likelihood of fostering a girl rather than a boy by 0.17 percentage points. This effect is entirely driven by mothers opposed to the practice and residing in high-prevalence areas, and is not explained by child marriage, differences in women's bargaining power or beliefs about other gender norms. The mobility of girls is observed mainly among mothers with valid "exit options" for their daughter(s), i.e., belonging to ethnic groups with important geographic variation in cutting prevalence, and being part of an inter-ethnic couple where the prevalence associated with each spouse's group differs substantially. The use of fostering as an exit strategy appears effective: fostered girls from high-prevalence groups are significantly less likely to be cut than their non-fostered counterparts. I further find, in the case of Senegal, that a decline in FGC prevalence due to the ban was followed by a reduction in the proportion of girls fostered in groups where the practice was entrenched.

These results shed light on an overlooked role of child fostering in Sub-Saharan Africa, namely as a means of protection against norms that individuals oppose. More generally, they suggest that mobility may not only respond to harmful social norms, but also shape their persistence or transformation. Collecting more precise data on child fostering, including whether fostered children eventually return and the extent to which they maintain ties with their communities of origin, would make it possible to better understand these dynamics. A complementary approach would be to study the adult trajectories of women who were fostered during childhood, examining whether they are more likely to live outside their community of origin, remain uncut, or marry outside their ethnic group or region. Such evidence would help distinguish between two channels. If daughters of non-complying parents leave permanently, marry into lower-prevalence environments, and weaken their ties with their communities of origin, selective mobility may contribute to the persistence of FGC by maintaining local social environments in which compliance with the norm remains dominant. Conversely, if daughters return and/or maintain strong ties with their families and communities of origin, mobility may instead contribute to broader norm change through direct diffusion (Diabate and Mesplé-Somps, 2019) or cultural remittances (Barsbai et al., 2017).

Bibliography

Ahinkorah, B. O., J. E. Hagan Jr, A.-A. Seidu, O. A. Bolarinwa, E. Budu, C. Adu, J. Okyere, A. G. Archer, and T. Schack (2023): "Association between Female Genital Mutilation and Girl-Child Marriage in Sub-Saharan Africa," *Journal of Biosocial Science*, 55, 87–98.

Akresh, R. (2009): "Flexibility of Household Structure: Child Fostering Decisions in Burkina Faso," *Journal of Human Resources*, 44, 976–997.

Alber, E. (2004): "Grandparents as Foster-Parents: Transformations in Foster Relations between Grandparents and Grandchildren in Northern Benin," *Africa*, 74, 28–46.

Alradie-Mohamed, A., R. Kabir, and S. M. Y. Arafat (2020): "Decision-Making Process in Female Genital Mutilation: A Systematic Review," *International Journal of Environmental Research and Public Health*, 17, 3362.

Arbatli, C. E., Q. H. Ashraf, O. Galor, and M. Klemp (2020): "Diversity and conflict," *Econometrica*, 88, 727–797.

Barsbai, T., H. Rapoport, A. Steinmayr, and C. Trebesch (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.

Beck, S., P. De Vreyer, S. Lambert, K. Marazyan, and A. Safir (2015): "Child Fostering in Senegal," *Journal of Comparative Family Studies*, 46, 57–73.

Bertoli, S., M. Clerc, J. Loper, and È. R. Fernández (2025): "Migration and the epidemiological approach: time and self-selection into foreign ancestries matter," *Journal of Development Economics*, 176, 103505.

Bjalkander, O., S. Bergstrom, L. Almroth, and B. Leigh (2012): "Female genital mutilation in Sierra Leone: who are the decision makers?" *African Journal of Reproductive Health*, 16, 119–131.

Bocquier, P., A. Menashe-Oren, and W. Nie (2023): "Migration's contribution to the urban transition," *Demographic Research*, 48, 681–732.

Boyden, J., A. Pankhurst, and Y. Tafere (2013): *Harmful Traditional Practices and Child Protection: Contested Understandings and Practices of Female Child Marriage and Circumcision in Ethiopia*, Young Lives.

Bursztyn, L., A. L. González, and D. Yanagizawa-Drott (2020): "Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia," *American Economic Review*, 110, 2997–3029.

Cappa, C., C. Thomson, and C. Murray (2020): "Understanding the association between parental attitudes and the practice of female genital mutilation among daughters," *PLoS One*, 15, e0233344.

Cetorelli, V., B. Wilson, E. Batyra, and E. Coast (2020): "Female Genital Mutilation/Cutting in Mali and Mauritania: Understanding Trends and Evaluating Policies," *Studies in Family Planning*, 51, 51–69.

Chesnokova, T. and R. Vaithianathan (2010): "The Economics of Female Genital Cutting," *The BE Journal of Economic Analysis & Policy*, 10.

Corno, L. and D. de Walque (2007): "The Determinants of HIV Infection and Related Sexual Behaviors: Evidence from Lesotho," World Bank Policy Research Working Paper No. 4421.

Cotton, C. (2021): "An Enduring Institution? Child Fostering in Sub-Saharan Africa," *Population and Development Review*, 47, 1179–1206.

——— (2024): "Child Fostering and Family Size Preferences in Uganda," *Studies in Family Planning*, 55, 23–43.

Diabate, I. and S. Mesplé-Somps (2019): "Female Genital Mutilation and Migration in Mali: Do Return Migrants Transfer Social Norms?" *Journal of Population Economics*, 32, 1125–1170.

Engebretsen, S., M. Gueye, A. J. Melnikas, S. Fofana, B. Fané, and S. Amin (2020): "Adolescent girls' migration and its impact on early marriage: Qualitative findings in Mali," *PLoS One*, 15, e0230370.

Fafchamps, M. and A. R. Quisumbing (2007): *Household Formation and Marriage Markets in Rural Areas*, vol. 4, Elsevier.

Fernández, R. (2007): "Women, Work, and Culture," *Journal of the European Economic Association*, 5, 305–332.

Ferreira, P., S. Gulesci, E. La Ferrara, D. Smerdon, and M. Sulaiman (2024): "Changing Harmful Norms through Information and Coordination: Experimental Evidence from Somalia," Preliminary draft.

García-Hombrados, J. and E. Salgado (2023): "Legal Bans, Female Genital Cutting, and Education: Evidence from Senegal," *The World Bank Economic Review*, 37, 74–92.

Giavazzi, F., I. Petkov, and F. Schiantarelli (2019): "Culture: Persistence and evolution," *Journal of Economic Growth*, 24, 117–154.

Gibson, M. A., E. Gurmu, R. Chua, H. Van Bavel, and S. Myers (2023): "Abandoning Female Genital Mutilation/Cutting (FGMC) Is an Emerging but Costly Parental Investment Strategy in Rural Ethiopia," *Social Science & Medicine*, 335, 116170.

Gichohi, M., A. Kotsadam, M. Løvgren, C. Ringdal, and E. Villanger (2025): "Does it Matter What Others Think? Information, Norms, and Female Genital Mutilation in Ethiopia," CMI Working Paper WP 2026:6.

Grant, M. J. and S. Yeatman (2012): "The Relationship between Orphanhood and Child Fostering in Sub-Saharan Africa, 1990s–2000s," *Population Studies*, 66, 279–295.

——— (2014): "The Impact of Family Transitions on Child Fostering in Rural Malawi," *Demography*, 51, 205–228.

Gulesci, S., S. Jindani, E. La Ferrara, D. Smerdon, M. Sulaiman, and H. P. Young (2023): "A Stepping Stone Approach to Norm Transitions," Harvard Kennedy School Working Paper No. RWP23-013.

Gyimah, S. O. (2006): "Migration and fertility behavior in Sub-Saharan Africa: The case of Ghana," *Journal of Comparative Family Studies*, 235–252.

Hayford, S. R. and J. Trinitapoli (2011): "Religious Differences in Female Genital Cutting: A Case Study from Burkina Faso," *Journal for the Scientific Study of Religion*, 50, 252–271.

Hernlund, Y. and B. Shell-Duncan (2007): "Contingency, context, and change: Negotiating female genital cutting in the Gambia and Senegal," *Africa Today*, 43–57.

Hirschman, A. O. (1972): *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*, Harvard University Press.

Kahn, S. (2016): ""You see, one day they cut": The evolution, expression, and consequences of resistance for women who oppose female genital cutting," *Journal of Human Behavior in the Social Environment*, 26, 622–635.

Käkelä, E. (2022): "Strategies of denial: Women's experiences of culture of disbelief and discreditation in the treatment of asylum claims on the grounds of female genital cutting (FGC)," *Journal of Ethnic and Migration Studies*, 48, 560–577.

Khalifa, S. (2022): "Female Genital Cutting and Bride Price," Unpublished working paper.

Klomegah, R. (2000): "Child Fostering and Fertility: Some Evidence from Ghana," *Journal of Comparative Family Studies*, 31, 107–115.

Lindskog, A., H. Congdon Fors, and A.-S. Isaksson (2024): "Harmful Norms: Social Interdependence and Persistence of Female Genital Cutting," SSRN Working Paper No. 4834497.

Lindstrom, D. P. and S. G. Saucedo (2007): "The interrelationship between fertility, family maintenance, and Mexico-US migration," *Demographic Research*, 17, 821–858.

Mackie, G. (1996): "Ending Footbinding and Infibulation: A Convention Account," *American Sociological Review*, 61, 999–1017.

Marazyan, K. (2015): "Foster Children's Educational Outcomes when Child Fostering Is Motivated by Norms: The Case of Cameroon," *Journal of Comparative Family Studies*, 46, 85–103.

McGavock, T. (2018): "Schooling and Work Away from Home: Gender Differences in Child Fostering in West Africa," in *PAA 2018 Annual Meeting*, Population Association of America.

——— (2023): "More than a Village: Why Endogenous Living Arrangements Affect 40 Percent of Non-Orphaned Children in Sub-Saharan Africa," Working paper.

Middelburg, A. and A. Balta (2016): "Female Genital Mutilation/Cutting as a Ground for Asylum in Europe," *International Journal of Refugee Law*, 28, 416–452.

Novak, L. (2020): "Persistent Norms and Tipping Points: The Case of Female Genital Cutting," *Journal of Economic Behavior & Organization*, 177, 433–474.

OFPRA (2023): "Rapport d'activité," Annual activity report.

Omigbodun, O., T. Bella-Awusah, N. Emma-Echiegu, J. Abdulmalik, A. Omigbodun, M.-H. Doucet, and D. Groleau (2022): "Escaping Social Rejection, Gaining Total Capital: The Complex Psychological Experience of Female Genital Mutilation/Cutting (FGM/C) among the Izzi in Southeast Nigeria," *Reproductive Health*, 19, 41.

Pohl, M., E. Batyra, and A. Esteve (2025): "The Stability of Child Fostering in Sub-Saharan Africa: The Case of Senegal," *Studies in Family Planning*, 56, 113–134.

Serra, R. (2009): "Child Fostering in Africa: When Labor and Schooling Motives May Coexist," *Journal of Development Economics*, 88, 157–170.

Shell-Duncan, B., A. Moreau, K. Wander, and S. Smith (2018): "The Role of Older Women in Contesting Norms Associated with Female Genital Mutilation/Cutting in Senegambia: A Factorial Focus Group Analysis," *PLOS ONE*, 13, e0199217.

Shell-Duncan, B., K. Wander, Y. Hernlund, and A. Moreau (2011): "Dynamics of Change in the Practice of Female Genital Cutting in Senegambia: Testing Predictions of Social Convention Theory," *Social Science & Medicine*, 73, 1275–1283.

Sonobe, Y. (2025): "Female genital mutilation/cutting as a ground for asylum in France," *Japanese Journal of Sociology*, 34, 171–185.

UNICEF (2016): *Female Genital Mutilation/Cutting: A Global Concern*, UNICEF.

Vissandjée, B., S. Denetto, P. Migliardi, and J. Proctor (2014): "Female Genital Cutting (FGC) and the Ethics of Care: Community Engagement and Cultural Sensitivity at the Interface of Migration Experiences," *BMC International Health and Human Rights*, 14, 1–10.

Vogt, S., C. Efferson, and E. Fehr (2017): "The risk of female genital cutting in Europe: comparing immigrant attitudes toward uncut girls with attitudes in a practicing country," *SSM-Population Health*, 3, 283–293.

Weny, K., R. Silva, and S. J. Klug (2025): "Self-report and proxy reports in survey data on female genital mutilation, Senegal," *Bulletin of the World Health Organization*, 103, 366.

Yount, K. M. (2002): "Like Mother, Like Daughter? Female Genital Cutting in Minia, Egypt," *Journal of Health and Social Behavior*, 43, 336–358.

3.A Sample construction

A.1 Country-wave selection based on the existence of a FGC module

1. Some countries have no FGC module at all: they do not appear in the database. **[26 SSA countries[27]]**

2. Some countries have an FGC module for at least one wave.[28] **[21 SSA countries, 99 country-year waves, 1,076,674 observations]**

   (a) A given wave is not covered by an FGC module: they appear as "`fcknow` == ." **[36 country-year waves, 278,414 observations]**

   (b) A given wave is covered by an FGC module **[63 country-year waves, 798,260 observations]**. The module can be administered to:

      i. All households: there is no "niu" value as a response of the variable `fcknow`. **[57 country-waves, 742,087 observations]**

      ii. A subsample of households.[29]: Non-selected households appear as "fcknow == niu". **[6 waves, 56,173 observations]**

A.2 Age limit at cutting

1. The question about age at cutting is asked to female respondents only in a subsample of the waves. Age at cutting cannot be computed for three countries for which this question is never asked: Ghana, Liberia and Uganda. Cameroon has only 44 observations of cut women, and is thus discarded by construction, as ethnicity-cohort-region groups smaller than 50 are not kept in the sample. The age at cutting is asked in all (or all but one) waves in nine countries[30], and in at least one wave in the remaining five countries[31].

[27] The content of the FGC module is not consistent across country-year waves. For instance, Ghana (in 2003), and Eswatini (in 2006) technically have a FGC module, but that solely includes a question on the cutting status of the surveyed women, and can thus be used only for part of the analysis.

[28] I keep waves without an FGC module as information about girl's mobility can still be used.

[29] This is the case for Cameroon 2004, Chad 2014, Ethiopia 2016, Kenya 2014, Mali 2018 and Nigeria 2018. See the selection rules below.

[30] CAR, Chad, Sierra Leone, Guinea, Ivory Coast, Mali, Niger, Nigeria and Burkina Faso.

[31] Benin, Ethiopia, Kenya, Senegal and Togo.

2. I require that at least fifty non-missing observations on age at cutting be available within a group—defined by ethnicity, region of residence, and 10-year birth decade—in order to compute a threshold. Groups with fewer than fifty non-missing observations on age at cutting are therefore not assigned a threshold. As a result, 154,008 female respondents are *not* associated with an age limit for not being at risk of cutting any more, and are thus discarded from the sample. [**15 countries, 62 country-waves, 593,450 women**].

A.3 Construction of the sample for the main analysis

Starting from the sample of 1,076,674 women described at point 2 of the first paragraph:

1. I keep only women who can be linked to a precise ethnicity[32] [**18 countries, 77 country-waves, 747,458 women**].

2. Among these women, I keep only mothers [**18 countries, 77 country-waves, 548,387 mothers, 2,268,500 children**].

3. Among the children of these mothers, I keep only those who can be linked to an age limit for cutting [**13 countries, 62 country-waves, 323,151 mothers, 1,055,748 children**].

4. Among these children, I use for the estimation only those who are under the age limit for cutting. [**13 countries, 57 country-waves, 201,076 mothers, 419,999 children**] and who can be linked to a prevalence at the ethnicity, region of residence, and 10-year birth decade level[33] [**13 countries, 54 country-waves, 162,970 mothers, 384,752 children**].[34]

3.B Key variable construction

B.1 Daughter's cutting status

[32] Eswatini, Egypt and Tanzania are discarded as there is no question on ethnicity for any wave. Cameroon 2018, Ethiopia 2019, Côte d'Ivoire 1994, Niger 2012, Uganda 2001 and Uganda 2006 also include no question on ethnicity.

[33] Here again, groups with fewer than fifty non-missing observations have a missing value for the prevalence.

[34] The number of observations of Table 3.3 is lower because mother fixed effects creates singletons, which are automatically dropped. The number of observations of the estimation with a less demanding structure of fixed effects (Table D.3) is however almost equal to this figure.

Until 1999, respondents with at least one living daughter were asked whether their eldest daughter had been cut. These waves account for approximately 12% of the daughters. From 1999 to 2010, DHS surveys collected information on the cutting status of either the eldest daughter or the most recently cut daughter, depending on the wave, as well as on the total number of cut daughters. These waves account for approximately 54% of the daughters. From 2010 to 2019, women were asked, in addition to the total number of cut daughters, about the cutting status of each of their daughters below an age cut-off—usually 15, but sometimes 17 or 19. These waves account for the remaining 34% of the daughters. To construct a harmonized measure, I first restrict the sample to daughters aged 15 or younger. I then apply the following procedure: (i) when the individual cutting status of a daughter is available, I use it directly; (ii) when it is not available, I compare the number of cut daughters with the total number of daughters reported by the mother. In 23% of cases, these numbers coincide, indicating that all daughters are cut, while in 62% of cases, the number of cut daughters is zero. In the remaining cases, the exact cutting status of each daughter is unknown. As it is very rare that some daughters are cut and other daughters are not, given the younger daughters are old enough to have undergone the practice (Hayford and Trinitapoli, 2011), I assume that the n eldest daughters of each mother are cut, where n is equal to the reported number of cut daughters. This approximation is likely to be valid, as if a younger daughter is cut, it is unlikely that an older daughter is not, especially given that the prevalence of FGC tends to decrease over time (Lindskog et al., 2024).

B.2 Child fostering

Two distinct approaches are used when studying child fostering through the DHS. The first approach exploits information from the mother's questionnaire in the household of *origin* (Cotton, 2021; Pohl et al., 2025; Klomegah, 2000). For all living children, mothers are asked whether the child coresides in her household or resides elsewhere.[35] Children who coreside with their mother are classified as "non-fostered," whereas those living elsewhere are classified as "fostered". This measure has several advantages. First, it varies at the child level, allowing for differences in living arrangements among siblings within the same family. Second, and crucially for this analysis, it allows each child to be linked to a precise cutting risk, since I

[35] Although the wording of questions and response categories varies slightly across survey waves, it is always possible to determine whether a child coresides with the mother.

observe the mother's ethnicity[36] and region of residence. However, a major drawback of this approach is that it relies on a strong assumption: that non-coresidence necessarily implies fostering. In practice, non-coresidence may reflect a range of situations unrelated to fostering, such as attendance at boarding school,[37] coresidence with the father only, or marriage outside the parental household.

The second approach relies on the household questionnaire from the household of *destination* of fostered children (Marazyan, 2015; Grant and Yeatman, 2012). For each child residing in the household, the questionnaire records whether the mother and father are alive and whether each parent resides in the household.[38] Following McGavock (2023), I classify a child as fostered if both parents are reported to be alive but neither resides in the household, and if the child is not listed as the household head, nor as the spouse, son-in-law, or daughter-in-law of the household head, in order to reduce the likelihood of capturing child marriage. This definition is consistent with DHS documentation, which indicates that foster children can be identified using household questionnaire information. Relative to the household-of-origin approach, this method rules out several forms of non-coresidence unrelated to fostering, notably temporary absence for schooling, coresidence with the father only, and most cases of child marriage.[39] However, this approach has important limitations for the purposes of this study. Because it relies on the household of destination, it does not allow me to recover information on the child's exposure to cutting risk at origin, nor to observe the fostering status of siblings. For these reasons, my preferred measure of child fostering is the one based on the household of origin. The final estimation sample thus consists of all children under 18 and under the age limit at cutting observed in the mother questionnaire (i.e., in the origin household). Table D.1 presents the distribution across countries of this sample, as described in point 4 of Section A.3.

To validate this measure of fostering, I compare the number of fostered children identified

[36] As well as the father's ethnicity for part of the sample, as detailed in Section 3.4.3.

[37] The DHS enumerator's manual for the women's questionnaire states: "If a child is away at boarding school or lives with other people on a permanent basis, record NO. If the child is away for a short while but usually lives with the mother, record YES." See https://dhsprogram.com/pubs/pdf/DHSM1/DHS6_Interviewer_Manual_French_19Oct2012_DHSM1.pdf, last accessed on May 27, 2026.

[38] The target population varies across surveys—children under 15 in most samples and under 18 in some. For harmonization, I restrict the sample to children under 15.

[39] While this possibility cannot be entirely ruled out—for instance, if the fostered child is married to a household member who is not the child of the household head—it is substantially less likely.

using the mother's questionnaire with the number identified using the household-of-destination approach, which should provide the closest approximation to the true number of fostered children. I construct four groups: non-fostered children identified through the mother's questionnaire; fostered children identified through the mother's questionnaire (origin household); non-fostered children identified through the household questionnaire; and fostered children identified through the household questionnaire (destination household). As discussed above, the two populations of fostered children may differ for several reasons. Moreover, the two populations of *non*-fostered children also differ for two main reasons. First, the mother's questionnaire is administered only to eligible mothers residing in the household, i.e., aged 15 to 49 in most surveys. Children from a mother outside this age range will thus not be identified through the mother's questionnaire, but will appear in the household roster. For comparability, I therefore restrict the sample of children identified through the mother's questionnaire to those under age 15. I also discard children from the mother questionnaire who are old enough to be married (whose non-coresidence may reflect marriage rather than fostering) to further reduce discrepancies between the two measures.[40] I cannot impose a similar restriction on children identified through the household questionnaire, as the ages of parents or caregivers of fostered children in the destination household are not observed.

As a result, the number of children identified through the household questionnaire is mechanically slightly higher: 403,012 children, compared to 381,251 identified through the mother's questionnaire. Among these, 45,985 (12.06%) are classified as fostered using the mother's questionnaire, compared to 44,580 (11.06%) using the household's questionnaire. The higher proportion identified using the mother's questionnaire likely reflects the fact that non-coresidence does not always correspond to fostering (e.g., boarding school, coresidence with the father, or residual cases of child marriage). Nevertheless, the similarity of these proportions is reassuring regarding the ability of the mother's questionnaire to identify actual fostered children.

[40]Note that these restrictions just ensure comparability and are not applied to the estimation sample.

3.C Mother-child migration

Another strategy that naturally comes to mind when thinking about how mothers may protect their daughters from cutting is migration. Mothers opposed to cutting belonging to communities where it is the norm could migrate to lower prevalence areas. Such migration could occur at different stages of the life cycle relative to marriage and childbirth. This timing is important, since migration, fertility, and marriage arrangements are closely interconnected (Gyimah, 2006; Lindstrom and Saucedo, 2007; Bocquier et al., 2023; Engebretsen et al., 2020). Ideally, I would observe each woman's full migration history, as well as her place of birth, which would allow me to identify the timing of migration relative to union formation and childbirth. However, the DHS does not provide such information. I can only use information from the women's questionnaire on the respondent's current place of residence and the number of years she has lived there. I therefore focus on migration occurring *after* the birth of a child (which only captures part of the broader migration process), comparing migration after the birth of a girl to migration after the birth of a boy.

First, I conduct the same descriptive analysis as for child fostering, comparing children who moved with their mother to those who did not, from the household of origin and destination perspective. The picture differs markedly from the one for child fostering. Table C.1 presents similar statistics for children who moved with their mother.[41] There are no systematic differences between boys and girls in observable characteristics, either in the full sample or among firstborn children (Table 3.C). Children who move with their mother, however, tend to reside in more urban, lower-prevalence areas and have mothers who are more educated, less likely to be cut, and less supportive of the practice.

If migration is used as protection against cutting, I should observe migration after the birth of a girl, and not a boy, in communities where cutting is the norm and the mother is opposed. To explore this, I estimate the following equation:

$$Y_i = \alpha G_i \times P_{g(i)} + \beta G_i + \gamma P_{g(i)} + \delta_{a(i)} + \eta X_{m(i)} + \delta_e + \delta_c + \delta_y + \epsilon_i \qquad (3.4)$$

This equation is similar to Eq. (3.1), except that the dependent variable is a dummy for mother-child migration. This variable takes the value one if the mother still coresides with

[41] The number of observations is larger because the sample includes all children under 18, not only those below the age limit for cutting.

her child and moved between the birth of her child and the child's group-specific age limit for cutting. In concrete terms, the dummy takes the value one when the difference between the child's age and the number of years the respondent has lived in her current residence is positive and less than or equal to the child's age limit[42]. Importantly, this equation is estimated only for firstborns, because the interpretation is more straightforward for them: a mother from a group with a strong tradition of excision may be more likely to move after giving birth to a daughter, in an attempt to protect her from the practice. In contrast, it is more difficult to draw clear conclusions in cases where a mother did not move after the birth of her first daughter, but did move after the birth of her second or third. Since only one child per mother is included in the sample, mother fixed effects would absorb all identifying variation. For this reason, mother fixed effects are replaced by ethnicity (δ_e), cluster[43] (δ_c), and year of enumeration fixed effects (δ_y), as well as the vector $X_{m(i)}$ including the following control variables at the mother's level: her level of education, her cutting status, her age, the number of births she has had, and a dummy for whether she resides in a rural or urban area. Results are presented in Table C.5. The coefficient on the interaction term remains largely insignificant regardless of the mother's beliefs, indicating no differential migratory response to the birth of a boy or girl between high- and low-prevalence groups. These results do not support the hypothesis that mothers respond to cutting risk through mother-child migration.

[Descriptive statistics – moved vs. not moved, firstborn]Descriptive Statistics – Children

[42] For groups in which girls are cut at birth, the age limit is zero; in these cases, the dummy takes the value one only if the mother moves during the first year after the child's birth.

[43] According to the DHS Sampling Manual, "A cluster is the smallest geographical survey statistical unit for DHS surveys. It consists of a number of adjacent households in a geographical area. For DHS surveys, a cluster corresponds either to an enumeration area, or a segment of a large enumeration area."

Table C.1: Descriptive Statistics – Children who moved vs. not-moved

	Moved			Not moved		
Variable	Girls	Boys	Diff.	Girls	Boys	Diff.
Age of the child	9.830	9.810	0.020	8.773	8.801	-0.028***
	(0.022)	(0.022)	(0.032)	(0.008)	(0.007)	(0.011)
Urban status	0.390	0.392	-0.002	0.276	0.276	0.000
	(0.003)	(0.003)	(0.004)	(0.001)	(0.001)	(0.001)
Education	0.624	0.614	0.010	0.379	0.374	0.005***
	(0.005)	(0.005)	(0.007)	(0.001)	(0.001)	(0.002)
Mother's age	33.578	33.568	0.009	34.698	34.668	0.030*
	(0.035)	(0.035)	(0.049)	(0.011)	(0.011)	(0.016)
Heard about FGC	0.903	0.907	-0.005	0.871	0.870	0.001
	(0.002)	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)
FGC should continue	0.407	0.409	-0.002	0.485	0.486	-0.001
	(0.004)	(0.004)	(0.005)	(0.001)	(0.001)	(0.001)
Woman cut	0.626	0.628	-0.003	0.674	0.674	-0.001
	(0.003)	(0.003)	(0.005)	(0.001)	(0.001)	(0.001)
Regional prevalence of FGC	0.421	0.425	-0.004	0.482	0.482	-0.001
	(0.002)	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)
Ethnic-region prev. of FGC	0.401	0.405	-0.003	0.473	0.474	-0.001
	(0.002)	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)
Number of observations	32,847	32,792	65,639	358,382	368,474	726,856

Notes: The unit of observation is a child (boy or girl) under 18 who can be linked to a group-level prevalence and age at cutting. Children are identified through the mother's survey, which means that mothers with multiple children appear multiple times in the dataset. Standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

who moved vs. not-moved, firstborn only

	Moved			
Variable	Girls	Boys	Diff.	C
Age of the child	9.836	9.787	0.049	9
	(0.039)	(0.039)	(0.055)	(0
Urban status	0.392	0.392	-0.000	0
	(0.005)	(0.005)	(0.007)	(0
Education	0.743	0.728	0.015	0
	(0.008)	(0.008)	(0.012)	(0
Mother's age	29.398	29.351	0.047	29
	(0.049)	(0.049)	(0.069)	(0
Heard about FGC	0.901	0.903	-0.002	0
	(0.004)	(0.004)	(0.005)	(0
FGC should continue	0.396	0.400	-0.005	0
	(0.006)	(0.006)	(0.009)	(0
Woman cut	0.603	0.607	-0.004	0
	(0.006)	(0.006)	(0.008)	(0
Regional prevalence of FGC	0.415	0.422	-0.006	0
	(0.003)	(0.003)	(0.005)	(0
Ethnic-region prev. of FGC	0.395	0.399	-0.004	0
	(0.004)	(0.004)	(0.005)	(0
Number of observations	11,294	11,172	22,466	72

Table C.5: FGC prevalence and mobility of the firstborn

	Dep. var.: Mother moved			
	Sample			
	All	Att. known	Anti-FGC	Pro-FGC
	(1)	(2)	(3)	(4)
Fb. girl × Prevalence	-0.004	-0.005	-0.017	0.031
	(0.007)	(0.009)	(0.015)	(0.027)
FGC prevalence	-0.142***	-0.157***	-0.228**	-0.033
	(0.045)	(0.056)	(0.089)	(0.077)
Firstborn is a girl	0.000	0.001	0.005	-0.027
	(0.005)	(0.006)	(0.006)	(0.022)
R^2	0.225	0.233	0.256	0.217
Year FE	Yes	Yes	Yes	Yes
Cluster FE	Yes	Yes	Yes	Yes
Ethnic group FE	Yes	Yes	Yes	Yes
Mean of dep. var.	0.21	0.21	0.19	0.24
Number of countries	10	10	10	10
Number of ethnic gr.	87	87	86	75
Observations	57,904	33,071	18,489	14,576

Notes: The unit of observation is a firstborn child younger than the age limit for cutting and under 18. The dependent variable equals one if the mother moved between the birth of her firstborn and the year the child reached the age limit for cutting. All specifications include the following control variables: the mother's age fixed effects, age of the firstborn child fixed effects, the number of births she has had, her level of education, a dummy for whether she resides in a rural or urban area. All specifications include ethnic-group fixed effects δ_e, region fixed effects δ_r, and year fixed effects δ_y. Standard errors clustered at the mother level are in parentheses. $^{*}\,p < 0.1$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

3.D Additional tables and figures

D.1 Descriptive statistics

Table D.1: Countries of the estimation sample

Country	Frequency	Percent
Benin	21,438	7.00
Burkina Faso	8,955	2.92
Chad	4,018	1.31
Côte d'Ivoire	5,976	1.95
Ethiopia	27,646	9.02
Guinea	21,162	6.91
Kenya	51,489	16.81
Mali	27,142	8.86
Niger	15,317	5.00
Nigeria	17,359	5.67
Senegal	62,709	20.47
Sierra Leone	31,229	10.19
Togo	11,944	3.90
Total	306,384	100

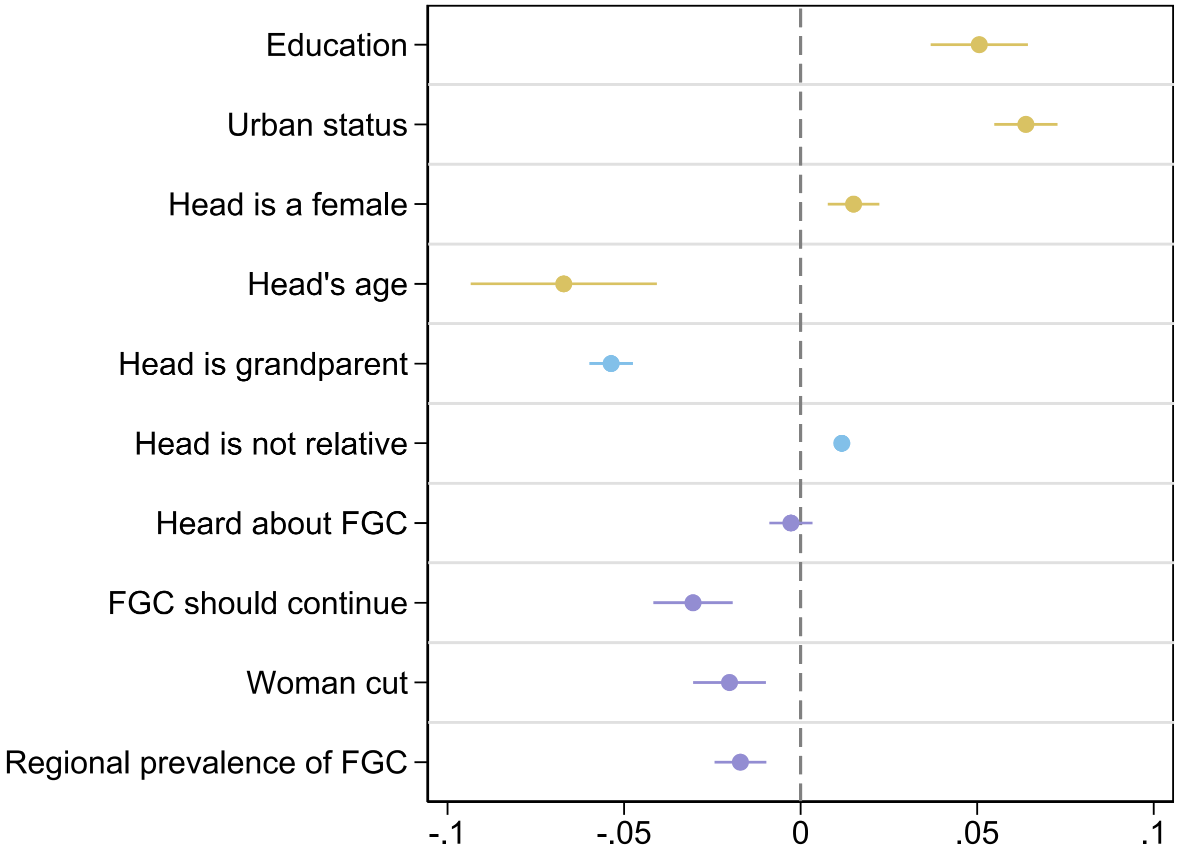
Notes: The unit of observation is a child younger than the age limit for cutting and under 18.

Table D.2: Group-specific age limit at cutting

Age limit (90% threshold)	Frequency	Percent	Cumulative percentage
0	194	0.06	0.06
1	179	0.06	0.12
2	572	0.19	0.31
3	761	0.25	0.56
4	753	0.25	0.80
5	7,068	2.31	3.11
6	16,081	5.25	8.36
7	14,269	4.66	13.02
8	10,208	3.33	16.35
9	10,515	3.43	19.78
10	35,227	11.50	31.28
11	5,666	1.85	33.13
12	46,782	15.27	48.40
13	48,320	15.77	64.17
14	15,425	5.03	69.20
15	12,118	3.96	73.16
16	32,228	10.52	83.67
17	26,904	8.78	92.46
18	13,134	4.29	96.74
19	5,900	1.93	98.67
20	339	0.11	98.78
22	52	0.02	98.80
25	3,689	1.20	100.00
Total	306,384	100	

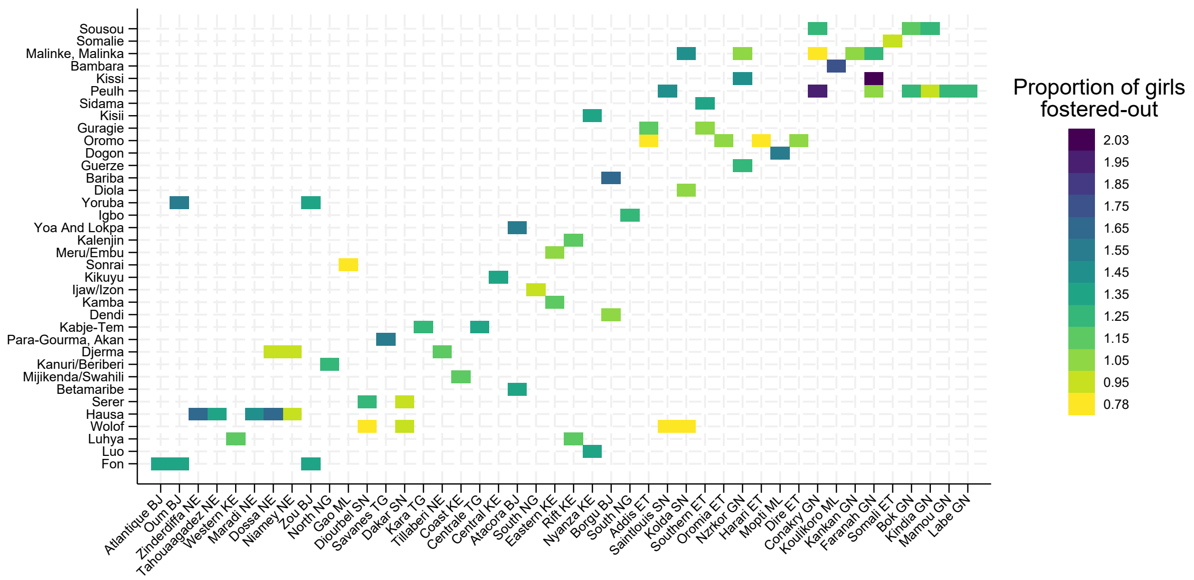
Notes: The unit of observation is a child younger than the age limit for cutting and under 18.

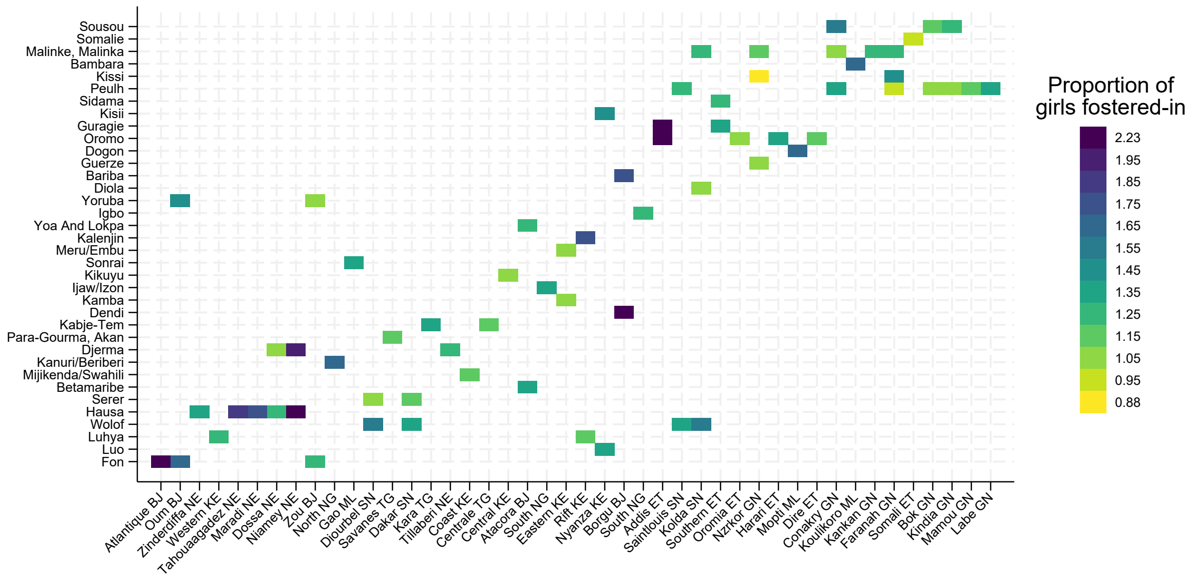
Figure D.1: Characteristics of the destination household of fostered children



Notes: The figure plots the coefficient α from the following equation: $Y_i = \alpha G_i \times F_i + \beta G_i + \gamma F_i + \delta_{a(i)} + \epsilon_i$. The unit of observation is a child under 15 identified through the household survey. Education, FGC knowledge, FGC beliefs and cutting status are calculated as the mean of all women in the household. Head's age has been divided by 10 to fit in the graph, which needs to be taken into account when considering the marginal effects. Demographic outcomes are presented in yellow, relationship to the household head in blue, and FGC-related in purple. 95% confidence intervals are displayed.

Figure D.2: Proportion of girls fostered-out and fostered-in, by ethnic group and region





Data sources: Author's elaboration on the DHS.

Notes: In the first (second) graph, each square represents the proportion of girls under 15 fostered-out (fostered-in) compared to boys under 15 fostered-out (fostered-in) for a given ethnic group in a given region. Only ethnicity–region pairs with at least 100 children fostered-in are kept in the sample.

D.2 Regressions

Table D.3: FGC prevalence and child fostering, less demanding FEs

	Dep. var.: Child is fostered							
	All Mothers	Att. known	Anti-FGC	Pro-FGC	All Mothers	Att. known	Anti-FGC	Pro-FGC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Child girl × Prev.	0.015***	0.029***	0.035***	0.006	0.016***	0.028***	0.038***	0.007
	(0.003)	(0.004)	(0.006)	(0.009)	(0.003)	(0.004)	(0.006)	(0.009)
FGC prevalence	-0.010	-0.019**	-0.039***	0.020	-0.076***	-0.102***	-0.189***	-0.015
	(0.006)	(0.009)	(0.011)	(0.017)	(0.013)	(0.018)	(0.024)	(0.031)
Child is a girl	0.016***	0.005**	0.002	0.025***	0.016***	0.007***	0.004	0.024***
	(0.002)	(0.003)	(0.003)	(0.008)	(0.002)	(0.003)	(0.003)	(0.008)
R^2	0.113	0.123	0.092	0.157	0.193	0.212	0.225	0.255
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year & Ethnic gr. FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster FE	No	No	No	No	Yes	Yes	Yes	Yes
Mean of dep. var.	0.14	0.15	0.15	0.16	0.14	0.15	0.15	0.16
Number of countries	13	13	13	13	13	13	13	13
Number of ethnic gr.	116	108	107	102	116	108	105	101
Observations	384,734	222,215	112,569	109,641	383,976	221,278	111,067	108,760

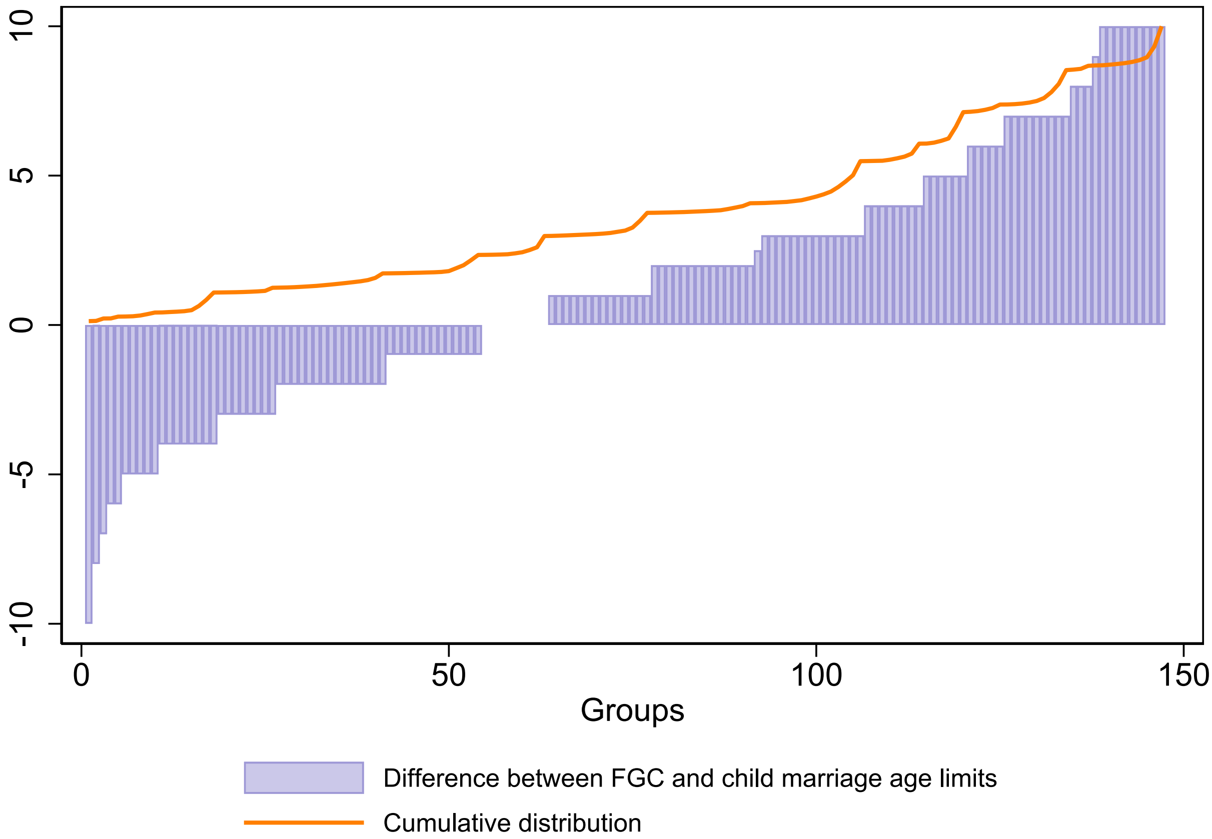
Notes: The unit of observation is a child younger than the age limit for cutting and under 18. The dependent variable is a binary indicator equal to one if the child does not coreside with his or her mother. All specifications include year fixed effects, ethnic-group fixed effects, and cluster fixed effects. Individual controls include the following control variables: child's age fixed effect, mother's age fixed effects, her level of education, the number of births she has had, and a dummy for whether she resides in a rural or urban area. Mother fixed effects are *not* introduced to the regression, meaning that single child are also included to sample. The difference between Columns 1-4 and 5-8 is the introduction of cluster fixed effects. Clusters are defined for all observations, the discrepancy in the number of observations between Columns 1-4 and 5-8 rather stems from singletons that are automatically dropped from the sample before the estimation. Standard errors clustered at the mother level are in parentheses. $^{*}\, p < 0.1$, $^{**}\, p < 0.05$, $^{***}\, p < 0.01$.

Table D.4: Alternative explanation: child marriage

	Dep. var.: Child is fostered			
	Sample			
	All Mothers	Att. known	Anti-FGC	Pro-FGC
	(1)	(2)	(3)	(4)
Child girl × Prev.	0.018***	0.024***	0.026***	0.013
	(0.003)	(0.005)	(0.007)	(0.010)
FGC prevalence	-0.097***	-0.149***	-0.274***	-0.071**
	(0.019)	(0.025)	(0.036)	(0.036)
Child is a girl	0.014***	0.010***	0.008**	0.019**
	(0.002)	(0.003)	(0.003)	(0.009)
R^2	0.587	0.605	0.607	0.603
Mother FE	Yes	Yes	Yes	Yes
Mean of dep. var.	0.13	0.15	0.14	0.15
Number of countries	13	13	13	13
Number of ethnic gr.	111	100	97	96
Observations	286,003	161,779	80,008	81,771

Notes: The unit of observation is a child younger than the age limit for cutting, the age limit for marriage, and under 18. The dependent variable is a binary indicator equal to one if the child does not coreside with his or her mother. All specifications include the following control variables: child's age and mother fixed effects. Standard errors clustered at the mother level are in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure D.3: Age limits for cutting and child marriage



Data sources: Author's elaboration on the DHS.

Notes: This figure plots the difference between the (upper) age limit of cutting and the (lower) age limit for child marriage. A negative difference thus refers to groups where girls can be at risk of child marriage and cutting at the same time.

Table D.5: Alternative explanation: gender roles within the household

	Dep. var.: Child is fostered					
	Sample: Gender role index is defined					
	All	Anti-FGC	Pro-FGC	All	Anti-FGC	Pro-FGC
	(1)	(2)	(3)	(4)	(5)	(6)
Child girl × Prev.	0.034***	0.045***	0.010	0.031***	0.049***	0.017
	(0.004)	(0.007)	(0.012)	(0.004)	(0.007)	(0.012)
FGC prevalence	-0.105***	-0.195***	-0.018	-0.128***	-0.175***	-0.008
	(0.022)	(0.028)	(0.040)	(0.009)	(0.028)	(0.040)
Child is a girl	0.006**	0.003	0.025**	0.008***	0.003	0.021**
	(0.003)	(0.003)	(0.011)	(0.003)	(0.003)	(0.011)
WBP Index = 1				-0.001	0.000	0.000
				(0.003)	(0.004)	(0.004)
WBP Index = 2				-0.002	-0.000	-0.004
				(0.003)	(0.005)	(0.005)
WBP Index = 3				-0.002	-0.004	-0.004
				(0.003)	(0.004)	(0.005)
R^2	0.213	0.220	0.256	0.218	0.224	0.260
Year, eth. & cl. FEs	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep. var.	0.16	0.15	0.17	0.16	0.15	0.17
Number of countries	12	12	12	12	12	12
Number of ethnic gr.	89	87	80	89	87	80
Observations	162,459	82,927	78,450	162,459	82,927	78,450

Notes: The unit of observation is a child younger than the age limit for cutting and under 18. The dependent variable is a binary indicator equal to one if the child does not coreside with his or her mother. All specifications include year fixed effects, ethnic-group fixed effects and cluster fixed effects, as well as the following control variables: child's age fixed effects, mother's age fixed effects, her level of education, the number of births she has had, and a dummy for whether she resides in a rural or urban area. Mother fixed effects are *not* introduced to the regression, meaning that single child are also included to sample. Standard errors clustered at the mother level are in parentheses. $^{*}\,p < 0.1$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

Table D.6: Alternative explanation: Beliefs regarding IPV as a placebo

	Dep. var.: Child is fostered								
	Sample								
	Beliefs about IPV & FGC defined	Anti-FGC	Pro-FGC	Anti-IPV	IPV Justified	Anti-FGC & Anti-IPV	Anti-FGC & IPV Justified	Pro-FGC & Anti-IPV	Pro-FGC & IPV Justified
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Child girl × Prev.	0.031***	0.039***	0.012	0.044***	0.027***	0.060***	0.029***	0.007	0.025*
	(0.004)	(0.007)	(0.011)	(0.009)	(0.005)	(0.014)	(0.008)	(0.025)	(0.014)
FGC prevalence	-0.104***	-0.166***	-0.047	-0.027	-0.157***	-0.098*	-0.212***	0.140*	-0.123***
	(0.020)	(0.027)	(0.033)	(0.041)	(0.023)	(0.051)	(0.033)	(0.077)	(0.036)
Child is a girl	0.005*	0.003	0.021**	-0.002	0.007*	-0.008	0.007*	0.029	0.009
	(0.003)	(0.003)	(0.010)	(0.005)	(0.004)	(0.005)	(0.004)	(0.020)	(0.012)
R^2	0.214	0.223	0.255	0.301	0.225	0.316	0.244	0.351	0.259
Mother FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep. var.	0.16	0.15	0.17	0.18	0.15	0.16	0.15	0.22	0.15
Number of countries	11	11	11	11	11	11	11	10	11
Number of ethnic gr.	91	87	85	88	90	83	86	73	83
Observations	178,689	88,671	88,836	51,387	125,653	32,059	55,112	18,511	69,445

Notes: The unit of observation is a child younger than the age limit for cutting and under 18. The dependent variable is a binary indicator equal to one if the child does not coreside with his or her mother. All specifications include the following control variables: child's age fixed effects and mother fixed effects. Standard errors clustered at the mother level are in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table D.7: Exit option: geographic variability at the ethnic level

	Dep. var.: Child is fostered							
	Sample							
	Top 50% lower variability				Top 50% higher variability			
	All	Att. known	Anti-FGC	Pro-FGC	Att. known	All	Anti-FGC	Pro-FGC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Child girl × Prev.	0.014***	0.018**	0.014	0.018	0.019***	0.032***	0.039***	0.010
	(0.005)	(0.008)	(0.011)	(0.017)	(0.004)	(0.006)	(0.010)	(0.016)
FGC prevalence	-0.109***	-0.184***	-0.358***	-0.110***	-0.197***	-0.056	-0.149**	0.017
	(0.022)	(0.029)	(0.041)	(0.041)	(0.034)	(0.052)	(0.070)	(0.080)
Child is a girl	0.020***	0.018***	0.018***	0.019	0.014***	0.002	0.000	0.022
	(0.003)	(0.005)	(0.005)	(0.012)	(0.003)	(0.004)	(0.004)	(0.014)
R^2	0.596	0.623	0.612	0.628	0.578	0.591	0.593	0.588
Mother FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep. var.	0.13	0.15	0.13	0.19	0.15	0.16	0.17	0.14
Number of countries	12	11	11	11	13	12	12	12
Number of ethnic gr.	55	52	50	50	60	53	52	49
Observations	147,325	74,832	40,232	34,600	159,059	96,959	45,466	51,493

Notes: The unit of observation is a child younger than the age limit for cutting and under 18. The dependent variable is a binary indicator equal to one if the child does not coreside with his or her mother. All specifications include the following control variables: child's age fixed effects and mother fixed effects. Standard errors clustered at the mother level are in parentheses. $^{*}\,p < 0.1$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

Table D.8: FGC prevalence and child fostering: role of the couple's ethnic composition

	Dep. var.: Child is fostered					
	Sample					
	All children	Father's ethn. known	Inter-ethnic couple	Intra-ethnic couple	Small gap	Large gap
	(1)	(2)	(3)	(4)	(5)	(6)
Child girl × Prev.	0.017***	0.016**	0.041**	0.011	0.034*	0.063**
	(0.003)	(0.008)	(0.018)	(0.009)	(0.023)	(0.027)
FGC prevalence	-0.134***	-0.042	0.201	-0.288***	2.307	0.509
	(0.018)	(0.080)	(0.127)	(0.091)	(1.536)	(1.345)
Child is a girl	0.017***	0.019***	-0.004	0.022***	-0.037	-0.030
	(0.002)	(0.005)	(0.015)	(0.005)	(0.039)	(0.034)
R^2	0.587	0.556	0.540	0.562	0.621	0.611
Mother FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep. var.	0.14	0.12	0.12	0.12	0.10	0.15
Number of countries	13	12	12	11	8	9
Number of ethnic gr.	114	83	67	76	31	42
Observations	306,384	42,030	10,587	31,443	5,189	4,954

Notes: The unit of observation is a child younger than the age limit for cutting and under 18. The dependent variable is a binary indicator equal to one if the child does not coreside with his or her mother. All specifications include the following control variables: child's age fixed effects and mother fixed effects. Standard errors clustered at the mother level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D.9: Child fostering and girl's cutting status, excluding girls at risk of being married

	Dep. var.: Girl is cut							
	Sample: Girls							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Girl is fostered	-0.005	-0.025***	-0.010**	-0.010**	-0.005	-0.019***	-0.001	-0.001
	(0.004)	(0.004)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
Fostered × Prev.					-0.002	-0.007	-0.020**	-0.023**
					(0.009)	(0.009)	(0.010)	(0.010)
FGC prevalence	0.664***	0.508***	0.177***	0.121**	0.664***	0.509***	0.179***	0.123**
	(0.004)	(0.007)	(0.030)	(0.060)	(0.004)	(0.007)	(0.030)	(0.060)
R^2	0.290	0.305	0.659	0.663	0.290	0.305	0.659	0.663
Basic controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Child age FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
C. age x Cluster FE	No	No	Yes	Yes	No	No	Yes	Yes
C. age x E. group FE	No	No	No	Yes	No	No	No	Yes
Mean of dep. var.	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14
Number of countries	12	12	12	12	12	12	12	12
Number of ethnic gr.	94	94	94	94	94	94	94	94
Observations	85,849	85,849	85849	85,849	85,849	85849	85,849	85,849

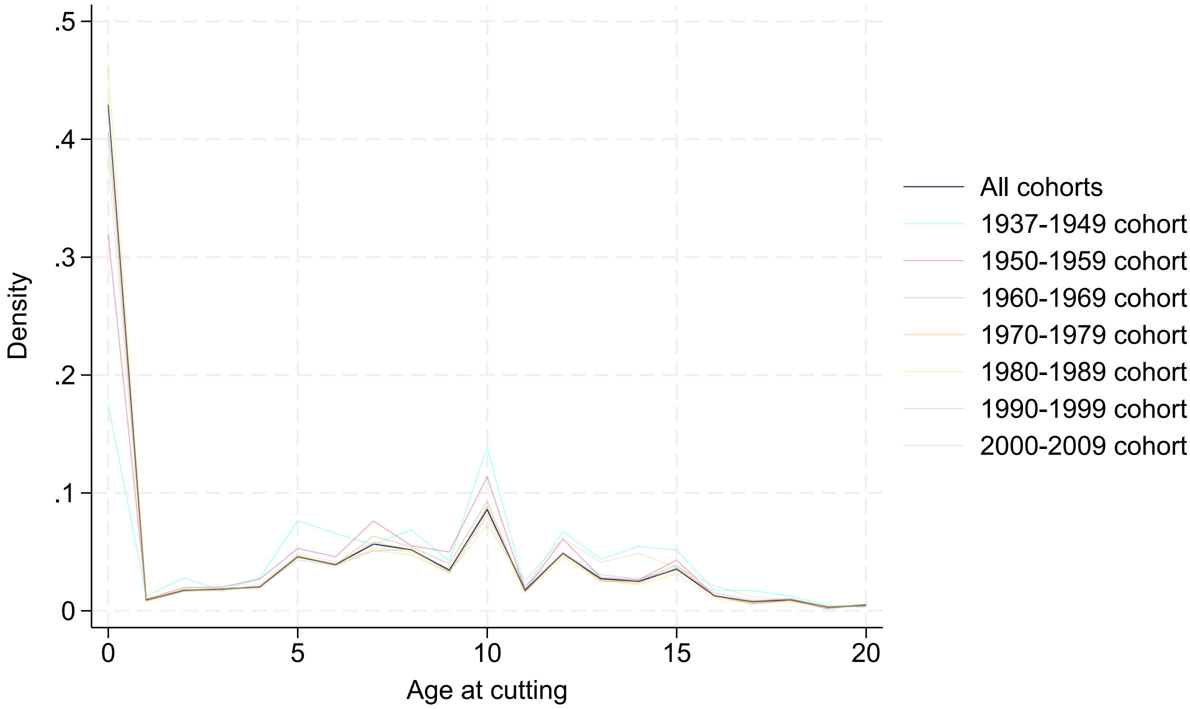
Notes: The unit of observation is a girl younger than the age limit for cutting and under 15. The dependent variable is a binary indicator equal to one if the girl is cut. Basic controls include year of enumeration, mother's level of education, mother's cutting status, and a dummy for urban residence. Additional fixed effects are indicated in the table. Standard errors clustered at the mother level are in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table D.10: FGC ban and child fostering in Senegal

	Dep. var.: Child is fostered				
	Sample: children of mothers interviewed in Senegal				
	All	Boys	Girls	Boys	Girls
	(1)	(2)	(3)	(4)	(5)
Born after ban × Girl	-0.020***				
	(0.005)				
Born after ban × Prev.		-0.013	-0.021**		
		(0.016)	(0.016)		
Born after ban × Practicing gp.				-0.012	-0.022**
				(0.011)	(0.011)
R^2	0.127	0.200	0.131	0.200	0.131
Cluster FE	Yes	Yes	Yes	Yes	Yes
Ethnic group FE	Yes	Yes	Yes	Yes	Yes
Mean of dep. var.	0.12	0.12	0.11	0.12	0.11
Number of ethnic gr.	6	6	6	6	6
Observations	79,352	40,094	39,130	40,094	39,130

Notes: The unit of observation is a child younger than the age limit for cutting and under 18. The dependent variable is a binary indicator equal to one if the child does not coreside with his or her mother. All specifications include the following control variables: child's age fixed effects and mother fixed effects. The difference in the number of observations between this table and Table 3.6 is due to the inclusion of mother fixed effects in the latter, which exclude mothers without within-mother variation in the relevant comparison. In Column 1, this requires children born both before and after the ban. In Columns 2-5, which are estimated separately by child sex, it requires children of the relevant sex born both before and after the ban. Standard errors clustered at the mother level are in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

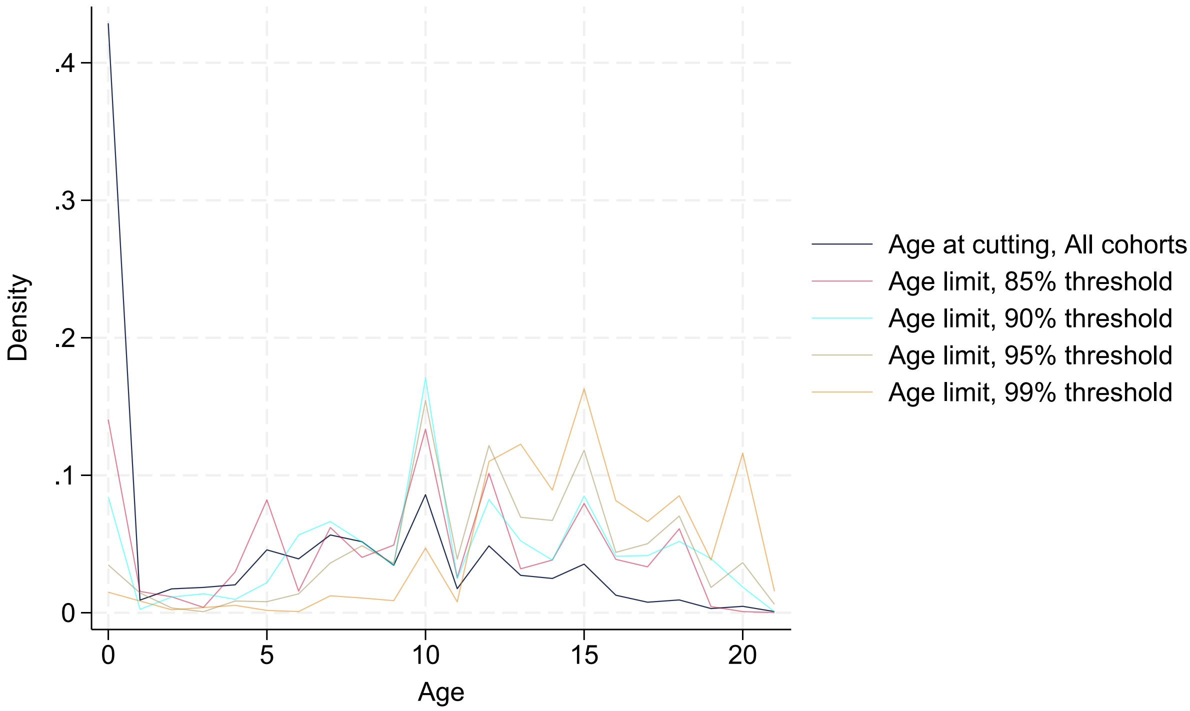
Figure D.4: Distribution of declared age at cutting among cohorts



Data sources: Author's elaboration on the DHS.

Notes: This figure plots the distribution of own age at cutting declared by mothers. The black solid line represents the distribution for the whole sample, while lighter lines give the distribution for each 10-year cohort.

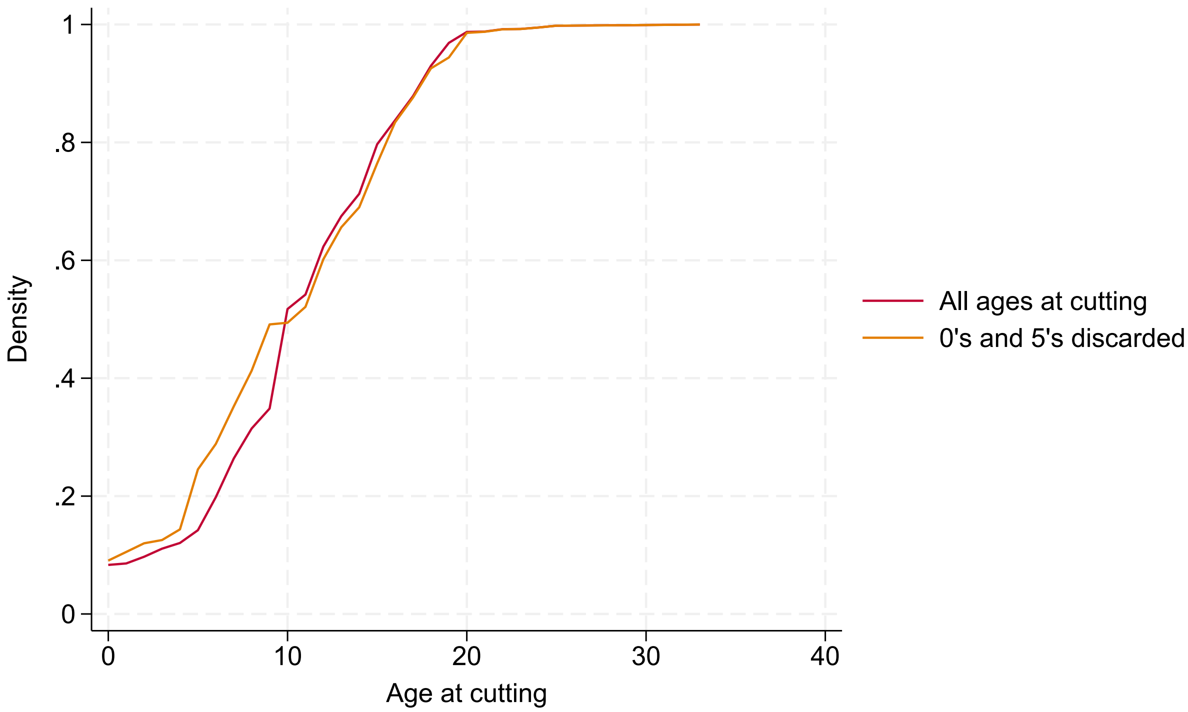
Figure D.5: Distribution of age limits at cutting



Data sources: Author's elaboration on the DHS.

Notes: This figure plots the distribution of the age limit at cutting, computed as a percentile of the own age at cutting declared by mothers. The black solid line again represents the distribution of age at cutting for the whole sample, while lighter lines give the distribution of various age percentiles limit for cutting.

Figure D.6: Distribution of age limits at cutting



Data sources: Author's elaboration on the DHS.

Notes: This figure plots the distribution of the age limit at cutting using the 90^{th} percentile and keeping all values of declared ages at cutting in red, and discarding declared ages ending with a 0 or a 5.

Table D.11: FGC prevalence and child fostering, excluding round ages

	Dep. var.: Child is fostered			
	Sample			
	All	Att. known	Anti-FGC	Pro-FGC
	(1)	(2)	(3)	(4)
Child girl × Prev.	0.017***	0.029***	0.032***	0.009
	(0.003)	(0.005)	(0.008)	(0.011)
FGC prevalence	-0.130***	-0.117***	-0.267***	-0.036
	(0.019)	(0.025)	(0.037)	(0.035)
Child is a girl	0.018***	0.009***	0.006*	0.025***
	(0.002)	(0.003)	(0.004)	(0.009)
R^2	0.596	0.611	0.608	0.614
Mother FE	Yes	Yes	Yes	Yes
Mean of dep. var.	0.14	0.16	0.15	0.17
Number of countries	13	13	13	13
Number of ethnic gr.	111	102	100	96
Observations	270,388	153,467	76,172	77,295

Notes: The unit of observation is a child younger than the age limit for cutting and under 18. The dependent variable is a binary indicator equal to one if the child does not coreside with his or her mother. All specifications include the following control variables: child's age fixed effects and mother fixed effects. To ensure that age heaping does not drive the results, ages at cutting ending in 0 or 5 have been excluded from the sample for the computation of age limit at cutting. The number of observations is lower than in the main specification because the age limits resulting from discarding round ages are on average lower than baseline age limits. Standard errors clustered at the mother level are in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table D.12: FGC prevalence (alternative measure) and child fostering

	Dep. var.: Child is fostered			
	Sample			
	All Mothers	Att. known	Anti-FGC	Pro-FGC
	(1)	(2)	(3)	(4)
Child girl × Prev.	0.016***	0.026***	0.032***	0.005
	(0.003)	(0.005)	(0.007)	(0.010)
FGC prevalence	-0.103***	-0.082***	-0.269***	0.008
	(0.017)	(0.023)	(0.035)	(0.031)
Child is a girl	0.017***	0.010***	0.007**	0.027***
	(0.002)	(0.003)	(0.003)	(0.008)
R^2	0.587	0.605	0.602	0.608
Mother FE	Yes	Yes	Yes	Yes
Mean of dep. var.	0.14	0.16	0.15	0.16
Number of countries	13	13	13	13
Number of ethnic gr.	114	104	101	98
Observations	306,384	171,791	85,698	86,093

Notes: The unit of observation is a child younger than the age limit for cutting and under 18. The dependent variable is a binary indicator equal to one if the child does not coreside with his or her mother. All specifications include the following control variables: child's age fixed effects and mother fixed effects. Standard errors clustered at the mother level are in parentheses. $^{*}\,p < 0.1$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

Table D.13: Mother's FGC status and child fostering

	Dep. var.: Child is fostered			
	Sample			
	All	Att. known	Anti-FGC	Pro-FGC
	(1)	(2)	(3)	
Child girl × Mother cut	0.016***	0.021***	0.021***	0.007
	(0.003)	(0.004)	(0.005)	(0.012)
Child is a girl	0.014***	0.009***	0.008***	0.023*
	(0.003)	(0.003)	(0.003)	(0.012)
R^2	0.599	0.599	0.599	0.599
Mother FE	Yes	Yes	Yes	Yes
Mean of dep. var.	0.15	0.15	0.15	0.15
Number of countries	13	13	13	13
Number of ethnic gr.	109	108	106	104
Observations	234,147	204,829	98,428	106,401

Notes: The unit of observation is a child younger than the age limit for cutting and under 18. The dependent variable is a binary indicator equal to one if the child does not coreside with his or her mother. All specifications include the following control variables: child's age fixed effects and mother fixed effects. Standard errors clustered at the mother level are in parentheses. $^{*}\, p < 0.1$, $^{**}\, p < 0.05$, $^{***}\, p < 0.01$.

Table D.14: FGC prevalence and child fostering, by FGC knowledge

	Dep. var.: Child is fostered			
	Sample			
	All mothers	FGC know. asked	Mother does not know FGC	Mother knows FGC
	(1)	(2)	(3)	(4)
Child girl × Prev.	0.017***	0.024***	-0.020	0.030***
	(0.003)	(0.005)	(0.028)	(0.005)
FGC prevalence	-0.134***	-0.124***	0.044	-0.120***
	(0.018)	(0.025)	(0.156)	(0.026)
Child is a girl	0.017***	0.012***	0.033***	0.008**
	(0.002)	(0.003)	(0.007)	(0.003)
R^2	0.587	0.607	0.608	0.607
Mother FE	Yes	Yes	Yes	Yes
Mean of dep. var.	0.14	0.16	0.17	0.16
Number of countries	13	12	12	12
Number of ethnic gr.	114	100	69	99
Observations	306,384	178,227	15,566	162,661

Notes: The unit of observation is a child younger than the age limit for cutting and under 18. The dependent variable is a binary indicator equal to one if the child does not coreside with his or her mother. All specifications include the following control variables: child's age fixed effects and mother fixed effects. Standard errors clustered at the mother level are in parentheses. $^{*}\,p < 0.1$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

Table D.15: FGC prevalence and child fostering, alternative thresholds

	Dep. var.: Child is fostered								
	Sample								
	All	Anti-FGC	Pro-FGC	All	Anti-FGC	Pro-FGC	All	Anti-FGC	Pro-FGC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Child girl × Prev., 85	0.018***	0.037***	-0.003						
	(0.003)	(0.008)	(0.011)						
FGC prevalence, 85	-0.058***	-0.265***	0.143***						
	(0.022)	(0.051)	(0.052)						
Child girl × Prev., 95				0.011***	0.028***	0.000			
				(0.003)	(0.007)	(0.009)			
FGC prevalence, 95				-0.036*	-0.114***	0.165***			
				(0.019)	(0.040)	(0.043)			
Child girl × Prev., 99							0.005*	0.020***	0.008
							(0.003)	(0.006)	(0.008)
FGC prevalence, 99							-0.092***	-0.103***	0.048
							(0.016)	(0.030)	(0.036)
Child is a girl	0.016***	0.007*	0.035***	0.020***	0.011***	0.030***	0.020***	0.011***	0.019***
	(0.002)	(0.004)	(0.009)	(0.002)	(0.003)	(0.008)	(0.002)	(0.003)	(0.006)
R^2	0.590	0.605	0.606	0.570	0.586	0.584	0.553	0.568	0.558
Mother FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep. var.	0.14	0.14	0.16	0.14	0.14	0.15	0.13	0.14	0.15
Number of countries	13	13	13	13	13	13	13	13	13
Number of ethnic gr.	111	90	90	120	105	98	120	104	99
Observations	284,377	76,015	80,014	404,271	108,513	120,267	566,197	160,161	162,414

Notes: The unit of observation is a child younger than the age limit for cutting and under 18. The dependent variable is a binary indicator equal to one if the child does not coreside with his or her mother. All specifications include the following control variables: child's age fixed effects and mother fixed effects. Standard errors clustered at the mother level are in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table D.16: Child fostering and girl's cutting status, excluding round ages

	Dep. var.: Girl is cut							
	Sample: Girls							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Girl is fostered	-0.004	-0.024***	-0.015***	-0.015***	-0.005	-0.020***	-0.000	0.001
	(0.004)	(0.004)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
Fostered × Prev.					0.001	-0.007	-0.028***	-0.030***
					(0.009)	(0.009)	(0.011)	(0.011)
FGC prevalence	0.640***	0.499***	0.223***	0.087	0.639***	0.500***	0.227***	0.090
	(0.004)	(0.008)	(0.033)	(0.068)	(0.004)	(0.008)	(0.033)	(0.068)
R^2	0.278	0.291	0.645	0.651	0.278	0.291	0.645	0.651
Basic controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Child age FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
C. age x Cluster FE	No	No	Yes	Yes	No	No	Yes	Yes
C. age x E. group FE	No	No	No	Yes	No	No	No	Yes
Mean of dep. var.	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
Number of countries	13	13	13	13	13	13	13	13
Number of ethnic gr.	102	102	102	102	102	102	102	102
Observations	82,806	82,806	82,806	82,806	82,806	82,806	82,806	82,806

Notes: The unit of observation is a girl younger than the age limit for cutting and under 15. The dependent variable is a binary indicator equal to one if the girl is cut. Basic controls include year of enumeration, mother's level of education, mother's cutting status, and a dummy for urban residence. Additional fixed effects are indicated in the table. Standard errors clustered at the mother level are in parentheses. $^{*}\, p < 0.1$, $^{**}\, p < 0.05$, $^{***}\, p < 0.01$.

Table D.17: Child fostering and girl's cutting status, alternative measure of prevalence

	Dep. var.: Girl is cut							
	Sample: Girls							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Girl is fostered	-0.002	-0.020***	-0.013***	-0.013***	-0.010**	-0.024***	0.000	0.002
	(0.004)	(0.004)	(0.005)	(0.005)	(0.004)	(0.004)	(0.005)	(0.005)
Fostered × Prev.					0.016*	0.008	-0.025**	-0.030***
					(0.009)	(0.009)	(0.011)	(0.011)
FGC prevalence	0.680***	0.472***	0.147***	-0.778**	0.678***	0.471***	0.150***	-0.775**
	(0.004)	(0.007)	(0.020)	(0.366)	(0.004)	(0.007)	(0.020)	(0.366)
R^2	0.277	0.300	0.660	0.665	0.277	0.300	0.660	0.665
Basic controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Child age FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
C. age x Cluster FE	No	No	Yes	Yes	No	No	Yes	Yes
C. age x E. group FE	No	No	No	Yes	No	No	No	Yes
Mean of dep. var.	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
Number of countries	12	12	12	12	12	12	12	12
Number of ethnic gr.	103	103	103	103	103	103	103	103
Observations	92,042	92,042	92,042	92,042	92,042	92,042	92,042	92,042

Notes: The unit of observation is a girl younger than the age limit for cutting and under 15. The dependent variable is a binary indicator equal to one if the girl is cut. Basic controls include year of enumeration, mother's level of education, mother's cutting status, and a dummy for urban residence. Additional fixed effects are indicated in the table. Standard errors clustered at the mother level are in parentheses. $^{*}\, p < 0.1$, $^{**}\, p < 0.05$, $^{***}\, p < 0.01$.

Table D.18: Child fostering and girl's cutting status, depending on mother's cutting status

	Dep. var.: Girl is cut							
	Sample: Girls							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Girl is fostered	-0.018***	-0.023***	-0.015***	-0.015***	0.001	-0.003	0.002	0.003
	(0.004)	(0.004)	(0.005)	(0.005)	(0.002)	(0.003)	(0.003)	(0.003)
Fostered × Mother cut					-0.026***	-0.027***	-0.024***	-0.025***
					(0.006)	(0.006)	(0.007)	(0.007)
Mother is cut	0.499***	0.480***	0.240***	0.232***	0.503***	0.484***	0.244***	0.236***
	(0.003)	(0.003)	(0.007)	(0.008)	(0.003)	(0.003)	(0.008)	(0.008)
R^2	0.215	0.228	0.643	0.649	0.215	0.228	0.643	0.649
Basic controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Child age FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
C. age x Cluster FE	No	No	Yes	Yes	No	No	Yes	Yes
C. age x E. group FE	No	No	No	Yes	No	No	No	Yes
Mean of dep. var.	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
Number of countries	12	12	12	12	12	12	12	12
Number of ethnic gr.	106	106	106	106	106	106	106	106
Observations	104,581	104,581	104,581	104,581	104,581	104,581	104,581	104,581

Notes: The unit of observation is a girl younger than the age limit for cutting and under 15. The dependent variable is a binary indicator equal to one if the girl is cut. Basic controls include year of enumeration, mother's level of education, mother's cutting status, and a dummy for urban residence. Additional fixed effects are indicated in the table. Standard errors clustered at the mother level are in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table D.19: Child fostering and girl's cutting status, alternative thresholds

	Dep. var.: Girl is cut											
	Sample: Girls											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Girl is fostered, 85	-0.009**	-0.016***	-0.004*	-0.005								
	(0.002)	(0.004)	(0.004)	(0.005)								
Fostered × Prev., 85			-0.024**	-0.031***								
			(0.010)	(0.011)								
Girl is fostered, 95					-0.006	-0.013***	-0.003	-0.004				
					(0.004)	(0.005)	(0.004)	(0.005)				
Fostered × Prev., 95							-0.022**	-0.027***				
							(0.009)	(0.010)				
Girl is fostered, 99									-0.004	-0.010**	-0.001	-0.002
									(0.003)	(0.004)	(0.003)	(0.004)
Fostered × Prev., 99											-0.020**	-0.025**
											(0.010)	(0.010)
R^2	0.302	0.671	0.302	0.671	0.286	0.649	0.286	0.649	0.271	0.630	0.271	0.630
Basic controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Child age FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
C. age x Cluster FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
C. age x E. group FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Mean of dep. var.	0.14	0.14	0.14	0.14	0.15	0.15	0.15	0.15	0.16	0.16	0.16	0.16
Number of countries	12	12	12	12	12	12	12	12	12	12	12	12
Number of e. group	97	97	97	97	106	106	106	106	111	111	111	111
Observations	78,412	78,412	78,412	78,412	104,735	104,735	104,735	104,735	126,984	126,984	126,984	126,984

Notes: The unit of observation is a girl younger than the age limit for cutting and under 15. The dependent variable is a binary indicator equal to one if the girl is cut. Basic controls include year of enumeration, mother's level of education, mother's cutting status, and a dummy for urban residence. Additional fixed effects are indicated in the table. Standard errors clustered at the mother level are in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table D.20: FGC ban and child fostering in Senegal, excluding round ages

	Dep. var.: Child is fostered		
	Sample: children of mothers interviewed in Senegal		
	All	Boys	Girls
	(1)	(2)	(3)
Born after ban × Girl	-0.037***		
	(0.007)		
Born after ban × Prevalence		-0.268	-1.427**
		(0.478)	(0.819)
R^2	0.538	0.653	0.538
Mother FE	Yes	Yes	Yes
Mean of dep. var.	0.12	0.12	0.12
Number of countries	1	1	1
Number of ethnic gr.	6	6	6
Observations	46,912	15,495	15,079

Notes: The unit of observation is a child younger than the age limit for cutting and under 18. The dependent variable is a binary indicator equal to one if the child does not coreside with his or her mother. All specifications include the following control variables: child's age fixed effects and mother fixed effects. Standard errors clustered at the mother level are in parentheses. $^{*}\,p < 0.1$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

Table D.21: FGC ban and child fostering in Senegal, alternative measure of prevalence

	Dep. var.: Child is fostered		
	Sample: children of mothers interviewed in Senegal		
	All	Boys	Girls
	(1)	(2)	(3)
Born after ban × Girl	-0.026***		
	(0.006)		
Born after ban × Prevalence		-0.094	-0.180**
		(0.065)	(0.079)
R^2	0.520	0.635	0.521
Mother FE	Yes	Yes	Yes
Mean of dep. var.	0.12	0.12	0.12
Number of countries	1	1	1
Number of ethnic gr.	6	6	6
Observations	63,789	21,980	21,264

Notes: The unit of observation is a child younger than the age limit for cutting and under 18. The dependent variable is a binary indicator equal to one if the child does not coreside with his or her mother. All specifications include the following control variables: child's age fixed effects and mother fixed effects. Standard errors clustered at the mother level are in parentheses. $^{*}\,p < 0.1$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

Table D.22: FGC ban and child fostering in Senegal, using mother's cutting status

	Dep. var.: Child is fostered		
		Sample	
	All	Boys	Girls
	(1)	(2)	(3)
Born after ban × Girl	-0.026***		
	(0.006)		
Born after ban × Mother cut		0.007	-0.072**
		(0.041)	(0.039)
R^2	0.520	0.661	0.530
Mother FE	Yes	Yes	Yes
Mean of dep. var.	0.12	0.13	0.13
Number of countries	1	1	1
Number of ethnic gr.	6	6	6
Observations	63,789	11,539	10,973

Notes: The unit of observation is a child younger than the age limit for cutting and under 18. The dependent variable is a binary indicator equal to one if the child does not coreside with his or her mother. All specifications include the following control variables: child's age fixed effects and mother fixed effects. Standard errors clustered at the mother level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D.23: FGC ban and child fostering in Senegal, mothers who never heard about cutting

	Dep. var.: Child is fostered		
		Sample	
	All	Boys	Girls
	(1)	(2)	(3)
Born after ban × Girl	-0.029		
	(0.068)		
Born after ban × Prevalence		-3.215	0.741
		(2.217)	(1.017)
R^2	0.529	0.718	0.494
Mother FE	Yes	Yes	Yes
Mean of dep. var.	0.12	0.13	0.11
Number of countries	1	1	1
Number of ethnic gr.	6	5	5
Observations	2,950	1,075	951

Notes: The unit of observation is a child younger than the age limit for cutting and under 18. The dependent variable is a binary indicator equal to one if the child does not coreside with his or her mother. All specifications include the following control variables: child's age fixed effects and mother fixed effects. Standard errors clustered at the mother level are in parentheses. $^{*}\,p < 0.1$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

Table D.24: FGC ban and child fostering in Senegal, alternative thresholds

	Dep. var.: Child is fostered								
	Sample: children of mothers interviewed in Senegal								
	All	Boys	Girls	All	Boys	Girls	All	Boys	Girls
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Born after ban × Girl, 85	-0.028**								
	(0.007)								
Born after ban × Prev., 85		-0.058	-0.151**						
		(0.065)	(0.075)						
Born after ban × Girl, 95				-0.024**					
				(0.006)					
Born after ban × Prev., 95					-0.061	-0.137**			
					(0.060)	(0.069)			
Born after ban × Girl, 99							-0.021**		
							(0.006)		
Born after ban × Prev., 99								-0.055	-0.124*
								(0.057)	(0.066)
R^2	0.531	0.645	0.534	0.518	0.632	0.520	0.506	0.619	0.508
Mother FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep. var.	0.11	0.11	0.11	0.12	0.12	0.12	0.12	0.12	0.12
Number of ethnic gr.	6	6	6	6	6	6	6	6	6
Observations	53,426	18,682	18,041	68,973	24,151	23,306	82,514	28,895	27,901

Notes: The unit of observation is a child younger than the age limit for cutting and under 18. The dependent variable is a binary indicator equal to one if the child does not coreside with his or her mother. All specifications include the following control variables: child's age fixed effects and mother fixed effects. Standard errors clustered at the mother level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

3.E Non-usable ethnicity

This section defines the cases in which the ethnicity information reported in the DHS cannot be used to assign respondents to an ethnic group, documents the prevalence of such cases, and discusses their implications for the analysis.

Respondents' ethnicity is retrieved from a dedicated question in the DHS questionnaire. I consider as non-usable any ethnicity response that corresponds either to a country name or to a supra-national denomination (such as "West African"), for two simple reasons. First, although some countries in the sample have a clearly dominant ethnic group, no country has a *unique* ethnicity. It is therefore not possible to identify a respondent's ethnicity solely from a country name, and even less so from a supra-national denomination. Second, country names or supra-national denominations reported as ethnicity generally correspond to residual categories similar to "other" (such as "other Togolese"), or arise from survey constraints where foreigners are asked to report a nationality rather than an ethnicity. For these reasons, I classify national- and supra-national-level responses as non-usable ethnicities (alongside residual or non-informative categories such as "other", "foreigner", and missing values, which are detailed below).

The wording of the question on ethnicity does not vary considerably across waves. However, the modalities of the answers do vary, even within a given country. The answer may include a list of ethnic groups to choose from, a blank, or both. Where there is no blank, there is always an "other" option for individuals who do not identify with any of the listed groups. However, where there is both a list and a blank, the situation is practically the same, as the DHS appears to recode these self-declared groups as "other" (except for Kenya, where some of these ethnicities can count as an answer, while there is still a sizeable "other" category). In these two, and most frequent, cases, people reporting a non-listed ethnicity will thus be classified as "other". Technically, these individuals could be reporting a small non-listed ethnic group, but also any non-usable group (such as "West African"). Country-waves with a list are those with a value "Yes" in Column 1 of Table E.1. Cases where there is a blank only (and no list) refer only to Ethiopia and Côte d'Ivoire (value "No" in Column 1). One could think that, in these cases, the actual full list of self-reported ethnicities would be available. However, the DHS still creates a residual "other" category for these waves, that can be more or less sizeable across waves. This suggests that there is also some reclassification by the DHS. Unlike in cases where there is a list, this reclassification does not seem to systematically affect smaller groups, as groups as small as one observation can be found in some of these waves. I thus interpret the category "other" in these waves where there is no list as actual non-usable ethnicities that the DHS reclassified, e.g., individuals reporting being West African.[44] These cases are actually very few: they represent at most 0.52 percent of the observations for Ethiopia 2016 (see the proportion of "other" in Column 3 for waves where the value of Column 1 is "No"). Please note that, as I only consider ethnicity-region-cohort groups larger than 50 individuals to compute prevalence at cutting, most of the small groups are in any case excluded from the analysis.

The case of foreigners also differs across waves. In some cases, they are required to state their nationality and not their ethnic group, or the enumerator is instructed to report "foreigner". These waves have a value of "No" in Column 2. In other cases, nothing is stated in particular and foreigners are not treated differently. These waves have a value of "Yes" in Column 2. In the first case, foreigners are thus discarded from the sample. In the second case, the outcome depends on the response format. If there is a list of groups (with or without blanks), foreigners are kept in the sample if they declare an ethnicity from the list, which is possible as numerous ethnic groups spread across borders. If there is a blank only, foreigners can report their ethnicity in the same way as natives do.

Table E.1: Non-usable ethnicities by country-wave

Country-wave	List (1)	F. asked (2)	% Other (3)	% Foreign (4)	% Don't know (5)	% Missing (6)	% Non-usable (7)	Obs. (8)
Nigeria 2013	Yes	Yes	29.813	0.000	0.005	0.148	29.966	81,032
Nigeria 2003	Yes	Yes	28.723	0.000	0.000	0.363	29.086	14,316
Nigeria 2008	Yes	Yes	28.401	0.000	0.000	0.609	29.009	69,650
Nigeria 2018	Yes	Yes	25.885	0.000	0.070	0.000	25.956	89,557
Côte d'Ivoire 1998	No	Yes	25.682	0.000	0.000	0.000	25.682	5,241
Côte d'Ivoire 2011	No	No	1.941	19.360	0.000	0.111	21.412	19,835
Nigeria 1990	Yes	Yes	21.137	0.000	0.000	0.185	21.322	19,473
Sierra Leone 2008	Yes	No	12.651	0.650	0.000	0.103	13.405	14,607
Mali 1995-6	Yes	No	10.212	0.610	0.000	0.114	10.937	23,609
Togo 1988	Yes	Yes	9.893	0.000	0.000	0.000	9.893	7,541
Senegal 1986	Yes	No	6.485	2.332	0.000	0.021	8.838	9,561

Continued on next page

[44]With the exception of Côte d'Ivoire 1998 where the proportion of "other" is very high and likely the result of a reclassification by the DHS.

Table E.1 (continued)

Country-wave	List (1)	F. asked (2)	% Other (3)	% Foreign (4)	% Don't know (5)	% Missing (6)	% Non-usable (7)	Obs. (8)
Togo 1998	Yes	No	2.631	6.043	0.000	0.092	8.765	18,550
Senegal 2010-11	Yes	No	6.417	2.327	0.000	0.000	8.743	31,247
Senegal 2015	Yes	No	5.294	3.208	0.000	0.000	8.502	18,078
Senegal 1997	Yes	No	5.474	2.502	0.000	0.031	8.008	19,144
Senegal 2017	Yes	No	5.020	2.617	0.000	0.000	7.637	33,665
Chad 2004	Yes	No	6.263	1.315	0.000	0.057	7.635	14,066
Senegal 2012-13	Yes	No	5.853	1.616	0.000	0.000	7.469	17,084
Kenya 2008-9	Yes	Yes	7.414	0.000	0.000	0.024	7.439	16,414
Burkina Faso 1998	Yes	No	6.472	0.702	0.000	0.000	7.174	14,386
Burkina Faso 1993	Yes	No	5.451	1.171	0.028	0.114	6.764	14,089
Mali 2012	Yes	No	5.778	0.617	0.000	0.000	6.395	26,254
Senegal 2005	Yes	No	4.625	1.607	0.000	0.050	6.282	28,002
Burkina Faso 2010	Yes	No	5.622	0.373	0.020	0.207	6.222	39,117
Sierra Leone 2019	Yes	No	5.784	0.235	0.000	0.000	6.019	28,059
Senegal 2016	Yes	No	3.578	2.434	0.000	0.000	6.012	17,830
Mali 2006	Yes	No	3.795	2.100	0.000	0.058	5.953	34,337
Sierra Leone 2013	Yes	No	5.239	0.342	0.000	0.224	5.805	31,302
Kenya 1989	Yes	Yes	4.762	0.000	0.000	0.000	4.762	18,816
Benin 1996	Yes	No	0.452	3.771	0.000	0.396	4.619	12,623
Mali 2001	Yes	No	2.916	1.398	0.000	0.194	4.509	30,894
Senegal 1992-93	Yes	No	1.969	2.404	0.000	0.055	4.427	14,478
Benin 2001	Yes	No	1.056	3.243	0.000	0.037	4.337	13,350
Benin 2006	Yes	No	0.548	3.287	0.000	0.280	4.114	40,710
Niger 1992	Yes	No	0.109	3.430	0.000	0.408	3.947	14,695
Kenya 1993	Yes	Yes	3.248	0.000	0.000	0.176	3.424	17,642
Senegal 2014	Yes	No	0.959	2.414	0.000	0.000	3.373	17,316
Burkina Faso 2003	Yes	Yes	2.895	0.000	0.000	0.215	3.110	27,460
Kenya 1998	Yes	Yes	2.271	0.000	0.000	0.096	2.367	16,601
Guinea 2012	Yes	Yes	2.131	0.000	0.000	0.000	2.131	19,194
Kenya 2014	Yes	Yes	2.098	0.000	0.000	0.022	2.119	64,171
Niger 1998	Yes	Yes	1.940	0.000	0.000	0.006	1.945	17,375
Benin 2011	Yes	No	0.447	1.131	0.000	0.000	1.578	36,945
Kenya 2003	Yes	Yes	1.393	0.000	0.000	0.000	1.393	15,719
Guinea 1999	Yes	No	0.196	1.093	0.000	0.000	1.289	14,818
Niger 2006	Yes	Yes	1.014	0.000	0.000	0.017	1.032	22,972
Guinea 2005	Yes	No	0.224	0.765	0.000	0.000	0.989	17,389
Ethiopia 2011	No	No	0.141	0.016	0.000	0.683	0.839	31,936
Ethiopia 2005	No	No	0.453	0.055	0.000	0.274	0.783	27,343

Continued on next page

Table E.1 (continued)

Country-wave	List (1)	F. asked (2)	% Other (3)	% Foreign (4)	% Don't know (5)	% Missing (6)	% Non-usable (7)	Obs. (8)
Ethiopia 2016	No	No	0.515	0.020	0.000	0.134	0.669	30,662
Guinea 2018	Yes	No	0.000	0.325	0.000	0.000	0.325	21,227
Ethiopia 2000	No	No	0.000	0.032	0.076	0.007	0.115	27,779
Mean			8.333	1.169	0.007	0.126	9.635	1,339,870

Notes: The unit of observation is a child (boy or girl) below 18. Children are identified through the mother's survey.

Finally, individuals can answer that they do not know their ethnicity, and then be coded as "don't know", or simply do not answer the question and be classified as "missing". Please note that "missing" answers also encompass cases where the question should have been asked but was not, for example because the enumerator inadvertently missed it or because the interview had to stop. Please also note that these interpretations rely on indirect evidence from the distribution of categories, as the precise DHS wave-by-wave recoding rules are not documented.

This means that individuals reporting a non-usable ethnic group, individuals who do not know their ethnicity, and individuals for whom no usable answer is recorded are excluded from the analysis in all waves. In waves where there is a list, natives belonging to smaller non-listed groups and foreigners not belonging to these groups are also excluded. In waves where foreigners are asked for their nationality rather than their ethnicity, or are recorded simply as "foreigner", all foreigners are discarded from the sample. The number of foreigners can be counted precisely only in such waves, and their proportion in the sample is displayed in Column 4.[45] In other waves, foreigners are divided between listed ethnicities, if they are part of one, and the "other" category. The number of individuals reporting a non-usable ethnicity can be counted precisely in waves without a list and where foreigners are not asked for their ethnicity (i.e., Column 1 is "No" and Column 2 is "No"), under the assumption that they are the only ones put in the "other" category. In waves where there is a list, the individuals reporting a non-usable ethnicity cannot be distinguished from those reporting a smaller non-

[45]Ethiopia is an exception here: it has a proportion of foreigners even though they are allowed to report an ethnicity. Ethiopia does not have a list of ethnicities, so foreigners can declare their group directly. The proportion in Column 4 likely represents foreigners declaring a non-usable ethnicity. They represent between 3 and 9 responses depending on the survey year.

listed ethnicity. Based on waves without a list, the former category is likely much smaller than the latter. This means that Column 3 is either only individuals reporting a non-usable ethnicity (Column 1 is "No" and Column 2 is "No"), a mix of individuals reporting a non-usable ethnicity and foreigners (Column 1 is "No" and Column 2 is "Yes"), a mix of individuals reporting a non-usable ethnicity and individuals belonging to a small non-listed ethnicity (Column 1 is "Yes" and Column 2 is "No"), or a mix of individuals reporting a non-usable ethnicity, foreigners, and individuals belonging to a small non-listed ethnicity (Column 1 is "Yes" and Column 2 is "Yes").

Individuals who do not know their ethnicity can be counted precisely in all waves. The proportion for each wave is displayed in Column 5. Individuals who did not want to answer the question cannot be counted precisely, as they are systematically mixed with cases where the question was not asked for various unrelated reasons. The proportions displayed in Column 6 are thus likely slightly inflated. Column 7 sums up the proportion of non-usable ethnicities for each country-wave, computed as the sum of each column, and Column 8 gives the number of observations for each country-wave. In the whole sample, 9.63 percent of respondents cannot be linked to a usable ethnicity. The disproportionately higher proportion of non-usable ethnicities for Nigeria comes from the fact that the DHS provides two variables for this country: a detailed version and a non-detailed version. The detailed version is based on a survey without a list, where individuals can report their ethnicity freely, and is not available for Nigeria 2018. The absence of a list, coupled with the high ethnic fractionalization of Nigeria, leads to a high number of small categories (i.e., under 50 individuals) listed for this variable. This, coupled with the absence of the 2018 wave, makes the use of this variable less relevant for the analysis, but explains the high level of "other" for this country: it is mostly made up of smaller non-listed groups in the non-detailed version of the variable. When excluding Nigeria, the proportion of non-usable ethnicities drops to 4.98 percent. This is illustrative of a more general fact: the non-random proportion of non-usable ethnicity is not primarily explained by individuals deliberately reporting no ethnicity, or larger groups not usable as ethnicities, but rather by the differential structure of the questionnaire across waves.

This distinction is worth noting, as Bertoli et al. (2025) show in the context of the United States that whether individuals report a foreign ancestry is correlated with other cultural traits.[46] Even if they represent a small share of the sample, respondents who deliberately

[46] The example is slightly different, as foreign ancestry is intended to capture migration history, whereas

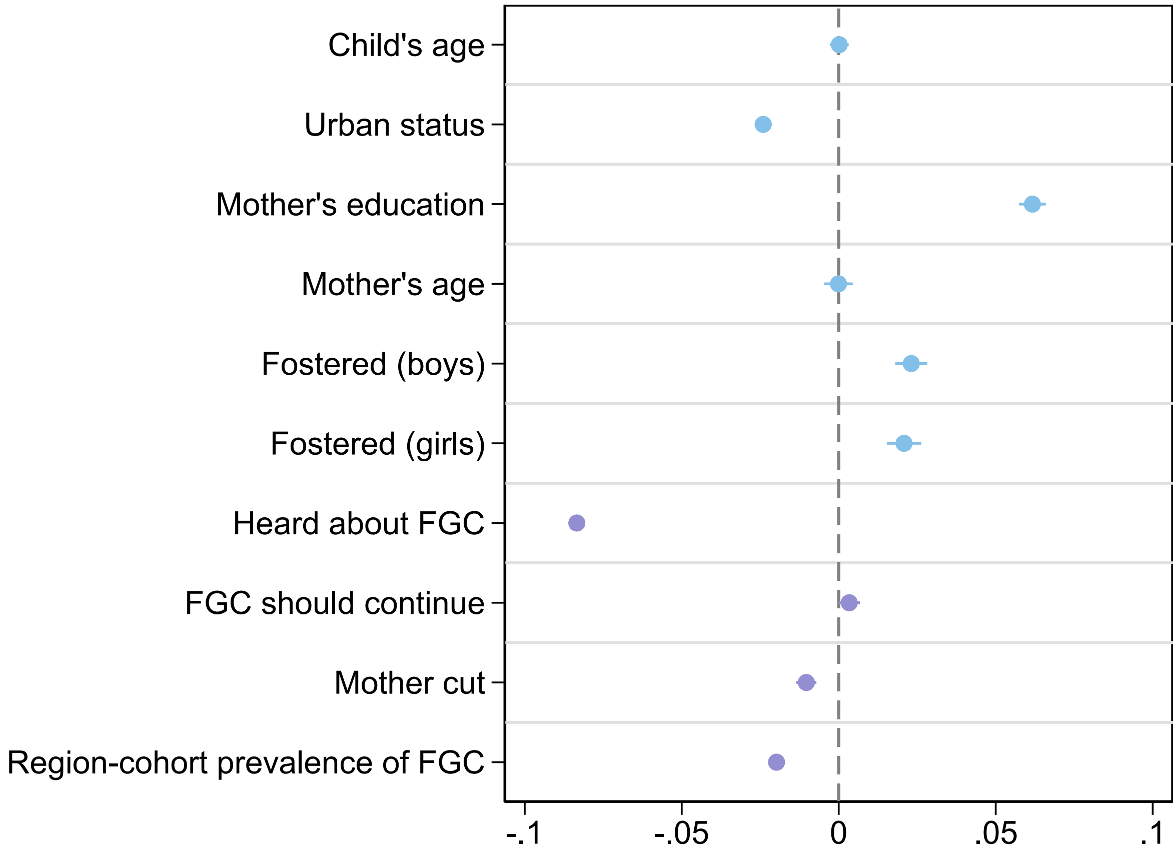
report no usable ethnicity may therefore generate bias if selection into reporting ethnicity is strong.

I thus explore whether children whose mothers do not report a usable ethnicity differ from the rest of the sample by regressing various outcomes on a dummy for reporting a non-usable ethnicity, and country and year-of-enumeration fixed effects. I present results in Figure E.1. Children of mothers declaring a non-usable ethnicity are more likely to be fostered, and their mothers are on average more educated, less aware of cutting, and slightly less likely to be cut. They are less likely to reside in urban areas and be cut. Unfortunately, even with country and year-of-enumeration fixed effects, it is not possible to disentangle self-selection into reporting ethnicity from questionnaire structure. For example, smaller groups not listed in the questionnaire are bundled under the "other" category, while possibly having characteristics distinct from the rest of the sample.

To check whether differences in questionnaire structure and/or self-selection into reporting a usable ethnicity affect the results, I re-estimate the main specifications (i) separately by the value of the variable *List*, (ii) separately by the value of the variable *Foreign asked*, and (iii) restricting the sample to country-waves in which the proportion of non-usable ethnicity is below 10%.

ethnicity applies to the whole population. It nevertheless provides a concrete example of how questions on subjective traits may generate selected samples.

Figure E.1: Characteristics of children from a mother declaring a non-usable vs. usable ethnicity



Notes: The unit of observation is a child under 18. Children are identified through the mother's survey, which means that mothers with multiple children appear multiple times in the dataset. All variables but child's age and fostering status refer to the mother. Child's and mother's ages have been divided by 10 to fit in the graph, which needs to be taken into account when considering the marginal effects. 95% confidence intervals are displayed.

Table E.2: FGC prevalence and child fostering: country-waves with a list for the question on ethnicity

	Dep. var.: Child is fostered			
	All Mothers	Att. known	Anti-FGC	Pro-FGC
	(1)	(2)	(3)	(4)
Child girl × Prev.	0.022***	0.031***	0.040***	0.009
	(0.003)	(0.005)	(0.008)	(0.011)
FGC prevalence	-0.111***	-0.048	-0.306***	0.069
	(0.023)	(0.036)	(0.054)	(0.049)
Child is a girl	0.017***	0.009***	0.006*	0.030***
	(0.002)	(0.003)	(0.003)	(0.010)
R^2	0.587	0.607	0.602	0.612
Mother FE	Yes	Yes	Yes	Yes
Mean of dep. var.	0.15	0.17	0.15	0.18
Number of countries	11	11	11	11
Number of ethnic gr.	92	82	81	78
Observations	275,460	150,372	76,533	73,839

Notes: The unit of observation is a child younger than the age limit for cutting and under 18. The dependent variable is a binary indicator equal to one if the child does not coreside with his or her mother. All specifications include the following control variables: child's age fixed effects and mother fixed effects. Standard errors clustered at the mother level are in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table E.3: FGC prevalence and child fostering: country-waves without a list for the question on ethnicity

	Dep. var.: Child is fostered			
	All Mothers	Att. known	Anti-FGC	Pro-FGC
	(1)	(2)	(3)	(4)
Child girl × Prev.	0.001	0.016	0.020*	0.014
	(0.014)	(0.019)	(0.011)	(0.028)
FGC prevalence	-0.036	-0.055	-0.233***	0.022
	(0.029)	(0.035)	(0.062)	(0.042)
Child is a girl	0.007	0.003	0.004	-0.001
	(0.011)	(0.014)	(0.018)	(0.023)
R^2	0.564	0.562	0.606	0.499
Mother FE	Yes	Yes	Yes	Yes
Mean of dep. var.	0.08	0.09	0.12	0.06
Number of countries	2	2	2	2
Number of ethnic gr.	22	22	20	20
Observations	30,924	21,419	9,165	12,254

Notes: The unit of observation is a child younger than the age limit for cutting and under 18. The dependent variable is a binary indicator equal to one if the child does not coreside with his or her mother. All specifications include the following control variables: child's age fixed effects and mother fixed effects. Standard errors clustered at the mother level are in parentheses. $^{*}\, p < 0.1$, $^{**}\, p < 0.05$, $^{***}\, p < 0.01$.

Table E.4: FGC prevalence and child fostering: country-waves where foreigners are asked for their ethnicity

	Dep. var.: Child is fostered			
	All Mothers	Att. known	Anti-FGC	Pro-FGC
	(1)	(2)	(3)	(4)
Child girl × Prev.	0.027***	0.035***	0.039***	0.016
	(0.007)	(0.009)	(0.015)	(0.017)
FGC prevalence	-0.311***	-0.269***	-0.311***	-0.051
	(0.029)	(0.057)	(0.062)	(0.146)
Child is a girl	0.018***	0.014***	0.011**	0.028**
	(0.003)	(0.005)	(0.005)	(0.012)
R^2	0.597	0.607	0.624	0.571
Mother FE	Yes	Yes	Yes	Yes
Mean of dep. var.	0.13	0.15	0.15	0.16
Number of countries	7	7	7	7
Number of ethnic gr.	54	48	46	44
Observations	98,447	46,661	32,049	14,612

Notes: The unit of observation is a child younger than the age limit for cutting and under 18. The dependent variable is a binary indicator equal to one if the child does not coreside with his or her mother. All specifications include the following control variables: child's age fixed effects and mother fixed effects. Standard errors clustered at the mother level are in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table E.5: FGC prevalence and child fostering: country-waves where foreigners are not asked for their ethnicity

	Dep. var.: Child is fostered			
	All Mothers	Att. known	Anti-FGC	Pro-FGC
	(1)	(2)	(3)	(4)
Child girl × Prev.	0.017***	0.030***	0.032***	0.009
	(0.004)	(0.006)	(0.009)	(0.014)
FGC prevalence	-0.056**	-0.112***	-0.301***	-0.042
	(0.023)	(0.028)	(0.044)	(0.036)
Child is a girl	0.015***	0.004	0.002	0.015
	(0.002)	(0.004)	(0.004)	(0.012)
R^2	0.583	0.604	0.590	0.615
Mother FE	Yes	Yes	Yes	Yes
Mean of dep. var.	0.15	0.16	0.15	0.16
Number of countries	11	9	9	9
Number of ethnic gr.	79	68	66	66
Observations	207,937	125,130	53,649	71,481

Notes: The unit of observation is a child younger than the age limit for cutting and under 18. The dependent variable is a binary indicator equal to one if the child does not coreside with his or her mother. All specifications include the following control variables: child's age fixed effects and mother fixed effects. Standard errors clustered at the mother level are in parentheses. $^{*}\,p < 0.1$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

Table E.6: FGC prevalence and child fostering: country-waves with at least 90% of usable ethnicities

	Dep. var.: Child is fostered			
	All Mothers	Att. known	Anti-FGC	Pro-FGC
	(1)	(2)	(3)	(4)
Child girl × Prev.	0.017***	0.029***	0.036***	0.012
	(0.003)	(0.005)	(0.008)	(0.012)
FGC prevalence	-0.181***	-0.225***	-0.361***	-0.135***
	(0.019)	(0.027)	(0.036)	(0.039)
Child is a girl	0.016***	0.007**	0.005	0.021**
	(0.002)	(0.003)	(0.004)	(0.010)
R^2	0.583	0.606	0.595	0.617
Mother FE	Yes	Yes	Yes	Yes
Mean of dep. var.	0.14	0.15	0.14	0.16
Number of countries	11	11	11	11
Number of ethnic gr.	99	90	86	84
Observations	266,636	140,009	71,055	68,954

Notes: The unit of observation is a child younger than the age limit for cutting and under 18. The dependent variable is a binary indicator equal to one if the child does not coreside with his or her mother. All specifications include the following control variables: child's age fixed effects and mother fixed effects. Standard errors clustered at the mother level are in parentheses. $^{*}\,p < 0.1$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

Table E.7: Daughter's fostering and cutting status: country-waves with a list for the question on ethnicity

	Dep. var.: Girl is cut							
	Sample: Girls							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Girl is fostered	-0.012***	-0.031***	-0.010**	-0.010**	-0.001	-0.016***	0.002	0.004
	(0.004)	(0.004)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
Fostered × Prev.					-0.021**	-0.028***	-0.022**	-0.025**
					(0.009)	(0.009)	(0.010)	(0.010)
FGC prevalence	0.684***	0.524***	0.178***	0.131**	0.687***	0.528***	0.181***	0.134**
	(0.004)	(0.007)	(0.029)	(0.058)	(0.004)	(0.008)	(0.029)	(0.058)
R^2	0.324	0.341	0.670	0.674	0.324	0.341	0.670	0.674
Basic controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Child age FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
C. age x Cluster FE	No	No	Yes	Yes	No	No	Yes	Yes
C. age x Ethnic gr. FE	No	No	No	Yes	No	No	No	Yes
Mean of dep. var.	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16
Number of countries	10	10	10	10	10	10	10	10
Number of e. group	80	80	80	80	80	80	80	80
Observations	82,079	82,079	82,079	82,079	82,079	82,079	82,079	82,079

Notes: The unit of observation is a girl younger than the age limit for cutting and under 15. The dependent variable is a binary indicator equal to one if the girl is cut. Basic controls include year of enumeration, mother's level of education, mother's cutting status, and a dummy for urban residence. Additional fixed effects are indicated in the table. Standard errors clustered at the mother level are in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table E.8: Daughter's fostering and cutting status: country-waves without a list for the question on ethnicity

	Dep. var.: Girl is cut							
	Sample: Girls							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Girl is fostered	-0.010	-0.030*	-0.054***	-0.051***	-0.125***	-0.107***	-0.059**	-0.067**
	(0.017)	(0.017)	(0.019)	(0.020)	(0.022)	(0.025)	(0.030)	(0.030)
Fostered × Prev.					0.199***	0.130***	0.008	0.025
					(0.044)	(0.045)	(0.056)	(0.058)
FGC prevalence	0.440***	0.155***	0.314***	-0.617*	0.412***	0.142***	0.312***	-0.620*
	(0.020)	(0.031)	(0.088)	(0.364)	(0.022)	(0.031)	(0.089)	(0.364)
R^2	0.050	0.101	0.584	0.595	0.051	0.101	0.584	0.595
Basic controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Child age FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
C. age x Cluster FE	No	No	Yes	Yes	No	No	Yes	Yes
C. age x Ethnic gr. FE	No	No	No	Yes	No	No	No	Yes
Mean of dep. var.	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08
Number of countries	2	2	2	2	2	2	2	2
Number of e. group	23	23	23	23	23	23	23	23
Observations	9,960	9,960	9,960	9,960	9,960	9,960	9,960	9,960

Notes: The unit of observation is a girl younger than the age limit for cutting and under 15. The dependent variable is a binary indicator equal to one if the girl is cut. Basic controls include year of enumeration, mother's level of education, mother's cutting status, and a dummy for urban residence. Additional fixed effects are indicated in the table. Standard errors clustered at the mother level are in parentheses. $^{*}\,p < 0.1$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

Table E.9: Daughter's fostering and cutting status: where foreigners are asked for their ethnicity

	Dep. var.: Girl is cut							
	Sample: Girls							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Girl is fostered	0.016**	0.012*	-0.003	-0.002	0.006	0.010	0.003	0.005
	(0.007)	(0.007)	(0.008)	(0.009)	(0.008)	(0.008)	(0.009)	(0.009)
Fostered × Prev.					0.027	0.004	-0.013**	-0.015**
					(0.019)	(0.009)	(0.006)	(0.006)
FGC prevalence	0.681***	0.507***	0.275***	0.560**	0.677***	0.506***	0.277***	0.561**
	(0.009)	(0.013)	(0.058)	(0.226)	(0.009)	(0.013)	(0.058)	(0.226)
R^2	0.302	0.341	0.671	0.678	0.302	0.341	0.671	0.678
Basic controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Child age FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
C. age x Cluster FE	No	No	Yes	Yes	No	No	Yes	Yes
C. age x Ethnic gr. FE	No	No	No	Yes	No	No	No	Yes
Mean of dep. var.	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
Number of countries	6	6	6	6	6	6	6	6
Number of e. group	46	46	46	46	46	46	46	46
Observations	23,915	23,915	23,915	23,915	23,915	23,915	23,915	23,915

Notes: The unit of observation is a girl younger than the age limit for cutting and under 15. The dependent variable is a binary indicator equal to one if the girl is cut. Basic controls include year of enumeration, mother's level of education, mother's cutting status, and a dummy for urban residence. Additional fixed effects are indicated in the table. Standard errors clustered at the mother level are in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table E.10: Daughter's fostering and cutting status: country-waves where foreigners are not asked for their ethnicity

	Dep. var.: Girl is cut							
	Sample: Girls							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Girl is fostered	-0.014***	-0.035***	-0.016***	-0.016***	-0.011**	-0.032***	-0.003	-0.002
	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.005)	(0.005)
Fostered × Prev.					-0.005	-0.005	-0.022**	-0.024**
					(0.010)	(0.010)	(0.011)	(0.011)
FGC prevalence	0.674***	0.539***	0.154***	0.079	0.675***	0.540***	0.157***	0.082
	(0.004)	(0.010)	(0.032)	(0.059)	(0.004)	(0.010)	(0.032)	(0.059)
R^2	0.278	0.291	0.652	0.657	0.278	0.291	0.652	0.657
Basic controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Child age FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
C. age x Cluster FE	No	No	Yes	Yes	No	No	Yes	Yes
C. age x Ethnic gr. FE	No	No	No	Yes	No	No	No	Yes
Mean of dep. var.	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
Number of countries	9	9	9	9	9	9	9	9
Number of e. group	70	70	70	70	70	70	70	70
Observations	68,124	68,124	68,124	68,124	68,124	68,124	68,124	68,124

Notes: The unit of observation is a girl younger than the age limit for cutting and under 15. The dependent variable is a binary indicator equal to one if the girl is cut. Basic controls include year of enumeration, mother's level of education, mother's cutting status, and a dummy for urban residence. Additional fixed effects are indicated in the table. Standard errors clustered at the mother level are in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table E.11: Daughter's fostering and cutting status: country-waves with at least 90% of usable ethnicities

	Dep. var.: Girl is cut							
	Sample: Girls							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Girl is fostered	-0.008*	-0.025***	-0.012**	-0.012**	-0.005	-0.020***	0.001	0.003
	(0.004)	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
Fostered × Prev.					-0.006	-0.009	-0.024**	-0.026**
					(0.009)	(0.009)	(0.010)	(0.010)
FGC prevalence	0.666***	0.527***	0.158***	0.120**	0.667***	0.529***	0.161***	0.123**
	(0.004)	(0.008)	(0.030)	(0.059)	(0.004)	(0.008)	(0.030)	(0.059)
R^2	0.306	0.320	0.664	0.670	0.306	0.320	0.664	0.670
Basic controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Child age FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
C. age x Cluster FE	No	No	Yes	Yes	No	No	Yes	Yes
C. age x Ethnic gr. FE	No	No	No	Yes	No	No	No	Yes
Mean of dep. var.	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
Number of countries	10	10	10	10	10	10	10	10
Number of e. group	89	89	89	89	89	89	89	89
Observations	76,227	76,227	76,227	76,227	76,227	76,227	76,227	76,227

Notes: The unit of observation is a girl younger than the age limit for cutting and under 15. The dependent variable is a binary indicator equal to one if the girl is cut. Basic controls include year of enumeration, mother's level of education, mother's cutting status, and a dummy for urban residence. Additional fixed effects are indicated in the table. Standard errors clustered at the mother level are in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

General Conclusion

This thesis has studied the links between culture, spatial mobility, and social norms through three empirical essays in development economics. Its starting point is a simple but important difficulty: cultural traits are rarely observed directly. Researchers must therefore rely on proxies, assign individuals to origins or groups, and interpret the resulting measures as capturing the norms, beliefs, or preferences to which individuals have been exposed. This operation becomes especially delicate when individuals move, or when their ancestors moved, because place of residence, place of origin, declared identity, and context of socialization may no longer coincide.

The first contribution of the thesis is methodological. The first chapter shows that empirical results on cultural persistence can be highly sensitive to the way data are constructed and individuals are assigned to an origin. By conducting a deep reproduction of Giuliano and Nunn (2021), it documents discrepancies between the specifications described in the article and the code used to produce the results. Correcting these discrepancies when possible leads to a more nuanced relationship between ancestral climatic variability and the importance attached to tradition. Beyond this specific case, the chapter highlights a blind spot in reproducibility checks: code may run correctly and reproduce published tables while remaining inconsistent with the empirical choices described in the text.

The second contribution concerns the identification of cultural effects using migration. The epidemiological approach treats migration as a methodological resource: by comparing individuals of different origins within a common destination country, it seeks to separate inherited culture from the contemporary environment. The second chapter shows that this strategy depends critically on how origin is measured. In the United States, self-reported foreign ancestry does not only capture ancestral origin. It also reflects the timing of ancestral migration and selection into continuing to report a foreign ancestry. The chapter proposes simple proxies for this heterogeneity and shows that some results attributed to cultural persistence may partly capture differences in migration timing or selection. This does not invalidate the epidemiological approach, but it clarifies the conditions under which it can be interpreted as evidence of cultural transmission.

The final chapter shifts the focus from mobility as a measurement problem or an identification tool to mobility as a possible response to social norms. In the context of female genital cutting in sub-Saharan Africa, it explores whether parents who are opposed to the norm, but reside in communities where it is deeply entrenched, tend to use child fostering as a protective strategy. The results show that girls are more likely than boys to be fostered out in high-prevalence groups, especially when their mothers oppose cutting. This pattern is stronger when families have credible exit options provided by inter-ethnic marriage and substantial geographic variation in prevalence, and fostered girls are also less likely to be cut. The chapter therefore suggests that child fostering can serve as a spatial form of exit from a locally binding norm.

Together, the three chapters show that mobility plays three distinct roles in the empirical study of culture. It can weaken identification by making origins harder to measure; it can help identify cultural effects when used carefully; and it can allow individuals to change their exposure to social pressure. More broadly, the thesis argues that culture should not be studied independently from the empirical operations that measure it, nor from the spatial movements that reshape the link between individuals, places, and groups.

Résumé substantiel

Cette thèse est composée de trois essais empiriques en économie du développement, et porte sur les liens entre culture, mobilité spatiale et normes sociales. Son fil directeur est double. Elle interroge d'une part la manière dont les chercheurs mesurent la culture et rattachent les individus à une origine culturelle lorsque la mobilité rompt le lien entre lieu de résidence et origine ancestrale. Elle analyse d'autre part la mobilité comme outil d'identification et comme réponse aux contraintes de normes sociales locales.

Le premier chapitre propose une réplication approfondie de l'article de Giuliano et Nunn (2021), qui relie la variabilité climatique ancestrale à une moindre importance accordée aujourd'hui à la tradition. L'article original avance l'hypothèse selon laquelle les sociétés historiquement exposées à une forte variabilité climatique intergénérationnelle auraient accordé moins de valeur à la tradition, parce que les connaissances transmises par les générations précédentes auraient été moins adaptées à des environnements changeants. À partir de cette intuition, Giuliano et Nunn (2021) testent tout au long du papier si les individus dont les ancêtres ont connu une plus forte variabilité climatique accordent aujourd'hui moins d'importance à la tradition.

Le chapitre va au-delà de la reproduction des résultats publiés à partir du code fourni par les auteurs (*computational reproducibility*). Il compare systématiquement les spécifications décrites dans le texte de l'article avec les codes et les données de réplication disponibles. Cette démarche permet de mettre en évidence plusieurs écarts entre ce qui est annoncé dans l'article et ce qui est effectivement estimé dans les fichiers de réplication, notamment concernant la construction de certaines variables, la définition des échantillons et la spécification de plusieurs estimations. Lorsque les données disponibles le permettent, le chapitre corrige ces écarts et ré-estime les modèles correspondants. La correction de ces écarts laisse apparaître une relation entre variabilité climatique ancestrale et importance de la tradition plus nuancée que dans les estimations initiales. Au-delà du papier de Giuliano et Nunn (2021), le chapitre met en lumière un angle mort du processus de publication en économie. Le texte est évalué par l'éditeur et les rapporteurs, tandis que le code peut être revu ou non en fonction des journaux, mais souvent par un acteur autre que ceux cités précédemment, typiquement le *Data Editor*. Ce dernier vérifie que le code fourni génère bien les résultats présentés, mais pas nécessairement qu'il teste les mêmes spécifications que celles présentées dans le papier. Cette absence de vérification de la cohérence entre le texte et le code peut laisser subsister des contradictions entre les deux. La contribution du chapitre est donc double. Il participe d'une part à la littérature sur la reproductibilité en économie, en insistant sur l'importance de vérifier la cohérence entre texte, code et données. Il montre d'autre part que les résultats portant sur la persistance culturelle peuvent être sensibles aux choix d'assignation des individus à une origine et aux décisions prises lors de la construction des variables.

Le deuxième chapitre étudie plus directement l'approche épidémiologique, une stratégie empirique très utilisée en économie de la culture. Cette approche consiste à comparer des migrants ou des descendants de migrants vivant dans un même pays de destination, mais issus de pays ou groupes d'origine différents. Son objectif est de maintenir constant, autant que possible, l'environnement institutionnel, économique et géographique contemporain, tout en faisant varier l'origine culturelle. Dans ce cadre, les différences observées entre individus vivant dans un même pays peuvent être interprétées comme reflétant des différences de normes, croyances ou préférences héritées de leur origine.

Cette stratégie repose cependant sur une condition forte : il faut pouvoir assigner correctement chaque individu à une origine culturelle pertinente. Le deuxième chapitre examine cette condition dans le cas des États-Unis, où de nombreux travaux utilisent l'ascendance étrangère déclarée dans les recensements américains pour identifier l'origine culturelle des individus nés aux États-Unis. L'ascendance déclarée est attractive parce qu'elle permet de relier des individus contemporains à des pays d'origine ancestrale. Mais elle ne mesure pas seulement cette origine. Elle reflète aussi la temporalité de la migration, c'est-à-dire le temps écoulé depuis l'arrivée des ancêtres aux États-Unis, ainsi que la sélection des individus qui continuent à déclarer une ascendance étrangère plusieurs générations après la migration.

Le chapitre documente cette hétérogénéité entre groupes d'ascendance à partir des recensements américains. Il construit plusieurs proxys du temps écoulé depuis la migration et de la sélection dans la déclaration d'ascendance, notamment la part des migrants de première génération au sein d'un groupe, la concentration spatiale des individus déclarant la même ascendance, et l'usage de la langue ancestrale. Ces mesures permettent d'identifier des groupes pour lesquels l'ascendance déclarée renvoie à une migration relativement récente et d'autres pour lesquels elle correspond à une origine plus ancienne, plus diffuse et potentiellement plus sélectionnée chez ceux qui la déclarent. Le chapitre montre ensuite que certains résultats attribués à la persistance culturelle peuvent être sensibles à cette hétérogénéité. Des différences interprétées comme des effets de la culture peuvent en partie refléter des différences de temporalité migratoire ou de sélection dans la déclaration d'ascendance.

Cette analyse ne remet pas en cause l'intérêt de l'approche épidémiologique, mais elle en précise les conditions de validité. La migration peut constituer une ressource méthodologique précieuse pour étudier la culture, à condition de reconnaître les biais inhérents à la mesure de l'origine culturelle : l'ascendance déclarée est une mesure utile, mais elle doit être interprétée avec prudence. La contribution du chapitre est donc de proposer une réflexion méthodologique sur l'usage de l'ascendance dans l'économie de la culture, ainsi que des diagnostics simples permettant d'évaluer la robustesse des résultats existants aux différences de temporalité migratoire et de sélection entre groupes d'origine.

Le troisième chapitre analyse la mobilité sous un angle différent. Il ne s'agit plus d'étudier comment la mobilité complique la mesure de la culture, ni comment elle peut être utilisée pour identifier empiriquement ses effets, mais de comprendre comment elle peut devenir une stratégie face à une norme sociale contraignante. Le chapitre s'intéresse au cas des mutilations génitales féminines en Afrique subsaharienne. Cette pratique est profondément ancrée dans les contextes locaux, et sa persistance dépend à la fois des préférences des parents et des attentes de la communauté. Des parents opposés à l'excision peuvent ainsi subir une pression sociale importante lorsqu'ils vivent dans un environnement où la pratique est la norme. Dans un tel contexte, une stratégie d'évitement consistant à éloigner temporairement ou durablement la fille de l'environnement social dans lequel la norme est appliquée peut émerger.

Le chapitre étudie cette hypothèse à travers le confiage des enfants, une forme courante de mobilité en Afrique subsaharienne par laquelle un enfant est envoyé vivre dans un autre ménage que celui de sa mère. À partir des Enquêtes Démographiques et de Santé menées dans treize pays d'Afrique subsaharienne, le chapitre construit des mesures de prévalence de l'excision au niveau de groupes définis par l'ethnicité, la région et la cohorte. Il examine ensuite si les filles issues de groupes où l'excision est fortement ancrée sont plus susceptibles d'être confiées que les garçons, et si cette différence dépend des préférences déclarées par la mère à l'égard de la pratique.

Les résultats montrent que les filles sont plus souvent confiées que les garçons dans les groupes où la prévalence de l'excision est élevée. En revanche, la naissance d'une fille dans un environnement à forte prévalence d'excision ne s'accompagne pas d'une augmentation significative de la migration du ménage. L'effet sur le confiage est particulièrement marqué lorsque la mère déclare s'opposer à la pratique. Ces résultats sont cohérents avec l'idée que le confiage peut être utilisé comme une stratégie de protection lorsque les préférences parentales entrent en conflit avec les attentes de la communauté. Le chapitre montre également que cette stratégie dépend de l'existence d'options de sortie crédibles. Elle est plus forte lorsque les groupes présentent une importante variation géographique de la prévalence de l'excision, ou lorsque les parents appartiennent à des groupes caractérisés par des niveaux de prévalence différents, ce qui peut élargir l'ensemble des contextes moins risqués vers lesquels les filles peuvent être envoyées.

Le chapitre examine ensuite plusieurs explications alternatives. Les résultats ne semblent pas s'expliquer par des normes générales de genre, par le mariage précoce ou par une plus grande propension générale à confier les filles dans certains groupes. Ils sont également cohérents avec deux résultats additionnels. D'une part, les filles confiées sont moins susceptibles d'être excisées. D'autre part, au Sénégal, le recul de la pratique après l'interdiction de l'excision en 1999 s'accompagne d'une baisse du confiage des filles dans les groupes où l'excision était initialement la plus ancrée. Ces éléments renforcent l'interprétation du confiage comme une forme d'exit spatial face à une norme sociale contraignante.

La contribution du troisième chapitre est de montrer que la mobilité peut être mobilisée par les individus comme une réponse à la pression normative locale. Elle ne constitue pas seulement un déplacement géographique, mais une manière de modifier l'exposition d'un enfant à un environnement social donné. Le chapitre contribue ainsi à la littérature sur les normes sociales et les pratiques néfastes en soulignant que les comportements observés peuvent refléter non seulement l'adhésion ou l'opposition à une norme, mais aussi les stratégies mises en œuvre pour y échapper lorsque la contestation directe est coûteuse.

Ces trois chapitres montrent que l'étude empirique de la culture dépend étroitement des choix méthodologiques qui permettent de la mesurer. La mobilité y apparaît sous trois formes différentes. Elle est d'abord une source de fragilité, car elle complique l'assignation des individus à une origine culturelle. Elle est ensuite une ressource méthodologique, car elle permet de comparer des individus vivant dans un même contexte contemporain mais issus d'origines différentes. Elle est enfin une stratégie concrète, par laquelle les individus peuvent modifier leur exposition aux normes sociales. Cette thèse vise ainsi à contribuer à une compréhension plus dynamique des liens entre culture, origine et mobilité spatiale en économie du développement.